CellWHISPER disentangles direct cell–cell communication from structural proximity

Anurendra Kumar[1], Felix Rivera[2], Bhavay Aggarwal[2], Nicholas Zhang[2], Ahmet Coskun[2], Saurabh Sinha[2,3*]

1 College of Computing, Georgia Institute of Technology, Atlanta, GA, 30332, USA
2 The Wallace H. Coulter Department of Biomedical Engineering, Georgia Institute of Technology, Atlanta, GA, 30332, USA
3 H. Milton Stewart School of Industrial & Systems Engineering, Georgia Institute of Technology, Atlanta, GA, 30318, USA

* Corresponding Authors
Corresponding Authors-
1. Saurabh Sinha, Email- saurabh.sinha@bme.gatech.edu

Abstract

Direct, contact-dependent cell–cell communication shapes tissue physiology and disease. Spatial transcriptomics offers a unique opportunity to infer such communication mechanisms, but the few existing tools for such inference are prone to high error rates due to spatial structure in distribution of cell types and gene expression. We present CellWHISPER, a statistical framework for inferring contact-mediated cell–cell communication, including gap junctions and ligand-receptor mechanisms, from single-cell resolution spatial data. CellWHISPER achieves strict error control via a specialized hypothesis-testing procedure that accounts for the confounding effects of cell-type-specific expression and spatial organization. Its use of an exact closed-form significance score obviates expensive permutation testing and enables scalable analysis of large tissues an extensive signaling compendia. To uncover salient communication patterns, CellWHISPER also includes a latent variable model that distills recurrent patterns into maps of mutual preferences among cell types and signaling genes. Applied to mouse brain STEREO-seq and Xenium datasets, CellWHISPER identifies cell type-specific gap-junction coupling and yields a comprehensive connexin-interaction map. We experimentally confirmed Connexin 43-mediated interfaces among astrocytes, microglia and endothelial cells, which were among the highest-ranked predictions; in contrast, existing approaches fail to prioritize these confirmed interactions. Differential analysis of an Alzheimer's disease model versus wild-type identifies preserved astroglial coupling but increased microglia-associated gap-junction communication. In summary, CellWHISPER disentangles direct cell–cell communication from structural proximity, enabling statistically robust and computationally scalable inference from spatial transcriptomics, leading to the discovery of tissue- and disease-associated contact-mediated signaling programs.

INTRODUCTION

Cell signaling plays a pivotal role in driving cellular heterogeneity, differentiation, and pathogenesis within tissues. Rapid proliferation of single-cell transcriptomics technologies has led to several computational tools for discovering cell-cell communication (CCC) mechanisms[1-4]. These tools, such as CellChat[5], CellPhoneDB[6] and NicheNet[7], typically reveal communication between specific cell types via curated ligand-receptor (LR) pairs. Broadly, CCC can be categorized into paracrine, autocrine, juxtacrine, and endocrine signaling[1]. Existing CCC methods focus on paracrine and autocrine signaling mediated by LR interactions, whereas direct juxtacrine signaling, particularly via gap junction communication (GJC), remains largely unexplored. Rapidly emerging spatial transcriptomics (ST) technologies including Stereoseq, VisiumHD, Nanostring and Xenium, provide a transformative opportunity to bridge this gap by combining gene expression with spatial proximity, enabling investigation of direct (contact-mediated) communication between cells. However, computational inference of direct CCC poses distinct statistical challenges: spatial organization of cell-types and heterogeneous gene expression act as strong confounders that can inflate apparent evidence for signaling if not explicitly controlled (**Fig. 1a**).

Direct cell-cell signaling operates primarily via two mechanisms: GJC[8] and LR interactions. Gap junctions are intercellular channels (**Supp. Fig. 7**) that enable direct cell-cell transfer of ions, metabolites, second messengers, and regulatory molecules[9]. GJC-mediated inter-cellular networks play vital roles in normal physiology[10-12] and are implicated in diverse pathological conditions[13-15]. Gap junction channels are formed by connexin proteins, assembled as hexamers of homotypic or heterotypic connexins. Over 20 connexins have been identified in mammals, and the existence of a compatibility "code" governing gap junction formation is an important open problem[16,17]. Similarly, direct LR pathways, such as Notch-Delta[18] and Eph-Ephrin[19] regulate differentiation, embryogenesis, and immune cell communication[20], motivating the search for additional instances of direct LR signaling in varying contexts[21].

Several existing tools allow CCC inference from single-cell transcriptomics, but very few of these (CellChatv2[22], SpaTalk[24], and COMMOT[25]) exploit inter-cellular proximity information available in ST data, a key requirement for inferring *contact-dependent* communication. We found these ST-based inference tools to exhibit surprisingly high false positive errors, motivating this study. Specifically, we performed a randomized control by permuting cell locations within each cell type, preserving cell-type-specific spatial organization and LR expression while destroying spatial proximity between ligand- and receptor-expressing cells. In this control, evidence for true direct CCC should be eliminated. Strikingly, the above-mentioned ST-based CCC inference tools reported comparable numbers of direct CCC mechanisms in randomized and original datasets, indicating insufficient statistical calibration for spatially constrained inference. This finding is consistent with published studies debating the error rates of existing tools [3,26] and suggestions of inflated false positives[1,4,27]. It motivated us to devise a statistical procedure for CCC inference that explicitly accounts for gene-expression heterogeneity, spatial organization, and statistical dependencies induced by shared cellular neighborhoods.

We present CellWHISPER (Workflow for HIgh-precision Spatial Proximity-mediated cEll-cell inteRactions), a toolkit to infer direct CCC from single-cell ST data. CellWHISPER makes three core computational contributions: (i) a proximity-constrained statistical test with a confounder-aware null that improves specificity for direct CCC, (ii) an analytically derived exact null (closed-form mean/variance of the CCC-measuring statistic) that avoids explicit permutations and enables scalable inference across large tissues and signaling compendia, and (iii) a latent variable model that analyzes the full tensor of tested CCC mechanisms to capture its higher-order interpretable

structures without first collapsing the tensor into information-losing matrices (as is currently done). Applying CellWHISPER to mouse brain datasets from STEREO-seq[28] and Xenium reveals known and novel GJC, reporting the first comprehensive brain “connexin code”[16,29]. We experimentally confirmed Connexin 43-mediated interfaces among astrocytes, microglia and endothelial cells, which were among the highest-ranked predictions; in contrast, existing approaches fail to prioritize these confirmed interactions. LR analysis recapitulates established CCC programs such as mossy fiber pathway while differential analyses identify communication mechanisms associated with Alzheimer’s Disease. In summary, CellWHISPER enables rigorous, scalable inference of direct inter-cellular communication, revealing tissue- and disease-specific signaling programs.

RESULTS

Overview of CellWHISPER

CellWHISPER infers direct cell-cell communication (CCC) involving a "quadruplet" of two cell types and two signaling genes. Given ST data and cell-type labels, it tests whether a cell-type pair shows evidence of direct communication via a signaling gene pair[3] (connexins or LR pairs), returning a list of cell type-signaling gene quadruplets with z-scores quantifying significance (**Fig. 1b,c,** Methods). A latent variable model (LVM) summarizes the resulting quadruplet set into interpretable modules and preference maps, and an optional differential analysis identifies condition-specific mechanisms (Methods, **Fig. 1c**).

At the core is a statistical test for whether cell types CT_i and CT_j communicate via genes SG_k and SG_l. CellWHISPER binarizes SG expression using a quantile-based threshold. Spatial information is represented by a k-nearest neighbors (KNN) graph[30], from which a subgraph ("whisper network") is derived to include edges between neighboring cells of types CT_i and CT_j expressing genes SG_k and SG_l respectively (Fig. 1b). The whisper network edge count (N_{ijkl}) quantifies the evidence for this quadruplet in the tissue sample. To assess significance, CellWHISPER shuffles locations among cells of same type (preserving cell-type expression and spatial distributions), recomputes N_{ijkl} and uses the resulting null to compute a z-score. For scalability, CellWHISPER derives an analytically computed exact null for the within-cell-type randomization scheme, avoiding explicit shuffling (Methods). This exact null utilizes closed-form expressions for the expectation and variance of the whisper network edge count, explicitly accounting for dependence induced by shared cells in the spatial graph. As a result, CellWHISPER computes calibrated z-scores and enables orders-of-magnitude improvements in computational efficiency relative to permutation-based approach (**Fig. 1e,f**). In practice, this enables direct CCC inference on datasets comprising tens of thousands of cells and thousands of signaling gene pairs on standard hardware, without exceeding memory limits or incurring prohibitive runtimes. To assess statistical control for direct CCC, we applied the within-type randomization control described above. On downsampled Stereo-seq data (~5,000 cells, 28 cell types, 367 LR pairs) competing tools (CellChat v2[22], COMMOT[25], SpaTalk[24]) reported similar numbers of significant quadruplets on real and randomized data whereas CellWHISPER reported far fewer quadruplets on randomized data, yielding an empirically estimated FPR <5%, underscoring its high specificity (**Fig. 1d,** Methods).

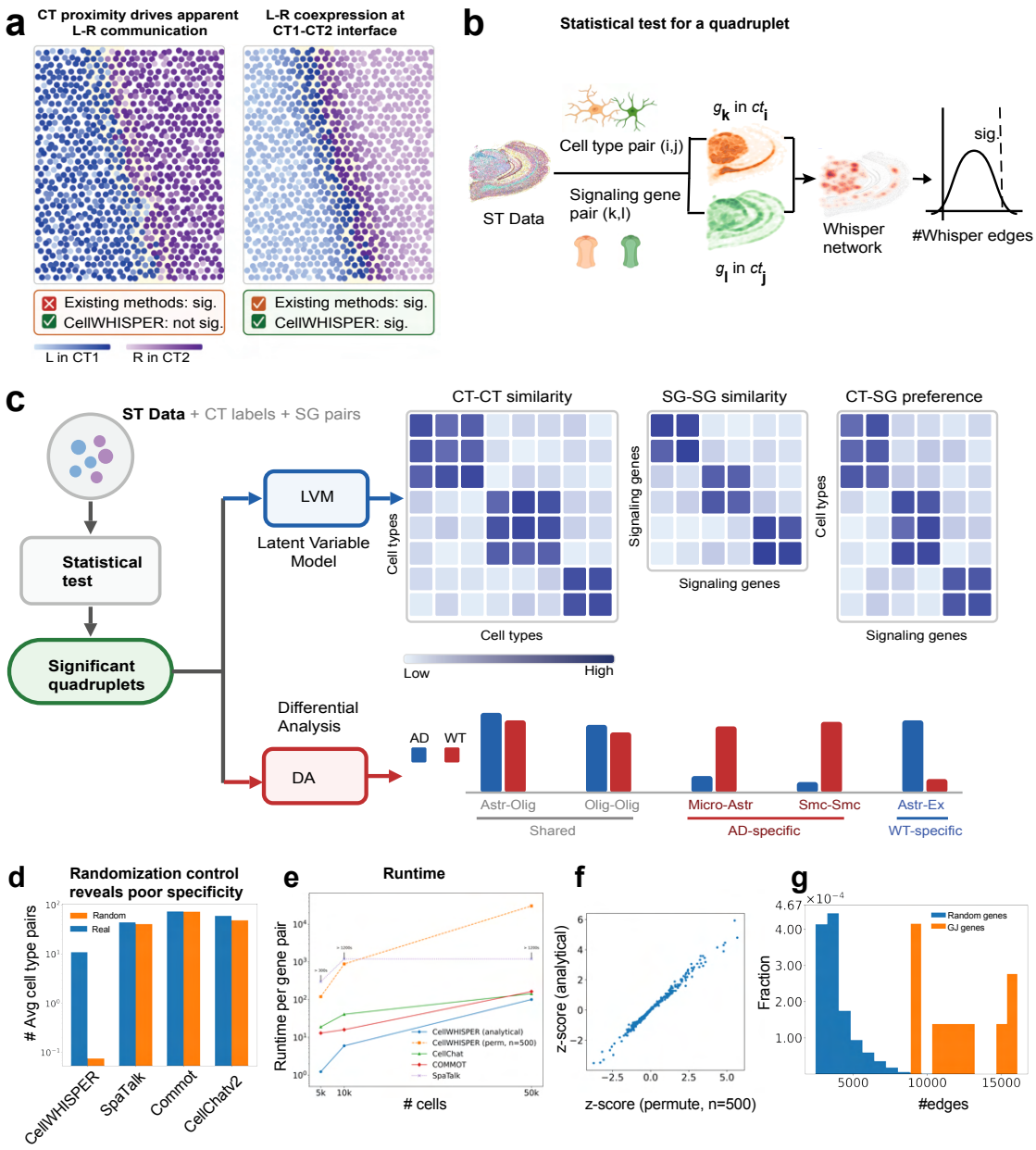

Figure 1. Overview of CellWHISPER. (**a**) **Cell-type confounding in ligand–receptor communication inference.** When ligand L is expressed broadly across cell type CT1 and receptor R across CT2, spatial proximity between the two cell types produces apparent L–R co-expression at their interface (left; yellow zone) even without specific L–R signaling. When the same number of high-expressing cells are enriched specifically at the CT1–CT2 interface, the co-expression reflects genuine signaling (right; yellow zone). Both scenarios appear identical to existing methods, which report significant communication in both cases. CellWHISPER evaluates co-expression against a null that randomizes cell identities within each cell type, preserving cell-type-specific expression distributions and spatial organization, and correctly distinguishes the two. (**b**) **Statistical testing for a CT-SG quadruplet**: For a given cell type pair (ct_i, ct_j) and signaling gene pair (g_k, g_l), CellWHISPER constructs a "whisper network", where each edge represents a pair of

adjacent cells of types ct_i, ct_j expressing genes g_k, g_l respectively, and the count of such "whisper edges" serves as the test statistic. Under the null model, cell locations are randomized within each cell type, generating a background distribution of edge counts, which is used to assess statistical significance. (**c**) **CellWHISPER framework**: CellWHISPER takes single-cell resolution ST data as input, along with cell type (CT) labels and signaling gene pairs, which may be ligand-receptor (LR) gene pairs or gap junction (GJ) gene pairs. It performs a statistical test (panel b) to compute a z-score for each cell type pair and signaling gene pair (a CT-SG quadruplet), reporting all quadruplets with significant z-scores. A latent variable model can then analyze the full complement of quadruplet-level results to infer higher-order communication patterns such as CT pairs that are similar in their SG usage (CT-CT similarity matrix), SG pairs that mediate communication between similar cell types (SG-SG similarity matrix), and quantify whether a certain cell types tends to communicate using a specific SG (CT-SG preference matrix). CellWHISPER can also be applied across conditions for differential analysis, identifying CT-SG quadruplets specific to one condition versus the other. (**d**) **Randomized control reveals poor specificity**: Shown is the average number of communicating cell-type pairs inferred per ligand–receptor by four different methods, on real versus randomized data (cell identities permuted within each cell type, preserving expression distributions and cell-type spatial organization while disrupting cross-type proximity). Competing methods show limited separation between real and randomized data, whereas CellWHISPER yields substantially fewer calls under randomization, indicating significantly improved specificity for direct CCC. We summarize this as an empirical FPR, defined as the number of significant CT–SG quadruplets detected on randomized data divided by the number detected on real data under fixed thresholds (CellWHISPER: <5%). **Benchmarking was performed on a downsampled hippocampal subset (~5,000 cells) due to computational constraints of competing methods** (Methods, Supp. Fig. 5). (**e**) **Runtime and scalability.** Wall-clock time to evaluate all cell-type pairs for a single signaling-gene pair on datasets of increasing size (~5,000; ~10,000; and ~50,000 cells), comparing CellWHISPER with the analytically derived exact null versus permutation testing (n = 500), and three competing tools (CellChat, COMMOT, SpaTalk) run with recommended settings. CellWHISPER's analytical null achieves ~100× speedup over permutation on ~5,000 cells and substantially larger speedups at ~10,000 and ~50,000 cells, enabling practical screening on full-resolution tissues. SpaTalk did not complete within the allotted runtime at one or more dataset sizes (shown as timeouts). **(f) Consistency between analytical and permutation-based significance**. Z-scores computed using the analytical null closely match those from permutation testing across CT–SG tests, confirming that the analytical formulation reproduces the permutation-based null while avoiding explicit shuffling. (**g**) **Whisper network specificity for connexins:** For each homotypic connexin gene, an aggregated whisper network was constructed as the union of whisper networks across all cell-type pairs in the Stereo-seq mouse brain dataset. Edge counts for connexin networks were significantly higher than those for networks constructed from randomly selected genes (n = 5000) (Mann–Whitney U test p = 4.6 × 10e-8). All connexins yielded ≥8,878 edges, whereas only 0.4% of random genes reached this threshold.

We focused initial validation on one of the best studied mechanisms of direct CCC: gap junctions (GJ). Gap junction communication provides a stringent stress test for direct CCC inference because they are purely contact-mediated and therefore particularly sensitive to spatial and expression confounders. GJ form when connexin protein in adjacent cells dimerize to establish intercellular channels[31] (Fig. 1a). To evaluate the plausibility of using mRNA levels as proxies for connexin-mediated communication, we performed targeted spatial profiling in an induced pluripotent stem cell model[12] and observed concordance between *Gja1* mRNA and Cx43 protein abundance, as well as substantial overlap between mRNA- and protein-derived whisper networks (**Supp. Figure 1, Methods**). In a Stereo-seq mouse brain dataset[28] (Methods), aggregated whisper networks for connexin pairs were significantly larger than those from random genes (**Fig. 1g**, Mann-Whitney U test p-value 4.6e-8, Methods): each of 10 connexins yielded a whisper network of ≥8878 edges, while only 22/5000 (0.4%) random genes met this criterion (**Supp. Figure 2**).

CellWHISPER identifies known and novel mechanisms of gap junction communication (GJC) in the mouse brain

To demonstrate the utility of CellWHISPER, we first applied it to discover GJC in mouse brain. From a computational perspective, GJC provides a stringent benchmark for direct CCC inference because it is purely contact-mediated and therefore maximally sensitive to spatial and expression confounders. GJC is believed to be widespread in the brain[32] and has been implicated in neurodevelopment and diseases [33-35]. Certain combinations of cell types and connexins are preferentially involved in brain GJC [31,36] and we sought to explore such a “connexin code”[16,29]. We analyzed a STEREO-seq mouse hemibrain dataset[28], comprising of ~50,000 cells across 28 cell types (CTs), using ten highly variable connexins as signaling genes (SG). (All 100 pairs of these genes were included in the SG compendium). CellWHISPER identified 1726 significant CT-SG quadruplets (z-score > 3, whisper network size > 30, **Fig. 2a**, **Supp. Data 1**) at an empirical FPR <1% estimated from randomized controls (**Supp. Figure 3**). This demonstrates that well-controlled inference over large quadruplet spaces is feasible at tissue scale (Fig 1e). Predictions were dominated by inter-glial GJC mechanisms (**Fig. 2b**).

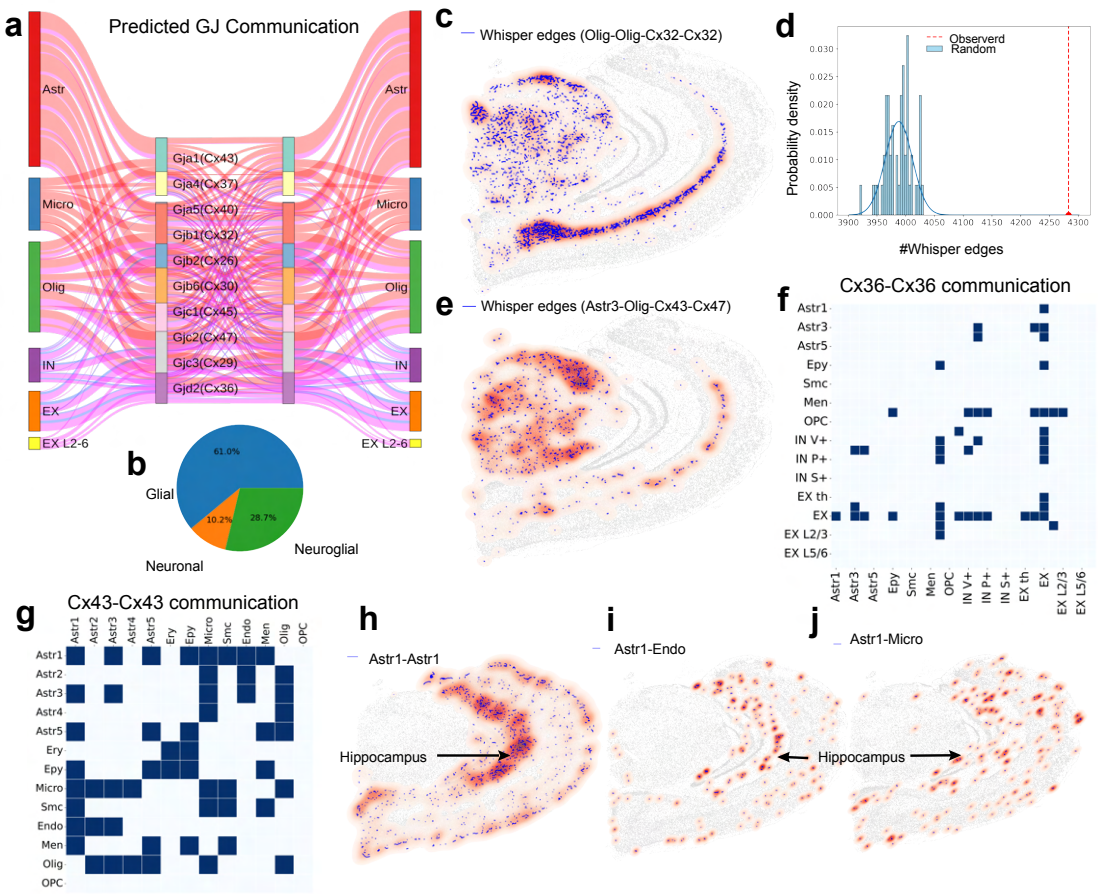

Figure 2. CellWHISPER Identifies Direct Gap Junction Communication in the Mouse Brain. (a) Gap junction communication identified in mouse brain: Sankey plot of the top 150 significant CT-SG quadruplets (a cell type pair communicating via a gap junction gene pair) ranked by whisper network size identified in Stereo-seq data (~50K cells). For visual ease, "Astr" represents all astrocyte subtypes (Astr1-5), "IN" represents all interneuron subtypes and "EX L2-6" represents all excitatory neuron layers. ("Olig" = oligodendrocyte, "Micro" = microglia.) Lines indicate CellWHISPER-predicted communication involving specific cell type pairs (or aggregated groups such as Astr) and specific gap junction genes: red for glial-glial communication, blue for neuronal-neuronal communication, and pink for neuroglial communication. Line thickness reflects the magnitude of communication for each CT–SG pair, while node height reflects the total communication associated with that cell type or signaling gene. (**b**) Pie chart of the three major classes of gap junction communication in the top 150 predicted CT-SG quadruplets shown in (a). (**c**) **Oligodendrocyte self-communication.** Whisper network of oligodendrocytes communicating via homotypic Cx32 gap junctions. Blue edges represent whisper network connections; red shading denotes kernel density smoothing at edge centers for visual clarity. (**d**) **Statistical significance of whisper edges.** Histogram of edge counts from randomized whisper networks (cell locations permuted within cell type), with

fitted normal distribution (blue line). The red dashed line marks the observed number of edges in the whisper network shown in (c) which is significantly higher than expected under the null. Y-axis shows the probability density of randomized edge counts. (**e**) **Heterotypic astroglial communication.** Whisper network of astrocytes predicted to communicate with oligodendrocytes via heterotypic gap junction formed by Cx43 (astrocytic side) and Cx47 (oligodendrocytic side). (**f**) Cell types predicted to communicate using homotypic Cx36 gap junctions. Blue entries in heatmap denote significant cell type pairs (z score>3, whisper edge count > 30) found by CellWHISPER. Neuronal communication is predicted more frequently compared to glial communication. (**g**) Glial cell types predicted to communicate using Cx43-Cx43 gap junctions. Only glial communication is shown since it is predicted more frequently compared to neuronal communication (Supp. Fig. 6) (**h–j**) Astrocytic communication enriched in hippocampus. Whisper networks showing (h) astrocyte–astrocyte, (i) astrocyte–endothelial, and (j) astrocyte–microglia interactions mediated by homotypic Cx43 (GJA1) gap junctions

We examined the highest-confidence mechanisms from two complementary perspectives. Ranking quadruplets by whisper network size revealed that the top six represent inter-oligodendrocyte GJC via combinations of Cx47, Cx32, and Cx29 (**Supp. Data 1, Supp. Data 2**) consistent with established roles of Cx47 and Cx32 in oligodendrocytic coupling[37,38]. **Fig. 2c** illustrates homotypic Cx32 whisper network between oligodendrocytes, comprising 4,283 edges, far exceeding chance expectation (**Fig. 2d**, z-score = 13.09). The top predictions also included Cx29-mediated oligodendrocytic GJC. Although Cx29 transcripts are enriched in oligodendrocytes[39], prior studies indicate that it primarily localizes to the adaxonal membrane rather than forming functional gap junctions[40], highlighting a known limitation of transcript-based inference when subcellular protein localization is critical. The next highest-ranking quadruplets involve astrocyte-astrocyte GJC via Cx43, Cx30, and Cx26, connexins known to couple astrocytes into an astrocytic syncytium essential for homeostasis and adult neurogenesis[17,41-44]. We also identified high-confidence astrocyte-oligodendrocyte (A/O) communication via Cx43 on the astrocyte side and Cx47, Cx32 or Cx29 on the oligodendrocytic side. **Fig. 2e** illustrates *Astr3*-oligodendrocyte communication via Cx43-Cx47, consistent with prior reports supporting functional A/O GJ coupling[45], though there is some unresolved debate about their functionality[44].

Examining CellWHISPER results ranked by z-score rather than network size (**Supp. Data 3**) highlights additional mechanisms supported by fewer communicating cell pairs, which may thus be less well-studied in prior work. Among these, we identified GJC between excitatory neurons and other cell types, including thalamic excitatory neurons, oligodendrocytes, and astrocytes (**Supp. Data 3**). Maubecin et al.[46] provided direct evidence of functional neuro-glial coupling using

265 heterotypic gap junctions in rodents. Gap junctions between excitatory neurons are known to be
266 primarily mediated by Cx36 (Gjd2)[47,48] and extensively studied in rodent models[49-52]. As shown in
267 **Fig. 2f**, excitatory neurons are a major site of homotypic Cx36-mediated communication predicted
268 by CellWHISPER.
269
270 **CellWHISPER identifies homotypic Cx43 gap junctions as a mechanism underlying**
271 **astrocytic syncytium**
272 Cx43 is the most studied connexin, with key roles in differentiation and homeostatic regulation[53-
273 55]. We examined the complete map of cell type pairs predicted by CellWHISPER to participate in
274 Cx43-mediated GJC (**Fig. 2g**), finding a predominance of glial over neuronal communication. This
275 analysis provides a focused validation setting, as Cx43-mediated interactions are well
276 characterized. Astrocyte subpopulation Astr1 is predicted to communicate with other astrocytes,
277 as well as endothelial, microglia, oligodendrocytes and ependymal cells. This aligns with known
278 biology: Cx43 mediates astrocyte-oligodendrocyte-ependymal cell interactions forming an
279 isopotential syncytium [54,56-58]. Predicted astroglial interactions exhibited a higher network density
280 in the hippocampus (**Figs. 2h-j**). Inter-astrocyte GJC (z-score 18.13, **Fig. 2h**) in the hippocampus
281 has been documented to be Cx43-mediated. Astrocyte-endothelial GJC (z-score 10.94, **Fig. 2i**)
282 has been implicated in maintaining blood–brain barrier integrity[59-61]. Additionally, CellWHISPER
283 predicted Cx43-mediated communication between microglia and astrocytes (**Fig. 2j**) as well other
284 microglia. Indeed, recent studies suggest that microglia upregulate Cx43 under pathological or
285 inflammatory conditions and form functional gap junctions – with each other and with astrocytes[62-
286 64].
287

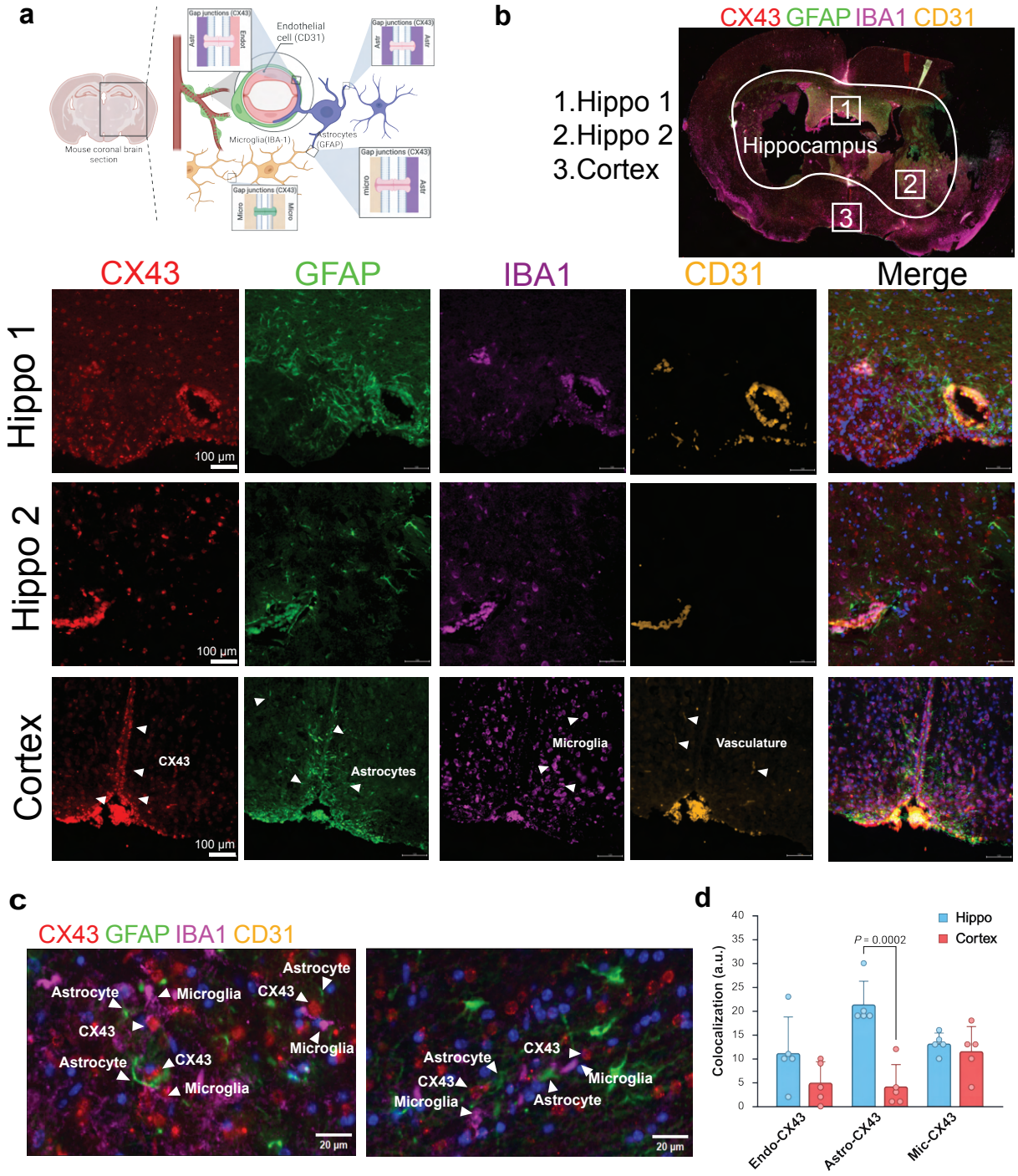

Figure 3. Spatially resolved analysis of Cx43-mediated cellular crosstalk across brain regions. (**a**) Schematic of gap junction–mediated interactions between astrocytes (GFAP), microglia (IBA1), and endothelial cells (CD31) via Cx43 (created with BioRender.com). (**b**) Immunofluorescence of coronal mouse brain section stained for Cx43, GFAP, IBA1, and CD31. Hippocampal boundaries were defined using the

Allen Mouse Brain Atlas. Regions of interest (ROI) within hippocampus (Hippo1, Hippo2) and cortex were outlined for higher-resolution imaging. Scale bar, 100 µm. (**c**) Single-cell–resolution images of hippocampal ROIs showing colocalization of astrocytes and microglia with Cx43. (**d**) Quantification of Cx43 colocalization across hippocampus and cortex. We selected 10 regions (478 × 262 pixels each) within and outside the hippocampus and quantified Cx43 colocalization with astrocytes, endothelial cells, and microglia. Colocalization was defined as instances where Cx43 puncta were positioned between or overlapping the indicated cell types, yielding 30 values (one per cell type per region) and a total of 333 interactions. Two-way ANOVA revealed a significant enrichment of astrocyte–Cx43 interactions in the hippocampus compared to the cortex ($p = 0.0002$).

To validate these computationally inferred predictions, we performed immunofluorescent staining on mouse coronal brain sections for Cx43 together with markers of astrocytes (GFAP), endothelial cells (CD31), and microglia (IBA1) (**Fig. 3a, b**; Methods). All channels showed robust signal across brain regions (**Fig. 3b**). Distinct Cx43 co-localization was evident at astrocyte–astrocyte, astrocyte–endothelial, astrocyte–microglial, and microglial–microglial interfaces (Fig 3b,c). Quantitative analysis showed higher co-localization of Cx43 with cell-type markers in the hippocampus compared to cortex (**Fig. 3d**), with astrocytic contacts enriched >5-fold. Representative examples of colocalization of astrocytes, microglia and Cx43 in the hippocampus are shown in **Fig. 3c**. Notably, these experimentally supported Cx43 interfaces (Astrocyte–Astrocyte, Astrocyte–Endothelium, and Astrocyte–Microglia) are prioritized among the highest-ranked CellWHISPER predictions, enabling targeted experimental follow-up. By contrast, CellChat v2 (the only ST-based tool that scales to this data set) assigns extremal significance scores (p-value = 0) to a large set of candidate cell-type pairs (~150 pairs), including these interfaces, limiting its usefulness for prioritizing specific mechanisms. Together, these observations support the ability of CellWHISPER to recover known, region-specific contact-mediated communication patterns using a rigorous confounder-aware framework.

CellWHISPER identifies direct ligand-receptor (LR) communication

Direct LR communication plays a fundamental role in organizing local tissue architecture[21]. We next used LR interactions as a test of generality of the CellWHISPER analytic framework. Unlike gap junctions, LR signaling spans diverse functional classes (adhesion, guidance, immune and ECM), while still requiring spatial proximity for direct interactions. To identify such interactions in the mouse brain, we re-analyzed STEREO-seq data using 367 curated LR pairs from the CellChat database[5] (Methods). CellWHISPER identified 4743 significant CT–SG quadruplets (z-score > 4,

whisper network size > 30) (**Fig. 4a, Supp. Data 4**), with an empirically estimated FPR < 1%, spanning inter-glial, neuron-glia and neuron-neuron communication (**Fig. 4b**).

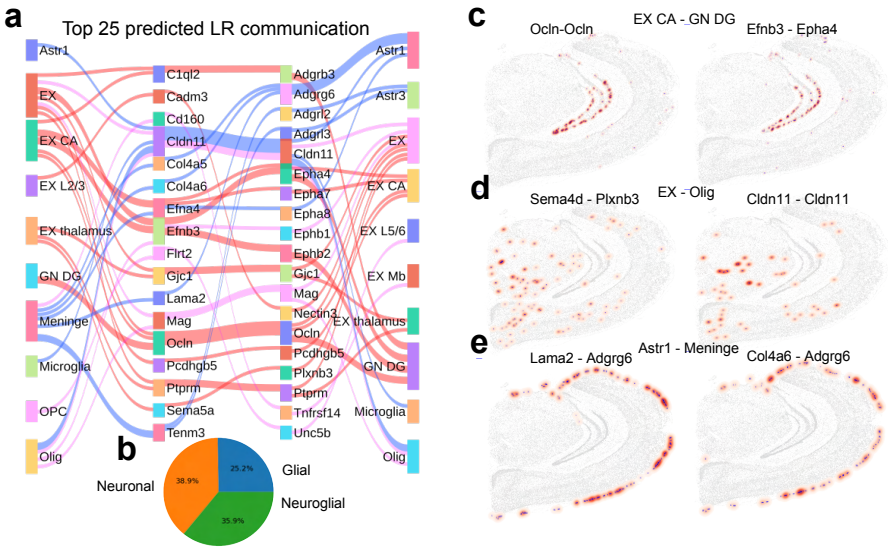

Figure 4. CellWHISPER Identifies direct Ligand-Receptor communication in the mouse brain. (a) Ligand–receptor communication mechanisms. Sankey plot showing the top 25 significant CT–SG quadruplets **ranked by z-score** (z-score > 4, whisper network size > 30) identified using 367 curated ligand–receptor pairs. Randomized control analysis estimated a false positive rate (FPR) < 1% for this compendium. **(b)** Communication classes. Distribution of major interaction types, including neuron–neuron, neuron–glia, and inter-glial communication. **(c–e)** Six representative whisper networks (two per panel, left/right). Selected ligand–receptor–mediated cell–cell interactions illustrating key modes of direct communication: (c) excitatory CA3 neurons (EX CA) and dentate gyrus granule neurons (GN DG) communicating via Efnb3-Epha4 (right) and tight junction Ocln (left); (d) excitatory neuron-oligodendrocyte interactions involving *Cldn* (right) and semaplexin family member (left); and (e) meningeal- astrocyte communication via extracellular matrix components (*Lama2*, *Col4a1–6*, *Adgrg6*).

The highest confidence predictions revealed communication between CA3 excitatory neurons (EX CA) and dentate gyrus granule neurons (GN DG) mediated by LR pairs of the mossy fiber pathway[65] (**Supp. Data 5**). These include multiple ephrin/Eph family members, C1ql2/3-Adgrb3 and homotypic interactions of occludin, protocadherin and claudin (**Supp. Data 5, Fig. 4c**). Ephrin–Eph and C1ql2/3–Adgrb3 (BAI3) interactions, both implicated in mossy fiber–CA3 synapse organization[66-68], ranked among the highest-confidence predictions, alongside multiple

348 adhesion-related pairs (including claudins/occludin). These findings are consistent with known
349 hippocampal wiring programs[69] and support the biological plausibility of CellWHISPER's
350 proximity-constrained LR inference.
351
352 CellWHISPER also predicted LR-mediated communication between oligodendrocytes and
353 excitatory neurons, involving homotypic MAG**,** claudin and semaplexin interactions (**Supp. Data**
354 **6, Fig. 4d**), consistent with known roles for these mechanisms in axon–glia recognition, myelin
355 stability[70], and activity-dependent modulation of myelination[71]. Additionally, we observed
356 interactions between astrocytes and meningeal cells, notably involving extracellular matrix
357 components (ECM) such as Lama2, type IV collagens (Col4a1–6), Adgrg6 (**Supp. Data 7, Fig.**
358 **4e**). These examples illustrate that CellWHISPER supports discovery across multiple LR
359 functional categories while maintaining stringent specificity for direct CCC.
360
361 **Latent Variable Modeling (LVM) infers general patterns of communication mechanisms**
362 We used CellWHISPER's "LVM" to infer higher-order patterns among significant quadruplets
363 (Methods). Unlike approaches that first collapse quadruplet-level information into CT x SG or CT
364 x CT matrices[5,6], an aggregation step that can be information-losing; CellWHISPER directly
365 models quadruplets to characterize communication profiles of cell types and signaling genes.
366 Specifically, the LVM models the presence or absence of statistically supported CT–SG
367 quadruplets using a shared low-dimensional embedding space for cell types and signaling genes
368 (Methods). Under this model, a quadruplet is more likely when each cell type exhibits higher
369 affinity for its corresponding signaling gene, yielding embedding geometry that captures CCC
370 preferences. The learned embeddings yield (1) CT–SG preferences, (2) CT–CT similarity, and
371 (3) SG–SG similarity. Model parameters are estimated via constrained log-likelihood
372 maximization. **Fig. 5b** shows that the optimization can accurately infer cell type similarity on
373 simulated data (Methods).
374
375 Applying LVM to GJ results in mouse brain Stereo-seq data revealed five dominant cell type
376 modules in CT-CT similarity map (**Fig. 5a, Supp. Data 8**) whereas a baseline method performing
377 naïve aggregation of quadruplet tensor did not reveal modular structure (**Supp. Figure 4**). LVM
378 also identified four GJ modules in SG-SG similarity map (**Fig. 5d**, Supp. Data 8). Cell type
379 modules showed preferences for specific GJ modules in the CT-SG preference matrix (**Fig. 5c**).
380 For example, excitatory neuron subtypes in cell type modules M1/M2 (EX Mb, EX L2/3, EX CA)
381 prefer GJ module M4 consisting of Gjd2 (Fig. 5c,d). Notably, DA neurons (excitatory or inhibitory

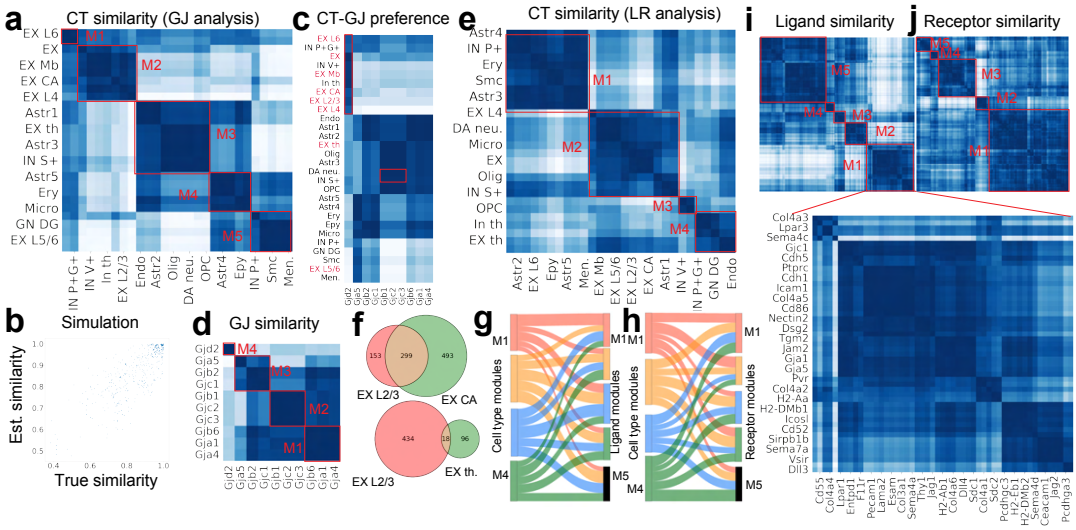

Figure 5. LVM infers module of signaling cell types and signaling genes. (**a**) Cell type similarity matrix. Cosine similarity of latent cell-type embeddings derived from the LVM routine of CellWHISPER, applied to gap junction communication in the Stereo-seq mouse brain dataset. Five dominant modules of cell types emerge, reflecting recurrent patterns of gap junction communication across the tissue. (**b**) Optimization. Constraint optimization within the latent variable model (Methods) identifies underlying cell type pair similarity, resolving higher-order structure in the data. (**c**) Cell type–gap junction preferences. Heatmap showing associations between cell type modules and gap junction gene modules. Distinct modules exhibit preferential coupling, e.g., Module M1 and M2 are enriched in excitatory neurons and prefer Gjd2 (GJ module M4). DA neurons and IN S+ prefer GJ module M2. **(d)** Cosine similarity of latent embeddings across gap junction genes, revealing four dominant connexin modules. (**e**) Cell-type similarity matrix for ligand–receptor communication. Cosine similarity of latent cell-type embeddings derived from CellWHISPER applied to ligand–receptor interactions (Stereo-seq mouse brain dataset). Four dominant cell-type modules emerge. Excitatory neurons in cortical layers and hippocampal CA region cluster together (M2), whereas excitatory neurons in thalamus is in module M4. (**f**) Communication overlap between excitatory neurons. Quantification of overlap in ligand–receptor communication patterns (shared CT–ligand–receptor triplets) shows that excitatory neurons in cortical layers 2/3 are more similar to CA excitatory neurons than to thalamic excitatory neurons, consistent with known laminar and neuromodulatory properties of neocortical and hippocampal neurons. (g–h) Cell type–ligand and cell type–receptor module specificity. Sankey plots showing preferential associations between cell-type modules and ligand modules (g), and between cell-type modules and receptor modules (h). Links were derived by averaging CT–SG preference scores within each module. (**i**) Ligand similarity map. Cosine similarity of ligand embeddings reveals five dominant modules (Supp. Data 9). Module M1 consists mainly of genes related to cell adhesion (e.g., PECAM1, CDH1/5, ICAM1), extracellular matrix components (COL4A1–6, LAMA2), and immune modulators (H2-Aa, H2-Eb1, H2-Ab1, CD52, CD55, CD86, PTPRC). Another ligand module (M5, Supp.

Data 9) is enriched for axon guidance and synaptic signaling genes , including protocadherins, ephrins, neurexins, semaphorins, Lrrc and Flrt family (**j**) Receptor similarity map. Cosine similarity of receptor embeddings identifies five receptor modules(Supp. Data 9).

depending on context[72]), clustered with inhibitory Sst+ interneurons[73] in cell type module M3 (**Fig. 5a**, Supp. Data 8) and shared preference for GJ module M2 (**Fig. 5c**), consistent with dye-coupling evidence of DA–interneuron electrical synapses[74] and suggesting candidate GJ mediating these synapses.

Next, we applied LVM to LR interactions (Fig. 4), identifying four dominant cell-type modules (**Fig. 5e, Supp. Data 9)**. Excitatory neurons in cortical layers and CA region co-cluster, whereas thalamic excitatory neurons lie in different module (**Fig. 5e,f**). This aligns with known features of excitatory neurons in the neocortex and hippocampus (of which CA is part), including laminar structure[75,76] and shared neuromodulatory response[77], different from excitatory neurons in thalamic neurons[78]. Analysis of ligand similarity map (**Fig. 5i**) identified five dominant modules, of which one (M1) is enriched for cell adhesion, ECM and immune modulators (**Supp. Data 9**). This ligand module preferentially associates with cell type module M1 (**Fig. 5g),** enriched for glial cell types located near the brain periphery, such as ependymal cells, meninges, and smooth muscle cells (Fig. 5e), suggesting a specialized glial–meningeal signaling axis potentially involved in barrier maintenance and neuroimmune communication. Another ligand module (M5, Supp. Data 9) is enriched for axon guidance and synaptic signaling genes[79,80], and prefers cell type module M2 (**Fig. 5g**), enriched in excitatory neurons in the cortical layers and hippocampus, consistent with their known role in neuronal cell-cell communication[81]. Similarly, we found five dominant modules in receptor similarity map (**Fig. 5j**, Supp Data 9) having preference for specific cell type modules (**Fig. 5h**). Together, LVM captures dominant communication patterns across GJC and LR signaling.

Differential analysis reveals Alzheimer's disease-associated GJC mechanisms

Gap junction-mediated communication (GJC) supports central nervous system homeostasis and is implicated in neurodegenerative diseases, including Alzheimer's disease (AD)[31,82,83]. Under normal conditions, GJC supports metabolic and synaptic function, especially through astrocytic networks[63,84] whereas in AD, aberrant gap junctions and hemichannel activity contributes to neuroinflammation and neurodegeneration[31,82,83] [31]. From a computational perspective, this analysis tests whether CellWHISPER can identify condition-associated rewiring using the same

statistically well-controlled, proximity-constrained quadruplet framework, while limiting false positives arising from differences in cell-type composition, spatial organization, or expression. To systematically interrogate these dynamics, we used CellWHISPER to decode GJC networks altered in AD.

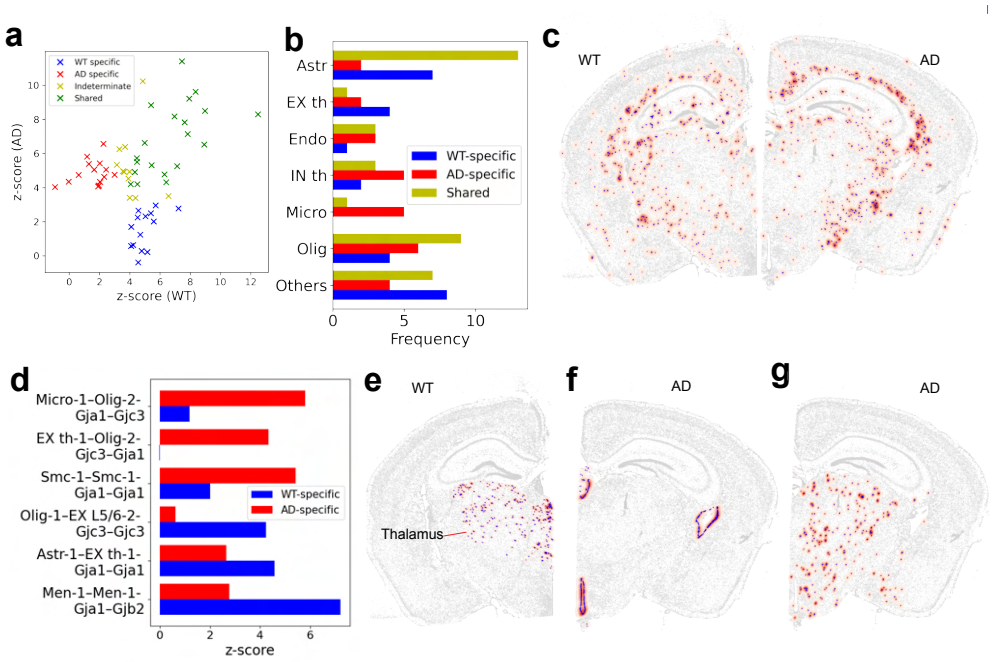

Figure 6. Differential Analysis infers conserved and condition-specific gap junction communication in Alzheimer's disease (AD). (**a**) Scatter plot of 60 high-confidence CT–SG quadruplets (z-score > 4, whisper network size > 30 in at least one condition) showing their z-scores in wild-type (WT) and TgCRND8 (AD) Xenium datasets for Gja1, Gjb2, and Gjc3. Quadruplets are classified as shared, condition-specific, or indeterminate. (**b**) Bar plot showing the frequency of each cell type across the 60 quadruplets. Astrocytes and oligodendrocytes are most enriched in shared interactions, microglia dominate AD-specific interactions, and astrocytes are more frequent in WT-specific interactions. (**c**) Whisper networks of Astr1–Olig1 communication via homotypic Cx43 in WT (left) and AD (right), showing preserved symmetry between conditions. (**d**) Selected condition-specific quadruplets, showing z-scores in WT versus AD. Top three WT-specific and AD-specific interactions were chosen based on whisper network size. (**e**) Whisper network of Astr1–Ex th1 communication via homotypic Cx43 in WT. (**f**) Whisper network of Smc1–Smc1 (smooth muscle cell self-communication) via homotypic Cx43 in AD, highlighting potential vascular remodeling. (**g**) Aggregated microglial whisper network in AD, constructed by taking the union of all AD-specific quadruplets involving microglia, revealing enrichment in thalamus and hypothalamus regions.

463 We analyzed Xenium mouse brain data from wild-type and TgCRND8 mice[85] (aged >13 months,
464 Methods) using three annotated GJ genes. After cross-sample integration to harmonize cell-type
465 labels (Methods), CellWHISPER's Differential Analysis (DA) function quantified quadruplets per
466 condition, and classified them as shared, condition-specific or indeterminate (Methods). This
467 revealed 60 high-confidence quadruplets (z-score > 4, whisper network size > 30, **Fig. 6a, Supp.**
468 **Data 10**). The twenty shared quadruplets are enriched for astrocytic (13/20) and oligodendrocytic
469 communication (9/20) (**Fig. 6b**) including Astr1-Olig1 homotypic Cx43 (**Fig. 6c**). Despite known
470 Cx43 upregulation near Aβ plaques, prior work suggests structural preservation of astrocytic GJC
471 in AD brain[86]; consistent with this, CellWHISPER identifies conserved GJC mechanisms despite
472 change in signaling gene expression.
473
474 DA further identified 15 TgCRND8-specific and 14 wild-type-specific quadruplets (Fig. 6a,d,
475 **Supp. 10**), including reduced astrocyte-thalamic excitatory neuron GJC via homotypic GJA1 (**Fig.**
476 **6e**). The corresponding whisper network (Fig. 6e) shows enrichment of this communication in the
477 WT thalamus. In contrast, TgCRND8-specific mechanisms included increased smooth muscle
478 cells homotypic GJA1 (Fig. 6d,f), potentially reflecting vascular remodeling[87,88]. TgCRND8-
479 specific mechanisms were enriched for microglial interactions (5/17 vs. 0/23 in WT-specific, p-
480 value < 0.01, two-tailed proportions test), and WT-specific mechanisms to be marginally enriched
481 for astrocytes (6/23 vs. 2/17 in TgCRND8) (**Fig. 6b**). These trends align with existing literature:
482 astrocyte networks maintain homeostasis in healthy brain[89], while microglial GJC becomes more
483 prominent during AD neuroinflammation[31,90,91]. Aggregated microglial whisper network (**Fig. 6g**)
484 showed enrichment in thalamus and hypothalamus. Connexin alterations have been reported in
485 thalamic regions in AD[82] while microglial changes in thalamus are linked to cognitive impairment
486 after cortical brain injury[92].
487
488
489 **DISCUSSION**
490 Direct cell–cell communication shapes tissue homeostasis and disease, yet existing spatial
491 transcriptomics tools conflate genuine signaling with cell type co-localization, limiting their utility
492 for rigorous inference. CellWHISPER addresses this by explicitly accounting for confounding
493 effects of gene-expression and spatial organization, CellWHISPER markedly reduces FPR
494 relative to competing tools. Two additional computational features broaden practicality and
495 interpretability: an analytical exact null that avoids explicit permutation testing and enables
496 scalable inference across large tissues and signaling compendia, and a quadruplet-native latent

variable model that extracts higher-order structure without collapsing interactions into information-losing pairwise summaries. A differential analysis module further extends the same testing framework to multi-condition settings to identify shared versus condition-associated mechanisms. The framework is agnostic to the signaling compendium (connexins, curated LR pairs, or user-defined contact genes) and scales to tens of thousands of cells, enabling routine application to emerging high-resolution spatial platforms.

We demonstrated CellWHISPER on mouse brain spatial datasets as validation settings spanning two major classes of direct CCC: gap junctions and contact-dependent ligand–receptor interactions. CellWHISPER recovered canonical glial networks and generated a large-scale map of predicted connexin-mediated interactions ("connexin code"). Immunostaining validated Cx43 interactions, demonstrating its colocalization at astrocyte–astrocyte/endothelial/microglial interfaces. LR analysis uncovered adhesion/synaptic programs, including the mossy fiber pathway and a glial–meningeal signaling axis. Differential analysis of an AD mouse model identified a shift toward microglial-associated GJC and region-specific remodeling in the thalamus and vasculature, consistent with emerging models of neuroinflammatory progression.

Several limitations motivate future work. CellWHISPER inference relies on mRNA as a proxy for membrane proteins, and post-transcriptional regulation or subcellular localization can yield false positives . It prioritizes specificity through a confounder-aware null model, which may come at the cost of sensitivity. In contrast, CellWHISPER excels at detecting communication events mediated by a smaller number of cell pairs. Extensions to lower-resolution platforms (e.g. Visium) may be possible via deconvolution approaches, and integrating pathway information could improve robustness of whisper networks. Finally, segmentation errors in ST[93] may propagate to CCC inference, a limitation shared across current methods.

In summary, CellWHISPER establishes a robust paradigm for inferring direct signaling from ST data, combining stringent statistical control with interpretable latent structure, to enable mechanistic dissection of tissue organization and for identifying disease-specific communication in development, homeostasis, and disease.

METHODS

Overview of CellWHISPER

CellWHISPER infers direct cell–cell communication (CCC) from spatial transcriptomics (ST) data. The input consists of gene expression profiles, spatial coordinates, and cell-type annotations for individual cells, together with a precompiled compendium of signaling gene (SG) pairs, such as connexins or ligand–receptor (LR) pairs. SG compendia can be obtained from existing resources or provided by the user as a custom list.

For each candidate combination of a cell-type pair (ct_i, ct_j) and a signaling gene pair (sg_k, sg_l), CellWHISPER tests whether the two cell types exhibit evidence of direct communication mediated by the signaling genes (Fig. 1b). Statistically supported interactions are reported as CT–SG quadruplets, each associated with a z-score quantifying significance (Fig. 1c). CellWHISPER includes three analysis modules (Fig. 1c).

(i) A statistical testing module identifies significant CT–SG quadruplets.

(ii) A latent variable model (LVM) summarizes the collection of significant quadruplets into interpretable preference and similarity maps, including CT–CT, CT–SG, and SG–SG relationships.

(iii) An optional differential analysis (DA) module identifies condition-specific or shared communication mechanisms across biological conditions.

Whisper network construction

Gene expression was binarized using a quantile-based threshold (75th percentile by default), yielding binary indicators of whether a signaling gene is expressed in a given cell.

Spatial proximity between cells was represented using a k-nearest neighbors (KNN) graph constructed from cell coordinates (k = 5 unless otherwise specified). This graph defines the set of potential physical contacts used for CCC inference. For datasets with irregular tissue geometry, distance-based thresholds should be used instead of fixed neighbor counts.

For each CT–SG quadruplet (ct_i, ct_j, sg_k, sg_l), a whisper network was derived by retaining only edges connecting neighboring cells of types ct_i and ct_j that express sg_k and sg_l, respectively (Fig. 1b). The number of edges in this subgraph, denoted N_{ijkl}, quantifies spatially supported evidence for the corresponding communication mechanism.

Statistical test for CT–SG quadruplets

To assess statistical significance, CellWHISPER constructs a null model by permuting cell identities within each cell type. This preserves cell-type–specific gene expression distributions and spatial organization while disrupting spatial proximity between signaling cells of different

types. The whisper network is recomputed under the null, yielding a null distribution for N_{ijkl}, from which a z-score is computed.

To enable scalability to large datasets, CellWHISPER derives an analytical approximation to the null distribution, avoiding explicit permutations while preserving control of confounding effects arising from expression heterogeneity and spatial cell-type structure.

Analytical approach for Statistical test of a quadruplet:

Let Z_u be a Boolean random variable indicating the presence of a whisper edge for a proximal cell pair corresponding to a quadruplet i, j, k, l. Our objective is to estimate the expectation and variance of the total number of such whisper edges: $E(\sum_u Z_u), Var\ (\sum_u Z_u)$.The probability that a specific edge u is present (i.e. $Z_u = 1$) is modeled as the product of probabilities that gene k is expressed in cell of type i and gene l is expressed in a cell of type j. Gene expression is determined by a global percentile threshold across all cells. Thus:

$$P(Z_u = 1) = p_i p_j$$

Where p_i is the probability that a cell of type i expresses gene k, and p_j is the probability that a cell of type j expresses gene l.

If Z_u are Independent (i.i.d.):

Assuming the Z_u are i.i.d. Bernoulli variables, the total follows a Binomial distribution with N_u WHISPER edges:

$$E\left(\sum_u Z_u\right) = \sum_u P(Z_u) = N_u P(Z_u = 1) = N_u p_i p_j$$

$$Var\left(\sum_u Z_u\right) = N_u P(Z_u = 1)\big(1 - P(Z_u = 1)\big) = N_u p_i p_j (1 - p_i p_j)$$

Accounting for Dependence Between Edges:

However, the Z_u variables are not independent because edges may share a common cell. While the expectation remains unchanged due to linearity, the variance must account for dependencies:

$$Var\left(\sum_u Z_u\right) = E\left(\left(\sum_u Z_u\right)^2\right) - \left(E\left(\sum_u Z_u\right)\right)^2$$

$$= E\left((\sum_u Z_u)^2 \right) - N_u^2 p_i^2 p_j^2$$

To expand the first term,

$$E\left((\sum_u Z_u)^2 \right) = E\left(\sum_u Z_u^2 \right) - E(\sum_{u,v} Z_u Z_v)$$

Since, $Z_u \in [\{0,1\}]$, $Z_u^2 = Z_u$, so the first sum reduces to $E(\sum_u Z_u) = N_u p_i p_j$. The second term depends on the structure of edge overlap and has the following cases (assuming $i \neq j$):

- Case1: No shared cells between u and v , therefore Z_u, Z_v are independent, $E(Z_u Z_v) = p_i^2 p_j^2$
- Case 2: Shared cell of type j but distinct cells of type i, therefore, $E(Z_u Z_v) = p_i p_j^2$
- Case 3: Shared cell of type i but distinct cells of type j, therefore $E(Z_u Z_v) = p_i^2 p_j$

If $i = j$, cases 2 and 3 collapse into a single case due to symmetry.

Stereoseq Data and Analysis

Mouse brain STEREO-seq data[28] were obtained from the Spatial Omics Database[94]. After quality control filtering, the dataset comprised 50,140 cells and 18,527 genes, including 10 gap junction (connexin) genes: *Gjd2* (Cx36), *Gja5* (Cx40), *Gjc1* (Cx45), *Gja4* (Cx37), *Gjb2* (Cx26), *Gja1* (Cx43), *Gjb6* (Cx30), *Gjb1* (Cx32), *Gjc2* (Cx47), and *Gjc3* (Cx29). We used the authors' annotated cell-type labels, encompassing 28 distinct cell types. Abbreviations include: CA3, cornu ammonis area 3; Mb, midbrain; EX, excitatory glutamatergic neuron; IN, GABAergic interneuron; DA, dopaminergic neuron; GN DG, granule cell of dentate gyrus; Astr, astrocyte; Micro, microglia; OPC, oligodendrocyte precursor cell; Oligo, oligodendrocyte; SMC, smooth muscle cell; Ery, erythrocyte; and Endo, endothelial cell.

Gene expression matrices were normalized and subsequently denoised using MAGIC imputation (t=2) to recover continuous expression profiles. To construct the *whisper network*, a gene was considered "expressed" in a given cell if its expression exceeded the 75th percentile of that gene's expression across all cells. For gap-junction analysis, we evaluated all possible connexin–connexin combinations. Although several connexins are known to preferentially form specific homo- or heterotypic channels, experimental evidence for many pairings remains incomplete or context-dependent; thus, interaction patterns were inferred empirically from the data.

Ligand–receptor database

We used the CellChat v2[22] ligand–receptor compendium, focusing specifically on pairs annotated as cell–cell contact interactions. After intersecting this database with the gene set retained from the STEREO-seq dataset, 367 ligand–receptor pairs were included for downstream analyses.

Randomization and False Positive Rate estimation:

We performed randomization by permuting cell IDs within each cell type. This procedure preserves the marginal spatial distribution and gene expression profiles specific to each cell type. False positive rate (FPR) is obtained by comparing the number of detected pairs obtained on randomized data with number of detected pairs on real data.

Benchmarking

We could not run Spatalk on full mouse hemibrain dataset containing 50K cells. Therefore, we selected part of hippocampal region (Supp. Fig. 5) containing 5K cells for benchmarking analysis in Fig. 1d. We randomized by permuting cell IDs within each cell type.

All comparison methods were downloaded from their official repositories and executed using recommended parameters. For CellChat v2, we used the computeCommunProb function with default parameters (type = "triMean", trim = 0.1, interaction.range = 250, contact.knn.k = 5), followed by filterCommunication with a minimum of 20 cells per cluster. For COMMOT, we applied commot.tl.cluster_communication to compute cell–cell communication for each ligand–receptor pair. For SpaTalk, we used the default parameters.

Connexin mRNA-protein concordance

CellWHISPER relies on cellular mRNA abundance as a proxy for membrane-bound connexin proteins. To evaluate this assumption, we performed targeted spatial profiling of connexin 43 (Cx43) protein and its corresponding mRNA (Gja1) at single-molecule resolution in an induced pluripotent stem cell (iPSC) model, where gap junction communication is critical for spatial patterning.

We observed a significant correlation between Gja1 mRNA expression and average cytosolic Cx43 protein intensity at the single-cell level (Spearman's $R = 0.47$, $p = 5 \times 10^{-20}$; Supplementary Fig. 1). Consistently, whisper networks inferred from Gja1 mRNA showed strong overlap with those inferred from Cx43 protein measurements (Fig. 1e, Supp. Fig. 1), supporting the use of mRNA as a proxy for connexin-mediated communication in this context.

Latent Variable Model (LVM):

LVM stipulates that the binary variable $I(ct_i, ct_j, g_k, g_l)$ indicating whether the CT-SG quadruplet ct_i, ct_j, g_k, g_l has statistical support in the data follows a Bernoulli distribution:

$$I(ct_i, ct_j, g_k, g_l) \sim Ber\ (p_{ijkl})$$

where u_i , u_j denote to-be-learnt low-dimensional representations (embeddings) for CTs ct_i , ct_j and v_k , v_l denote embeddings for SGs g_k, g_l, respectively, all embeddings being in a common latent space. The dot product $u_i . v_k$ captures the general propensity of ct_i for participating in CCC using g_k. Thus, the model posits that the quadruplet ct_i, ct_j, g_k, g_l is more likely if there is a higher tendency for ct_i communicating via g_k and ct_j communicating via g_l.

Using this formulation CellWHISPER learn embeddings that enable quantification of (1) cell type-signaling gene preferences ($u_i . v_k$), (2) cell type similarities ($u_i . u_j$), i.e., which pairs of CTs have similar communication mechanisms, and (3) signaling gene similarities ($v_k . v_l$), i.e., which SGs are used by the same cell types. It uses cosine similarity rather than dot products between embeddings to emphasize specificity and mitigate abundance bias inherent in dot product.

In the null model, p_{ijkl} is assumed to be constant for all indices i, j, k, l. In contrast, the alternate model factorizes p_{ijkl} into two terms:

- the propensity of cell type i to communicate via signaling gene k, and
- the propensity of cell type j to communicate via signaling gene l

Formally, this is expressed as:

$$p_{ijkl} = (u_i . v_k)(u_j . v_l)$$

where u_i and u_j are low dimensional embeddings for cell types i, j, and v_k, v_l are embeddings for signaling genes k and l. All embeddings lie in a shared latent space, and we enforce the constraints:

$$0 < u_{ir} < 1, \qquad 0 < v_{ir} < 1, \qquad \left|\left|u_{ir}\right|\right|_1 < 1, \left|\left|v_{kr}\right|\right|_1 < 1$$

This formulation is invariant under simultaneous permutation of both CT and SG pairs, i.e.

$$I(ct_i, ct_j, g_k, g_l) = I(ct_j, ct_i, g_l, g_k) \Rightarrow p_{ijkl} = p_{jilk}$$

We optimize the model using constrained log-likelihood maximization, implemented via `scipy.optimize` using the SLQSP method. The constraint is posed using the L1 norm of the embeddings that ensures that the problem is a linear optimization and probabilities $0 < p_{ijkl} < 1$.

To confirm identifiability and recovery, we simulated Gap Junction Communication results with the same orders as the STEREO-seq analysis (29 cell types, 10 connexins, d=3). Ground-truth embeddings were drawn from numpy.random.rand with unit norm cell type and signaling gene embeddings. The optimizer recovered CT–CT similarity, SG–SG similarity, and CT–SG preference structure with high agreement (Fig. 5b).

For ligand–receptor analysis, only a subset of ligand–receptor pairs are biologically plausible. To ensure model learnability, we explicitly set the probability $p_{ijkl} = 0$ for incompatible ligand–receptor combinations. For computational feasibility, we randomly subsample a small number of such incompatible pairs (2000 in our experiments) to include as negative examples during optimization. This approach enables efficient training while maintaining biological specificity in the learned embeddings.

Differential Analysis:

We performed Differential Analysis (DA) at the terminal time point at which amyloid pathology is well established. Wild-type (WT) and TgCRND8 (AD) Xenium datasets were jointly integrated using residualPCA to achieve conservative alignment across conditions while preserving biological differences. A joint embedding was constructed and cell-type clustering was performed on the integrated space yielding 25 clusters.

Label transfer from reference. Cell-type labels were transferred from the annotated STEREO-seq reference (analyzed above). For this step, the WT, AD, and STEREO-seq datasets were co-embedded with reciprocal PCA[95]. For each query cell (WT or AD), we identified its k = 7 nearest neighbors in the reference and recorded their labels. For each cluster in the joint WT+AD

724 embedding, neighbor votes were aggregated across all member cells and normalized by the total
725 number of votes in that cluster to yield a label-probability vector. The cluster label was assigned
726 as the cell type with the maximum normalized vote. Clusters containing multiple subpopulations
727 of the same broad class (e.g., oligodendrocytes) were further subdivided and relabeled as
728 subtypes (e.g., Olig-1, Olig-2).
729
730 DA quantification. Using these harmonized annotations, CellWHISPER quantified CT–SG
731 quadruplets separately in WT and AD datasets. Quadruplets were retained if they satisfied
732 significance thresholds (*z*-score > 4 and whisper-network size > 30). This procedure yielded 60
733 high-confidence quadruplets across the two conditions. A quadruplet was classified as shared if
734 both WT and AD satisfied these thresholds, condition-specific if one condition satisfied the
735 thresholds while the other did not (*z*-score < 3 or whisper-network size < 20), and indeterminate
736 otherwise.
737
738 **Experimental details for GJA1 mRNA-protein correlation**
739 **iPSC culture**
740 iPSCs were obtained and cultured according to Culture and Freezing Methods for WTC Derived
741 AICS hiPSC Lines: https://www.coriell.org/0/PDF/Allen/iPSC/AICS_SOP_WTC_CellCulture.pdf.
742 Coverslips (22 × 22 mm) were placed flat in sterile 6-well plates and coated with growth factor–
743 reduced, phenol-red–free Matrigel (Corning, cat. no. 356231) at a final protein concentration of
744 0.337 mg/mL. Briefly, Matrigel aliquots thawed at 4 °C were diluted on ice in cold DMEM/F12 to
745 1.5 mL per well, mixed thoroughly with pre-chilled pipettes to ensure homogeneity, and
746 immediately dispensed to fully cover each coverslip. Plates were incubated at room temperature
747 for 1 h, after which excess coating solution was aspirated and wells were gently rinsed with room-
748 temperature mTeSR1 medium supplemented with 1 μM Y-27632 (ROCK inhibitor) before cell
749 seeding. Coated coverslips were used for experiments within 6 h or, if needed, sealed with
750 Parafilm and stored at 4 °C for up to 2 weeks. Once cells were confluent, they were fixed in 4%
751 formaldehyde for staining.
752
753 **iPSC RNA and protein multiplexing**
754 Fixed cells on coverslip with permeabilized with 70% EtOH for 30 mins at -20 C. Cycle 1 RNA
755 was labeled using Molecular Instruments hybridized chain reaction RNA fluorescence in situ
756 hybridzation (HCR RNA FISH) v3.0 probes: Human Cx43 B3 555, Human MALAT1 B2 647. All
757 probes were amplified 75 mins at room temperature. Fluorescence images were acquired on

758 adapted Nikon Ti 2000-U microscope with appropriate filters. Then, samples were blocked and
759 incubated with Cx43 Ab (Invitrogen 71-0700, 1;150) overnight at 4 C followed by secondary
760 antibody staining (anti-rabbit 488) for 1 hr at room temperature. This Cycle 2 Cx43 Ab was imaged
761 again. DAPI counterstain was used to register the images from both cycles. Images were
762 segmented with Cellpose and quantified for downstream analysis.
763
764 **Immunofluorescence staining**
765 Mouse brain sections were fixed in 4% paraformaldehyde (PFA) and sectioned using Slides were
766 deparaffinized and rehydrated by subsequent rinses in Xylene, 100% EtOH, 95% EtOH, 80%
767 EtOH, 70% EtOH and deionized water (diH_2O). Then we performed antigen retrieval by placing
768 slides in a Coplin jar with 10% sodium citrate buffer in 1x phosphate-buffered saline (PBS). Coplin
769 jars were placed in a pressure cooker with 300 mL of tap water and set on high pressure for 15
770 minutes. The slides were let to cool for 20 minutes and then washed with diH_2O. They were then
771 blocked using 50 mg/mL bovine serum albumin (BSA) in 1x PBS. They were then washed with
772 .1% Tween 20 in 1x PBS (PBS-T). Primary antibodies were diluted based on manufacturers'
773 recommendation in 5% BSA and left to incubate on slides at 4°C, overnight. A list of antibodies
774 and their dilutions are below. Slides were washed the next day with PBS-T. For unconjugated
775 primary antibodies, slides were washed with three times with PBS-T for 2 minutes and incubated
776 in 1:5000 secondary antibody in PBS for 1 hour at room temperature. Slides were washed were
777 washed three times with PBS-T and counterstained by a 1:500 Dapi in 1x PBS solution for 10
778 minutes. Slides were washed with PBS-T for 2 minutes and nuclear staining was done using
779 1:2000 DAPI (Thermo Scientific, Ref:62248). Slides were mounted using 10% glycerol in 1x PBS.
780 Coverslips were placed over top of the slides, and the edges were coated with nail polish to
781 prevent dehydration. Slides were imaged using the Keyence BZ-X800 fluorescence microscope.
782
783 **Image preprocessing and quantification of colocalization**
784 Images of cells in tissue all underwent the following preprocessing before further analysis. Each
785 image contains multiple channels. The maximum intensity projection of each channel was then
786 generated. The projected image was the final image used for cellular counting analysis. For the
787 DAPI channel, the maximum intensity projection was generated from the raw image. For
788 visualization consistency, brightness and contrast of each fluorescent channel were adjusted
789 uniformly across all images using the "Brightness/Contrast" tool without altering the underlying
790 pixel intensity values. To measure intercellular colocalization, the straight-line tool was used to
791 draw a line between the centers of identified cells of interest (e.g., nuclei or marker-specific

792 regions). The length of each line was recorded using the “Measure” function, which reports
793 distances in calibrated units (μm) based on the microscope’s pixel-to-micrometer conversion. At
794 least 629 cells per field were analyzed, and distances were exported for downstream statistical
795 analysis. Previous steps were conducted in ImageJ/Fiji (National Institutes of Health). Statistical
796 significance was calculated using a regular two-way ANOVA in *BioRender Graph*. Statistical
797 levels are provided as follows: ns for $0.05 < p \leq 1$, * for $0.01 < p \leq 0.05$, ** for $0.001 < p \leq 0.01$, ***
798 for $0.0001 < p < 0.001$, and **** for $p \leq 0.0001$.
799
800 **Code Availability**
801 The code is available at https://github.com/anurendra/CellWHISPER.
802
803 **Data Availability**
804 Mouse brain STEREO-seq data[28] were obtained from the Spatial Omics Database[94]. Our iPSC-
805 derived and mouse brain immunostaining data will be made available upon publication. Xenium
806 *TgCRND8* Alzheimer’s-disease mouse-model data were obtained from wild-type and *TgCRND8*
807 mice[85], as provided by 10x Genomics. Only the final time point (age > 13 months) was used for
808 analysis.
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825

REFERENCES

1. Armingol, E., Officer, A., Harismendy, O. & Lewis, N.E. Deciphering cell–cell interactions and communication from gene expression. *Nature Reviews Genetics* **22**, 71-88 (2021).
2. Wilk, A.J., Shalek, A.K., Holmes, S. & Blish, C.A. Comparative analysis of cell–cell communication at single-cell resolution. *Nature biotechnology* **42**, 470-483 (2024).
3. Dimitrov, D. *et al.* LIANA+ provides an all-in-one framework for cell–cell communication inference. *Nature Cell Biology* **26**, 1613-1622 (2024).
4. Dimitrov, D. *et al.* Comparison of methods and resources for cell-cell communication inference from single-cell RNA-Seq data. *Nature communications* **13**, 3224 (2022).
5. Jin, S. *et al.* Inference and analysis of cell-cell communication using CellChat. *Nature communications* **12**, 1088 (2021).
6. Efremova, M., Vento-Tormo, M., Teichmann, S.A. & Vento-Tormo, R. CellPhoneDB: inferring cell–cell communication from combined expression of multi-subunit ligand–receptor complexes. *Nature protocols* **15**, 1484-1506 (2020).
7. Browaeys, R., Saelens, W. & Saeys, Y. NicheNet: modeling intercellular communication by linking ligands to target genes. *Nature methods* **17**, 159-162 (2020).
8. Hervé, J.-C. & Derangeon, M. Gap-junction-mediated cell-to-cell communication. *Cell and tissue research* **352**, 21-31 (2013).
9. Goodenough, D.A. & Paul, D.L. Gap junctions. *Cold Spring Harbor perspectives in biology* **1**, a002576 (2009).
10. Battefeld, A., Klooster, J. & Kole, M.H. Myelinating satellite oligodendrocytes are integrated in a glial syncytium constraining neuronal high-frequency activity. *Nature Communications* **7**, 1-13 (2016).
11. Rash, J.E. Molecular disruptions of the panglial syncytium block potassium siphoning and axonal saltatory conduction: pertinence to neuromyelitis optica and other demyelinating diseases of the central nervous system. *Neuroscience* **168**, 982-1008 (2010).
12. Glen, C.M., McDevitt, T.C. & Kemp, M.L. Dynamic intercellular transport modulates the spatial patterning of differentiation during early neural commitment. *Nature communications* **9**, 4111 (2018).
13. Harris, A. & Locke, D. *Connexins: a guide*, (Springer Science & Business Media, 2008).
14. Nakase, T. & Naus, C.C. Gap junctions and neurological disorders of the central nervous system. *Biochimica Et Biophysica Acta (BBA)-Biomembranes* **1662**, 149-158 (2004).
15. Jongsma, H.J. & Wilders, R. Gap junctions in cardiovascular disease. *Circulation research* **86**, 1193-1197 (2000).
16. Koval, M., Molina, S.A. & Burt, J.M. Mix and match: investigating heteromeric and heterotypic gap junction channels in model systems and native tissues. *FEBS letters* **588**, 1193-1204 (2014).
17. Stephan, J., Eitelmann, S. & Zhou, M. Approaches to study gap junctional coupling. *Frontiers in Cellular Neuroscience* **15**, 640406 (2021).
18. Zhou, B. *et al.* Notch signaling pathway: architecture, disease, and therapeutics. *Signal transduction and targeted therapy* **7**, 95 (2022).
19. Kania, A. & Klein, R. Mechanisms of ephrin–Eph signalling in development, physiology and disease. *Nature reviews Molecular cell biology* **17**, 240-256 (2016).
20. Nakandakari-Higa, S. *et al.* Universal recording of immune cell interactions in vivo. *Nature* **627**, 399-406 (2024).
21. Mattes, B. & Scholpp, S. Emerging role of contact-mediated cell communication in tissue development and diseases. *Histochemistry and cell biology* **150**, 431-442 (2018).

22. Jin, S., Plikus, M.V. & Nie, Q. CellChat for systematic analysis of cell–cell communication from single-cell transcriptomics. *Nature Protocols* **20**, 180-219 (2025).
23. Troulé, K. *et al.* CellPhoneDB v5: inferring cell–cell communication from single-cell multiomics data. *Nature Protocols*, 1-29 (2025).
24. Shao, X. *et al.* Knowledge-graph-based cell-cell communication inference for spatially resolved transcriptomic data with SpaTalk. *Nature communications* **13**, 4429 (2022).
25. Cang, Z. *et al.* Screening cell–cell communication in spatial transcriptomics via collective optimal transport. *Nature methods* **20**, 218-228 (2023).
26. Cesaro, G. *et al.* Advances and challenges in cell–cell communication inference: a comprehensive review of tools, resources, and future directions. *Briefings in Bioinformatics* **26**, bbaf280 (2025).
27. Zohora, F.T. *et al.* NEST: Spatially-mapped cell-cell communication patterns using a deep learning-based attention mechanism. *bioRxiv*, 2024.03. 19.585796 (2024).
28. Chen, A. *et al.* Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell* **185**, 1777-1792. e21 (2022).
29. Evans, W.H. & Martin, P.E. Gap junctions: structure and function. *Molecular membrane biology* **19**, 121-136 (2002).
30. Palla, G. *et al.* Squidpy: a scalable framework for spatial omics analysis. *Nature methods* **19**, 171-178 (2022).
31. Takeuchi, H. & Suzumura, A. Gap junctions and hemichannels composed of connexins: potential therapeutic targets for neurodegenerative diseases. *Frontiers in cellular neuroscience* **8**, 189 (2014).
32. Söhl, G., Maxeiner, S. & Willecke, K. Expression and functions of neuronal gap junctions. *Nature reviews neuroscience* **6**, 191-200 (2005).
33. Dere, E. & Zlomuzica, A. The role of gap junctions in the brain in health and disease. *Neuroscience & Biobehavioral Reviews* **36**, 206-217 (2012).
34. Dong, A., Liu, S. & Li, Y. Gap junctions in the nervous system: probing functional connections using new imaging approaches. *Frontiers in Cellular Neuroscience* **12**, 320 (2018).
35. Kajiwara, Y. *et al.* GJA1 (connexin43) is a key regulator of Alzheimer's disease pathogenesis. *Acta neuropathologica communications* **6**, 1-20 (2018).
36. Sánchez, O.F. *et al.* Role of connexins 30, 36, and 43 in brain tumors, neurodegenerative diseases, and neuroprotection. *Cells* **9**, 846 (2020).
37. Nualart-Marti, A., Solsona, C. & Fields, R.D. Gap junction communication in myelinating glia. *Biochimica et Biophysica Acta (BBA)-Biomembranes* **1828**, 69-78 (2013).
38. Papaneophytou, C., Georgiou, E. & Kleopa, K.A. The role of oligodendrocyte gap junctions in neuroinflammation. *Channels* **13**, 247-263 (2019).
39. Nagy, J.I., Ionescu, A.V., Lynn, B.D. & Rash, J.E. Connexin29 and connexin32 at oligodendrocyte and astrocyte gap junctions and in myelin of the mouse central nervous system. *Journal of Comparative Neurology* **464**, 356-370 (2003).
40. Kleopa, K.A., Orthmann, J.L., Enriquez, A., Paul, D.L. & Scherer, S.S. Unique distributions of the gap junction proteins connexin29, connexin32, and connexin47 in oligodendrocytes. *Glia* **47**, 346-357 (2004).
41. Orthmann-Murphy, J.L., Abrams, C.K. & Scherer, S.S. Gap junctions couple astrocytes and oligodendrocytes. *Journal of Molecular Neuroscience* **35**, 101-116 (2008).
42. Xing, L., Yang, T., Cui, S. & Chen, G. Connexin hemichannels in astrocytes: role in CNS disorders. *Frontiers in molecular neuroscience* **12**, 23 (2019).
43. Li, M. *et al.* Astrocytic Gap Junctions protein Cx43/Cx30 modulate EAAT1 and glutamate to mediate cerebral ischemia–reperfusion injury. *Brain Research* **1847**, 149306 (2025).

44. Nagy, J., Ionescu, A.V., Lynn, B. & Rash, J. Coupling of astrocyte connexins Cx26, Cx30, Cx43 to oligodendrocyte Cx29, Cx32, Cx47: Implications from normal and connexin32 knockout mice. *Glia* **44**, 205-218 (2003).
45. Orthmann-Murphy, J.L., Freidin, M., Fischer, E., Scherer, S.S. & Abrams, C.K. Two distinct heterotypic channels mediate gap junction coupling between astrocyte and oligodendrocyte connexins. *Journal of Neuroscience* **27**, 13949-13957 (2007).
46. Alvarez-Maubecin, V., García-Hernández, F., Williams, J.T. & Van Bockstaele, E.J. Functional coupling between neurons and glia. *Journal of Neuroscience* **20**, 4091-4098 (2000).
47. Srinivas, M. *et al.* Functional properties of channels formed by the neuronal gap junction protein connexin36. *Journal of Neuroscience* **19**, 9848-9855 (1999).
48. Lee, S.-N. *et al.* Cryo-EM structures of human Cx36/GJD2 neuronal gap junction channel. *Nature communications* **14**, 1347 (2023).
49. Nagy, J. & Rash, J. Cx36, Cx43 and Cx45 in mouse and rat cerebellar cortex: Species-specific expression, compensation in Cx36 null mice and co-localization in neurons vs. glia. *European Journal of Neuroscience* **46**, 1790-1804 (2017).
50. Feigenspan, A. *et al.* Expression of connexin36 in cone pedicles and OFF-cone bipolar cells of the mouse retina. *Journal of Neuroscience* **24**, 3325-3334 (2004).
51. Van Bockstaele, E., Garcia-Hernandez, F., Fox, K., Alvarez, V. & Williams, J. Expression of connexins during development and following manipulation of afferent input in the rat locus coeruleus. *Neurochemistry international* **45**, 421-428 (2004).
52. Kay, C.W., Ursu, D., Sher, E. & King, A.E. The role of Cx36 and Cx43 in 4-aminopyridine-induced rhythmic activity in the spinal nociceptive dorsal horn: an electrophysiological study in vitro. *Physiological Reports* **4**, e12852 (2016).
53. Zhou, M., Zhong, S. & Verkhratsky, A. Astrocyte syncytium: from neonatal genesis to aging degeneration. *Neural regeneration research* **19**, 395-396 (2024).
54. Li, Z.A. & Zhou, M. Astrocyte syncytium–a biopower grid system in the brain. *Journal of integrative neuroscience* **23**, 71 (2024).
55. Ma, B. *et al.* Gap junction coupling confers isopotentiality on astrocyte syncytium. *Glia* **64**, 214-226 (2016).
56. Serra, R. & Simard, J.M. Adherens, tight, and gap junctions in ependymal cells: A systematic review of their contribution to CSF-brain barrier. *Frontiers in neurology* **14**, 1092205 (2023).
57. Fabbiani, G. *et al.* Connexin signaling is involved in the reactivation of a latent stem cell niche after spinal cord injury. *Journal of Neuroscience* **40**, 2246-2258 (2020).
58. Bouillé, C., Mesnil, M., Barriere, H. & Gabrion, J. Gap junctional intercellular communication between cultured ependymal cells, revealed by lucifer yellow CH transfer and freeze-fracture. *Glia* **4**, 25-36 (1991).
59. De Bock, M. *et al.* Targeting gliovascular connexins prevents inflammatory blood-brain barrier leakage and astrogliosis. *Jci Insight* **7**, e135263 (2022).
60. Jiang, H., Zhang, Y., Wang, Z.-Z. & Chen, N.-H. Connexin 43: an interface connecting neuroinflammation to depression. *Molecules* **28**, 1820 (2023).
61. Johnson, A.M. *et al.* Connexin 43 gap junctions contribute to brain endothelial barrier hyperpermeability in familial cerebral cavernous malformations type III by modulating tight junction structure. *The FASEB Journal* **32**, 2615 (2018).
62. Eugenín, E.A. *et al.* Microglia at brain stab wounds express connexin 43 and in vitro form functional gap junctions after treatment with interferon-γ and tumor necrosis factor-α. *Proceedings of the National Academy of Sciences* **98**, 4190-4195 (2001).
63. Nakase, T., Fushiki, S. & Naus, C.C. Astrocytic gap junctions composed of connexin 43 reduce apoptotic neuronal damage in cerebral ischemia. *Stroke* **34**, 1987-1993 (2003).

64. Abudara, V. *et al.* Activated microglia impairs neuroglial interaction by opening C x43 hemichannels in hippocampal astrocytes. *Glia* **63**, 795-811 (2015).
65. Henze, D.A., Urban, N.N. & Barrionuevo, G. The multifarious hippocampal mossy fiber pathway: a review. *Neuroscience* **98**, 407-427 (2000).
66. Triplett, J.W. & Feldheim, D.A. Eph and ephrin signaling in the formation of topographic maps. in *Seminars in cell & developmental biology* Vol. 23 7-15 (Elsevier, 2012).
67. Yang, J.S., Wei, H.X., Chen, P.P. & Wu, G. Roles of Eph/ephrin bidirectional signaling in central nervous system injury and recovery. *Experimental and therapeutic medicine* **15**, 2219-2227 (2018).
68. Martinelli, D.C. *et al.* Expression of C1ql3 in discrete neuronal populations controls efferent synapse numbers and diverse behaviors. *Neuron* **91**, 1034-1051 (2016).
69. Newrzella, D. *et al.* The functional genome of CA1 and CA3 neurons under native conditions and in response to ischemia. *BMC genomics* **8**, 1-18 (2007).
70. Quarles, R.H. Myelin-associated glycoprotein (MAG): past, present and beyond. *Journal of neurochemistry* **100**, 1431-1448 (2007).
71. Gow, A. *et al.* CNS myelin and sertoli cell tight junction strands are absent in Osp/claudin-11 null mice. *Cell* **99**, 649-659 (1999).
72. Hutchison, M. *et al.* Genetic inhibition of neurotransmission reveals role of glutamatergic input to dopamine neurons in high-effort behavior. *Molecular psychiatry* **23**, 1213-1225 (2018).
73. Song, Y.-H., Yoon, J. & Lee, S.-H. The role of neuropeptide somatostatin in the brain and its application in treating neurological disorders. *Experimental & Molecular Medicine* **53**, 328-338 (2021).
74. Vandecasteele, M., Glowinski, J. & Venance, L. Electrical synapses between dopaminergic neurons of the substantia nigra pars compacta. *Journal of Neuroscience* **25**, 291-298 (2005).
75. Radnikow, G. & Feldmeyer, D. Layer-and cell type-specific modulation of excitatory neuronal activity in the neocortex. *Frontiers in neuroanatomy* **12**, 1 (2018).
76. Yan, J. & Gao, L. Organizing principles of the prefrontal cortex from a single-neuron projectome. Vol. 25 408-409 (NATURE PORTFOLIO HEIDELBERGER PLATZ 3, BERLIN, 14197, GERMANY, 2022).
77. Gulledge, A.T. & Kawaguchi, Y. Phasic cholinergic signaling in the hippocampus: functional homology with the neocortex? *Hippocampus* **17**, 327-332 (2007).
78. Sherman, S.M. Thalamus plays a central role in ongoing cortical functioning. *Nature neuroscience* **19**, 533-541 (2016).
79. Yuasa-Kawada, J., Kinoshita-Kawada, M., Tsuboi, Y. & Wu, J.Y. Neuronal guidance genes in health and diseases. *Protein & cell* **14**, 238-261 (2023).
80. Südhof, T.C. Synaptic neurexin complexes: a molecular code for the logic of neural circuits. *Cell* **171**, 745-769 (2017).
81. Lv, X. *et al.* Patterned cPCDH expression regulates the fine organization of the neocortex. *Nature* **612**, 503-511 (2022).
82. Angeli, S. *et al.* Altered expression of glial gap junction proteins Cx43, Cx30, and Cx47 in the 5XFAD model of Alzheimer's disease. *Frontiers in Neuroscience* **14**, 582934 (2020).
83. Maulik, M. *et al.* Amyloid-β regulates gap junction protein connexin 43 trafficking in cultured primary astrocytes. *Journal of Biological Chemistry* **295**, 15097-15111 (2020).
84. Pechlivanidou, M. *et al.* Glial gap junction pathology in the spinal cord of the 5xFAD mouse model of early-onset Alzheimer's disease. *International journal of molecular sciences* **23**, 15597 (2022).
85. Genomics, x. Xenium In Situ Analysis of Alzheimer's Disease Mouse Model Brain Coronal Sections from One Hemisphere Over a Time Course. (2024).

86. Yi, C. *et al.* Astroglial connexin43 contributes to neuronal suffering in a mouse model of Alzheimer's disease. *Cell Death & Differentiation* **23**, 1691-1701 (2016).
87. Kozberg, M.G. *et al.* Blood–brain barrier leakage and perivascular inflammation in cerebral amyloid angiopathy. *Brain Communications* **4**, fcac245 (2022).
88. Figueroa, X.F. & Duling, B.R. Gap junctions in the control of vascular function. *Antioxidants & redox signaling* **11**, 251-266 (2009).
89. Huang, X. *et al.* Astroglial connexins in neurodegenerative diseases. *Frontiers in molecular neuroscience* **14**, 657514 (2021).
90. Caruso, G., Di Pietro, L. & Caraci, F. Gap junctions and connexins in microglia-related oxidative stress and neuroinflammation: perspectives for drug discovery. *Biomolecules* **13**, 505 (2023).
91. Anwar, M.M. *et al.* Assessing the role of primary healthy microglia and gap junction blocker in hindering Alzheimer's disease neuroinflammatory type: early approaches for therapeutic intervention. *Frontiers in Neuroscience* **16**, 1041461 (2023).
92. Matoba, K. *et al.* FcγR-and CD9-dependent synapse engulfing microglia in the thalamus drives cognitive impairment following cortical brain injury. *bioRxiv*, 2024.09. 19.609743 (2024).
93. Yu, H. *et al.* Segmentation Matters: Cell Segmentation Challenge in Spatial Transcriptomics. *NEURON-D-25-01846*.
94. Yuan, Z. *et al.* SODB facilitates comprehensive exploration of spatial omics data. *Nature Methods* **20**, 387-399 (2023).
95. Stuart, T. *et al.* Comprehensive integration of single-cell data. *cell* **177**, 1888-1902. e21 (2019).