Supplementary Notes

1. Analysis of CC motif ‘CCC’ and pattern growth using TrimNN in AD study

Similar to analyzing CC motif ‘MMM’, we also performed analysis on CC motif ‘CCC’. Comparing ‘CCC’ motif-enriched and complementary regions, differentially expressed genes (DEGs) were identified as significant (p-value<0.05) using DESeq2, including *Plekha1*, *Bin1*, *Clu*, *Ctsb*, *Sort1* in eight-month-old samples, and *Trem2*, *App*, *Ptk2b*, *Bin1*, *Ctsb* in thirteen-month-old samples. On DEGs in thirteen-month-old samples, Gene Ontology (GO) enrichment analysis showed significant vesicle-mediated transport in synapse (q-value 9.09e-47), synaptic vesicle cycle (q-value 6.87e-41), regulation of synapse structure (q-value 7.79e-32), and neural transmitter transport (q-value 8.48e-30) (**Supplementary Fig. 19E**). Neural systems (q-value 2.91e-20), Transmission across Chemical Synapses (q-value 4.33e-16), Neurotransmitter receptors and postsynaptic signal transmission (q-value 1.20e-07), Apoptosis (q-value 4.62e-05), and Norepinephrine Neurotransmitter Release Cycle (q-value 4.62e-05) were enriched with pathway enrichment analysis (**Supplementary Fig. 20E**).

Based on the identified size-3 motif ‘CCC’, we performed pattern growth to identify size-4 motifs using TrimNN. Among all the size-4 ‘CCC’ expanded motifs, ‘CCCM’ showed the most significant difference between AD and control samples upon the motif ‘CCC’ with *p-value* 1.78e-09 and 5.04e-21 using Benjamini-Hochberg adjusted Fisher’s exact test. The ‘CCCM’ motif occurred 434 and 252 times in the sum of two replicates of eight-month-old samples and 368 and 266 in the sum of two replicates of thirteen-month-old samples, respectively. Similar to size-3 analysis, CellChat identified the cell-cell communications of motif ‘CCCM’ and its 3-hop regions, and we found genes GRN, CCL, VIP, VEGF, PDGF, PTN, and CXCL are consistent with the results from the size-3 motifs. The top AD risk genes intersected with DEGs in ‘CCCM’ is Ctsb in eight-month-old samples and Sort1, Bin1, Ctsb, App, and Ptk2b in thirteen-month-old samples. Using these DEGs in motif ‘CCCM’ in thirteen-month-old samples, GO enrichment analysis (**Supplementary Fig. 22D**) showed that post-synapse organization (q-value 7.59e-23), learning or memory (q-value 2.63e-14), cognition (q-value 1.93e-13), and receptor localization to synapse (q-value 1.22e-15) were enriched biological processes. Also, Pathway enrichment analysis (**Supplementary Fig. 23D**) showed Neuronal System (q-value 1.76e-08), Transmission across Chemical Synapses (q-value 2.98e-06), Synaptic adhesion-like molecules (q-value 8.27e-04), Neurexins and Neuroligins gene enrichment (q-value 2.93e-03).

2. Revisit AD case study with Shifted Interaction Motifs and Homeostatic Interaction Motifs

After introducing the concepts of Shifted Interaction Motifs and Homeostatic Interaction Motifs, we revisited the AD case study by analyzing size-3 and their related size-4 CC motifs with TrimNN. We performed a proportion test on size-3 motif ‘CCC’ and size-4 motif

‘CCCM’ among disease and control samples, where ‘C’ denotes cortex excitatory neuron and ‘M’ denotes microglia. All eight-month-old replicates 1 and 2 samples and thirteen-month-old replicates 1 and 2 samples obtained significant results with *p*-values 1.87E-05, 1.32e-06, 1.19E-15, and 5.84E-12, respectively. Meanwhile, on size-3 motif ‘MMM’ and size-4 motif ‘MMMC’ (*p*-values 0.006913, 1, 1, and 1), and size-3 motif ‘MMM’ and size motif ‘MMMM’ (*p*-values 0.2359, 1, 0.1225, and 0.9093), both were not significant using proportion test in eight-month-old replicates 1 and 2 samples and thirteen-month-old replicates 1 and 2 samples (**Supplementary Data 64**). These results demonstrated that the transition from motif ‘CCC’ to ‘CCCM’ was aligned with Shifted Interaction Motifs, and the other two types of motifs were aligned with Homeostatic Interaction Motifs, further validating the differences between these CC motifs.

3. TrimNN analysis on size-3 and size-4 CC motifs in the colorectal carcinoma study compared to NCEM

Compared to cell type coupling analysis in NCEM which estimates cell interactions of the effects of niche composition, CD4 T-cell type has trends to consistently couple across different cell types in colorectal samples (**Fig. 5L**), i.e., CD4 T-cells have top ranks (1st, 2nd, 3rd out of 8) in edge-wise analysis and it also showed that half of the effect values in NCEM were the highest (except CD8 T-cells as receiver) in CD4 T-cells sender effect. Nevertheless, the motif analysis was stable within specific cell types, leading to more consistent rankings. On the other hand, the results from NCEM appeared to have more variations among receivers for a single sender. They did not show a pattern getting higher rank values towards inner single cell types, i.e., endothelial cells as senders, they have ranks 2nd and 8th out of 8 and also rank 8th for the homogeneous endothelial cells.

NCEM showed that CD8 T-cells as senders and epithelial cells as receivers can differentiate diseases and controls (**Fig. 5R**). However, more details such as CC motifs can provide more insights into spatial tendencies. Targeted the size-2 couplings of CD8 T-cells and epithelial cells, only triangle patterns combining CD4 T-cells were considered significant (*p-value* 3.58e-12, Fisher’s exact test adjusting by Benjamini-Hochberg method) among all eight possible combinations in size-3 CC motifs (**Supplementary Data 63**). This observation demonstrated that the triangle-like size-3 CC motif of the cell type CD4 T-cells, CD8 T-cells, and Epithelial (‘ABC’) might be essential to capture the main difference among disease and control samples in colorectal carcinoma. Based on motif ‘ABC’, TrimNN with pattern growth search also identified one size-4 CC motif with a cell type Other Immune cells (other CD45+) (‘ABCD’) significant in both disease and control samples. This was also a Shifted Interaction Motifs shown in **Fig. 5A**.