**Title: Human single-cell atlas analysis reveals heterogeneous endothelial signaling**

Zimo Zhu[1,2,*], Rongbin Zheng[1,2,*,#] , Yang Yu[1,2], Lili Zhang[1,2], Kaifu Chen[1,2,#]

[1] Department of Cardiology, Boston Children's Hospital.

[2] Department of Pediatrics, Harvard Medical School.

[*] Equal Contribution.

[#] Corresponding: Kaifu.Chen@childrens.harvard.edu; Rongbin.Zheng@childrens.harvard.edu

**Abstract**

Endothelial cells (ECs) play complex roles across tissues and vessel types. Yet, systematic investigations of EC heterogeneity in the combined context of vessel type and tissue microenvironment are still largely lacking. We integrated over three million single cells of scRNA-seq datasets in 15 human tissues and found that ECs in some tissues (e.g., heart and kidney) exhibited greater tissue specificity, while others displayed more substantial vessel specificity. We developed a computational pipeline to analyze cell-cell communications (CCC) mediated by metabolites or proteins to explore microenvironmental regulation. Interestingly, our results showed that CCC events involving ECs varied vastly across tissues, highlighting tissue-specific EC interactions. Using topic modeling, we identified CCC patterns, termed CCC topics, representing metabolite- and protein-mediated interactions between ECs and other tissue-resident cells. Most CCC topics exhibited high tissue specificity, potentially explaining microenvironmental regulations for EC heterogeneity. The work systematically investigates EC heterogeneity and provides insights into how EC heterogeneity was regulated across diverse tissue microenvironments.

**Keywords:**

Cell-cell communication, tissue specificity, human endothelial, heterogeneity, single-cell

## Introduction

Endothelial cells (EC) line the cardiovascular vessels throughout the human body, performing essential roles in various biological processes, including tissue homeostasis, immune responses, and vascular health[1–4]. Despite ECs' common presence, they exhibit remarkable diversity in structure and function that varies by vessel types, anatomic location, and developmental stage [5], demonstrating heterogeneity within the vascular system. Major vessel types include the arterial, venous, capillary, and lymphatic vessels. For example, arterial ECs express markers such as ephrin-B2, which marks arterial identity in the arteries [6]. Venous ECs express EphB4, establishing venous characteristics in the veins [7]. EC also shows differences between tissue types. These differences allow ECs to perform tissue-specific molecular mechanisms, such as cardiac vascular morphogenesis in the heart, gas exchange in the lung, or filtration in the kidney [8]. Studies using single-cell RNA sequencing (scRNA-seq) have shown that ECs exhibit cellular heterogeneity and metabolic plasticity across tissues in mice [9–11] and humans [12–14]. However, there is still a lack of systematic comparison of EC heterogeneity across combined contexts of different vessels and tissues.

It has been reported that the heterogeneity of ECs is regulated by various molecular factors, including transcription factors (TFs) and other signaling molecules [15]. Early in embryogenesis, transcriptional programs driven by TFs, such as GATA2 and ETV2 [16,17], establish the endothelial identity and guide early EC lineage commitment. The differentiated ECs are later refined based on the tissue context in embryonic development. Besides transcriptional regulation, the microenvironment regulates cellular heterogeneity and functions [18]. It is known that microenvironmental factors, including extracellular matrix composition, mechanical forces, and neighboring cell types, modulate gene expression regulation in ECs [19,20]. These regulations make them adapt to the functional needs of respective tissues [20,21]. For instance, ECs in the heart undergo unique shear stress and mechanical forces that upregulate KLF2, promoting an anti-inflammatory gene expression profile [22]. These studies together suggested that the tissue microenvironment is pivotal in establishing tissue-specific EC transcriptomes. Yet, the comprehensive investigation of microenvironmental mechanisms for human tissue-specific ECs remains elusive.

We hypothesized that tissue-specific microenvironments mediated by CCC may regulate human EC heterogeneity across tissue types. CCC is a fundamental process through which cells cross-talk via signaling molecules, such as cytokines and metabolites. The intact CCC network

coordinates cell functions, including differentiation, development, and metabolism [23]. For instance, the blood-brain barrier in the brain explicitly influences EC response to neuronal signaling and inflammatory cues [24]. A unique Car4-high EC population was identified in the lung primed to communicate with other alveolar cells during alveolar regeneration after lung injury [25].

This study integrated over 3 million single cells from human scRNA-seq datasets in 15 tissues. First, we systematically assessed the EC heterogeneity across tissues and vessel types. Intriguingly, a higher heterogeneity was observed across vessel types compared to tissue types in most tissues, except the heart, kidney, and intestine. Further, we identified tissue-specific EC topics for metabolite and protein-mediated CCC using a machine learning algorithm, Latent Dirichlet Allocation (LDA). Diving into specific CCC topics, we found that most identified metabolite-sensor and ligand-receptor pairs were tissue-specific and have been reported as vital in carrying out tissue-specific functions. Our work showed a systematic EC heterogeneity analysis across vessel and tissue types. Moreover, the identified tissue-specific CCC topics hold the potential to explain the EC heterogeneity regulated by tissue microenvironmental factors and further provide insights into therapeutic target discovery for endothelial-related diseases.

## Results

### Heterogeneity of EC across tissue and vessel types

We integrated scRNA-seq data from 15 human tissue types (**Supplementary Table 1**) and visualized the EC population using UMAP to investigate the human EC heterogeneity. During the integration, we observed that batch effects may confound the cell clustering as different projects dominate the separation into different clusters (**Supplementary Fig.1**). Exemplified by cells from the heart tissue, the cells in the same tissue type or vessel type from different projects were separated into different clusters (**Fig. 1A, B**). However, the ECs of the same tissue type from different projects were no longer clustered separately after correcting the batch effects (**Supplementary Fig.2**). Instead, ECs became clustered by tissue or EC vessel types (**Fig. 1C, D**). Thereafter, we performed all follow-up analyses based on batch-corrected data. Interestingly, the principal component analysis indicated that all ECs in the heart and kidney, regardless of vessel types, appeared to be very different from those in the other tissue types (**Fig. 1E**). However, most ECs tended to be grouped by vessel types rather than by tissue types, such as the ones grouped into arterial and lymphatic ECs (**Fig. 1E**). To further confirm this observation, we next quantified the EC similarity among tissue and vessel types using

97 unsupervised hierarchical clustering method. Two primary clusters (cluster 1 and cluster 2) were
98 observed, as shown by the largest cophenetic distance based on the height of the dendrogram
99 (**Fig. 1F**). In cluster 1, the distance between ECs was significantly greater within tissue type
100 than within vessel type, indicating that ECs were more similar within vessel type than within
101 tissue type. However, the ECs in cluster 2 showed a larger distance within a vessel type than
102 within a tissue type, meaning that ECs within a tissue type are more similar. These results
103 suggested that the human EC posed significant heterogeneity across vessel and tissue types.
104 Furthermore, some ECs were more heterogeneous across vessel types than tissue types, while
105 some ECs were more heterogeneous across vessel types.
106
107 **A computational workflow to unravel tissue-specific microenvironmental signaling**
108 While the EC population in cluster 1 was predominantly heterogeneous across vessel types,
109 they still showed substantial heterogeneity across tissue types within a vessel type. This raised
110 the question of how tissue-specific signals were associated with EC tissue heterogeneity.
111 Previous reports show that CCC can regulate tissue-specific EC functions [26]. To study tissue-
112 specific CCC patterns, we developed a computational workflow by integratively analyzing two
113 types of CCC events, namely, metabolite-mediated and protein-mediated CCC (**Fig.2A**). First,
114 cutting-edge bioinformatics tools were implemented to detect CCC events in scRNA-seq data of
115 each tissue. Specifically, MEBOCOST [27] was used for metabolite-mediated CCC detection,
116 while LIANA+ [28,29] was used to detect protein-mediated CCC events. Next, we applied topic
117 modeling, an unsupervised machine learning algorithm, to detect CCC topics among all CCC
118 events across tissues. Topic modeling was extensively used in natural language processing
119 tasks, such as summarizing documents based on groups of words. Here, we defined
120 “document” as tissue and “words” as a collection of CCC events. The output “topics” represent
121 CCC patterns that could be either housekeeping or tissue-specific across tissue types (**Fig. 2B**).
122 The model also outputs two matrices reflecting the probability of a CCC belonging to a topic and
123 the probability of a tissue including a topic, respectively (**Fig. 2C**). We identify tissue-specific
124 CCC topics based on the CCC topic probability vector across tissues. Using probability vectors
125 across CCC events, we further determined if any groups of CCC were highly concentrated on a
126 tissue-specific topic. Finally, the workflow further evaluates gene expression specificities of
127 CCC-related genes across tissue types.
128
129 **Great heterogeneity of EC-related CCC across tissue types**

130 We sought to explore the number of EC-associated CCC events across tissue types.
131 Interestingly, the number of CCC for EC varied across different tissues (**Fig. 3A**). Some tissues,
132 such as the ovary and intestine, showed a much higher number of CCC events than other
133 tissues, such as kidney and adipose. The observation was highly consistent between
134 metabolite-mediated and protein-mediated CCC, suggesting tissue specificity of EC
135 communication with resident cells. Since the number of CCC might be affected by EC
136 population size and tissue cell type number, we computed the EC percentage and total cell type
137 number across different tissues (**Fig. 3B**). Surprisingly, the EC percentage in a tissue showed
138 little correlation (R=-0.05) with the total CCC number (**Fig. 3C**), while the number of cell types
139 was strongly correlated with the number of CCC events in the tissue (R=0.76, **Fig. 3D**). This
140 was consistent with the prior knowledge that the diversity of cell types and their spatial
141 organization within a tissue contribute to the richness of CCC networks [23]. However, the
142 observation of varying CCC numbers across tissue types remained even when it was
143 normalized by the number of cell types in the tissue (**Fig. 3E**). Taken together, these results
144 showed the diverse capability of EC to receive CCC signaling in different tissues, indicating
145 tissue-specific CCC patterns.
146
147 **Topic modeling revealed tissue-specific metabolite-mediated CCC patterns**
148 Functional metabolite-sensing machinery facilitates the crosstalk between the cellular biological
149 network and its tissue microenvironment cues [30]. Recent studies have highlighted dynamic
150 mechanisms with metabolites acting as signaling mediators to modulate processes like
151 inflammation and cancer progression [31]. To study EC tissue-specific heterogeneity under the
152 context of metabolite-mediated CCC, we applied topic modeling using communications detected
153 in each tissue and identified 15 metabolite CCC topics (**Supplementary Table 2**). Of 15 topics,
154 9 topics were highly tissue-specific. For instance, topic 11 was highly correlated with thymus.
155 Topic 5 was liver-dominant based on the topic-tissue probability vectors (**Fig. 4A**). Other less
156 tissue-specific topics were also identified, such as topic 7, topic 2, and topic 10, suggesting
157 some shared CCC networks for ECs across different tissues. Investigating each metabolite-
158 sensor pair in each topic, we observed that they were highly correlated with the specificity of
159 CCC topics (**Fig.4B**). To confirm this specificity further, we checked the number of
160 communications inferred by MEBOCOST for each tissue and found a matched pattern. The
161 metabolite-sensor pairs that appeared more frequently in a topic also tend to show a higher
162 number of CCC in the tissue related to that topic (**Fig.4C**). For example, we identified topic 13 to
163 be highly associated with the ovary (**Fig.4A**). Five metabolite-sensor pairs, including the

164 Cholesterol (metabolite)~RORA (sensor), L-Lysine~SLC7A1, Ornithine~SLC7A1,
165 Adenine~SLC35F5, and L-Cysteine~SLC43A2, exclusively showed high probability in topic 13
166 (**Fig.4B**), and showed more CCC events in the ovary (**Fig. 4C**). To inspect the results at the
167 expression level, we further assessed the communication score, which was inferred by the
168 metabolite enzyme and sensor expression. We found that all these pairs were highly expressed
169 in the ovary (**Fig.4D**). Interestingly, L-Lysine~SLC7A1 and Ornithine~SLC7A1 were the two
170 pairs that showed the highest specificity as they nearly solely expressed in the ovary but not in
171 other tissue. Consistently, many previous studies have demonstrated the importance of the
172 selected metabolite and sensor in carrying out tissue-specific functions. For example, a study
173 found lysine succinylation (Ksuc) as a newly discovered post-translational modification (PTM)
174 that plays an essential role in ovarian aging [32]. The paper pointed out that in the ovaries of
175 aged mice and those with premature ovarian insufficiency (POI), increased levels of Ksuc were
176 observed. Further, SLC7A1 was reported to be a key player in the ovarian tumor
177 microenvironment, influencing both cancer cell behavior and stromal fibroblast activation [33].
178 You et al. also found that in ovarian cancer (EOC), SLC7A1 inhibited the expression of the
179 CCL4 that was related to the immune infiltration microenvironment, suggesting a great
180 candidate for EOC therapy [34]. Together, we identified highly tissue-specific CCC topics
181 through topic modeling. The identified CCC events pose the potential to regulate EC
182 heterogeneity across tissue types.
183
184 **Topic modeling revealed tissue-specific protein-mediated CCC patterns**
185 Numerous studies have shown that protein ligand-receptor interactions are crucial in regulating
186 cell function and cell fate [35,36]. By applying topic modeling on protein-mediated CCC
187 landscapes inferred by LIANA+ [28,29], we identified 15 CCC topics involving ECs
188 (**Supplementary Table 2**). Interestingly, 13 of 15 protein-CCC topics were uniquely associated
189 with one tissue type (**Fig. 5A**). For example, topic 13 was strongly enriched with the thymus,
190 while topic 15 was strongly associated with the heart. Visualizing the ligand-receptor pairs that
191 participated in those communication events, we spotted that the ligand-receptor pairs were
192 highly topic-specific (**Fig. 5B**), further suggesting the tissue-specific CCC pattern. By checking
193 the number of communications associated with each pair (**Fig. 5C**), many ligand-receptor pairs
194 indeed showed a much higher number of communications in one tissue compared to other
195 tissues, consistent with the patterns observed above (**Fig. 5B**). These results suggested that the
196 identified ligand-receptor pairs in topics were tissue-specific, which indicated their unique,
197 important roles in regulating tissue-specific EC functions. Here, we have dived deeper into topic

198 15, which was a topic highly related to the heart (**Fig. 5A**). We focused on six ligand-receptor
199 pairs based on their significantly high probability score (topic-related) and high communication
200 number (tissue-related) (**Fig. 5B-C**), respectively. Those ligand-receptor pairs were
201 ADAM17~ITGB1, CALM1~RYR2, ADAM17~ITGA5, LAMA2~RPSA, ADAM17~NOTCH1, and
202 MAML2~NOTCH1. Interestingly, those pairs' ligand and receptor gene expression were also
203 highly expressed in the heart compared to other tissues (**Fig. 5D**). These identified ligands and
204 receptors have been reported to have important functions in the heart with in vivo and in vitro
205 experiments. For example, ADAM17 was essential to normal cardiovascular functions [37].
206 Elevated ADAM17 levels have been found in the lipid core of human atherosclerotic plaques
207 and mouse models with atherosclerotic lesions progression [38]. ADAM17 also plays a crucial
208 role in cardiac inflammation by cleaving and activating various proinflammatory cytokines and
209 their receptors [39]. It shows that ADAM17 is an important ligand that regulates normal heart
210 functions. Moreover, NOTCH1, which interacts with ADAM17 and MAML2 in the heart, was also
211 a crucial receptor in tissue formation and homeostasis. Ye et al. have revealed that the
212 NOTCH1 deficiency could cause defects in the proliferation of early human cardiomyocytes [40].
213 Roifman et al. found reduced NOTCH1 levels in both ventricular outflow tracts, suggesting that
214 NOTCH1 haploinsufficiency may cause bicuspid aortic valve and related cardiac anomalies [41].
215 A recent clinical study also showed an association between the NOTCH1 variants and cardiac
216 and extracardiac vascular anomalies [42]. This analysis provided insights into tissue-specific
217 communication patterns to EC for the ligand and receptor pairs and further poses the potential
218 to explain how EC heterogeneity was regulated at the tissue niche level.
219
220 **Discussion**
221 Tissue-specific EC functions have been widely studied within individual tissues and compared
222 between vessel types. It is unclear how human EC heterogeneity across different tissue types
223 was regulated systematically. An intriguing question is whether ECs should continue to be
224 grouped by vessel types or should they be further grouped by tissue types.
225
226 In this study, we investigated the microenvironments led by CCC as a potential factor to
227 regulate EC differences across vessel and tissue types. Analyzing human EC scRNA-seq data
228 specifically, we first identified different patterns of EC heterogeneity, which were shown by
229 tissue type and vessel type separately. Aside from the well-known knowledge that the EC
230 vessel types differed due to their specific locations and functions, ECs in some tissues, e.g., the
231 heart and kidney, were more different under the tissue context than vessel types. This finding

was comparable with other studies. Chen et al. showed that the embryonic EC heterogeneity came from tissue origin and vascular-bed position [43]. They also found that embryonic ECs in some tissues, such as the heart and liver, expressed tissue-selective markers.

In contrast, ECs in other tissues were more similar within each vessel type. Other groups also identified EC heterogeneity, as shown by tissue-specific marker genes and enriched functions, but with the mouse model [9,10]. In this study, we dived deeper into the microenvironment regulation by CCC and found tissue-specific CCC patterns for EC in humans. We investigated EC communications under various tissue contexts using an unsupervised machine-learning algorithm and topic modeling. For metabolite-mediated and protein-mediated communications, many CCC events inferred *in silico* were exclusively strong in one tissue compared to other tissues. Based on the previous reports, we further verified that the identified ligands and receptors were associated with essential functions in specific tissues. In summary, the identified CCC events in each topic were promising candidates for further investigation to provide insights into EC tissue-specific differences regulated by metabolite-mediated and protein-mediated CCC.

## Method and materials

### Data acquisition and process

We downloaded over 3,000,000 single cells of 15 human tissues from the DISCO database[44], including adipose, bladder, breast, gut, heart, intestine, kidney, liver, lung, ovary, skeletal muscle, skin, stomach, testis, and thymus. The downloaded data in each tissue includes raw gene expression counts and cell metadata annotation, such as cell type, project identification number, disease status, etc. First, we filtered out disease samples to focus on data from normal tissues, which resulted in over 2,000,000 single cells of 15 tissue types. The pre-processing, including normalization, was done in Python by Scanpy (version 1.8.2) based on the raw count matrix. The normalized data was used for the downstream analysis in this study, including EC heterogeneity analysis and cell-cell communication analysis. In the EC heterogeneity analysis, we retrieved single cells with cell types labeled as four major EC vessel types, namely, arterial, capillary, lymphatic, and venous EC. We finally removed the batch effect using the Harmony algorithm wrapped in the Scanpy package.

### Dimensionality reduction and clustering

We did uniform manifold approximation and projection (UMAP), reducing the dimensionality to show the clustering patterns across EC vessels and tissue types. We further investigated the difference between EC vessel type versus tissue type by performing principal component (PC) analysis using the Harmony-corrected PCs (50 in total). Results were visualized using PC1 and PC2, which explained the most variance. Finally, hierarchical clustering based on tissue type and vessel type was done with Scipy using 50PCs; the ward variance minimization algorithm and Euclidean metric were used to build the model. Pairwise distance (Euclidean) was calculated using the scikit-learn package for two major clusters for further quantifications.

**Cell-cell communication analysis**

To increase the computation efficiency, we down-sampled data from each tissue to contain only 10,000 cells and did normalization. Then, we utilized two bioinformatics tools to infer cell-cell communications. MEBOCOST [27] predicted metabolite-mediated CCC events. LIANA+ [28,29] was used to predict protein-mediated CCC. Each inference was performed on cell types within each tissue type. Default parameters were used when running the framework. The input data contained all cell types. We focused on communication events with a false discovery rate (FDR) of less than 0.05. Then, we restricted the predicted communications with the type of sender/source or receiver/target as ECs to further focus on EC-related CCC for the downstream analysis.

**Topic modeling**

Focusing on how EC was regulated, we further filtered our data to only allow EC as the receiver/target cells in a single communication event. Then, topic modeling was performed with Latent Dirichlet Allocation (LDA) in R. The input matrix was tissue type by communication events, filling with the number of communication events predicted in each tissue. The number of topics in the LDA analysis was set at 15 based on the assumption that each tissue formed its topic. The top 10 communication pairs (with the highest beta score from topic modeling) were selected to visualize topic-specific patterns due to their strong probability in each topic. The topic's word distribution, beta $\beta$, and communication number for each metabolite-sensor/protein-receptor pair were normalized for each topic and tissue, respectively. To validate the tissue-specific communication events, the communication score (product of the average metabolite and sensor expression) from the MEBOCOST and the average expression of ligand-receptor pairs from the CellPhoneDB were used, respectively.

## Quantification and statistical analysis

Statistical analyses were performed using Python 3.10. The Euclidean distance matrix was used to calculate the correlation. The data were presented as mean in bar plots, and the box plots displayed quantiles and median.

## Code and data availability

All the analyses were done using Python (version 3.8 for raw data processing and 3.10 for downstream analysis), except for the LDA analysis, which was performed using R (version 4.4.0). The script to reproduce all analyses in this study can be found at https://github.com/zhuzimoo/EC_project.

## Acknowledgments

During the period of this project, K. C. is supported in part by the grants R01GM125632, R01GM138407, R01HL148338, R01HL133254, R01CA278832, and R01HL174928 to K. C., L. Z. is supported in part by the grants R01HL155632 to L. Z. and Additional Ventures (Single Ventricle Research Foundation) to L. Z, R. Z is supported in part by the grants K99HG013662 to R. Z. and NextGen Accelerator Award (Cell Discovery Network).

## Author Contributions

K. C. and R. Z. conceived the project. Z. Z. performed all the analyses in this study and drafted the manuscript. All other co-authors read and revised the manuscript.

## References


[1] Krüger-Genge A, Blocki A, Franke R-P, Jung F. Vascular endothelial cell biology: An update. Int J Mol Sci 2019;20:4411.

[2] Pi X, Xie L, Patterson C. Emerging roles of vascular endothelium in metabolic homeostasis. Circ Res 2018;123:477–94.

[3] Jalkanen S, Salmi M. Lymphatic endothelial cells of the lymph node. Nat Rev Immunol 2020;20:566–78.

[4] Iglesias MJ, Kruse LD, Sanchez-Rivera L, Enge L, Dusart P, Hong M-G, et al. Identification of endothelial proteins in plasma associated with cardiovascular risk factors. Arterioscler Thromb Vasc Biol 2021;41:2990–3004.

[5] Aird WC. Phenotypic heterogeneity of the endothelium: I. Structure, function, and mechanisms. Circ Res 2007;100:158–73.

[6] Shin D, Garcia-Cardena G, Hayashi S, Gerety S, Asahara T, Stavrakis G, et al. Expression of ephrinB2 identifies a stable genetic difference between arterial and venous vascular smooth muscle as well as endothelial cells, and marks subsets of microvessels at sites of adult neovascularization. Dev Biol 2001;230:139–50.

[7] Jadlowiec CC, Feigel A, Yang C, Feinstein AJ, Kim ST, Collins MJ, et al. Reduced adult endothelial cell EphB4 function promotes venous remodeling. Am J Physiol Cell Physiol 2013;304:C627–35.

[8] Hellmut G A, Gou Young K. Organotypic vasculature: From descriptive heterogeneity to functional pathophysiology. Science 2017;357. https://doi.org/10.1126/science.aal2379.

[9] Kalucka J, de Rooij LPMH, Goveia J, Rohlenova K, Dumas SJ, Meta E, et al. Single-cell transcriptome atlas of Murine endothelial cells. Cell 2020;180:764–79.e20.

[10] Paik DT, Tian L, Williams IM, Rhee S, Zhang H, Liu C, et al. Single-cell RNA sequencing unveils unique transcriptomic signatures of organ-specific endothelial cells. Circulation 2020;142:1848–62.

[11] Bondareva O, Rodríguez-Aguilera JR, Oliveira F, Liao L, Rose A, Gupta A, et al. Single-cell profiling of vascular endothelial cells reveals progressive organ-specific vulnerabilities during obesity. Nat Metab 2022;4:1591–610.

[12] Schupp JC, Adams TS, Cosme C Jr, Raredon MSB, Yuan Y, Omote N, et al. Integrated single-cell atlas of endothelial cells of the human lung. Circulation 2021;144:286–302.

[13] Li Q, Zhu Z, Wang L, Lin Y, Fang H, Lei J, et al. Single-cell transcriptome profiling reveals vascular endothelial cell heterogeneity in human skin. Theranostics 2021;11:6461–76.

354 [14] Geldhof V, de Rooij LPMH, Sokol L, Amersfoort J, De Schepper M, Rohlenova K, et al.
355 Single cell atlas identifies lipid-processing and immunomodulatory endothelial cells in
356 healthy and malignant breast. Nat Commun 2022;13:5511.
357 [15] De Val S, Black BL. Transcriptional control of endothelial cell development. Dev Cell
358 2009;16:180–95.
359 [16] Lugus JJ, Chung YS, Mills JC, Kim S-I, Grass J, Kyba M, et al. GATA2 functions at multiple
360 steps in hemangioblast development and differentiation. Development 2007;134:393–405.
361 [17] Shi X, Richard J, Zirbes KM, Gong W, Lin G, Kyba M, et al. Cooperative interaction of Etv2
362 and Gata2 regulates the development of endothelial and hematopoietic lineages. Dev Biol
363 2014;389:208–18.
364 [18] Potente M, Mäkinen T. Vascular heterogeneity and specialization in development and
365 disease. Nat Rev Mol Cell Biol 2017;18:477–94.
366 [19] Majewska A, Wilkus K, Brodaczewska K, Kieda C. Endothelial cells as tools to model tissue
367 microenvironment in hypoxia-dependent pathologies. Int J Mol Sci 2021;22:520.
368 [20] Leone P, Malerba E, Susca N, Favoino E, Perosa F, Brunori G, et al. Endothelial cells in
369 tumor microenvironment: insights and perspectives. Front Immunol 2024;15:1367875.
370 [21] Aird WC, Edelberg JM, Weiler-Guettler H, Simmons WW, Smith TW, Rosenberg RD.
371 Vascular bed-specific expression of an endothelial cell gene is programmed by the tissue
372 microenvironment. J Cell Biol 1997;138:1117–24.
373 [22] Dekker RJ, van Thienen JV, Rohlena J, de Jager SC, Elderkamp YW, Seppen J, et al.
374 Endothelial KLF2 links local arterial shear stress levels to the expression of vascular tone-
375 regulating genes. Am J Pathol 2005;167:609–18.
376 [23] Su J, Song Y, Zhu Z, Huang X, Fan J, Qiao J, et al. Cell-cell communication: new insights
377 and clinical implications. Signal Transduct Target Ther 2024;9:196.
378 [24] Armulik A, Genové G, Mäe M, Nisancioglu MH, Wallgard E, Niaudet C, et al. Pericytes
379 regulate the blood-brain barrier. Nature 2010;468:557–61.
380 [25] Niethamer TK, Stabler CT, Leach JP, Zepp JA, Morley MP, Babu A, et al. Defining the role
381 of pulmonary endothelial cell heterogeneity in the response to acute lung injury. Elife
382 2020;9. https://doi.org/10.7554/eLife.53072.
383 [26] Amersfoort J, Eelen G, Carmeliet P. Immunomodulation by endothelial cells - partnering up
384 with the immune system? Nat Rev Immunol 2022;22:576–88.
385 [27] Zheng R, Zhang Y, Tsuji T, Gao X, Wagner A, Yosef N, et al. MEBOCOST: Metabolite-
386 mediated Cell Communication Modeling by Single Cell Transcriptome. Bioinformatics 2022.
387 [28] Dimitrov D, Türei D, Garrido-Rodriguez M, Burmedi PL, Nagai JS, Boys C, et al.

388 Comparison of methods and resources for cell-cell communication inference from single-
389 cell RNA-Seq data. Nat Commun 2022;13:3224.
390 [29] Dimitrov D, Schäfer PSL, Farr E, Rodriguez-Mier P, Lobentanzer S, Badia-I-Mompel P, et
391 al. LIANA+ provides an all-in-one framework for cell-cell communication inference. Nat Cell
392 Biol 2024;26:1613–22.
393 [30] Wang Y-P, Lei Q-Y. Metabolite sensing and signaling in cell metabolism. Signal Transduct
394 Target Ther 2018;3:30.
395 [31] Wang Y-P, Li J-T, Qu J, Yin M, Lei Q-Y. Metabolite sensing and signaling in cancer. J Biol
396 Chem 2020;295:11938–46.
397 [32] Le M, Li J, Zhang D, Yuan Y, Zhou C, He J, et al. The emerging role of lysine succinylation
398 in ovarian aging. Reprod Biol Endocrinol 2023;21:38.
399 [33] You S, Han X, Xu Y, Sui L, Song K, Yao Q. High expression of SLC7A1 in high-grade
400 serous ovarian cancer promotes tumor progression and is involved in MAPK/ERK pathway
401 and EMT. Cancer Med 2024;13:e7217.
402 [34] You S, Zhu X, Yang Y, Du X, Song K, Zheng Q, et al. SLC7A1 Overexpression is involved
403 in energy metabolism reprogramming to induce tumor progression in epithelial ovarian
404 cancer and is associated with immune-infiltrating cells. J Oncol 2022;2022:5864826.
405 [35] An C, Wang X, Song F, Hu J, Li L. Insights into intercellular receptor-ligand binding kinetics
406 in cell communication. Front Bioeng Biotechnol 2022;10:953353.
407 [36] Funk MA, Leitner J, Gerner MC, Hammerler JM, Salzer B, Lehner M, et al. Interrogating
408 ligand-receptor interactions using highly sensitive cellular biosensors. Nat Commun
409 2023;14:7804.
410 [37] Xu J, Mukerjee S, Silva-Alves CRA, Carvalho-Galvão A, Cruz JC, Balarini CM, et al. A
411 disintegrin and metalloprotease 17 in the cardiovascular and central nervous systems.
412 Front Physiol 2016;7:469.
413 [38] Oksala N, Levula M, Airla N, Pelto-Huikko M, Ortiz RM, Järvinen O, et al. ADAM-9, ADAM-
414 15, and ADAM-17 are upregulated in macrophages in advanced human atherosclerotic
415 plaques in aorta and carotid and femoral arteries--Tampere vascular study. Ann Med
416 2009;41:279–90.
417 [39] Adu-Amankwaah J, Adzika GK, Adekunle AO, Ndzie Noah ML, Mprah R, Bushi A, et al.
418 ADAM17, A key player of cardiac inflammation and fibrosis in heart failure development
419 during chronic catecholamine stress. Front Cell Dev Biol 2021;9:732952.
420 [40] Ye S, Wang C, Xu Z, Lin H, Wan X, Yu Y, et al. Impaired human cardiac cell development
421 due to NOTCH1 deficiency. Circ Res 2023;132:187–204.

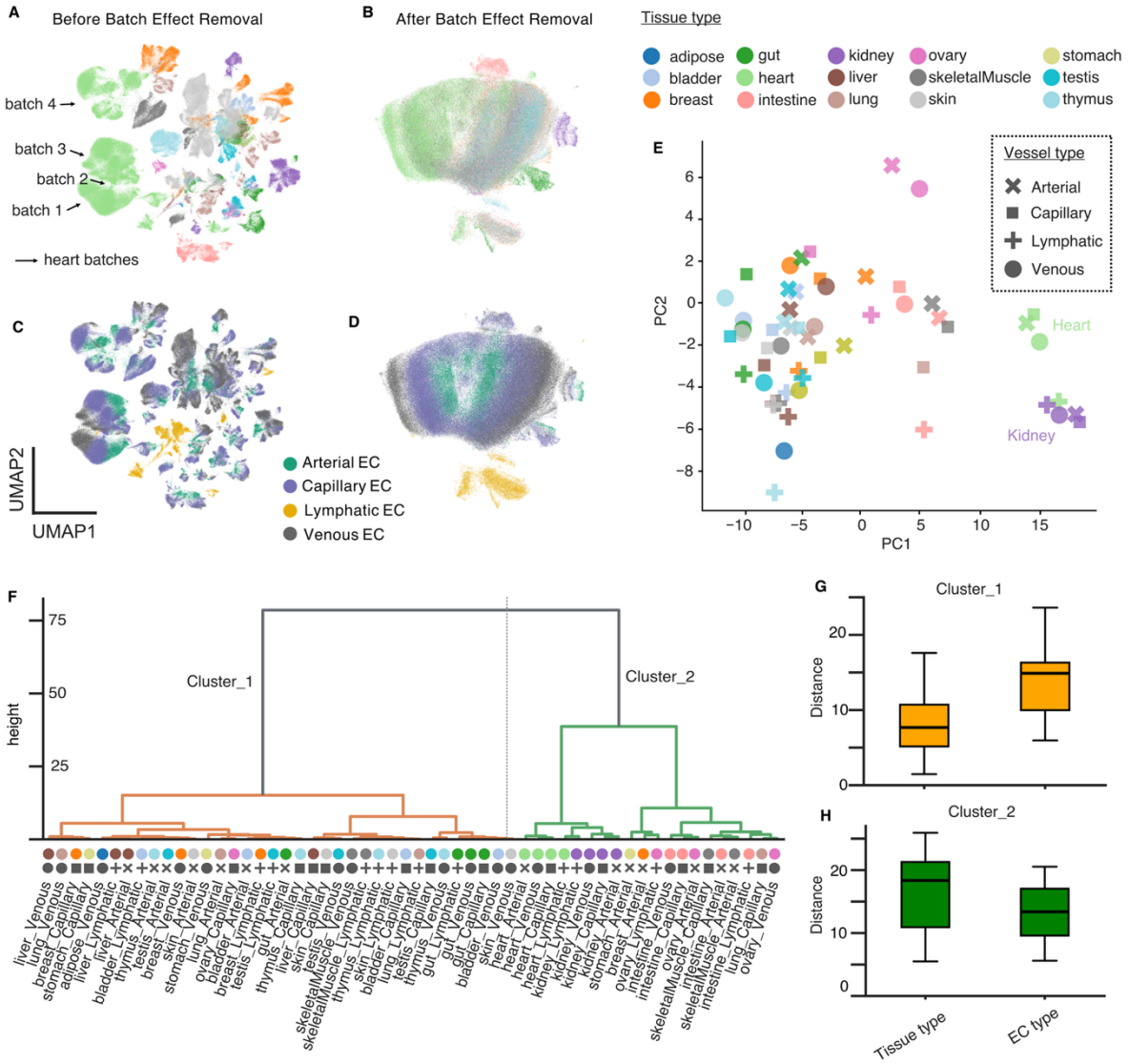
[41] Roifman M, Chung BHY, Reid DM, Teitelbaum R, Martin N, Nield LE, et al. Heterozygous NOTCH1 deletion associated with variable congenital heart defects. Clin Genet 2021;99:836–41.

[42] Stanley KJ, Kalbfleisch KJ, Moran OM, Chaturvedi RR, Roifman M, Chen X, et al. Expanding the phenotypic spectrum of NOTCH1 variants: clinical manifestations in families with congenital heart disease. Eur J Hum Genet 2024;32:795–803.

[43] Chen J, Zhang X, DeLaughter DM, Trembley MA, Saifee S, Xiao F, et al. Molecular and spatial signatures of mouse embryonic endothelial cells at single-cell resolution. Circ Res 2024;134:529–46.

[44] Li M, Zhang X, Ang KS, Ling J, Sethi R, Lee NYS, et al. DISCO: a database of Deeply Integrated human Single-Cell Omics data. Nucleic Acids Res 2022;50:D596–602.

434 **Figure 1**




435

436 **Figure 1: EC demonstrated heterogeneity across tissue and vessel types.**

437 (**A-D**) UMAP clustering visualization of EC across tissue and vessel types before (**A**, **B**) and

438 after (**C, D**) batch effect removal, with color-coded for tissue type (**A, C**) and vessel types (**B, D**).

439 (**E**) PCA projection of EC across tissue type and vessel type. Color-coded for tissue type

440 (consistent with **1A** and **1B**) and shape-indicated EC vessel type.

441 (**F**) Hierarchical clustering showing tissue type and vessel type on the x-axis and cophenetic

442 distance on the y-axis. The dotted line separated the two major clusters. Colors and shapes

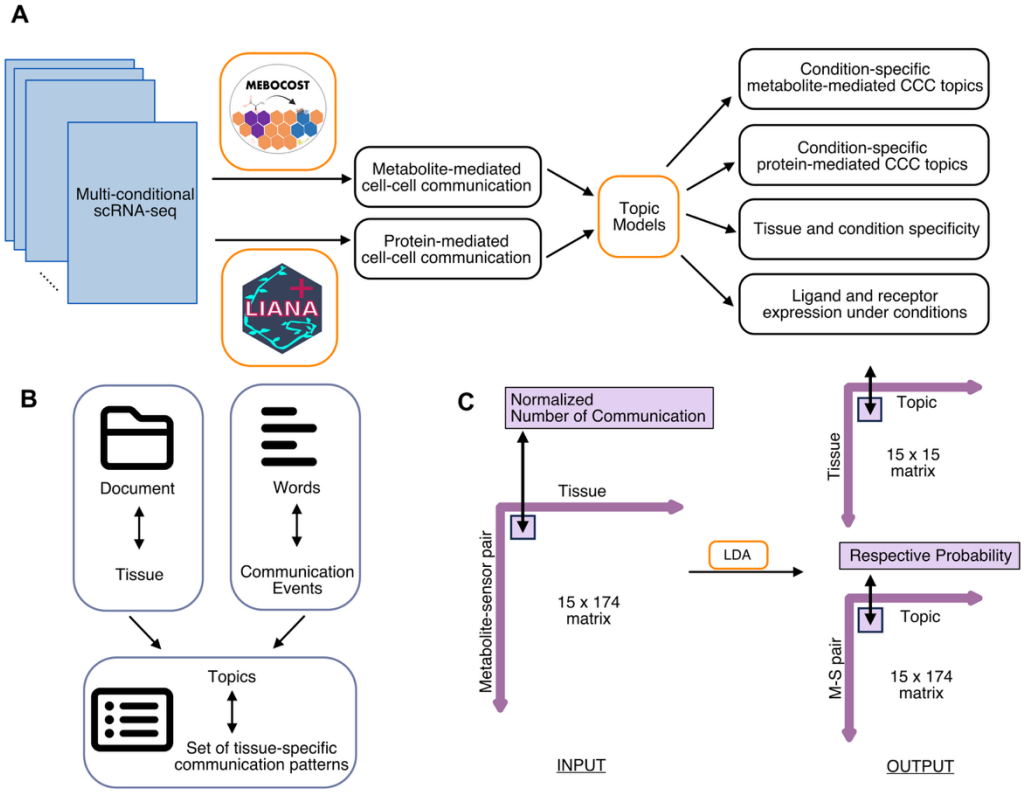
443 were used according to the PCA project in **E**.

444 (**G-H**) Pairwise distances were calculated for cluster 1 on the top (**G)** and cluster 2 on the

445 bottom (**H**). The box on the left shows the results grouping by tissue type and calculating the

446 vessel type distances, and the box on the right shows the results grouping by vessel types and
447 calculating the tissue distances. The median was shown with the bold line in the middle.
448

449 **Figure 2**




450

451 **Figure 2: Workflow for processing and analyzing multi-conditional scRNA-seq data**

452 (**A**) Scheme of analysis design.

453 (**B**) Explanation of topic modeling under the context of this study.

454 (**C**) Input and output matrix visualizations for the LDA algorithm.

455

456

**Figure 3**

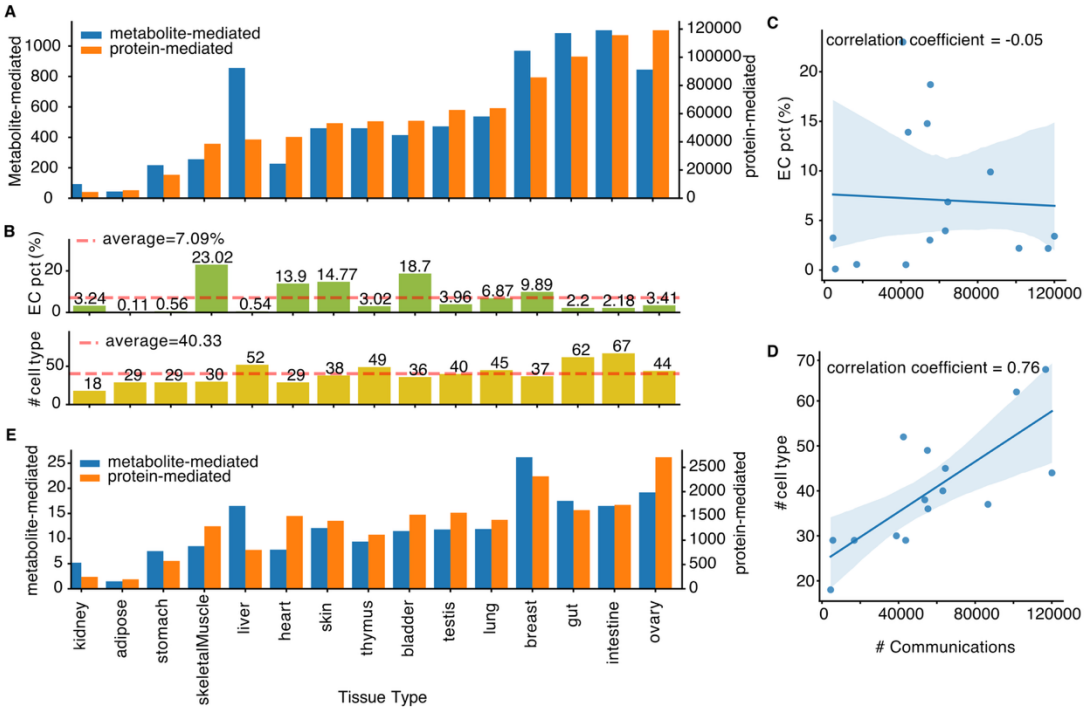



**Figure 3: Analysis of metabolite-sensor and ligand-receptor CCC.**

(**A**) A bar chart shows the number of communications inferred *in silico*. The blue color indicates the number of metabolite-mediated CCC (refers to the left y-axis scale). The orange color indicates the number of ligand-receptor CCC (refers to the right y-axis scale).

(**B**) Bar chart showing the percentage of EC in the whole cell population and the number of cell types included in a tissue. The red-dotted line indicates the average value under each category.

(**C**) The scatter plot shows the number of communications against EC percentage. Each dot represents a tissue type. The regression line is shown in blue.

(**D**) The scatter plot shows the number of communications against the number of cell types. Each dot represents a tissue type. The regression line is shown in blue.

(**E**) The bar chart shows the number of CCC events of every hundred cell types in each tissue to normalize the effect of total cell type numbers across tissues.

**Figure 4**

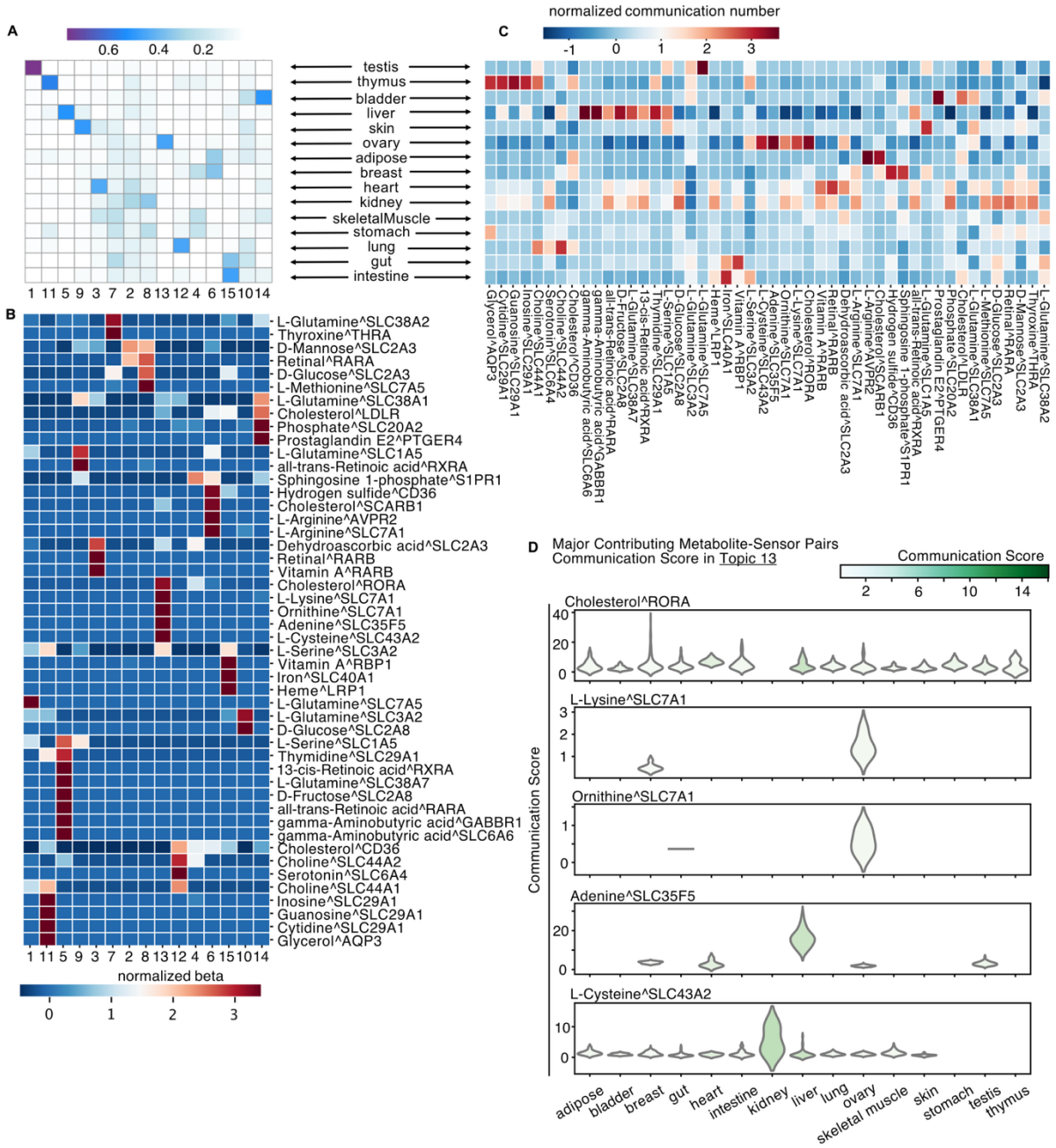



**Figure 4: LDA inferred topics of metabolite-mediated CCC networks**

(**A**) Clustered heatmap showing the correlation of the LDA results using MEBOCOST inferred CCC results. The X-axis indicated inferred topics and the y-axis labeled tissue type. The darker the color, the stronger the correlation between the corresponding topic and tissue.

(**B**) Clustered heatmap of the normalized beta score of selected metabolite-sensor pairs (y-axis) in each topic (x-axis). The x-axis label was aligned with that in **4A**. The warmer the color, the higher the specificity, and vice versa.

(**C**) Clustered heatmap of the normalized number of communications for tissue-specific metabolite-sensor pairs (x-axis). Tissue labels (y-axis) were aligned with that in **4A**. The warmer the color, the higher the communication number, and vice versa.

(**D**) Violin plots of highly expressed metabolite-sensor pairs in topic 13. The color represented the expression level based on the communication score. A missing violin indicated the metabolite-sensor pair was not expressed in that tissue.

**Figure 5**

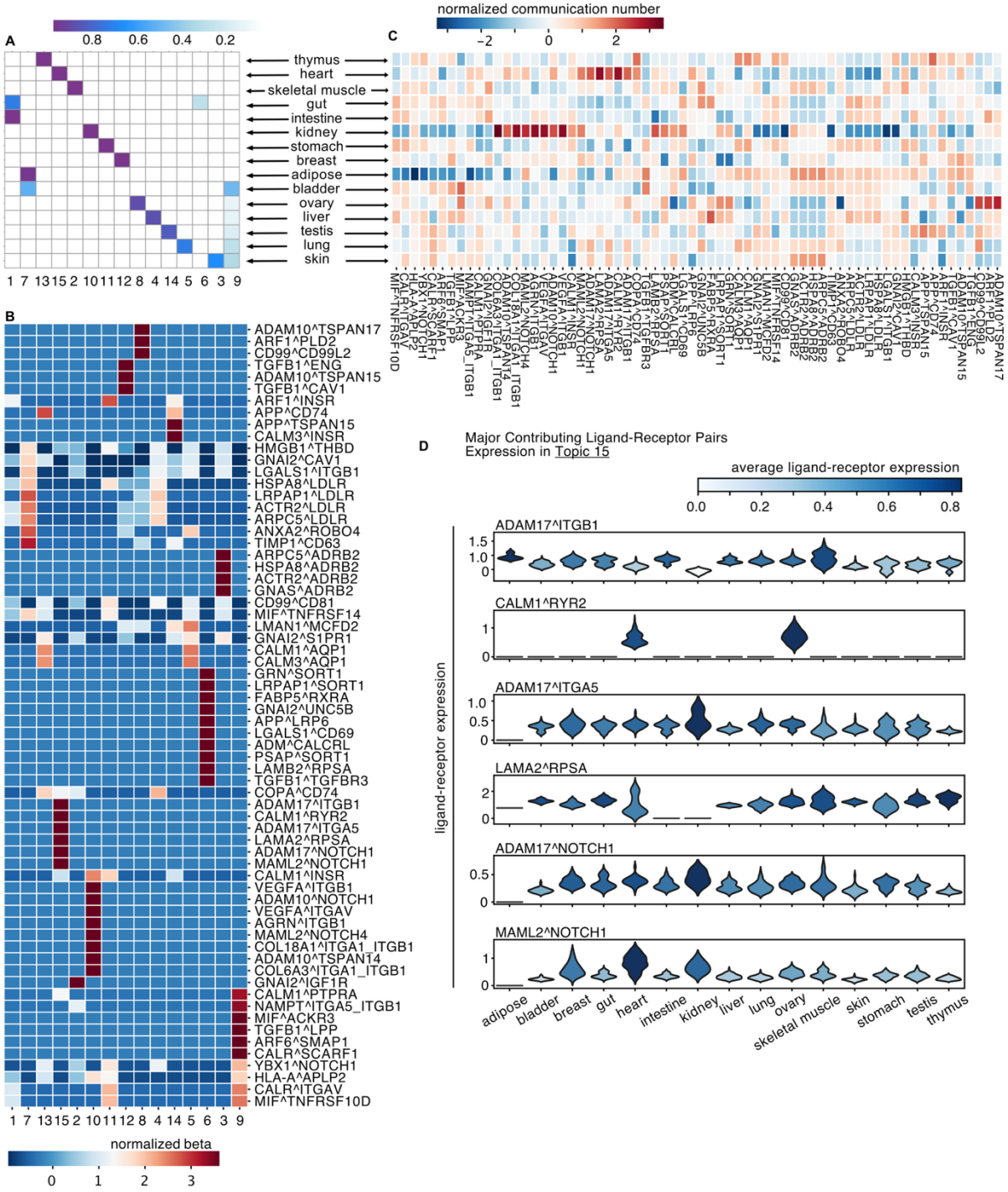



**Figure 5: LDA inferred topics of ligand-receptor CCC networks**

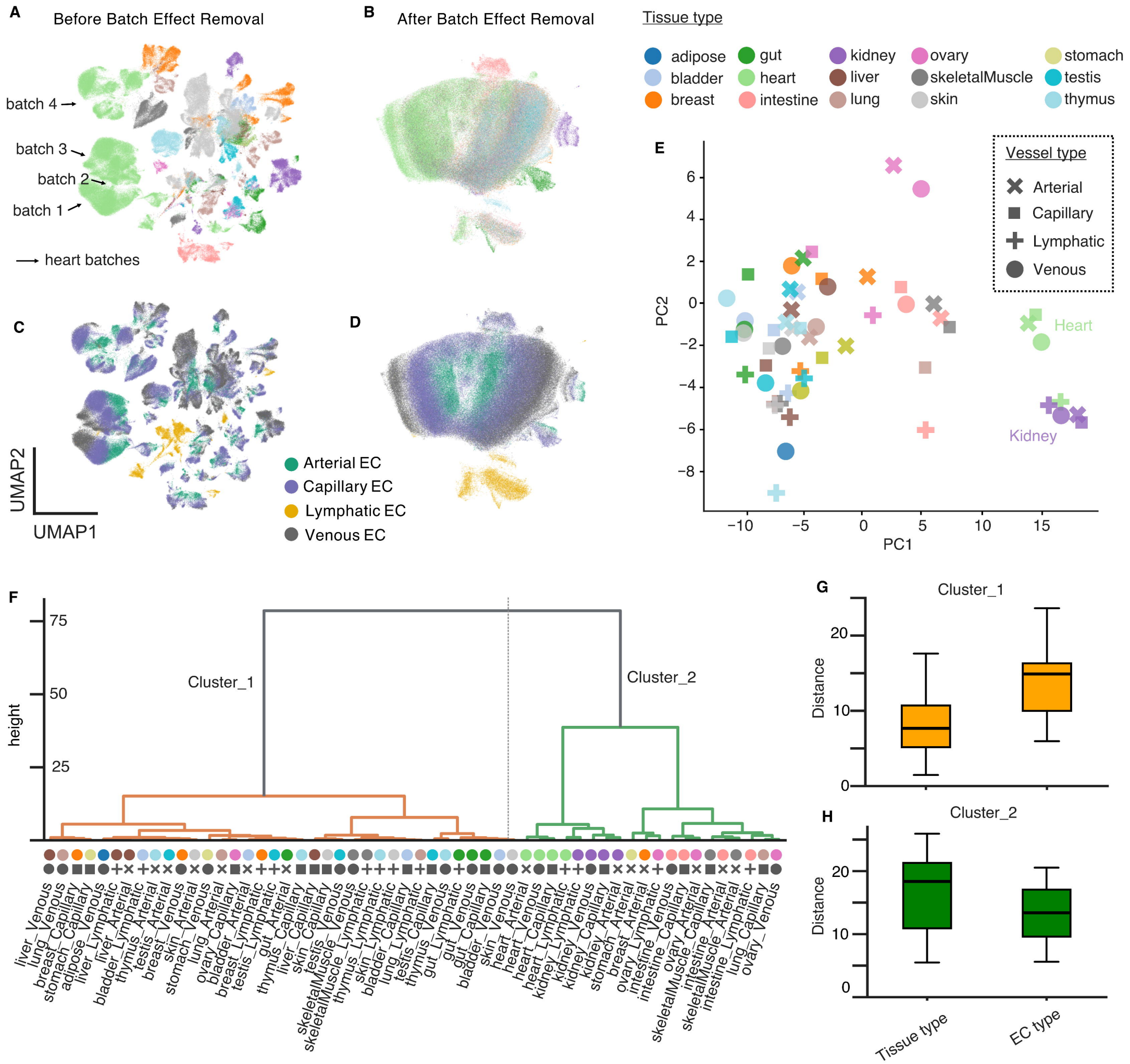
(**A**) Clustered heatmap showing the correlation of the LDA results using LIANA+ inferred CCC results. The X-axis indicated inferred topics and the y-axis labeled tissue type. The darker the color, the stronger the correlation between the corresponding topic and tissue.

(**B**) Clustered heatmap of the normalized beta score of selected ligand-receptor pairs (y-axis) in each topic (x-axis). The x-axis label was aligned with that in **5A**. The warmer the color, the higher the specificity, and vice versa.

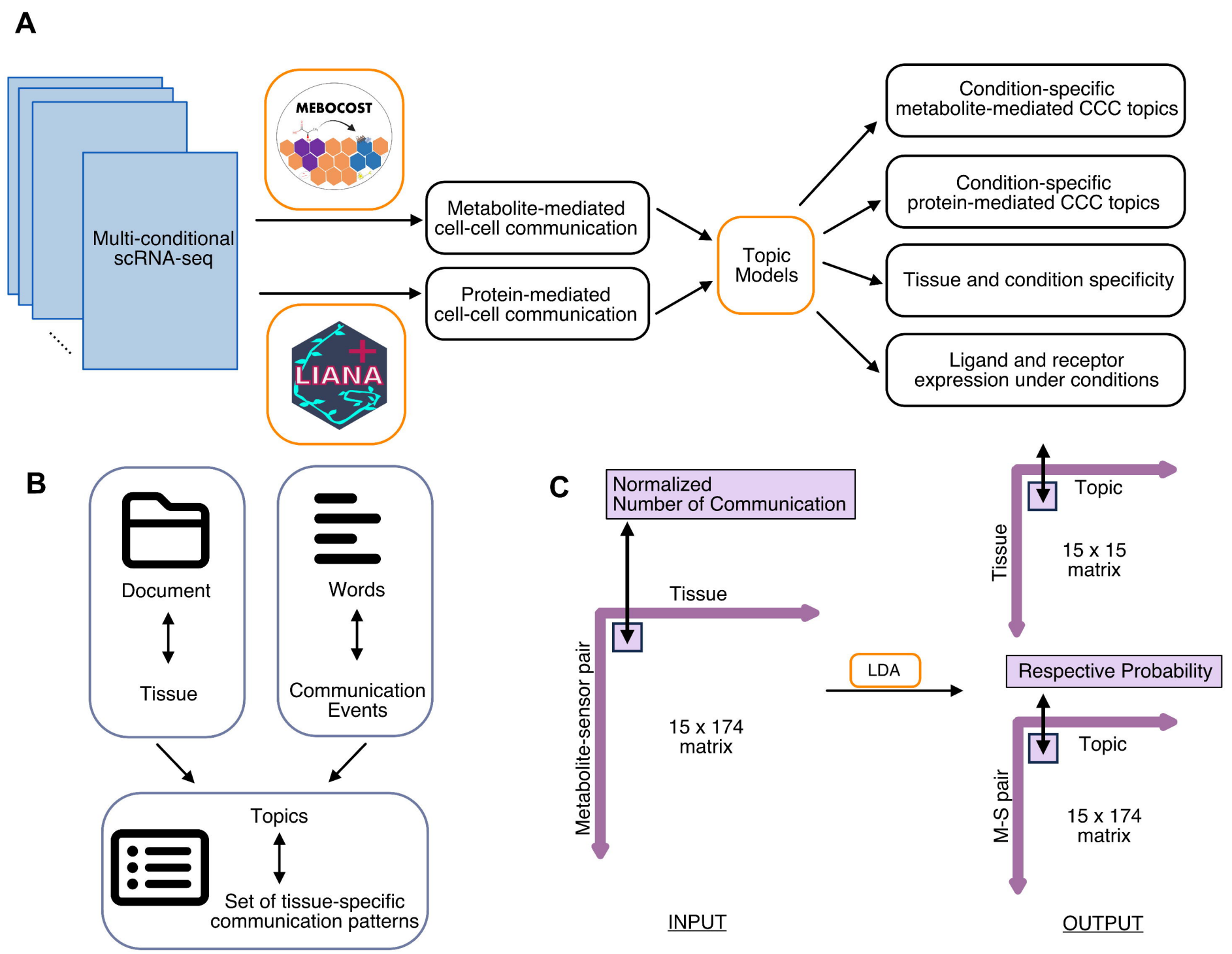
(**C**) Clustered heatmap of the normalized number of communications for tissue-specific ligand-receptor pairs (x-axis). Tissue labels (y-axis) were aligned with that in **5A**. The warmer the color, the higher the communication number, and vice versa.

(**D**) Violin plots of highly expressed ligand-receptor pairs in topic 15. The color represented the average of the pair expression level. A missing violin indicated the ligand-receptor pair was not expressed in that tissue.

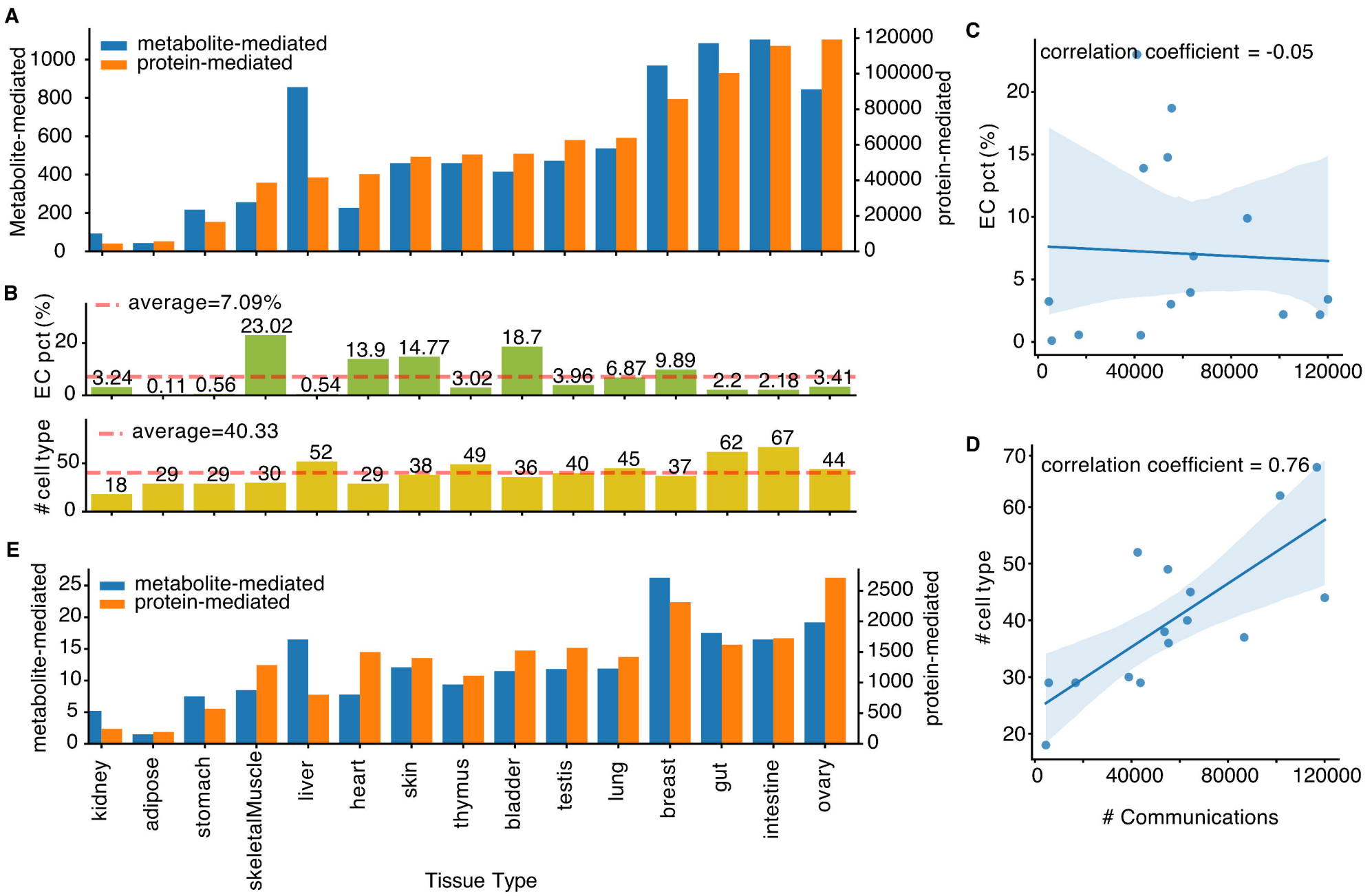
# Fig.1


A
Before Batch Effect Removal
batch 4
batch 3
batch 2
batch 1
heart batches
B
After Batch Effect Removal
Tissue type
adipose
bladder
breast
gut
heart
intestine
kidney
liver
lung
ovary
skeletalMuscle
skin
stomach
testis
thymus
C
D
UMAP2
UMAP1
Arterial EC
Capillary EC
Lymphatic EC
Venous EC
E
Vessel type
Arterial
Capillary
Lymphatic
Venous
Heart
Kidney
PC2
PC1
F
height
Cluster_1
Cluster_2
G
Cluster_1
Distance
H
Cluster_2
Distance
Tissue type
EC type

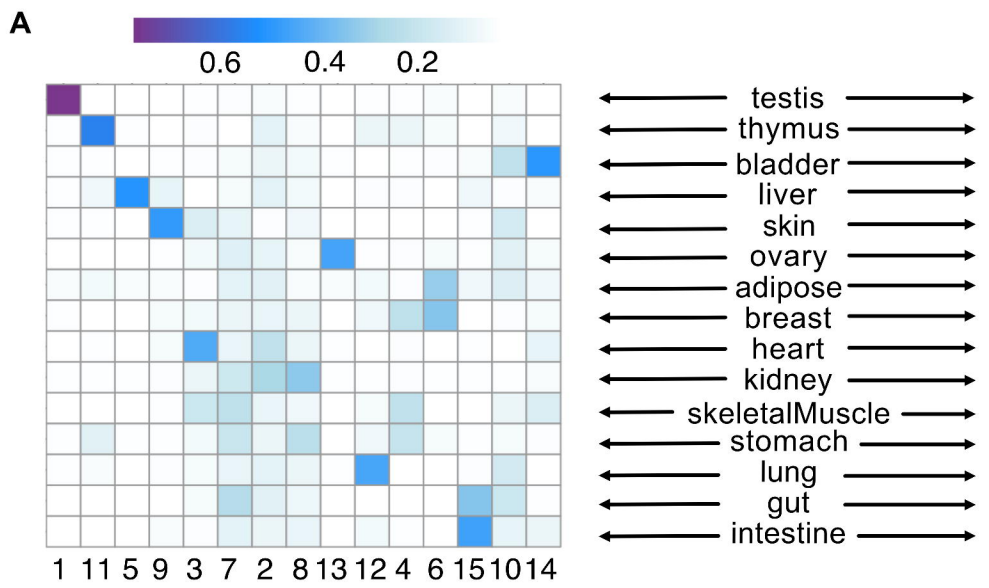
# Fig.2

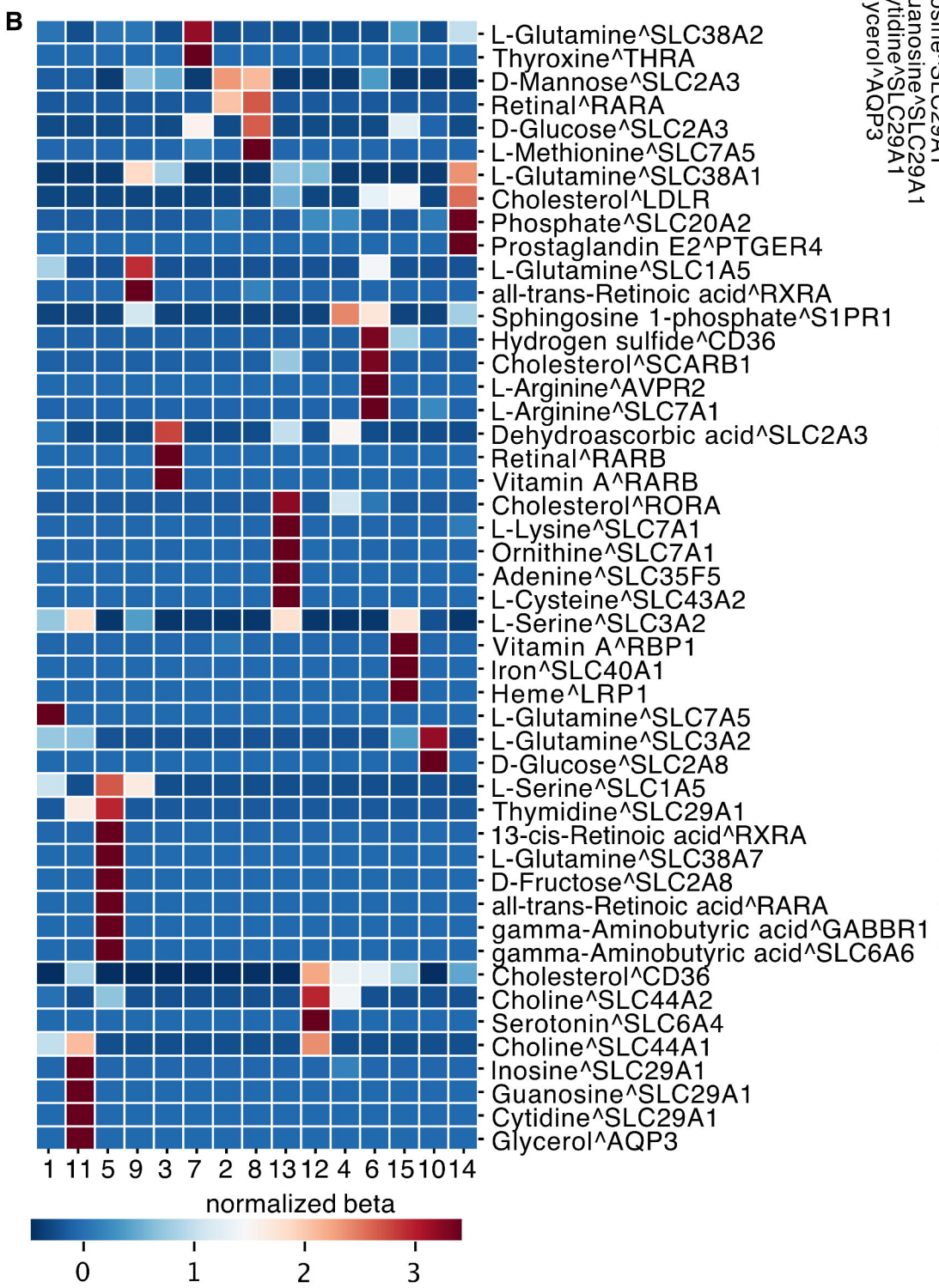

A
Multi-conditional scRNA-seq
MEBOCOST
LIANA
Metabolite-mediated cell-cell communication
Protein-mediated cell-cell communication
Topic Models
Condition-specific metabolite-mediated CCC topics
Condition-specific protein-mediated CCC topics
Tissue and condition specificity
Ligand and receptor expression under conditions
B
Document
Tissue
Words
Communication Events
Topics
Set of tissue-specific communication patterns
C
Normalized Number of Communication
Tissue
Metabolite-sensor pair
15 x 174 matrix
INPUT
LDA
Respective Probability
Topic
Tissue
15 x 15 matrix
Topic
M-S pair
15 x 174 matrix
OUTPUT

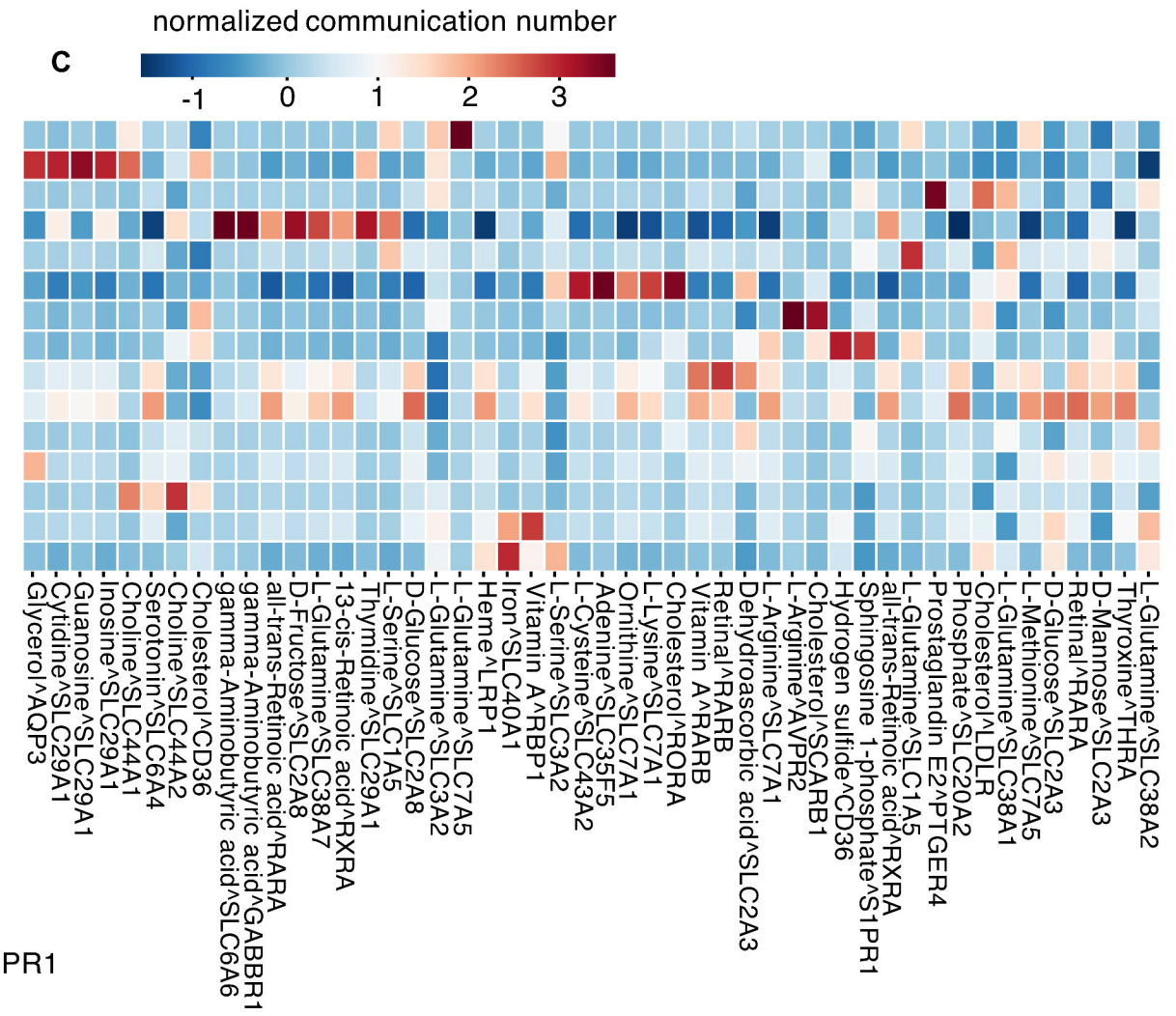
# Fig.3

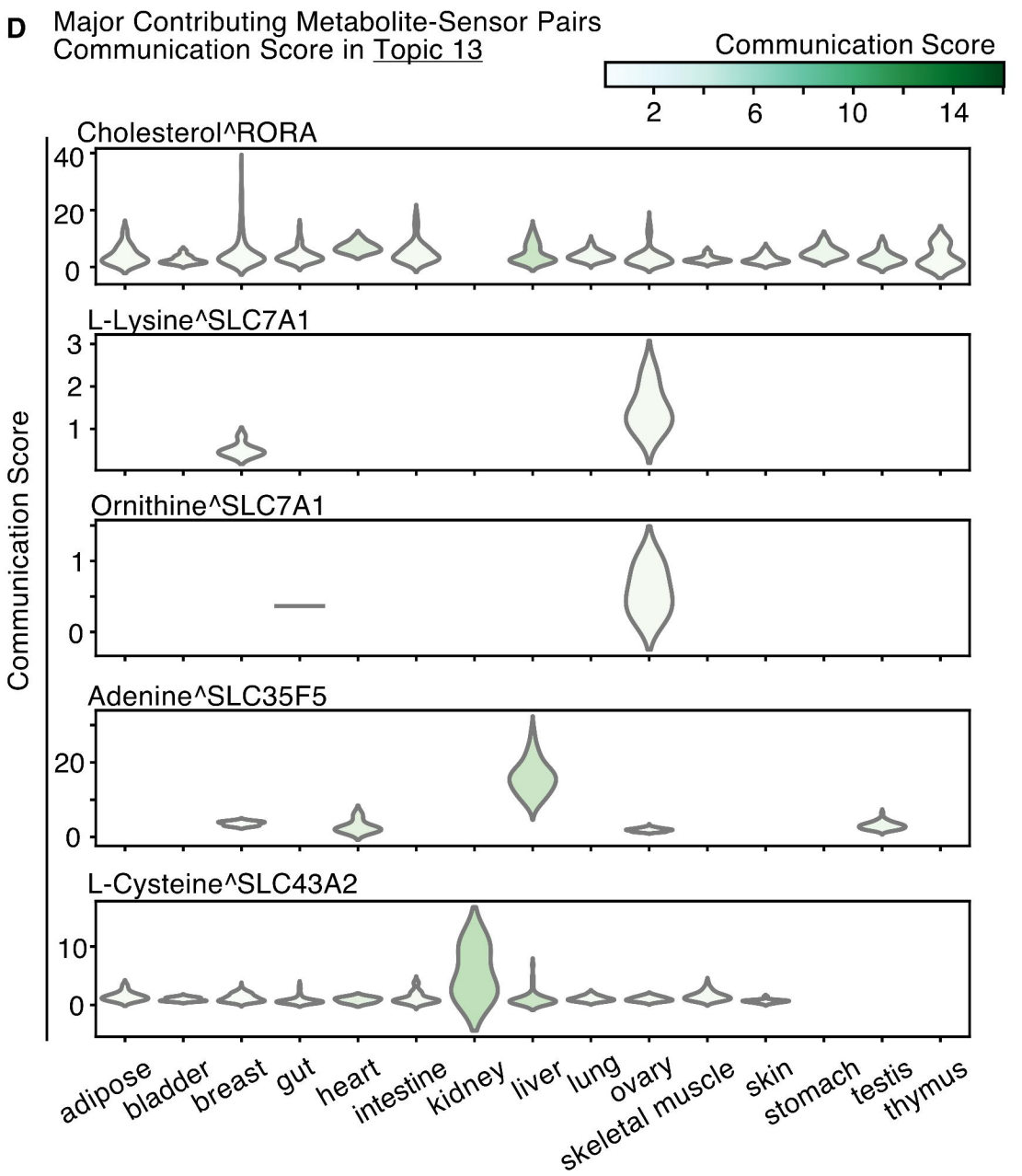

A
metabolite-mediated
protein-mediated
Metabolite-mediated
protein-mediated
B
average=7.09%
EC pct (%)
3.24 0.11 0.56 23.02 0.54 13.9 14.77 3.02 18.7 3.96 6.87 9.89 2.2 2.18 3.41
average=40.33
# cell type
18 29 29 30 52 29 38 49 36 40 45 37 62 67 44
C
correlation coefficient = -0.05
EC pct (%)
D
correlation coefficient = 0.76
# cell type
# Communications
E
metabolite-mediated
protein-mediated
kidney
adipose
stomach
skeletalMuscle
liver
heart
skin
thymus
bladder
testis
lung
breast
gut
intestine
ovary
Tissue Type

# Fig.4

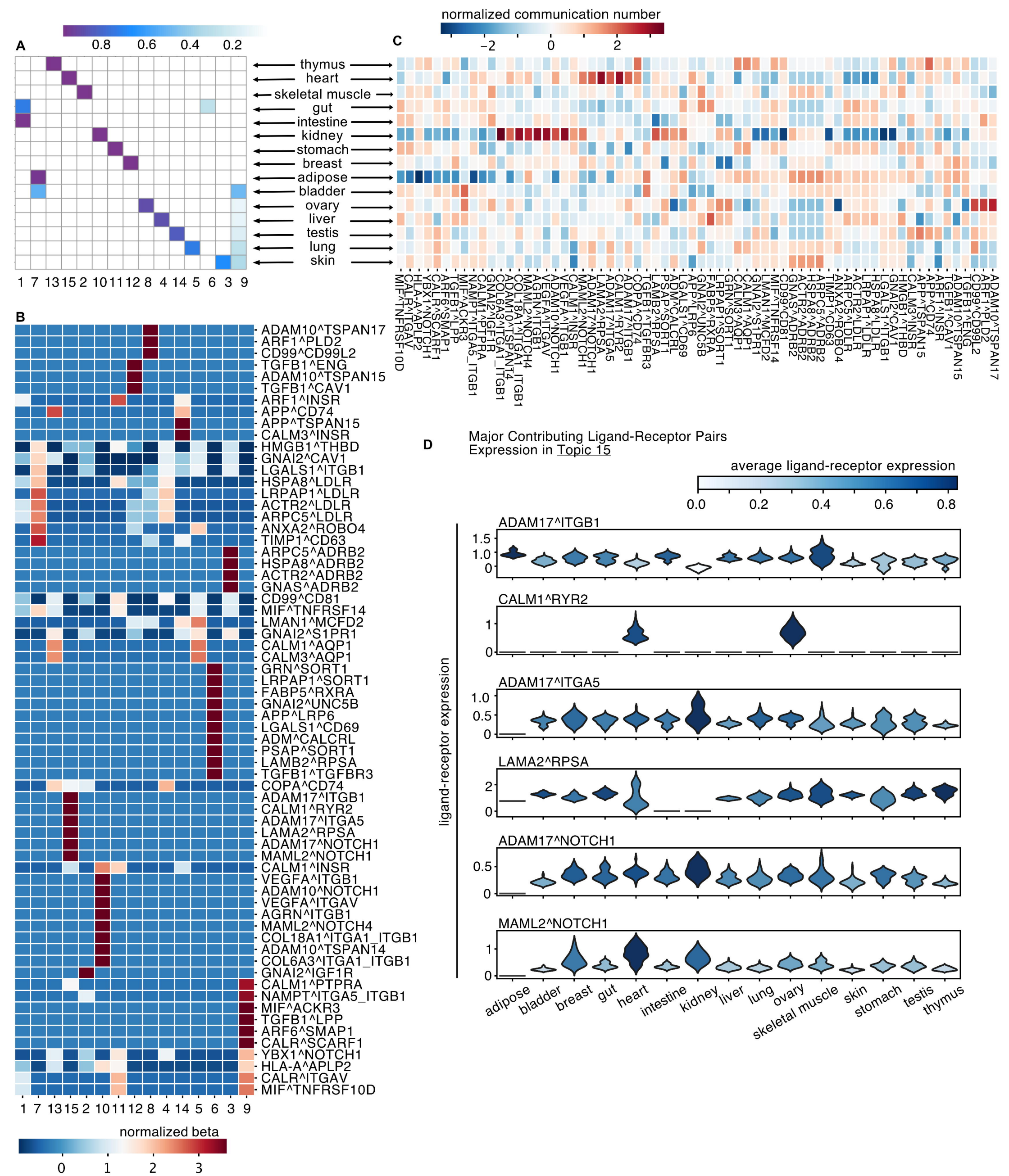

A
0.6
0.4
0.2
testis
thymus
bladder
liver
skin
ovary
adipose
breast
heart
kidney
skeletalMuscle
stomach
lung
gut
intestine
1 11 5 9 3 7 2 8 13 12 4 6 15 10 14


B
L-Glutamine^SLC38A2
Thyroxine^THRA
D-Mannose^SLC2A3
Retinal^RARA
D-Glucose^SLC2A3
L-Methionine^SLC7A5
L-Glutamine^SLC38A1
Cholesterol^LDLR
Phosphate^SLC20A2
Prostaglandin E2^PTGER4
L-Glutamine^SLC1A5
all-trans-Retinoic acid^RXRA
Sphingosine 1-phosphate^S1PR1
Hydrogen sulfide^CD36
Cholesterol^SCARB1
L-Arginine^AVPR2
L-Arginine^SLC7A1
Dehydroascorbic acid^SLC2A3
Retinal^RARB
Vitamin A^RARB
Cholesterol^RORA
L-Lysine^SLC7A1
Ornithine^SLC7A1
Adenine^SLC35F5
L-Cysteine^SLC43A2
L-Serine^SLC3A2
Vitamin A^RBP1
Iron^SLC40A1
Heme^LRP1
L-Glutamine^SLC7A5
L-Glutamine^SLC3A2
D-Glucose^SLC2A8
L-Serine^SLC1A5
Thymidine^SLC29A1
13-cis-Retinoic acid^RXRA
L-Glutamine^SLC38A7
D-Fructose^SLC2A8
all-trans-Retinoic acid^RARA
gamma-Aminobutyric acid^GABBR1
gamma-Aminobutyric acid^SLC6A6
Cholesterol^CD36
Choline^SLC44A2
Serotonin^SLC6A4
Choline^SLC44A1
Inosine^SLC29A1
Guanosine^SLC29A1
Cytidine^SLC29A1
Glycerol^AQP3
1 11 5 9 3 7 2 8 13 12 4 6 15 10 14
normalized beta
0
1
2
3


normalized communication number
C
-1
0
1
2
3
Glycerol^AQP3
Cytidine^SLC29A1
Guanosine^SLC29A1
Inosine^SLC29A1
Choline^SLC44A1
Serotonin^SLC6A4
Choline^SLC44A2
Cholesterol^CD36
gamma-Aminobutyric acid^SLC6A6
gamma-Aminobutyric acid^GABBR1
all-trans-Retinoic acid^RARA
D-Fructose^SLC2A8
L-Glutamine^SLC38A7
13-cis-Retinoic acid^RXRA
Thymidine^SLC29A1
L-Serine^SLC1A5
D-Glucose^SLC2A8
L-Glutamine^SLC3A2
L-Glutamine^SLC7A5
Heme^LRP1
Iron^SLC40A1
Vitamin A^RBP1
L-Serine^SLC3A2
L-Cysteine^SLC43A2
Adenine^SLC35F5
Ornithine^SLC7A1
L-Lysine^SLC7A1
Cholesterol^RORA
Vitamin A^RARB
Retinal^RARB
Dehydroascorbic acid^SLC2A3
L-Arginine^SLC7A1
L-Arginine^AVPR2
Cholesterol^SCARB1
Hydrogen sulfide^CD36
Sphingosine 1-phosphate^S1PR1
all-trans-Retinoic acid^RXRA
L-Glutamine^SLC1A5
Prostaglandin E2^PTGER4
Phosphate^SLC20A2
Cholesterol^LDLR
L-Glutamine^SLC38A1
L-Methionine^SLC7A5
D-Glucose^SLC2A3
Retinal^RARA
D-Mannose^SLC2A3
Thyroxine^THRA
L-Glutamine^SLC38A2


D
Major Contributing Metabolite-Sensor Pairs
Communication Score in Topic 13
Communication Score
2
6
10
14
Cholesterol^RORA
40
20
0
L-Lysine^SLC7A1
3
2
1
Ornithine^SLC7A1
1
0
Adenine^SLC35F5
20
0
L-Cysteine^SLC43A2
10
0
Communication Score
adipose
bladder
breast
gut
heart
intestine
kidney
liver
lung
ovary
skeletal muscle
skin
stomach
testis
thymus

# Fig.5


A
0.8
0.6
0.4
0.2
thymus
heart
skeletal muscle
gut
intestine
kidney
stomach
breast
adipose
bladder
ovary
liver
testis
lung
skin
1 7 13 15 2 10 11 12 8 4 14 5 6 3 9
B
ADAM10^TSPAN17
ARF1^PLD2
CD99^CD99L2
TGFB1^ENG
ADAM10^TSPAN15
TGFB1^CAV1
ARF1^INSR
APP^CD74
APP^TSPAN15
CALM3^INSR
HMGB1^THBD
GNAI2^CAV1
LGALS1^ITGB1
HSPA8^LDLR
LRPAP1^LDLR
ACTR2^LDLR
ARPC5^LDLR
ANXA2^ROBO4
TIMP1^CD63
ARPC5^ADRB2
HSPA8^ADRB2
ACTR2^ADRB2
GNAS^ADRB2
CD99^CD81
MIF^TNFRSF14
LMAN1^MCFD2
GNAI2^S1PR1
CALM1^AQP1
CALM3^AQP1
GRN^SORT1
LRPAP1^SORT1
FABP5^RXRA
GNAI2^UNC5B
APP^LRP6
LGALS1^CD69
ADM^CALCRL
PSAP^SORT1
LAMB2^RPSA
TGFB1^TGFBR3
COPA^CD74
ADAM17^ITGB1
CALM1^RYR2
ADAM17^ITGA5
LAMA2^RPSA
ADAM17^NOTCH1
MAML2^NOTCH1
CALM1^INSR
VEGFA^ITGB1
ADAM10^NOTCH1
VEGFA^ITGAV
AGRN^ITGB1
MAML2^NOTCH4
COL18A1^ITGA1_ITGB1
ADAM10^TSPAN14
COL6A3^ITGA1_ITGB1
GNAI2^IGF1R
CALM1^PTPRA
NAMPT^ITGA5_ITGB1
MIF^ACKR3
TGFB1^LPP
ARF6^SMAP1
CALR^SCARF1
YBX1^NOTCH1
HLA-A^APLP2
CALR^ITGAV
MIF^TNFRSF10D
1 7 13 15 2 10 11 12 8 4 14 5 6 3 9
normalized beta
0
1
2
3
C
normalized communication number
−2
0
2
D
Major Contributing Ligand-Receptor Pairs
Expression in Topic 15
average ligand-receptor expression
0.0
0.2
0.4
0.6
0.8
ligand-receptor expression
ADAM17^ITGB1
CALM1^RYR2
ADAM17^ITGA5
LAMA2^RPSA
ADAM17^NOTCH1
MAML2^NOTCH1
adipose
bladder
breast
gut
heart
intestine
kidney
liver
lung
ovary
skeletal muscle
skin
stomach
testis
thymus