Cell Reports, Volume *42*

Supplemental information

Emergence of division of labor in tissues through cell interactions and spatial cues

Miri Adler, Noa Moriel, Aleksandrina Goeva, Inbal Avraham-Davidi, Simon Mages, Taylor S. Adams, Naftali Kaminski, Evan Z. Macosko, Aviv Regev, Ruslan Medzhitov, and Mor Nitzan

Supplementary information

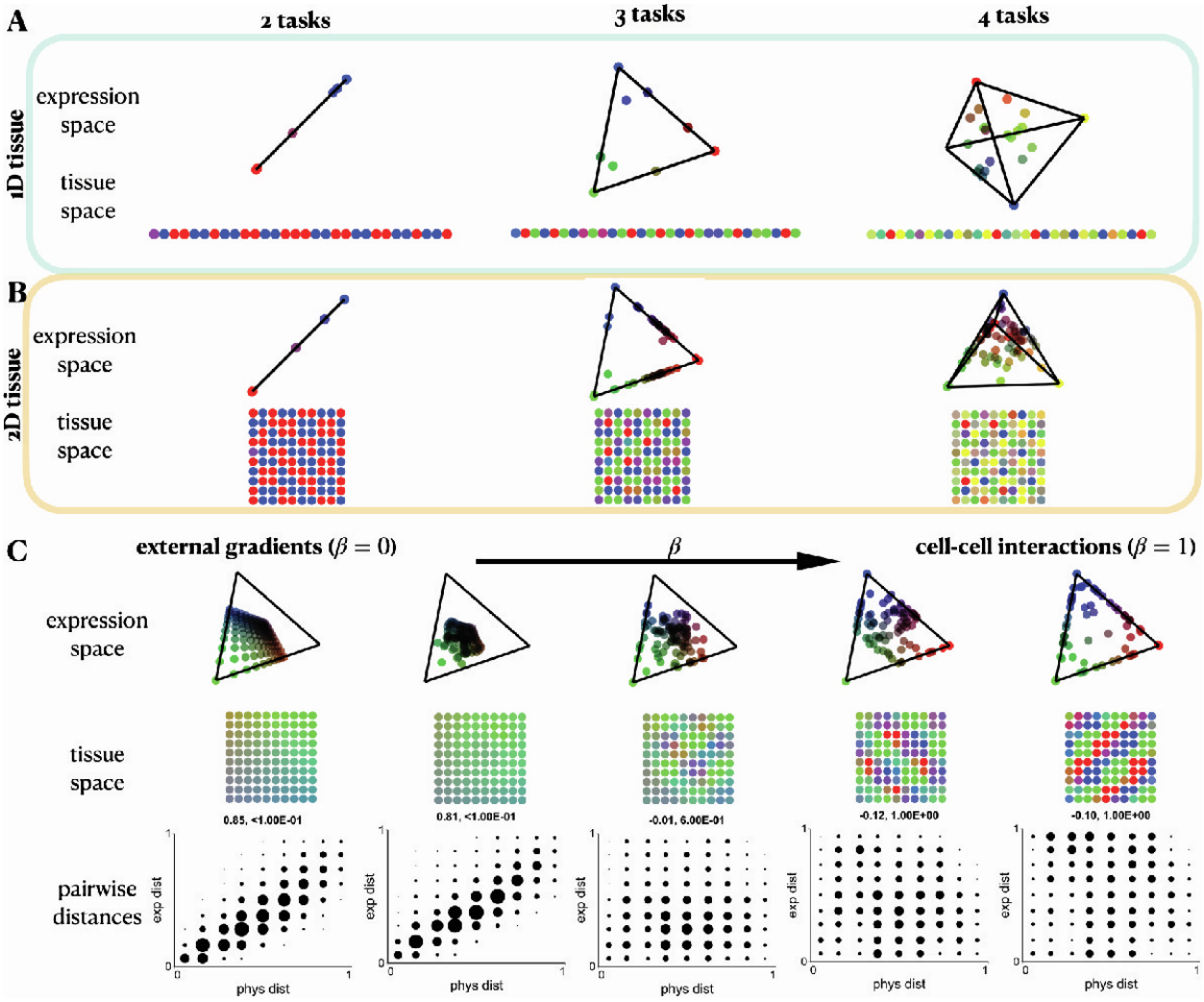


Figure S1: Simulation results with lateral inhibition and interpolation between gradients and interactions, Related to Figure 2. Columns describe the number of tasks (2, 3, and 4). Rows describe the dimension of the spatial organizations of cells in the tissue for 1D grid with 30 cells **(A)**, and 2D grid with 100 cells **(B)**. In all cases, a neighborhood of range 2 is considered. We plot the expression space (vertices represent archetypes' expression), and tissue space where cells' colors indicate their task allocation among the archetypes. **(C)** $\beta = (0, 0.5, 0.6, 0.75, 1)$ correspond to columns left to right, for 2D tissue with 3 tasks considering neighborhood of range=4. For each, expression and tissue spaces are plotted. Pairwise distance plots are titled with the Pearson correlation between expression and physical distances, and its p-val.

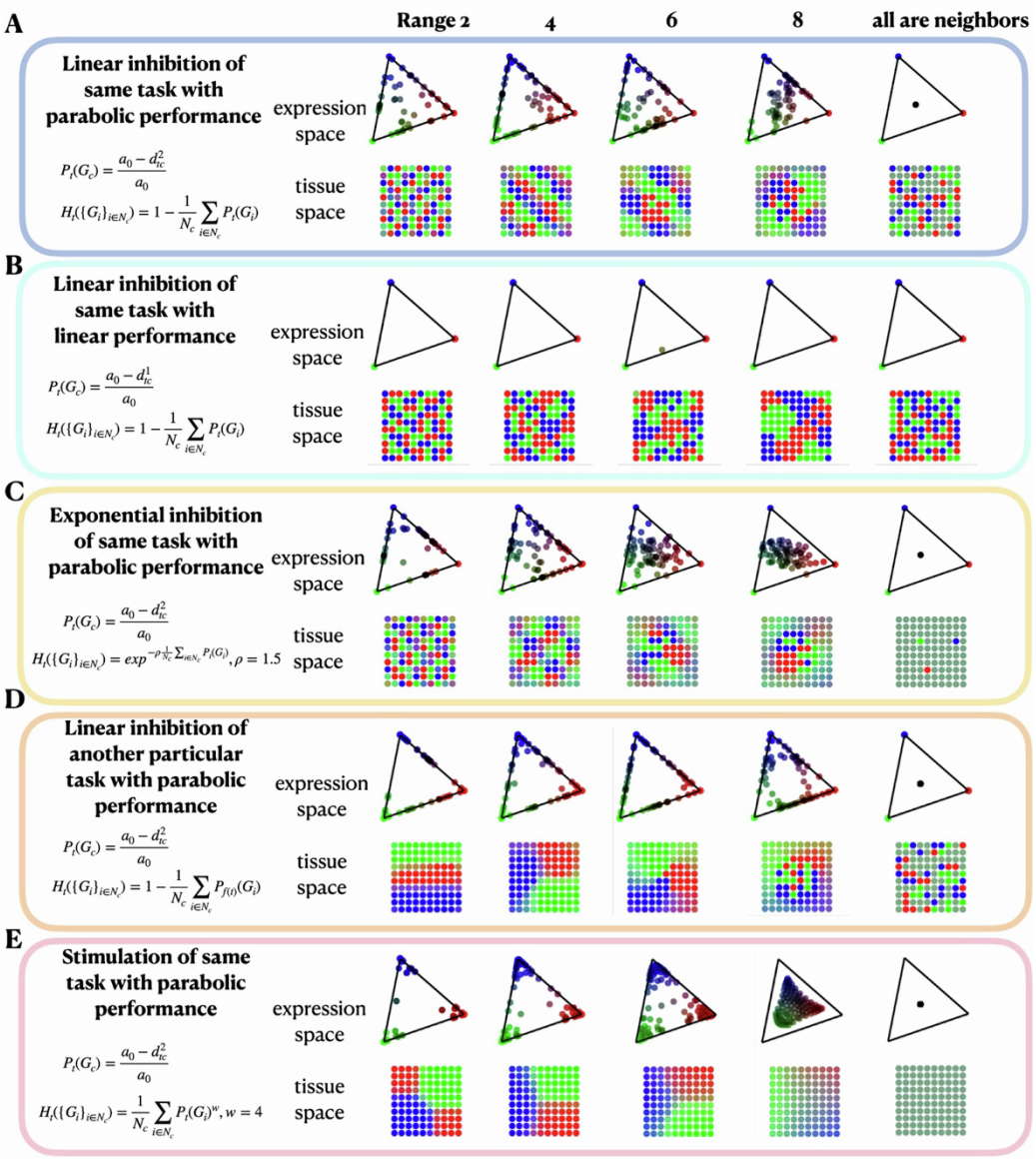
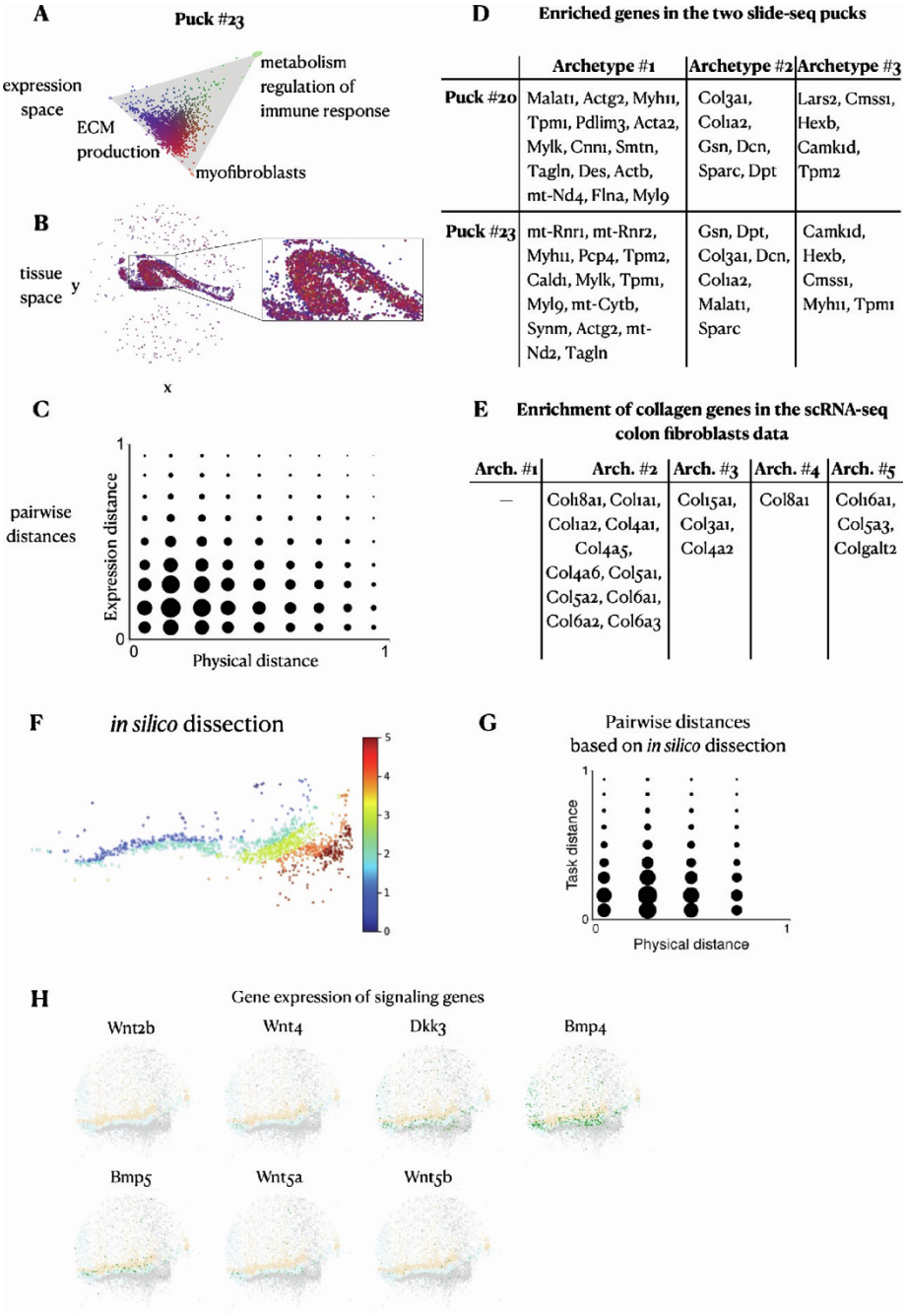


Figure S2: Simulation results for alternative models as a function of the range of interactions, Related to STAR Methods. (A) Performance and interaction functional forms used throughout the main text with the parameter $a_0 = 2$. **(B)** Performance linearly decreases as a function of the Euclidean distance from archetypes. **(C)** Interaction term, H_t, decays exponentially with the neighbors' mean performance in task t with rate factor $\rho = 1.5$. **(D)** Cyclic inhibition where task $f(t) = (t + 1)\ mod\ m$ inhibits the performance in task t, where m is the number of tasks. **(E)** Stimulating (positive) interactions weighted by a mean of the neighbors' performance to the power of $w = 4$.



	Archetype #1	Archetype #2	Archetype #3
Puck #20	Malat1, Actg2, Myh11, Tpm1, Pdlim3, Acta2, Mylk, Cnn1, Smtn, Tagln, Des, Actb, mt-Nd4, Flna, Myl9	Col3a1, Col1a2, Gsn, Dcn, Sparc, Dpt	Lars2, Cmss1, Hexb, Camk1d, Tpm2
Puck #23	mt-Rnr1, mt-Rnr2, Myh11, Pcp4, Tpm2, Cald1, Mylk, Tpm1, Myl9, mt-Cytb, Synm, Actg2, mt-Nd2, Tagln	Gsn, Dpt, Col3a1, Dcn, Col1a2, Malat1, Sparc	Camk1d, Hexb, Cmss1, Myh11, Tpm1

Arch. #1	Arch. #2	Arch. #3	Arch. #4	Arch. #5
—	Col18a1, Col1a1, Col1a2, Col4a1, Col4a5, Col4a6, Col5a1, Col5a2, Col6a1, Col6a2, Col6a3	Col15a1, Col3a1, Col4a2	Col8a1	Col16a1, Col5a3, Colgalt2

Figure S3: Analysis of additional puck from the Slide-seq data and in silico dissection, Related to Figure 3. (A) Analysis of puck #23 from the Slide-seq data of colon fibroblasts where cells' expression profile is projected on the first two PCs, and the cells are colored by their task specializations. **(B)** Spatial coordinates of the cells in tissue space. **(C)** The pairwise distances in physical versus expression space of enterocytes show no correlation (corr=0.004, p-val=0.338). **(D)** A table of enriched genes in the two Slide-seq pucks. **(E)** A table of collagen genes and their enrichment in the scRNA-seq data of mouse colon

fibroblasts. **(F)** *In silico* dissection into 6 bins of fibroblasts according to their inferred distance from the deep crypt layer (mimicking how the enterocytes in [12] were assayed). **(G)** Considering for each bead its *in silico* bin as its physical location, we compute the corresponding Pearson correlation of pairwise task distances versus physical distances (corr=0.02, p-val=0.037). **(H)** Distribution of signaling genes' expression across apical-basal colon layers – (light orange) apical plasma membrane, (light blue) deep crypt, (light gray) muscularis. Expression of each gene (in green) is log1p-transformed and truncated at the value of 0.5.

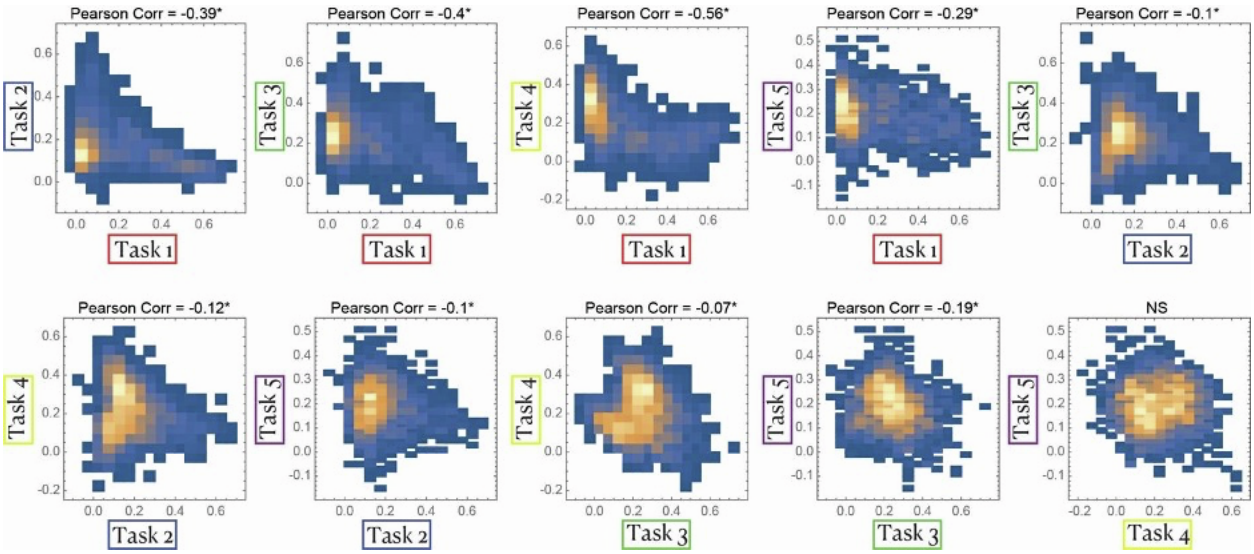


Figure S4: Quantitative analysis of task tradeoffs, Related to Figure 4. Task specialization scores of the Silde-seq fibroblast beads (inferred using the scRNA-seq data analyzed in Figure 4B) are plotted for each pair of archetypes. *p-val<0.001

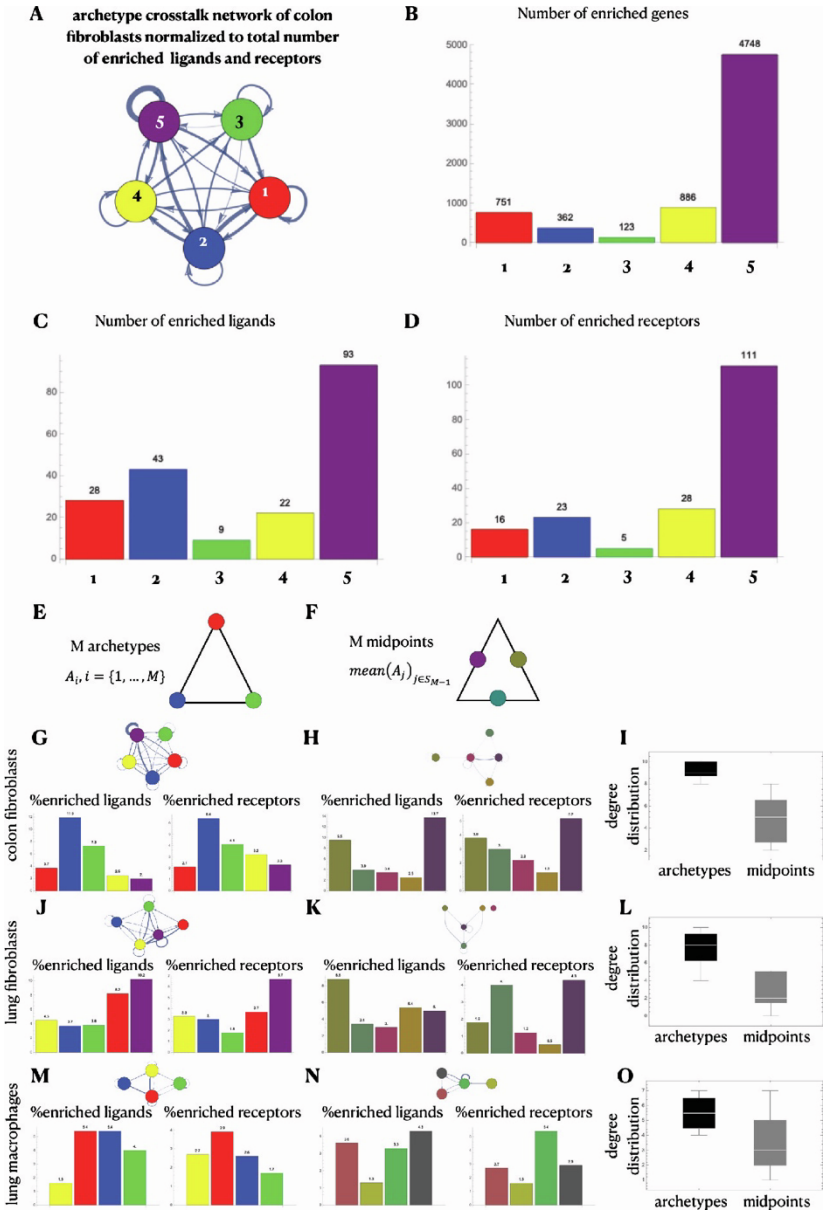


Figure S5: Ligand-receptor enrichments and crosstalk network for archetype midpoints, Related to Figure 5. (A) Archetype crosstalk network of the colon fibroblast scRNA-seq data analyzed in **Figure 4B** where the edge weight from archetype A to archetype B is normalized to the sum of the total number of enriched ligands near archetype A and the total number of enriched receptors near archetype B. **(B-D)** Total number of enriched genes (B), ligands (C) and receptors (D) near each of the five colon fibroblast archetypes. **(E-F)** Comparison of crosstalk networks between the M archetypes **(E)** and between M midpoints **(F)**. **(G-H)** Percentage of enriched ligands and receptors relative to total number of enriched genes near the colon fibroblast archetypes **(G)** and midpoints **(H)**. **(I)** Degree distribution of colon fibroblast

archetype crosstalk network (black) and midpoint crosstalk network (gray). The same only for the lung fibroblast data **(J, K, L)** and the lung macrophage data **(M, N, O)**.

Supplementary item titles

Table S1: Genes enriched towards enterocyte archetypes and ligand-receptor pairs enriched in the enterocytes data, Related to Figure 5.

Table S2: Genes enriched towards colon fibroblast scRNA-seq archetypes and ligand-receptor pairs enriched in the colon fibroblast scRNA-seq data, Related to Figure 4.

Table S3: Genes enriched towards lung fibroblast archetypes, Related to Figure 5.

Table S4: Genes enriched towards lung macrophage archetypes, Related to Figure 5.

Table S5: Ligand-receptor pairs enriched in the lung fibroblast and macrophage data, Related to Figure 5.