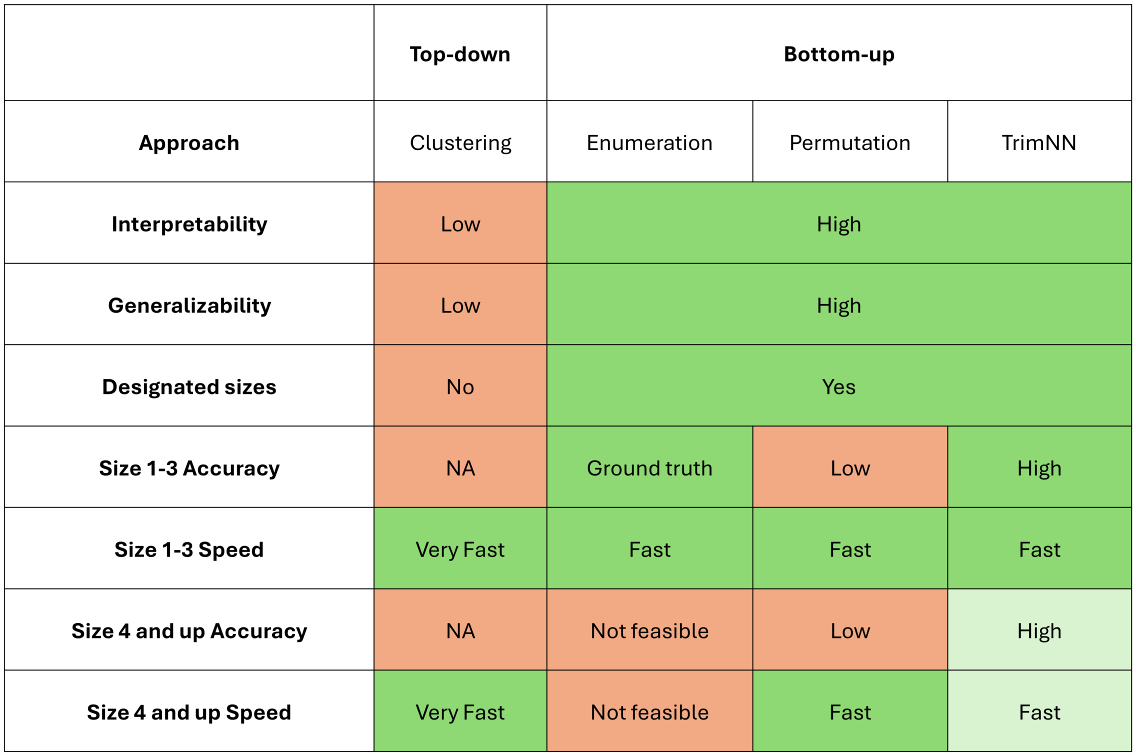
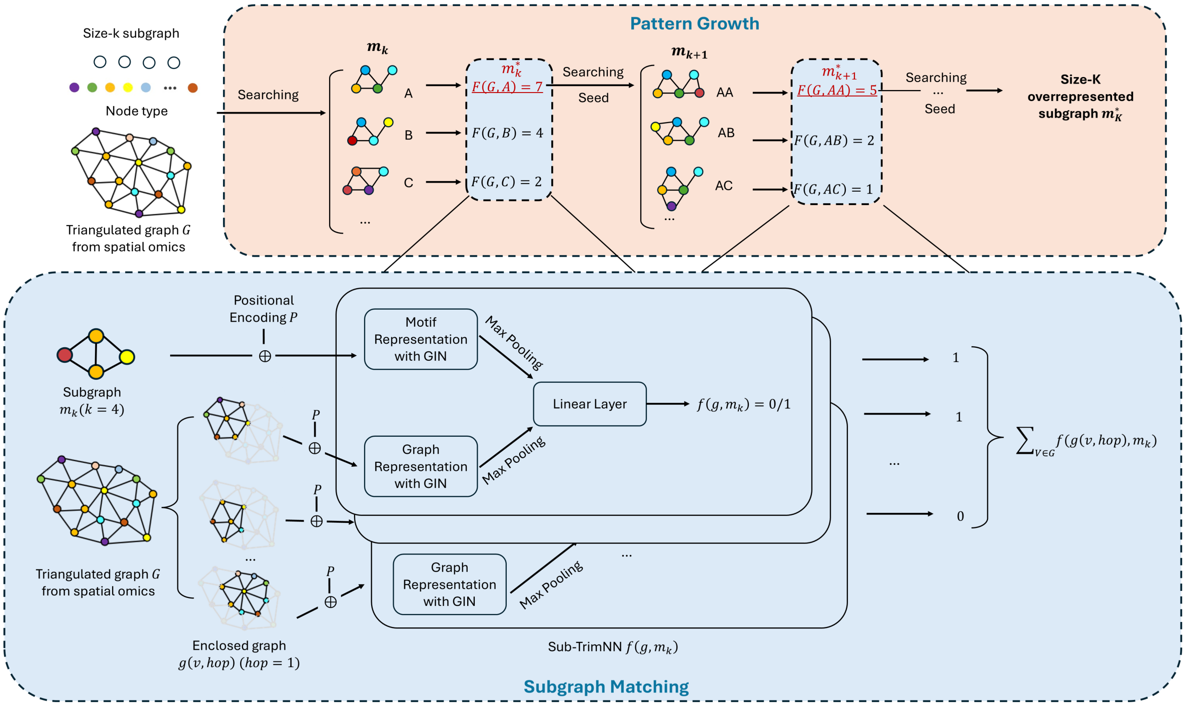
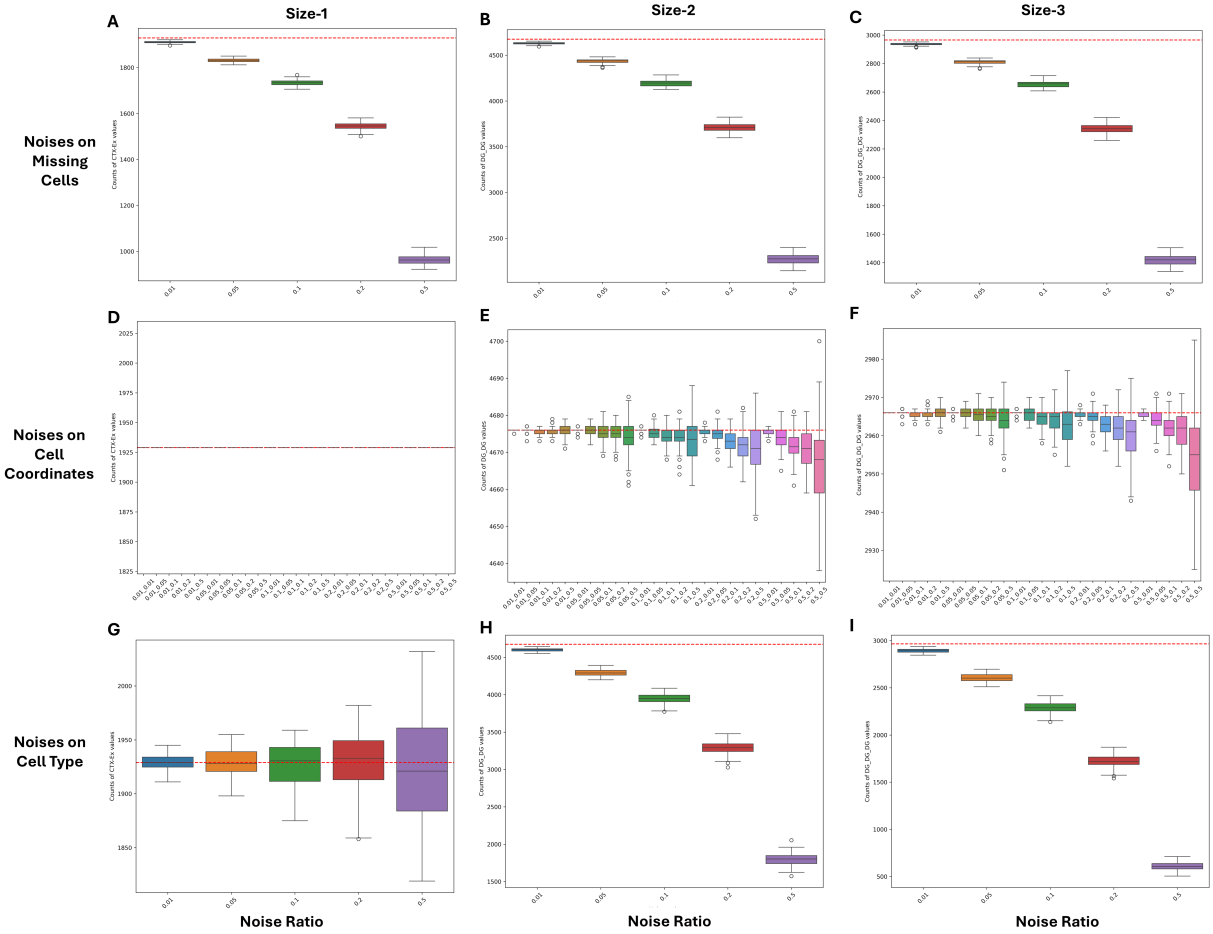


	Top-down	**Bottom-up**		
Approach	Clustering	Enumeration	Permutation	TrimNN
Interpretability	Low	High		
Generalizability	Low	High		
Designated sizes	No	Yes		
Size 1-3 Accuracy	NA	Ground truth	Low	High
Size 1-3 Speed	Very Fast	Fast	Fast	Fast
Size 4 and up Accuracy	NA	Not feasible	Low	High
Size 4 and up Speed	Very Fast	Not feasible	Fast	Fast

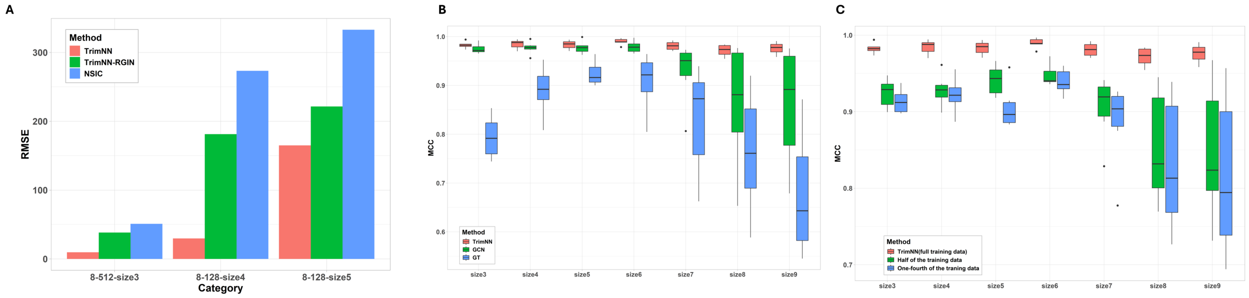
Supplementary Figure 1. Comparison of performance between top-down and bottom-up approaches**.**



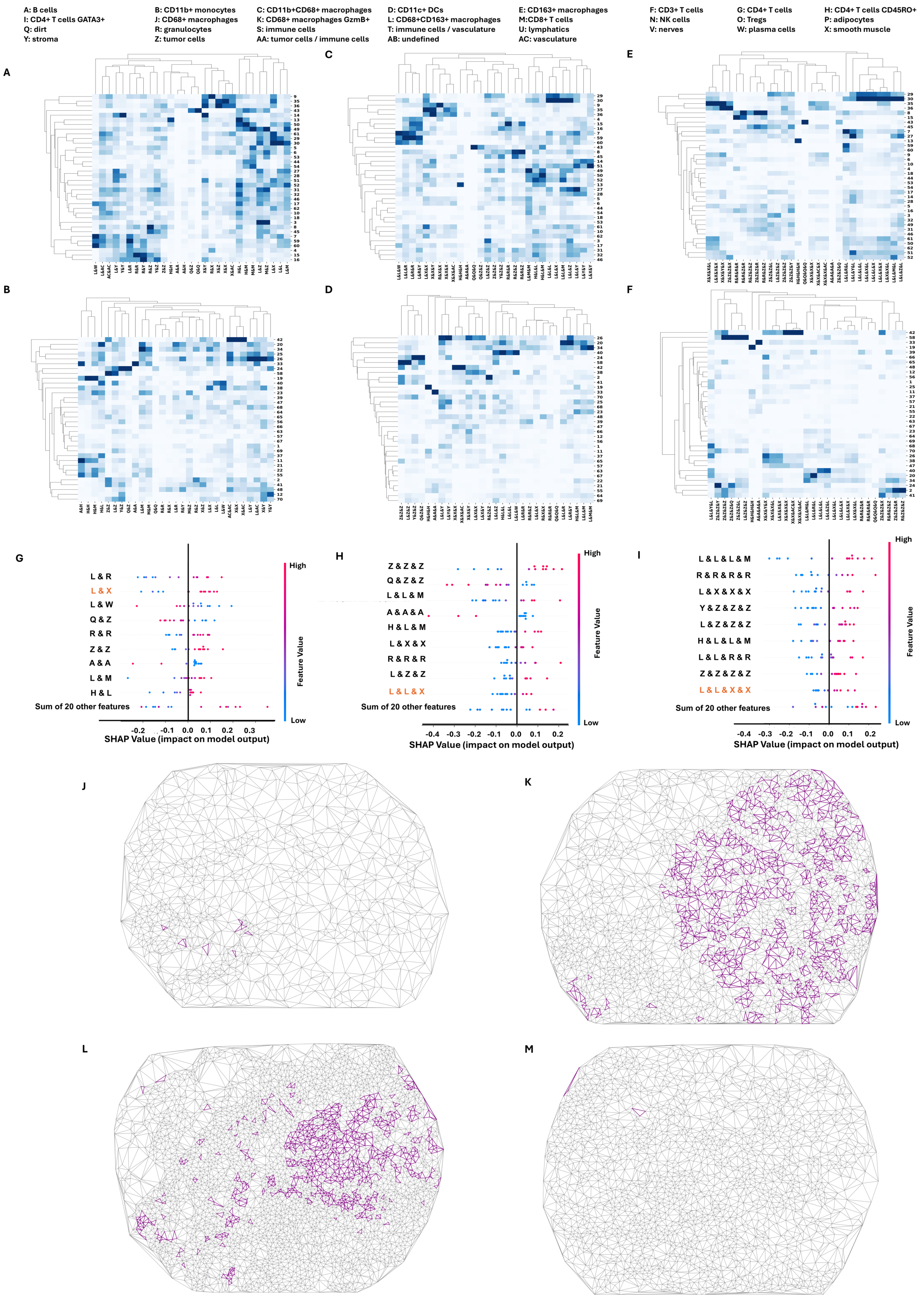
Supplementary Figure 2. The schema of TrimNN targets challenges of Pattern Growth (top) and Subgraph Matching (bottom). A size-4 subgraph m_k is taken as an example. In the challenge of Pattern Growth, all possible size-4 subgraphs are enumerated to check their occurrence in a triangulated graph G with trained TrimNN function $F(G, m_k)$ from Subgraph Matching. TrimNN summarizes many sub-TrimNN functions $f(g, m_k)$. For each sub-TrimNN, binary prediction on presence (1) or absence (0) is made by m_k on all possible enclosed subgraph g in G. The model uses GIN to learn the representations plus a positional encoding of the query motif and the target graph, then learns the relations between them.



Supplementary Figure 3. A-C. Simulations of cell missing effects on CC motifs at cell proportions of 0.01, 0.05, 0.1, 0.2, and 0.5, represented as absolute occurrence count of the most frequently occurring CC motif before and after simulated noises within **A.** size-1, **B**. size-2, and **C.** size-3. The red line indicates the absolute occurrence count of the most frequently occurring CC motif in the original data. **D-F**. Simulations of cell coordinate shifting effects on CC motifs at cell proportions of 0.01, 0.05, 0.1, 0.2, and 0.5 with different levels of noises of 0.01, 0.05, 0.1, 0.2, and 0.5. They are represented as the absolute occurrence count of the most frequently occurring CC motif before and after simulated noises within **D.** size-1, **E**. size-2, and **F.** size-3. **G-I**. Simulations of cell type misclassification effects on CC motifs at cell proportions of 0.01, 0.05, 0.1, 0.2, and 0.5, represented as absolute occurrence count of the most frequently occurring CC motif before and after simulated noises within **G.** size-1, **H**. size-2, and **I.** size-3.

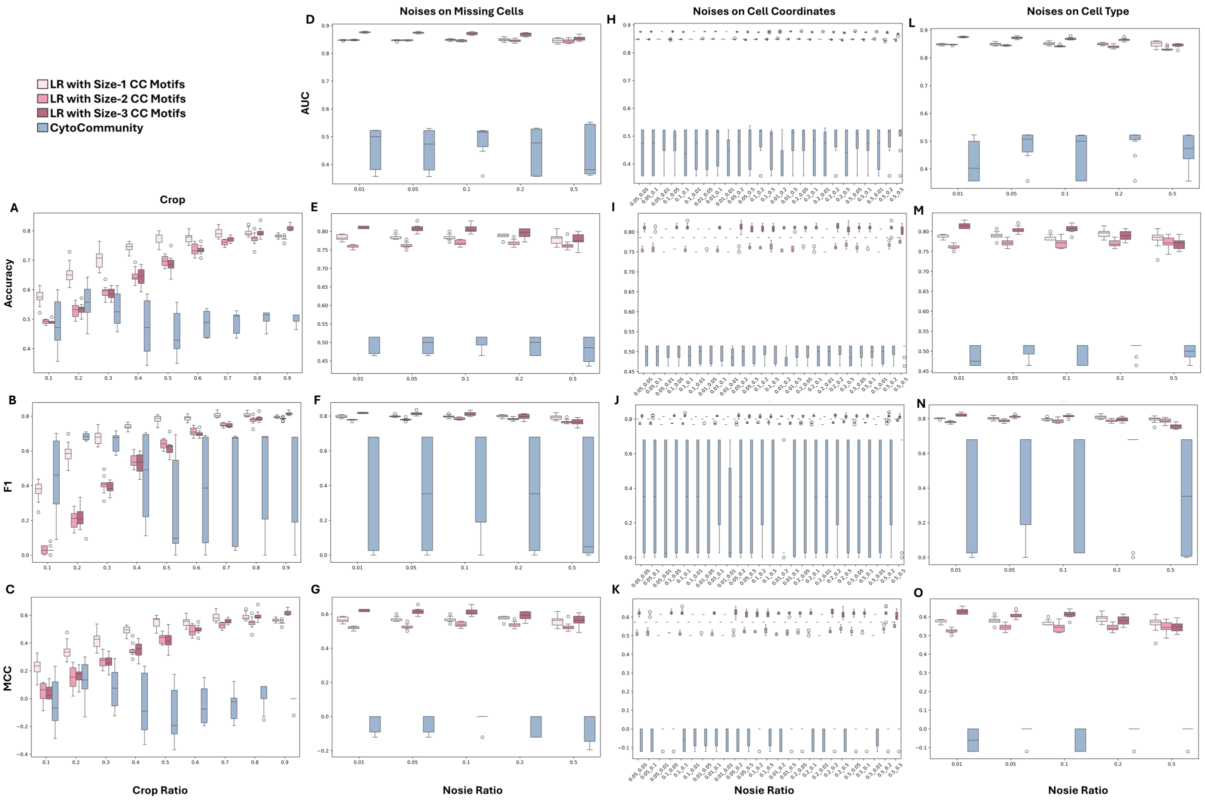


Supplementary Figure 4. A. Comparison of the performance of TrimNN, TrimNN-RGIN, and NSIC in the abundance of CC motifs with root mean square error (RMSE). **B.** Comparison of the model performances by replacing Graph Isomorphosim Network (GIN) in TrimNN with Graph Convolutional Networks (GCN) and Graph Transformer (GT). The X-axis represents different sizes of CC motifs, the Y-axis indicates Matthews Correlation Coefficient (MCC) values. **C.** Comparison of the performance of TrimNN trained with full, half, and one-quarter of the training data.

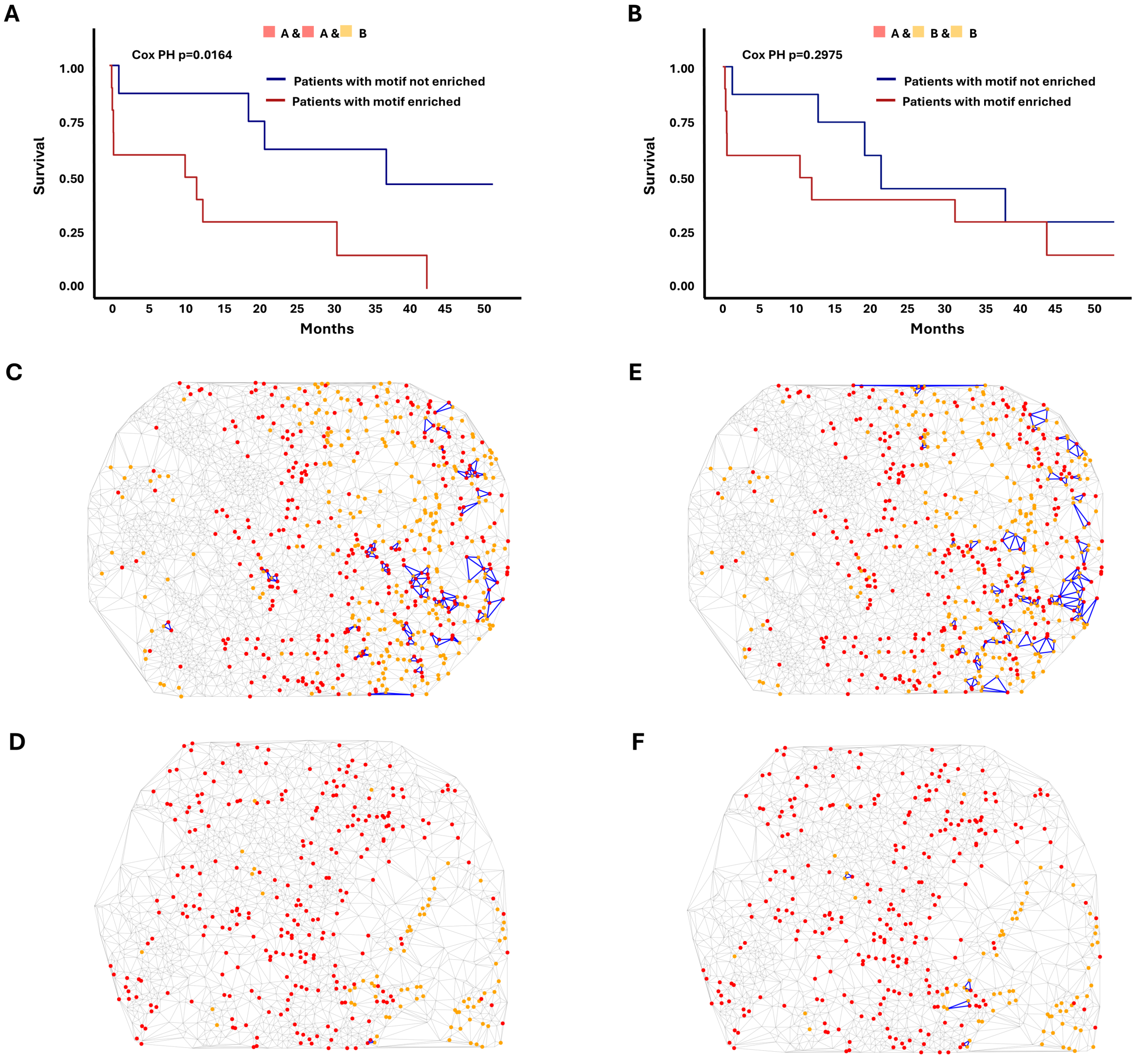


Supplementary Figure 5. The clustering map of the abundance of **A**. size-2, **C**. size-3, and **E**. size-4 CC motifs among DII patient samples and **B**. size-2, **D**. size-3, and **F**. size-

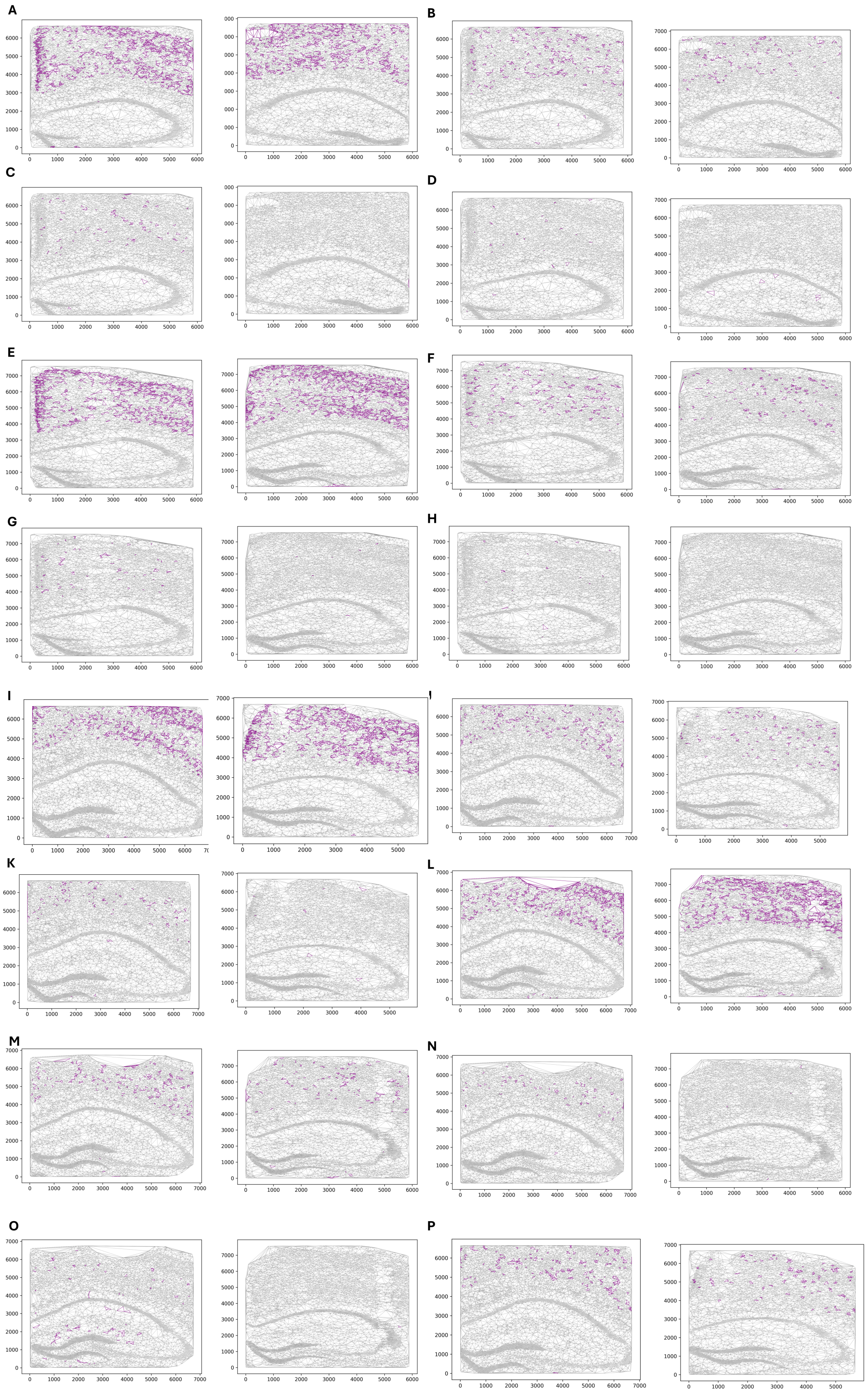
4 CC motifs among CLR patient samples. The SHAP value of Logistics Regression model using top 29 **G**. size-2, **H**. size-3, and **I**. size-4 CC motifs as features. The locations of the ‘tumor & immune & immune’ motif (marked as purple, others grey) are visualized in **J**. patient 35 (CLR) on spot 70 B and **K**. patient 9 (DII) on spot 18 B. The locations of the ‘B cells & B cells & B cells’ motif (marked as purple, others grey) are visualized in **L**. patient 28 (CLR) on spot 55 A and **M**. patient 14 (DII) on spot 28.



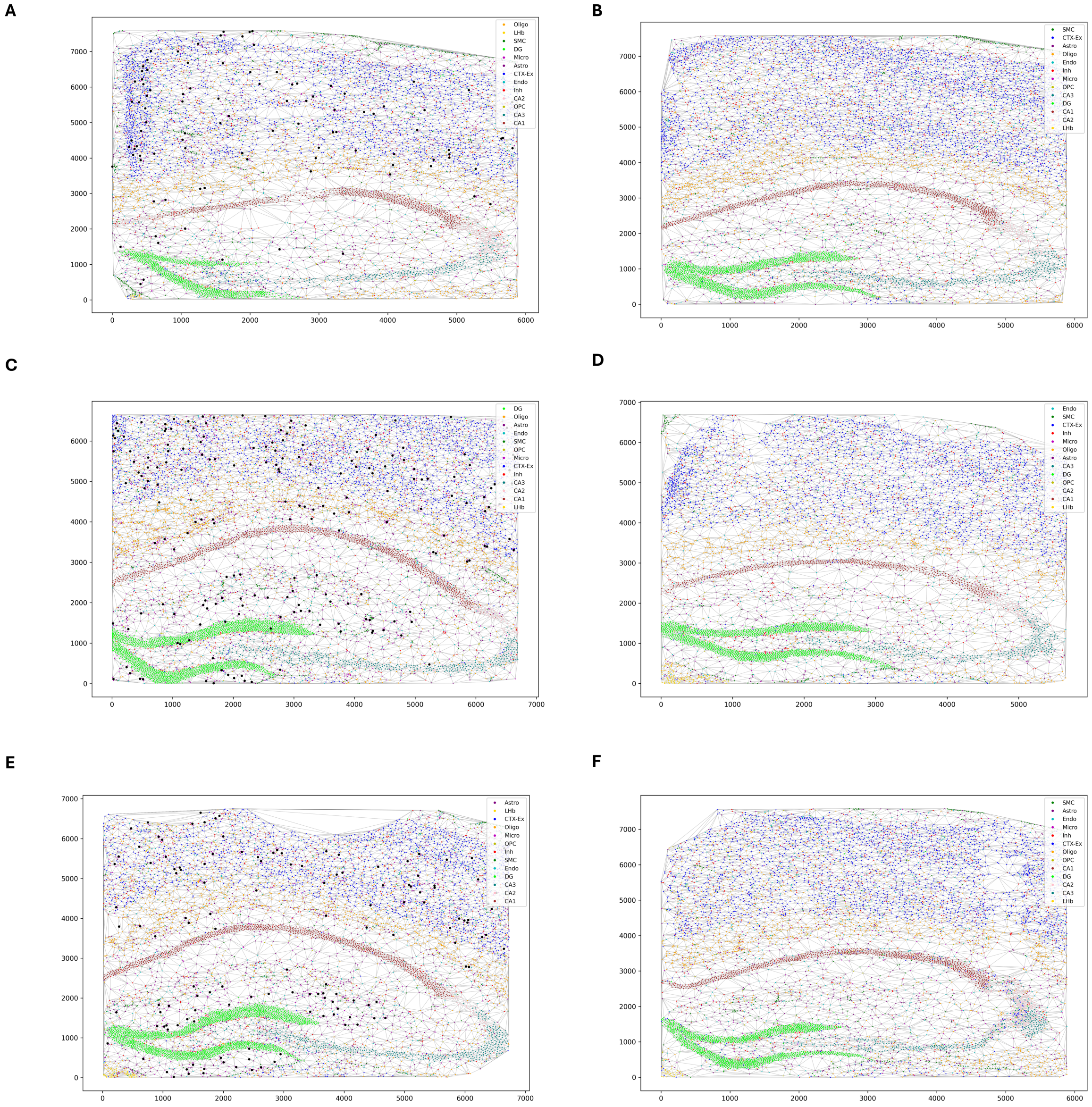
Supplementary Figure 6. Generalizability of the trained model testing on random cropping of the samples. X-axis is the ratio of width and height of the original ROI, and Y-axis is **A.** accuracy, **B.** F1 score, and **C**. MCC. AUC is Figure. 3C in the main text. Generalizability of the trained model testing on distorted samples with noises from cell missing. X-axis is the noise ratio, and Y-axis is **D.** AUC, **E.** accuracy, **F.** F1 score, and **G**. MCC. Generalizability of the trained model testing on distorted samples with noises from cell coordinates shifting. X-axis is the noise ratio, and Y-axis is **H.** AUC, **I.** accuracy, **J.** F1 score, and **K**. MCC. Generalizability of the trained model testing on distorted samples with noises from cell type misclassification. X-axis is the noise ratio, and Y-axis is **L.** AUC, **M.** accuracy, **N.** F1 score, and **O**. MCC.



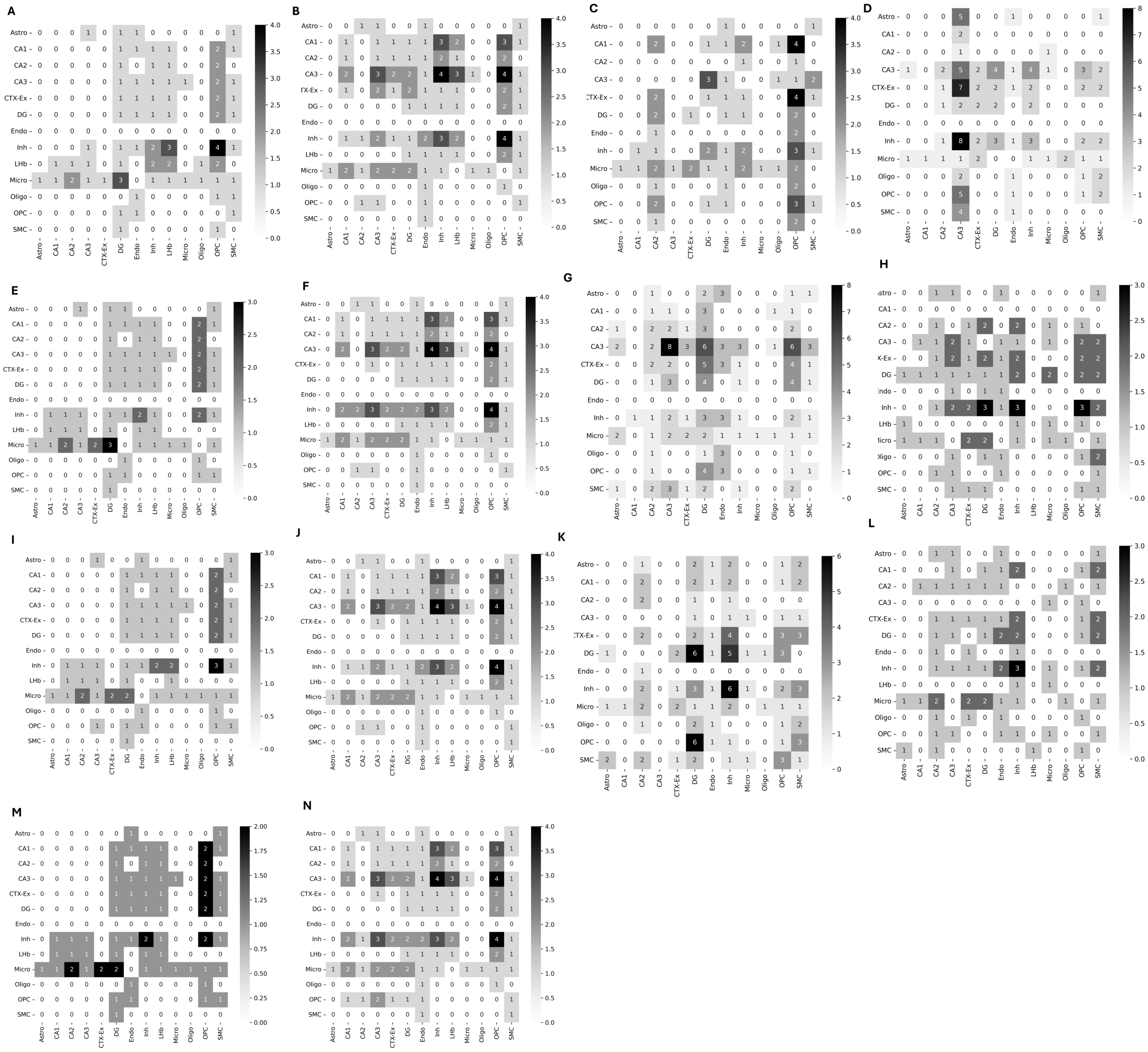
Supplementary Figure 7. Survival curves of DII patients with and without enriched CC motifs, including size-3 **A**. ‘A & A & B’ and **B**. ‘A & B & B’. Here cell type CD68+CD163+ macrophages are denoted as cell type ‘A’, and cell type smooth muscle are denoted as cell type ‘B’. The visualization of spatial locations of the size-3 motif ‘A & A & B’ in **C.** DII spot 5A and **D.** CLR spot 16A. The visualization of spatial locations of the size-3 motif ‘A & B & B’ in **E.** DII spot 5A and **F.** CLR spot 16A.


A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
0
1000
2000
3000
4000
5000
6000
7000

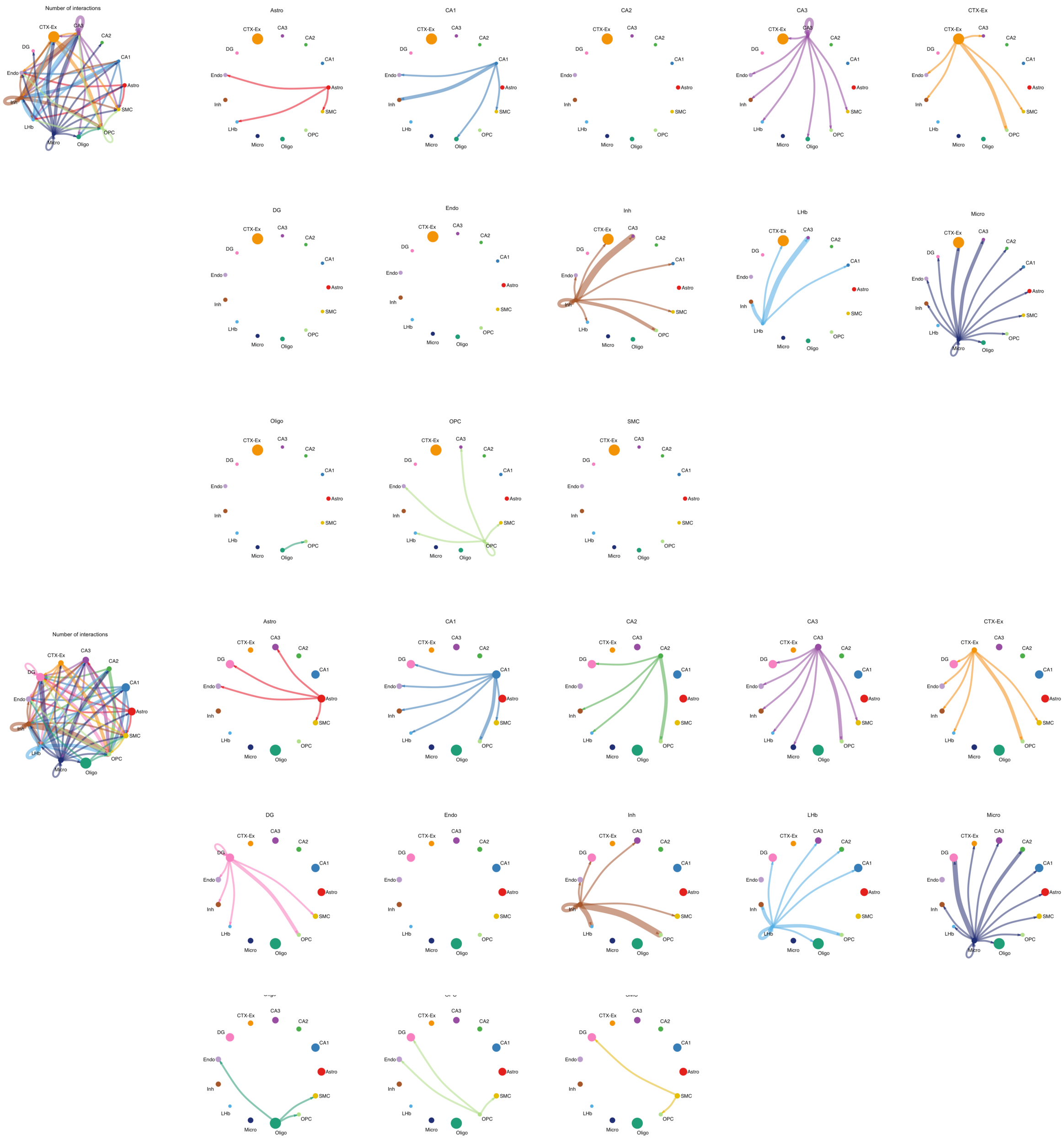
Supplementary Figure 8. Representative CC motifs in triangulation graph in disease (left column) and control (right column) samples, including localization of motifs **A.** 'CCC', **B.** 'CCM'**, C.** 'CMM', and **D.** 'MMM' in eight-month-old sample replicate 1, motifs **E.** 'CCC', **F.** 'CCM'**, G.** 'CMM', and **H.** 'MMM' in eight-month-old sample replicate 2, motifs **I.** 'CCC'**, J.** 'CCM', and **K.** 'CMM' in thirteen-month-old sample replicate 1, motifs **L.** 'CCC', **M.** 'CCM'**, N.** 'CMM', and **O.** 'MMM' in thirteen-month-old sample replicate 2, Size-4 CC motif **P**. 'CCCM' in thirteen-month-old sample replicate 1.



Supplementary Figure 9. The triangulated graphs as cellular communities of all other samples other than thirteen-month-old samples in the main figure, including eight-month-old replicate 1 in **A.** disease and **B.** control samples; eight-month-old replicate 2 **C.** disease and **D.** control samples; thirteen-month-old replicate 2 **E.** disease and **F.** control samples.



Supplementary Figure 10. CellChat results as a matrix between different cell types in CC motif regions (motif and its 3-hop) and non-motif regions. Number of links in non-motif ‘CCC’ regions in **A.** eight-month-old samples and **B.** thirteen-month-old samples. Number of links in motif ‘CCM’ regions in **C.** eight-month-old samples and **D.** thirteen-month-old samples. Number of links in non-motif ‘CCM’ regions in **E.** eight-month-old samples and **F.** thirteen-month-old samples. The number of links in motif ‘CMM’ regions in **G.** eight-month-old samples and **H.** thirteen-month-old samples. The number of links in non-motif ‘CMM’ regions in **I.** eight-month-old samples and **J.** thirteen-month-old samples. The number of links in motif ‘MMM’ regions in **K.** eight-month-old samples and **L.** thirteen-month-old samples. The number of links in non-motif ‘MMM’ regions in **M.** eight-month-old samples and **N.** thirteen-month-old samples. The number of links is defined in CellChat software by cellchat@net$count.

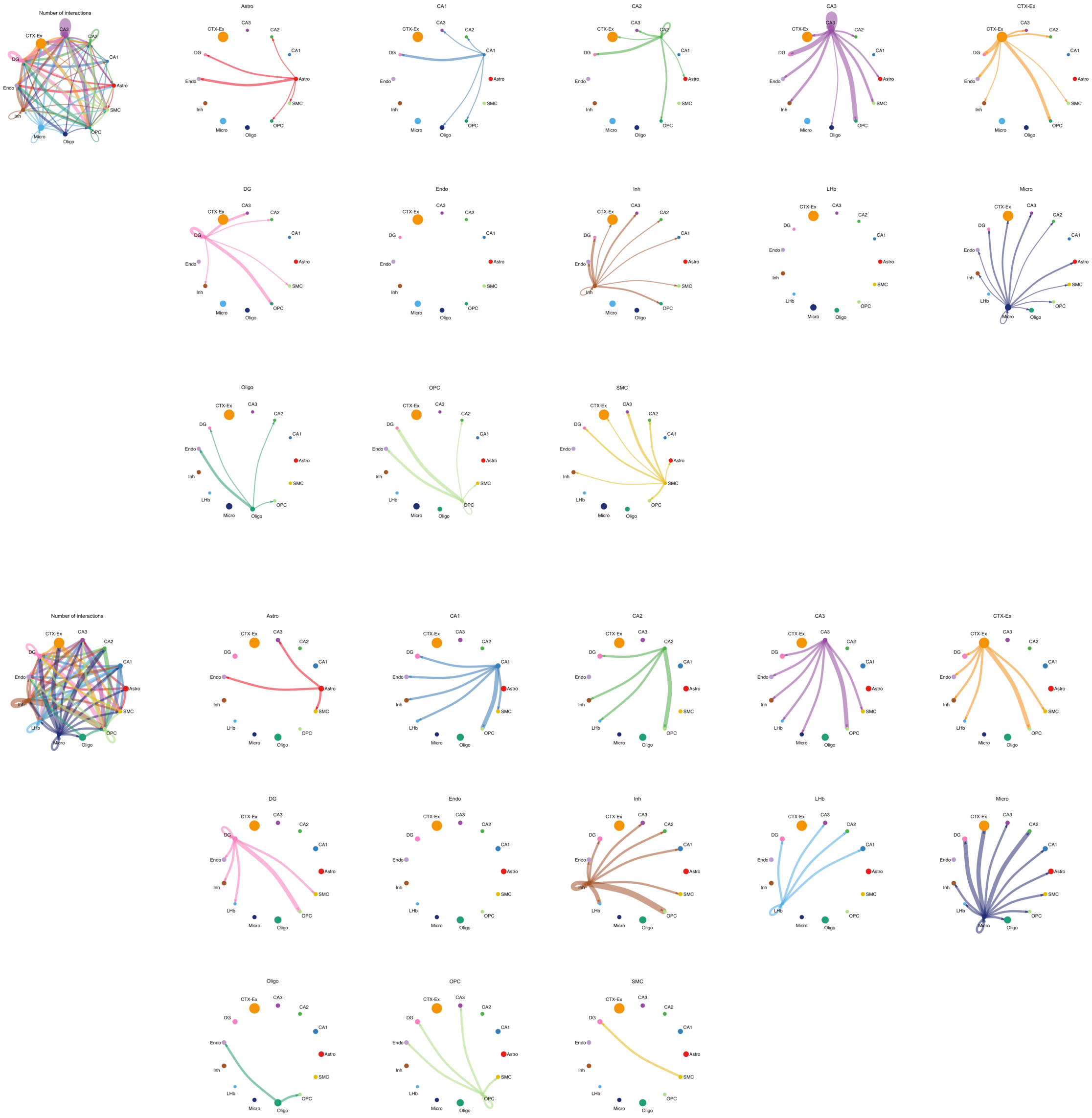


Supplementary Figure 11. Visualization of CellChat results on each cell type between in ‘CCC’ motif regions (motif and its 3-hop, top) and non-motif regions (bottom) in eight-month-old samples.

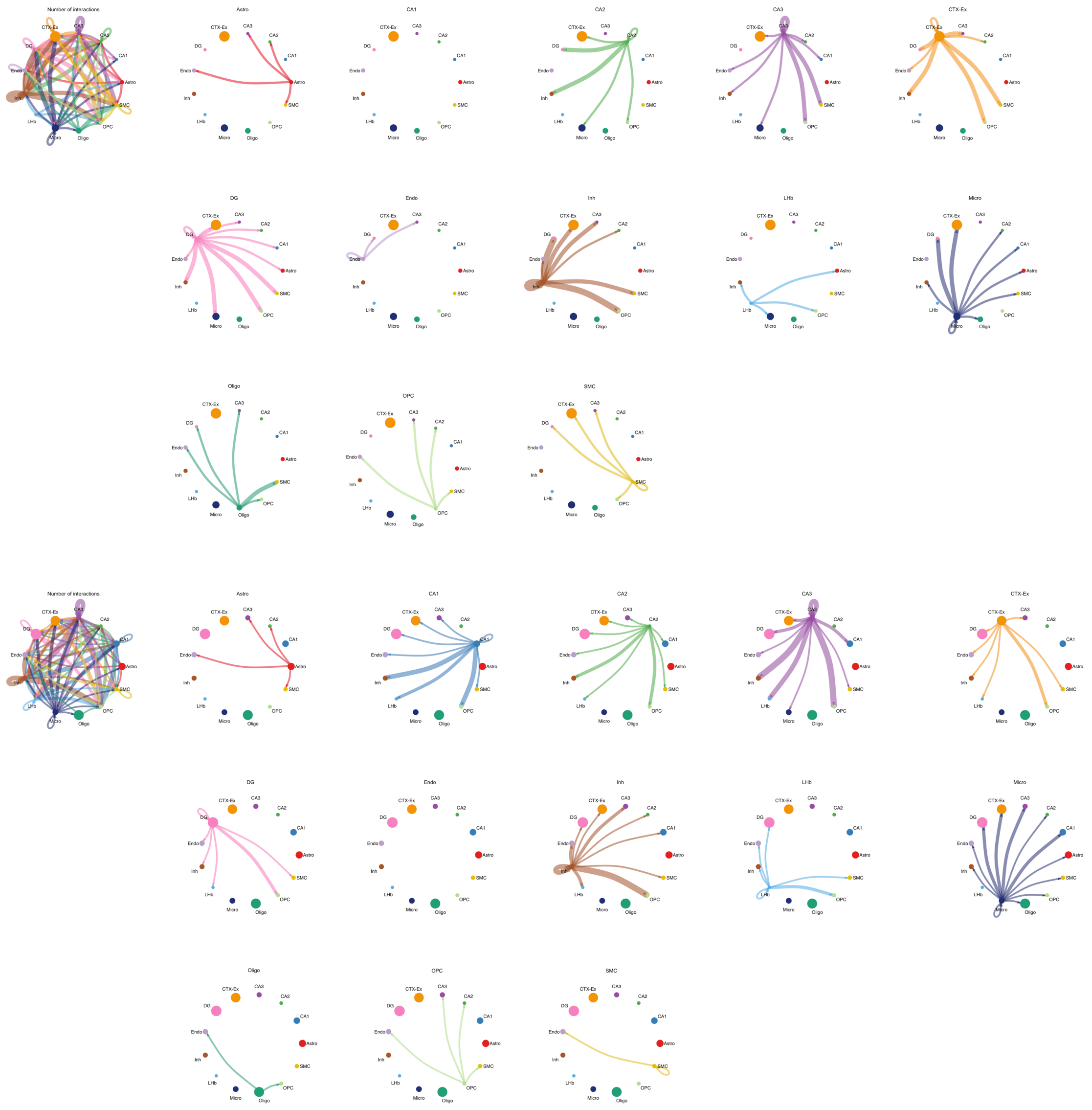
Supplementary Figure 12. Visualization of CellChat results on each cell type between in ‘CCC’ motif regions (motif and its 3-hop, top) and non-motif regions (bottom) in thirteen-month-old samples.

Supplementary Figure 13. Visualization of CellChat results on each cell type between in ‘CCM’ motif regions (motif and its 3-hop, top) and non-motif regions (bottom) in eight-month-old samples.

Supplementary Figure 14. Visualization of CellChat results on each cell type between in 'CCM' motif regions (motif and its 3-hop, top) and non-motif regions (bottom) in thirteen-month-old samples.



Supplementary Figure 15. Visualization of CellChat results on each cell type between in ‘CMM’ motif regions (motif and its 3-hop, top) and non-motif regions (bottom) in eight-month-old samples.

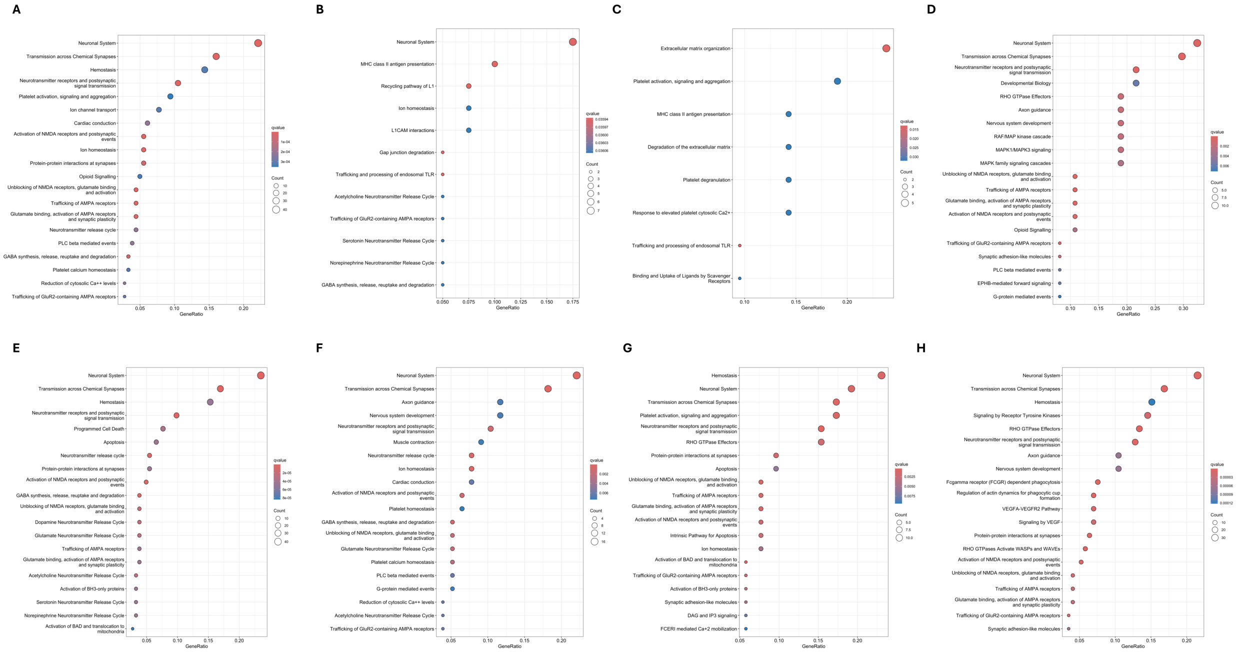


Supplementary Figure 16. Visualization of CellChat results on each cell type between in ‘CMM’ motif regions (motif and its 3-hop, top) and non-motif regions (bottom) in thirteen-month-old samples.

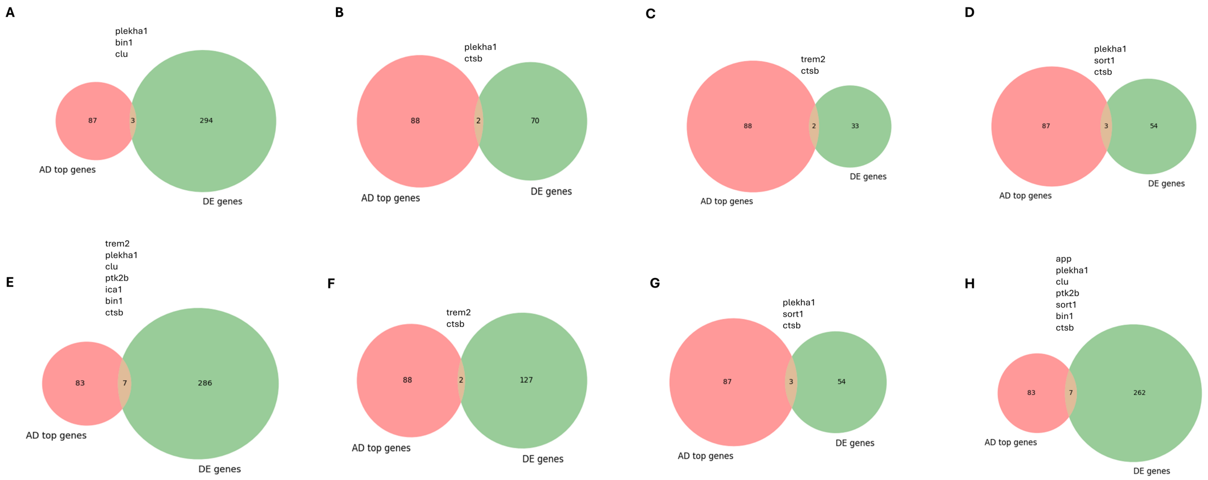
Supplementary Figure 17. Visualization of CellChat results on each cell type between in 'MMM' motif regions (motif and its 3-hop, top) and non-motif regions (bottom) in eight-month-old samples.

Supplementary Figure 18. Visualization of CellChat results on each cell type between in ‘MMM’ motif regions (motif and its 3-hop, top) and non-motif regions (bottom) in thirteen-month-old samples.

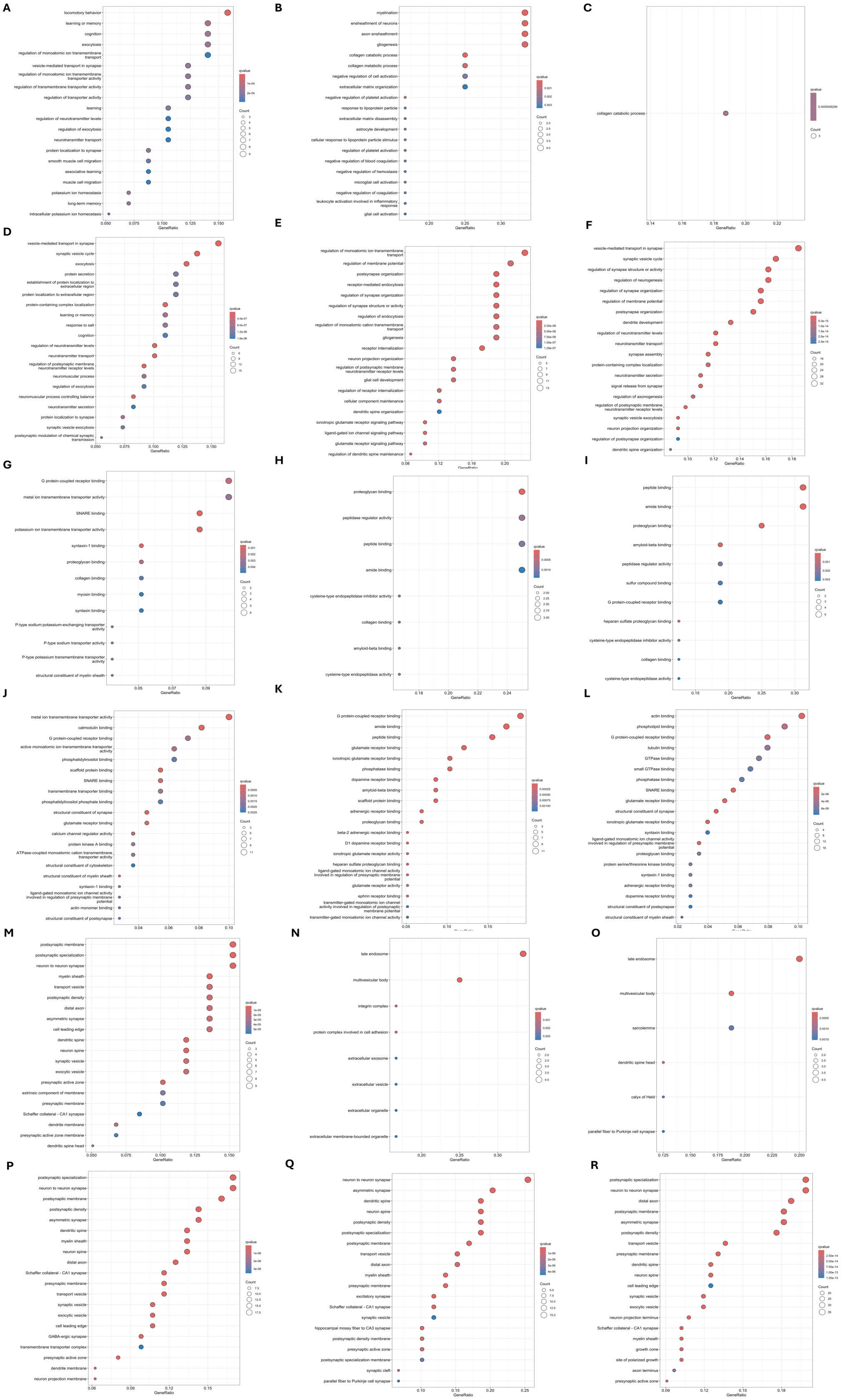
Supplementary Figure 19. GO enrichment analysis results for DEGs in size-3 CC motif regions. GO enrichment analysis on Biological Process results in eight-month-old samples for CC motifs **A.** ‘CCC’, **B.** ‘CCM’**, C.** ‘CMM’, and **D.** ‘MMM’. GO enrichment analysis on Biological Process results in thirteen-month-old samples for CC motifs **E.** ‘CCC’, **F.** ‘CCM’, **G.** ‘CMM’**,** and **H.** ‘MMM’. GO enrichment analysis on Molecular Function results in eight-month-old samples for CC motifs **I.** ‘CCC’, **J.** ‘CCM’**, K.** ‘CMM’, and **L.** ‘MMM’. GO enrichment analysis on Molecular Function results in thirteen-month-old samples for CC motifs **M.** ‘CCC’, **N.** ‘CCM’**, O.** ‘CMM’, and **P.** ‘MMM’. GO enrichment analysis on Cellular Component results in eight-month-old samples for CC motifs **Q.** ‘CCC’, **R.** ‘CCM’**, S.** ‘CMM’, and **T.** ‘MMM’. GO enrichment analysis on Cellular Component results in thirteen-month-old samples for CC motifs **U.** ‘CCC’, **V.** ‘CCM’**, W.** ‘CMM’, and **X.** ‘MMM’.



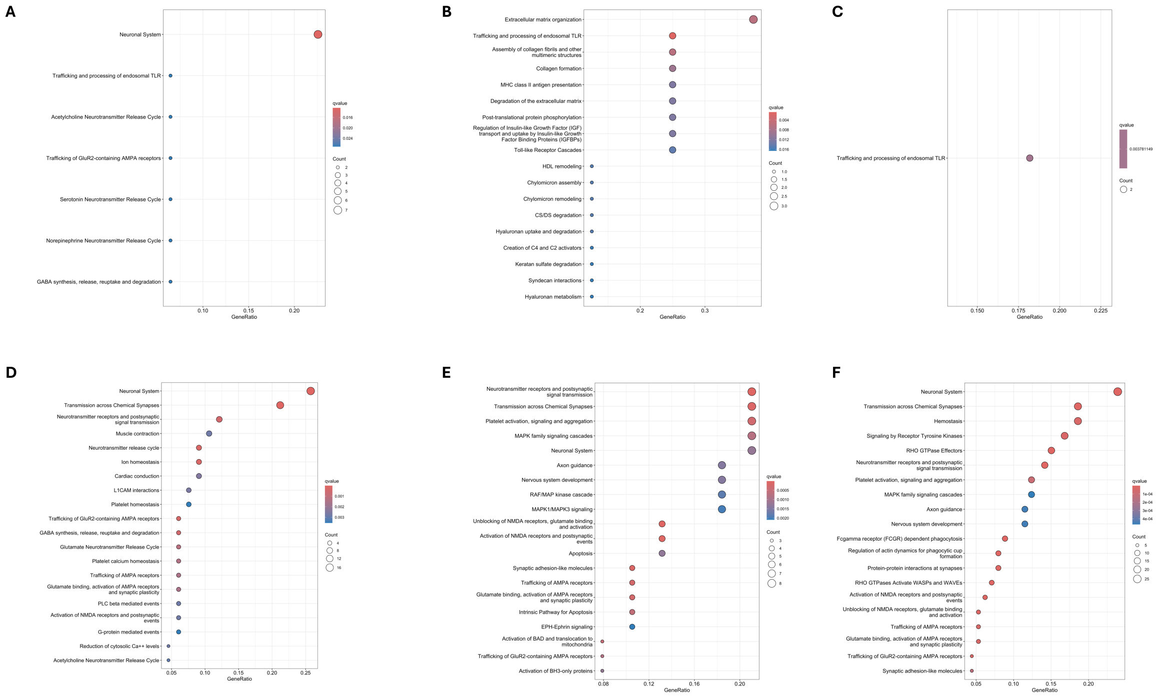
Supplementary Figure 20. Pathway enrichment analysis results for DEGs in size-3 CC motif regions. Pathway enrichment analysis results in eight-month-old samples for CC motifs **A.** ‘CCC’, **B.** ‘CCM’**, C.** ‘CMM’, and **D.** ‘MMM’. Pathway enrichment analysis results in thirteen-month-old samples for CC motifs **E.** ‘CCC’, **F.** ‘CCM’, **G.** ‘CMM’**,** and **H.** ‘MMM’.



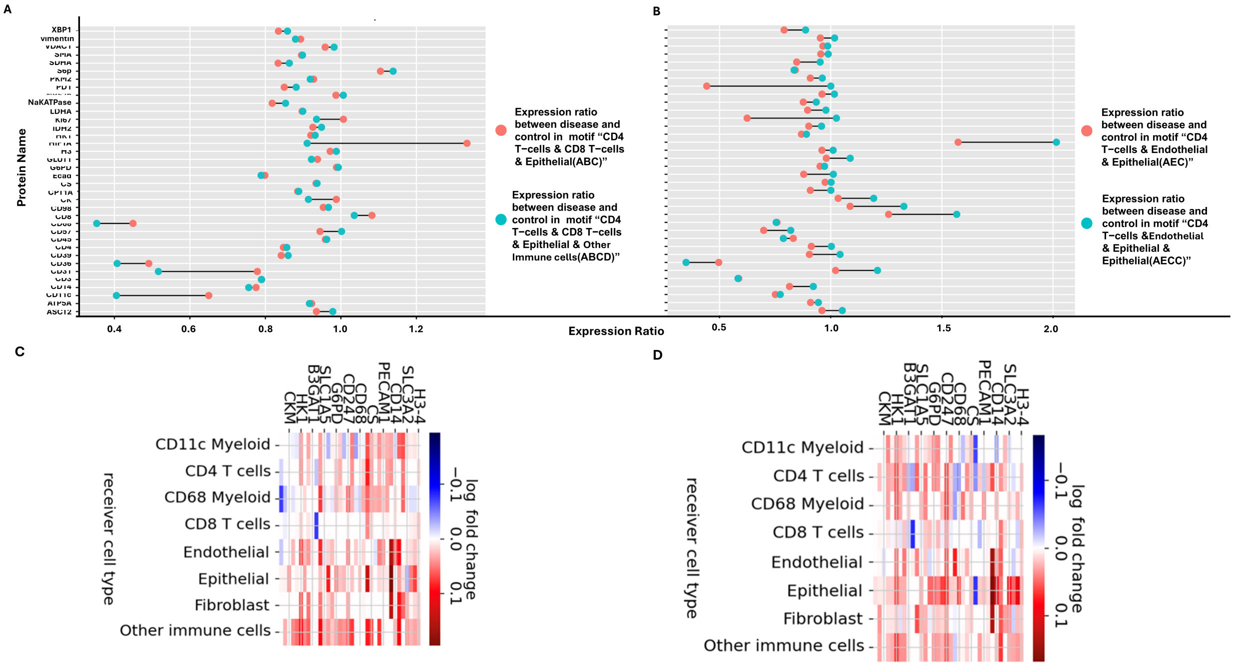
Supplementary Figure 21. The Venn diagrams of top genes by large-scale Alzheimer's disease GWAS analysis and DEGs in CC motif regions, including gene overlaps for motif 'CCC' in **A.** eight-month-old and **B.** 13-month-old samples; motif 'CCM' in **C.** eight-month-old and **D.** thirteen-month-old samples; motif 'CMM' in **E.** eight-month-old and **F.** thirteen-month-old samples; motif 'MMM' in **G.** eight-month-old and **H.** thirteen-month-old samples.


A
locomotory behavior
learning or memory
cognition
exocytosis
regulation of monoatomic ion transmembrane transport
vesicle-mediated transport in synapse
regulation of monoatomic ion transmembrane transporter activity
regulation of transmembrane transporter activity
regulation of transporter activity
learning
regulation of neurotransmitter levels
regulation of exocytosis
neurotransmitter transport
protein localization to synapse
smooth muscle cell migration
associative learning
muscle cell migration
potassium ion homeostasis
long-term memory
intracellular potassium ion homeostasis
0.050
0.075
0.100
0.125
0.150
GeneRatio
qvalue
1e-04
2e-04
Count
3
4
5
6
7
8
9
B
myelination
ensheathment of neurons
axon ensheathment
gliogenesis
collagen catabolic process
collagen metabolic process
negative regulation of cell activation
extracellular matrix organization
negative regulation of platelet activation
response to lipoprotein particle
extracellular matrix disassembly
astrocyte development
cellular response to lipoprotein particle stimulus
regulation of platelet activation
negative regulation of blood coagulation
negative regulation of hemostasis
microglial cell activation
negative regulation of coagulation
leukocyte activation involved in inflammatory response
glial cell activation
0.20
0.25
0.30
GeneRatio
qvalue
0.001
0.002
0.003
Count
2.0
2.5
3.0
3.5
4.0
C
collagen catabolic process
0.14
0.16
0.18
0.20
0.22
GeneRatio
qvalue
0.0005009256
Count
3
D
vesicle-mediated transport in synapse
synaptic vesicle cycle
exocytosis
protein secretion
establishment of protein localization to extracellular region
protein localization to extracellular region
protein-containing complex localization
learning or memory
response to salt
cognition
regulation of neurotransmitter levels
neurotransmitter transport
regulation of postsynaptic membrane neurotransmitter receptor levels
neuromuscular process
regulation of exocytosis
neuromuscular process controlling balance
neurotransmitter secretion
protein localization to synapse
synaptic vesicle exocytosis
postsynaptic modulation of chemical synaptic transmission
0.050
0.075
0.100
0.125
0.150
GeneRatio
qvalue
4.0e-07
8.0e-07
1.2e-06
1.6e-06
Count
6
9
12
15
E
regulation of monoatomic ion transmembrane transport
regulation of membrane potential
postsynapse organization
receptor-mediated endocytosis
regulation of synapse organization
regulation of synapse structure or activity
regulation of endocytosis
regulation of monoatomic cation transmembrane transport
gliogenesis
receptor internalization
neuron projection organization
regulation of postsynaptic membrane neurotransmitter receptor levels
glial cell development
regulation of receptor internalization
cellular component maintenance
dendritic spine organization
ionotropic glutamate receptor signaling pathway
ligand-gated ion channel signaling pathway
glutamate receptor signaling pathway
regulation of dendritic spine maintenance
0.08
0.12
0.16
0.20
qvalue
2.50e-08
5.00e-08
7.50e-08
1.00e-07
1.25e-07
Count
5
7
9
11
13
F
vesicle-mediated transport in synapse
synaptic vesicle cycle
regulation of synapse structure or activity
regulation of neurogenesis
regulation of synapse organization
regulation of membrane potential
postsynapse organization
dendrite development
regulation of neurotransmitter levels
neurotransmitter transport
synapse assembly
protein-containing complex localization
neurotransmitter secretion
signal release from synapse
regulation of axonogenesis
regulation of postsynaptic membrane neurotransmitter receptor levels
synaptic vesicle exocytosis
neuron projection organization
regulation of postsynapse organization
dendritic spine organization
0.10
0.12
0.14
0.16
0.18
GeneRatio
qvalue
5.0e-15
1.0e-14
1.5e-14
2.0e-14
2.5e-14
Count
16
20
24
28
32
G
G protein-coupled receptor binding
metal ion transmembrane transporter activity
SNARE binding
potassium ion transmembrane transporter activity
syntaxin-1 binding
proteoglycan binding
collagen binding
myosin binding
syntaxin binding
P-type sodium:potassium-exchanging transporter activity
P-type sodium transporter activity
P-type potassium transmembrane transporter activity
structural constituent of myelin sheath
0.05
0.07
0.09
GeneRatio
qvalue
0.001
0.002
0.003
0.004
Count
2
3
4
5
6
H
proteoglycan binding
peptidase regulator activity
peptide binding
amide binding
cysteine-type endopeptidase inhibitor activity
collagen binding
amyloid-beta binding
cysteine-type endopeptidase activity
0.18
0.20
0.22
0.24
GeneRatio
qvalue
0.0005
0.0010
Count
2.00
2.25
2.50
2.75
3.00
I
peptide binding
amide binding
proteoglycan binding
amyloid-beta binding
peptidase regulator activity
sulfur compound binding
G protein-coupled receptor binding
heparan sulfate proteoglycan binding
cysteine-type endopeptidase inhibitor activity
collagen binding
cysteine-type endopeptidase activity
0.15
0.20
0.25
0.30
GeneRatio
qvalue
0.001
0.002
0.003
Count
2
3
4
5
J
metal ion transmembrane transporter activity
calmodulin binding
G protein-coupled receptor binding
active monoatomic ion transmembrane transporter activity
phosphatidylinositol binding
scaffold protein binding
SNARE binding
transmembrane transporter binding
phosphatidylinositol phosphate binding
structural constituent of synapse
glutamate receptor binding
calcium channel regulator activity
protein kinase A binding
ATPase-coupled monoatomic cation transmembrane transporter activity
structural constituent of cytoskeleton
structural constituent of myelin sheath
syntaxin-1 binding
ligand-gated monoatomic ion channel activity involved in regulation of presynaptic membrane potential
actin monomer binding
structural constituent of postsynapse
0.04
0.06
0.08
0.10
qvalue
0.0005
0.0010
0.0015
0.0020
0.0025
Count
3
5
7
9
11
K
G protein-coupled receptor binding
amide binding
peptide binding
glutamate receptor binding
ionotropic glutamate receptor binding
phosphatase binding
dopamine receptor binding
amyloid-beta binding
scaffold protein binding
adrenergic receptor binding
proteoglycan binding
beta-2 adrenergic receptor binding
D1 dopamine receptor binding
ionotropic glutamate receptor activity
heparan sulfate proteoglycan binding
ligand-gated monoatomic ion channel activity involved in regulation of presynaptic membrane potential
glutamate receptor activity
ephrin receptor binding
transmitter-gated monoatomic ion channel activity involved in regulation of postsynaptic membrane potential
transmitter-gated monoatomic ion channel activity
0.05
0.10
0.15
qvalue
0.00025
0.00050
0.00075
0.00100
Count
3
5
7
9
11
L
actin binding
phospholipid binding
G protein-coupled receptor binding
tubulin binding
GTPase binding
small GTPase binding
phosphatase binding
SNARE binding
glutamate receptor binding
structural constituent of synapse
ionotropic glutamate receptor binding
syntaxin binding
ligand-gated monoatomic ion channel activity involved in regulation of presynaptic membrane potential
proteoglycan binding
protein serine/threonine kinase binding
syntaxin-1 binding
adrenergic receptor binding
dopamine receptor binding
structural constituent of postsynapse
structural constituent of myelin sheath
0.02
0.04
0.06
0.08
0.10
qvalue
2e-06
4e-06
6e-06
Count
4
8
12
16
M
postsynaptic membrane
postsynaptic specialization
neuron to neuron synapse
myelin sheath
transport vesicle
postsynaptic density
distal axon
asymmetric synapse
cell leading edge
dendritic spine
neuron spine
synaptic vesicle
exocytic vesicle
presynaptic active zone
extrinsic component of membrane
presynaptic membrane
Schaffer collateral - CA1 synapse
dendrite membrane
presynaptic active zone membrane
dendritic spine head
0.050
0.075
0.100
0.125
0.150
GeneRatio
qvalue
1e-05
2e-05
3e-05
4e-05
5e-05
Count
3
4
5
6
7
8
9
N
late endosome
multivesicular body
integrin complex
protein complex involved in cell adhesion
extracellular exosome
extracellular vesicle
extracellular organelle
extracellular membrane-bounded organelle
0.20
0.25
0.30
GeneRatio
qvalue
0.001
0.002
0.003
Count
2.0
2.5
3.0
3.5
4.0
O
late endosome
multivesicular body
sarcolemma
dendritic spine head
calyx of Held
parallel fiber to Purkinje cell synapse
0.125
0.150
0.175
0.200
0.225
0.250
GeneRatio
qvalue
0.0005
0.0010
0.0015
Count
2.0
2.5
3.0
3.5
4.0
P
postsynaptic specialization
neuron to neuron synapse
postsynaptic membrane
postsynaptic density
asymmetric synapse
dendritic spine
myelin sheath
neuron spine
distal axon
Schaffer collateral - CA1 synapse
presynaptic membrane
transport vesicle
synaptic vesicle
exocytic vesicle
cell leading edge
GABA-ergic synapse
transmembrane transporter complex
presynaptic active zone
dendrite membrane
neuron projection membrane
0.06
0.09
0.12
0.15
GeneRatio
qvalue
1e-06
2e-06
3e-06
Count
7.5
10.0
12.5
15.0
17.5
Q
neuron to neuron synapse
asymmetric synapse
dendritic spine
neuron spine
postsynaptic density
postsynaptic specialization
postsynaptic membrane
transport vesicle
distal axon
myelin sheath
presynaptic membrane
excitatory synapse
Schaffer collateral - CA1 synapse
synaptic vesicle
hippocampal mossy fiber to CA3 synapse
postsynaptic density membrane
presynaptic active zone
postsynaptic specialization membrane
synaptic cleft
parallel fiber to Purkinje cell synapse
0.10
0.15
0.20
0.25
GeneRatio
qvalue
1e-06
2e-06
3e-06
4e-06
Count
5.0
7.5
10.0
12.5
15.0
R
postsynaptic specialization
neuron to neuron synapse
distal axon
postsynaptic membrane
asymmetric synapse
postsynaptic density
transport vesicle
presynaptic membrane
dendritic spine
neuron spine
cell leading edge
synaptic vesicle
exocytic vesicle
neuron projection terminus
Schaffer collateral - CA1 synapse
myelin sheath
growth cone
site of polarized growth
axon terminus
presynaptic active zone
0.09
0.12
0.15
0.18
GeneRatio
qvalue
2.50e-14
5.00e-14
7.50e-14
1.00e-13
1.25e-13
Count
20
25
30
35

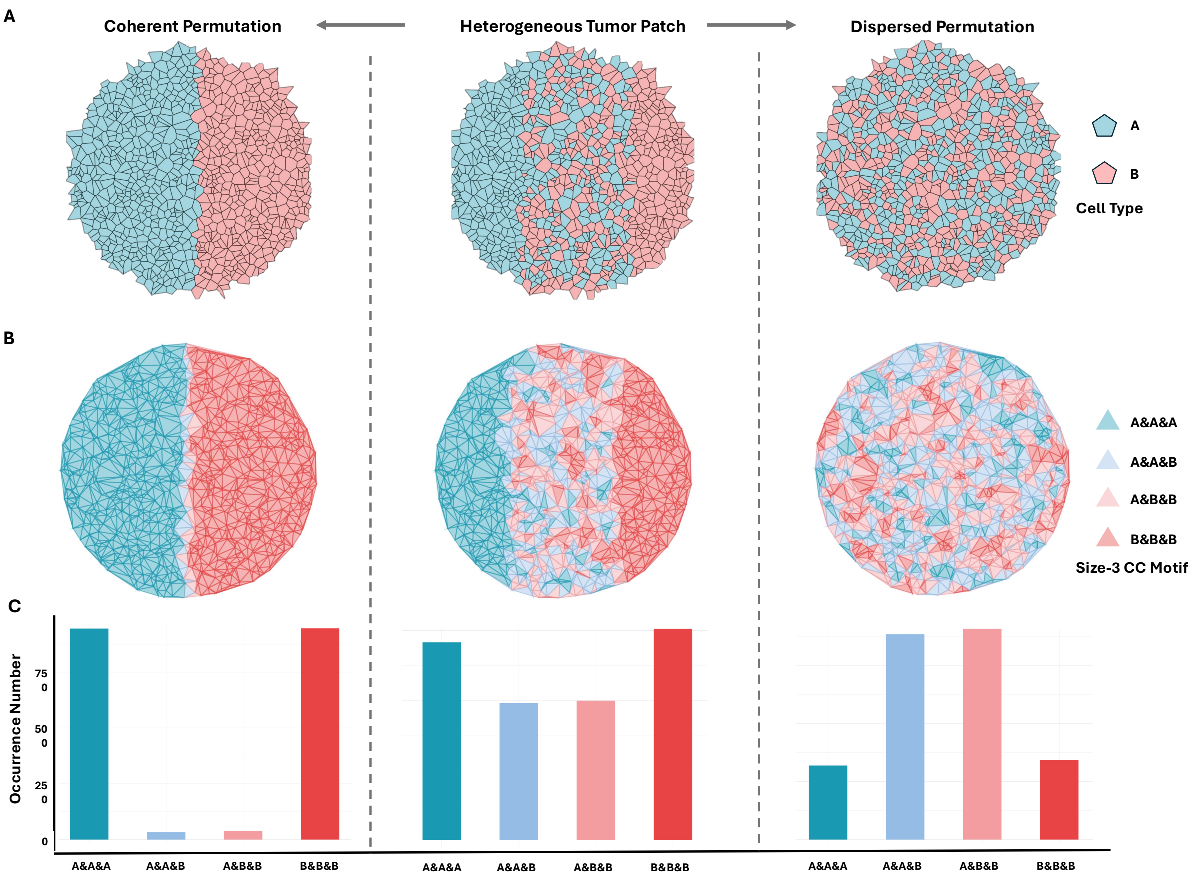
Supplementary Figure 22. GO enrichment analysis results for DEGs in size-4 CC motif regions. GO enrichment analysis on Biological Process results in eight-month-old samples for motifs **A**. ‘CCCM’, **B**. ‘MMMC’, and **C**. ‘MMMM’. GO enrichment analysis on Biological Process results in thirteen-month-old samples for motifs **D**. ‘CCCM’, **E**. ‘MMMC’, and **F**. ‘MMMM’. GO enrichment analysis on Molecular Function results in eight-month-old samples for motifs **G** ‘CCCM’, **H**. ‘MMMC’, and **I**. ‘MMMM’. GO enrichment analysis on Molecular Function results in thirteen-month-old samples for motifs **J**. ‘CCCM’, **K**. ‘MMMC’, and **L**. ‘MMMM’. GO enrichment analysis on Cellular Component results in eight-month-old samples for motifs **M**. ‘CCCM’, **N**. ‘MMMC’, and **O**. ‘MMMM’. GO enrichment analysis on Cellular Component results in thirteen-month-old samples for motifs **P**. ‘CCCM’, **Q**. ‘MMMC’, and **R**. ‘MMMM’.



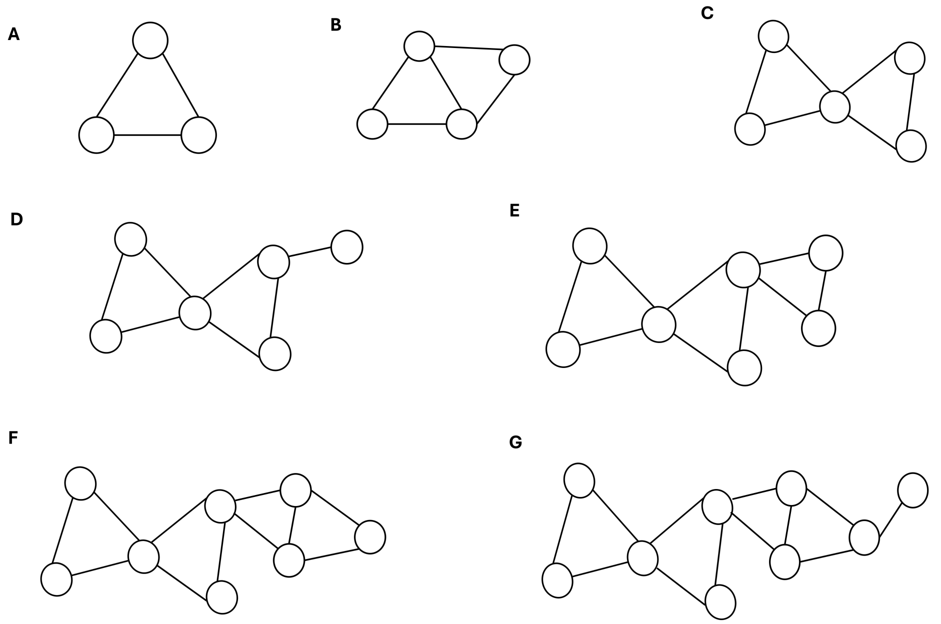
Supplementary Figure 23. Pathway enrichment analysis results for DEGs in size-4 CC motif regions. Pathway enrichment analysis results in eight-month-old samples for CC motifs **A.** ‘CCCM’, **B.** ‘MMMC’, and **C.** ‘MMMM’. Pathway enrichment analysis results in thirteen-month-old samples for CC motifs **D.** ‘CCCM’ , **E.** ‘MMMC’ , and **F.** ‘MMMM’.



Supplementary Figure 24. Dumbbell plot of expression ratio between disease and control samples in **A.** Shifted Interaction Motif and **B**. Homeostatic Interaction Motif. Red nodes denote the expression ratio in the size-3 motifs, and blue nodes denote its successor size-4 motifs. Sender effect on CD4 T-cells in **C**. Colon Carcinoma samples and **D**. Healthy control sample by NCEM gene-wise analysis.



Supplementary Figure 25. A demo system describes the dispersed and coherent cell organizations by the abundance of CC motifs. **A**. The spatial distribution of cell type A (Blue) and B (Red) in different cell organizations. **B**. The distribution of size-3 CC motif "A&A&A", "A&A&B", "A&B&B", and "B&B&B" across different cell organizations. **C**. The abundance of size-3 CC motifs "A&A&A", "A&A&B", "A&B&B", and "B&B&B" as occurrence numbers across different cell organizations.



Supplementary Figure 26. Seven CC motifs in different topologies used in the training dataset of TrimNN, including **A.** size-3. **B.** size-4. **C.** size-5. **D.** size6. **E.** size-7. **F.** size-8. **G.** size-9.