

Cell Genomics, Volume 4

Supplemental information

**Combinatorial expression motifs
in signaling pathways**

Alejandro A. Granados, Nivedita Kanrar, and Michael B. Elowitz

Table S1. List of single-cell RNA seq datasets used in this study (related to main figures 1 and 2)

Dataset	Time points	Reference	Cells	Mice sampled	Technology	No. of Cell States in Integrated Atlas
Forelimb atlas (The changing mouse embryo transcriptome at whole tissue and single-cell resolution)	E10.5, E11.0, E11.5, E12.0, E13.0, E13.5, E14.0, E15.0	[28]	90,637	Pair of forelimbs per time point	10X	90
A single-cell molecular map of mouse gastrulation and early organogenesis	E6.5, E6.75, E7.0, E7.25, E7.5, E7.75, E8.0, E8.25, E8.5	[63]	116,312	411 mouse embryos	10X	134
The emergent landscape of the mouse gut endoderm at single-cell resolution	E5.5	[30]	-	-	10X	20
Single-cell RNA-seq analysis unveils a prevalent epithelial/mesenchymal hybrid state during mouse organogenesis	E9.5, E10.5, E11.5	[31]	1916	7 embryos	Smart-seq2	28
Molecular recording of mammalian embryogenesis	E6.5, E7.0, E7.5, E8.0, E8.5	[27]	88,779	50 embryos	10X	181
Tabula Muris, Tabula Muris Senis	1mo, 3mo, 18mo, 21mo, 24mo, 30mo	[25,26]	450,000+	-	10X, Smart-seq2	753

Supplementary Figure 1

A

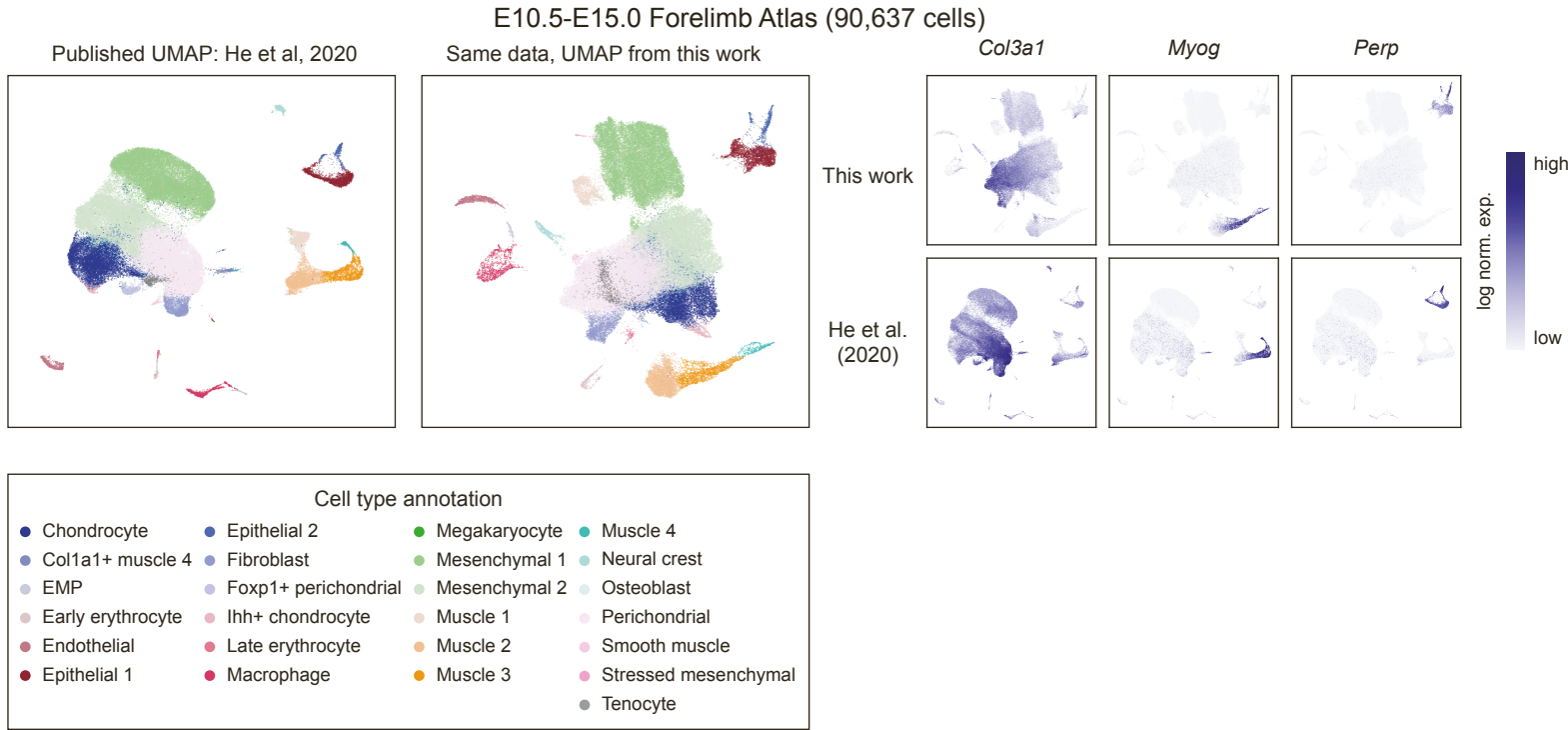


Figure S1. Single-cell RNA seq analysis and expression of TGF- β receptors (related to main figure 2)

- A. Analysis of scRNA-seq datasets using the standard Scanpy pipeline recapitulates published analyses, including [28]. Independent analysis of mouse forelimb over days E10.5-E15.0 captures similarity in cell types (colors, left) and gene expression (right).
- B. The integrated atlas captures cell type similarity across datasets. Cell clusters with similar annotations in different data sets remain similar to each other in the integrated atlas.
- C. Cluster-averaged profiles reflect co-expression in single-cells. Shown is an example of a single cluster from the forelimb epithelial tissue data set at day E15. Left, expression of TGF- β receptor genes averaged over all cells in the cluster corresponding to forelimb epithelial tissue at day E15.0. Right, pairwise conditional probability in single cells of gene 2 expression conditioned on gene 1 expression. Pairs of genes with significant entries (**) are co-expressed in the cluster-averaged profile. Higher-order conditional probabilities were not computed due to dropout effects in scRNA-seq data.

Supplementary Figure 2

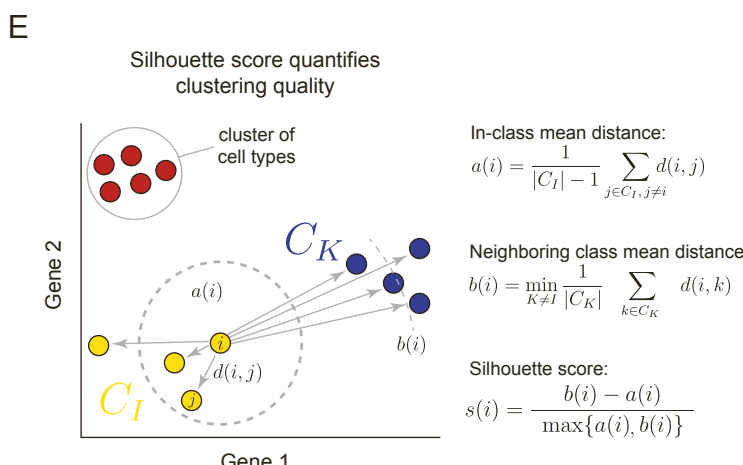
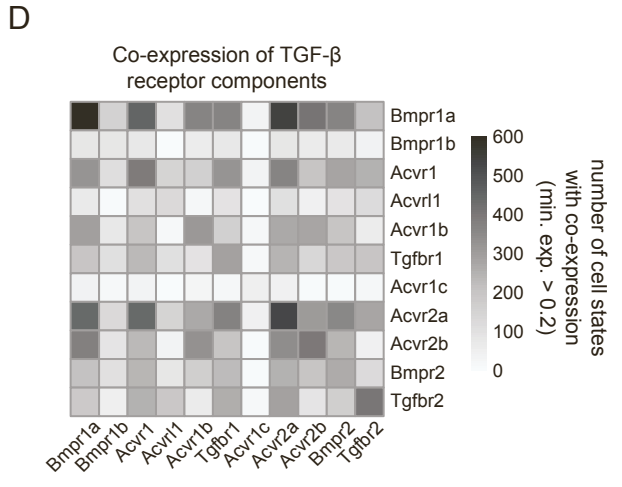
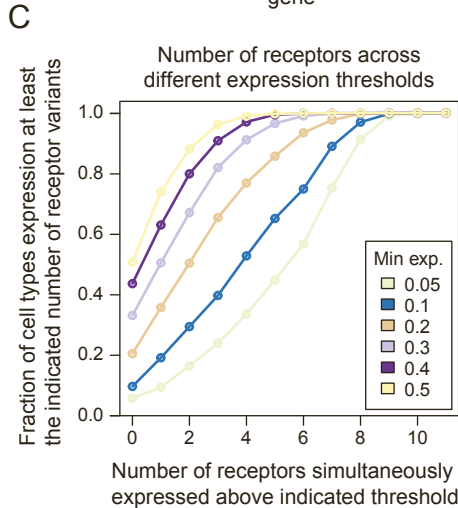
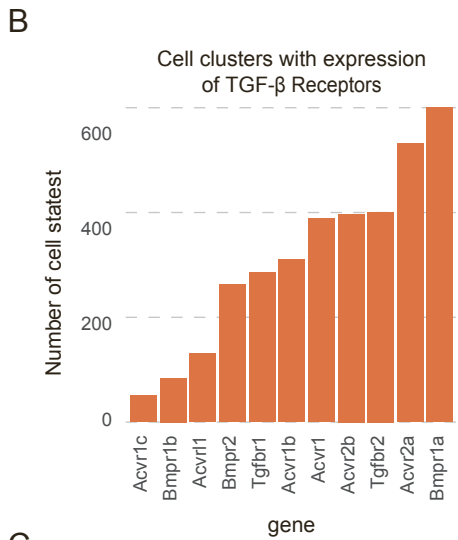


Figure S2. Co-expression statistics of TGF- β receptors across all cell states (related to main figure 3).

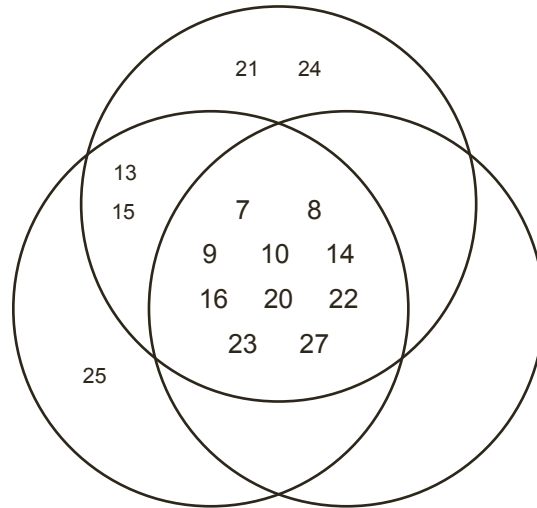
- A. A heatmap showing the average normalized expression of TGF- β receptors for the immune cell types not included in the clustering analysis due to low mRNA expression levels. Each gene's average expression level was normalized from 0 to 1 relative to the minimum and maximum values observed across all 1206 cell types (Figure 3B, Methods). A cell type was included if it showed normalized expression above 0.2 for more than two receptors. The color side bars indicate the tissue of origin and the broad immune cell class.
- B. Histogram showing the number of cell types in the integrated atlas with normalized expression of TGF- β receptors above a threshold of 0.2 in standardized expression units.
- C. The number of TGF- β receptor components simultaneously expressed for different values of the minimum expression threshold (colors).
- D. Pairwise co-expression of TGF- β receptor expression reveals broad receptor co-expression patterns. Off-diagonal elements indicate the number of cell states co-expressing, above threshold, the indicated pair of components. Diagonal elements indicate the number of cell states expressing the corresponding individual gene (as shown in A).
- E. The silhouette score quantifies clustering quality (schematic). For a given clustering, we compute the silhouette score on every data point i . We compute $a(i)$, the mean distance between i and every other point in the same cluster, and $b(i)$, the mean distance between i and the nearest neighboring cluster. The silhouette score for the data point i is then defined as the difference between the inter- and intra-cluster distances, normalized to the maximum of the two (equations). A silhouette score value close to 1 corresponds to well-defined clusters, where the data point i is similar to other members of its cluster and dissimilar to others. In contrast, a value close to -1 suggests poor cluster assignment. Finally, the silhouette score for a given clustering, is taken as the average of the individual scores for all data points. In our analysis, a data point corresponds to a vector of the scaled, normalized counts of pathway gene expression of a single cell state.

Supplementary Figure 3

A

TGF- β motifs obtained with different pairwise distance metrics

Mean of pairwise distances (in paper)

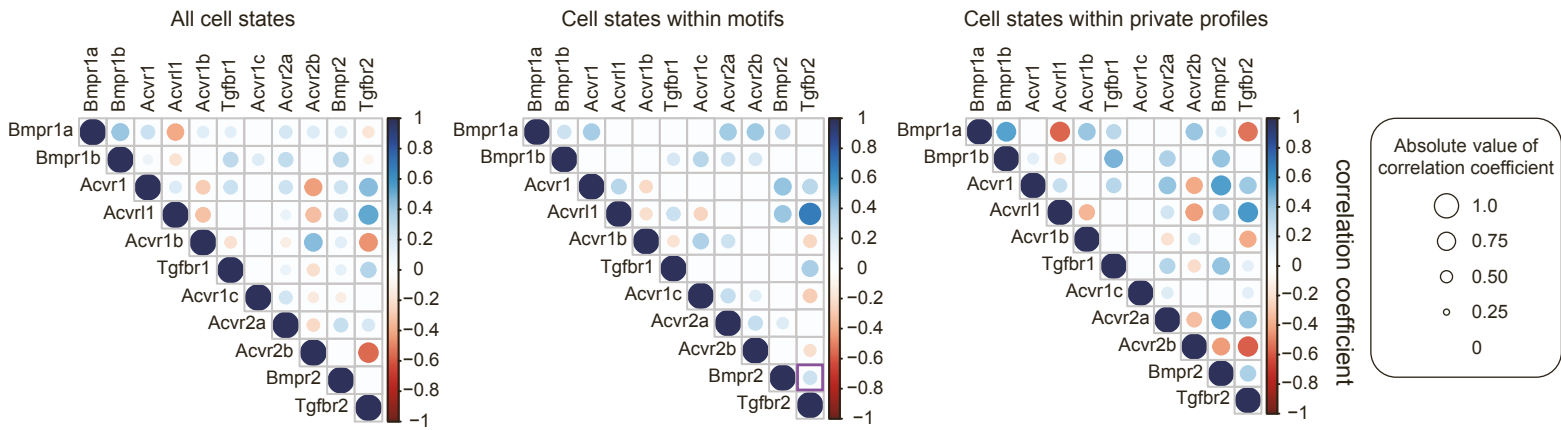


Top 10th percentile of pairwise distances

Maximum of pairwise distances

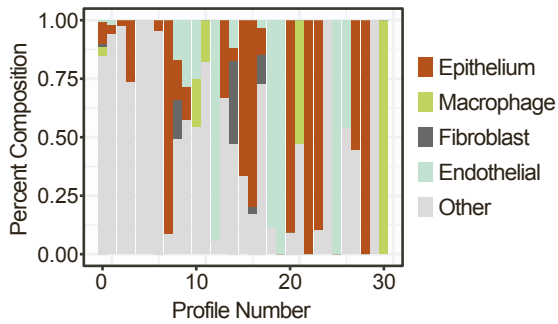
B

Pairwise correlations



C

Profile Tissue Composition



D

TGF- β 'private' profiles

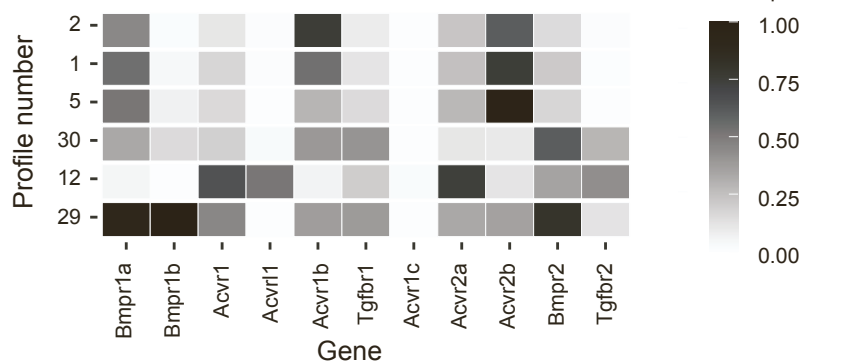


Figure S3. Dispersion metrics and private profiles for TGF- β (related to main figures 3 and 4).

- A. Alternative definitions of the dispersion metric recover similar sets of motifs. The mean of intra-class pairwise distances was used as the dispersion metric throughout this work. Still, we tested two additional dispersion metrics, one that uses the maximum of intra-class pairwise distances, the second that uses the top 10th percentile. The Venn diagram shows profiles identified as motifs from these three distinct definitions of the dispersion metric. Most profiles (shown in the intersection of the three circles) are robust to the definition of dispersion. Notably, the dispersion metric that utilizes the maximum pairwise distances only captures profiles in this intersection. The mean pairwise distance, however, captures two additional profiles as motifs, profiles 21 and 24. Profile 24 contains only two cell states, liver B cells and bone marrow NK cells. The top 10th percentile of pairwise distances captures the adult endothelium-specific profile, 25, as a motif. However, the maximum metric omits profiles 13 and 15, even though they appear to be motifs since they are both dispersed across the adult smooth muscle and adult kidney epithelium.
- B. TGF- β profiles exhibit unique pairwise receptor correlations. Each matrix represents the correlation coefficient for each pair of receptors across all cell states (left), cell states associated with motifs (middle), and cell states associated with private profiles (right).
- C. Cell class composition of each TGF- β profile, matching classes depicted in Figure 4F. “Other” includes all cell states in the atlas that do not fall into the epithelial, macrophage, fibroblast, or endothelial cell classes.
- D. TGF- β profiles with average dispersion less than the 50 percentile from the expected distribution were classified as *private profiles*.

Supplementary Figure 4

A

Well-defined peak (discrete profiles)

Broad peaks (continuous profiles)

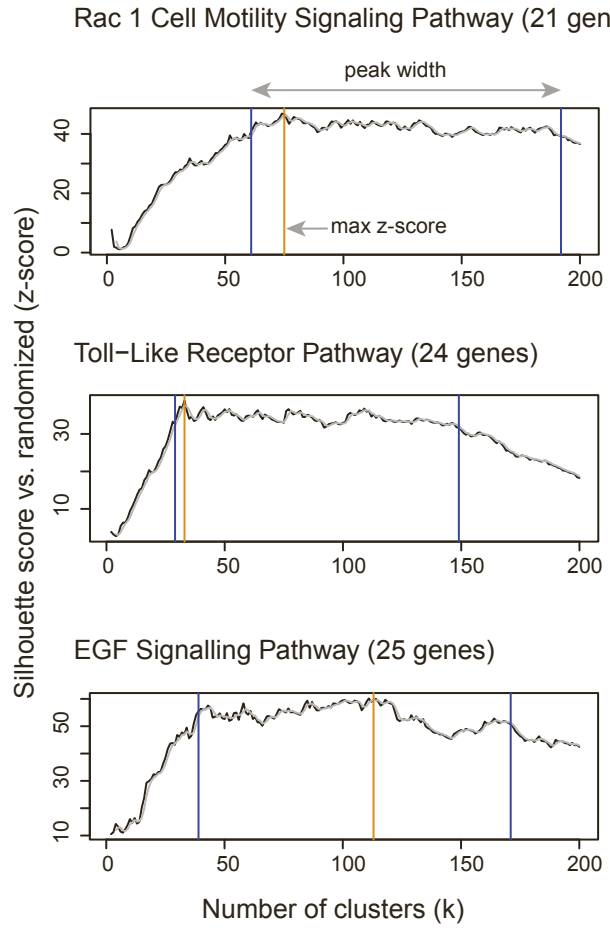
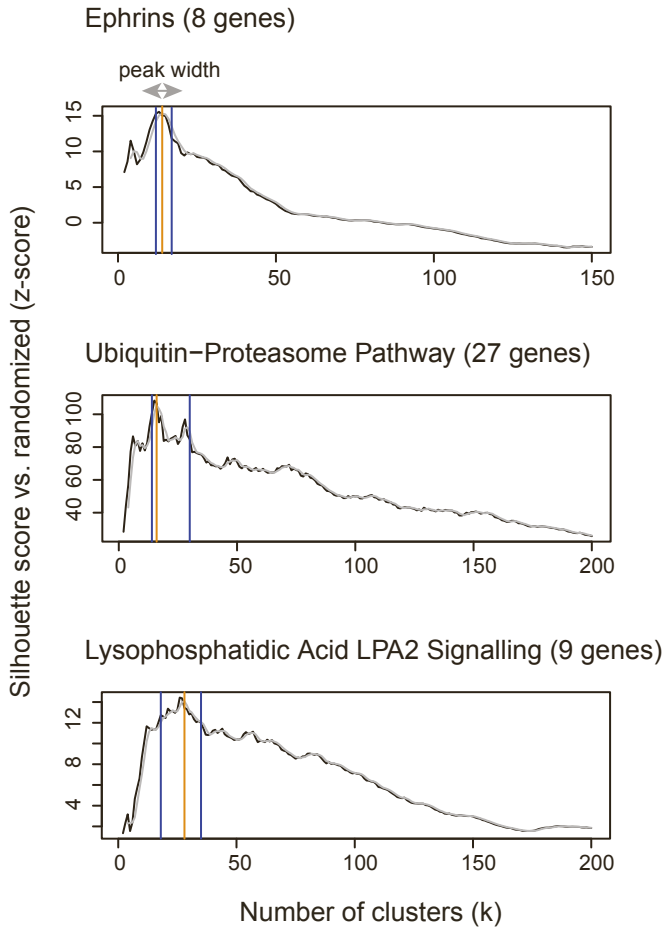
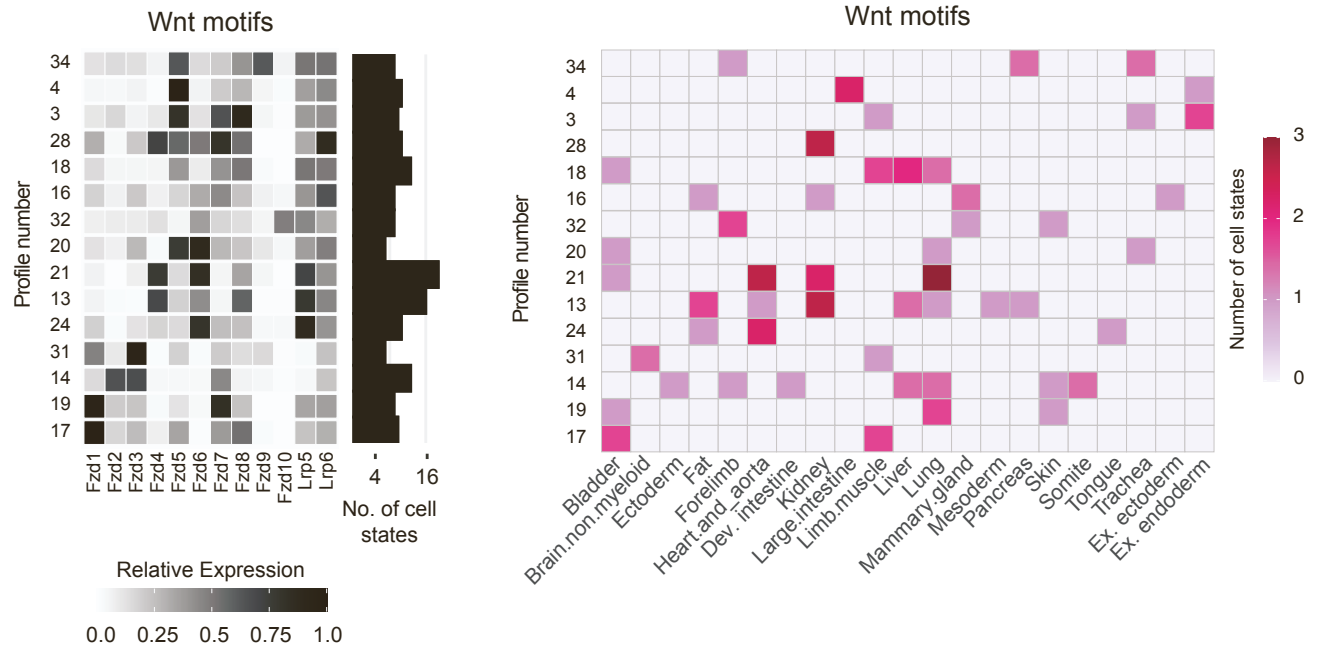


Figure S4. Silhouette analysis for other pathways (related to main figures 3 and 5).

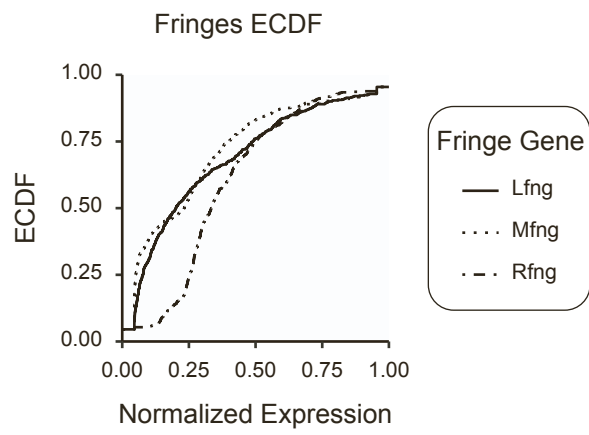
Silhouette analysis of indicated pathways, as in Figure 3A and Figure 5A. A vertical yellow line indicates the value of k (the proposed number of clusters) with the maximum silhouette z -score. The peak width is shown with two blue vertical lines (the minimum and maximum values of k with a 10% difference from the maximum value). Finally, the relative peak width is calculated as the ratio of the peak width divided by the maximum number of k considered. We show three pathways with the lowest (left) and highest (right) relative width.

Supplementary Figure 5

A



B



C

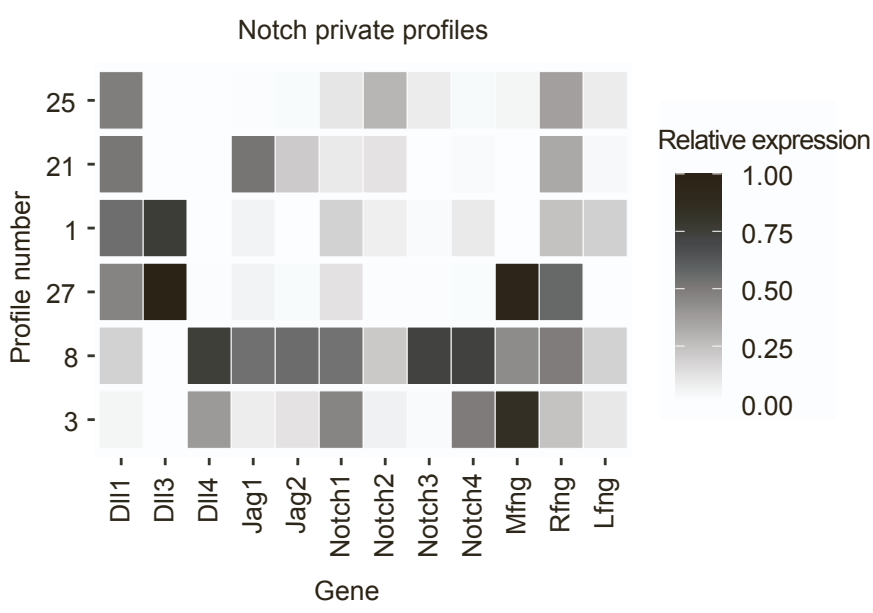
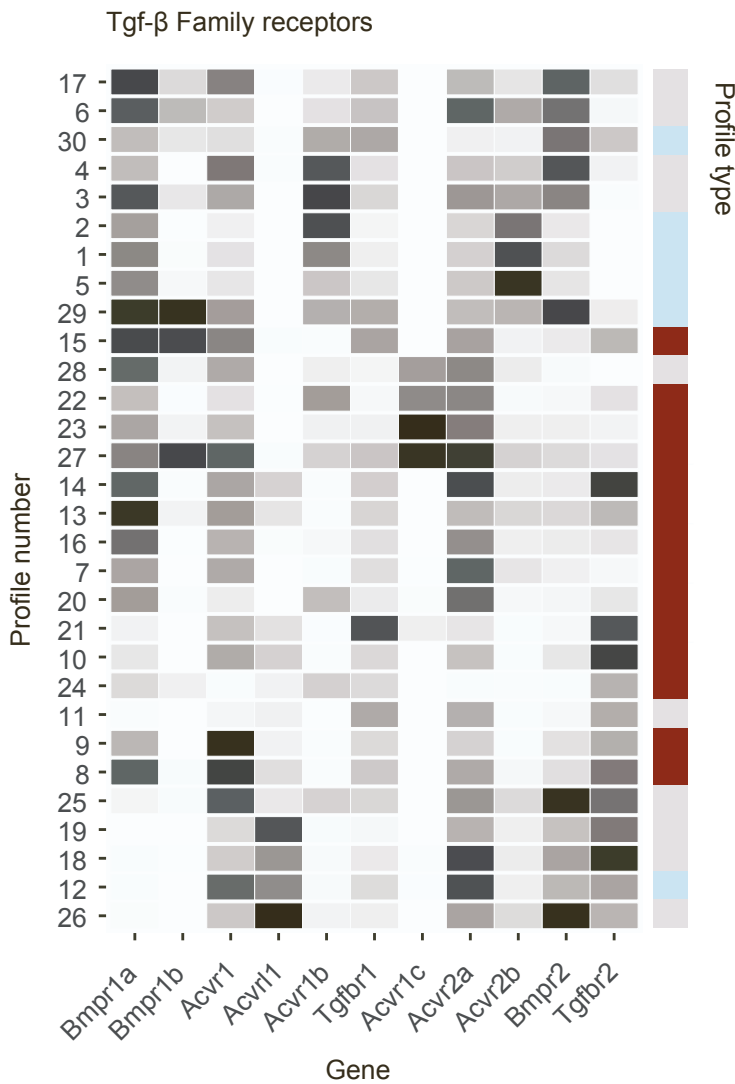


Figure S5. Motifs for Wnt and Notch pathways (related to main figures 4 and 5).

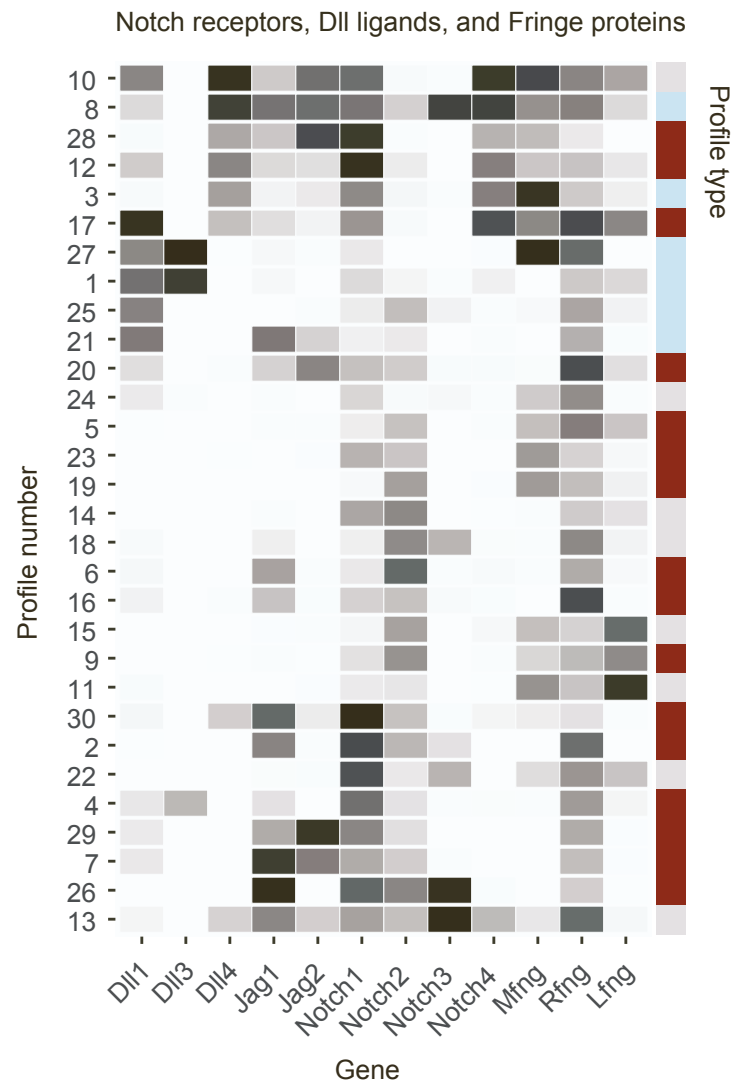
- A. Motifs in the Wnt pathway and their distribution across tissues and organs, similar to Figure 4D, E.
- B. Cumulative distribution histogram of Fringe gene expression across all cell states.
- C. Notch pathway private profiles.

Supplementary Figure 6

A



B



C

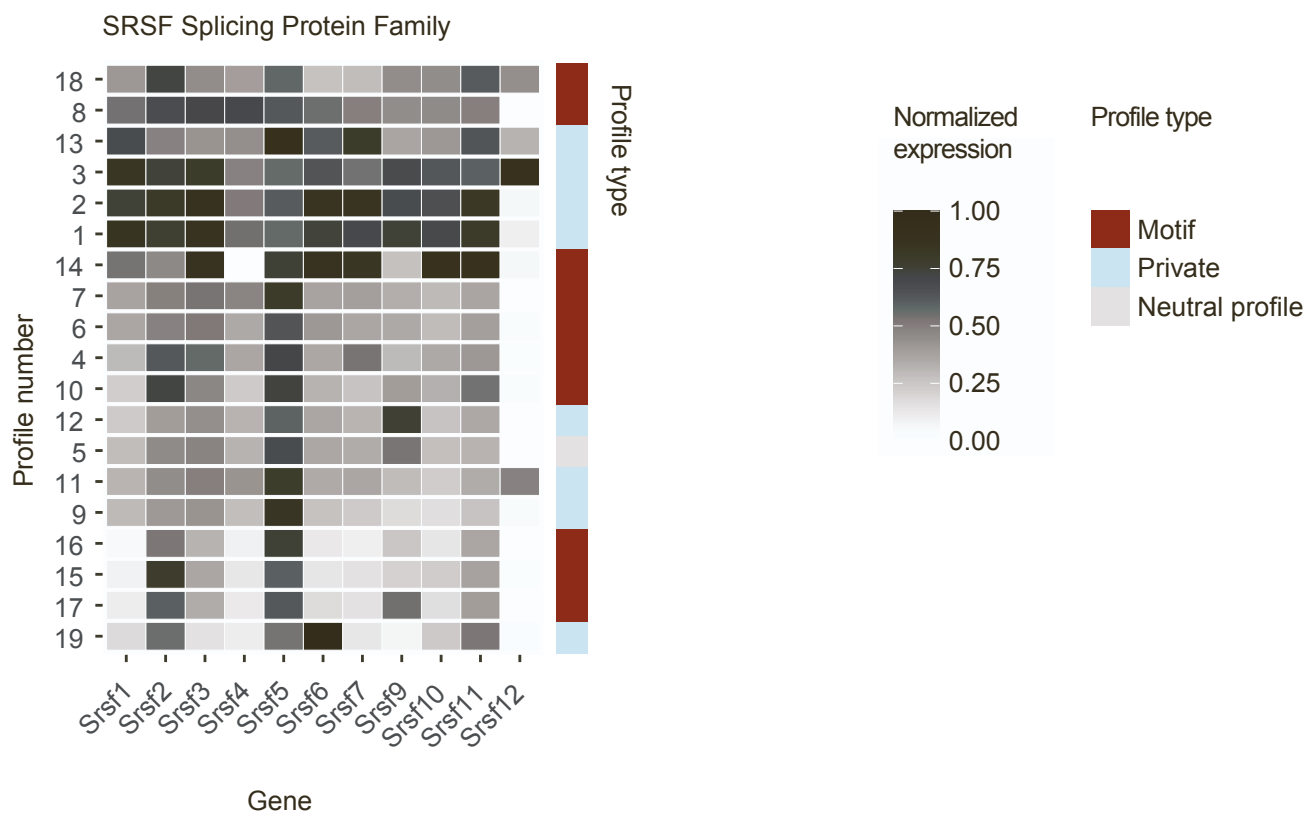


Figure S6. All profile classes for TGF- β , Notch, and SRSF (related to main figures 5 and 6).

All profiles are organized by similarity and annotated by their profile number and profile type (motif, private, or neutral) for TGF- β (A), Notch (B), and SRSF (C).

Supplemental data S1

Global cell type dendrogram, as shown in Figure 3C, along with a heatmap of the projection of each cell type on the top 20 principal components computed from the whole-transcriptome space.

Supplemental data S2

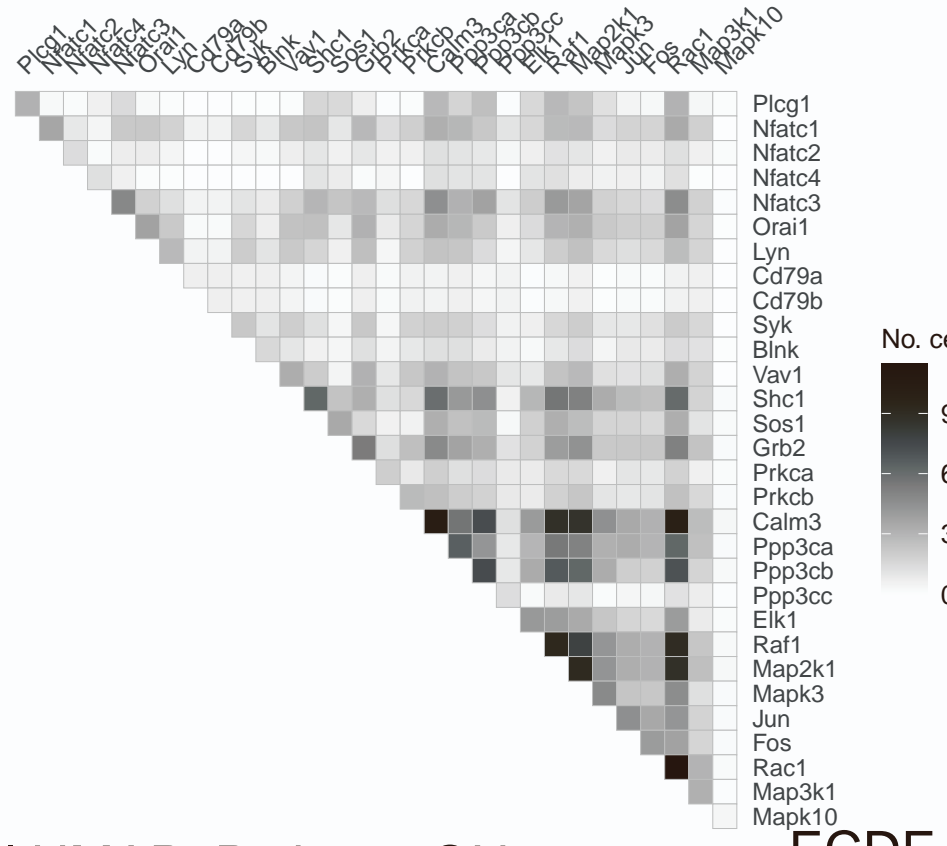
Analysis of pathways shown in Figure 5C. Each pathway has plots displayed on two pages. The first page has plots displayed for pathways as done for TGF-beta in Figures 3 and 4. The second page shows UMAPs highlighting each pathway motif (colored dots) and displays the dispersion score for the motif.

BCR Signaling Pathway (k_opt = 110)

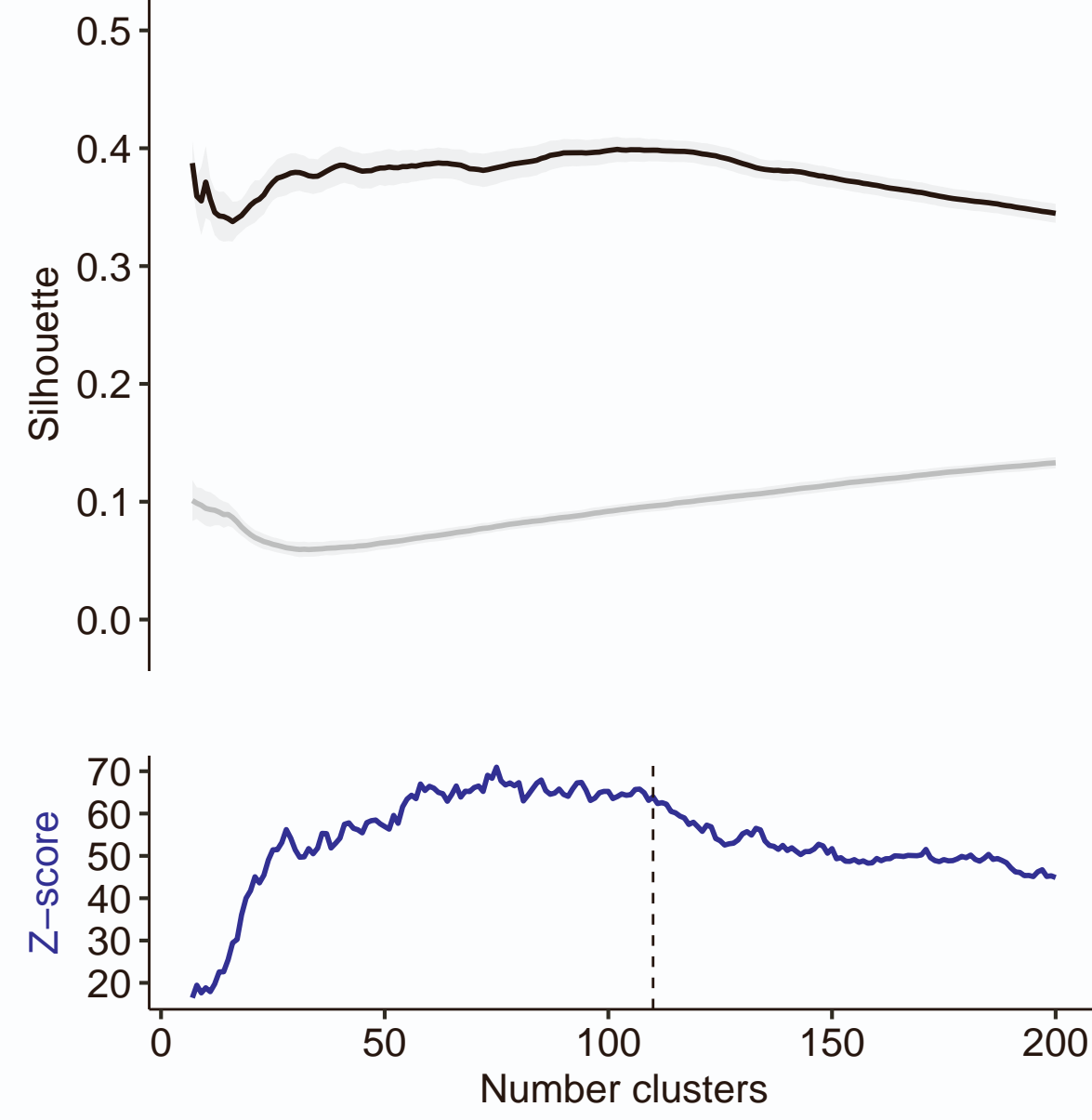
Pathway genes

Plcg1	Blnk	Ppp3cc
Nfatc1	Vav1	Elk1
Nfatc2	Shc1	Raf1
Nfatc4	Sos1	Map2k1
Nfatc3	Grb2	Mapk3
Orai1	Prkca	Jun
Lyn	Prkcb	Fos
Cd79a	Calm3	Rac1
Cd79b	Ppp3ca	Map3k1
Syk	Ppp3cb	Mapk10

Gene Co-Expression Above Threshold



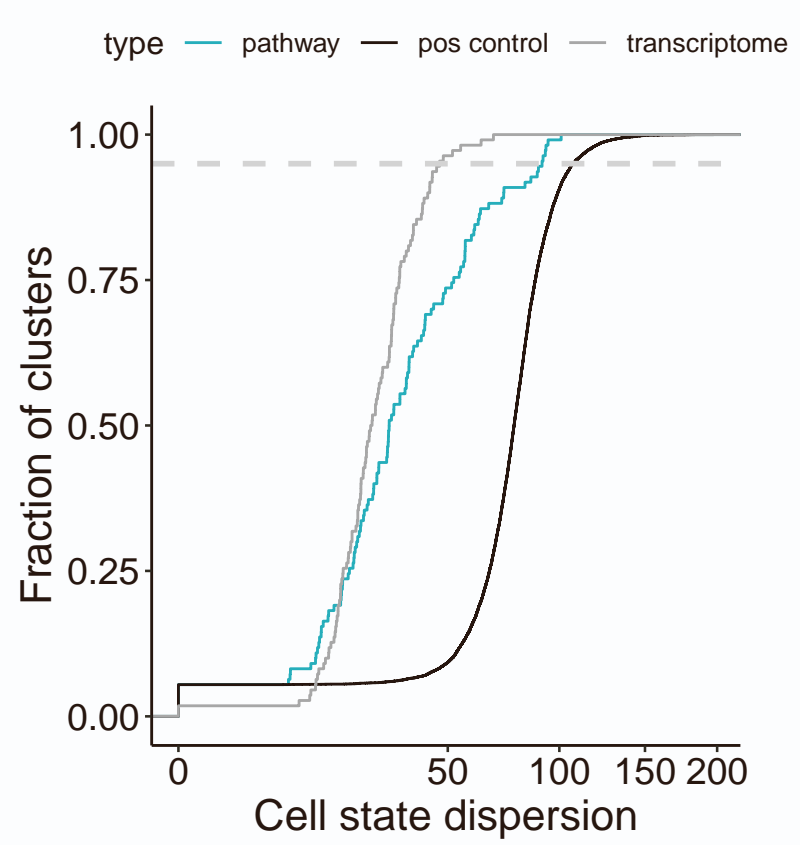
Silhouette and Z-score



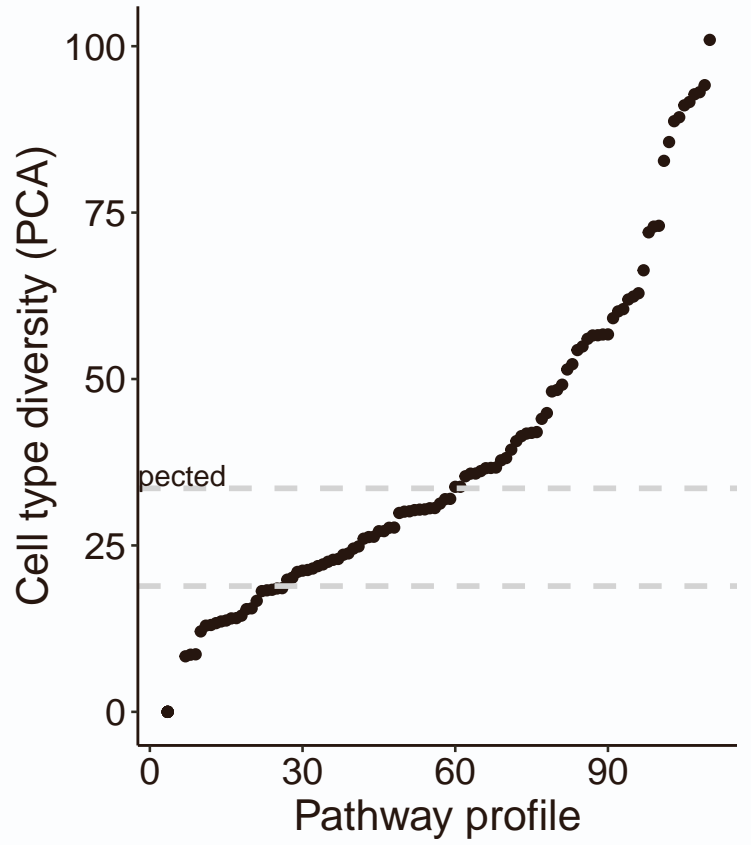
Global UMAP: Pathway ON



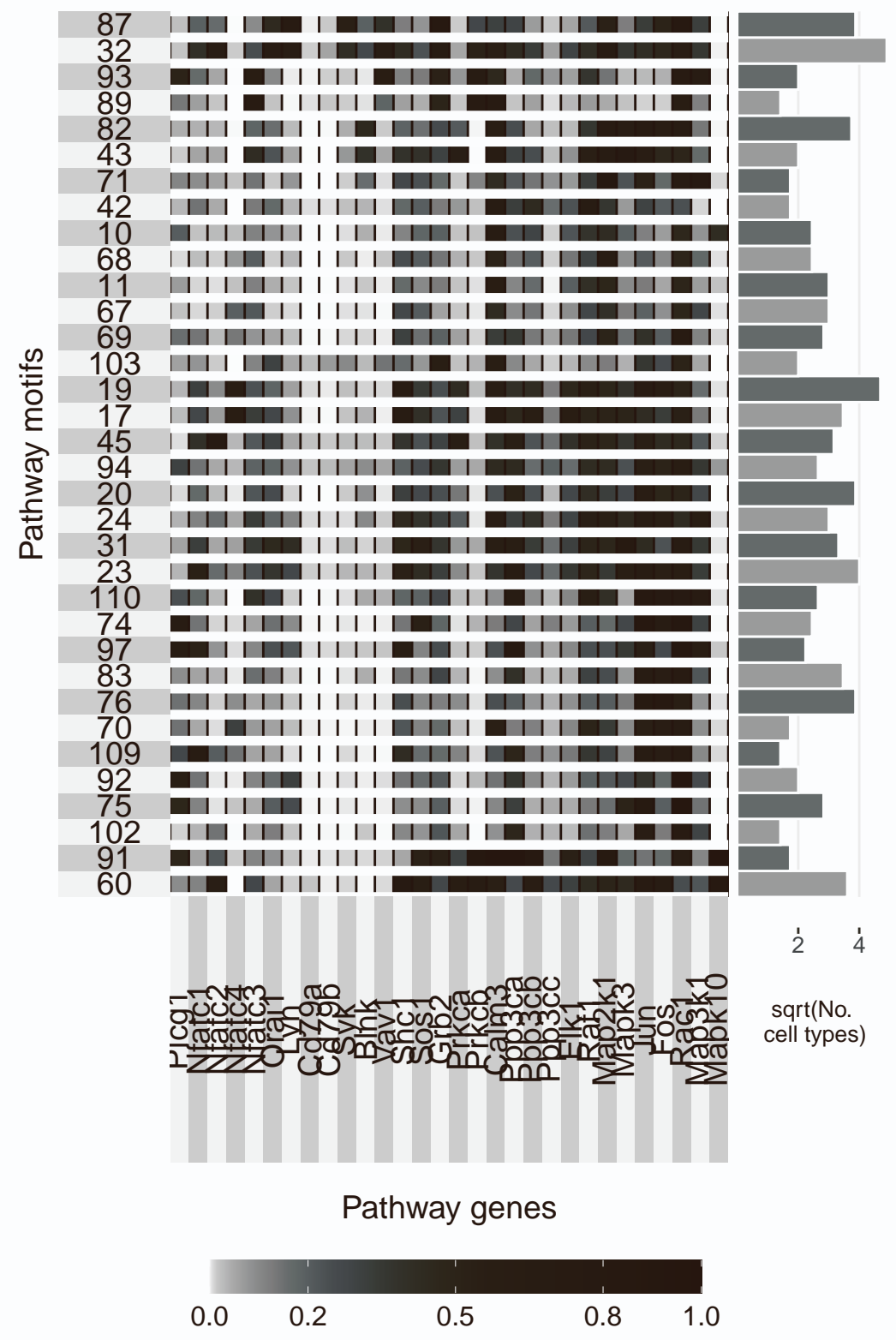
ECDF of Profile Dispersions



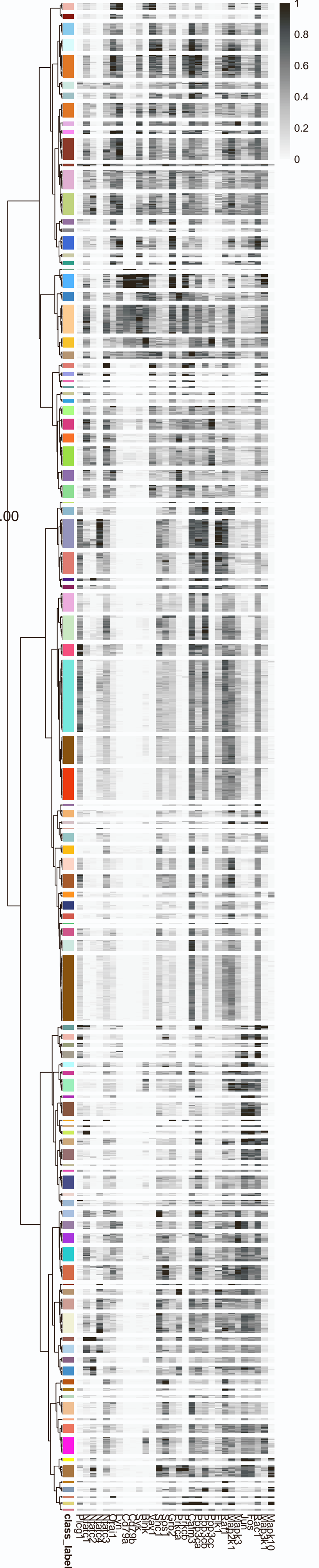
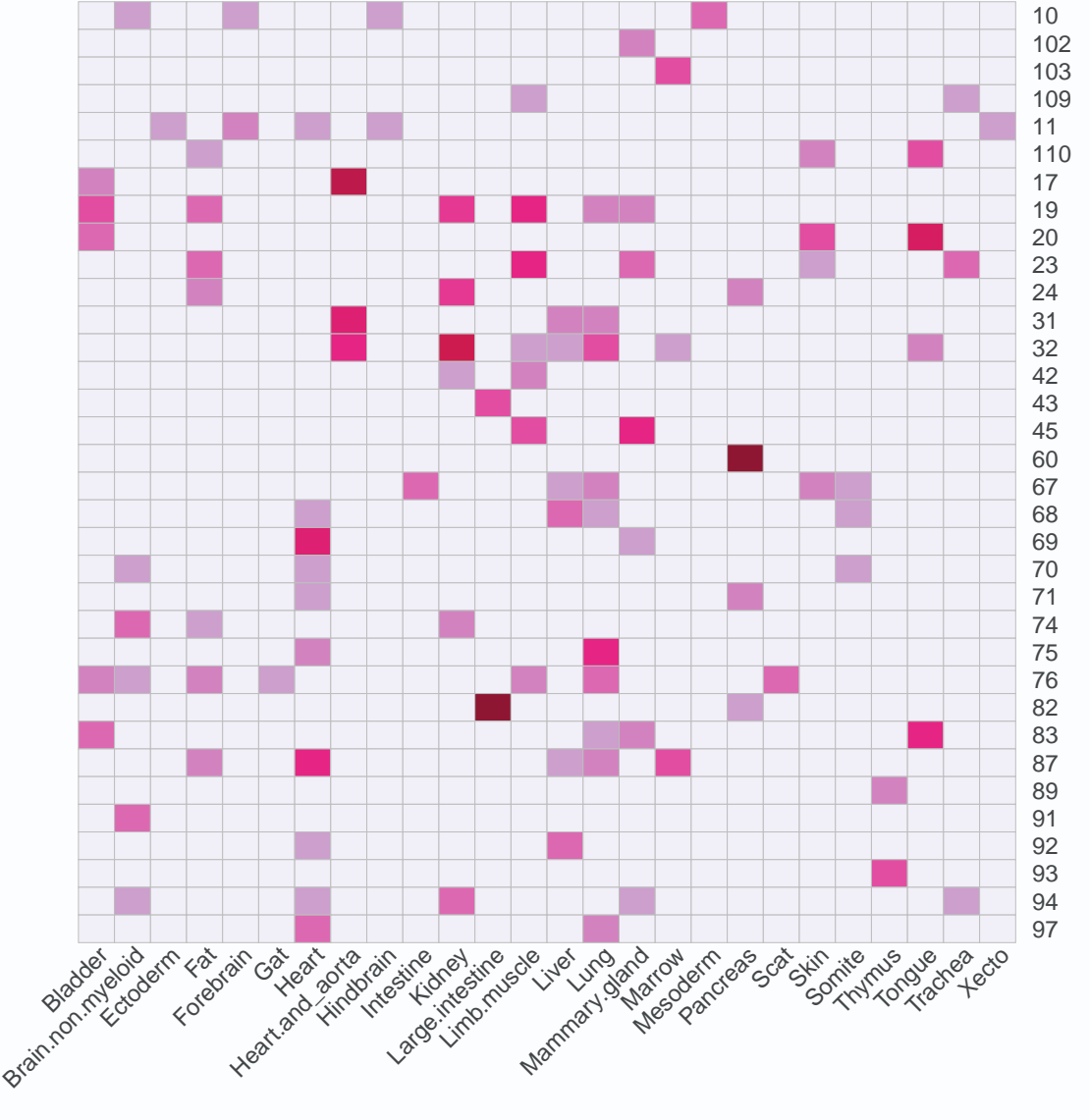
Profile Dispersion Plot



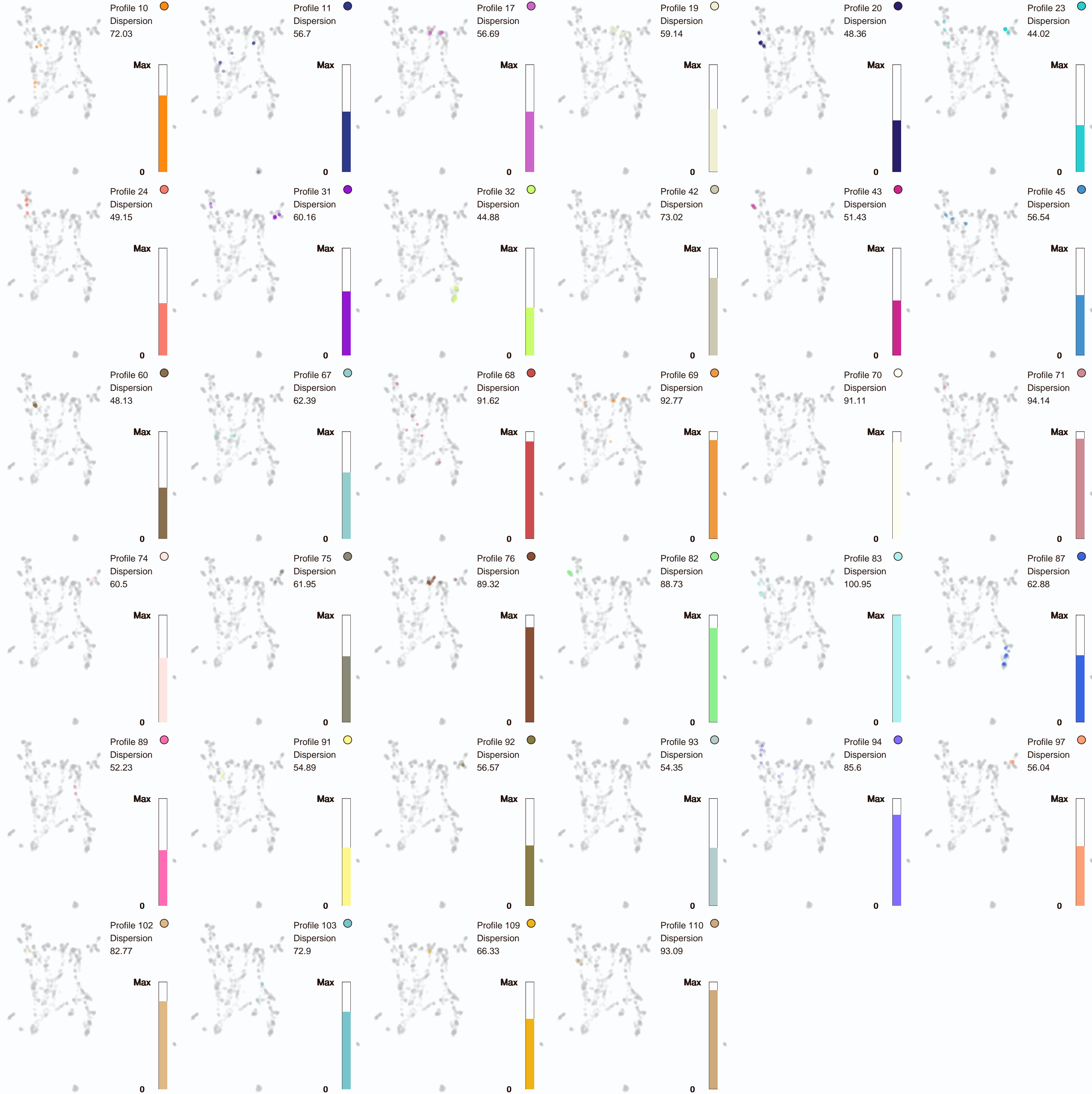
Motifs, Dispersion $\geq 90^{\text{th}}$ percentile



Motif Tissue Composition



BCR Signaling Pathway (k_opt = 110)

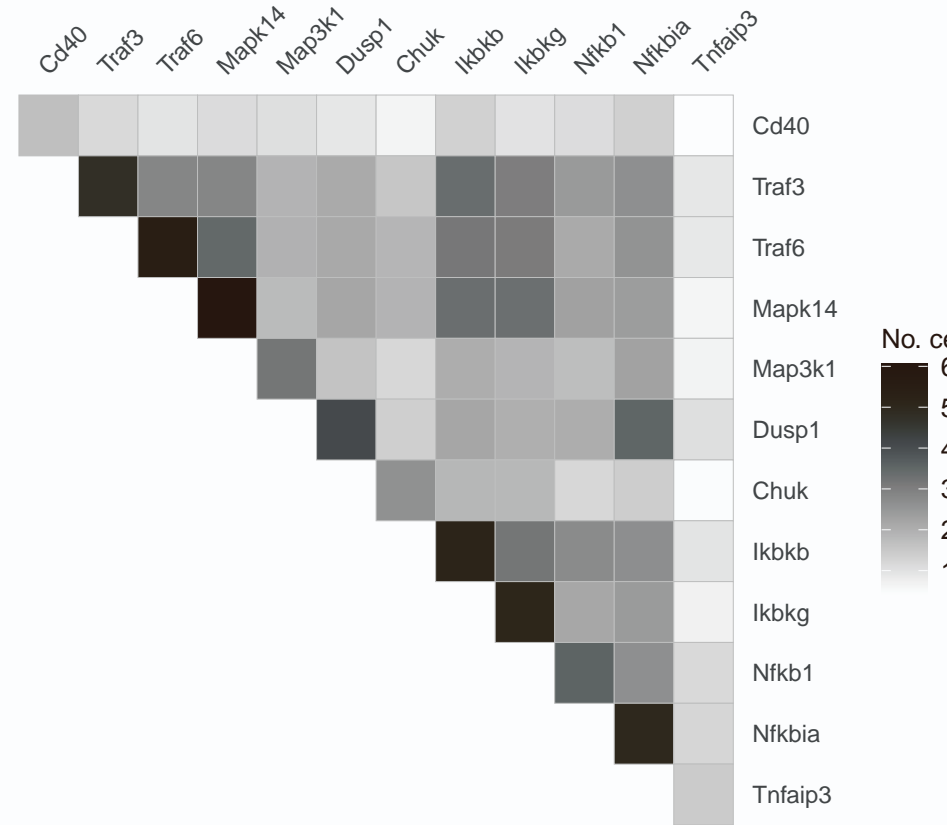


CD40L Signalling Pathway (k_opt = 67)

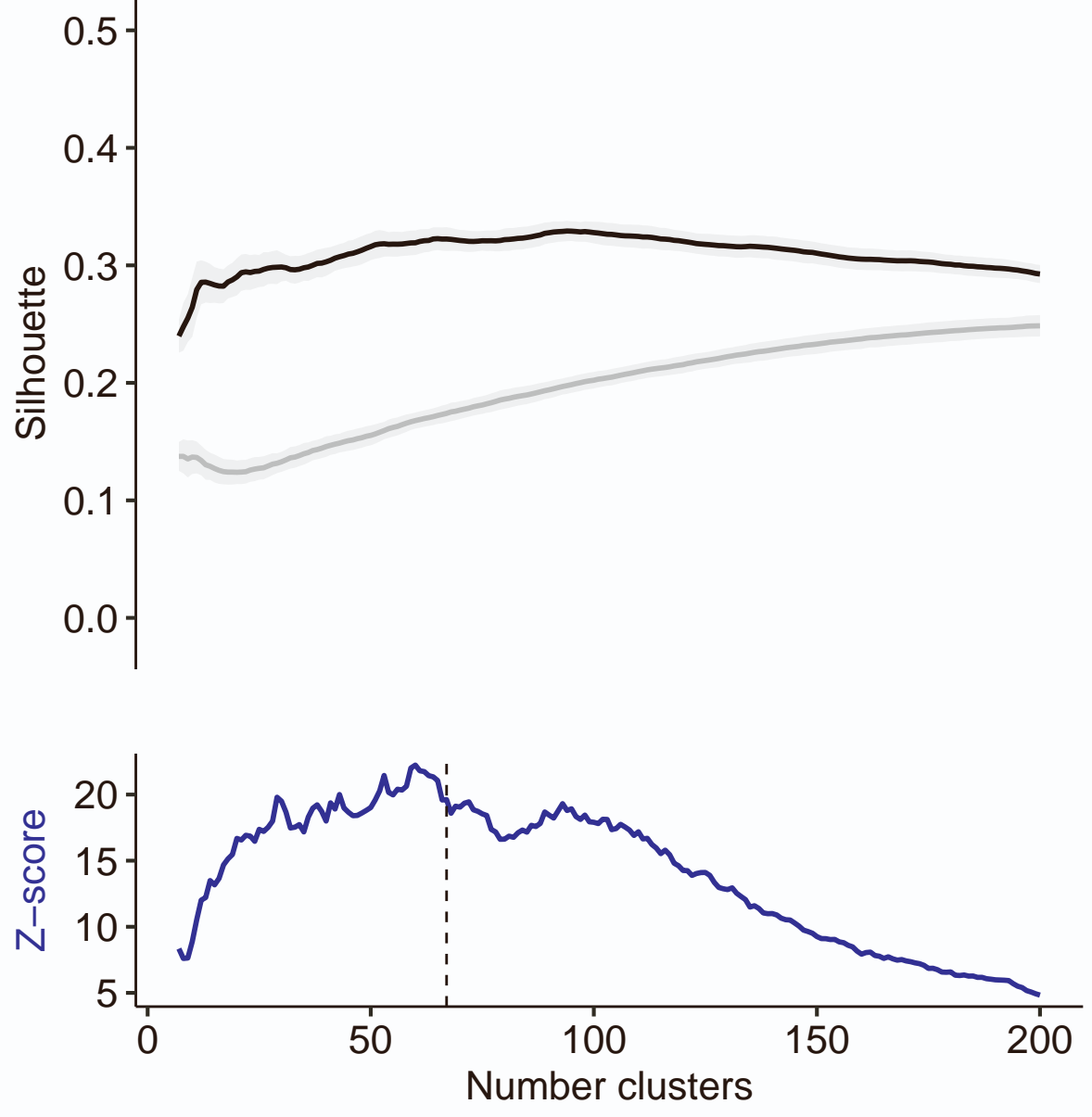
Pathway genes

Cd40	Map3k1	Ikbkg
Traf3	Dusp1	Nfkb1
Traf6	Chuk	Nfkbia
Mapk14	Ikbkb	Tnfaip3

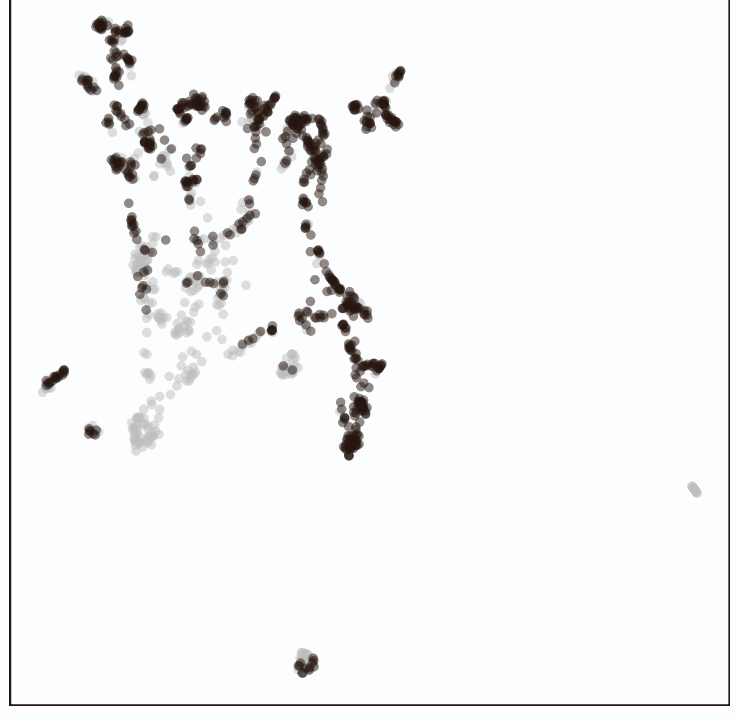
Gene Co-Expression Above Threshold



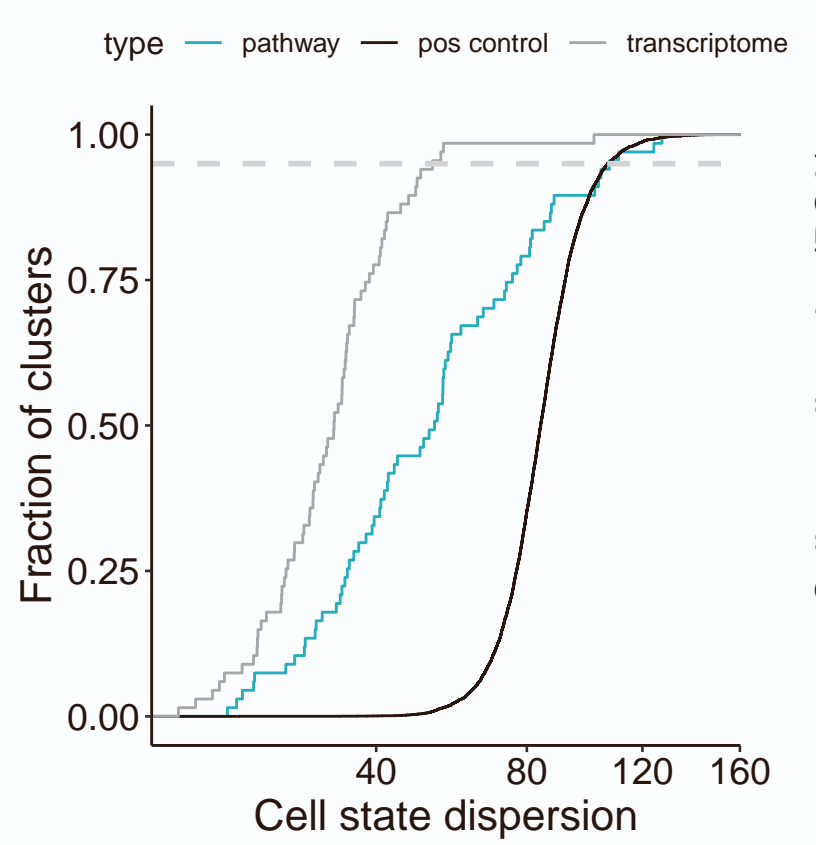
Silhouette and Z-score



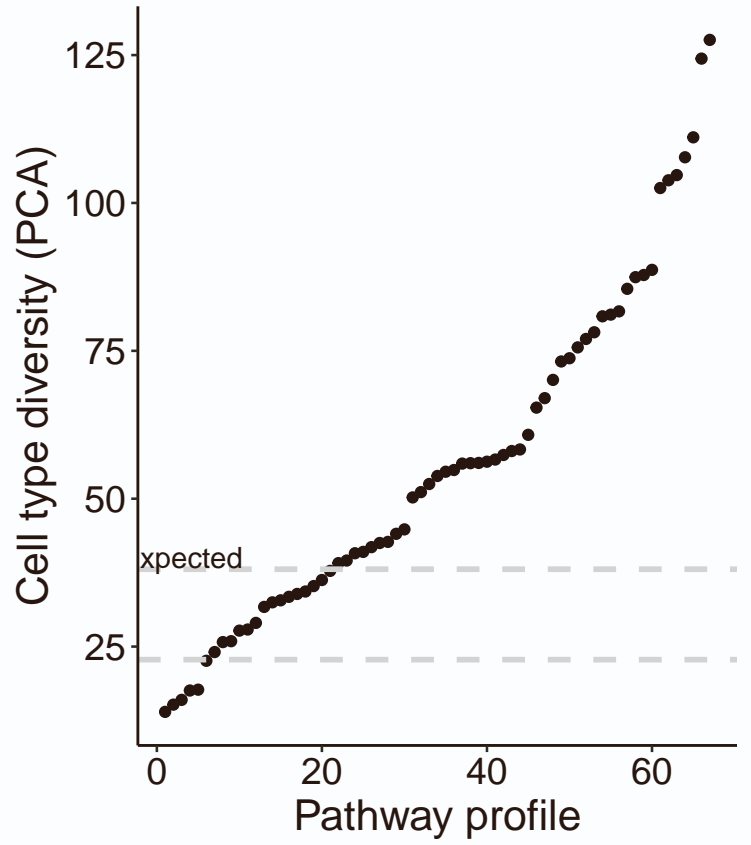
Global UMAP: Pathway ON



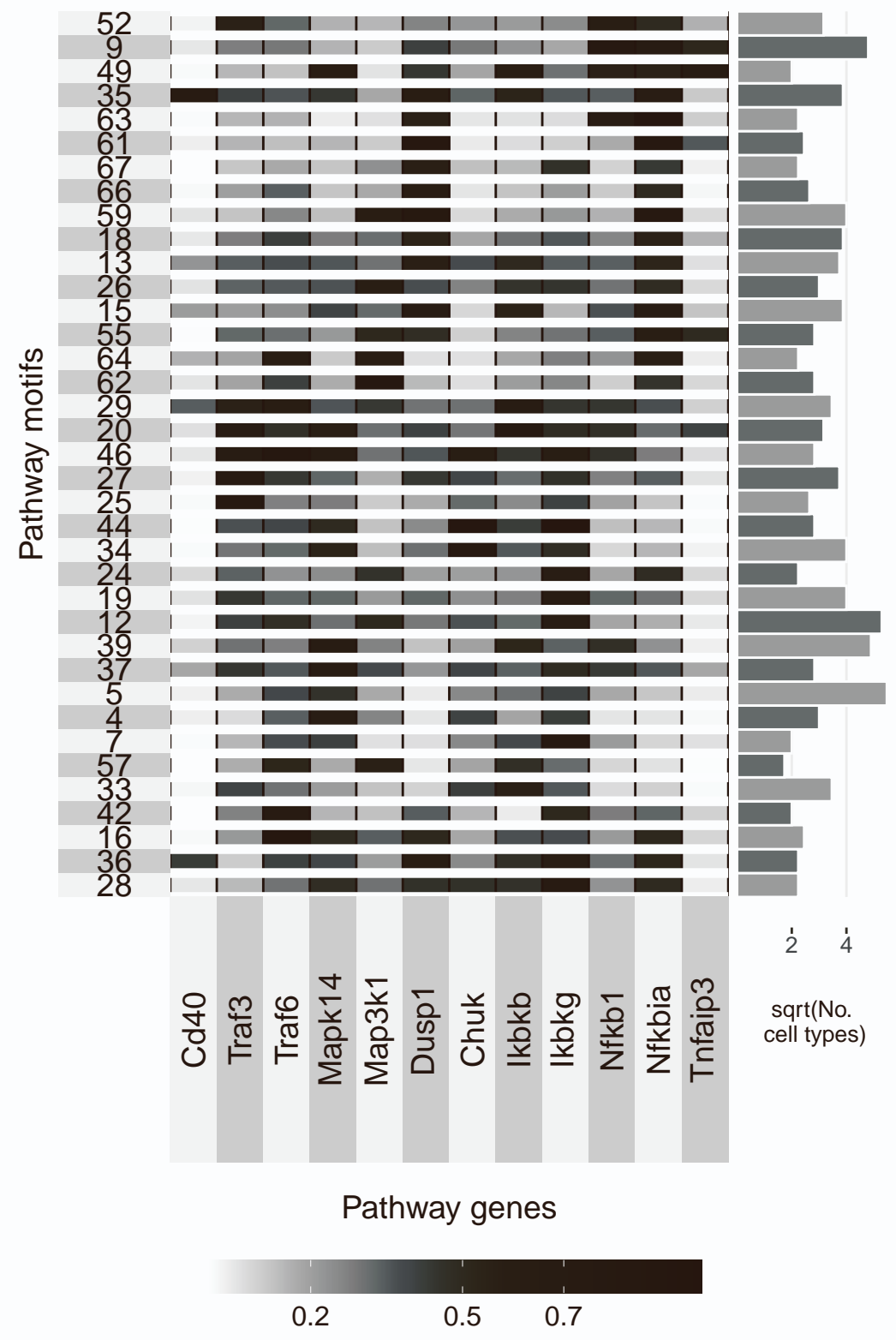
ECDF of Profile Dispersions



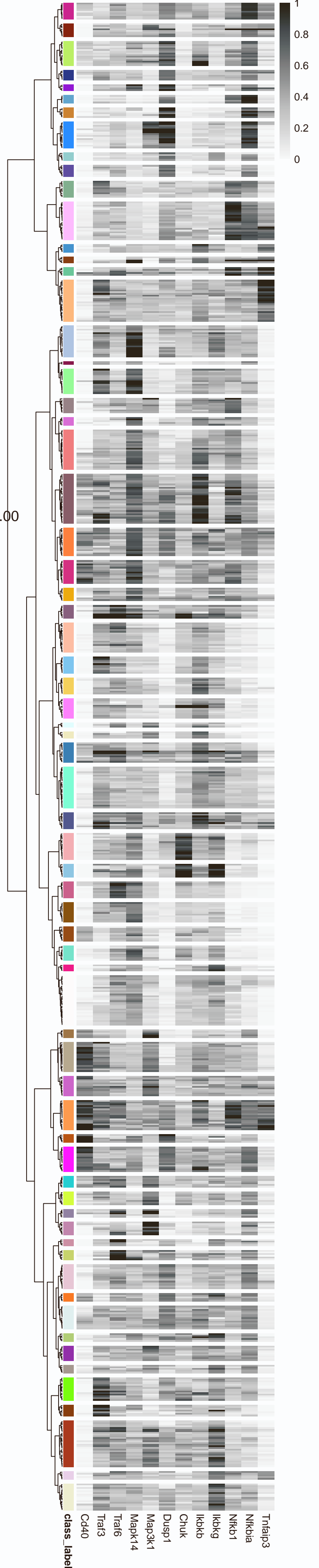
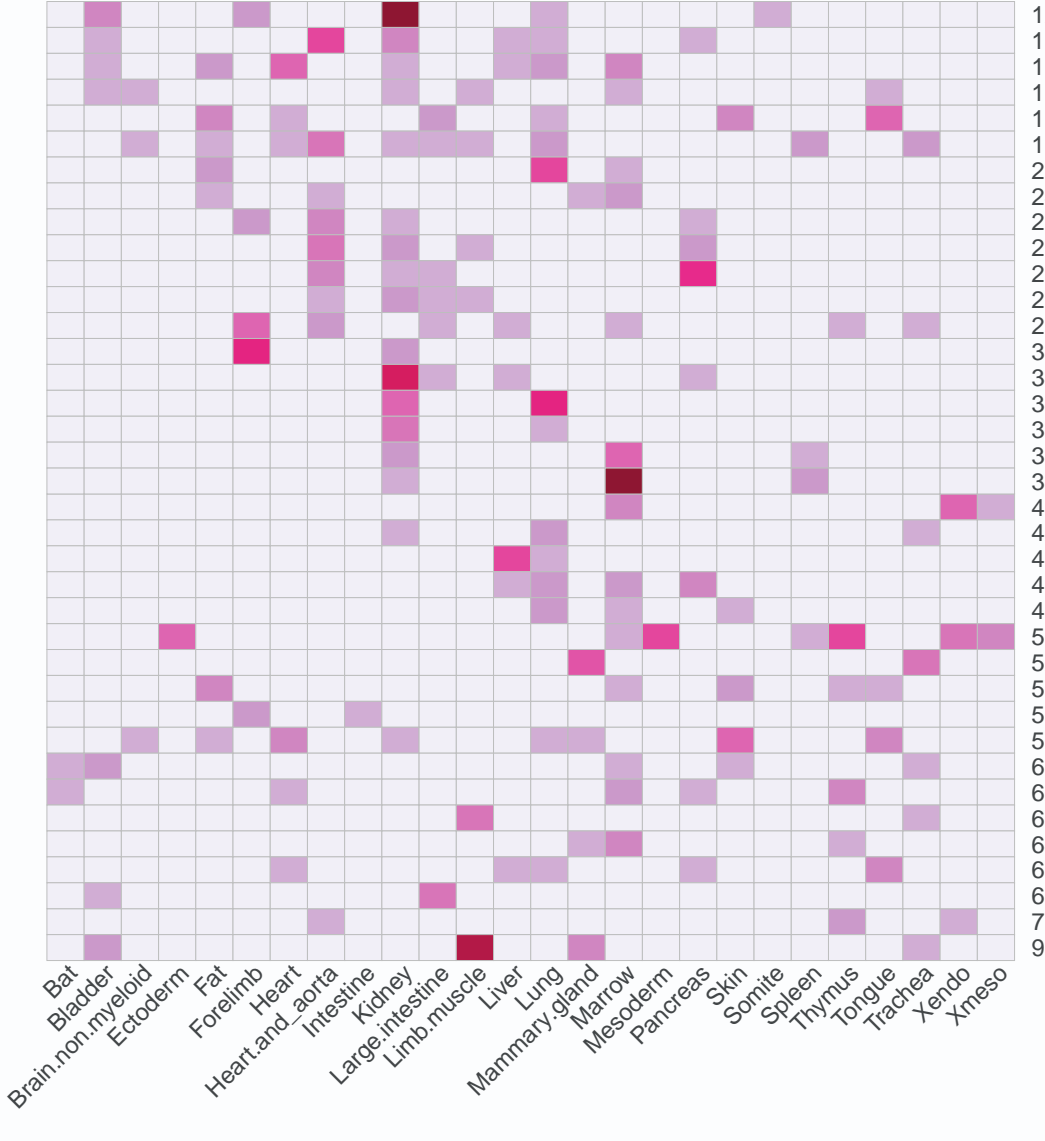
Profile Dispersion Plot



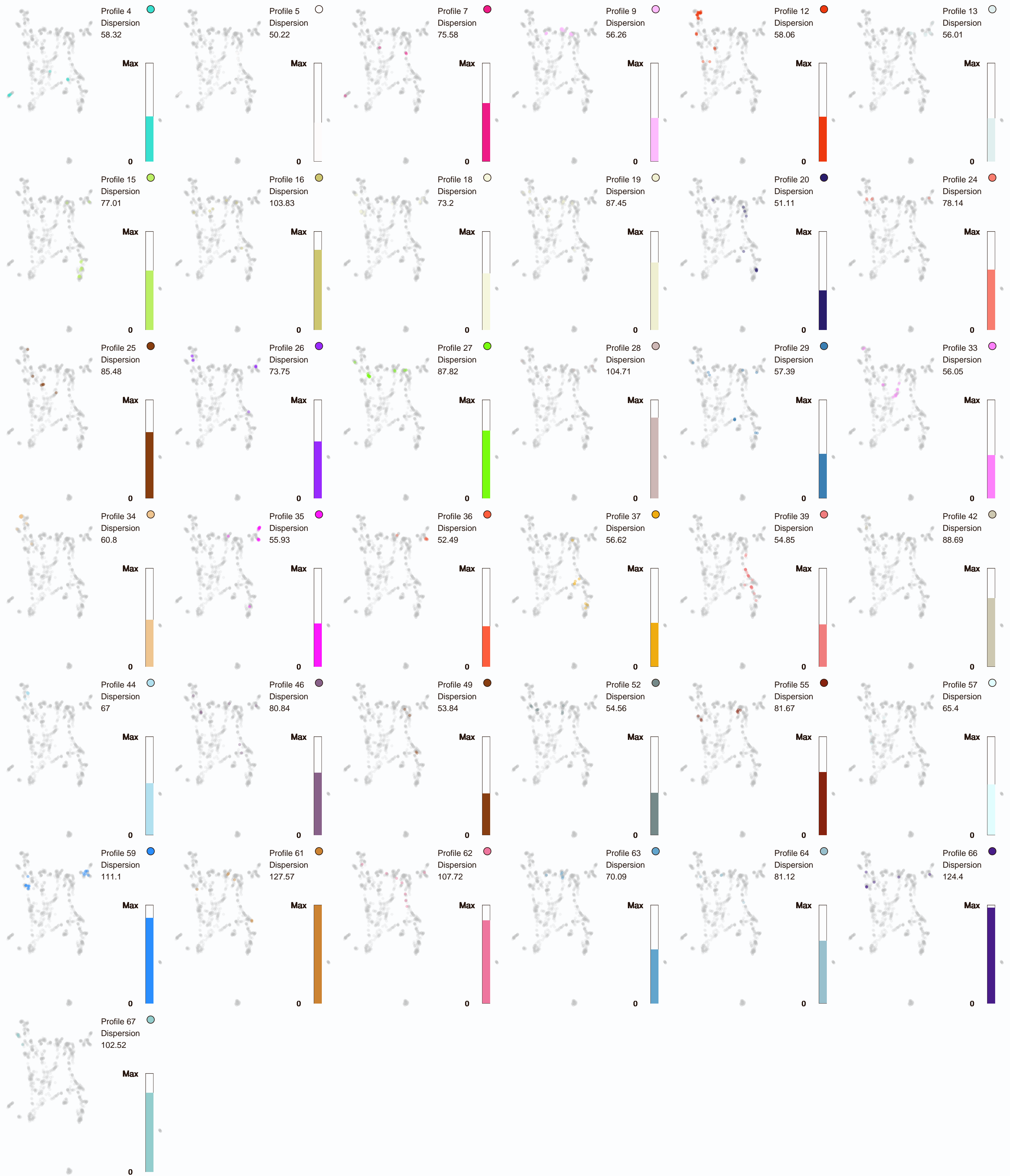
Motifs, Dispersion $\geq 90^{\text{th}}$ percentile



Motif Tissue Composition



CD40L Signalling Pathway (k_opt = 67)



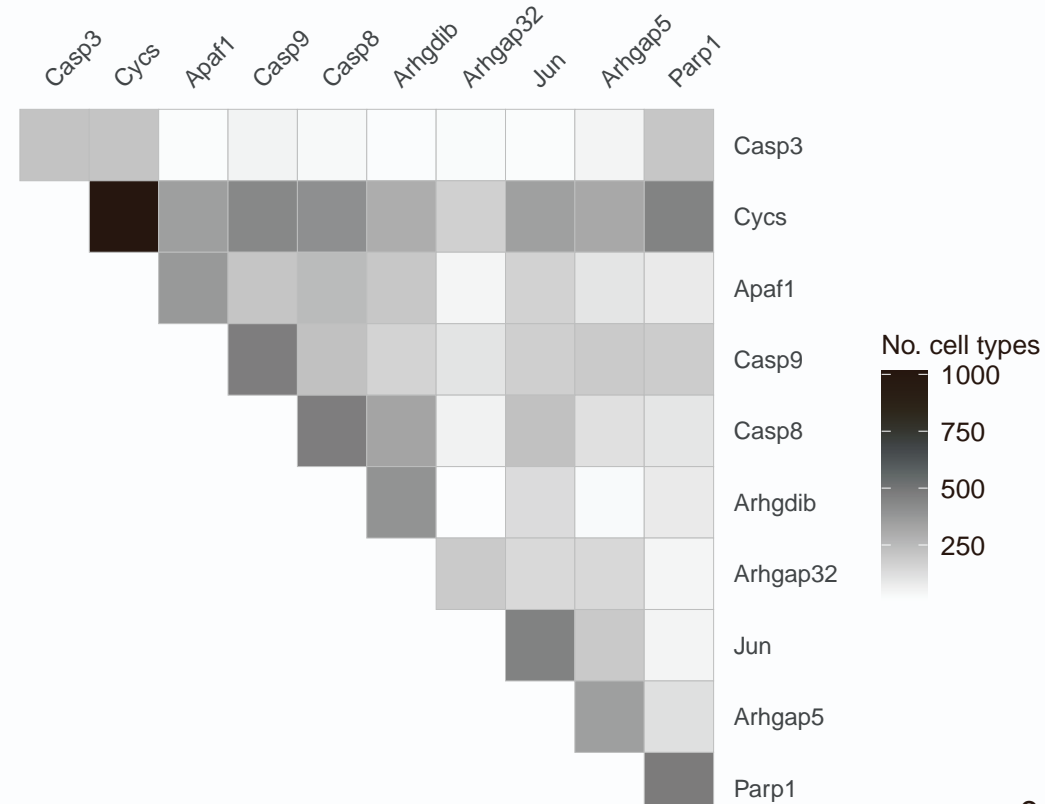
D4-GDI Signaling Pathway (k_opt = 77)

Casp3
Cycs
Apaf1
Casp9

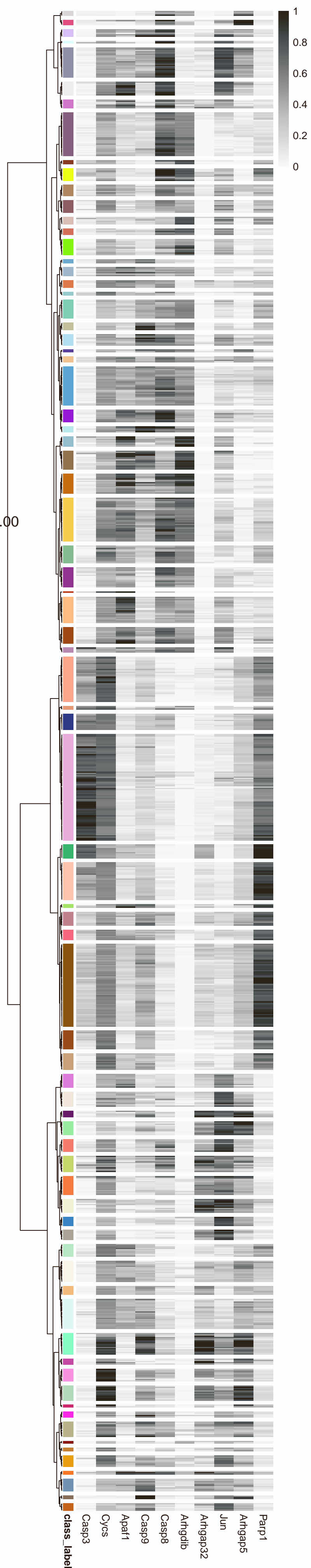
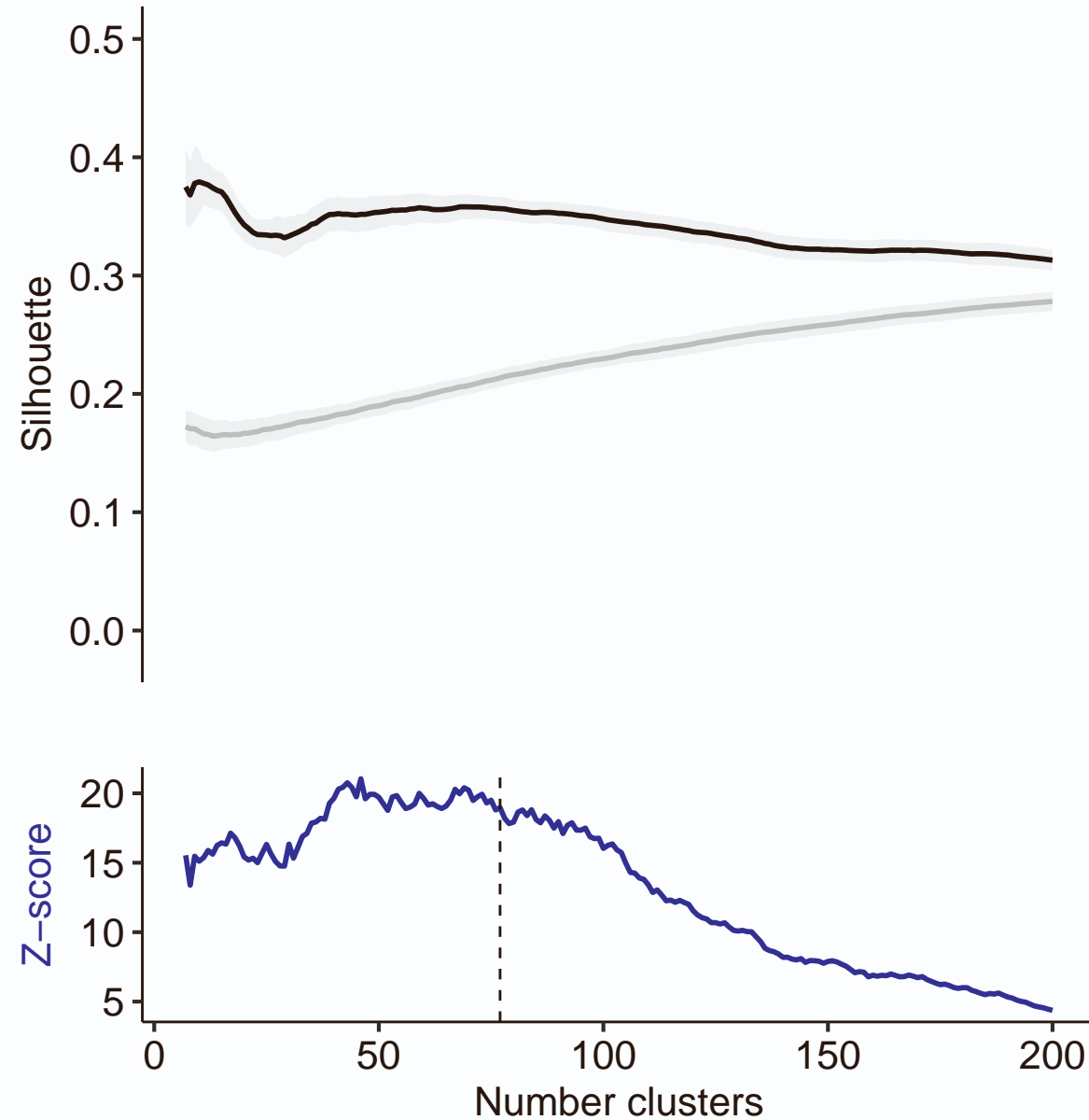
Pathway genes
Casp8
Arhgdib
Arhgap32

Jun
Arhgap5
Parp1

Gene Co-Expression Above Threshold



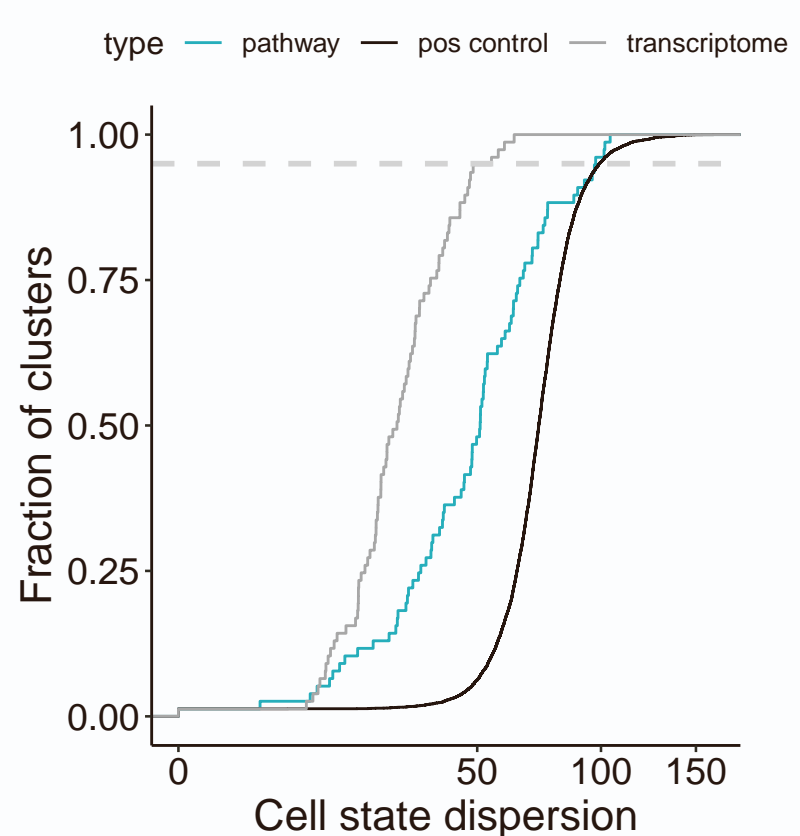
Silhouette and Z-score



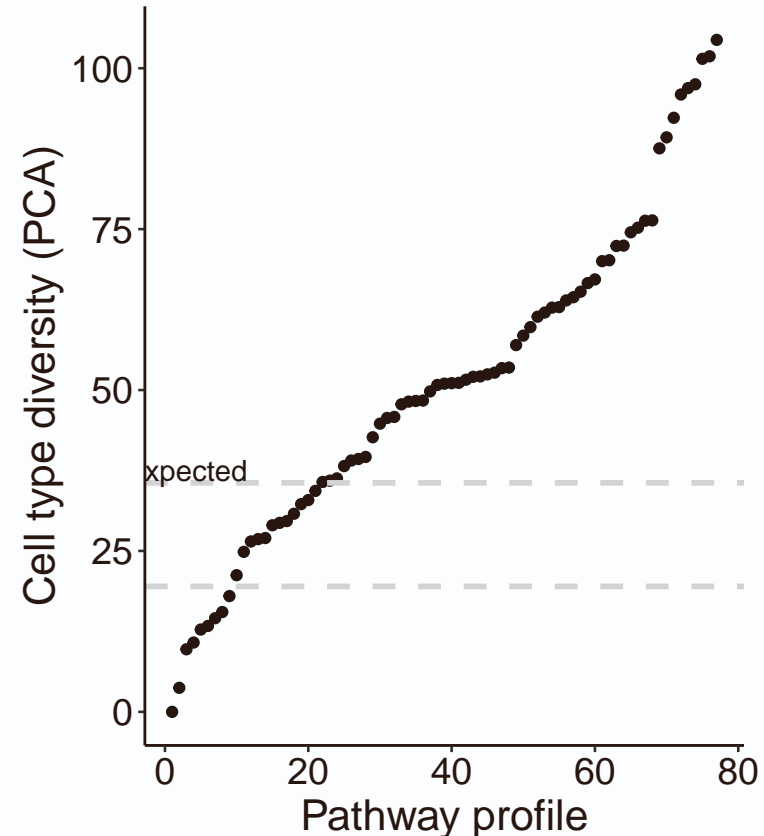
Global UMAP: Pathway ON



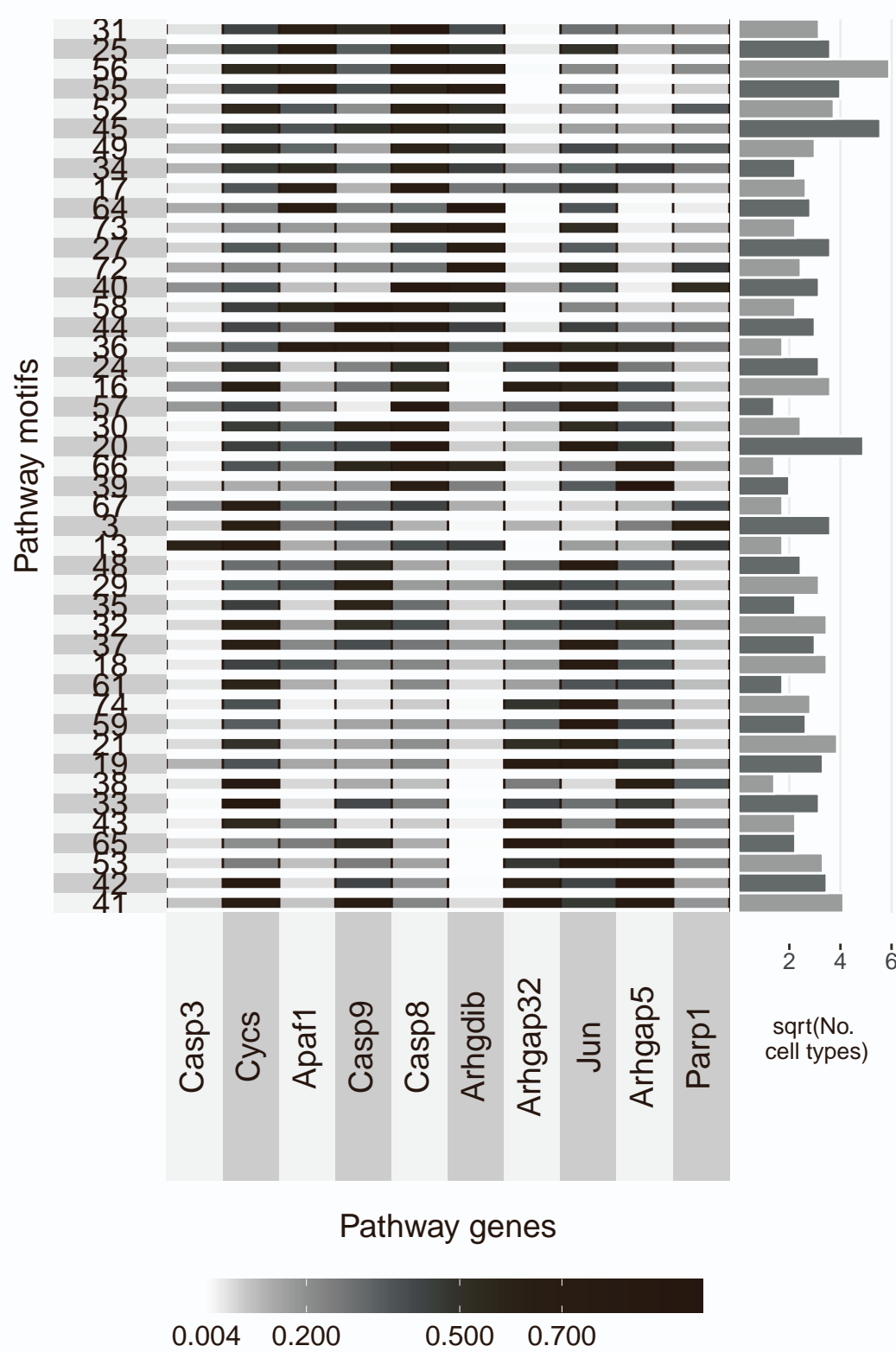
ECDF of Profile Dispersions



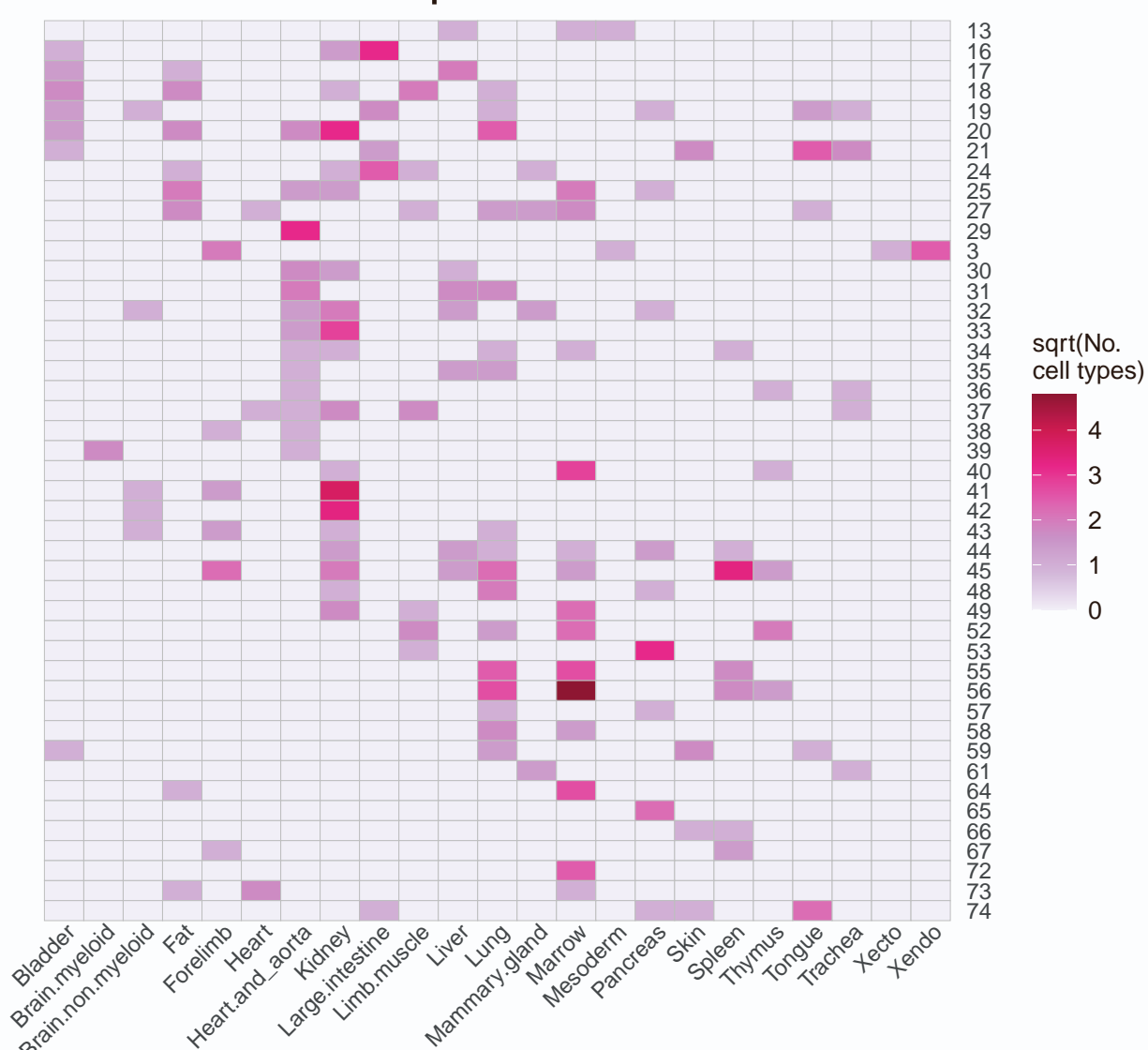
Profile Dispersion Plot



Motifs, Dispersion $\geq 90^{\text{th}}$ percentile



Motif Tissue Composition



D4-GDI Signaling Pathway (k_opt = 77)



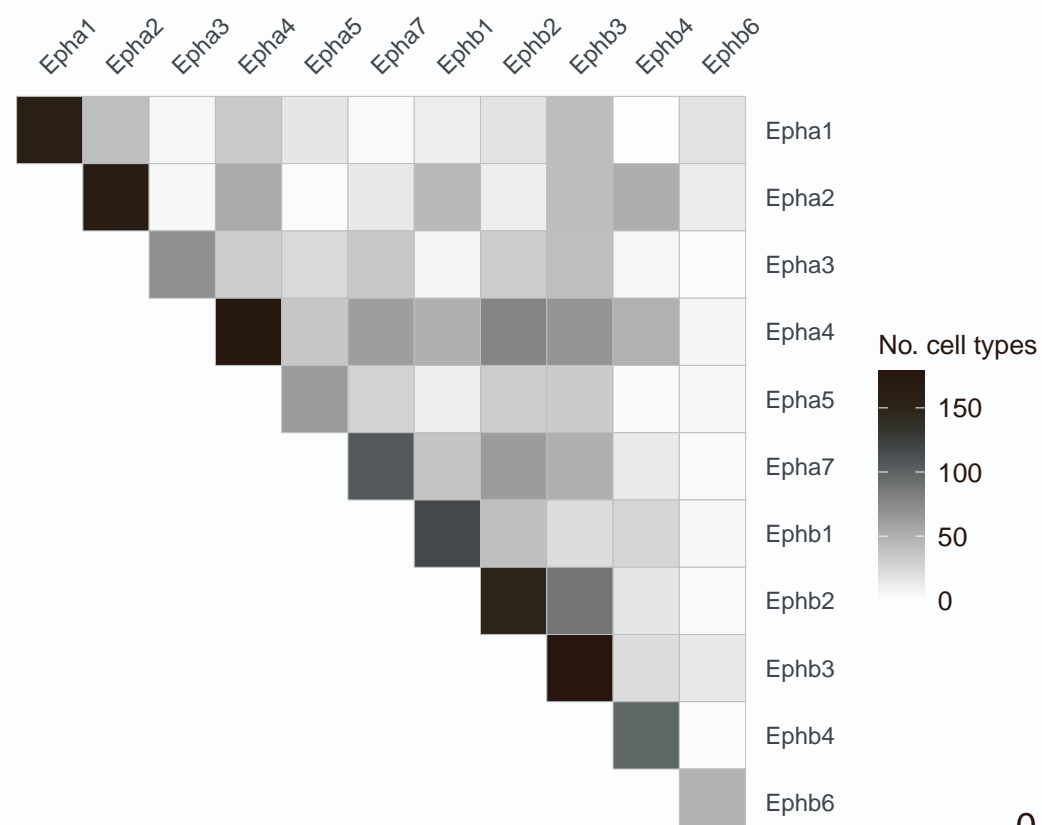
Eph A-B receptors (k_opt = 31)

Epha1
Epha2
Epha3
Epha4

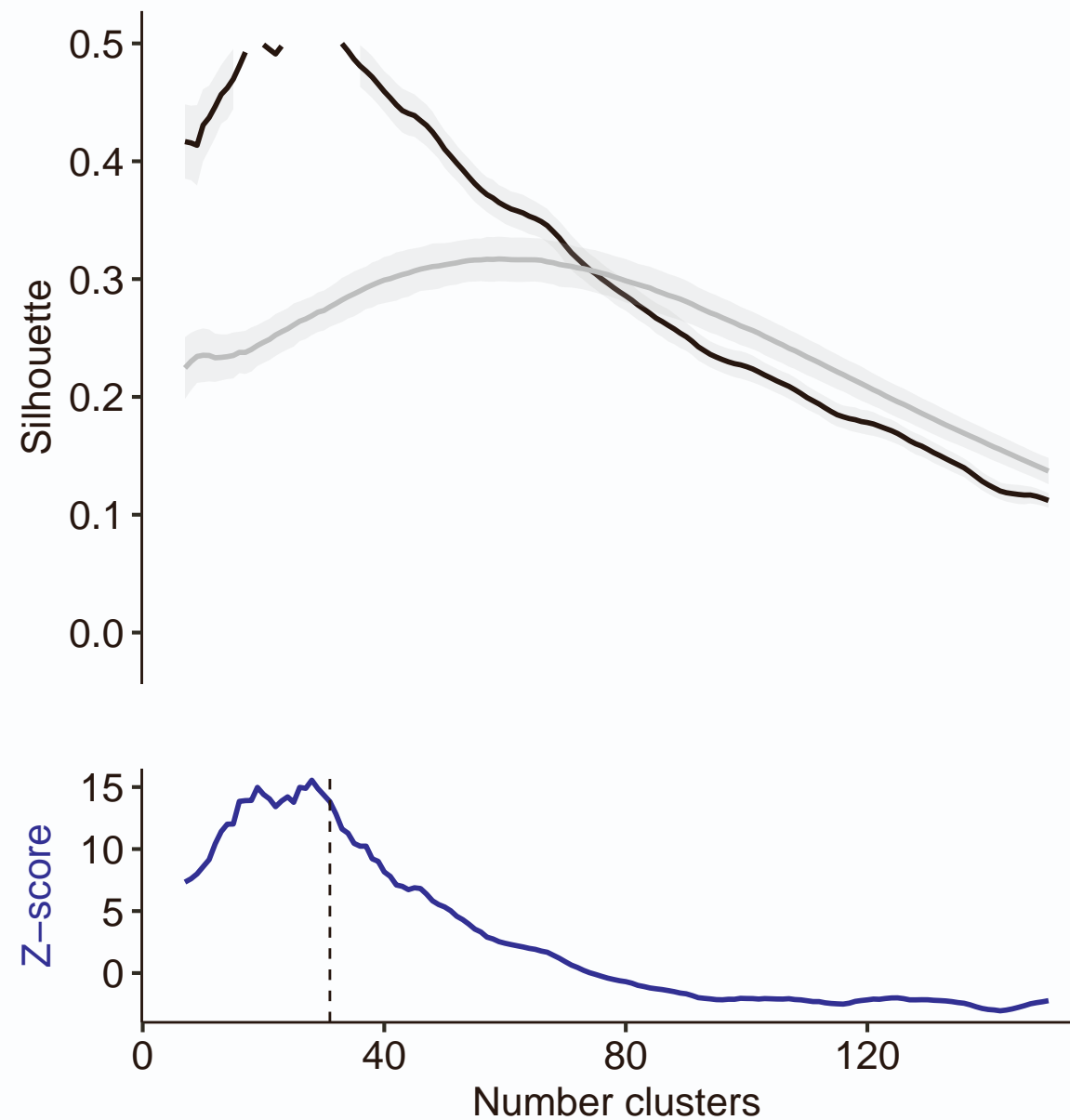
Pathway genes
Epha5
Epha7
Ephb1

Ephb2
Ephb3
Ephb4
Ephb6

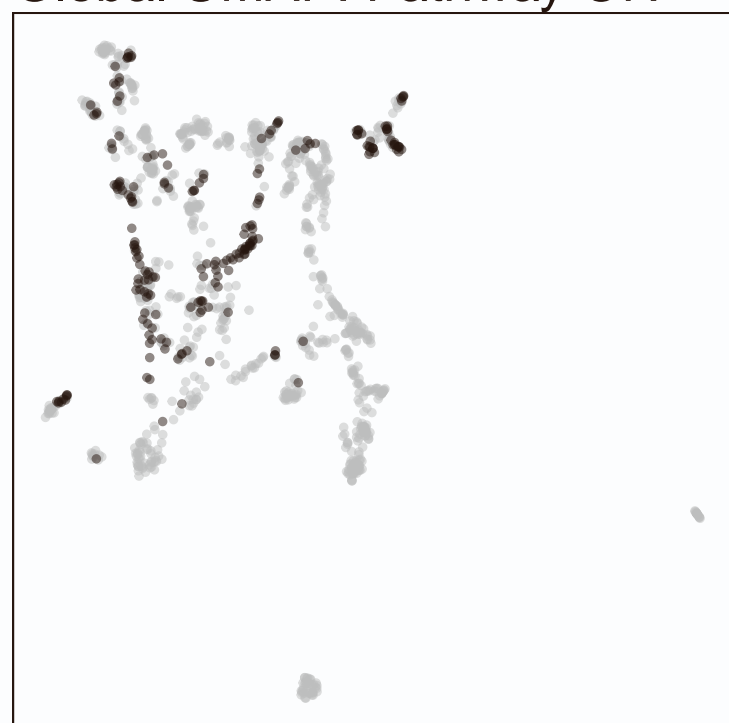
Gene Co-Expression Above Threshold



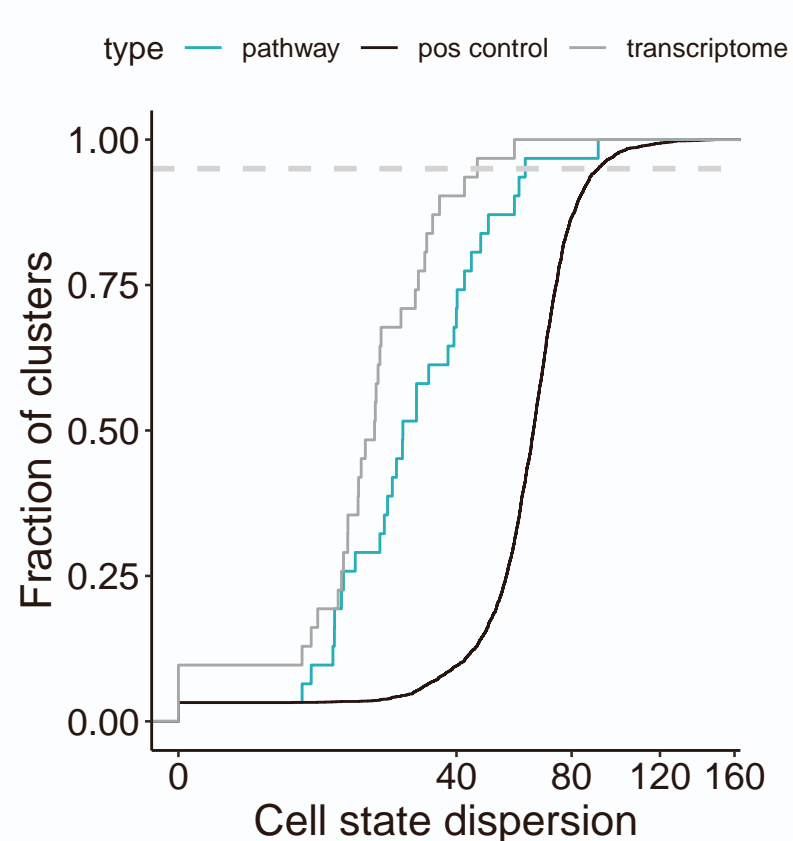
Silhouette and Z-score



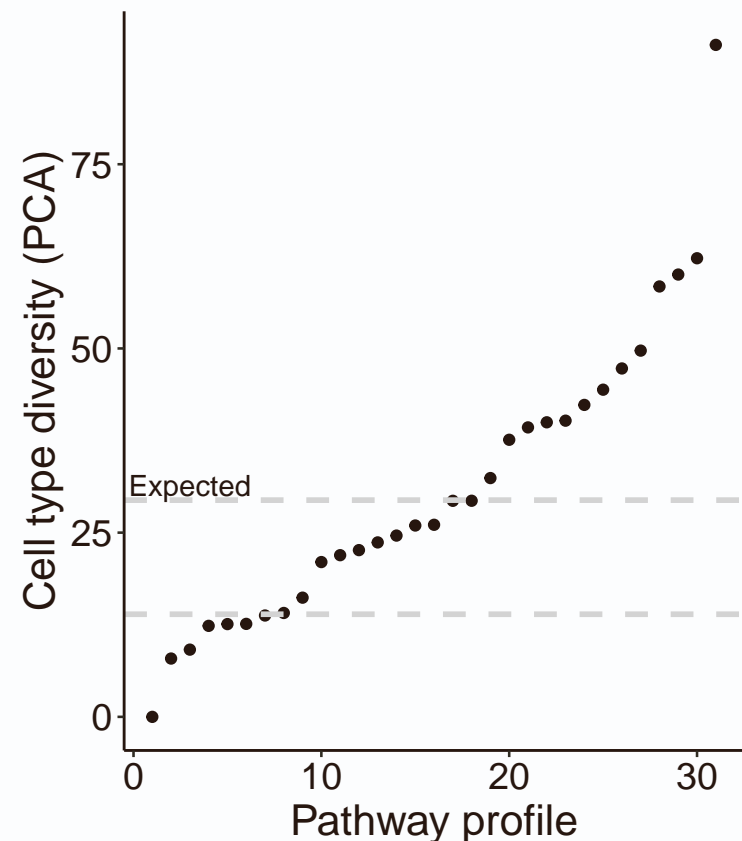
Global UMAP: Pathway ON



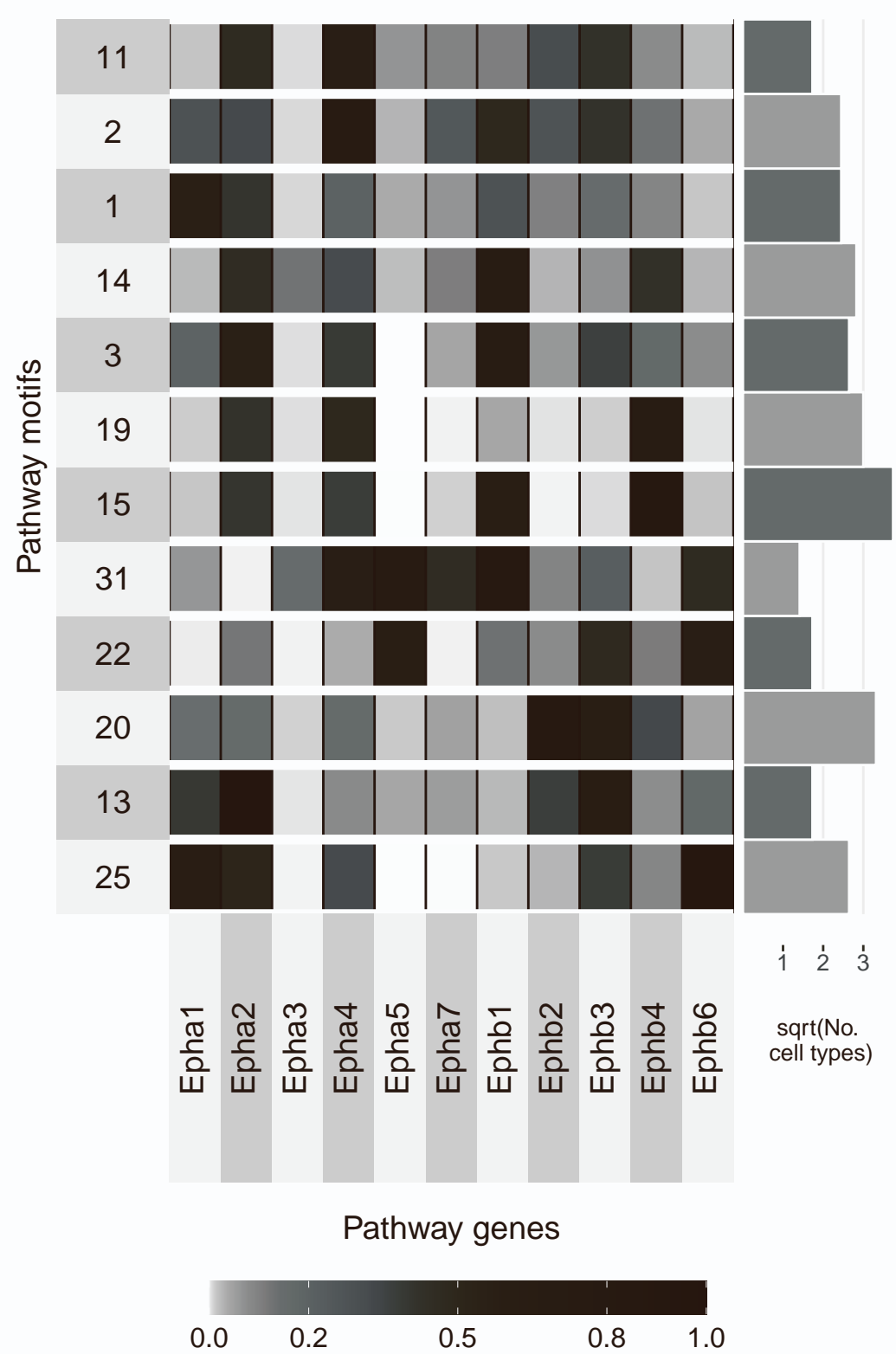
ECDF of Profile Dispersions



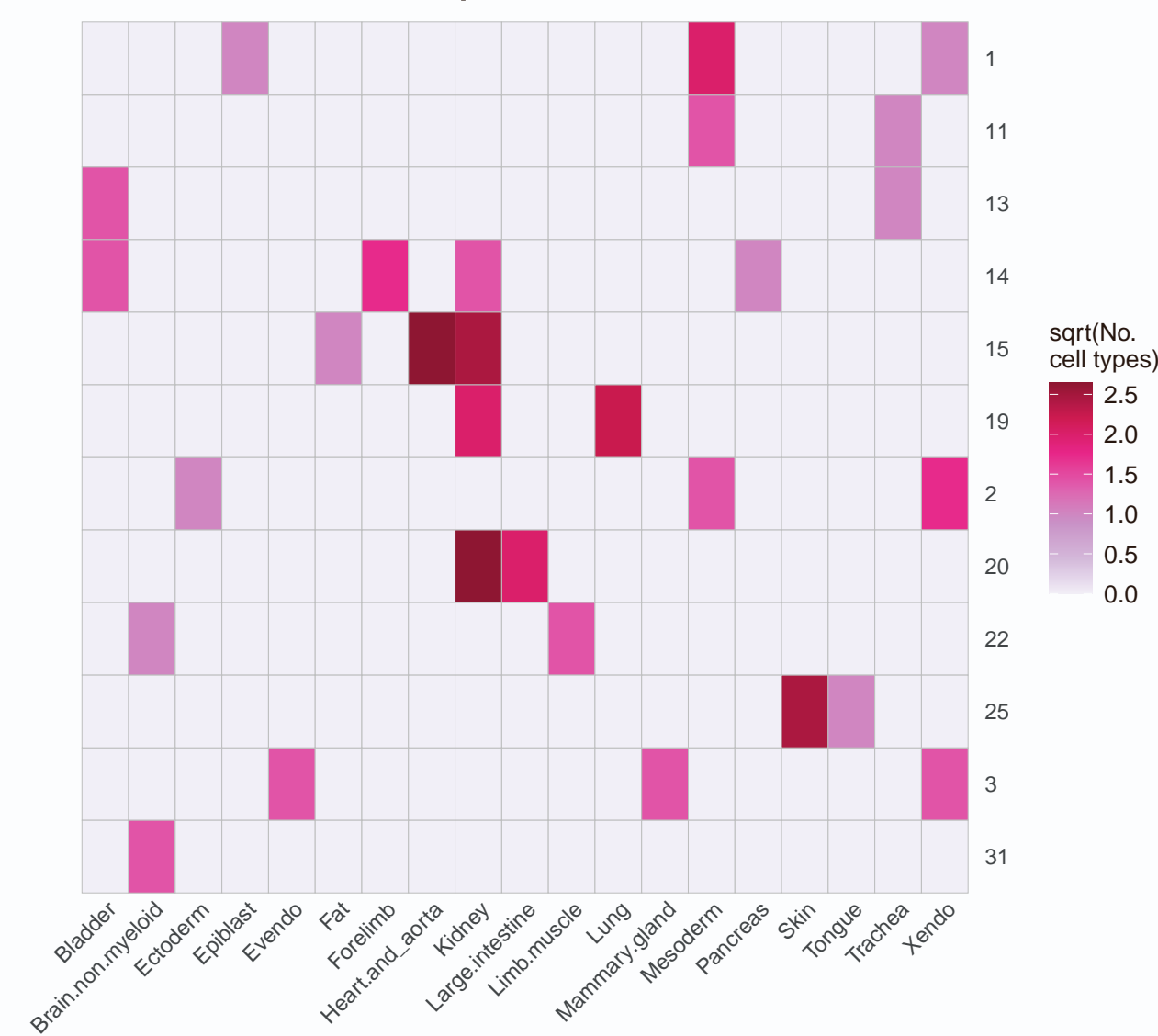
Profile Dispersion Plot



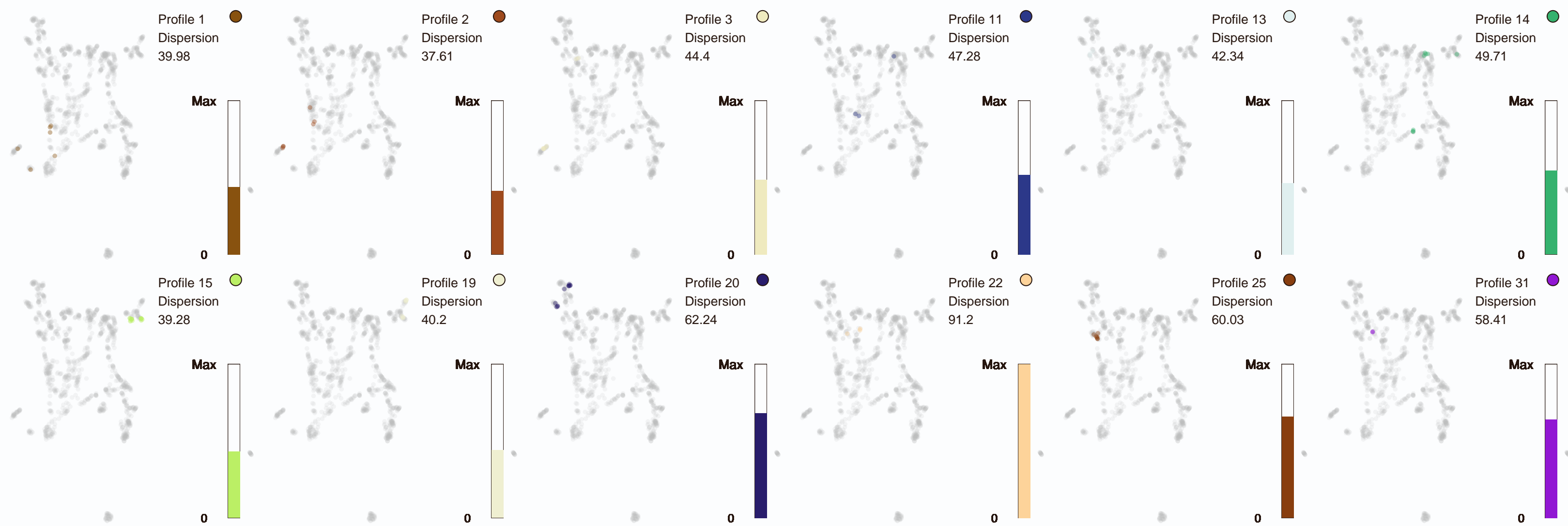
Motifs, Dispersion $\geq 90^{\text{th}}$ percentile



Motif Tissue Composition



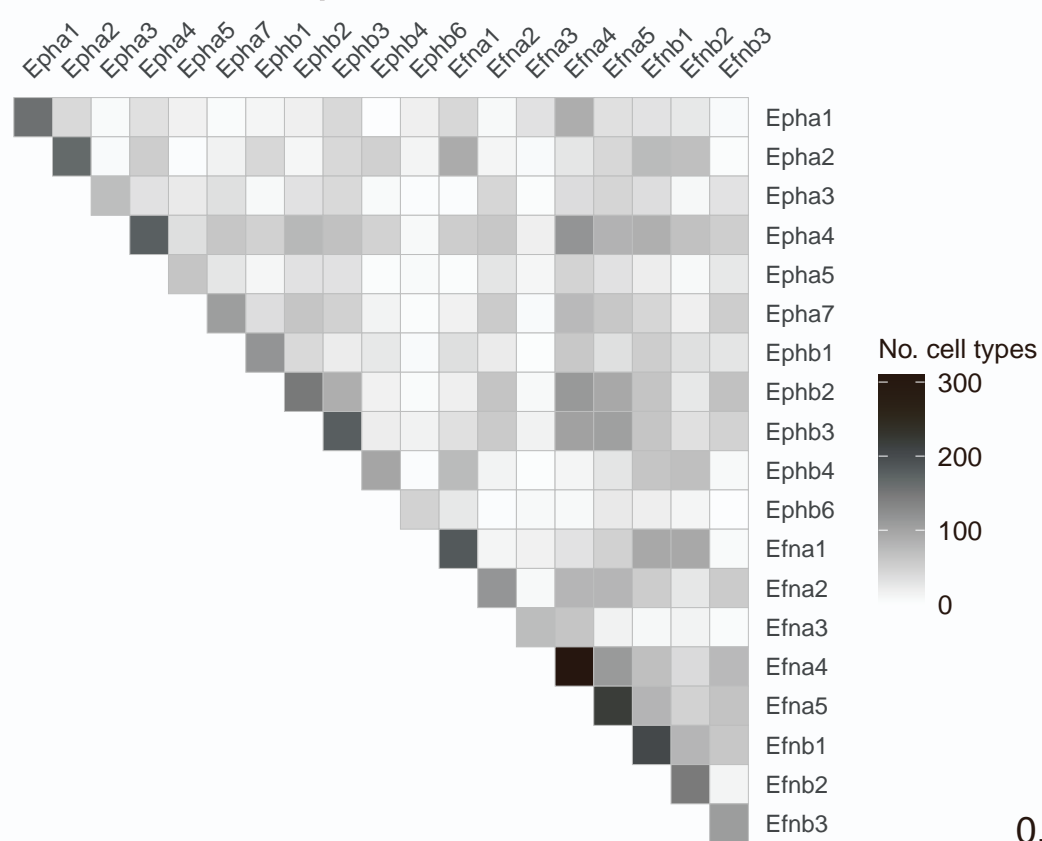
Eph A–B receptors (k_opt = 31)



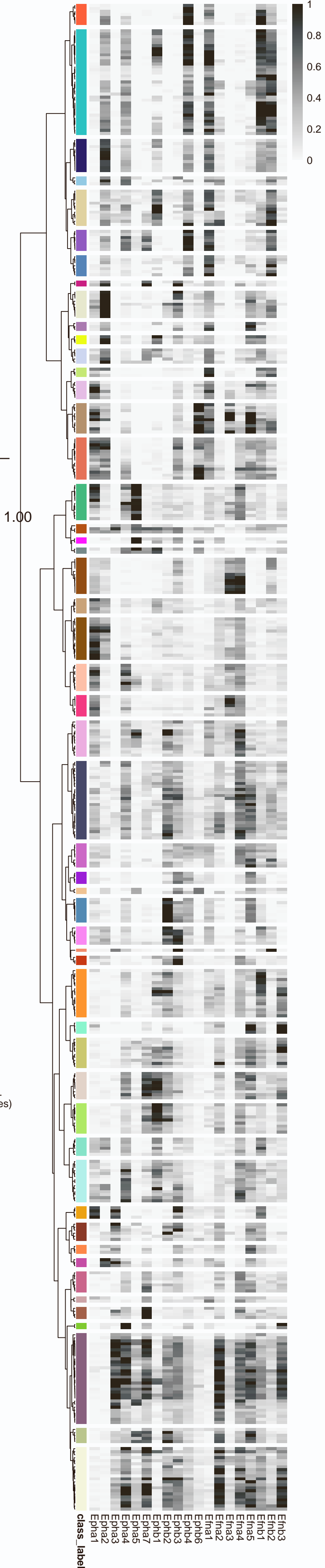
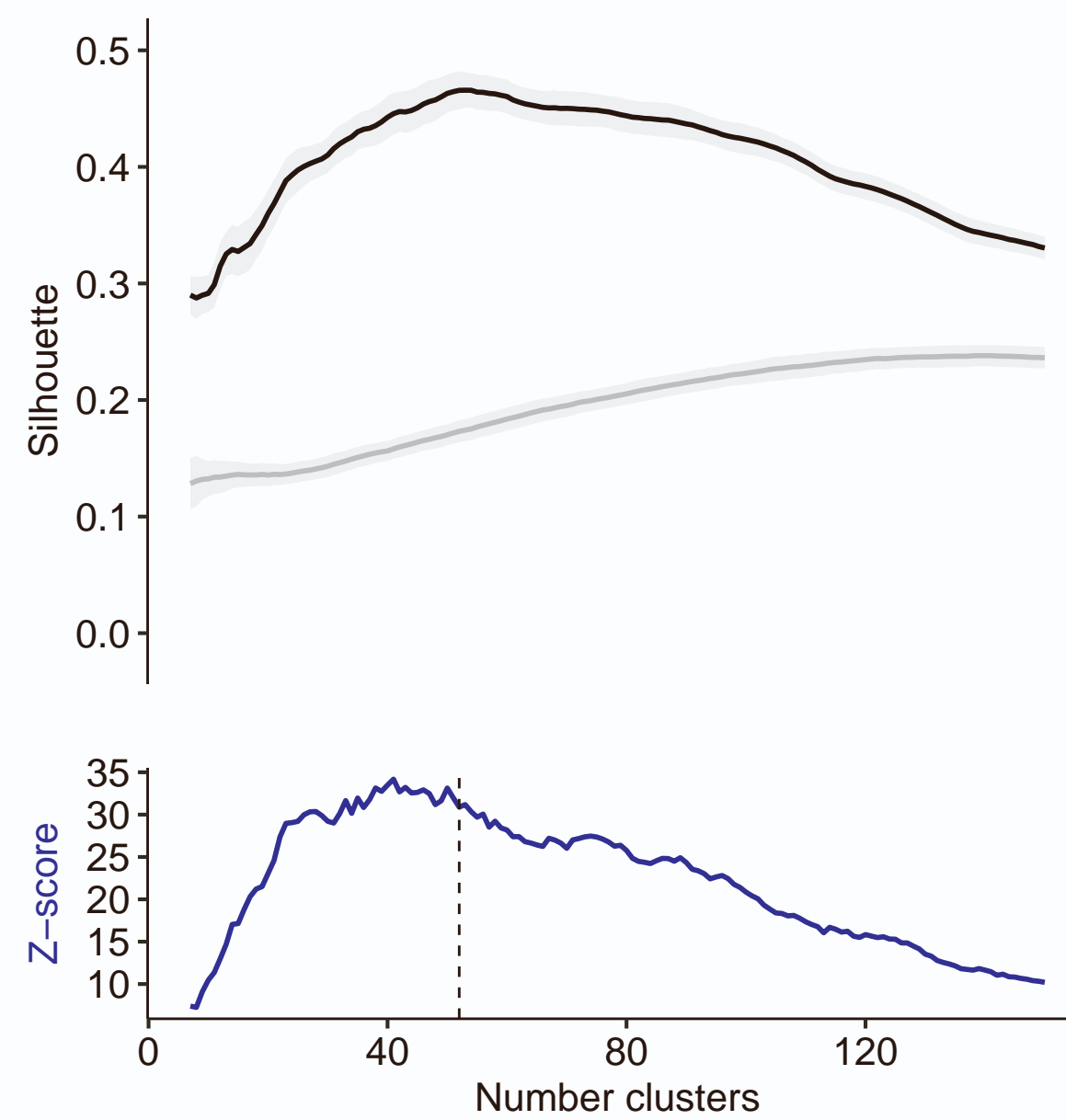
Eph receptors and ligands (k_opt = 52)

Epha1
 Epha2
 Epha3
 Epha4
 Epha5
 Epha7
 Ephb1
 Ephb2
 Ephb3
 Ephb4
 Ephb6
 Efna1
 Efna2
 Efna3
 Efna4
 Efna5
 Efnb1
 Efnb2
 Efnb3

Gene Co-Expression Above Threshold



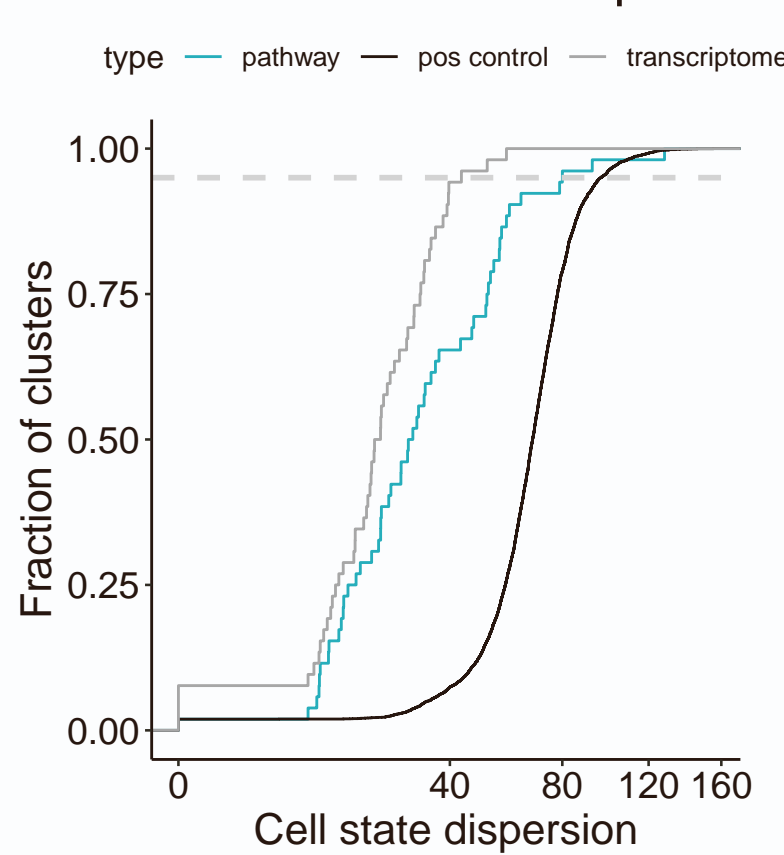
Silhouette and Z-score



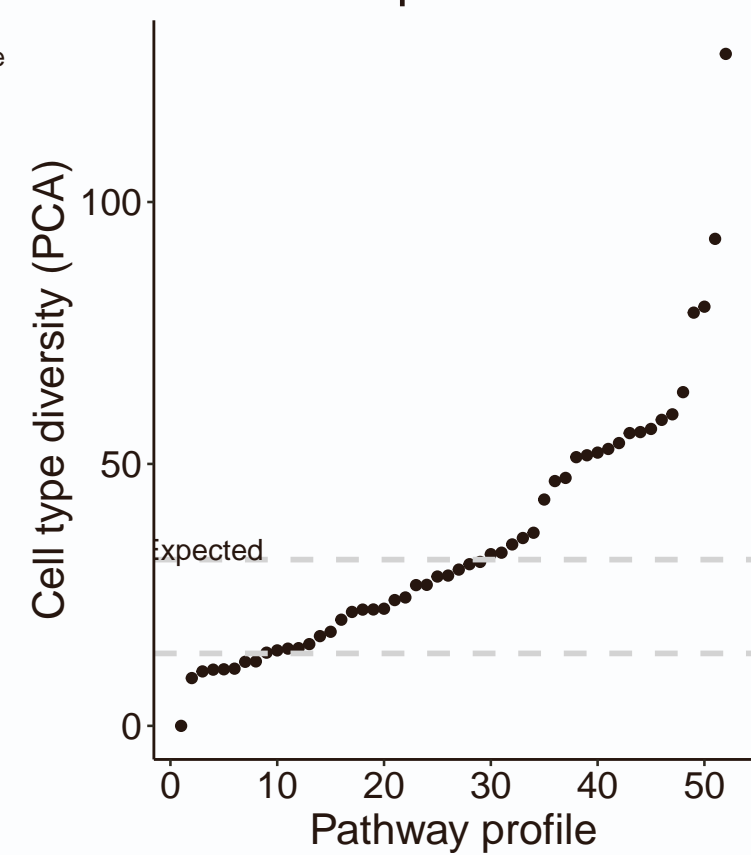
Global UMAP: Pathway ON



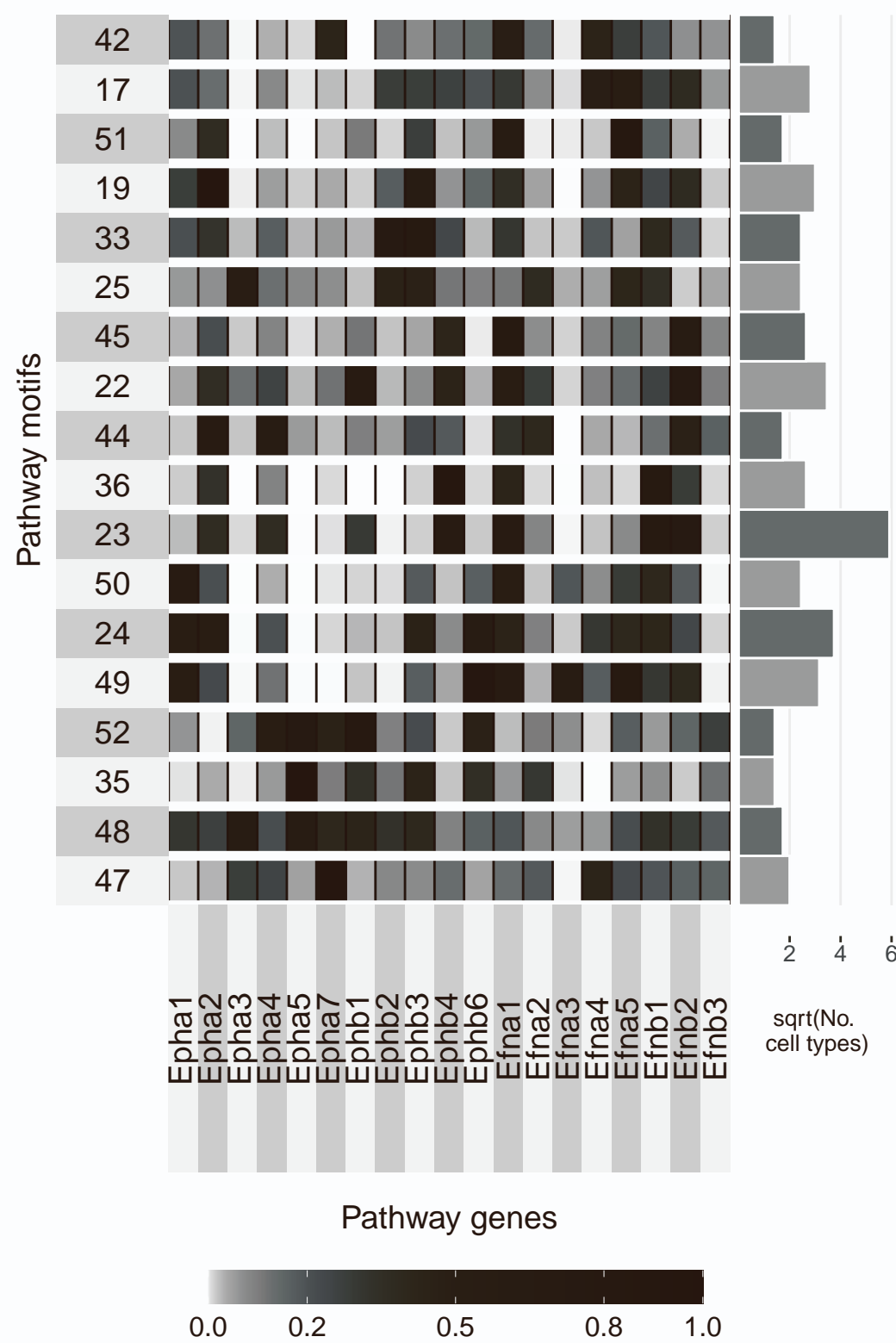
ECDF of Profile Dispersions



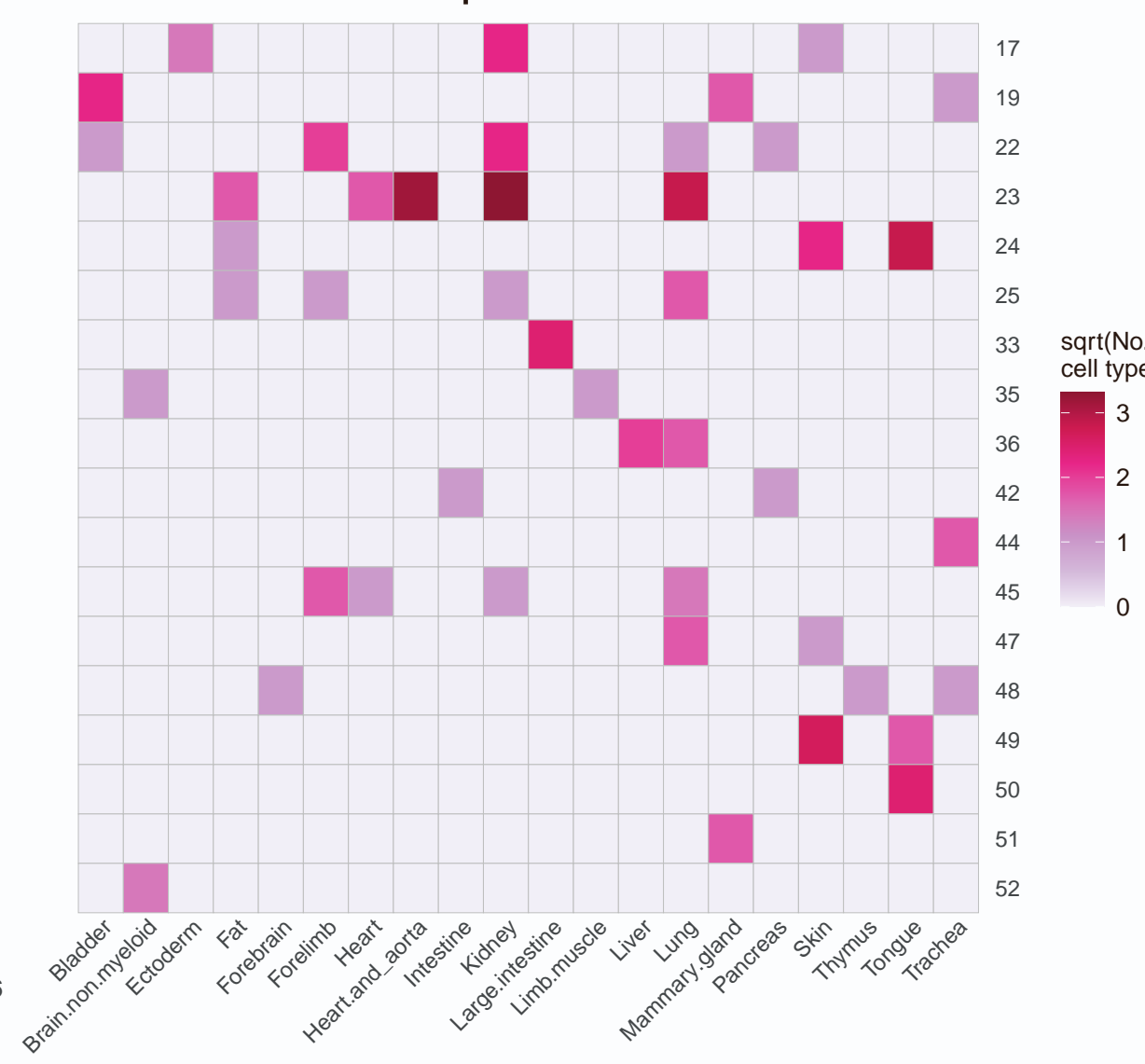
Profile Dispersion Plot



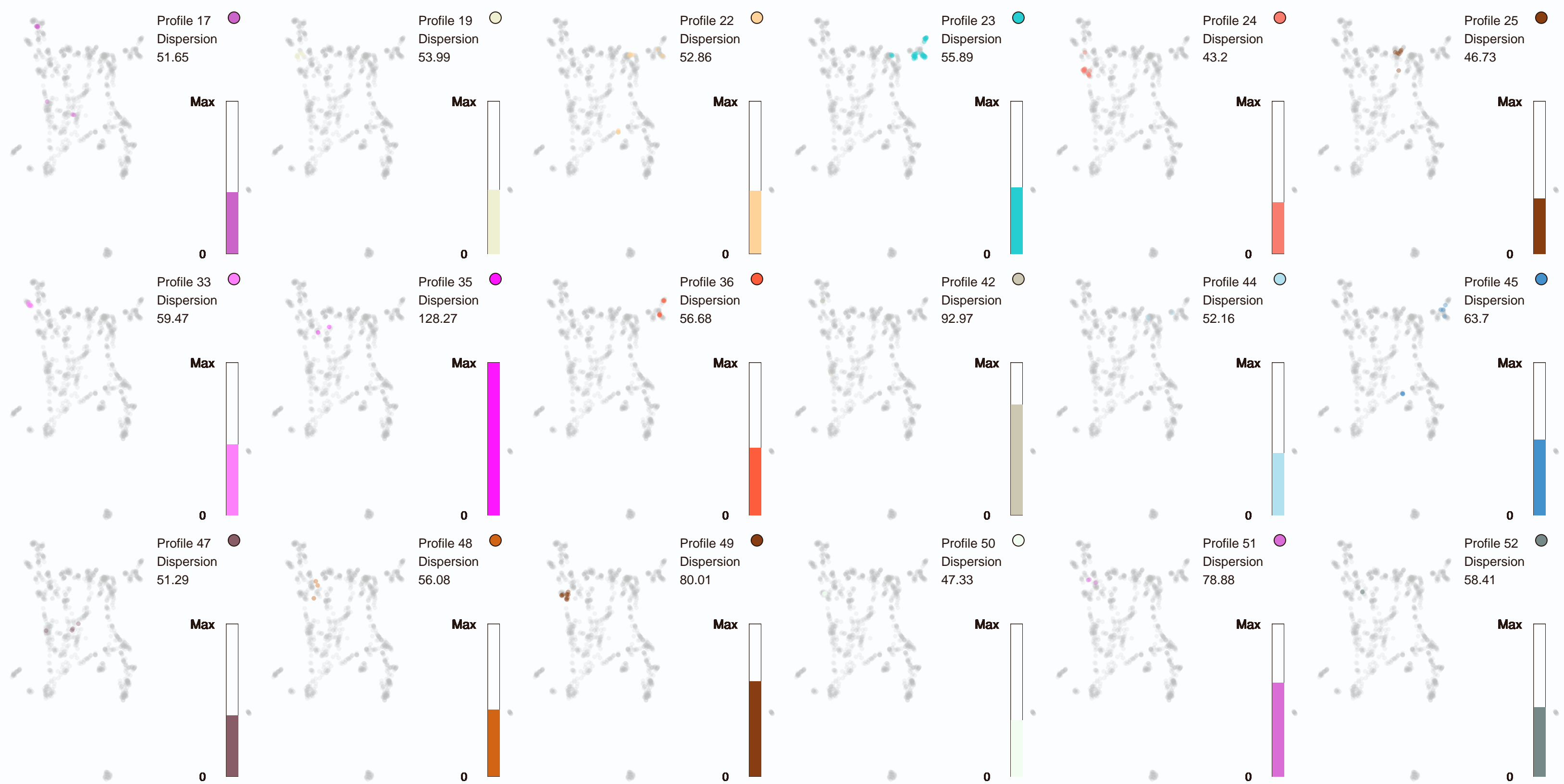
Motifs, Dispersion $\geq 90^{\text{th}}$ percentile



Motif Tissue Composition



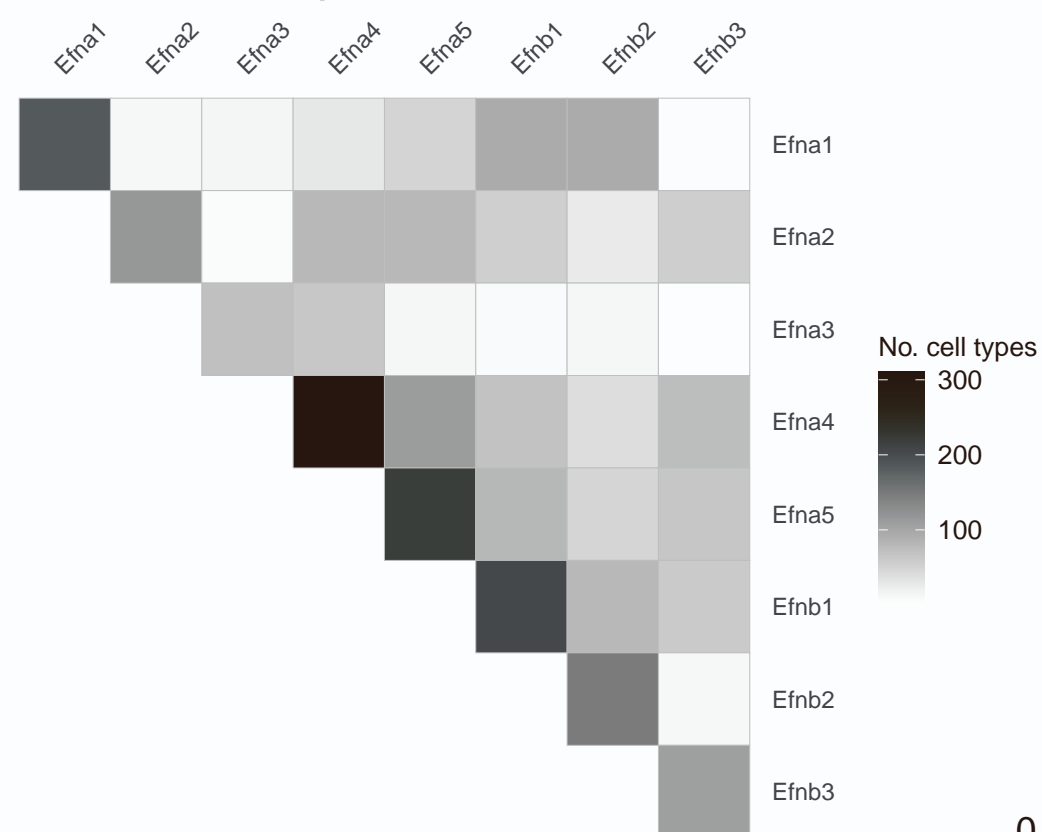
Eph receptors and ligands (k_opt = 52)



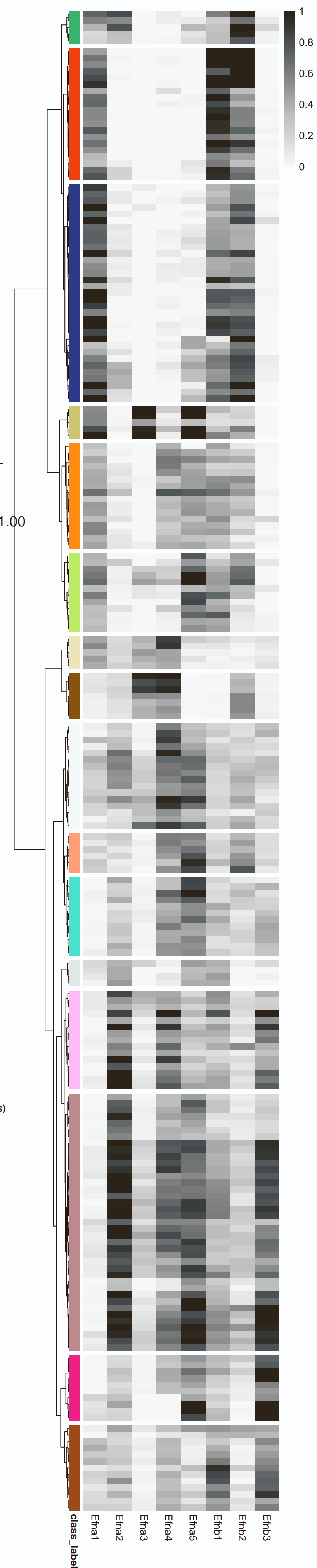
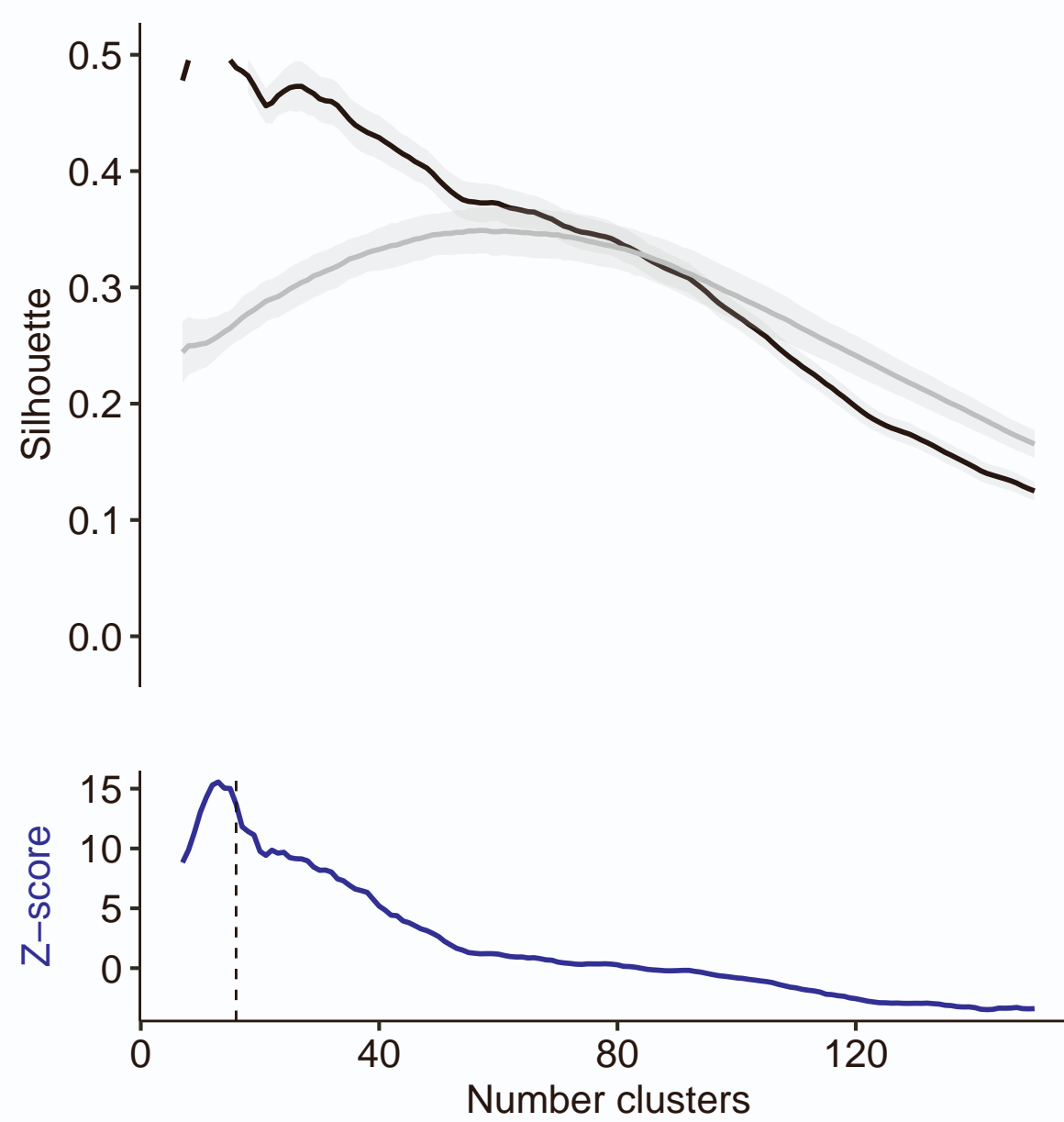
Eph_I (k_opt = 16)

Pathway genes
 Efna1 Efna2 Efna3 Efna4 Efna5 Efnb1 Efnb2 Efnb3

Gene Co-Expression Above Threshold



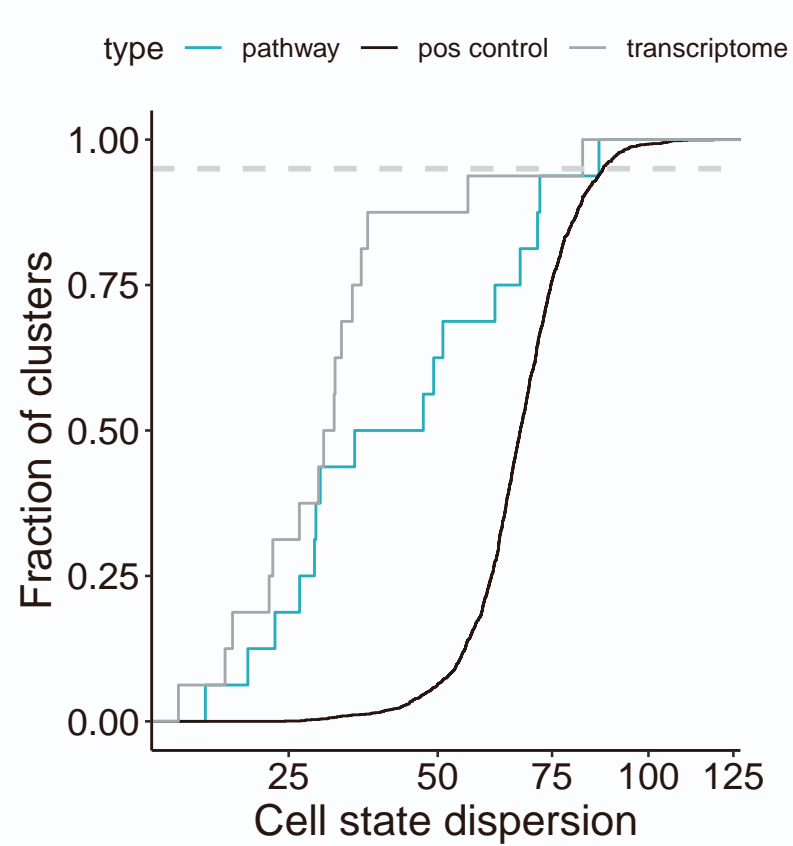
Silhouette and Z-score



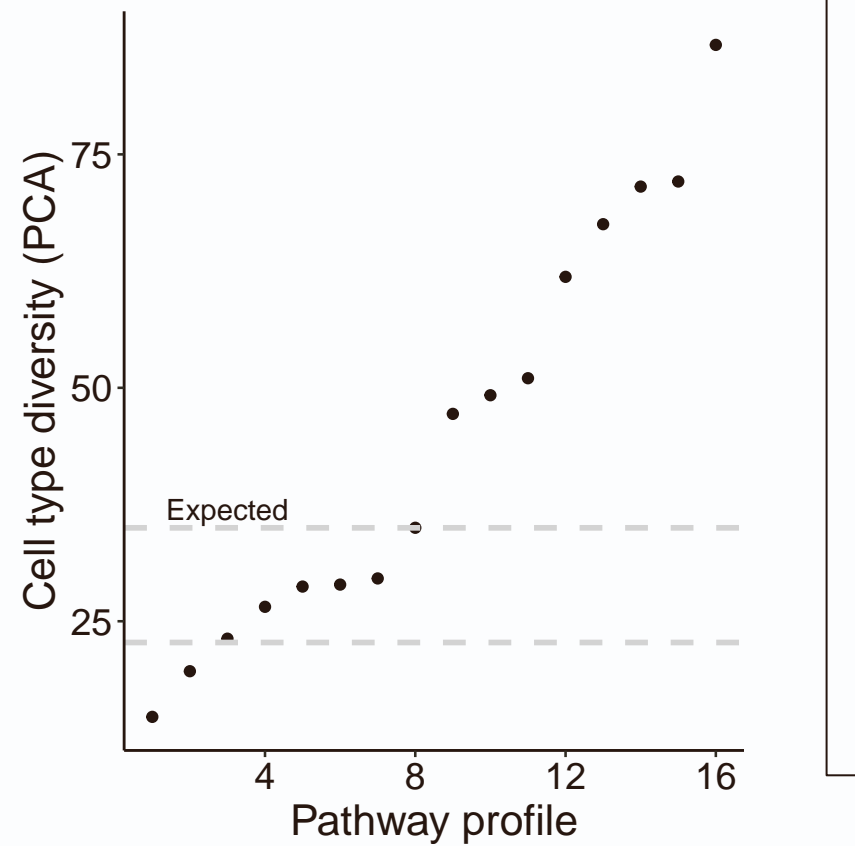
Global UMAP: Pathway ON



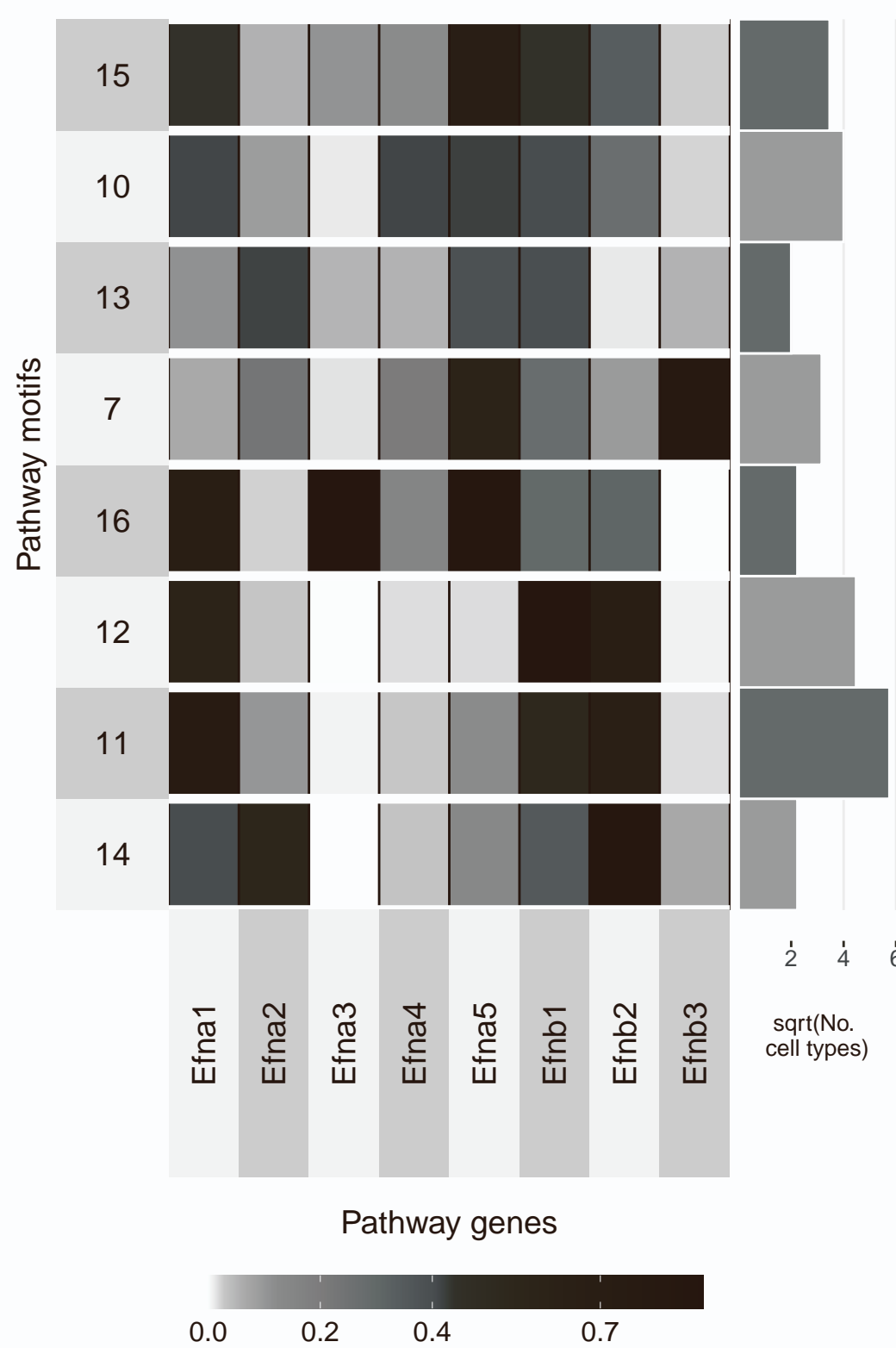
ECDF of Profile Dispersions



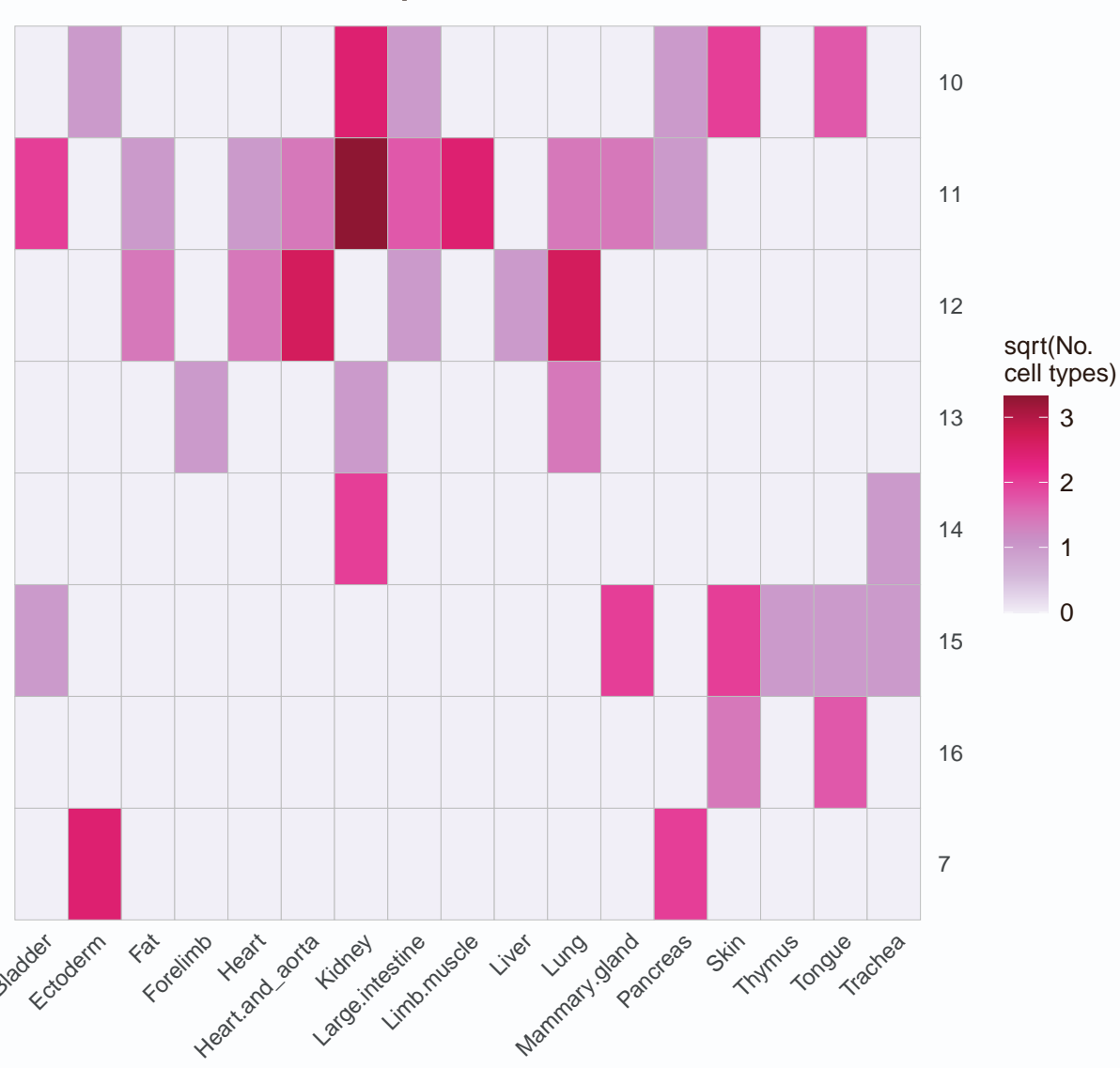
Profile Dispersion Plot



Motifs, Dispersion $\geq 90^{\text{th}}$ percentile

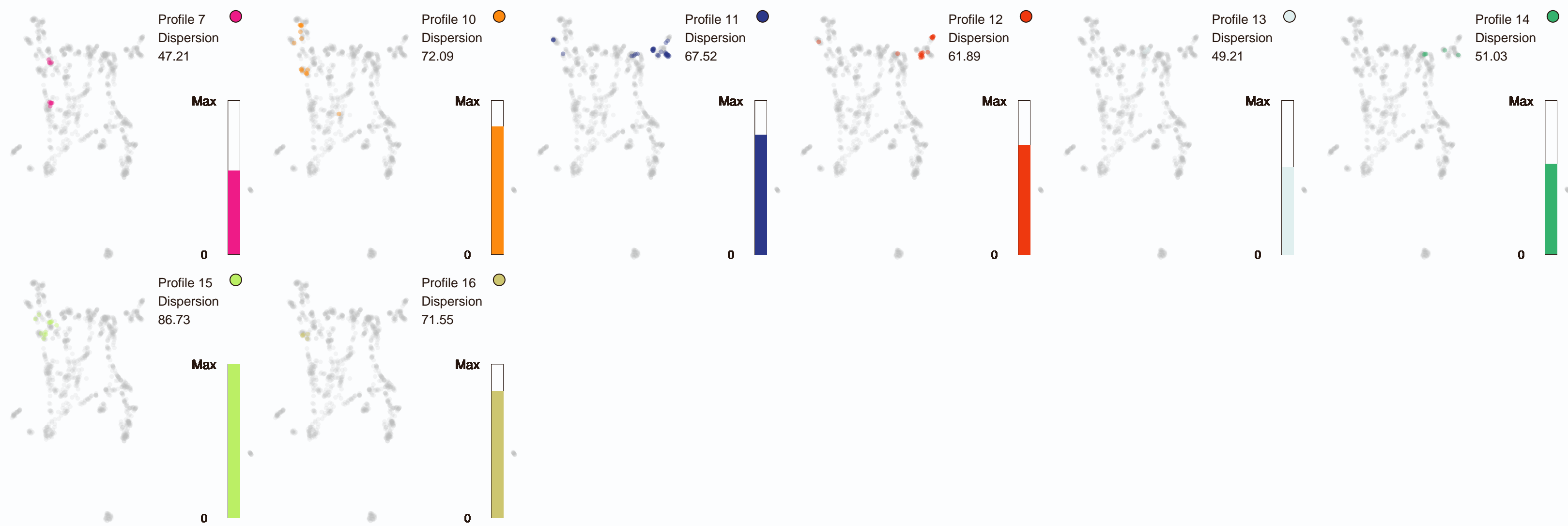


Motif Tissue Composition



class_label
 Efna1 Efna2 Efna3 Efna4 Efna5 Efnb1 Efnb2 Efnb3

Eph_I (k_opt = 16)



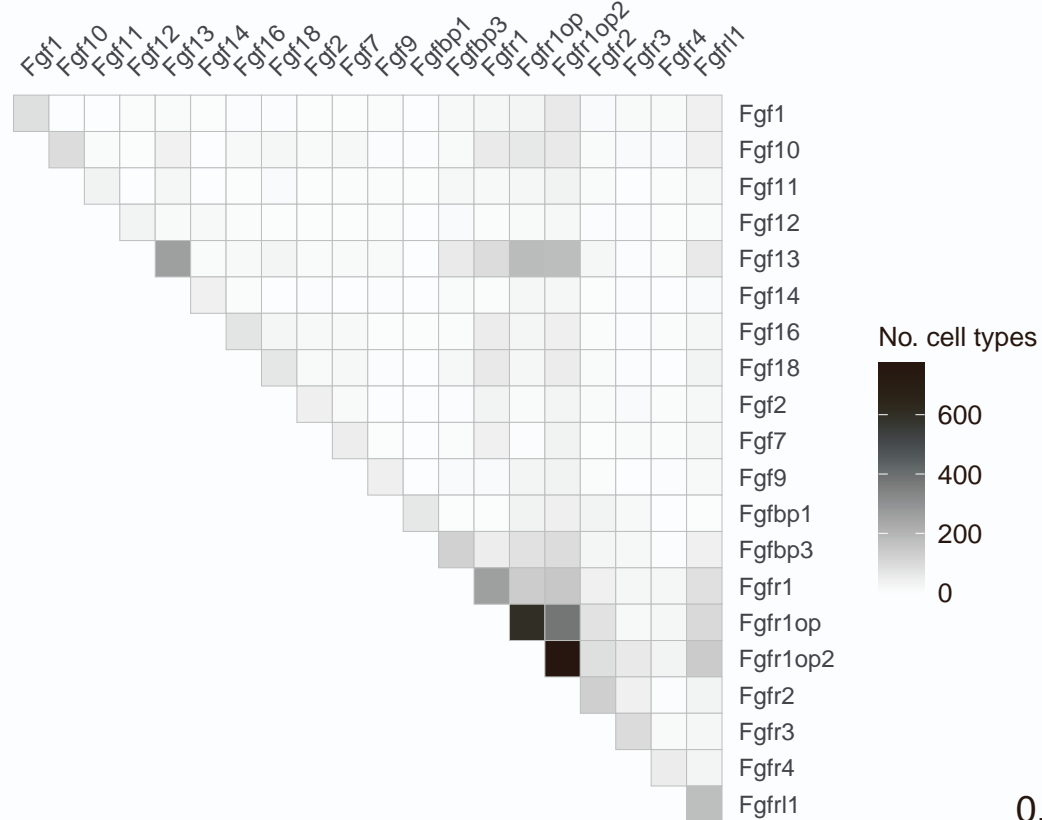
FGF cell signaling proteins (k_opt = 80)

Fgf1
 Fgf10
 Fgf11
 Fgf12
 Fgf13
 Fgf14
 Fgf16

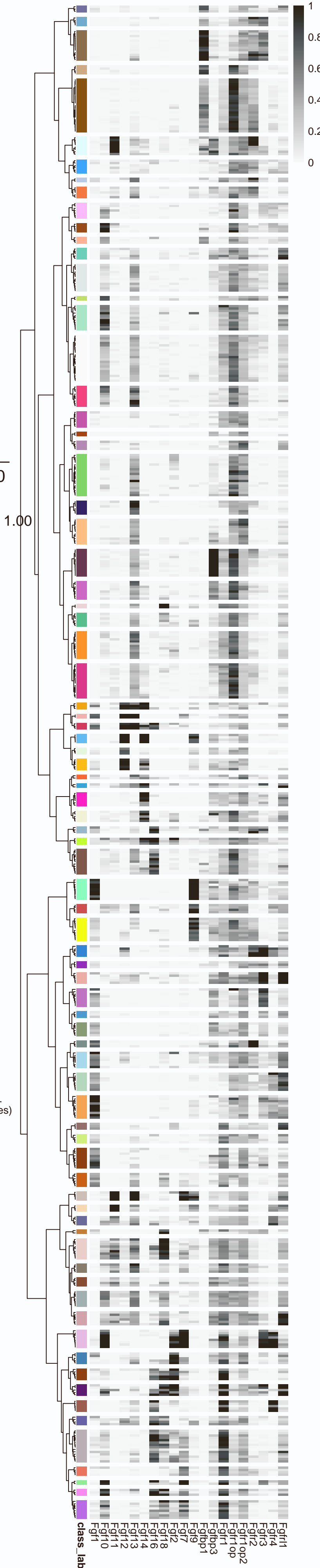
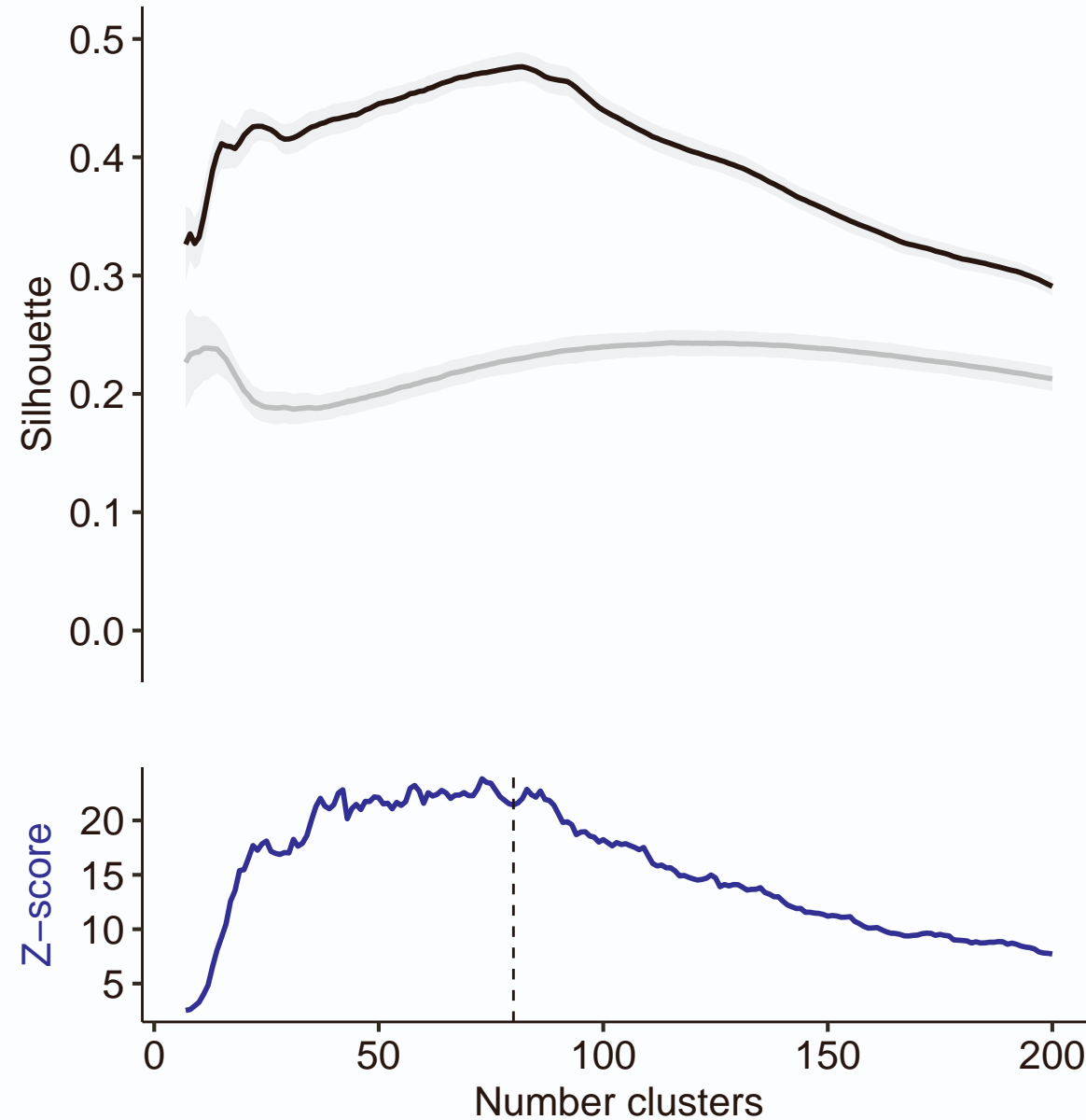
Pathway genes
 Fgf18
 Fgf2
 Fgf7
 Fgf9
 Fgfbp1
 Fgfbp3

Fgfr1
 Fgfr1op
 Fgfr1op2
 Fgfr2
 Fgfr3
 Fgfr4
 Fgfrl1

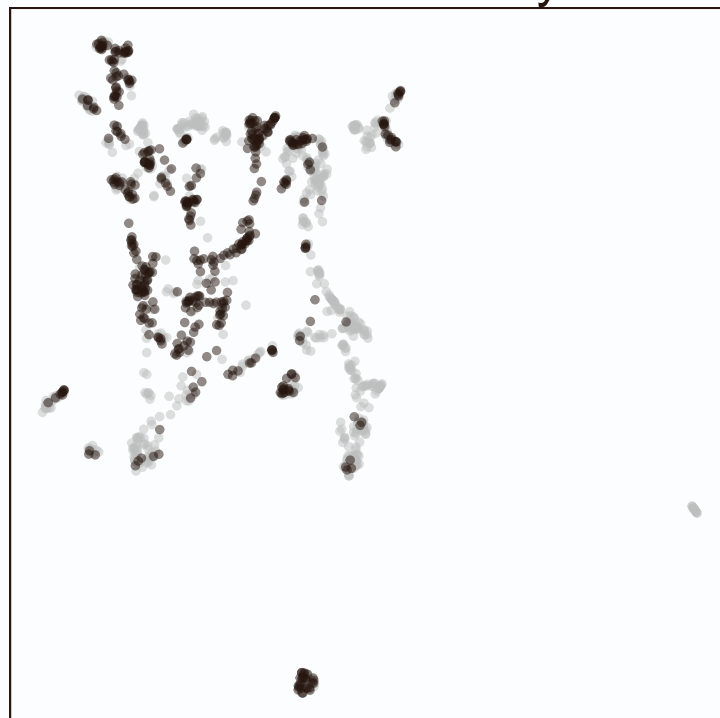
Gene Co-Expression Above Threshold



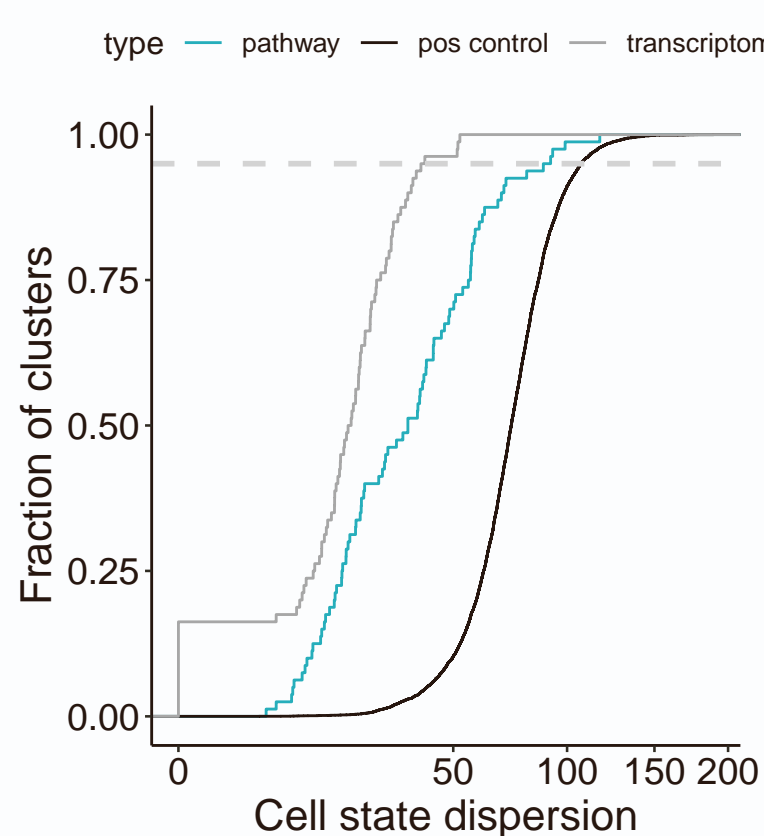
Silhouette and Z-score



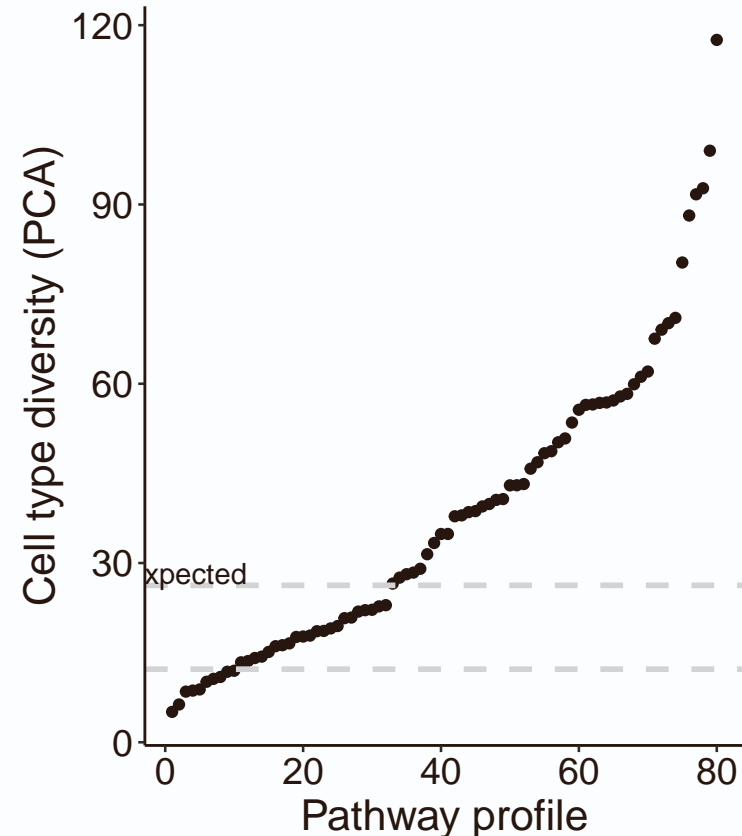
Global UMAP: Pathway ON



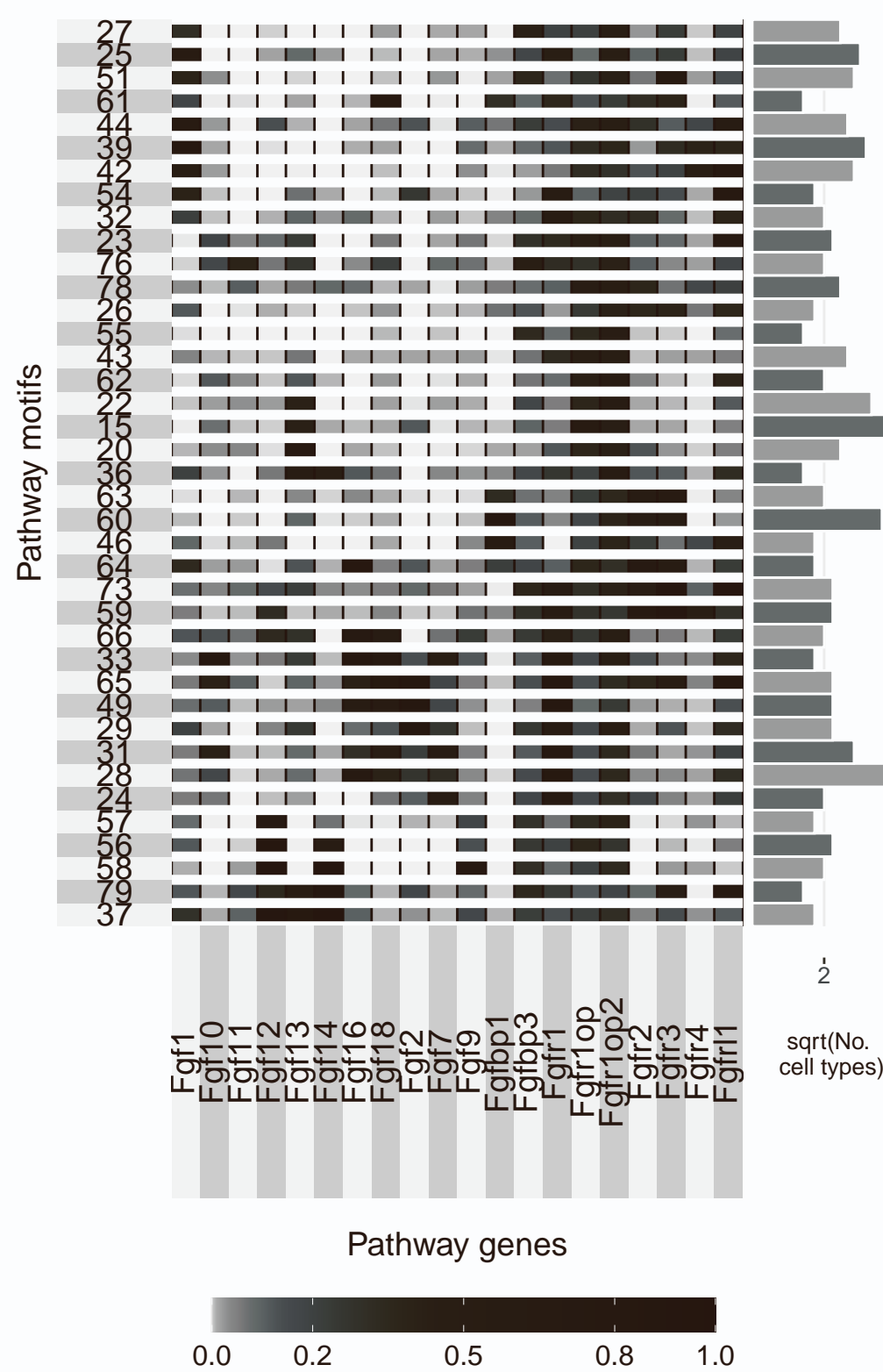
ECDF of Profile Dispersions



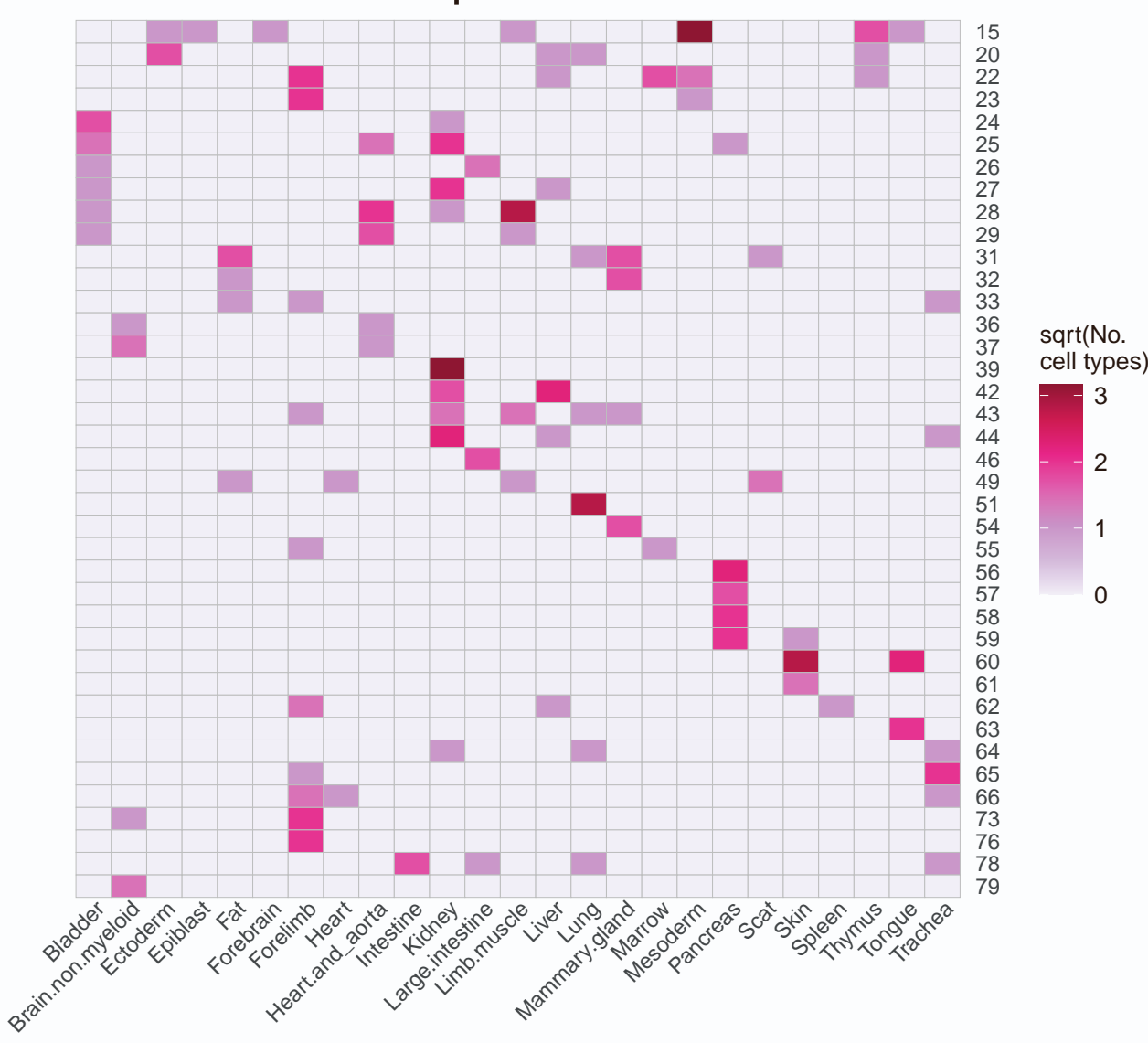
Profile Dispersion Plot



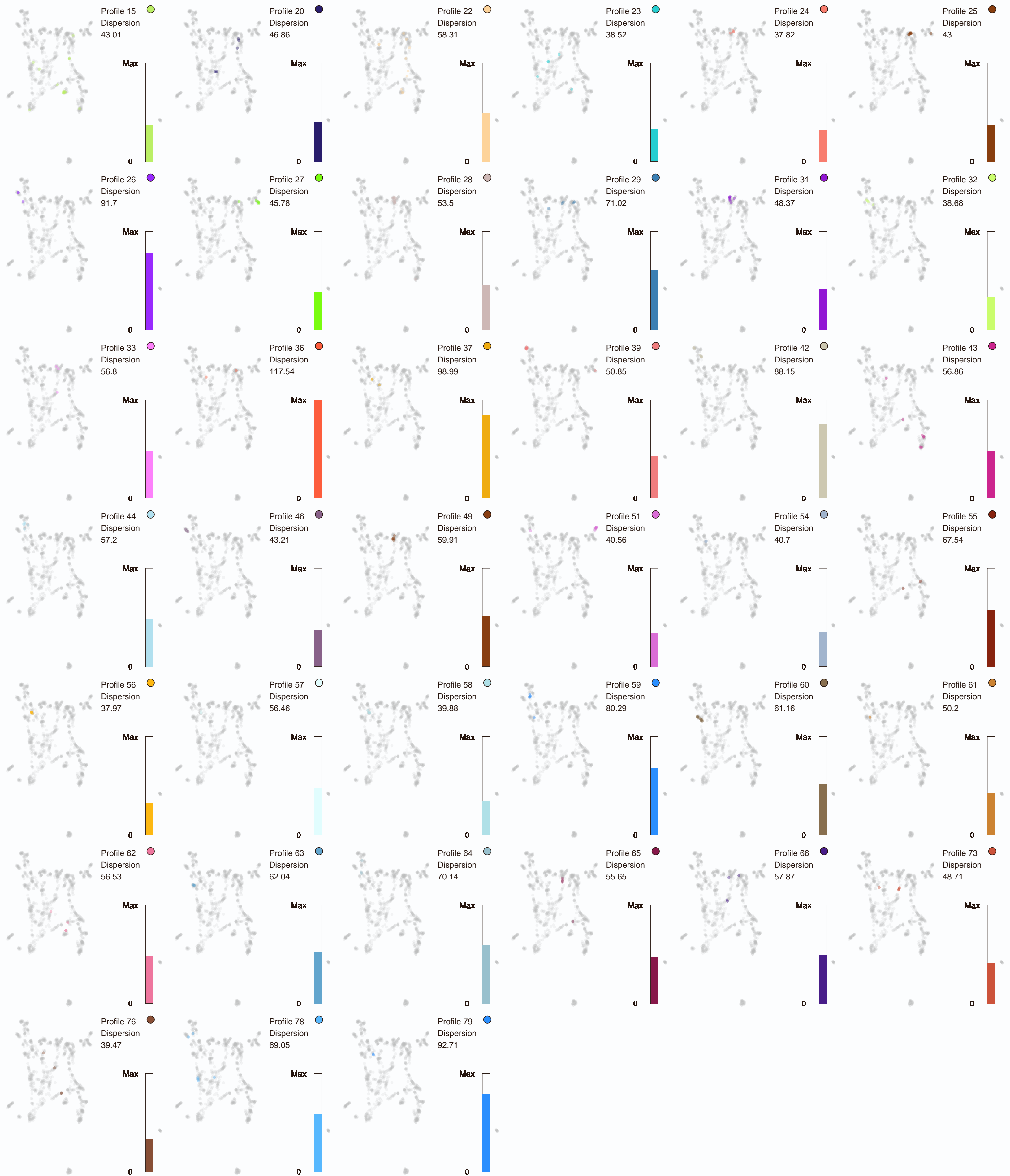
Motifs, Dispersion $\geq 90^{\text{th}}$ percentile



Motif Tissue Composition



FGF cell signaling proteins (k_opt = 80)

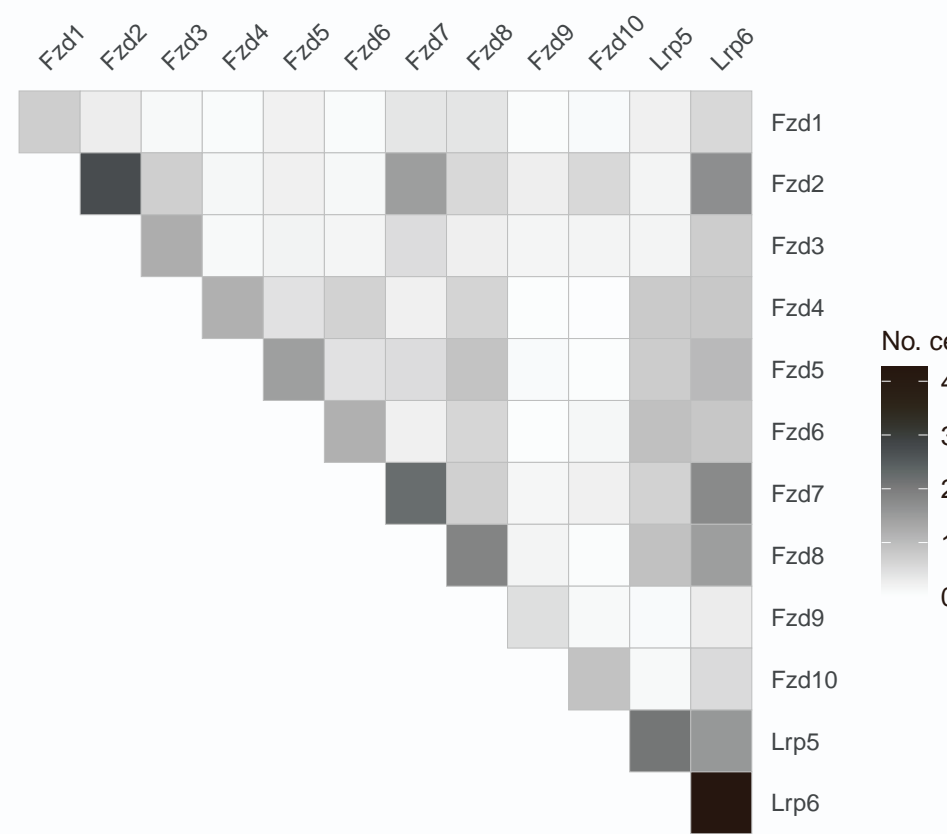


Frizzled and Lrp5 6 receptors for Wnt B Catenin Signaling (k_opt = 37)

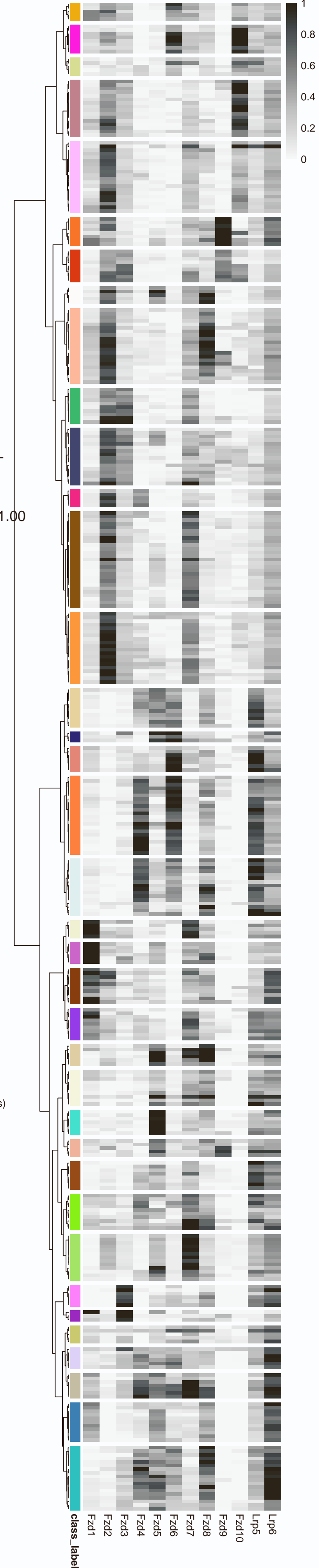
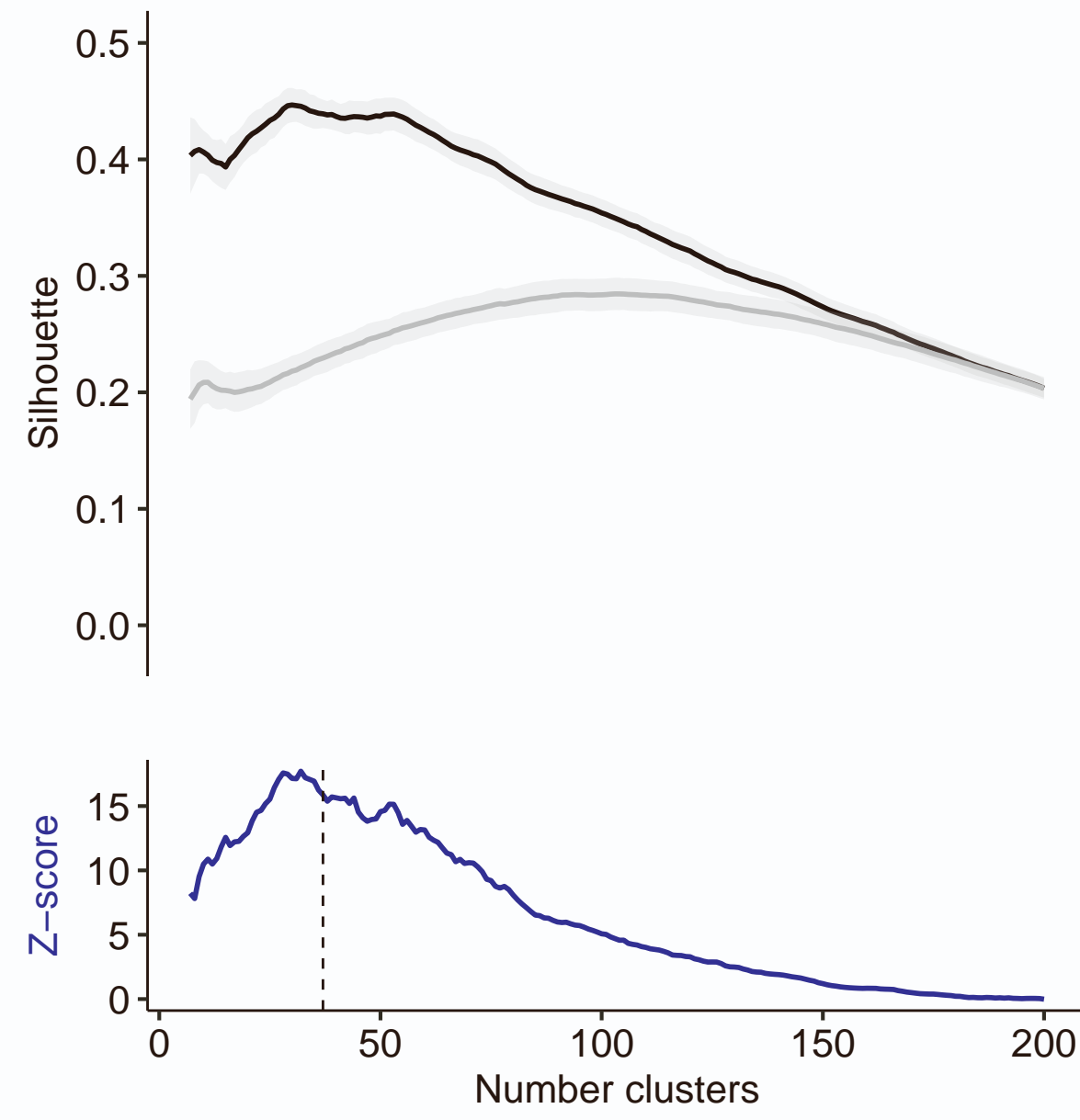
Pathway genes

Fzd1	Fzd5	Fzd9
Fzd2	Fzd6	Fzd10
Fzd3	Fzd7	Lrp5
Fzd4	Fzd8	Lrp6

Gene Co-Expression Above Threshold



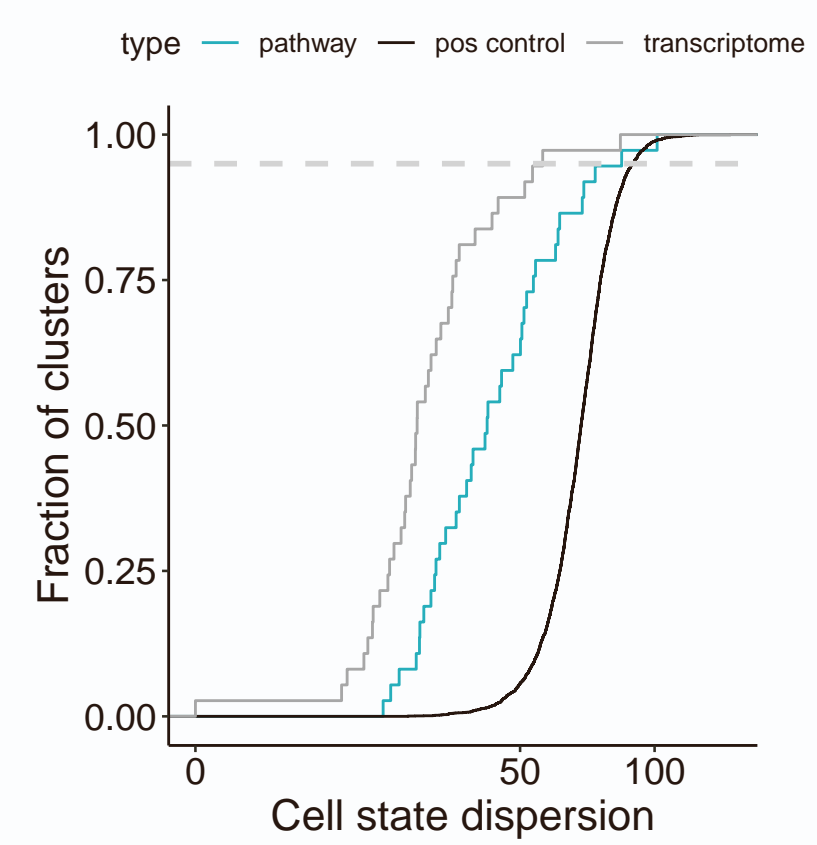
Silhouette and Z-score



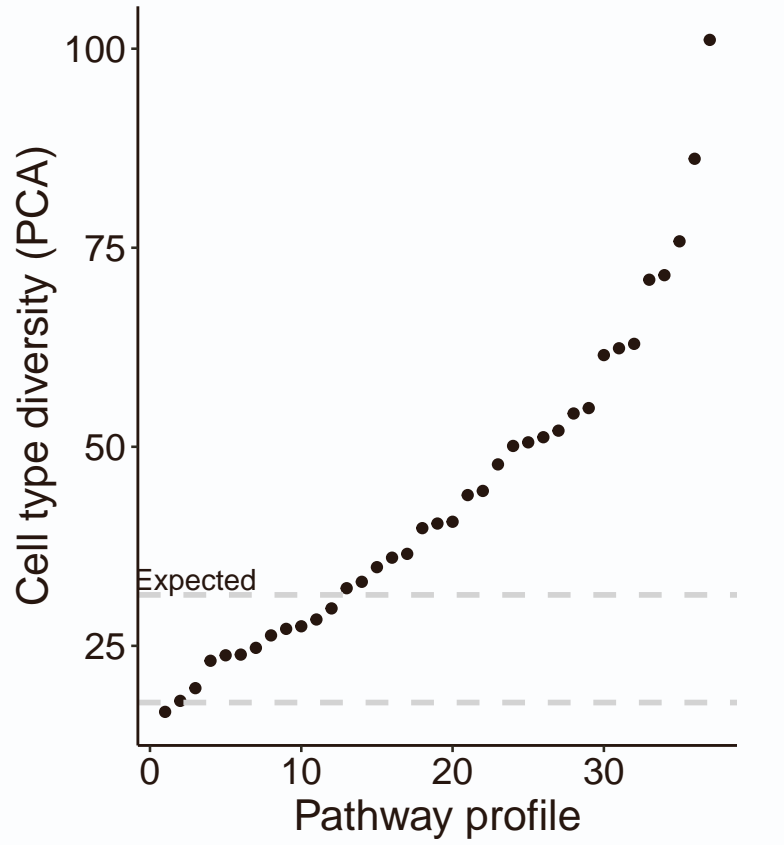
Global UMAP: Pathway ON



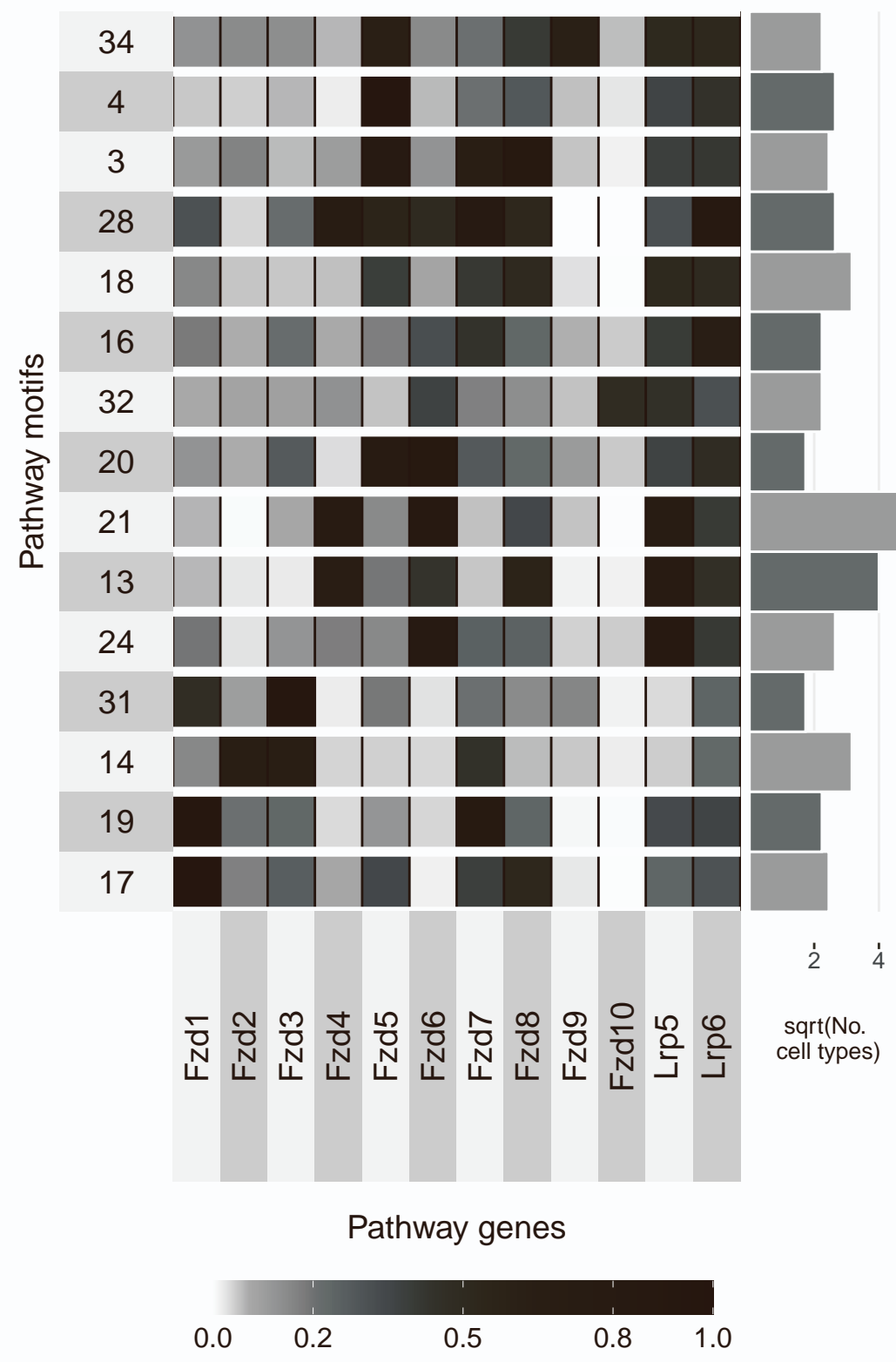
ECDF of Profile Dispersions



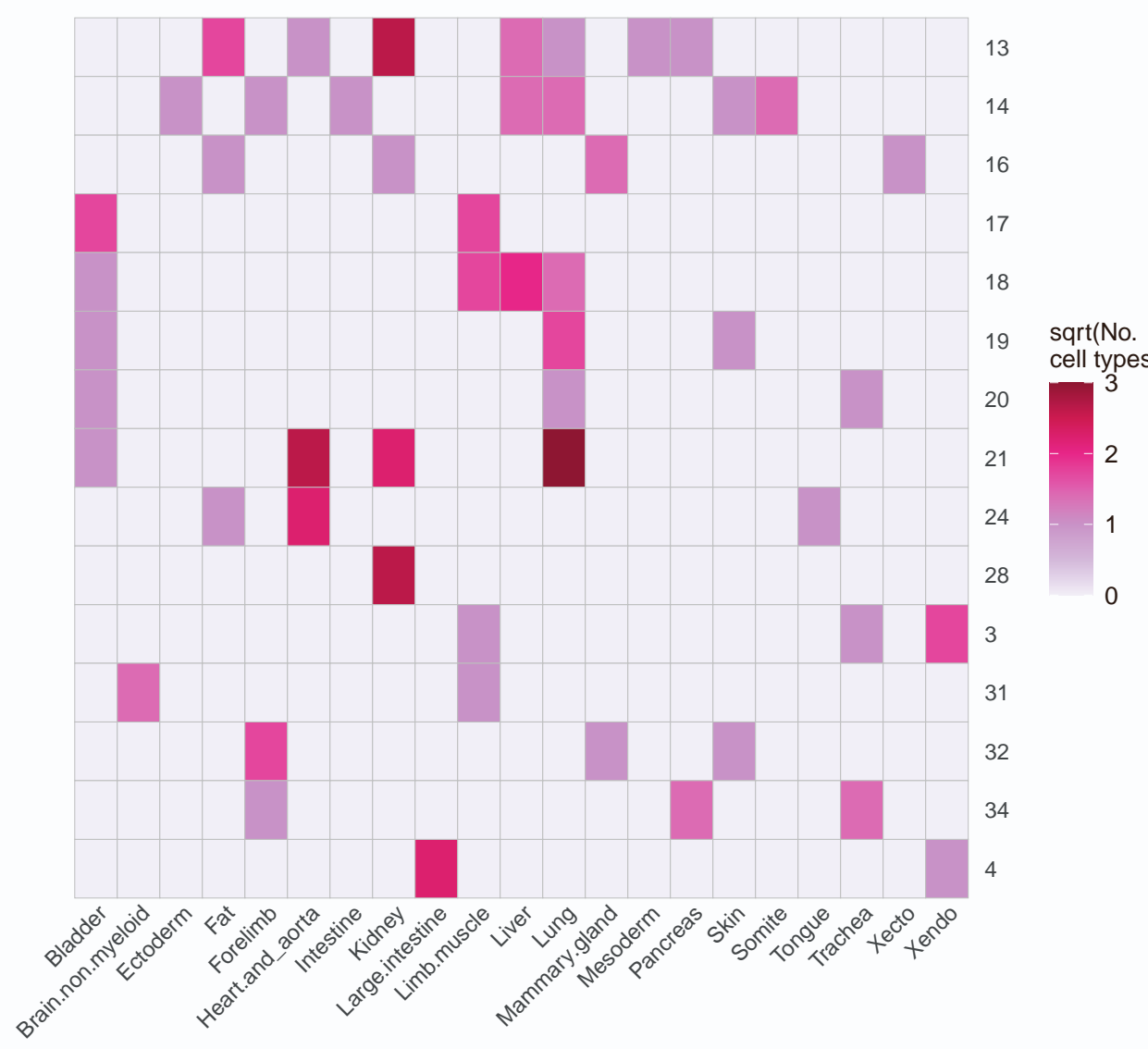
Profile Dispersion Plot



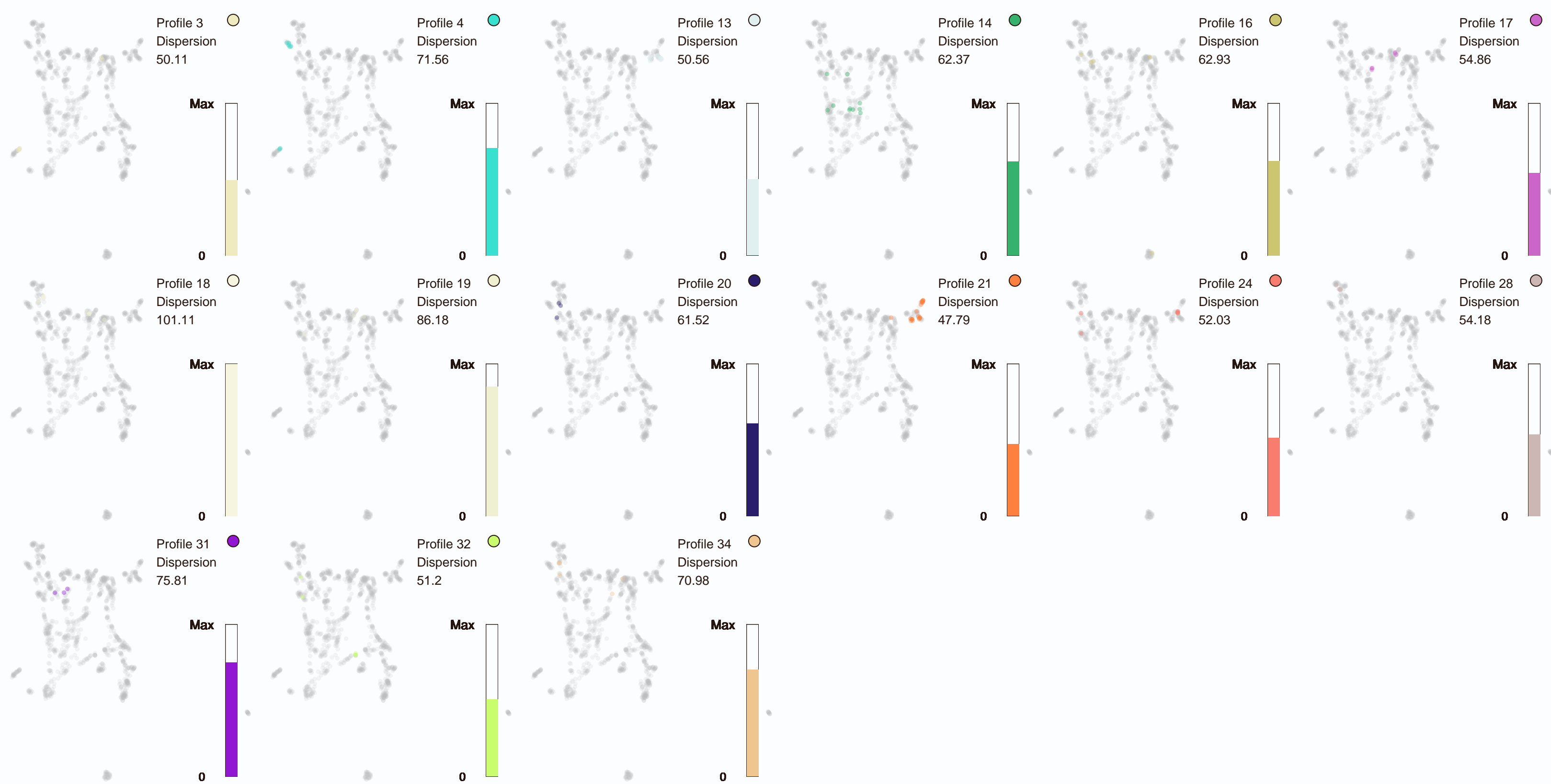
Motifs, Dispersion $\geq 90^{\text{th}}$ percentile



Motif Tissue Composition



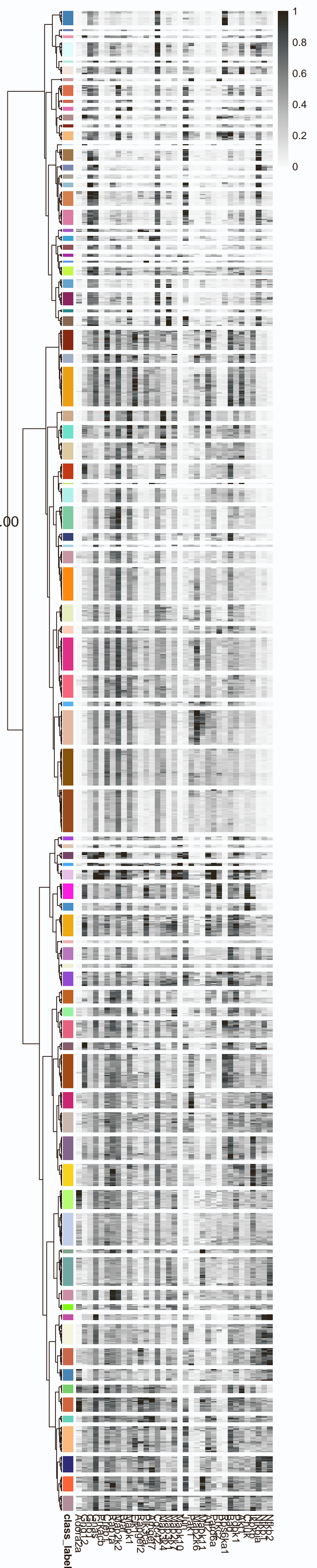
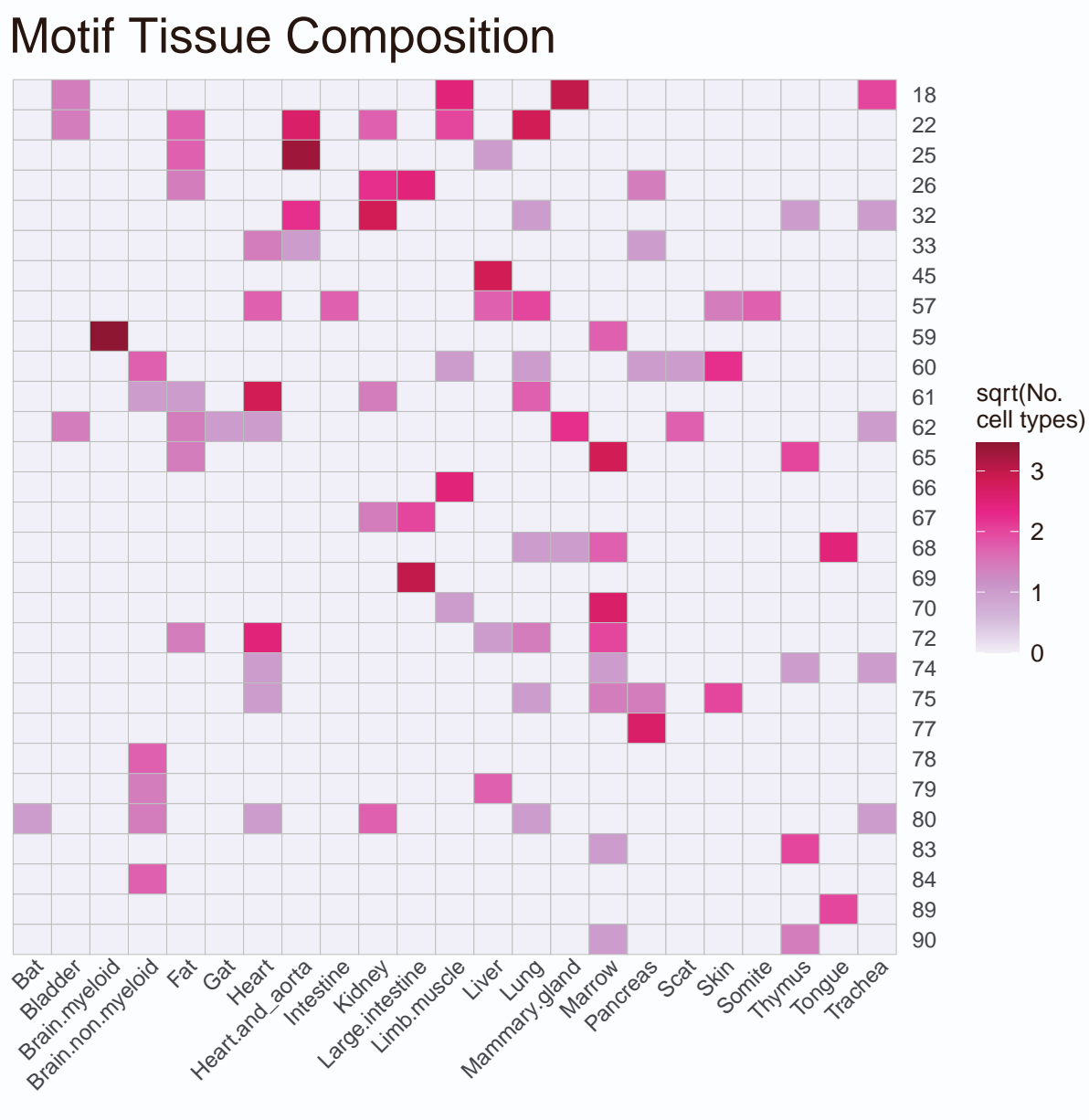
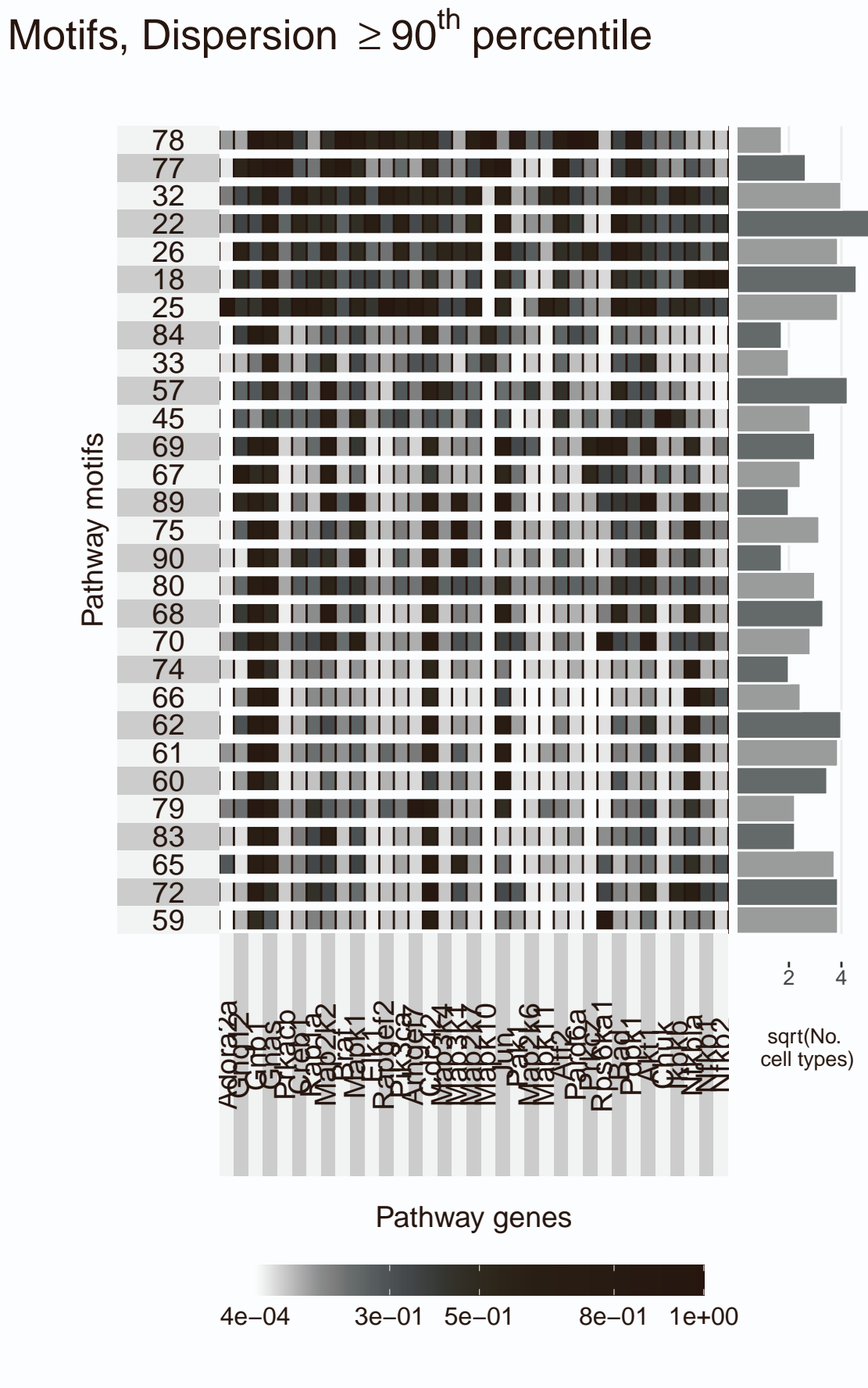
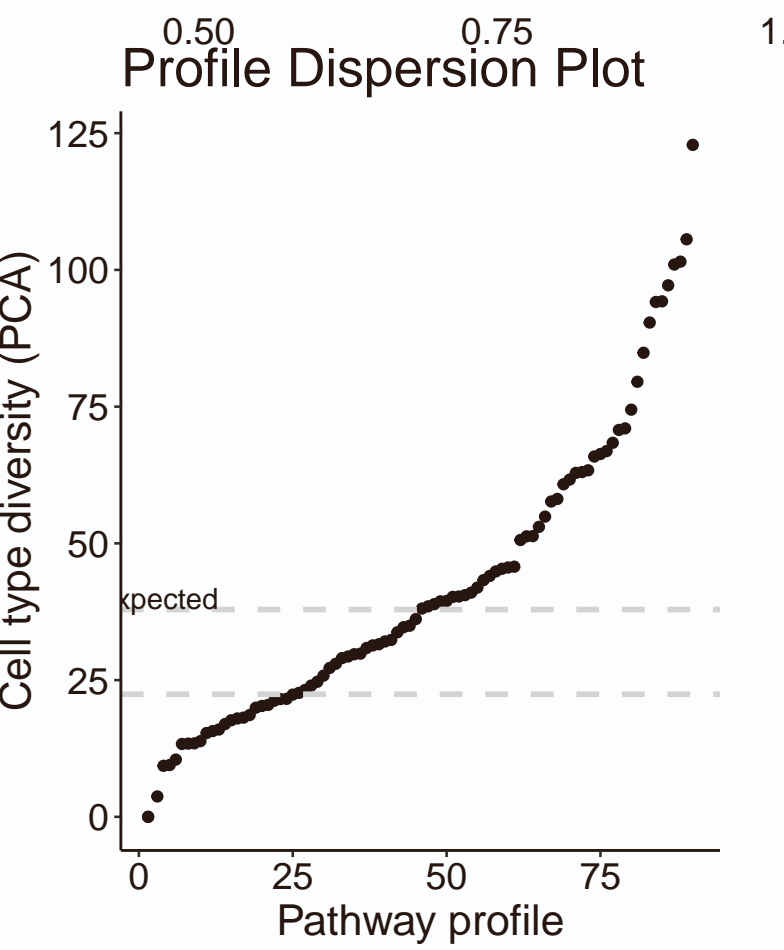
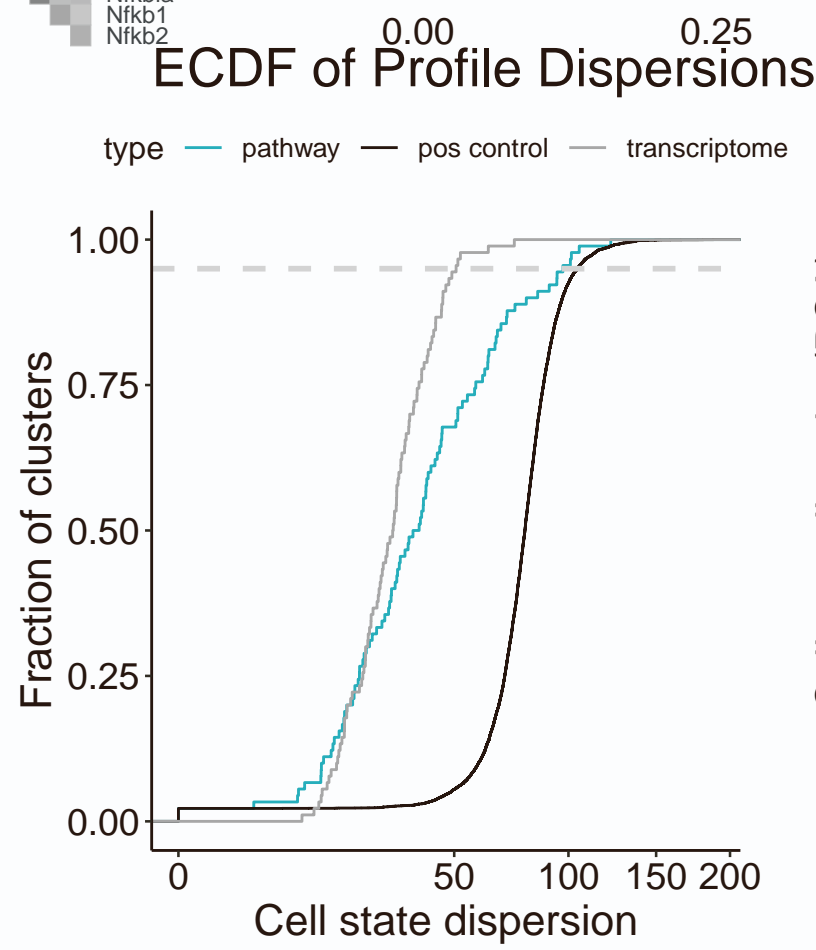
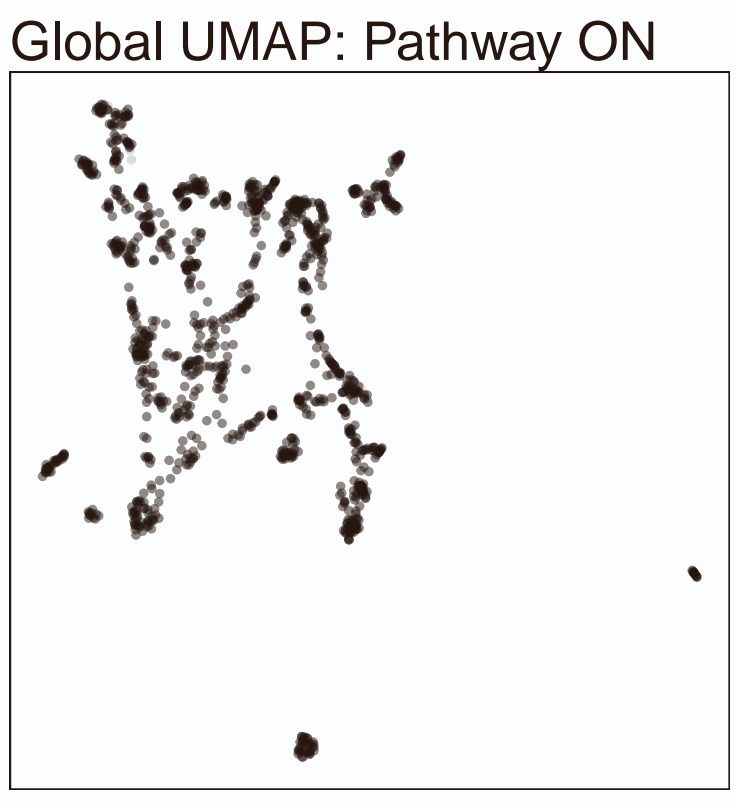
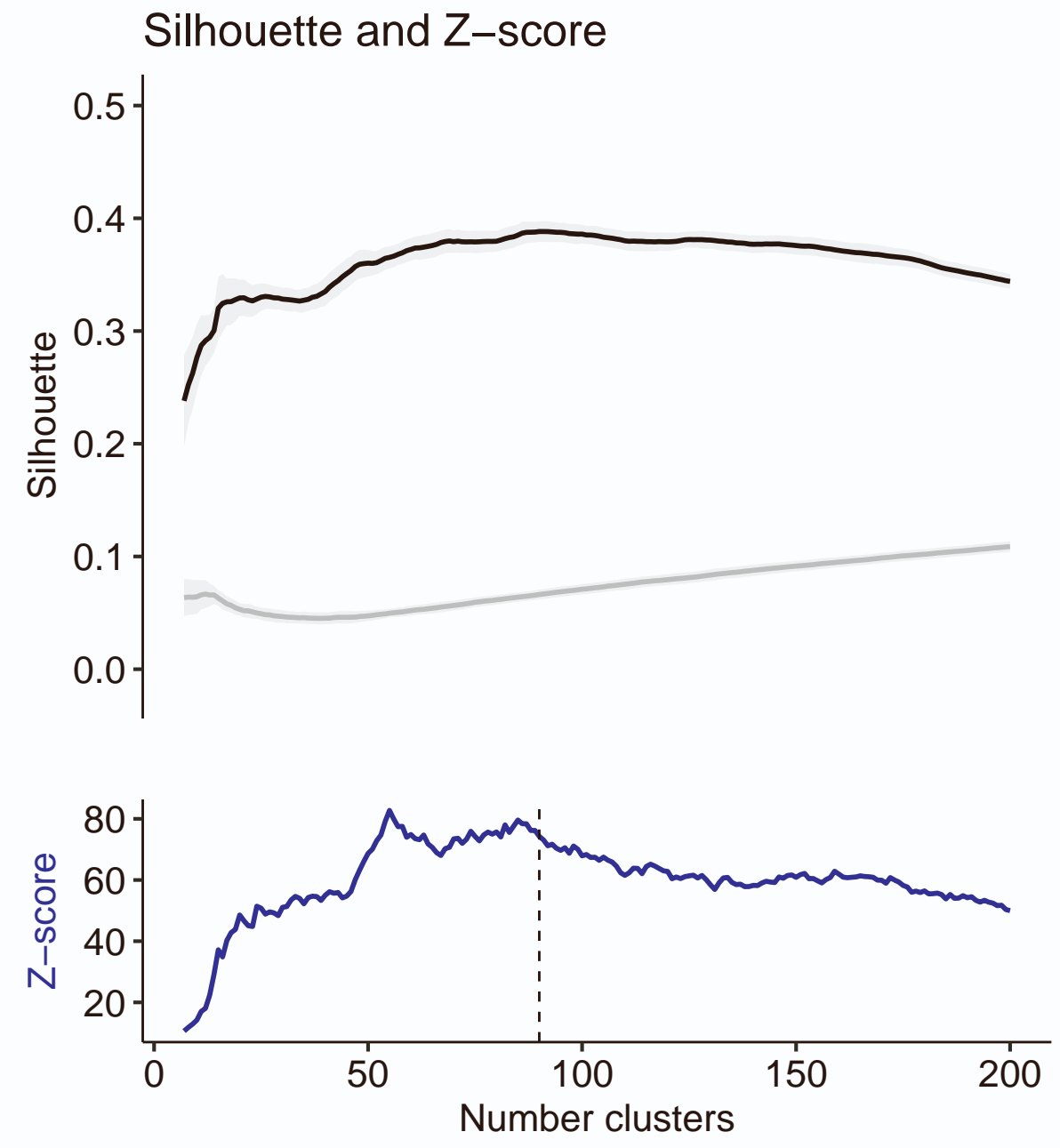
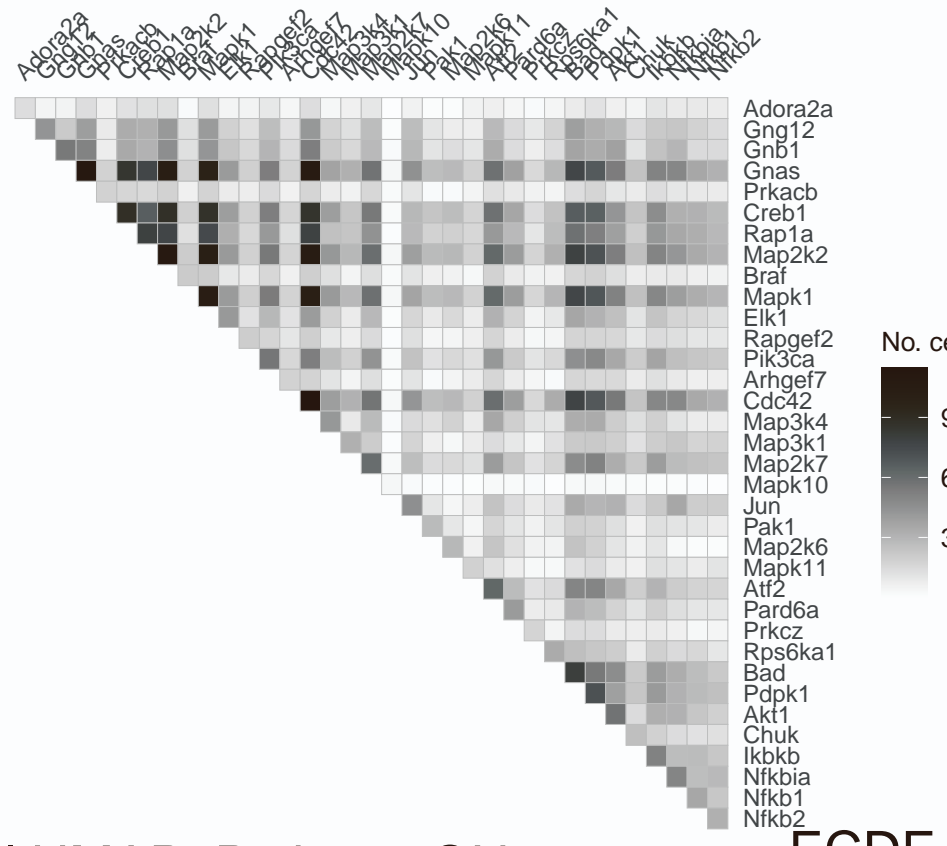
Frizzled and Lrp5 6 receptors for Wnt B Catenin Signaling (k_opt = 37)



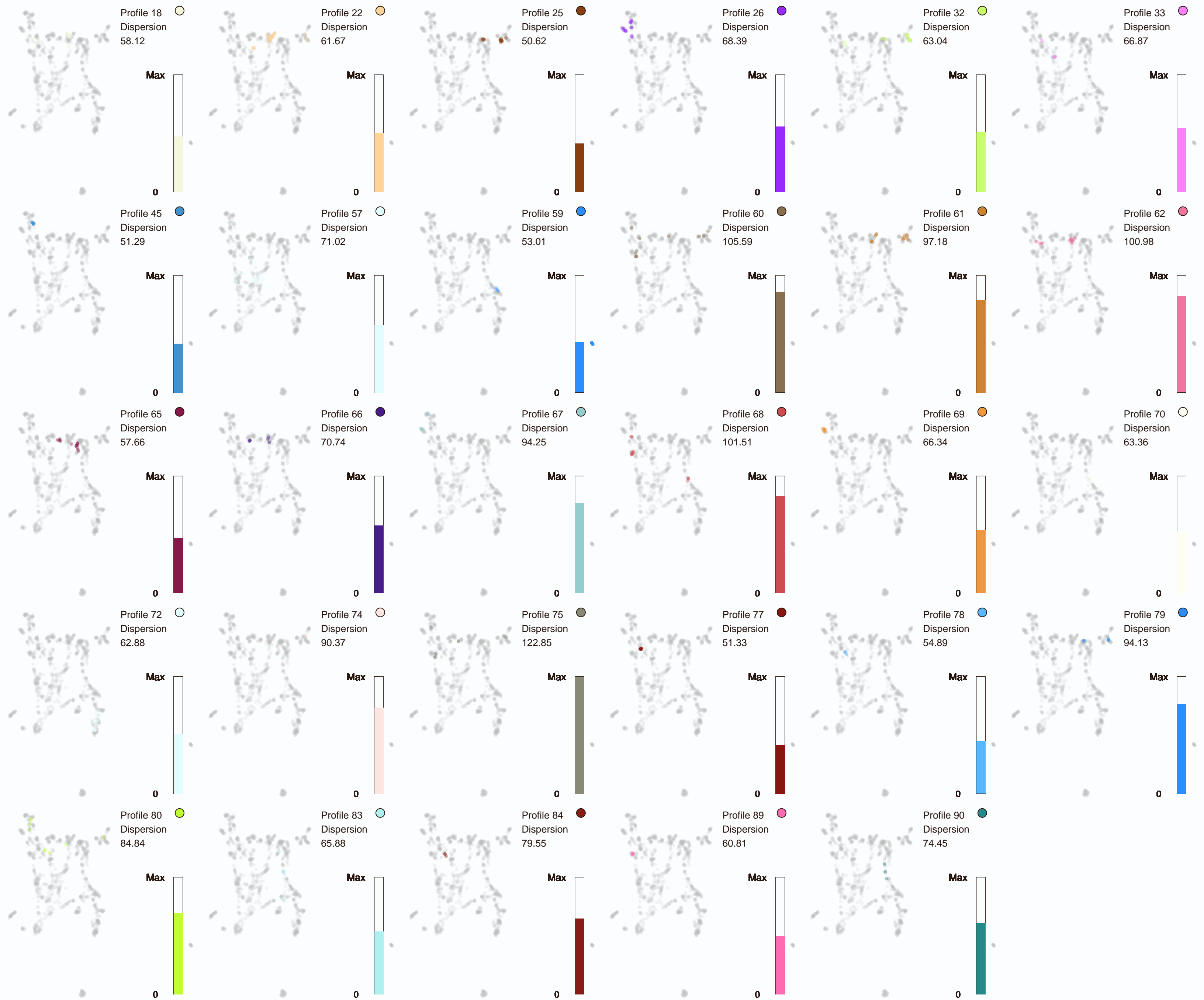
Cellular Signalling Through Adenosine Receptor A2a and Adenosine (k_opt = ...)

Gene Co-Expression Above Threshold

Adora2a, Gng12, Gnb1, Gnas, Prkacb, Creb1, Rap1a, Map2k2, Braf, Mapk1, Elk1, Pathway genes: Pik3ca, Arhgef7, Cdc42, Map3k4, Map3k1, Map2k7, Jun, Pak1, Map2k6, Mapk11, Atf2, Pard6a, Prkcz, Rps6ka1, Bad, Pdpk1, Akt1, Chuk, Ikbkb, Nfkb1, Nfkb2



Cellular Signalling Through Adenosine Receptor A2a and Adenosine (k_opt =



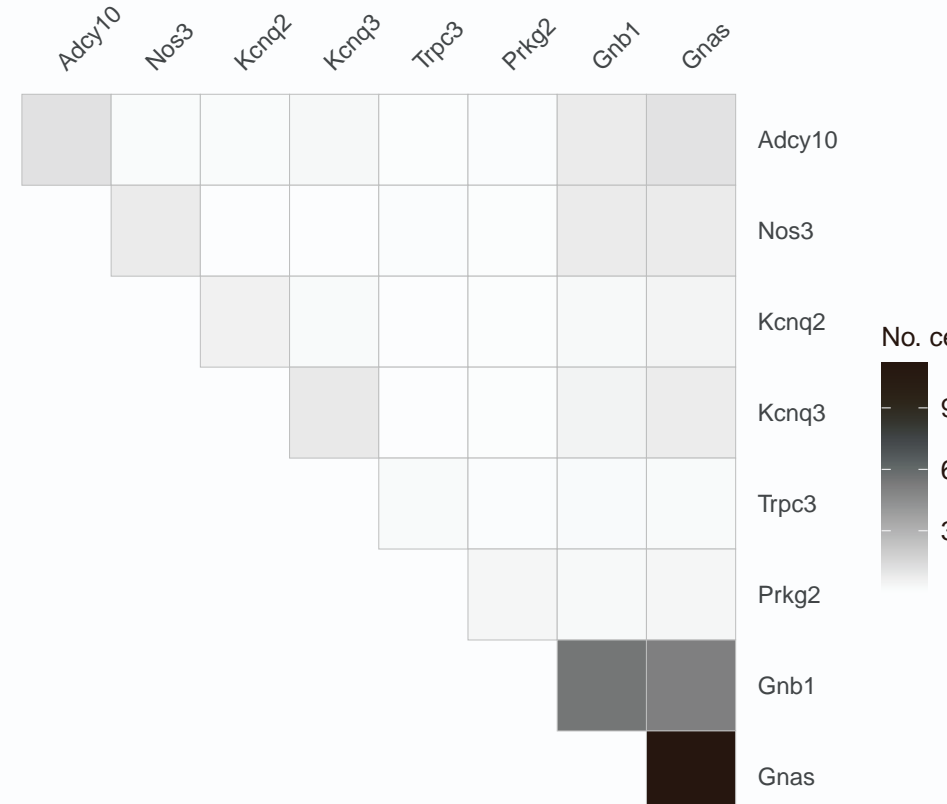
on Channels and Their Functional Role in Vascular Endothelium (k_opt = 32)

Adcy10
Nos3
Kcnq2

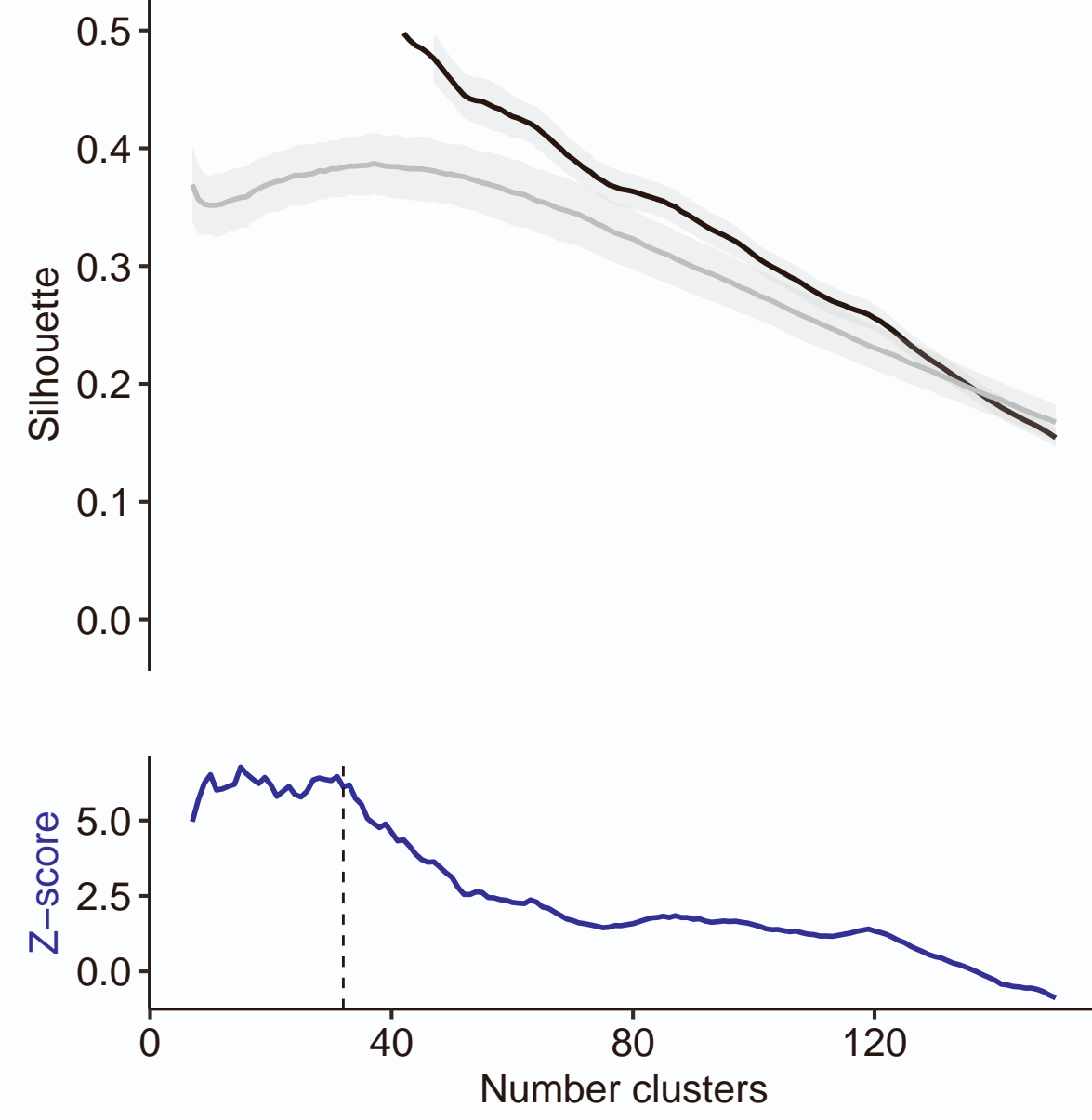
Pathway genes
Kcnq3
Trpc3

Prkg2
Gnb1
Gnas

Gene Co-Expression Above Threshold



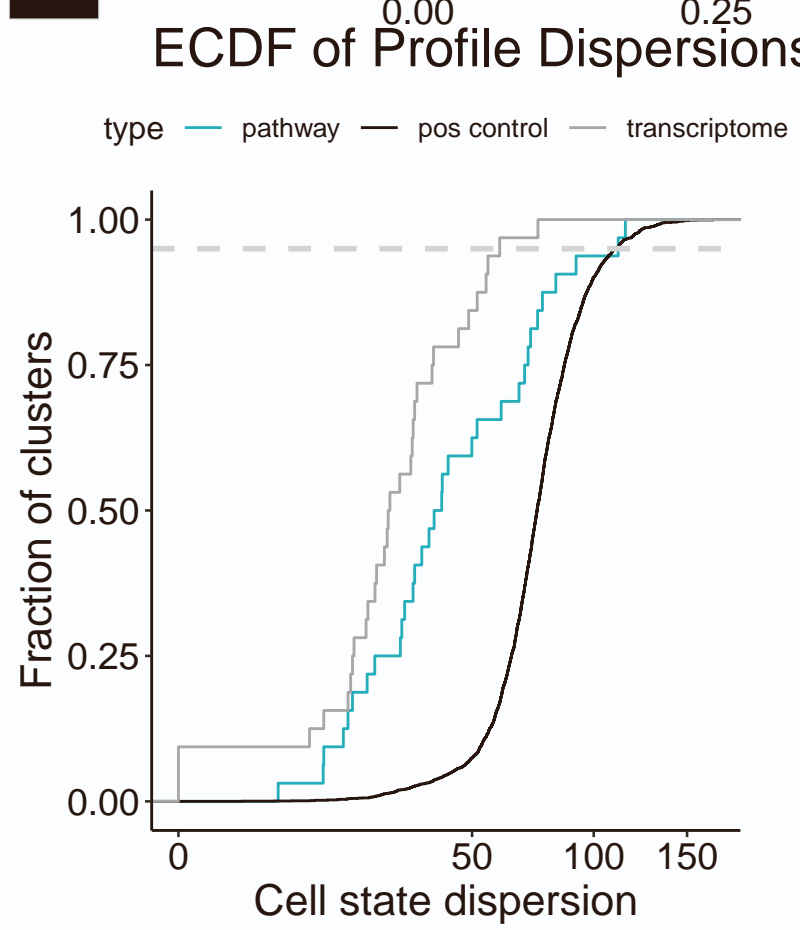
Silhouette and Z-score



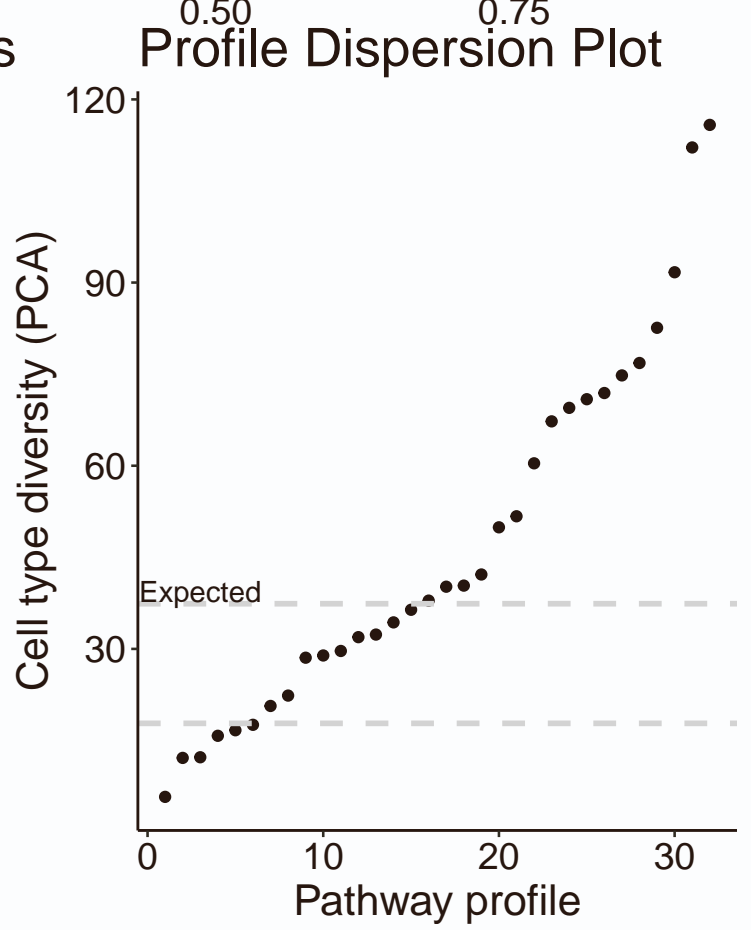
Global UMAP: Pathway ON



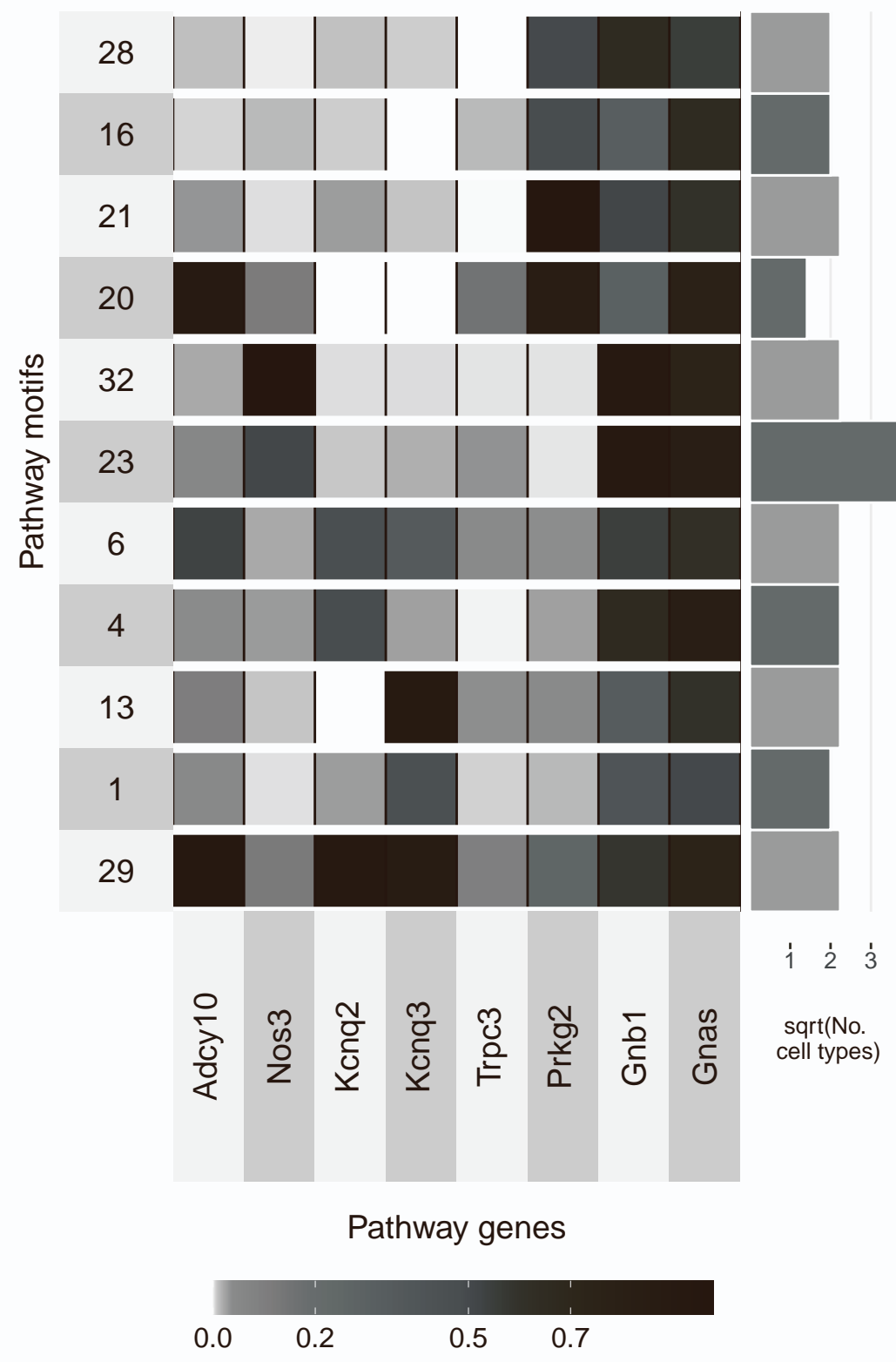
ECDF of Profile Dispersions



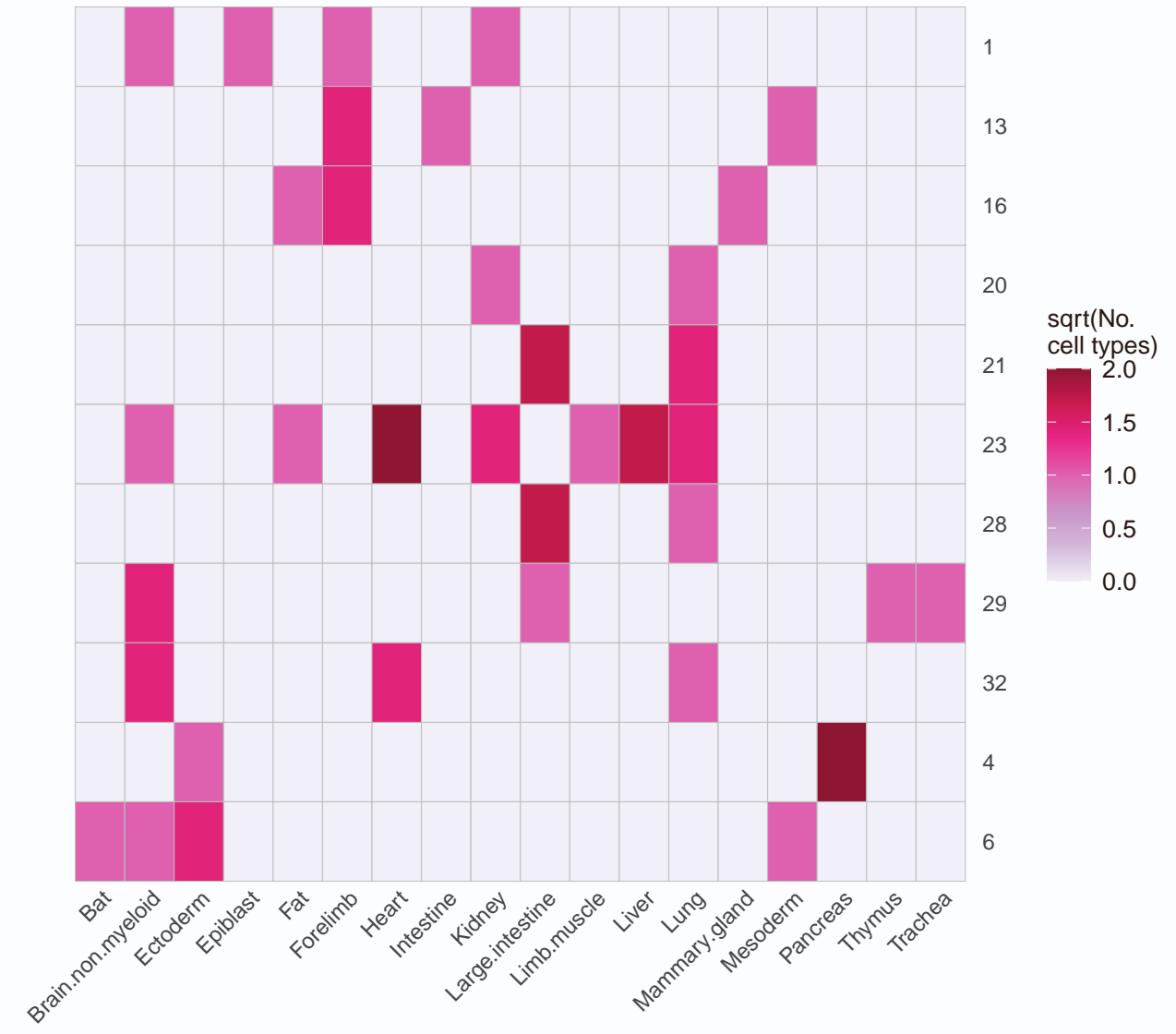
Profile Dispersion Plot



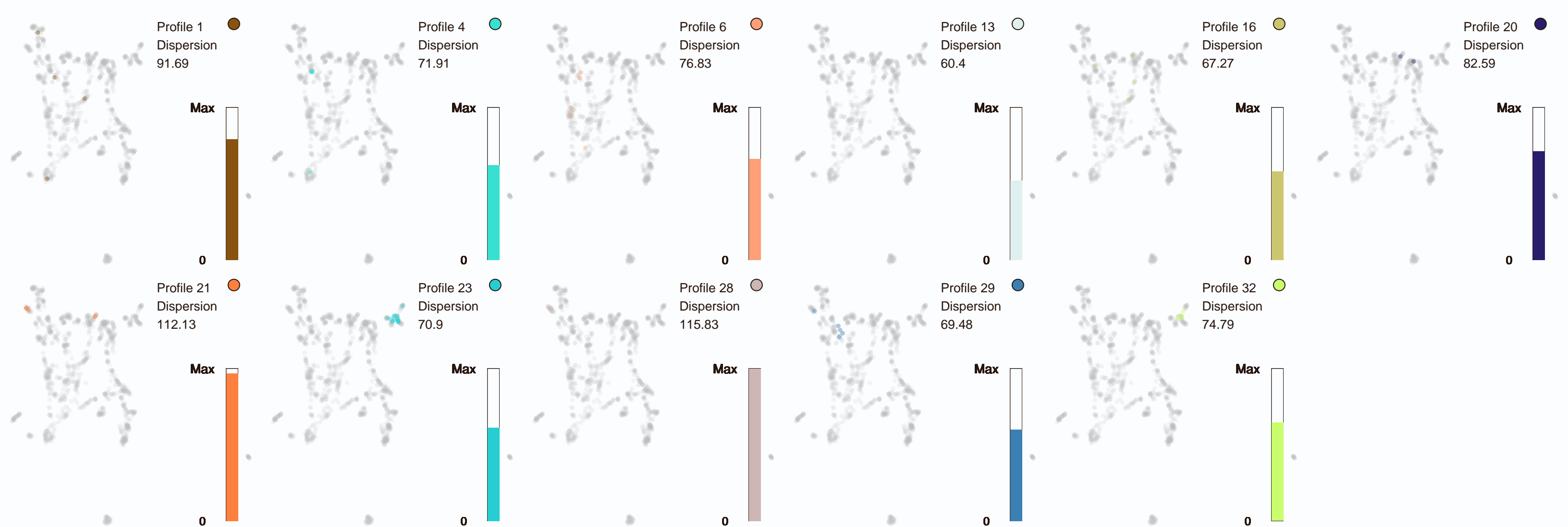
Motifs, Dispersion $\geq 90^{\text{th}}$ percentile



Motif Tissue Composition



ion Channels and Their Functional Role in Vascular Endothelium (k_opt = 32)



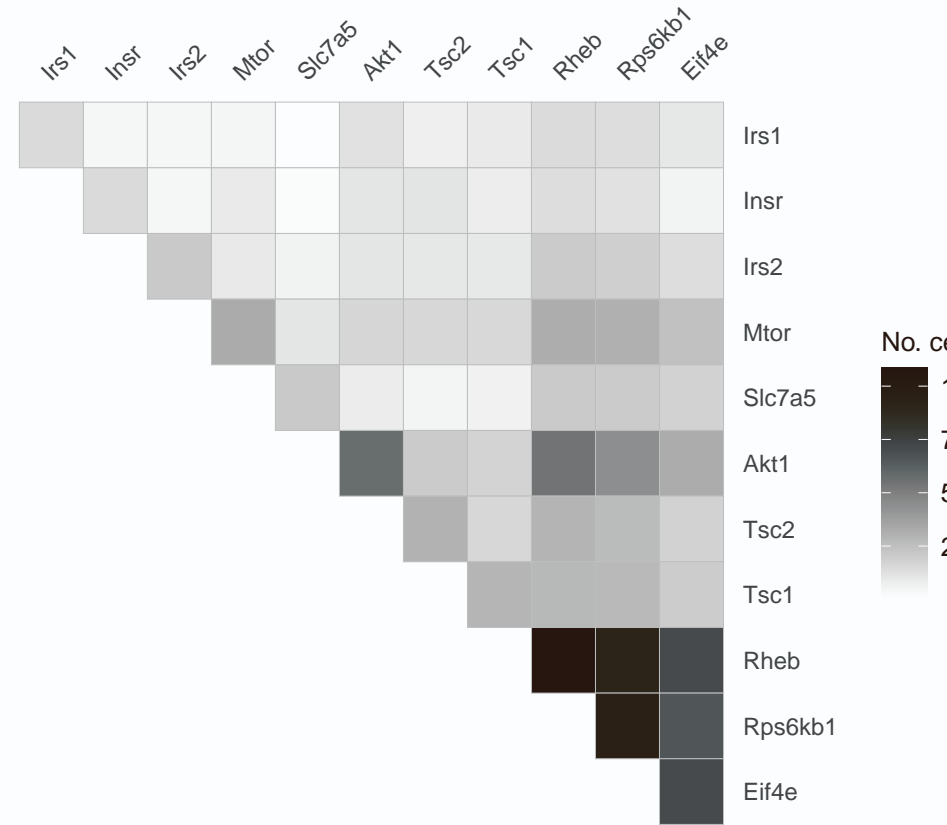
Leucine Stimulation on Insulin Signaling (k_opt = 34)

Irs1
Irsr
Irs2
Mtor

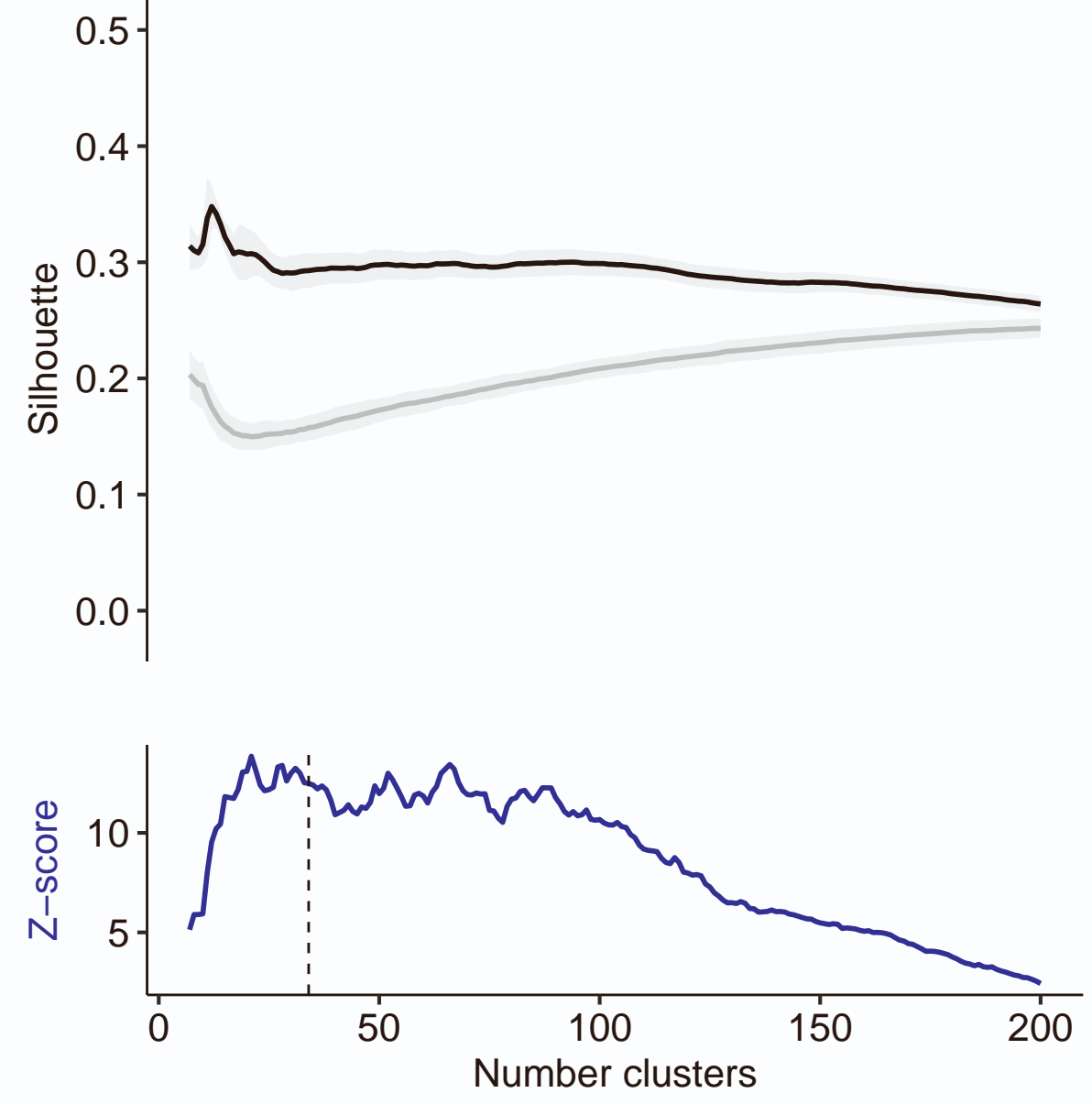
Pathway genes
Slc7a5
Akt1
Tsc2

Tsc1
Rheb
Rps6kb1
Eif4e

Gene Co-Expression Above Threshold



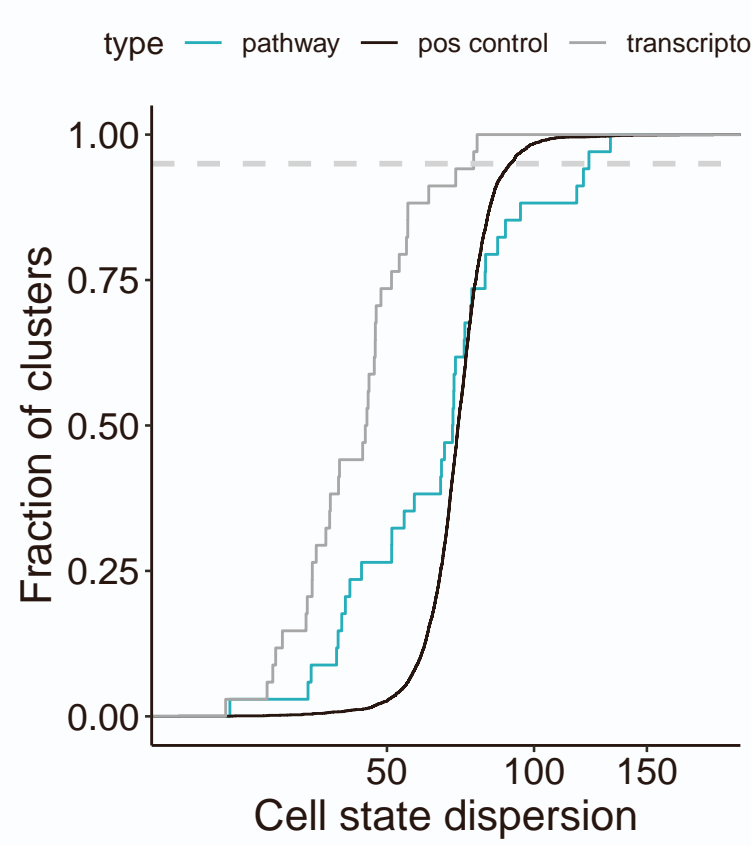
Silhouette and Z-score



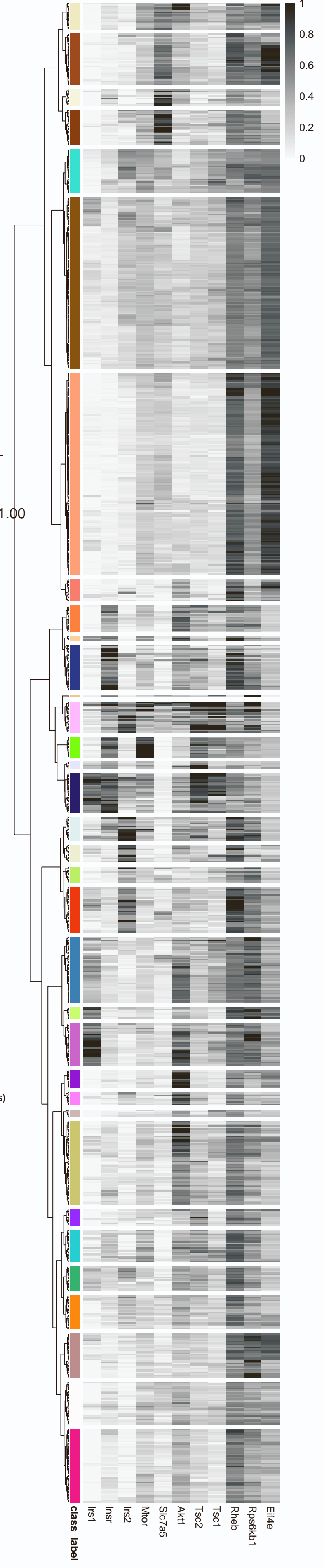
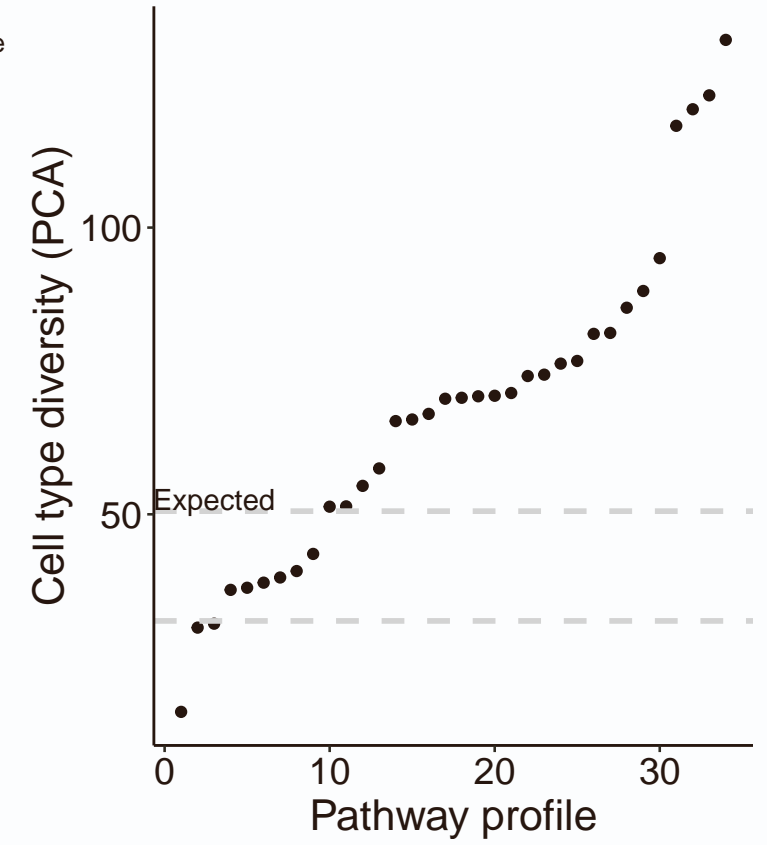
Global UMAP: Pathway ON



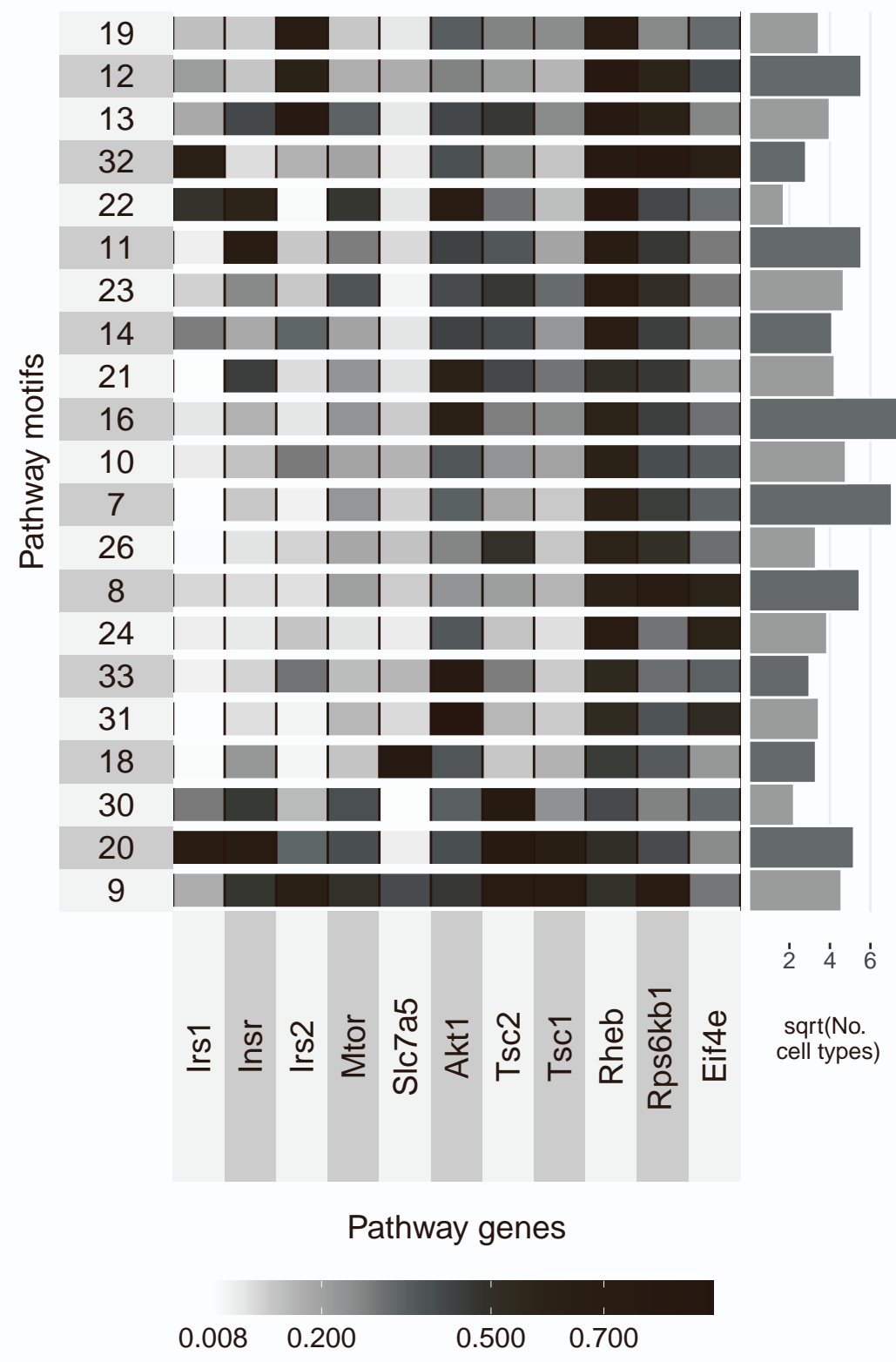
ECDF of Profile Dispersions



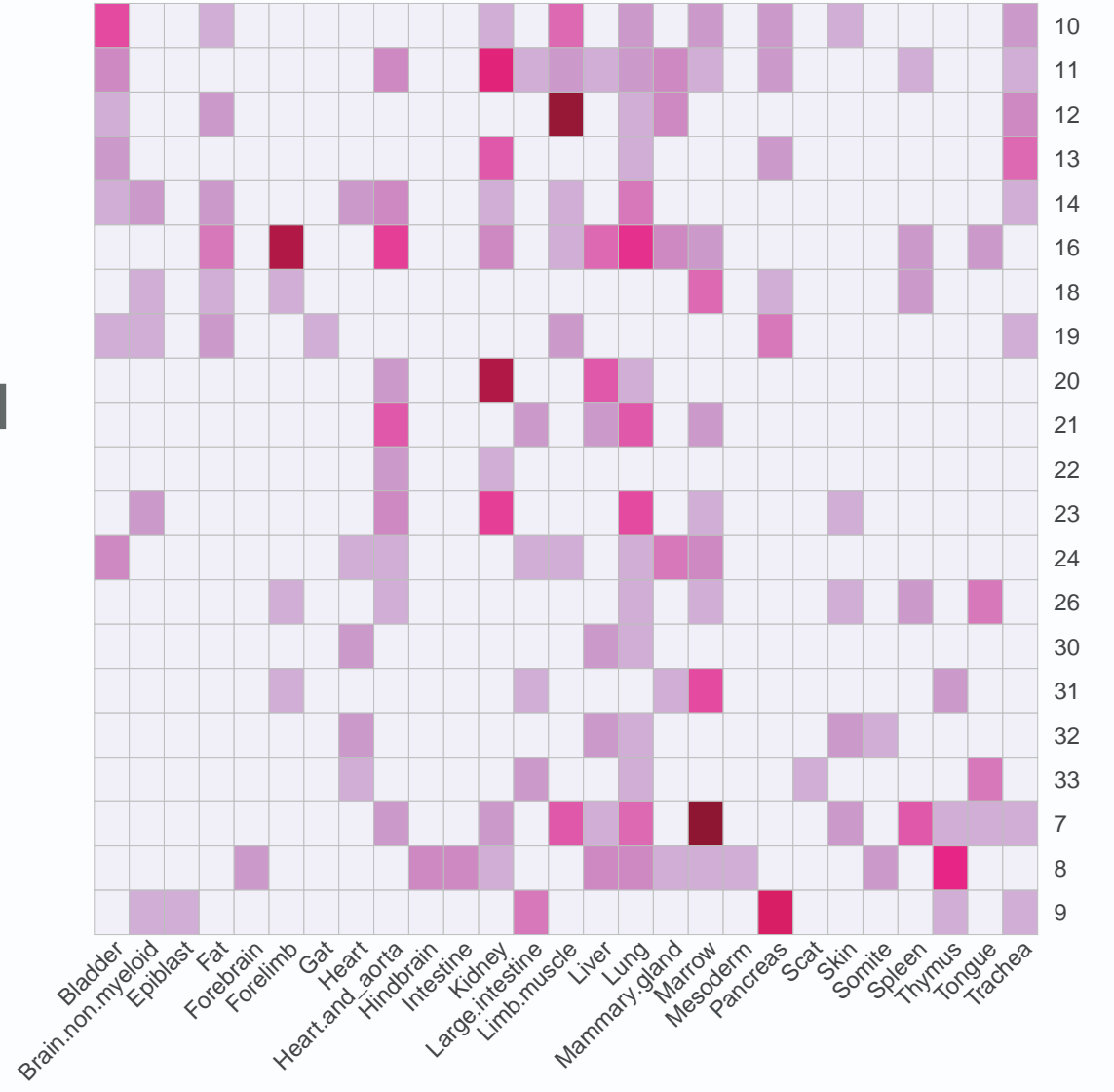
Profile Dispersion Plot



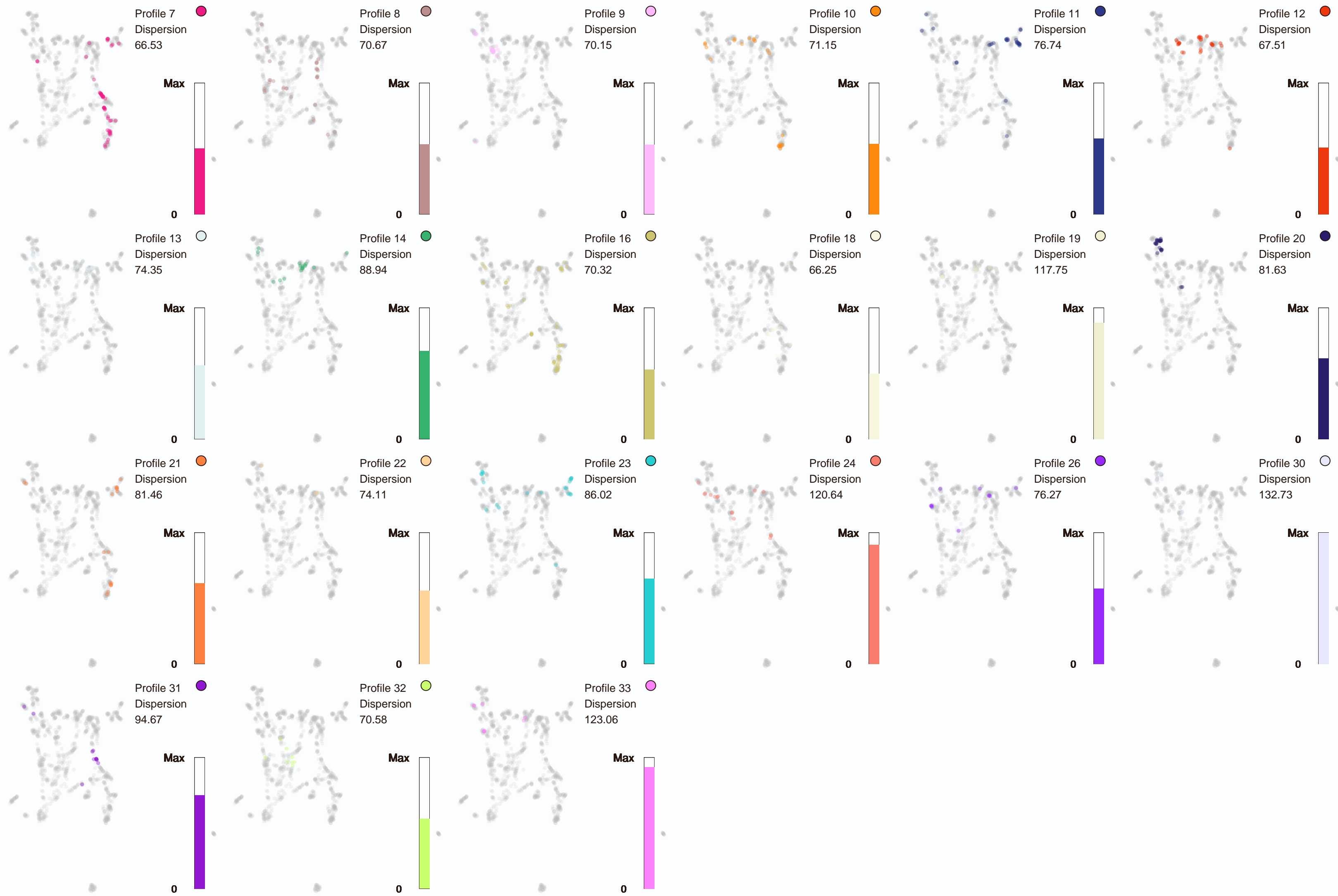
Motifs, Dispersion $\geq 90^{\text{th}}$ percentile



Motif Tissue Composition



Leucine Stimulation on Insulin Signaling (k_opt = 34)



LPS and Citrate Signaling and Inflammation (k_opt = 112)



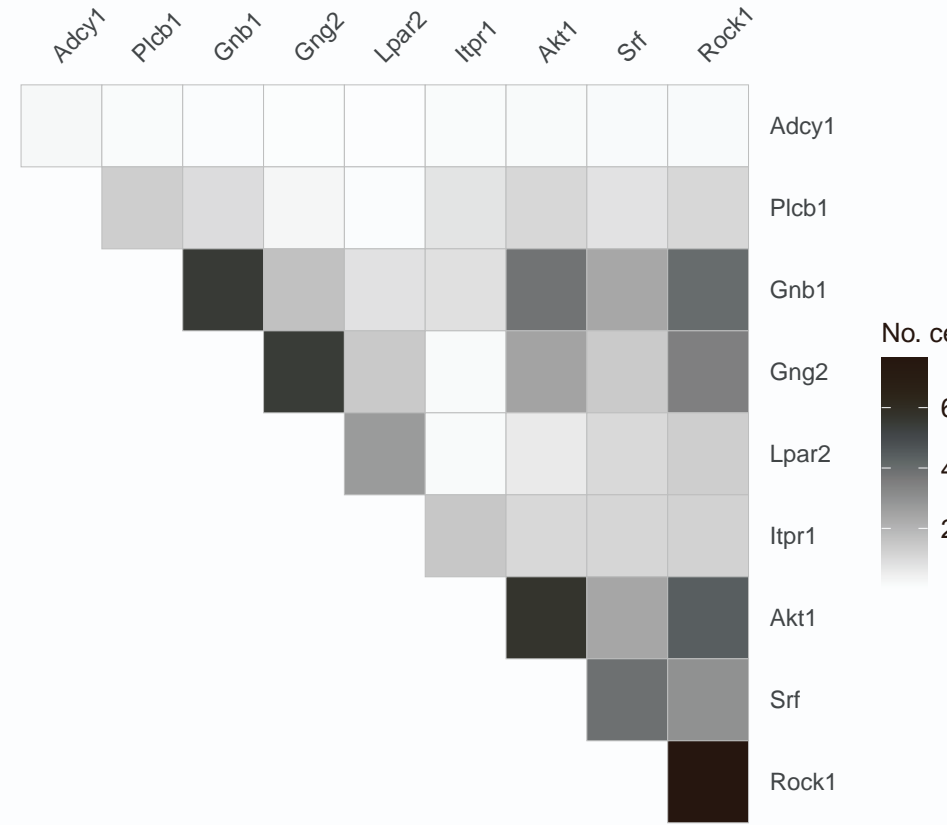
Lysophosphatidic Acid LPA2 Signalling (k_opt = 31)

Adcy1
Plcb1
Gnb1

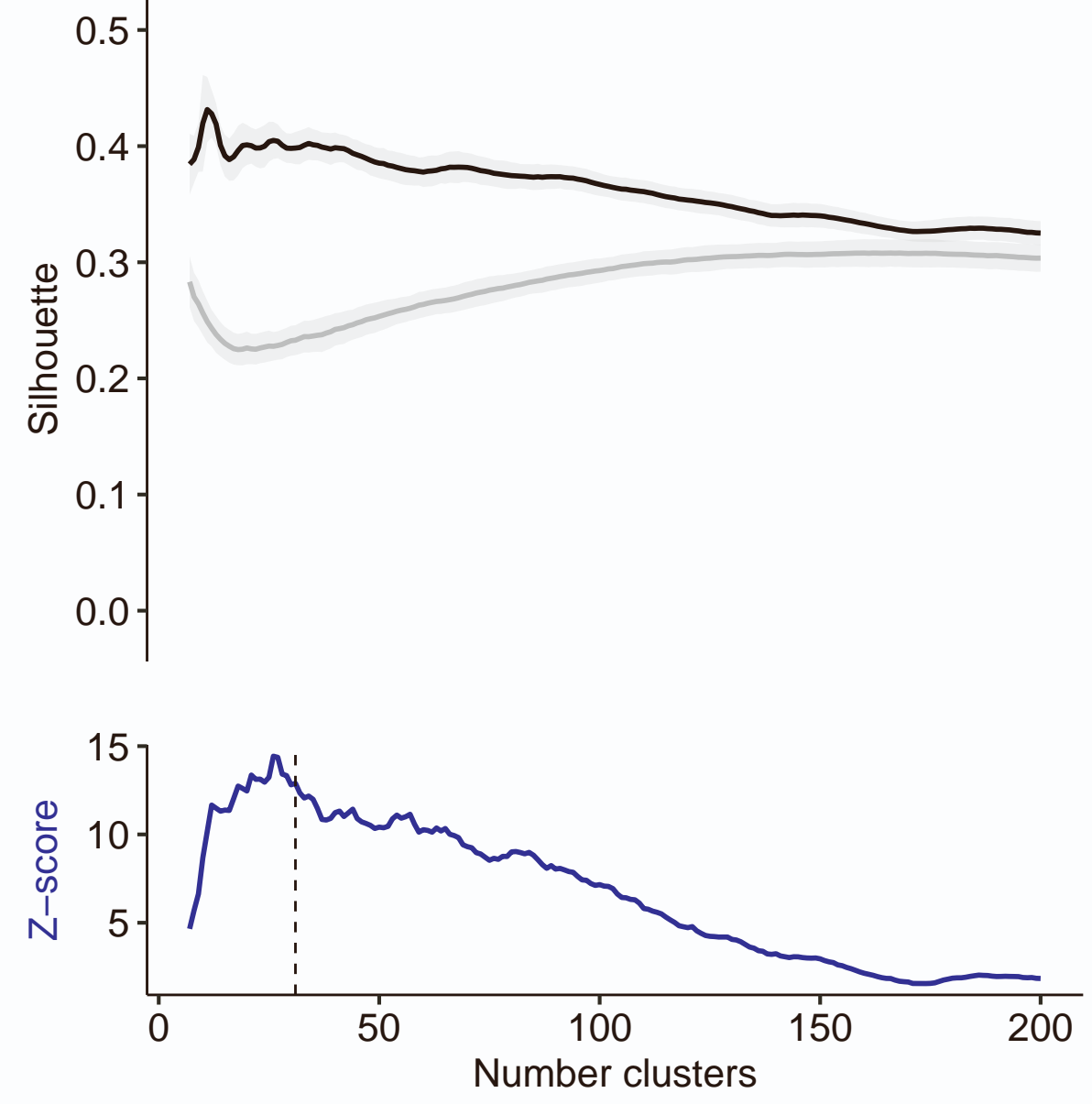
Pathway genes
Gng2
Lpar2
Itp1

Akt1
Srf
Rock1

Gene Co-Expression Above Threshold



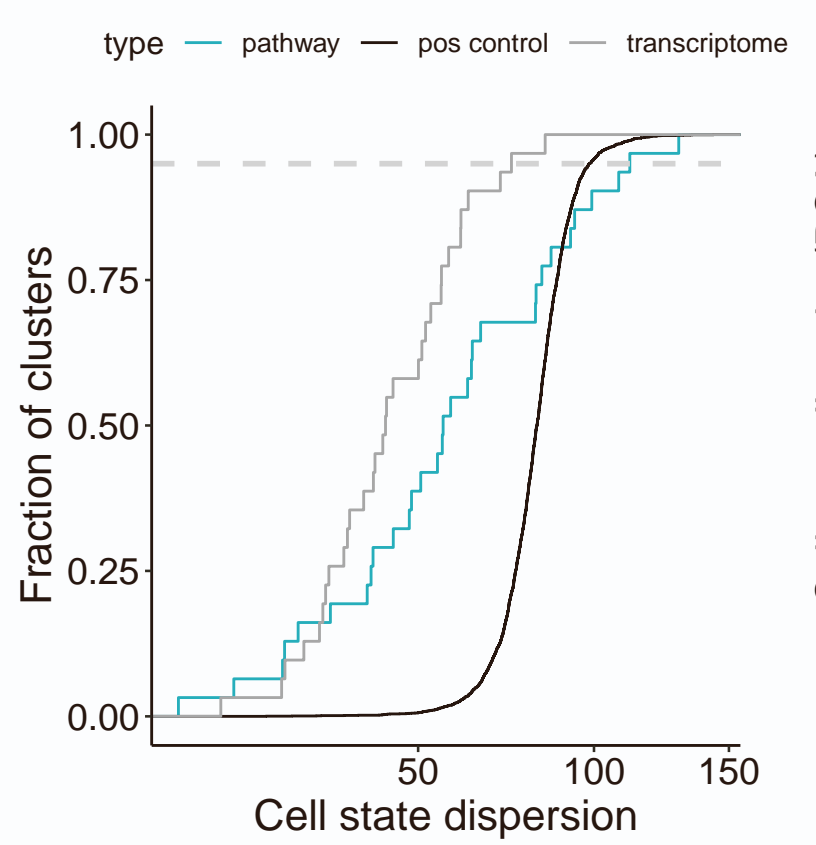
Silhouette and Z-score



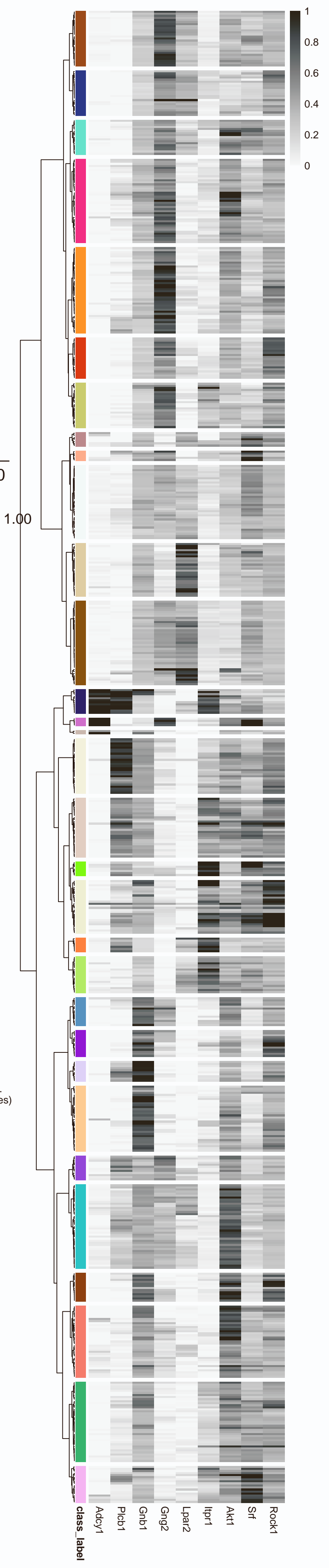
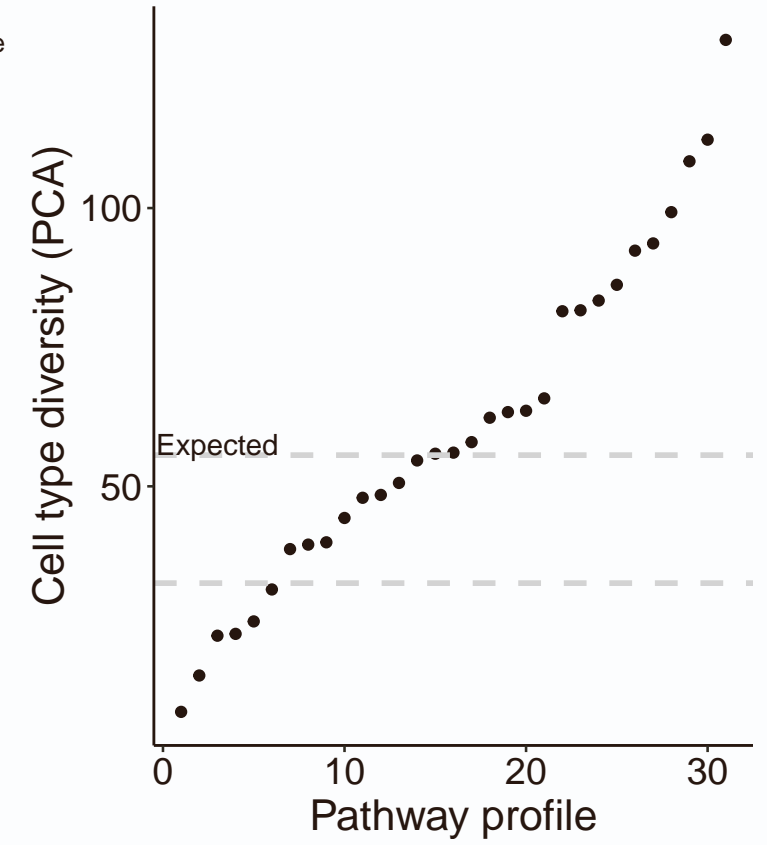
Global UMAP: Pathway ON



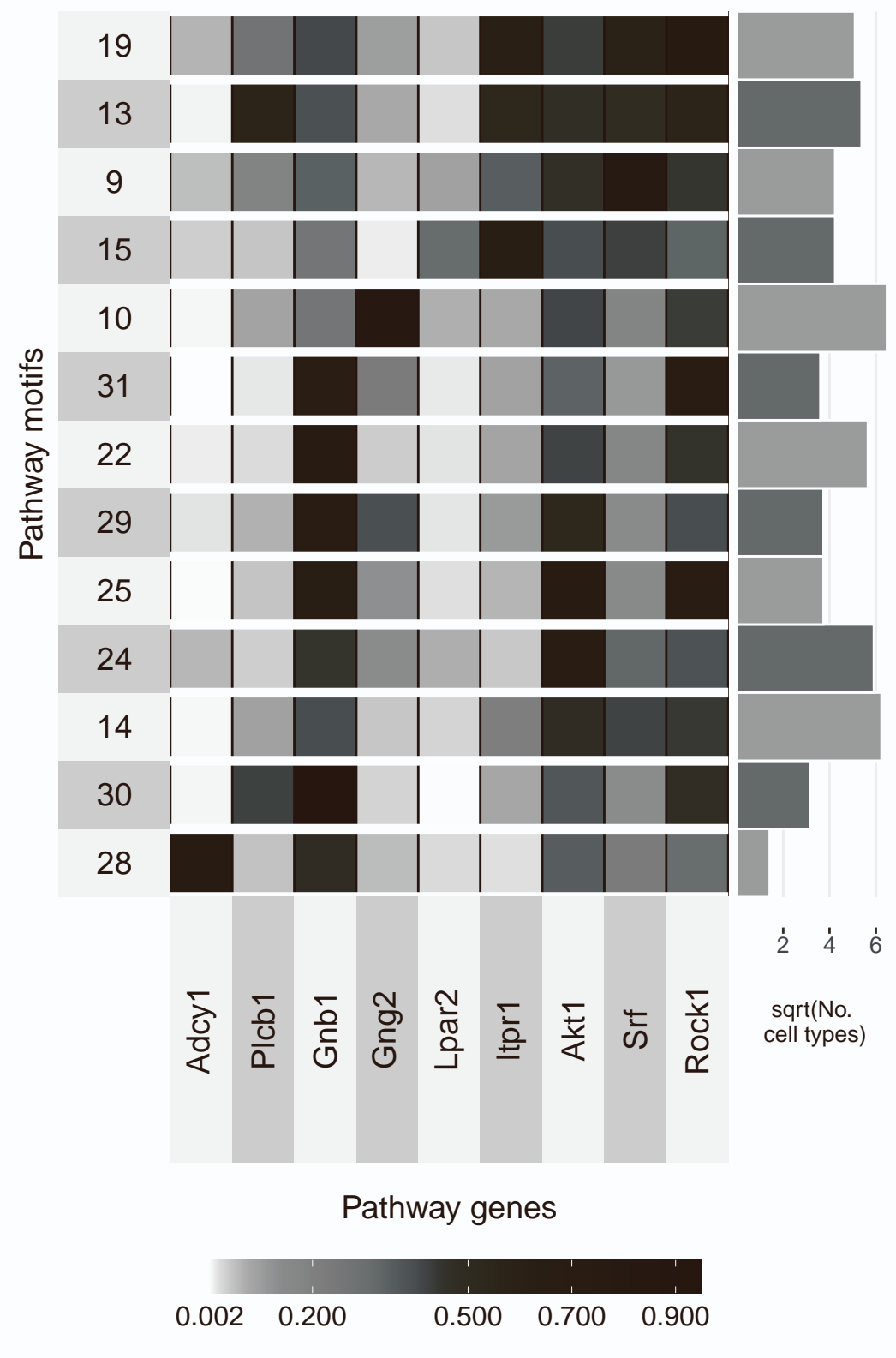
ECDF of Profile Dispersions



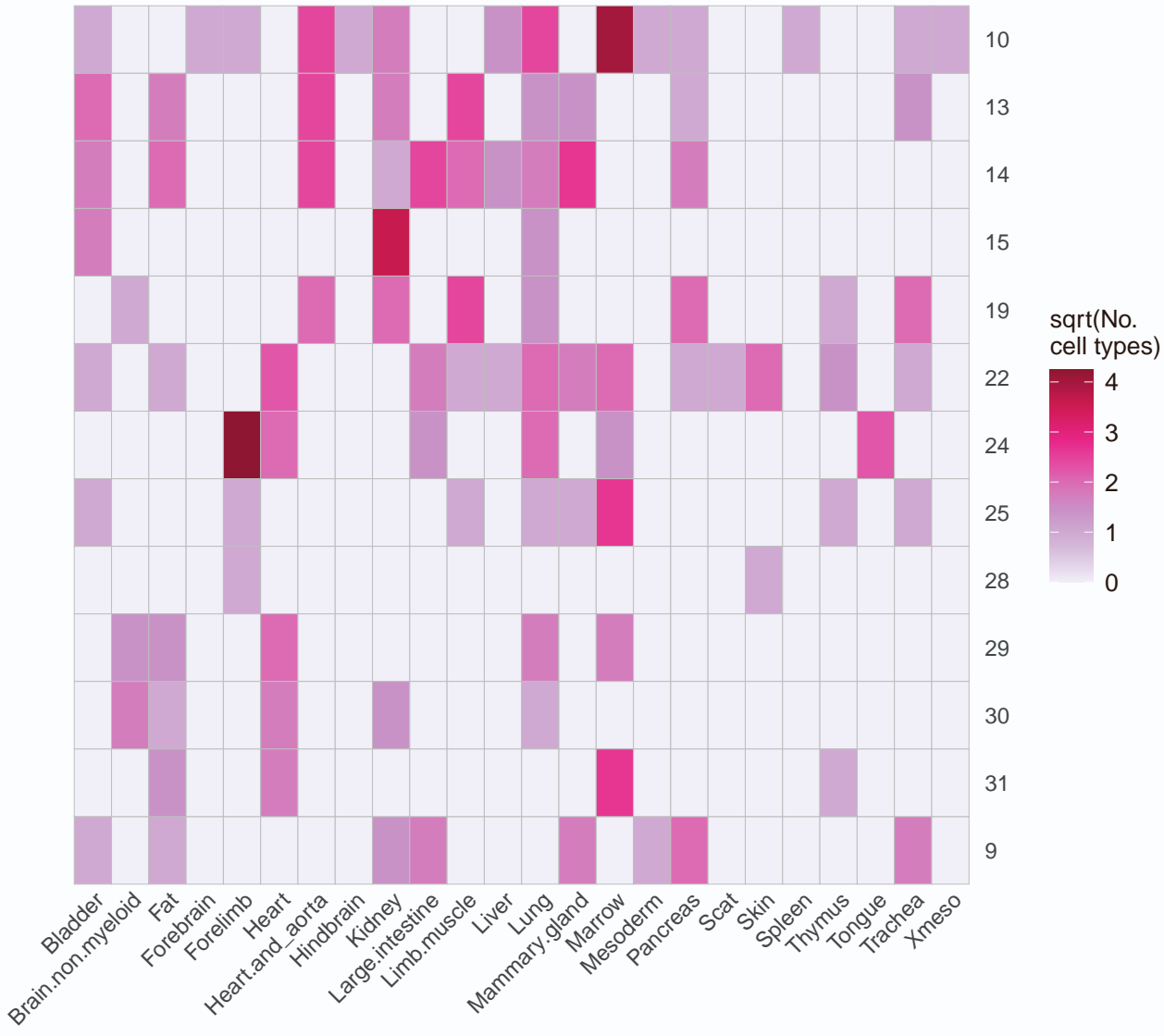
Profile Dispersion Plot



Motifs, Dispersion $\geq 90^{\text{th}}$ percentile



Motif Tissue Composition

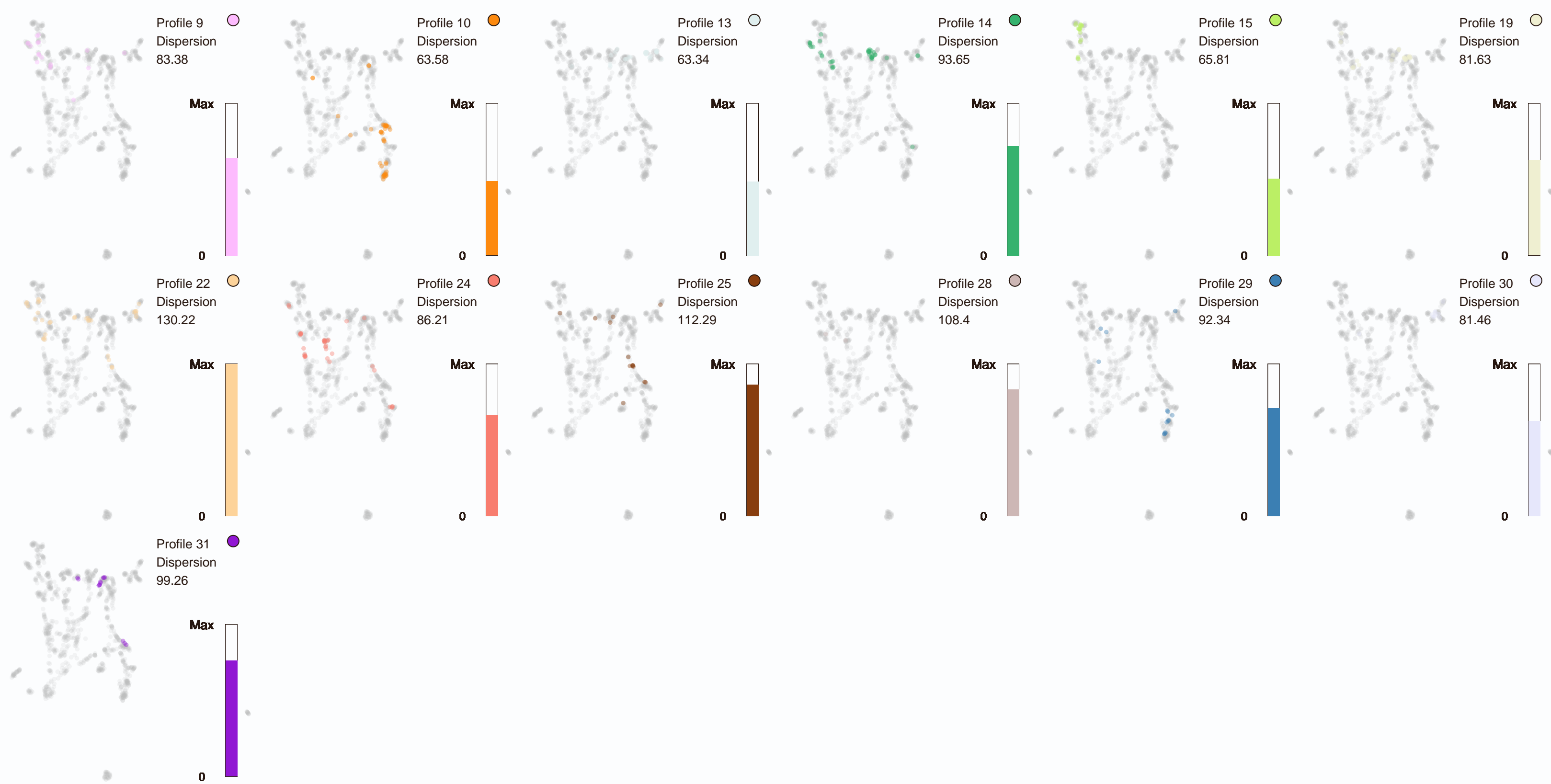


Pathway genes

Adcy1 Plcb1 Gnb1 Gng2 Lpar2 Itp1 Akt1 Srf Rock1

0.002 0.200 0.500 0.700 0.900

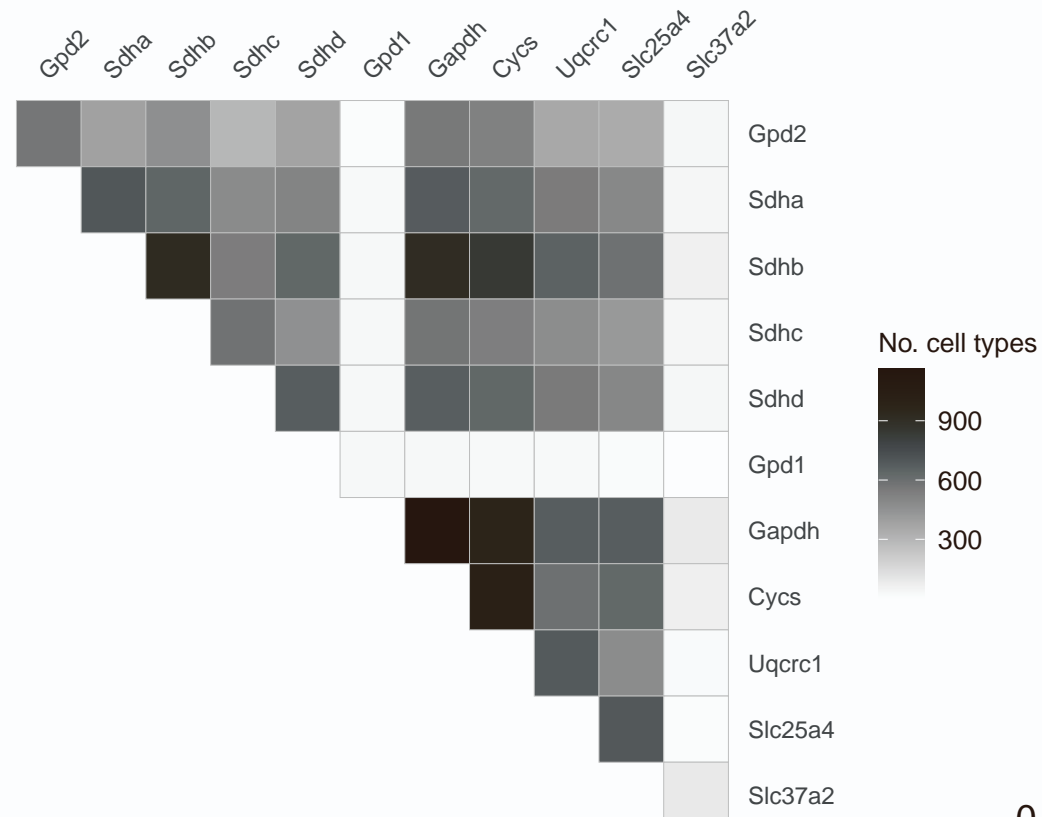
Lysophosphatidic Acid LPA2 Signalling (k_opt = 31)



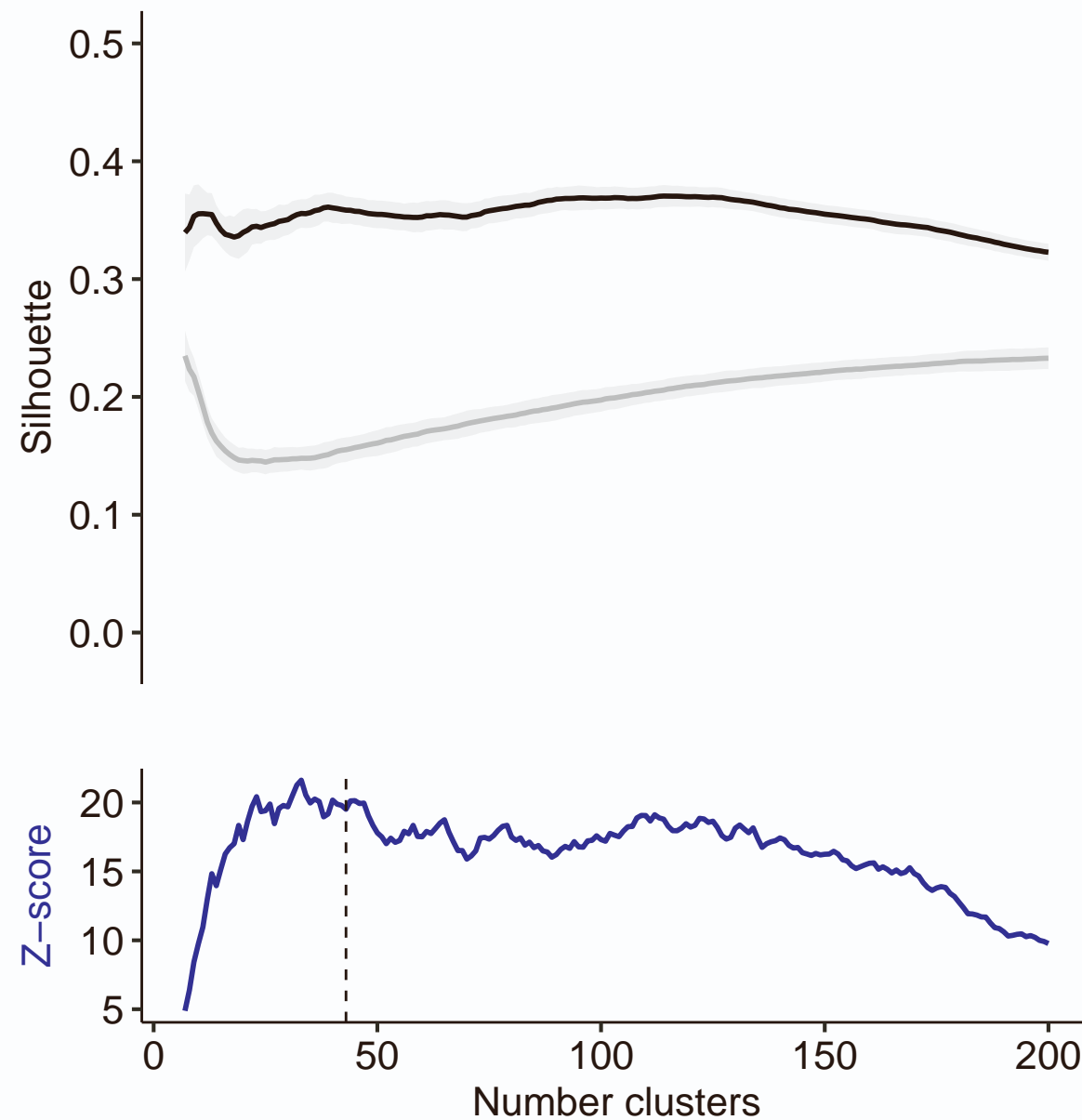
Mitochondrial Electron Transport Chain (k_opt = 43)

Gpd2
 Sdha
 Sdhb
 Sdhc
 Sdhd
 Gpd1
 Gapdh
 Cyts
 Uqcrc1
 Slc25a4
 Slc37a2

Gene Co-Expression Above Threshold



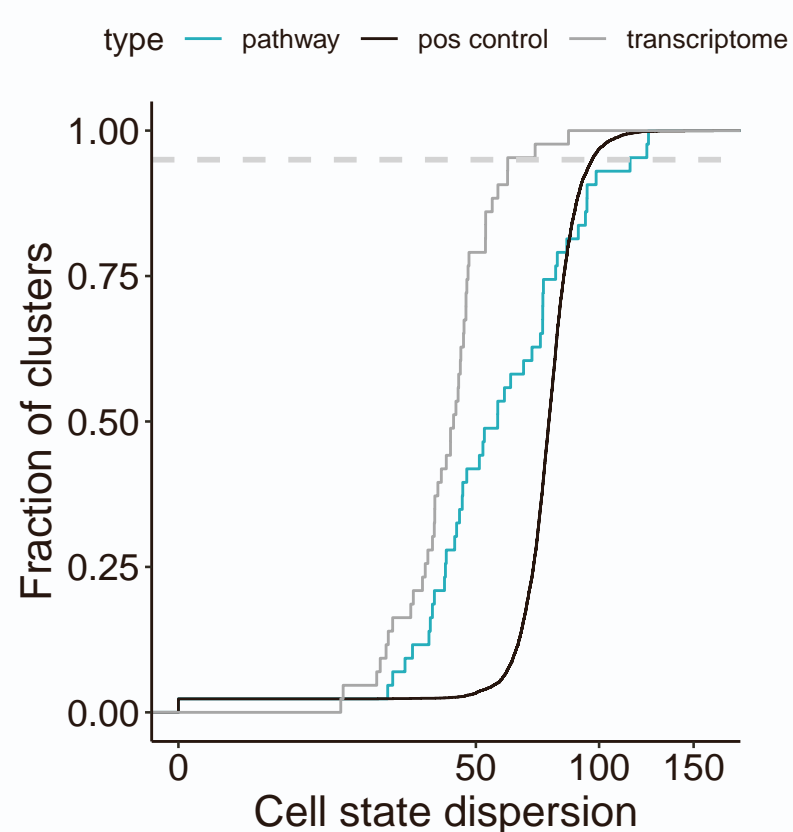
Silhouette and Z-score



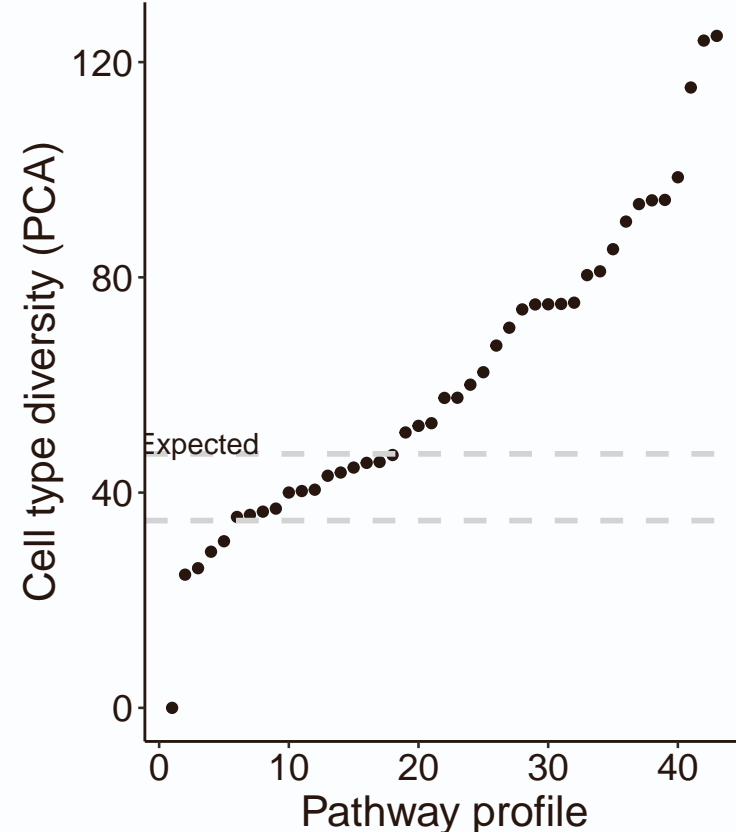
Global UMAP: Pathway ON



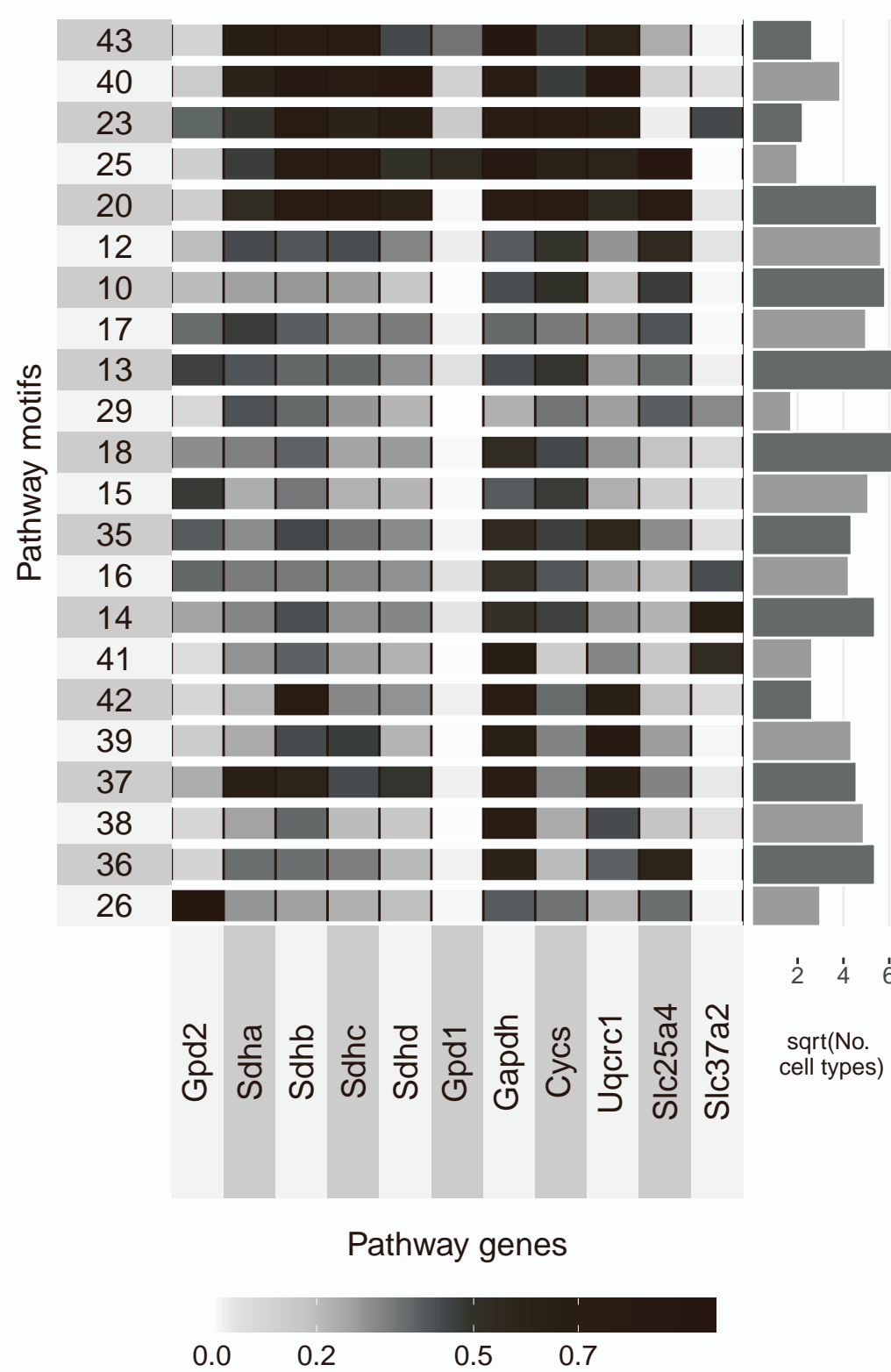
ECDF of Profile Dispersions



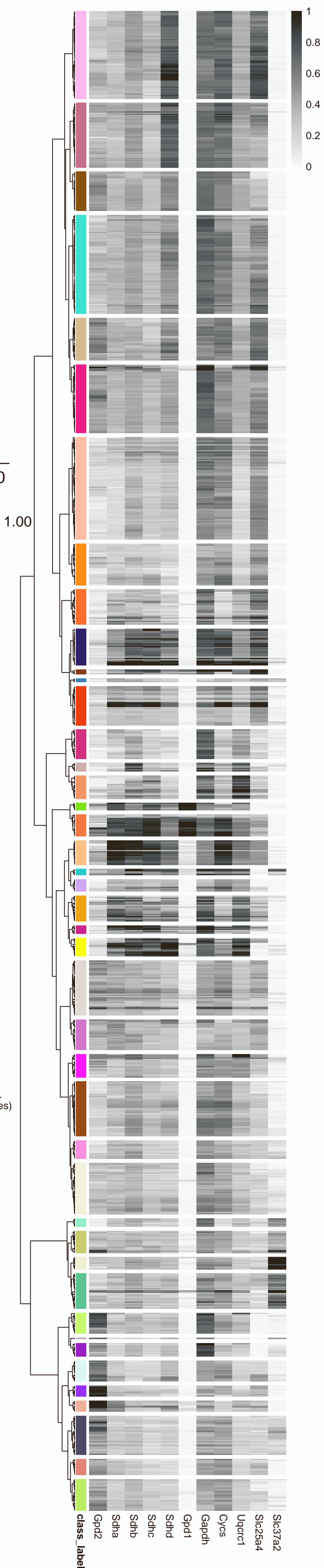
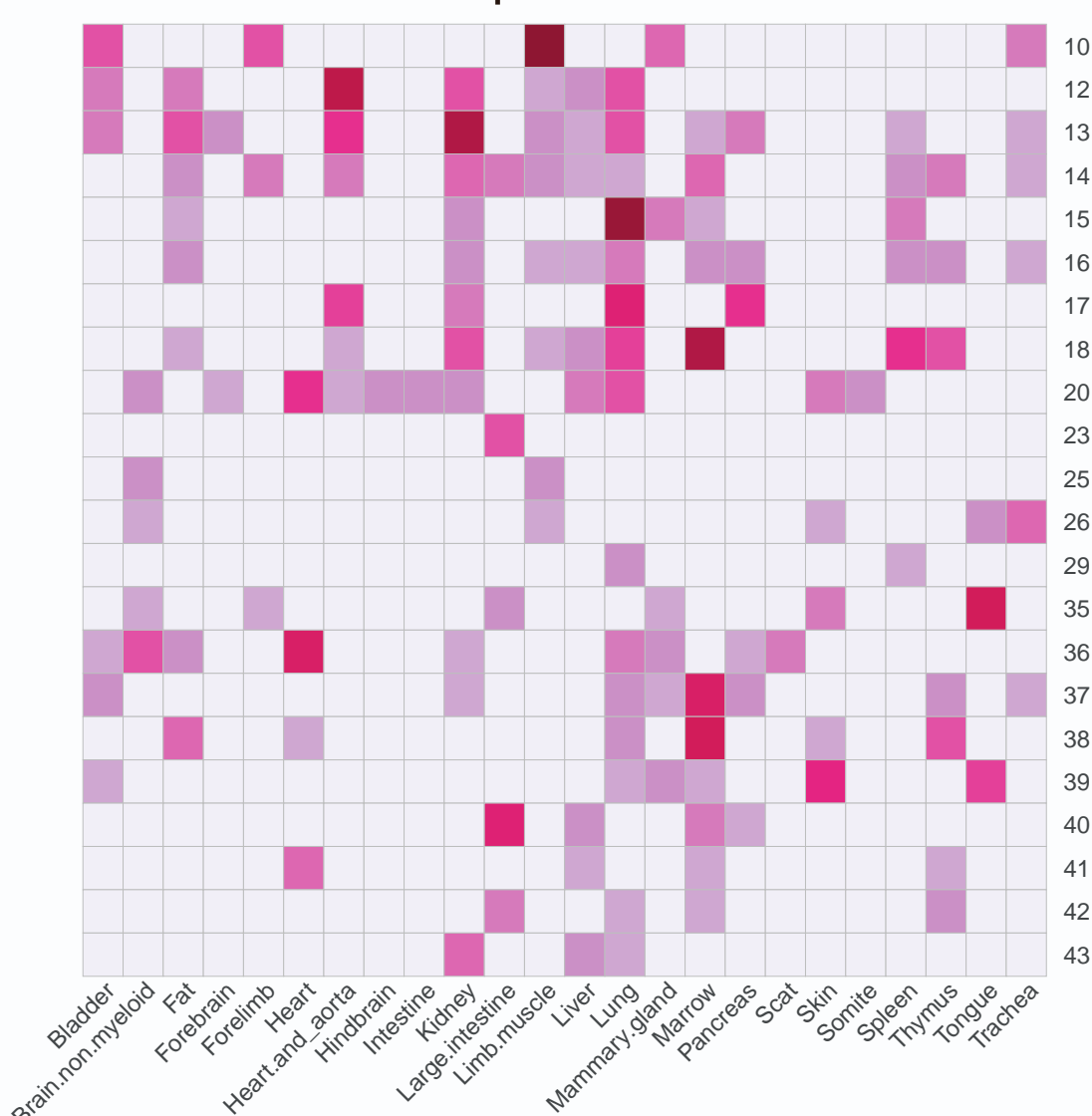
Profile Dispersion Plot



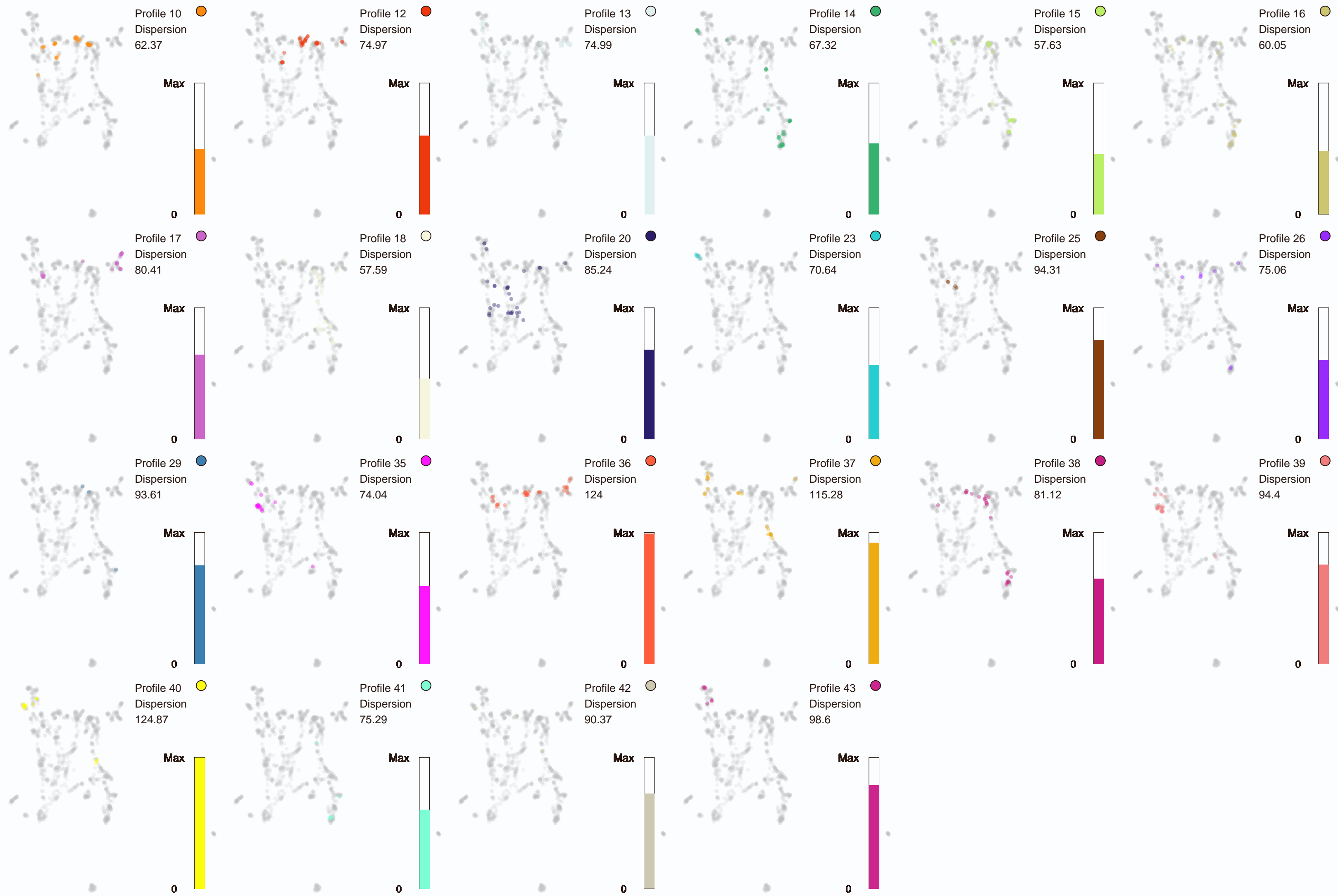
Motifs, Dispersion $\geq 90^{\text{th}}$ percentile



Motif Tissue Composition



Mitochondrial Electron Transport Chain (k_opt = 43)



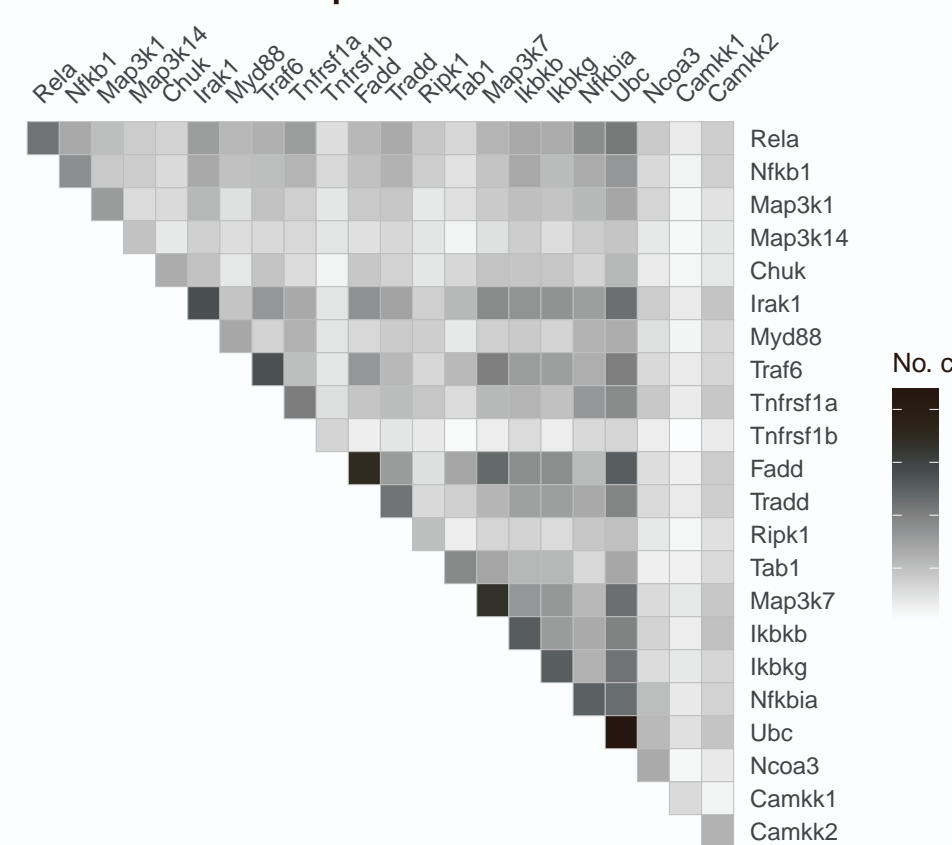
NF- κ B Signaling Pathway (k_opt = 113)

Rela
Nfkb1
Map3k1
Map3k14
Chuk
Irak1
Myd88
Traf6

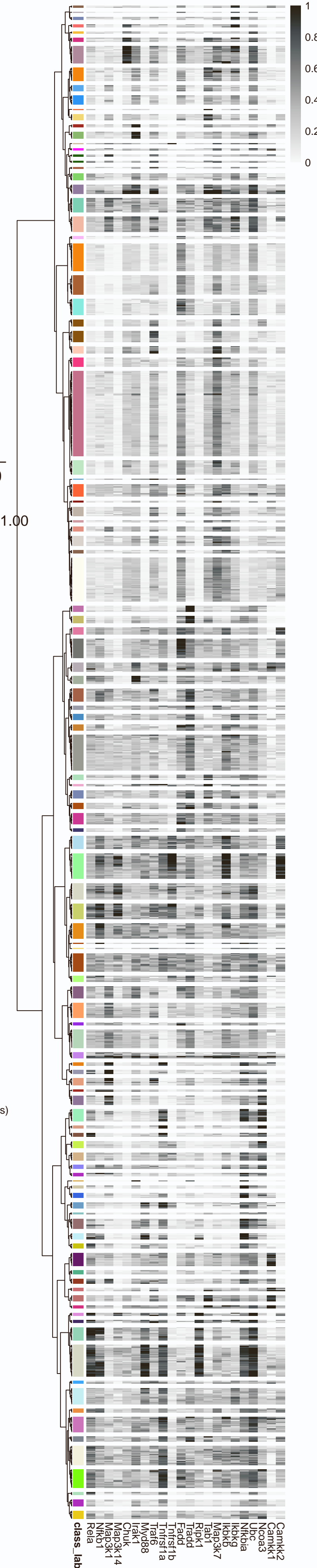
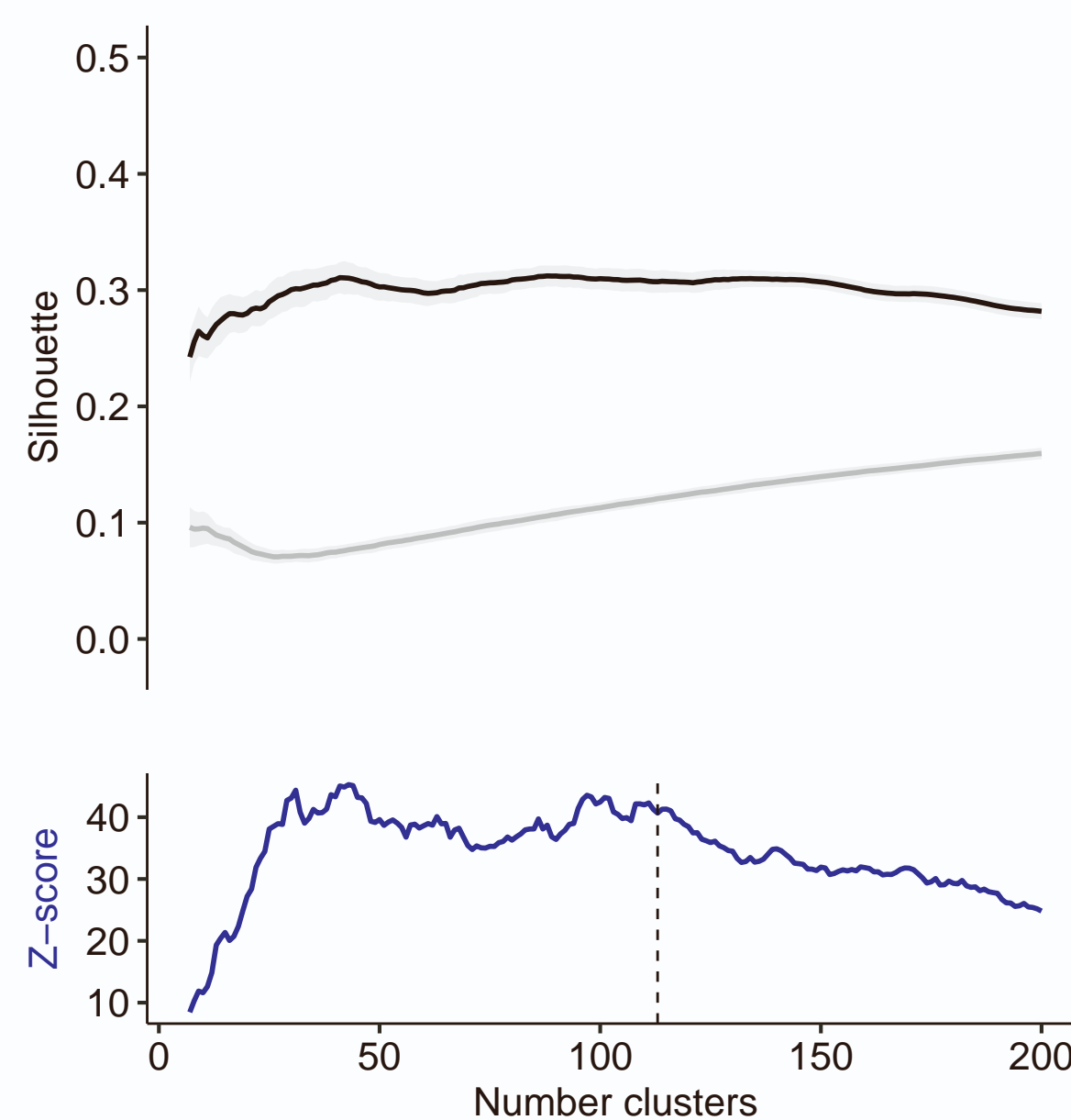
Pathway genes
Tnfrsf1a
Tnfrsf1b
Fadd
Tradd
Ripk1
Tab1
Map3k7

Ikbkb
Ikbkg
Nfkbia
Ubc
Ncoa3
Camkk1
Camkk2

Gene Co-Expression Above Threshold



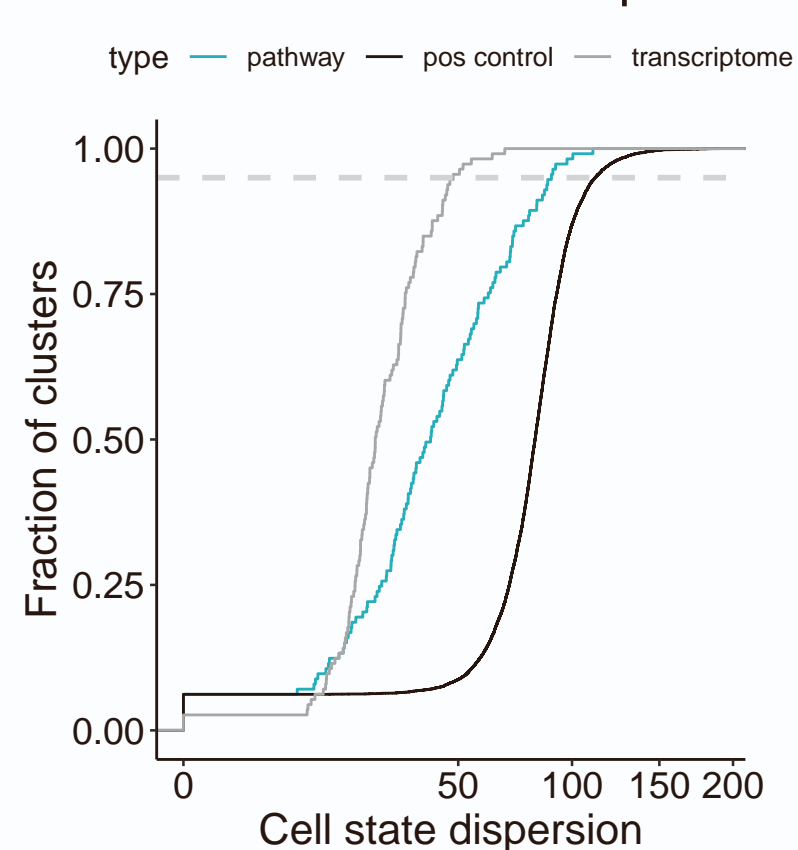
Silhouette and Z-score



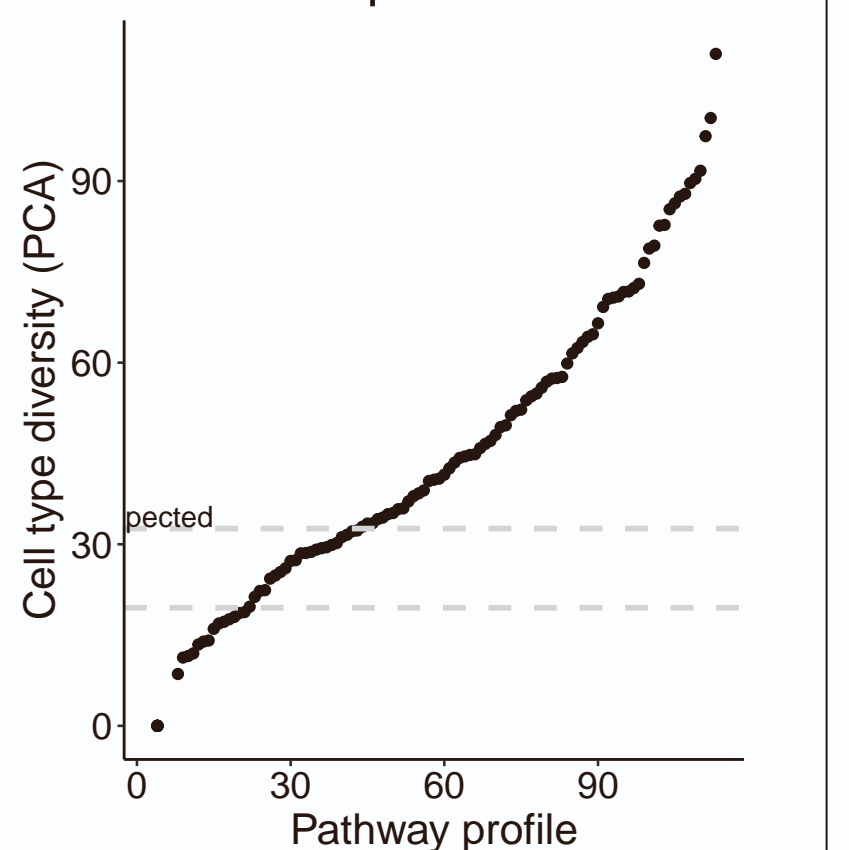
Global UMAP: Pathway ON



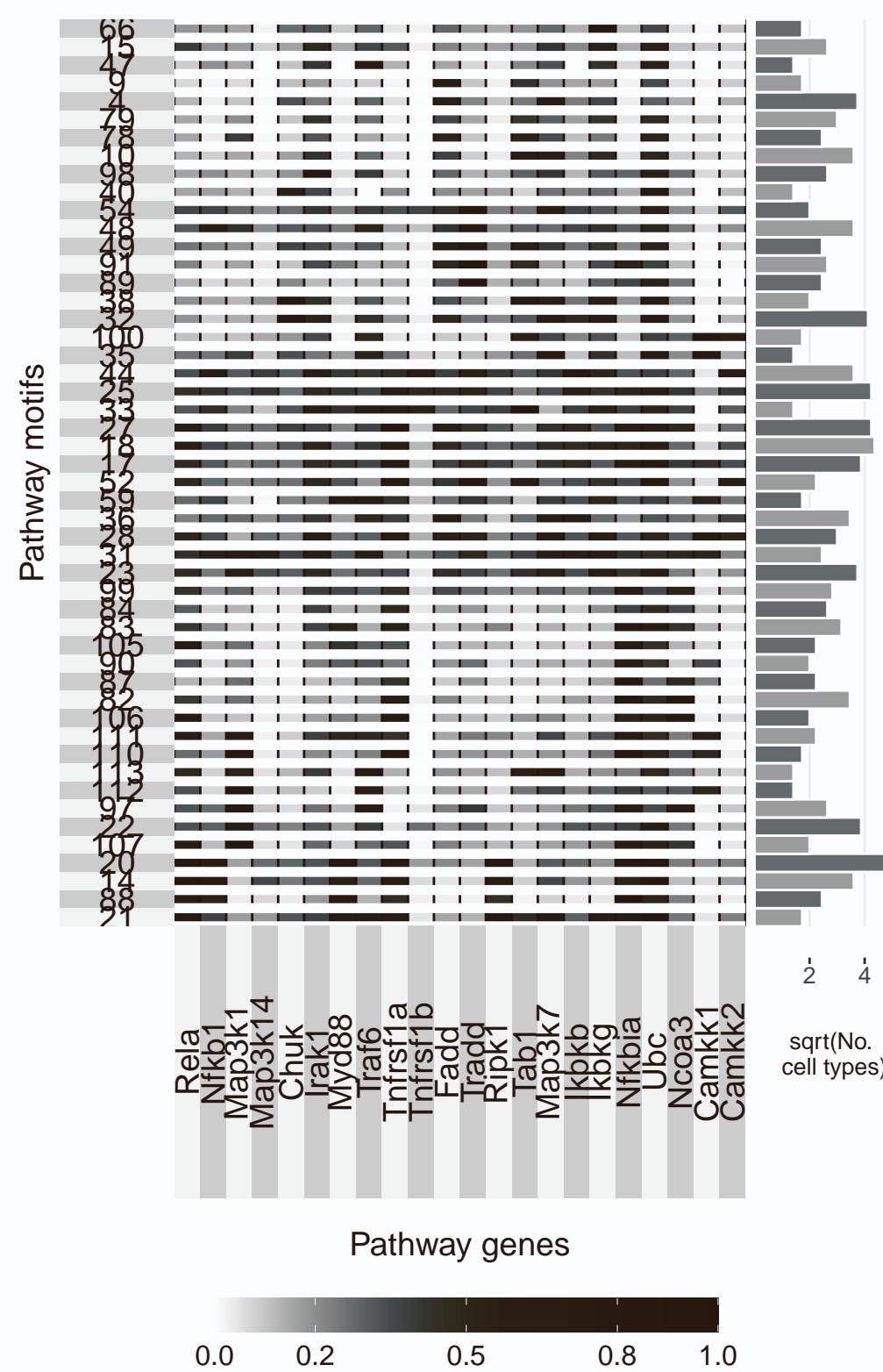
ECDF of Profile Dispersions



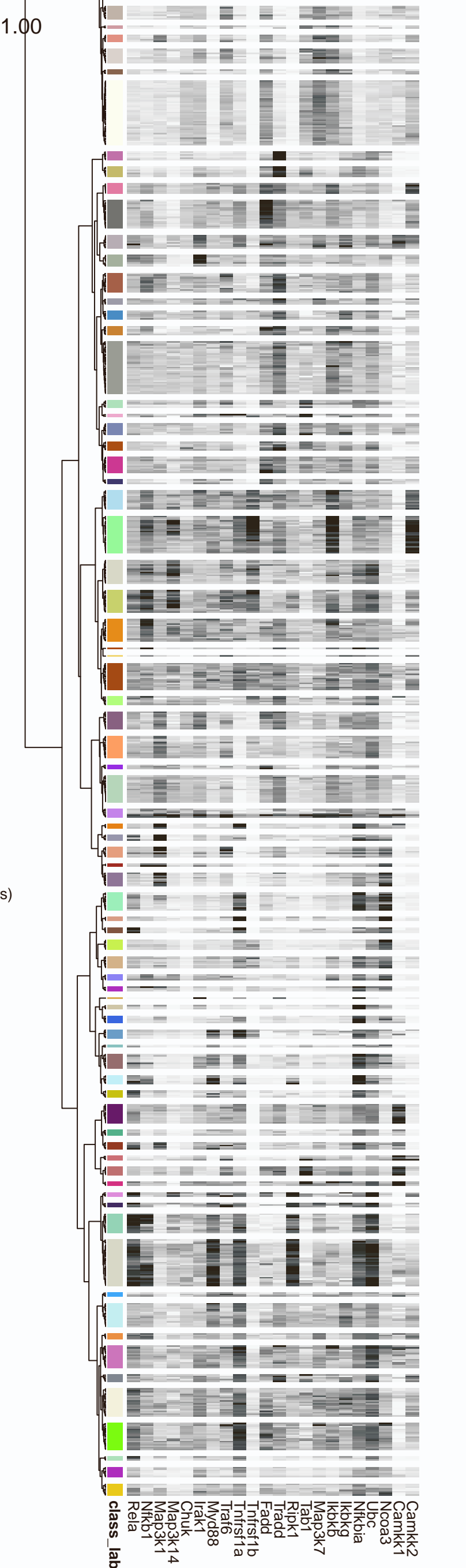
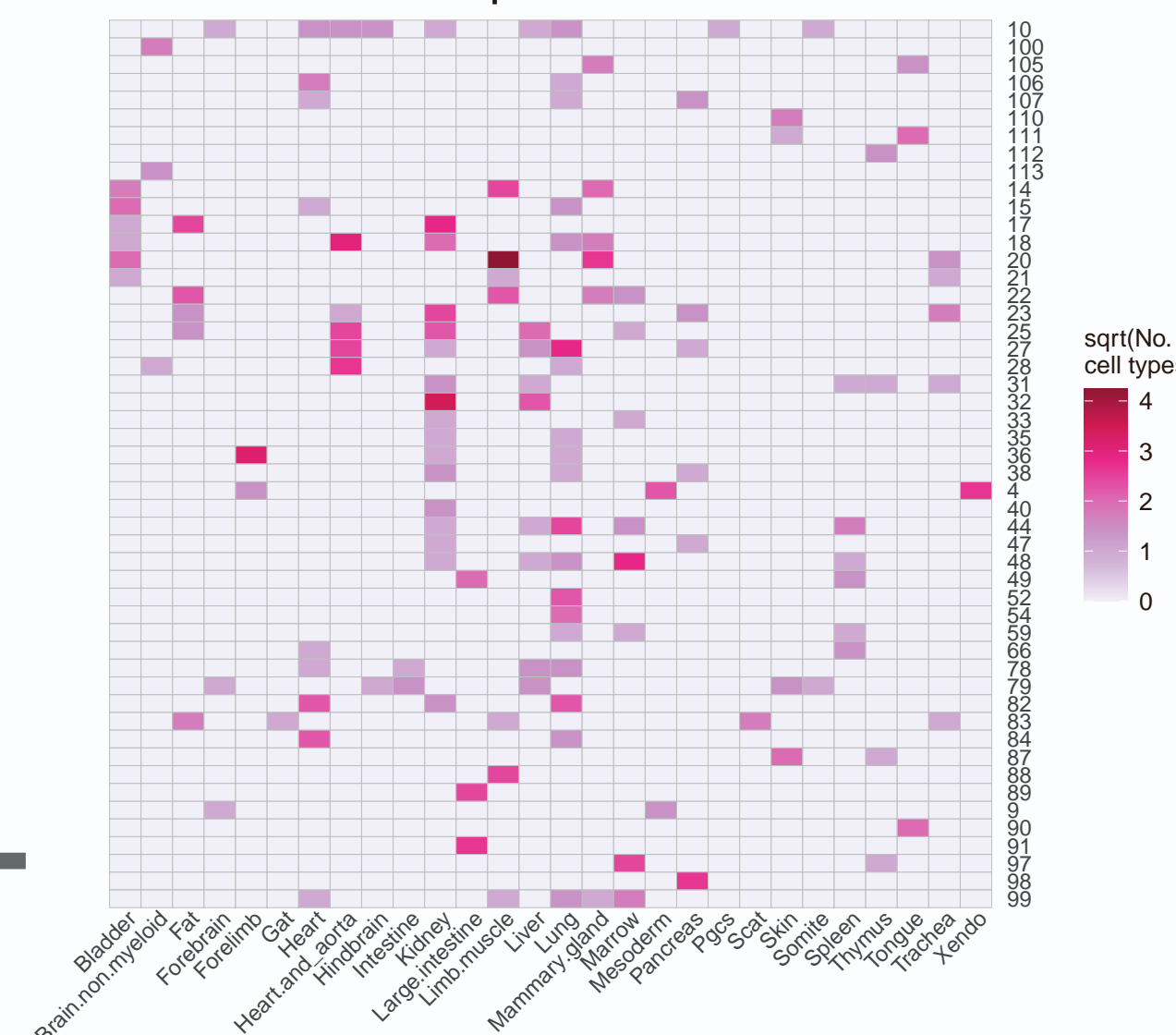
Profile Dispersion Plot



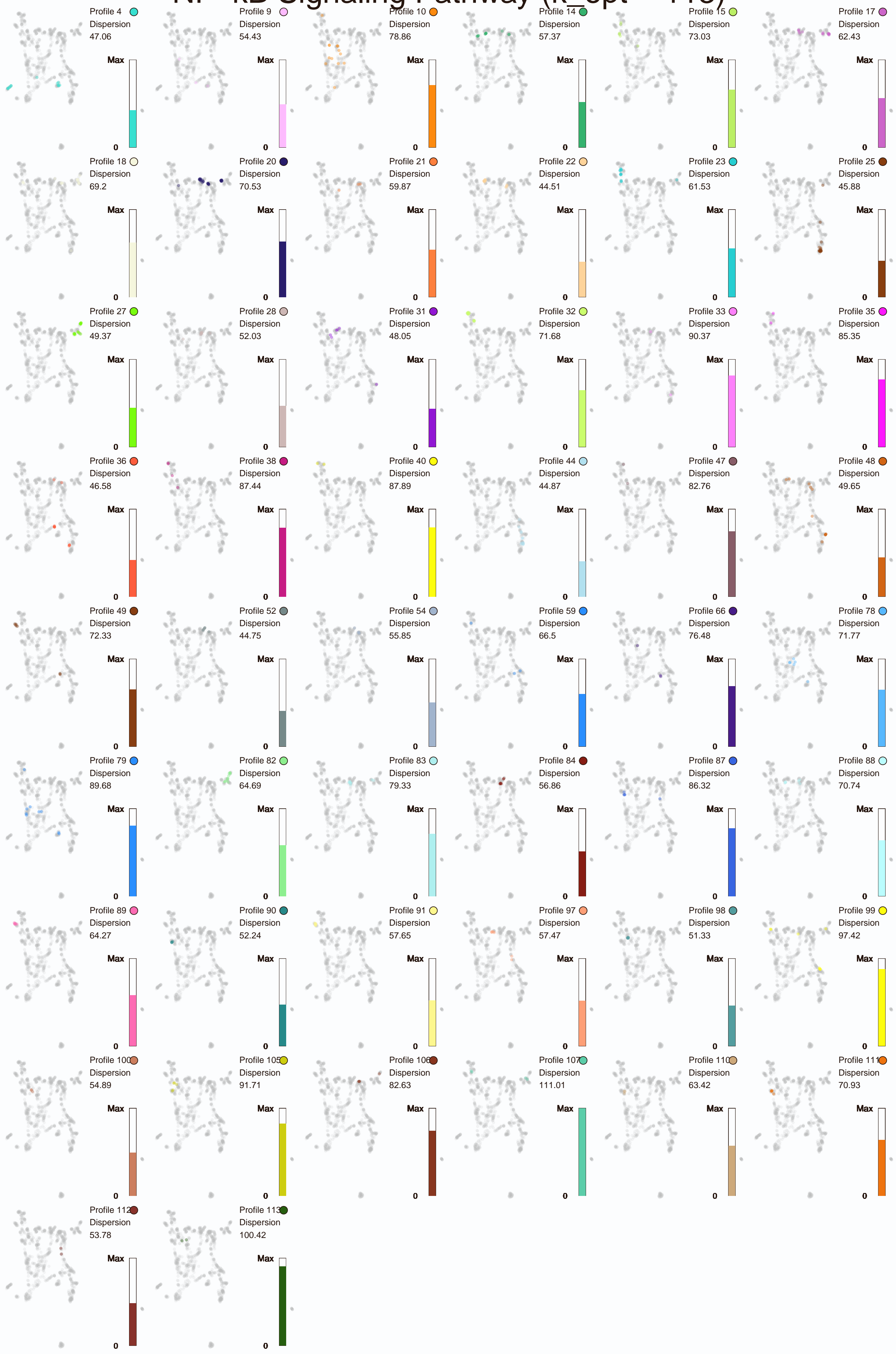
Motifs, Dispersion $\geq 90^{\text{th}}$ percentile



Motif Tissue Composition



NF- κ B Signaling Pathway ($k_{opt} = 113$)



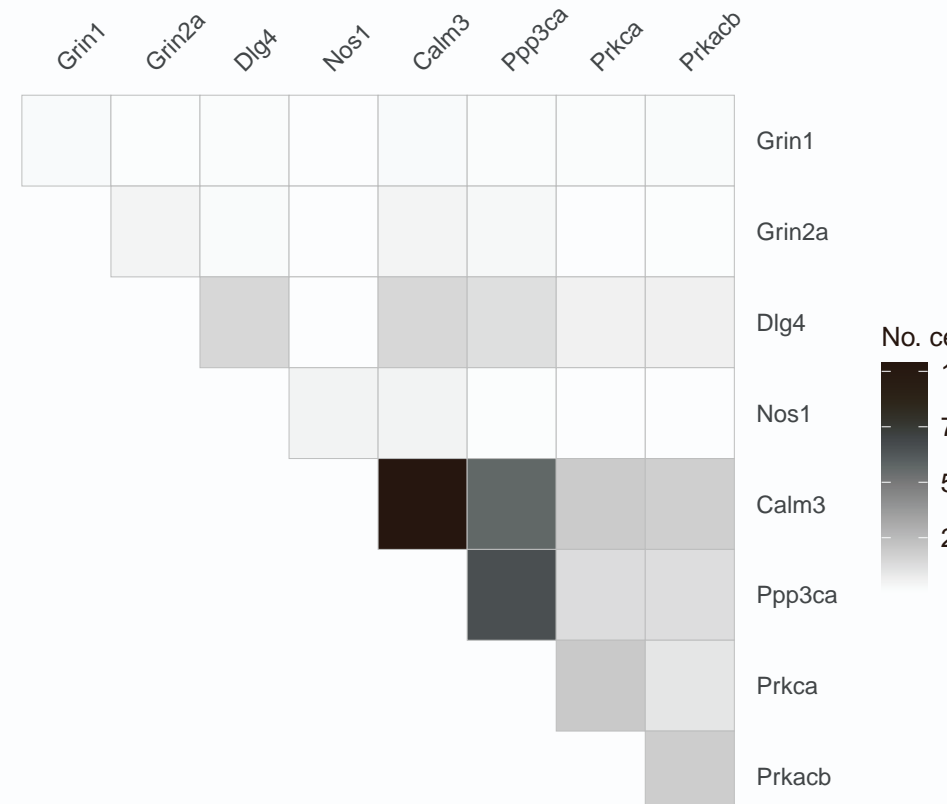
Nitric Oxide Signaling Pathway (k_opt = 28)

Grin1
Grin2a
Dlg4

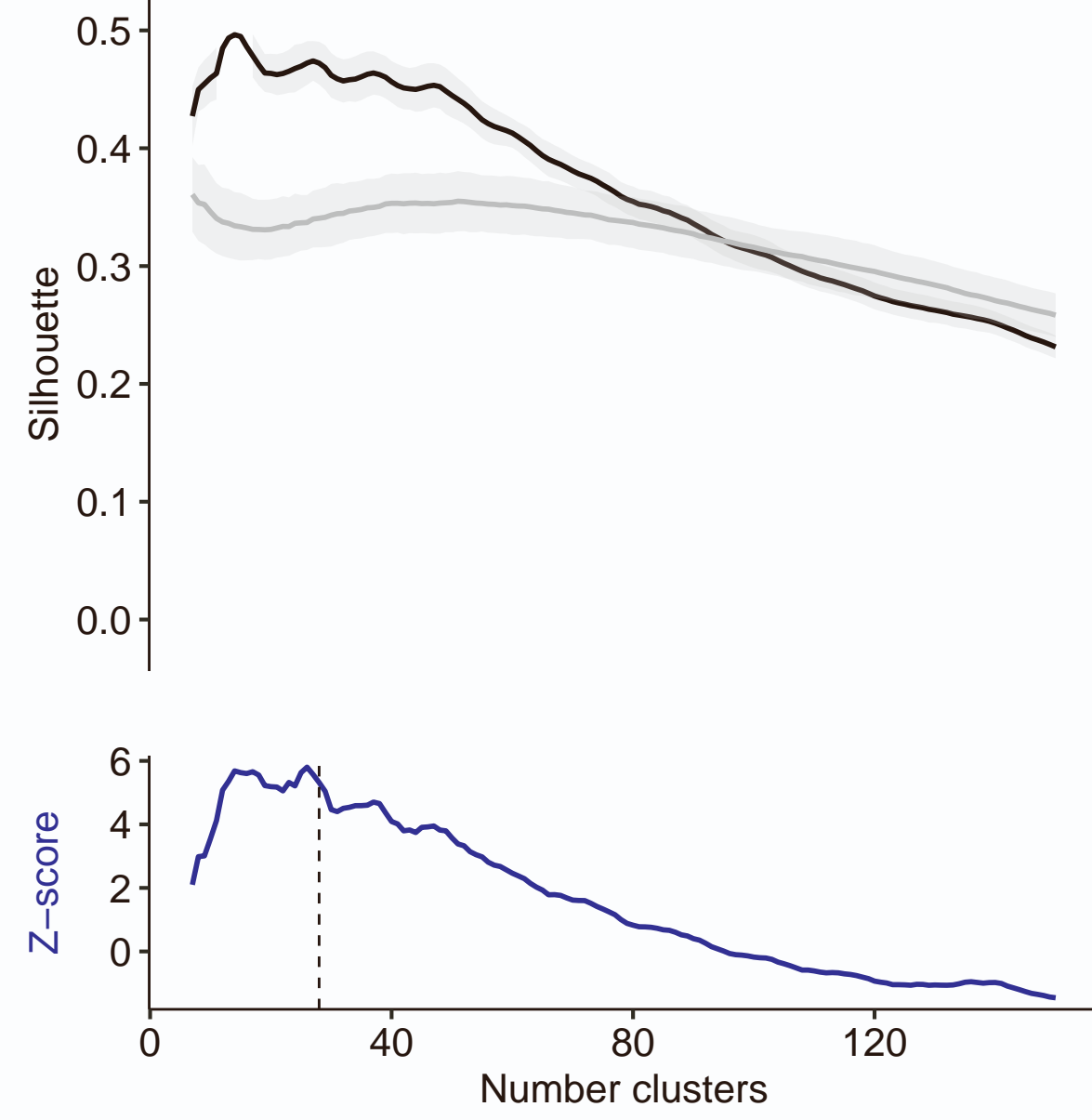
Pathway genes
Nos1
Calm3

Ppp3ca
Prkca
Prkacb

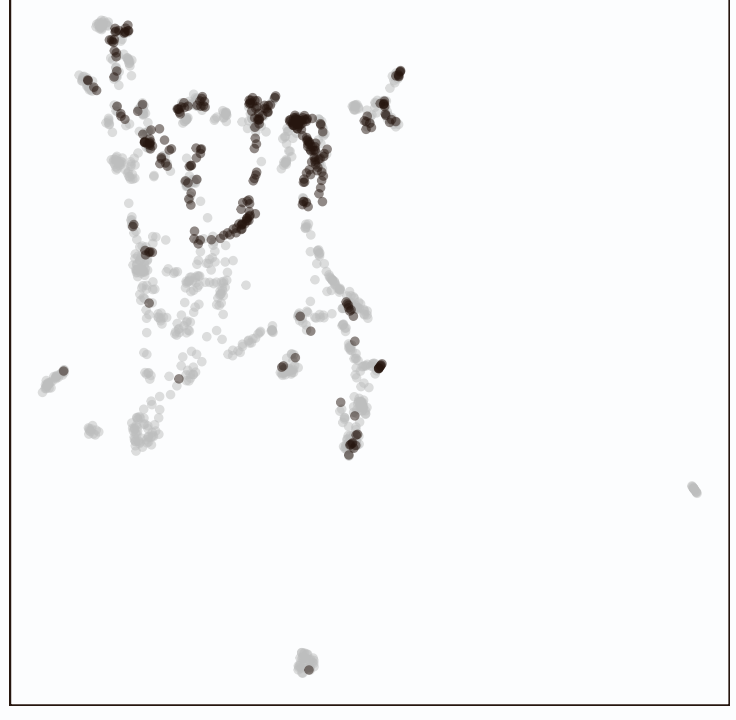
Gene Co-Expression Above Threshold



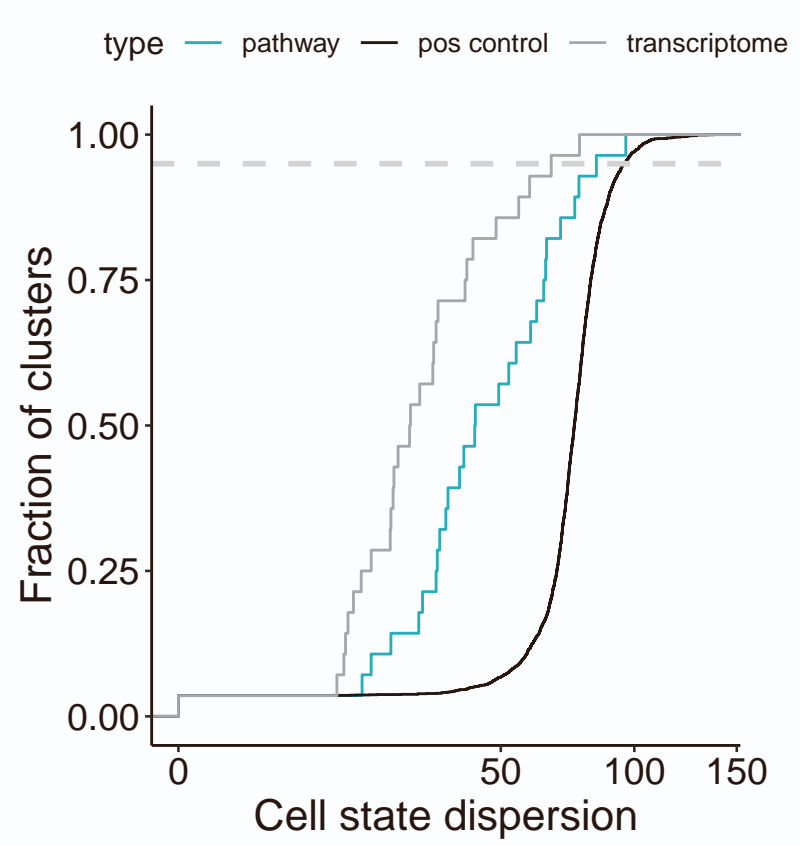
Silhouette and Z-score



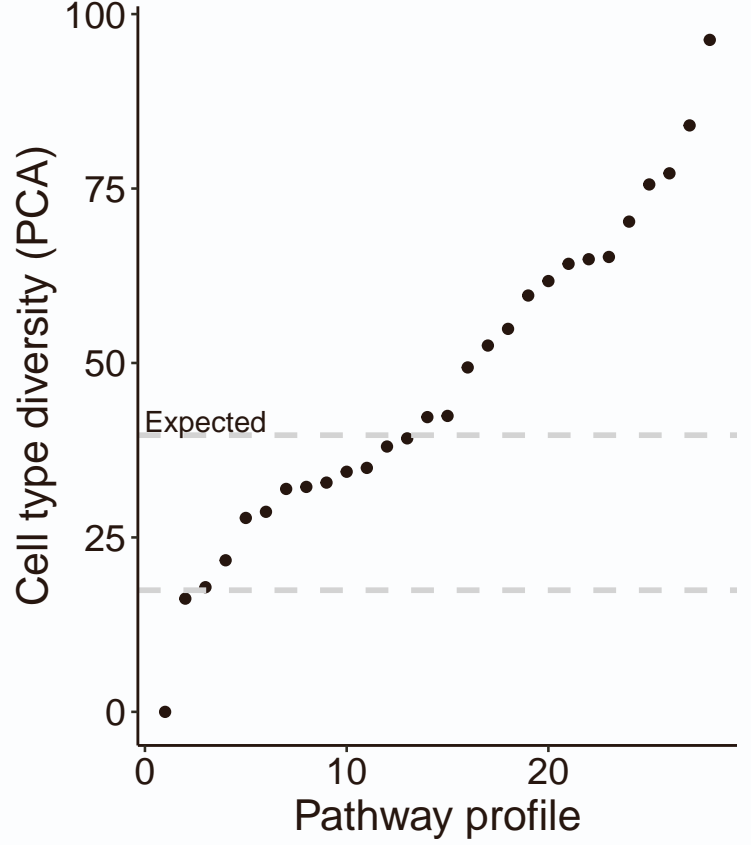
Global UMAP: Pathway ON



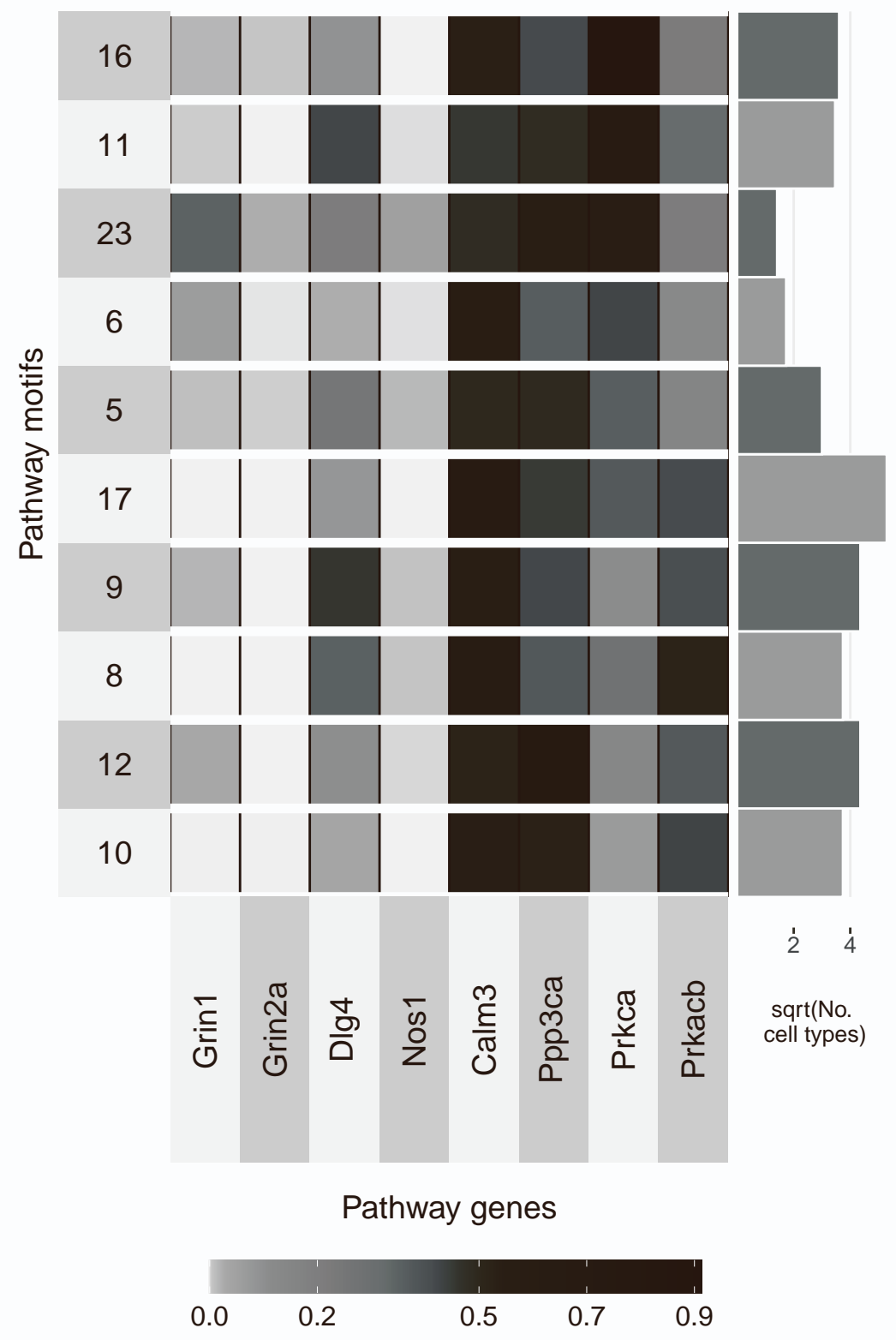
ECDF of Profile Dispersions



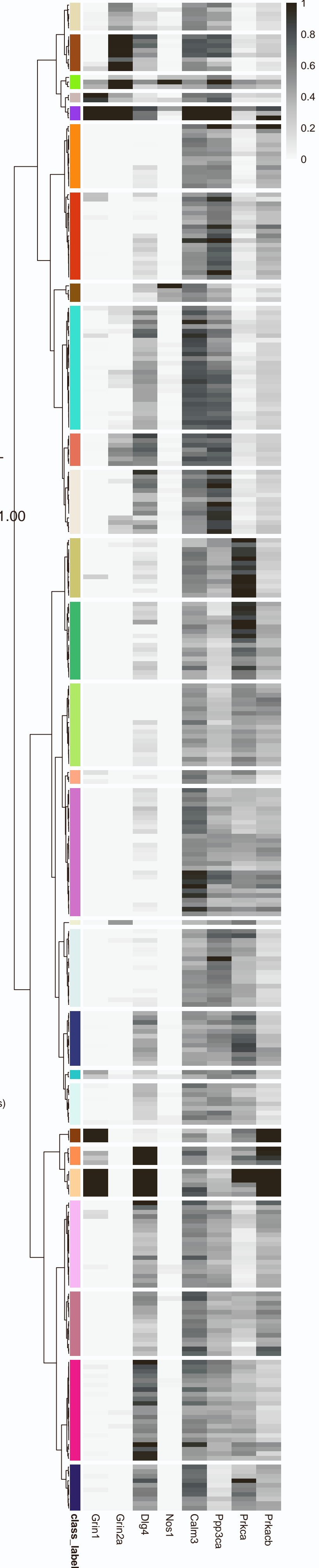
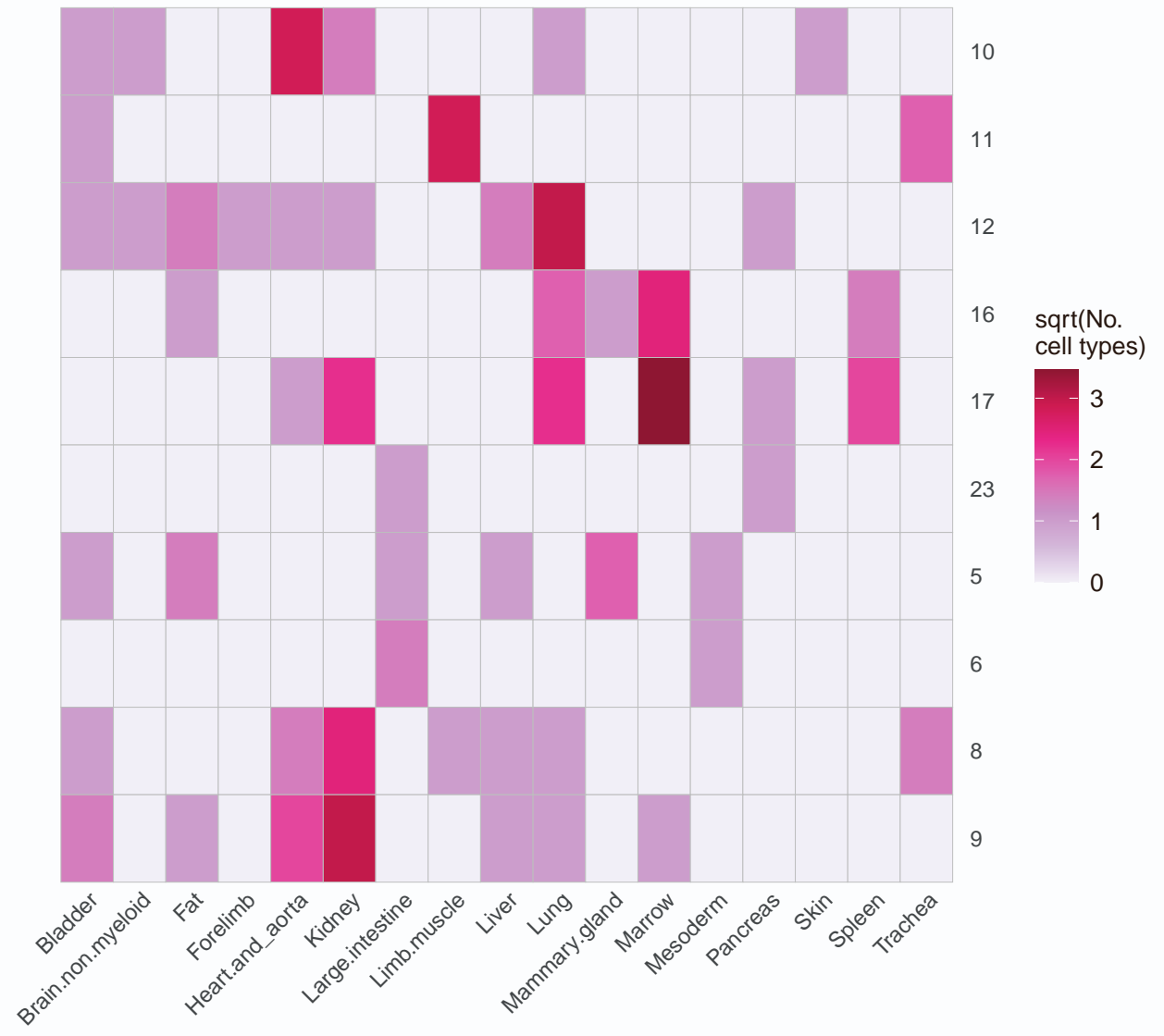
Profile Dispersion Plot



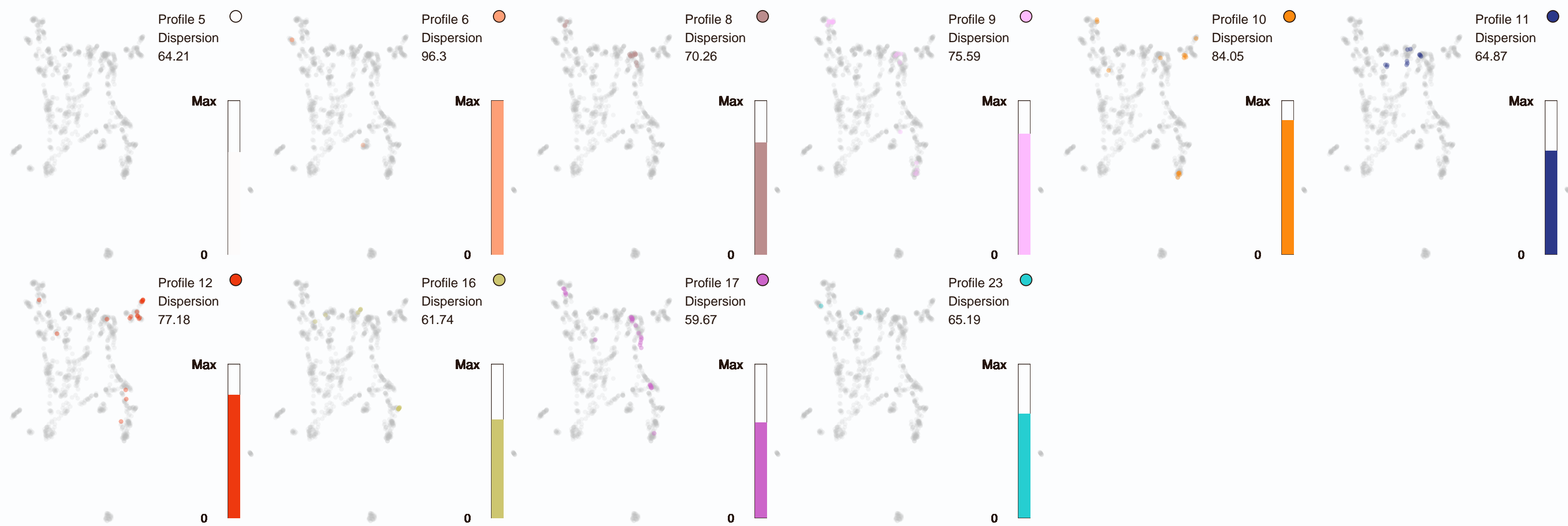
Motifs, Dispersion $\geq 90^{\text{th}}$ percentile



Motif Tissue Composition



Nitric Oxide Signaling Pathway (k_opt = 28)

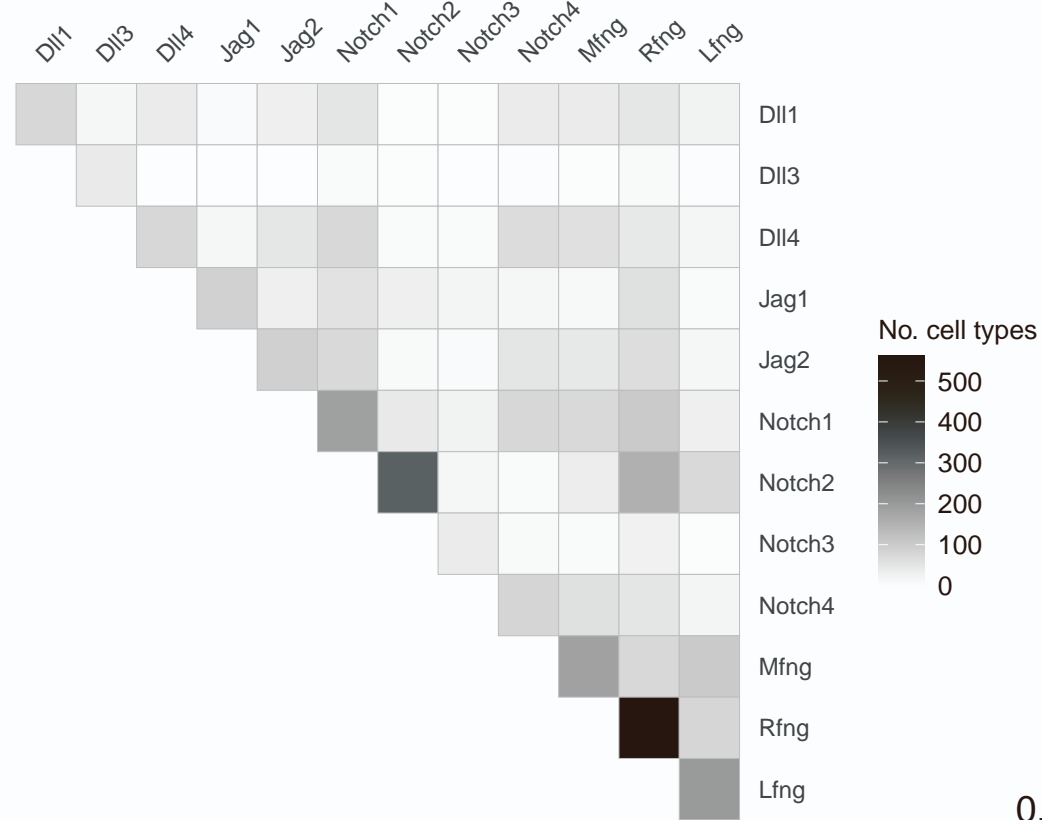


Notch receptors, Dll ligands and Fringe proteins (k_opt = 30)

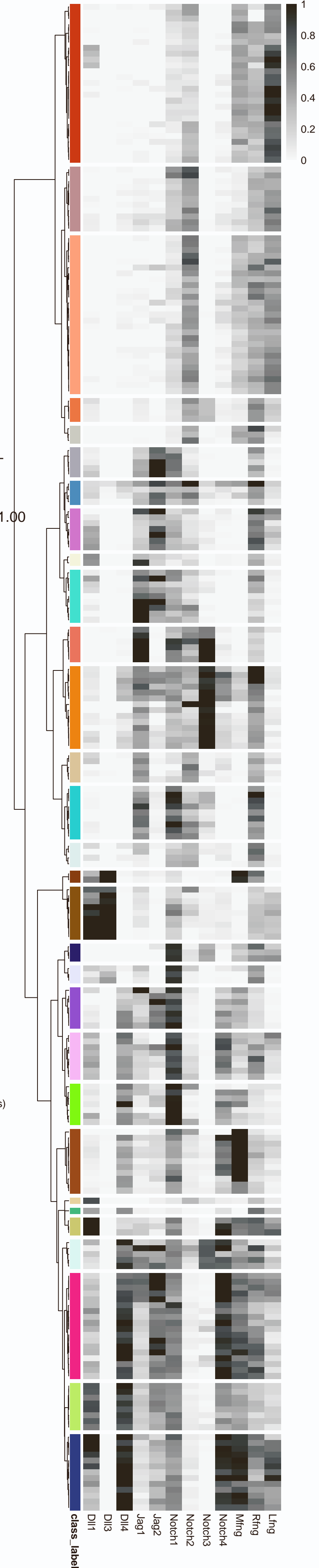
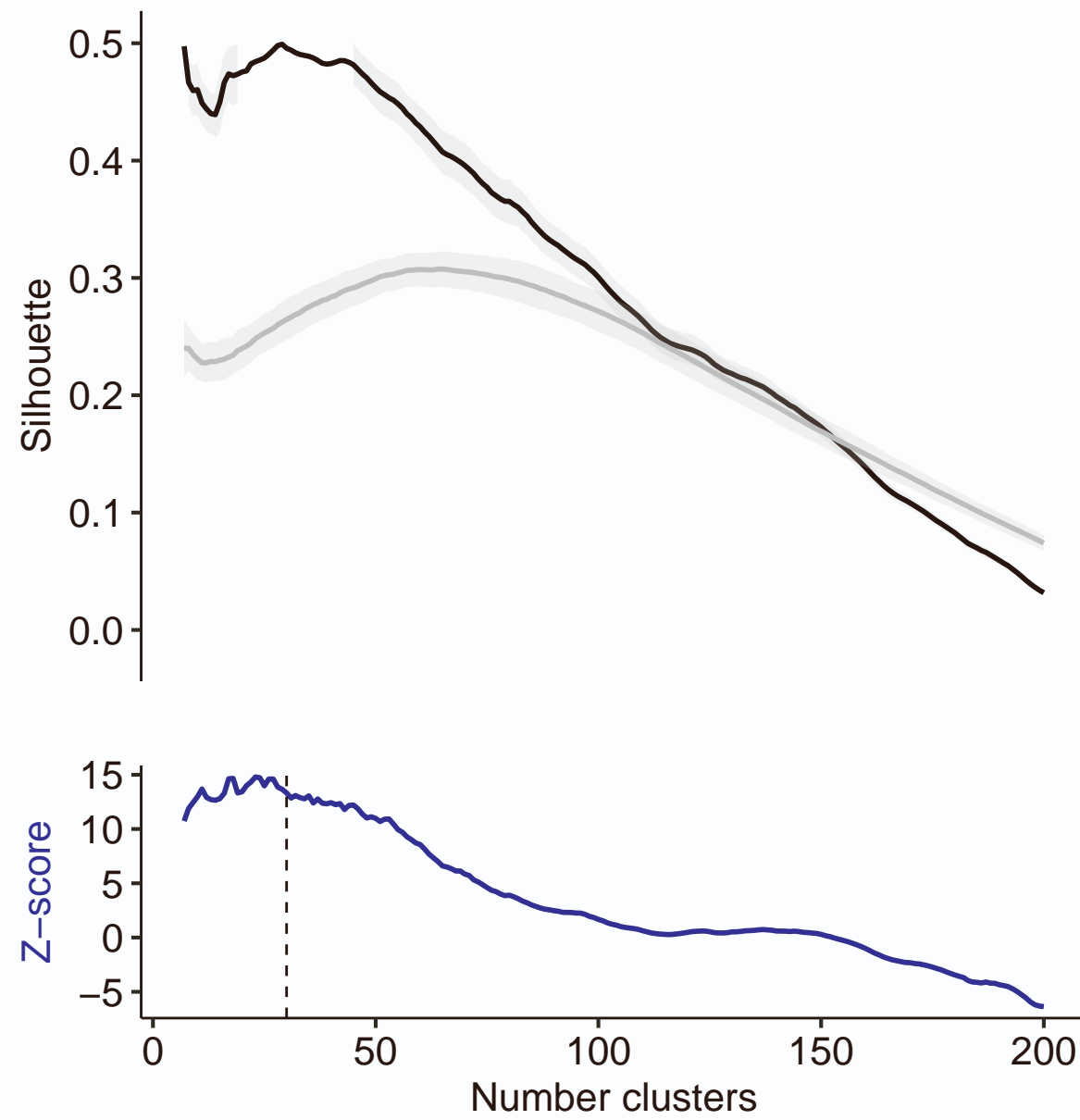
Pathway genes

- Dll1
- Dll3
- Dll4
- Jag1
- Jag2
- Notch1
- Notch2
- Notch3
- Notch4
- Mfng
- Rfng
- Lfng

Gene Co-Expression Above Threshold



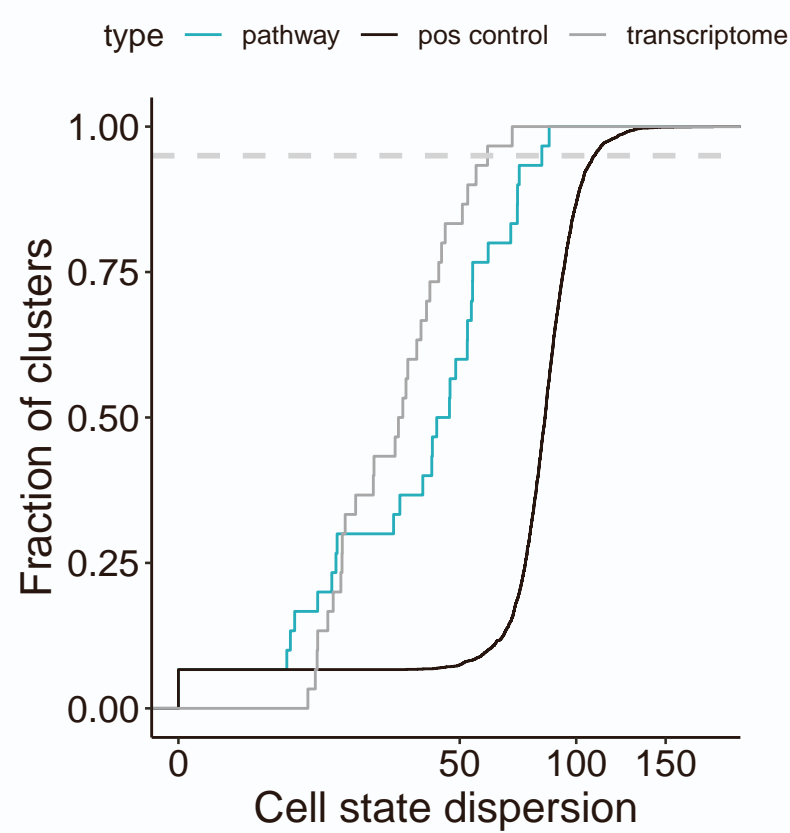
Silhouette and Z-score



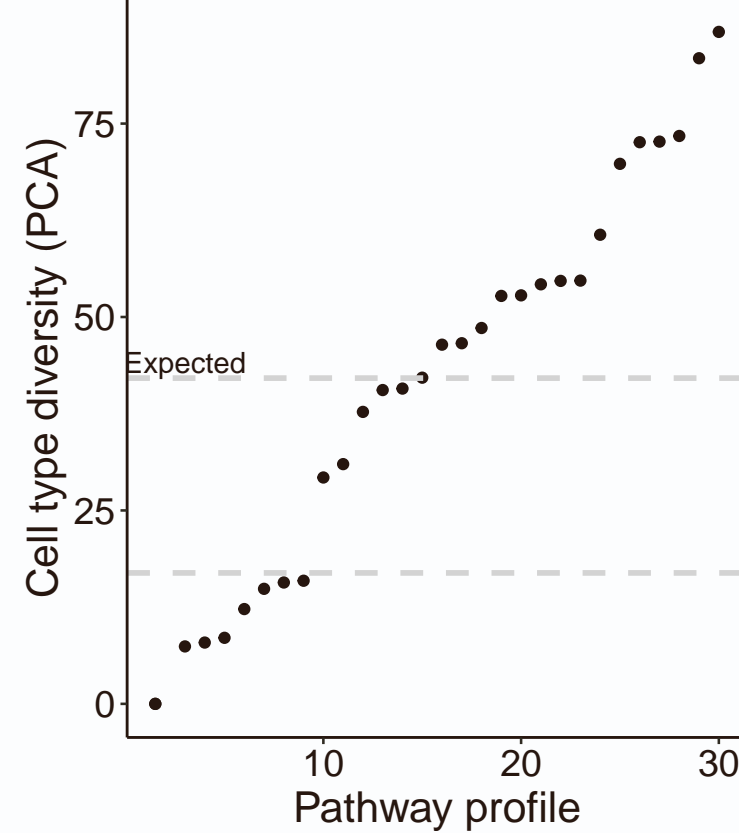
Global UMAP: Pathway ON



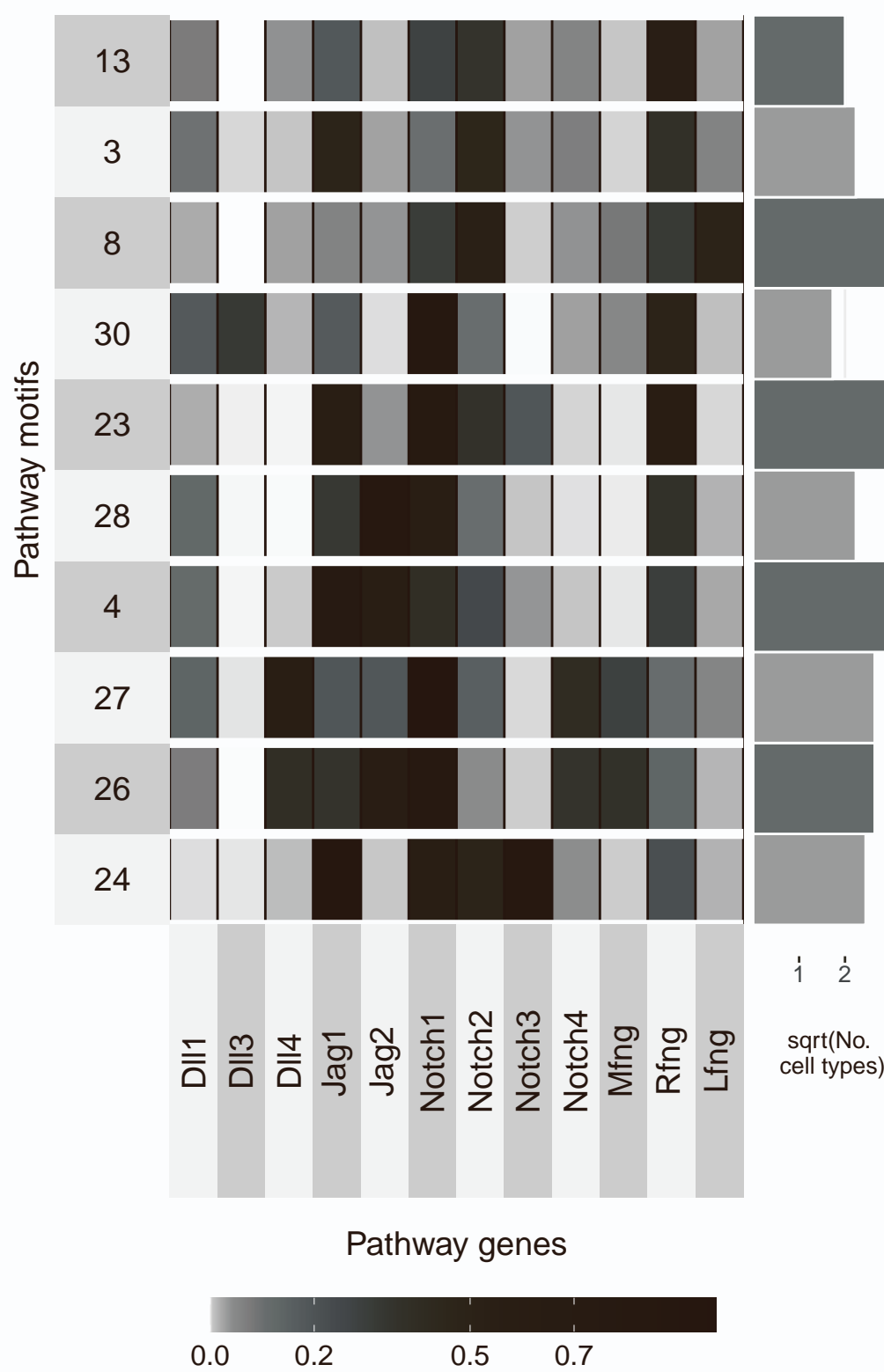
ECDF of Profile Dispersions



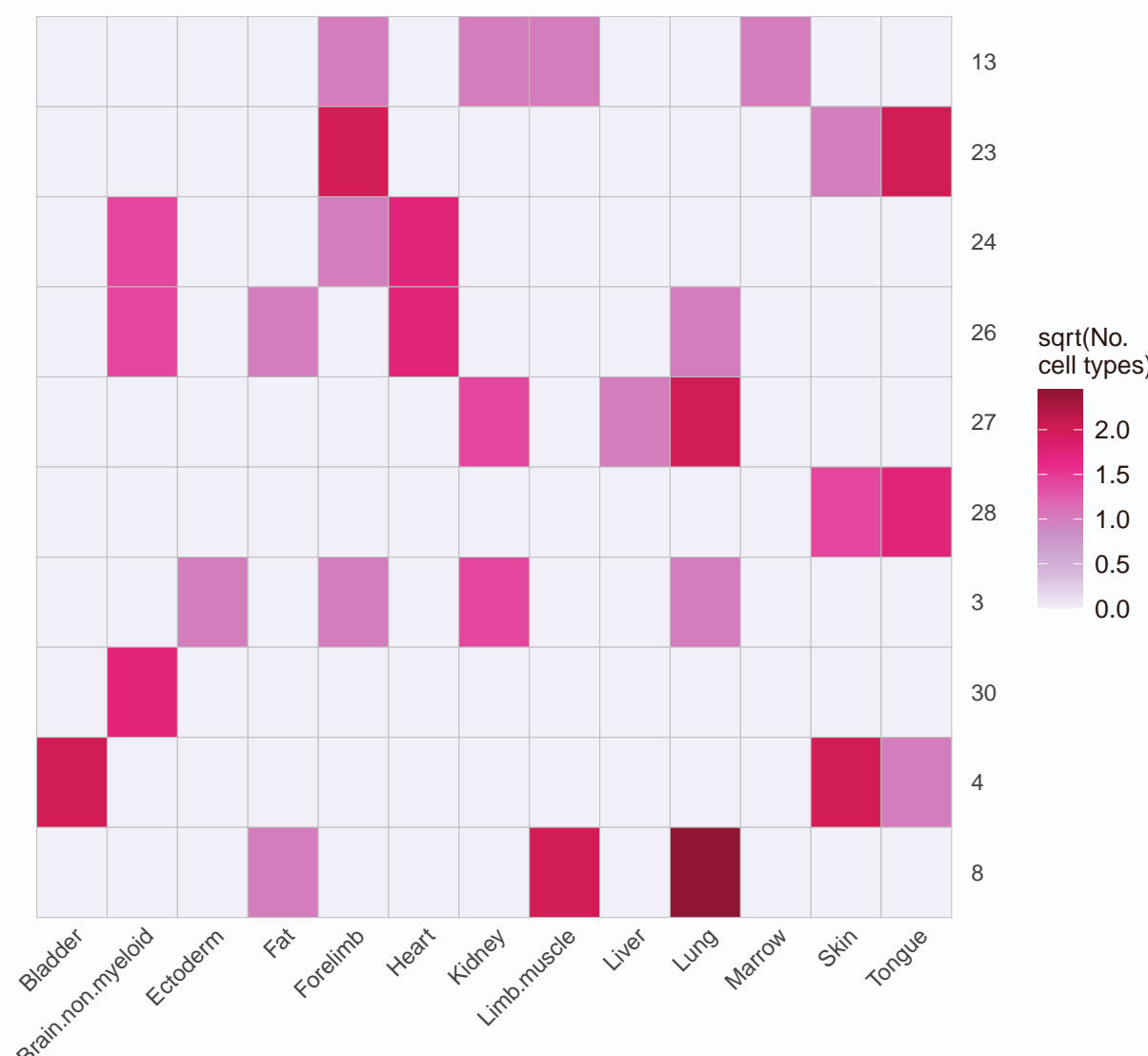
Profile Dispersion Plot



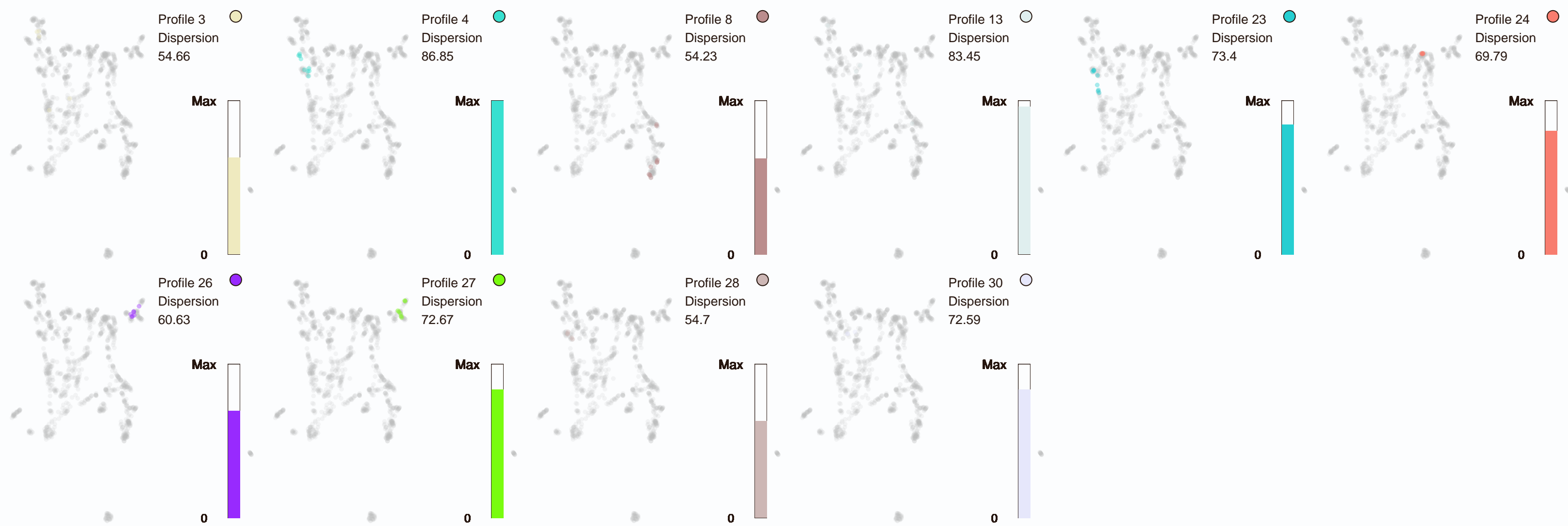
Motifs, Dispersion $\geq 90^{\text{th}}$ percentile



Motif Tissue Composition



Notch receptors, Dll ligands and Fringe proteins (k_opt = 30)

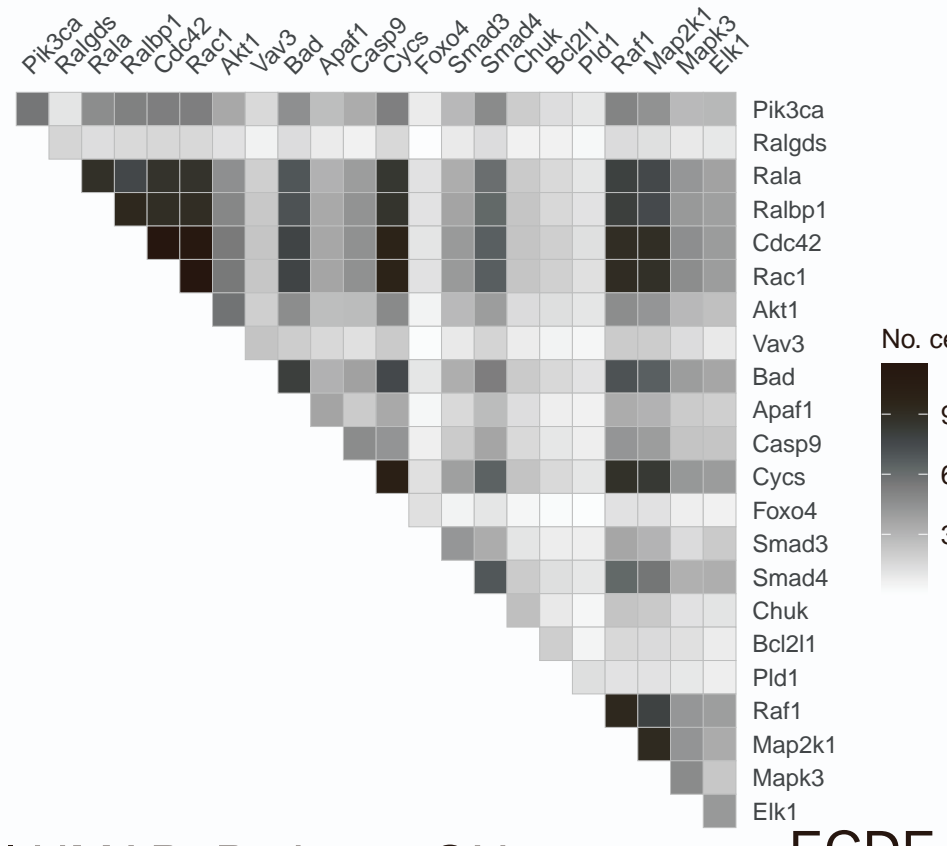


Ras Signaling Pathway (k_opt = 83)

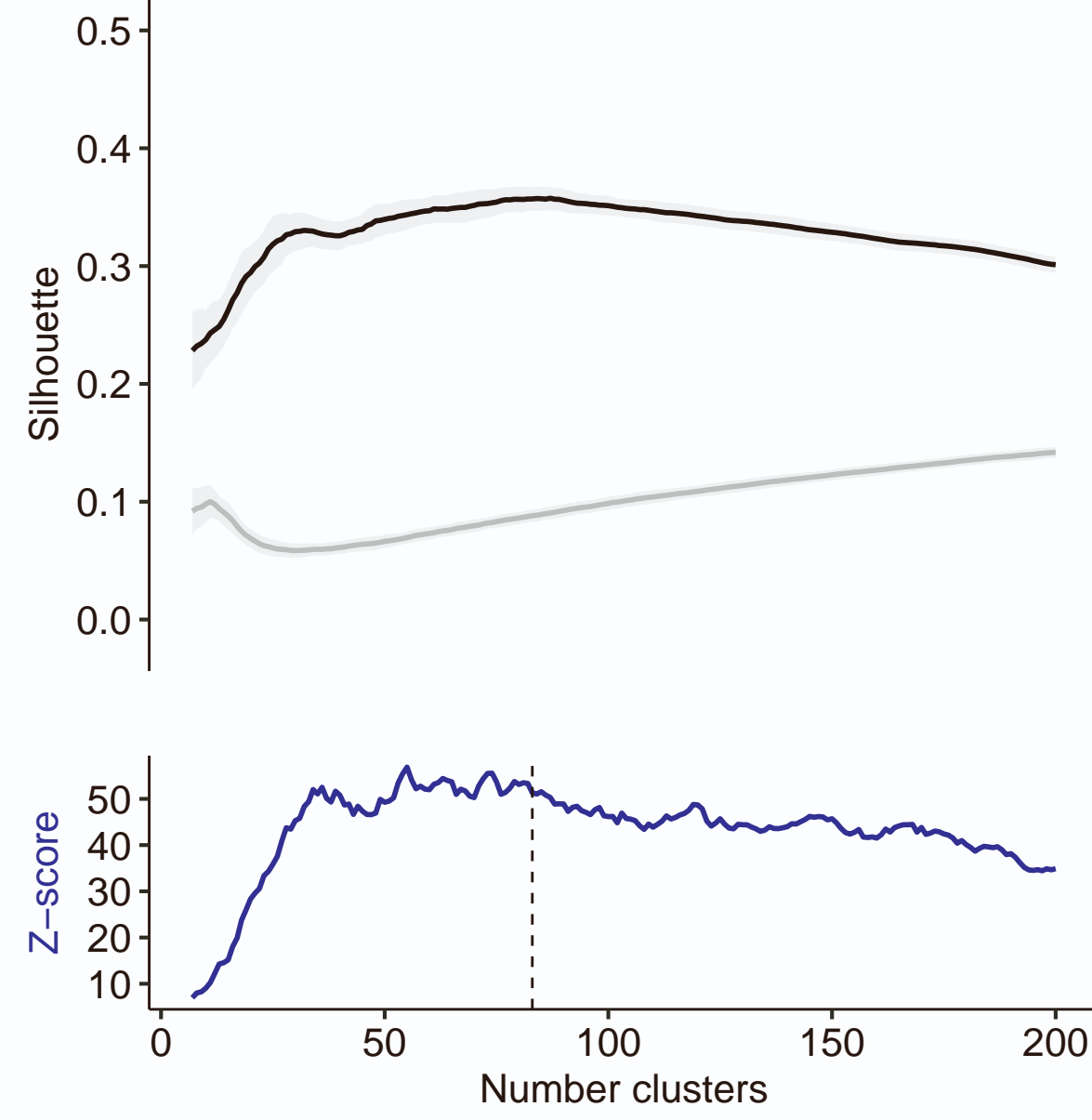
Pathway genes

Pik3ca	Bad	Chuk
Ralgsd	Apaf1	Bcl2l1
Rala	Casp9	Pld1
Ralbp1	Cyts	Raf1
Cdc42	Foxo4	Map2k1
Rac1	Smad3	Mapk3
Akt1	Smad4	Elk1
Vav3		

Gene Co-Expression Above Threshold



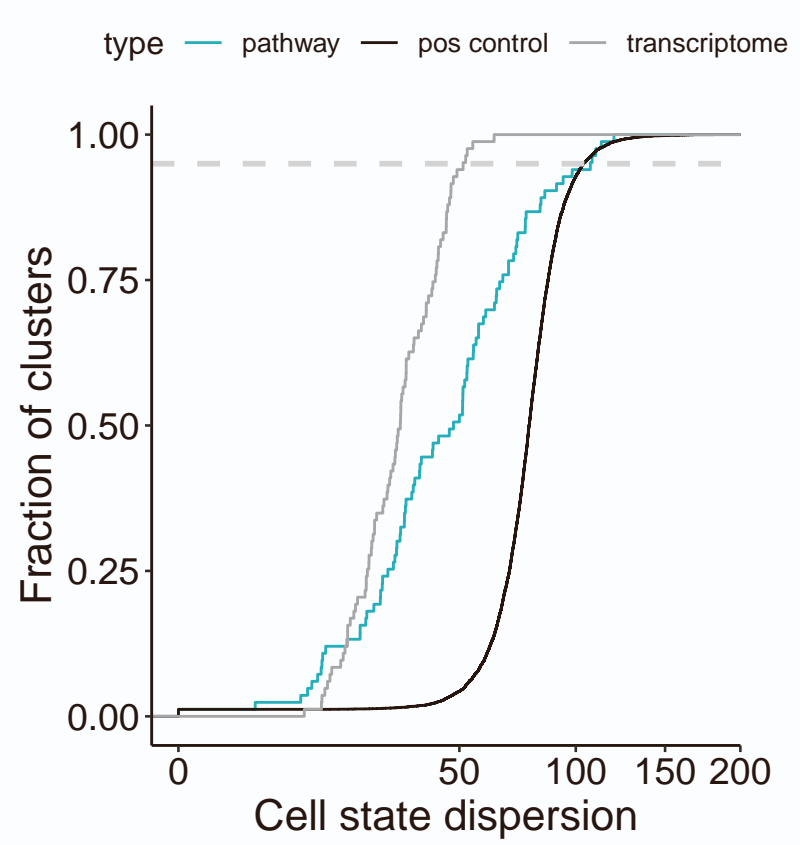
Silhouette and Z-score



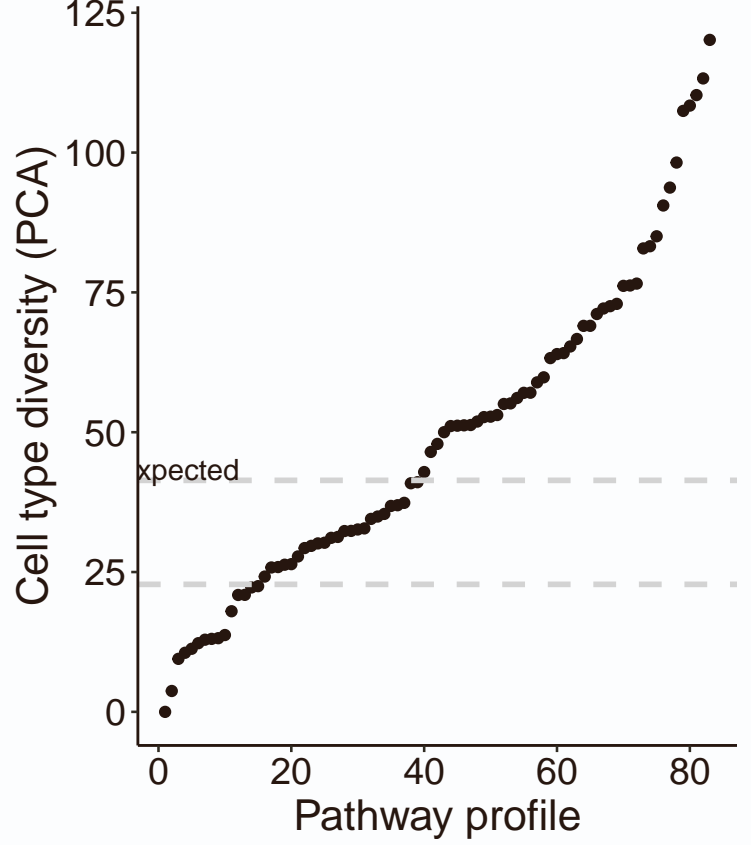
Global UMAP: Pathway ON



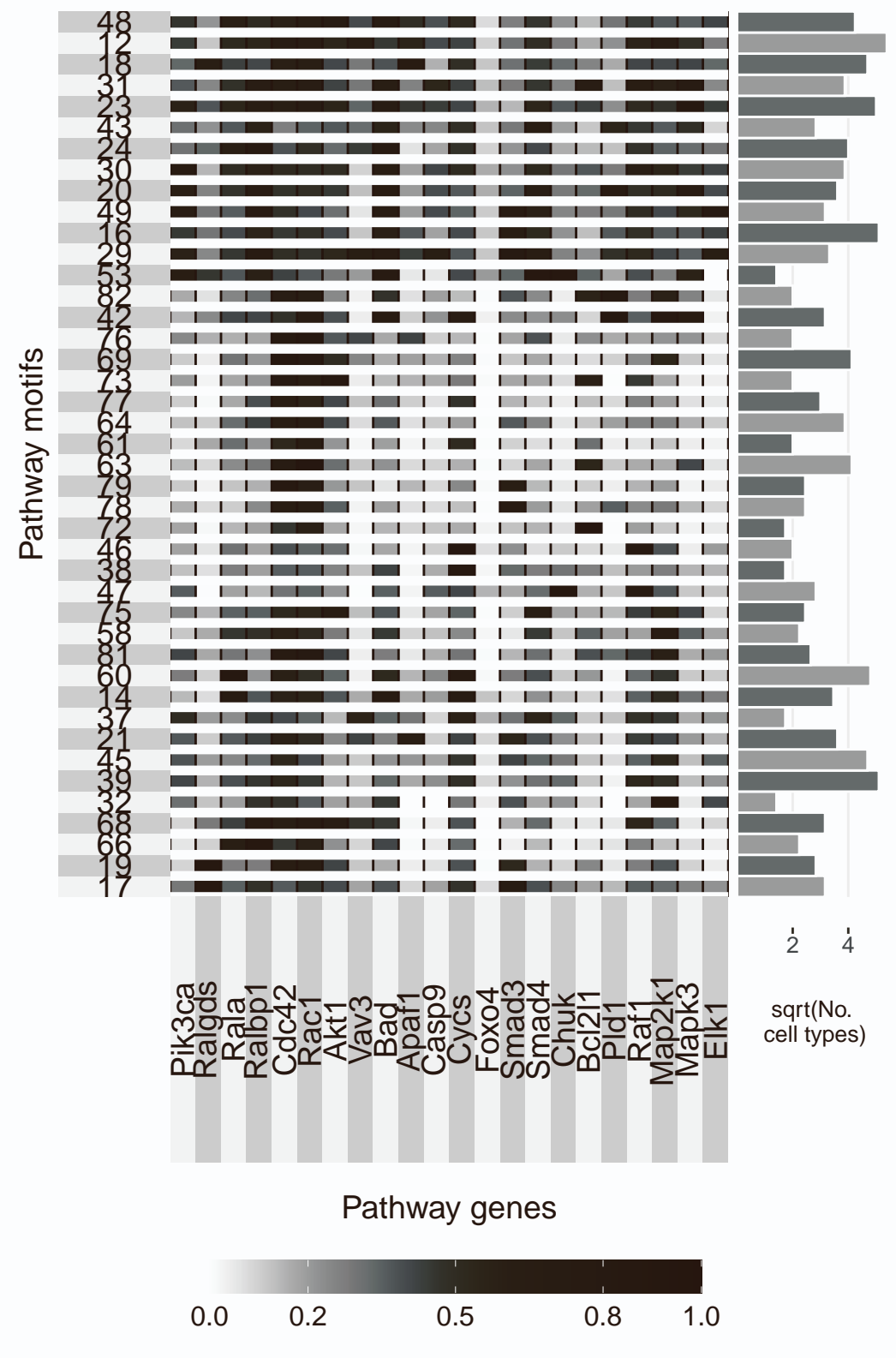
ECDF of Profile Dispersions



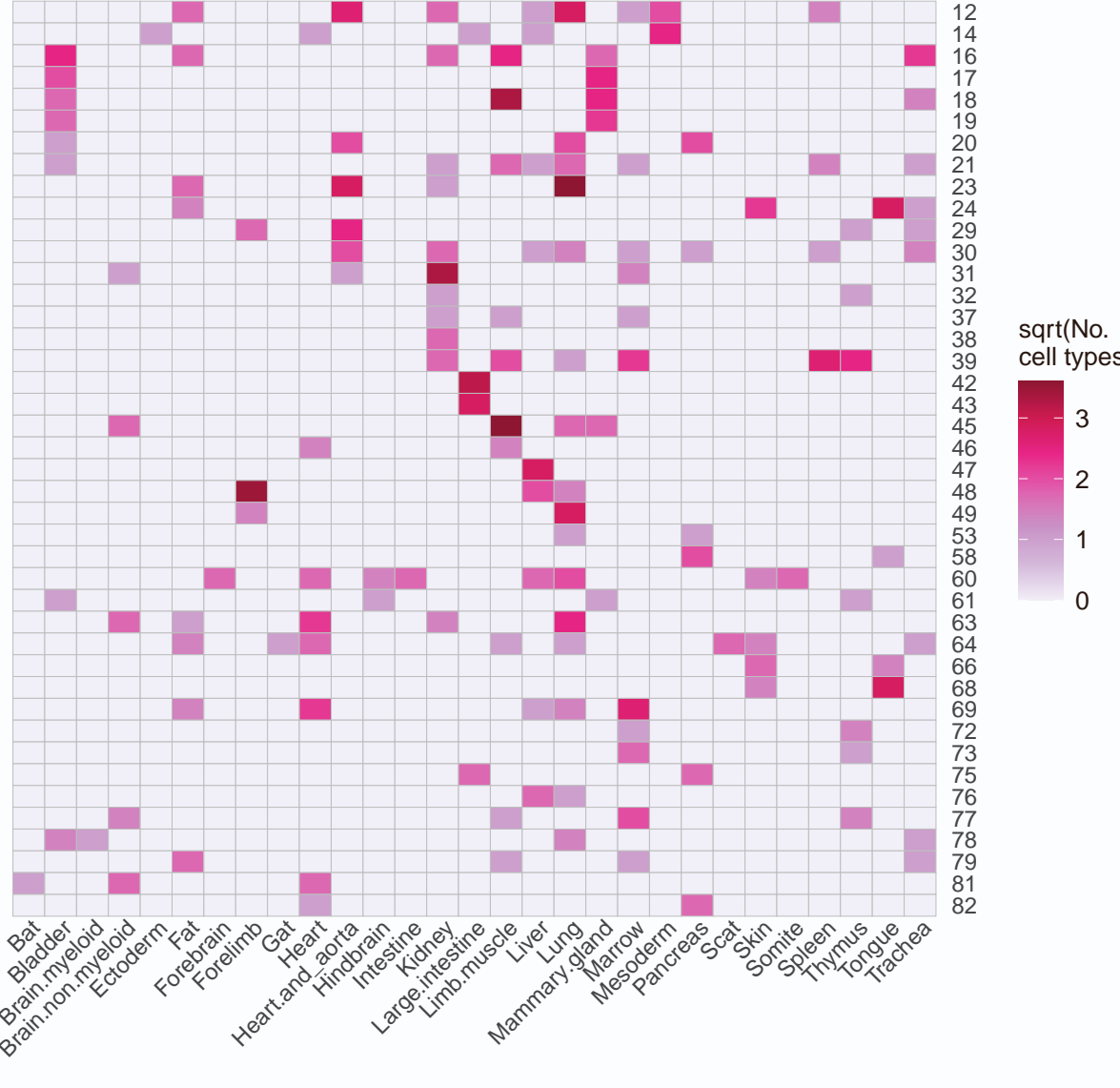
Profile Dispersion Plot



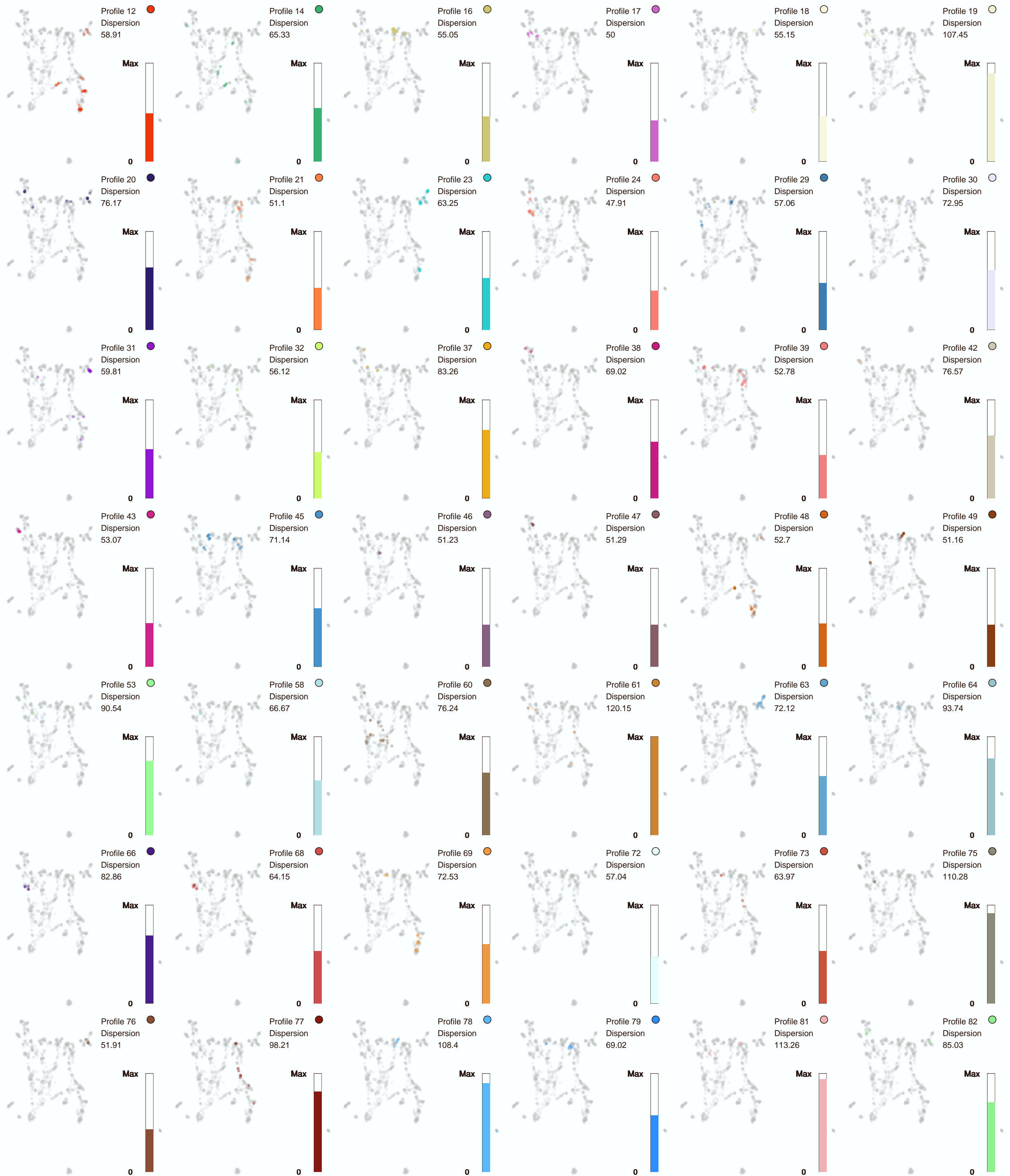
Motifs, Dispersion $\geq 90^{\text{th}}$ percentile



Motif Tissue Composition



Ras Signaling Pathway (k_opt = 83)



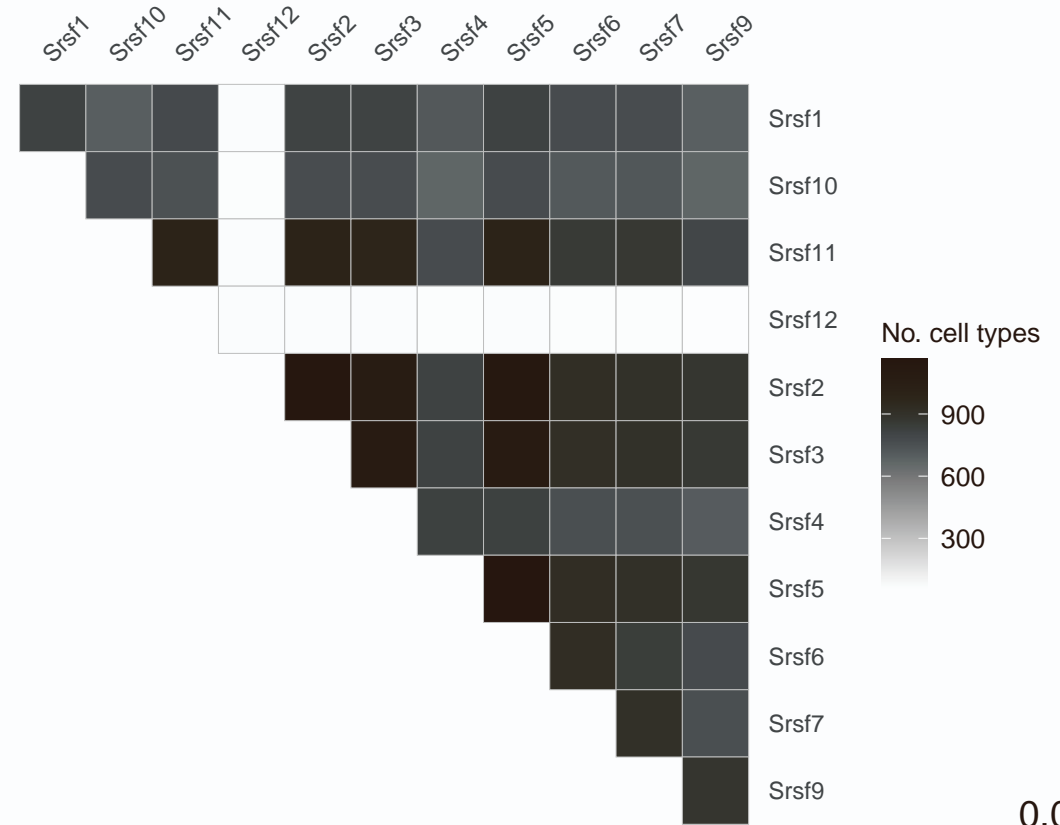
RNA-splicing by SR protein family (k_opt = 19)

Srsf1
Srsf10
Srsf11
Srsf12

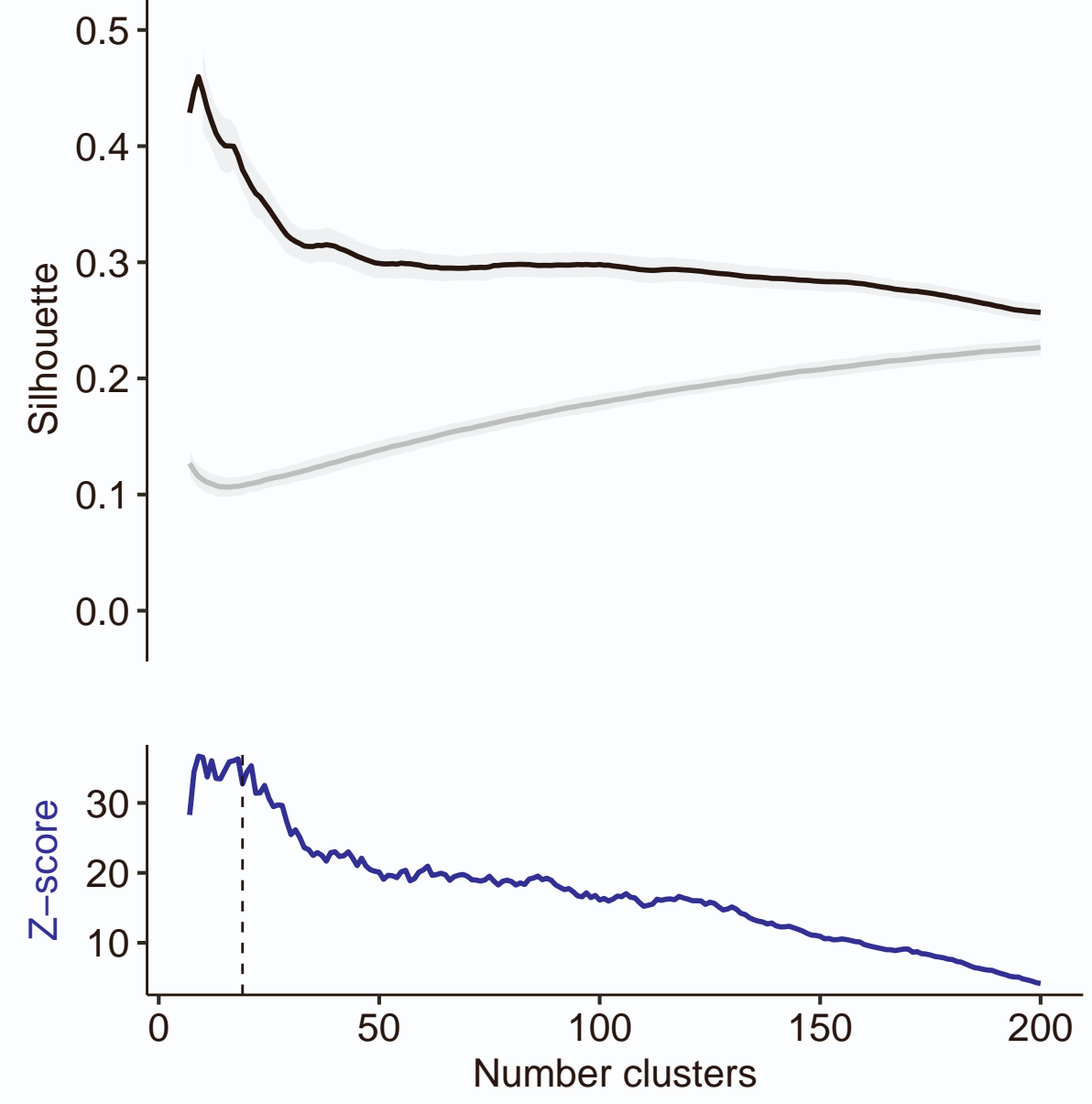
Pathway genes
Srsf2
Srsf3
Srsf4

Srsf5
Srsf6
Srsf7
Srsf9

Gene Co-Expression Above Threshold



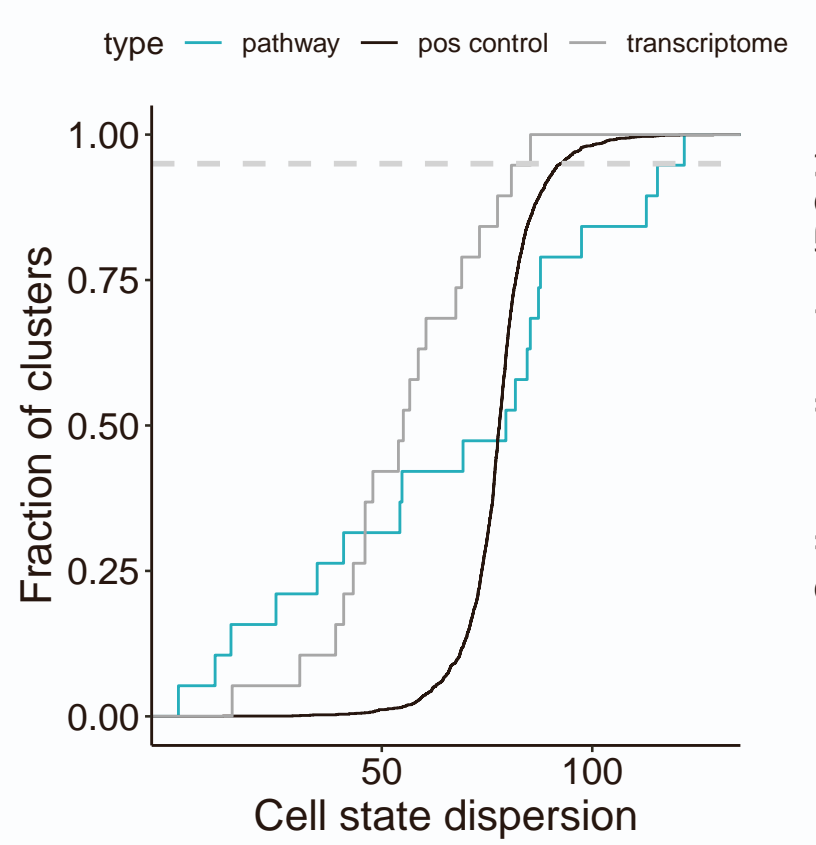
Silhouette and Z-score



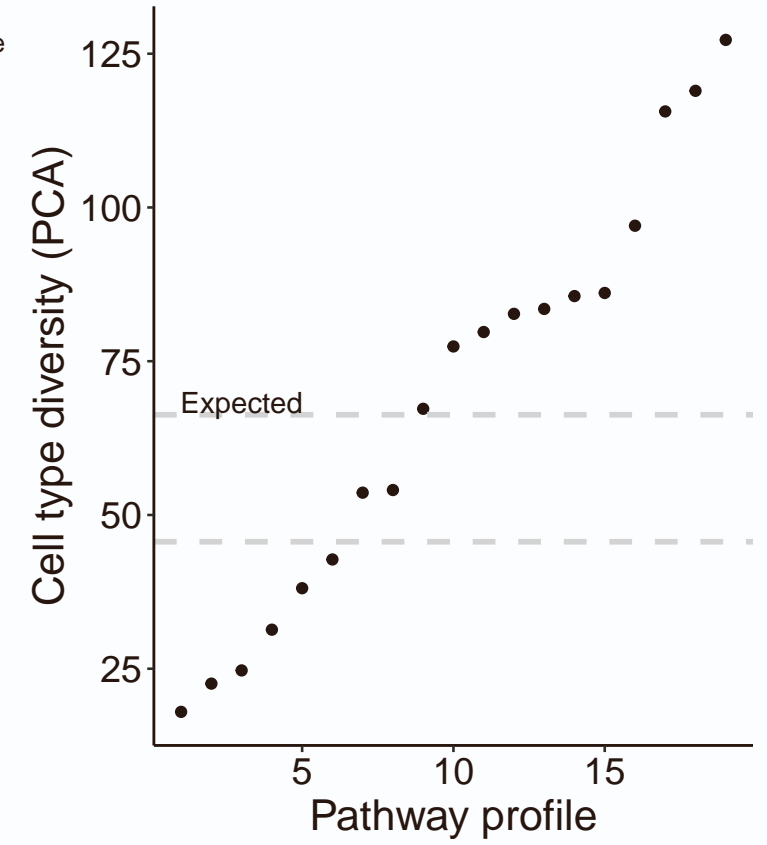
Global UMAP: Pathway ON



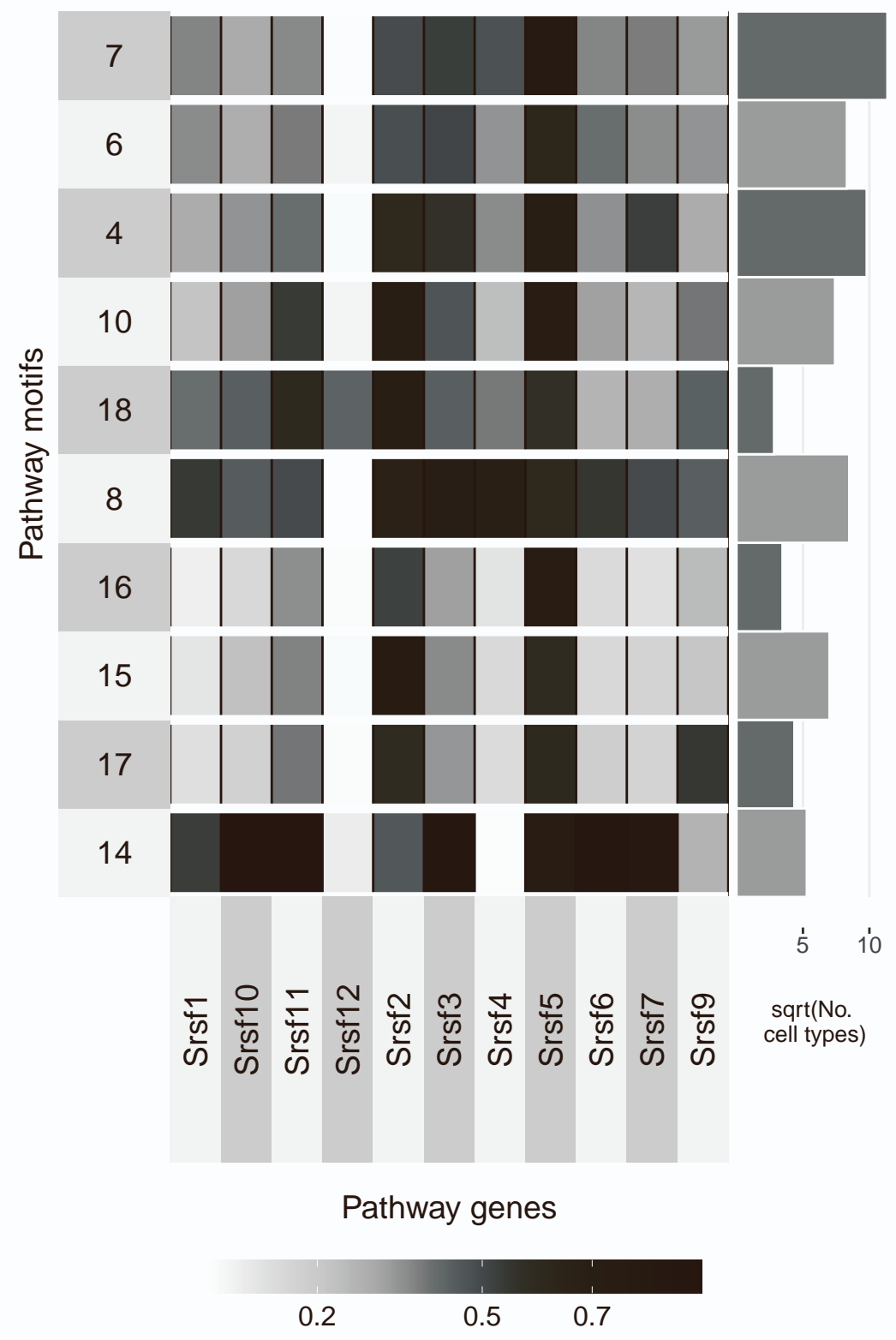
ECDF of Profile Dispersions



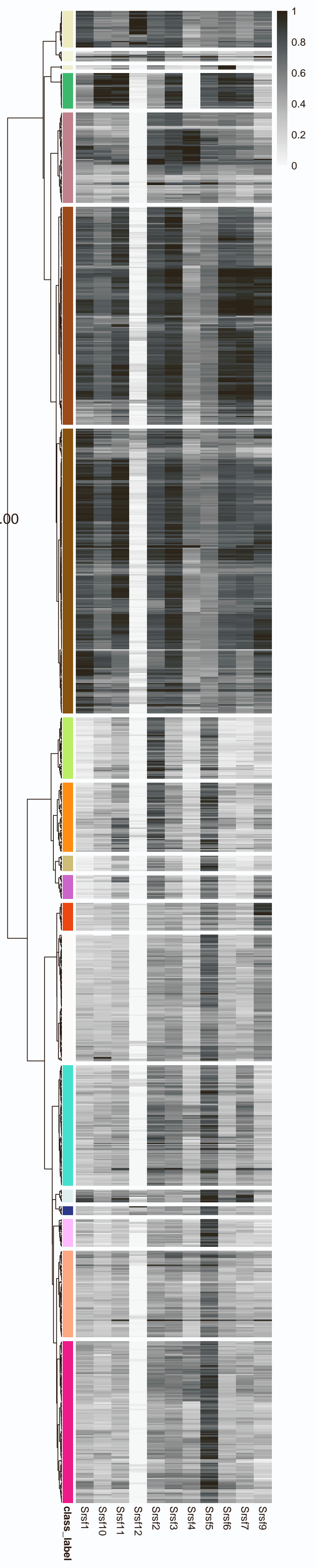
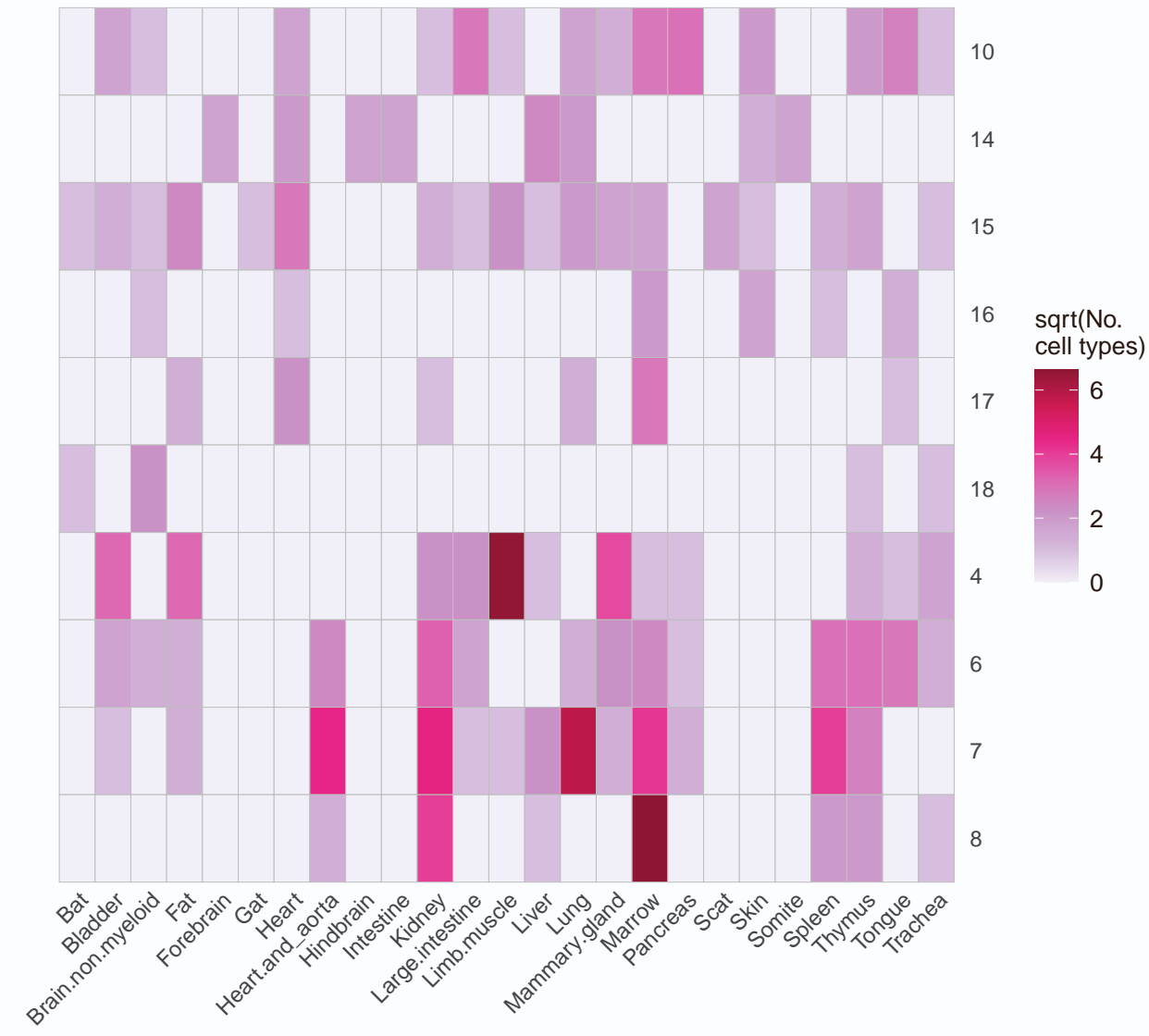
Profile Dispersion Plot



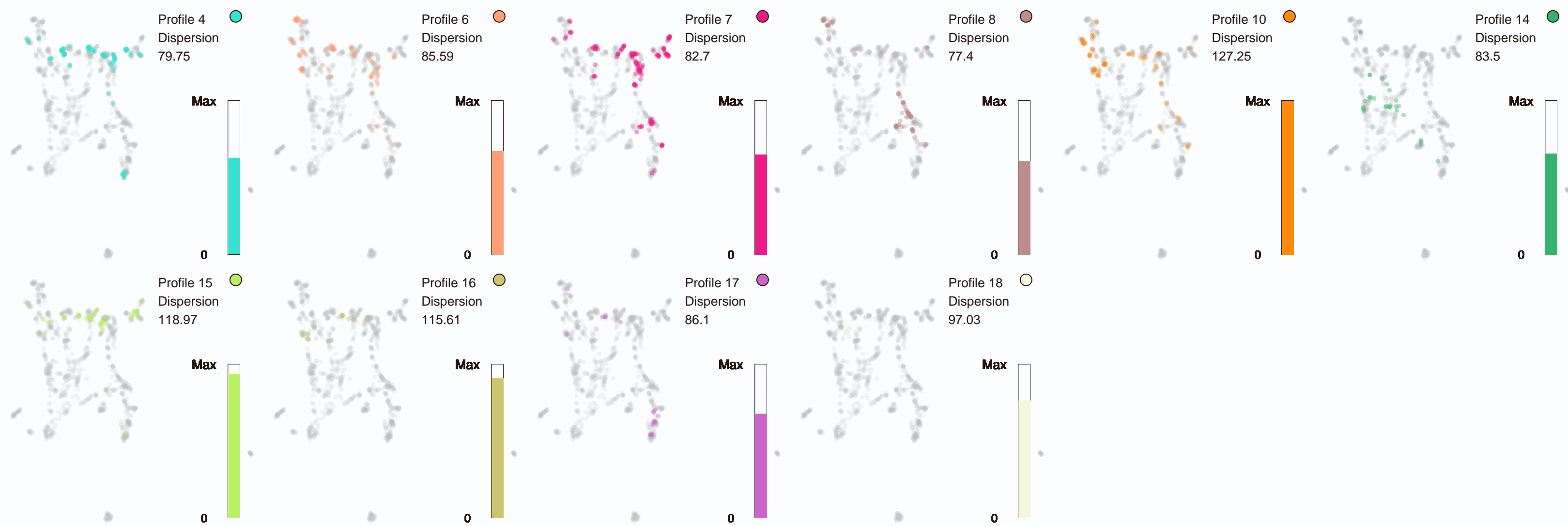
Motifs, Dispersion $\geq 90^{\text{th}}$ percentile



Motif Tissue Composition



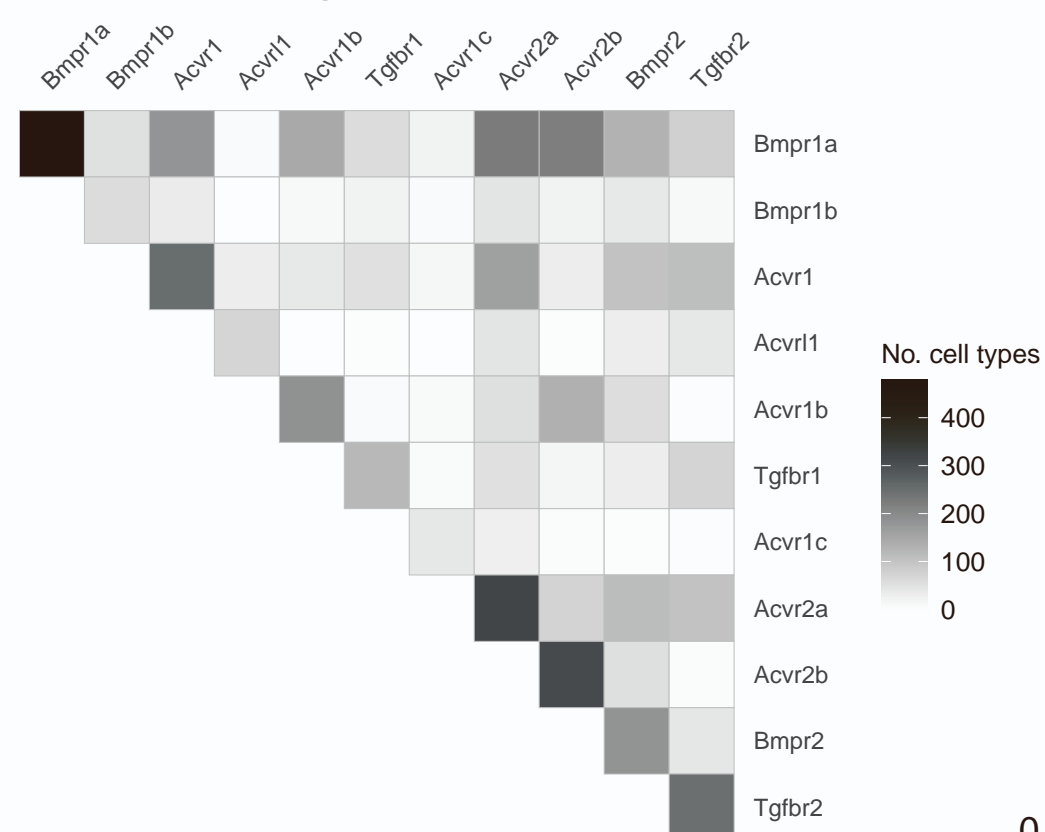
RNA-splicing by SR protein family (k_opt = 19)



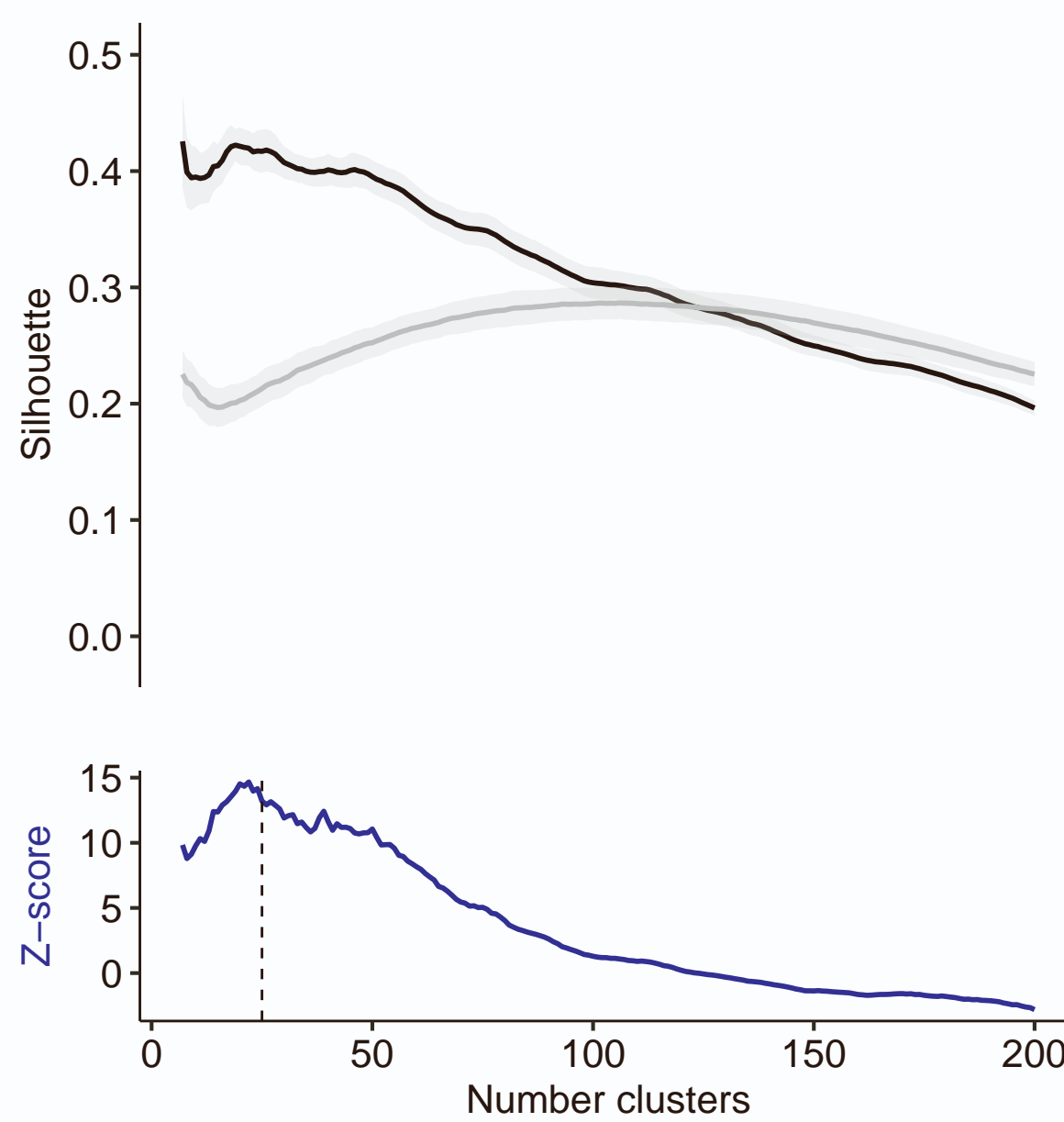
Tgf-beta family receptors (k_opt = 25)

Bmpr1a
 Bmpr1b
 Acvr1
 Acvr1c
 Acvr1b
 Tgfb1
 Acvr1c
 Acvr2a
 Acvr2b
 Bmpr2
 Tgfb2
 Acvr2a
 Acvr2b
 Bmpr2
 Tgfb2

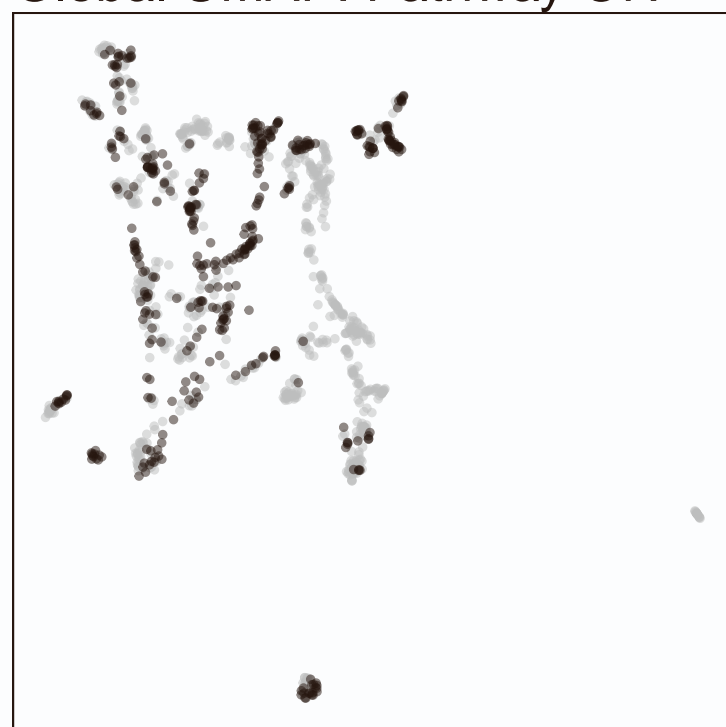
Gene Co-Expression Above Threshold



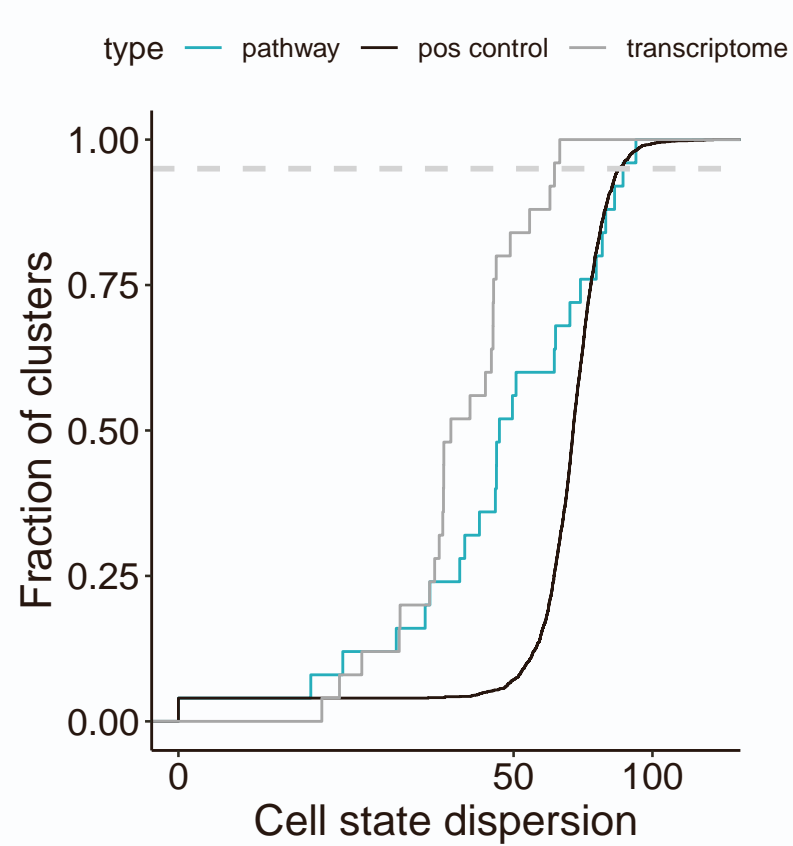
Silhouette and Z-score



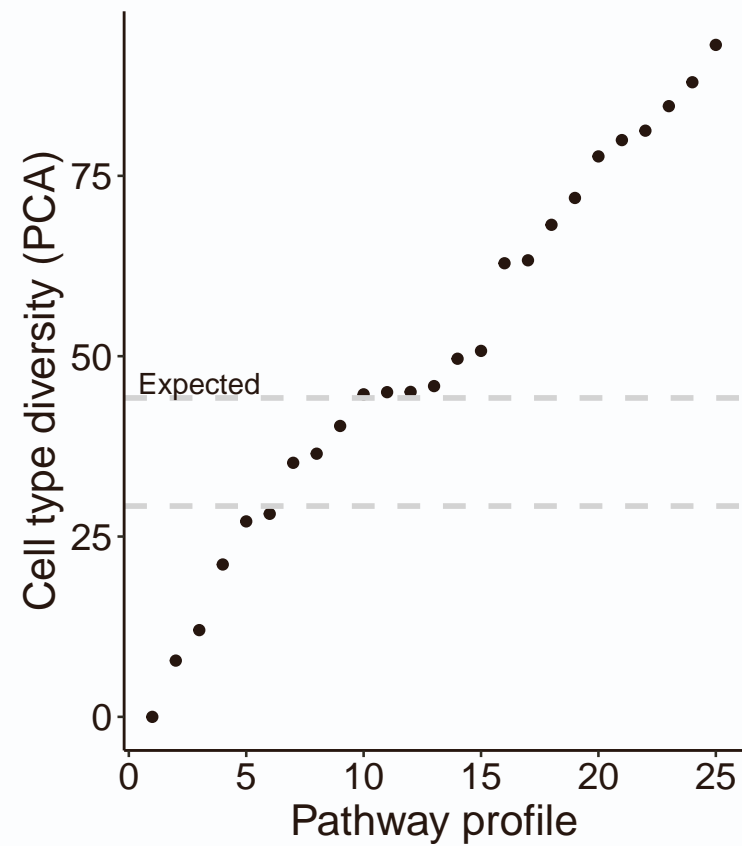
Global UMAP: Pathway ON



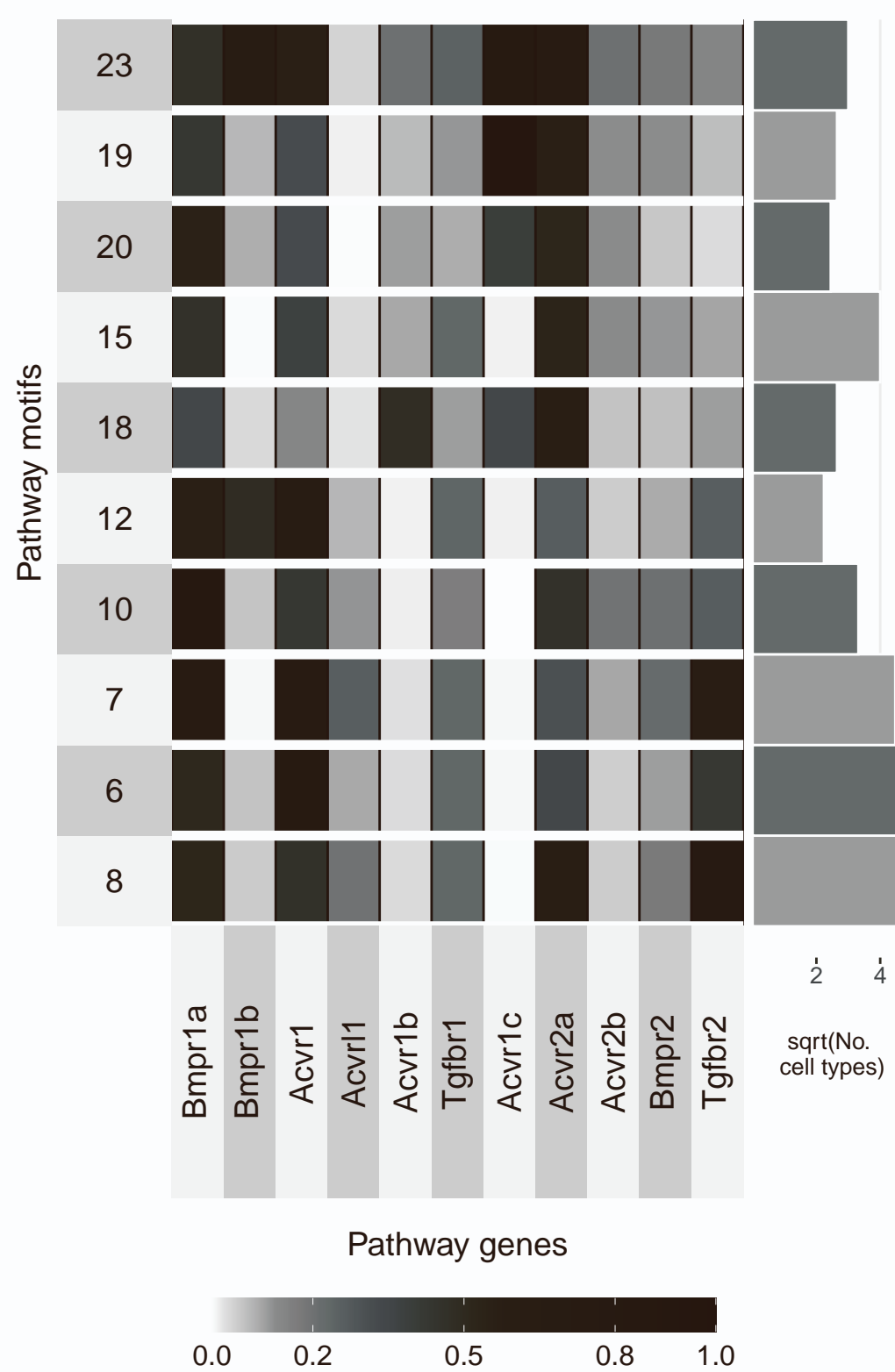
ECDF of Profile Dispersions



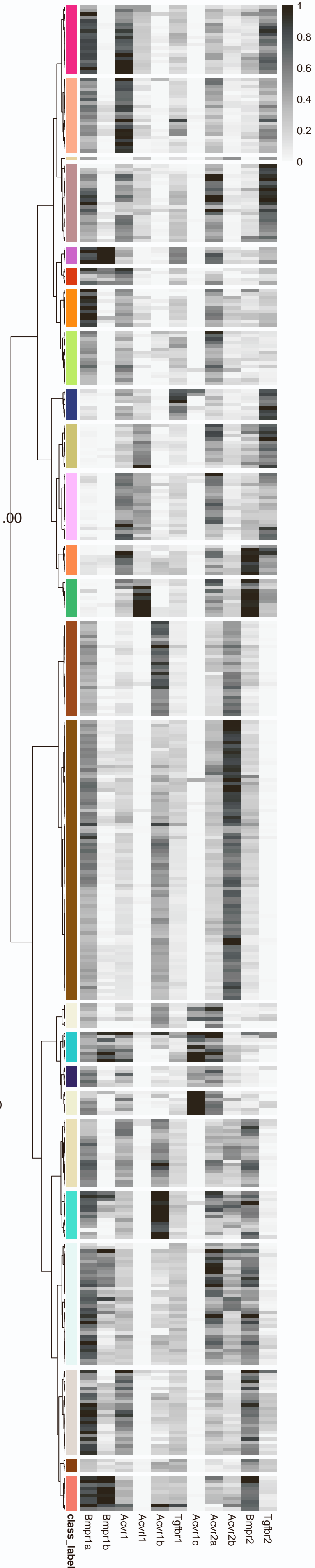
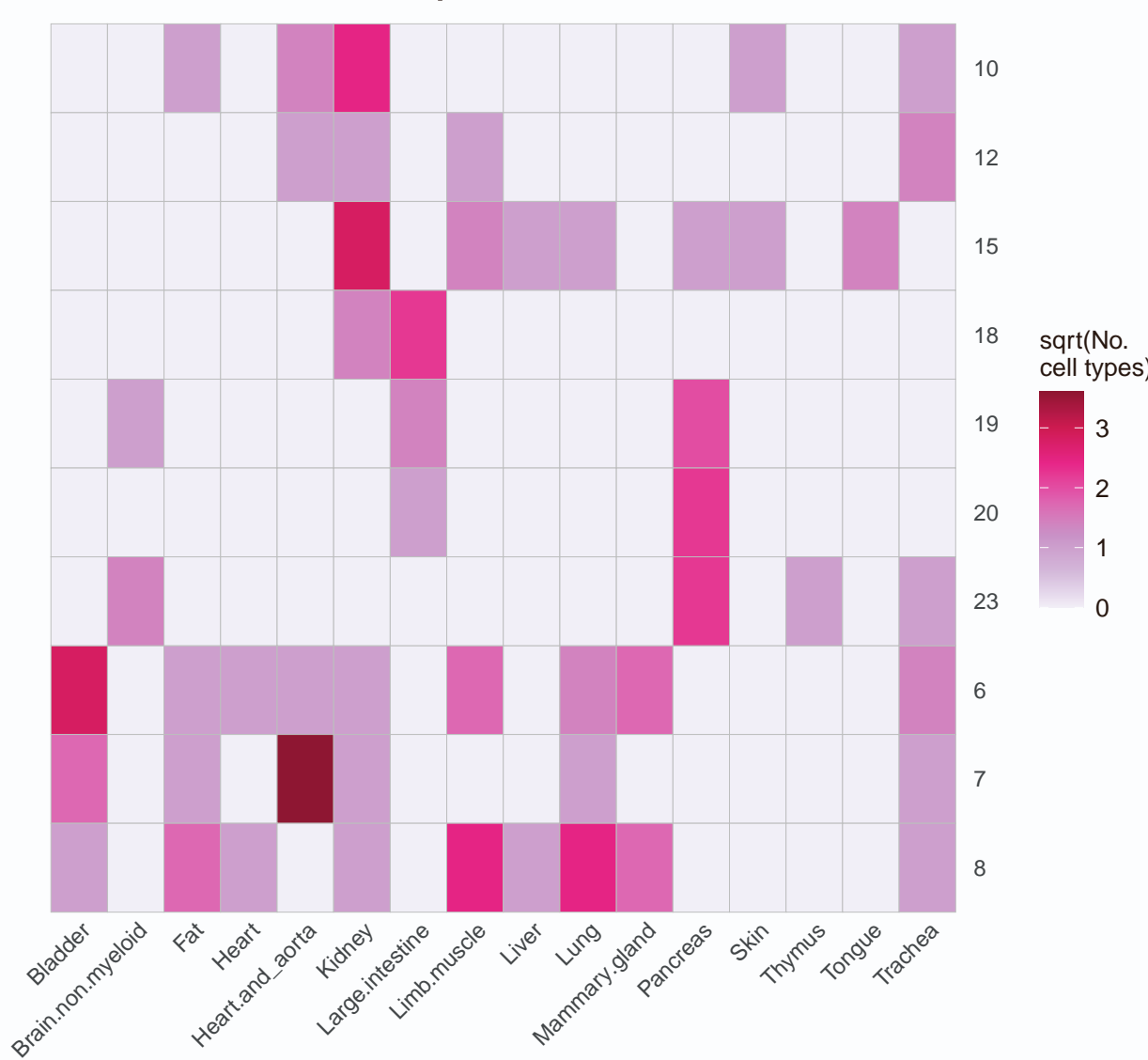
Profile Dispersion Plot



Motifs, Dispersion $\geq 90^{\text{th}}$ percentile



Motif Tissue Composition



Tgf-beta family receptors (k_opt = 25)

