Preservation of co-expression defines the primary tissue fidelity of human neural organoids

Jonathan M. Werner[1], Jesse Gillis[1,2]

[1]The Stanley Institute for Cognitive Genomics, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA

[2]Physiology Department and Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, ON, Canada

Abstract:

Human neural organoid models offer an exciting opportunity for studying often inaccessible human-specific brain development; however, it remains unclear how precisely organoids recapitulate fetal/primary tissue biology. Here, we characterize field-wide replicability and biological fidelity through a meta-analysis of single-cell RNA-sequencing data for first and second trimester human primary brain (2.95 million cells, 51 datasets) and neural organoids (1.63 million cells, 130 datasets). We quantify the degree to which primary tissue cell-type marker expression and co-expression are recapitulated in organoids across 12 different protocol types. By quantifying gene-level preservation of primary tissue co-expression, we show neural organoids lie on a spectrum ranging from virtually no signal to co-expression near indistinguishable from primary tissue data, demonstrating high fidelity is within the scope of current methods. Additionally, we show neural organoids preserve the cell-type specific co-expression of developing rather than adult cells, confirming organoids are an appropriate model for primary tissue development. Overall, quantifying the preservation of primary tissue co-expression is a powerful tool for uncovering unifying axes of variation across heterogeneous neural organoid experiments.

Introduction:

Pluripotent stem cells create self-organized multi-cellular structures, termed organoids, when cultured in a 3D *in vitro* environment[1,2]. The advantage of organoid models over 2D cell culture counterparts is their ability to generate structures that resemble endogenous tissues both in the differentiated cell-types produced and their 3D spatial organization[3,4]. The ability to model organogenesis in a controlled *in vitro* environment creates opportunities to study previously inaccessible developmental tissues from both humans and a range of model organisms[5,6,7]. As such, organoids are genetically accessible[8] and environmentally perturbable[9] models enabling the study of molecular, cellular, and developmental mechanisms behind tissue construction. However, the applicability of studies in organoids to *in vivo* biology hinges on how well these *in vitro* models recapitulate primary tissue developmental processes, which remains an open question.

Quantifying the degree to which organoid systems replicate primary tissue biological processes is a critical step toward understanding the strengths and limitations of these *in vitro* models[10–14]. However, studies that perform such primary tissue/organoid comparisons are inherently confounded by batch[15] (*in vivo* vs *in vitro*), making it difficult to disentangle batch effects from underlying primary tissue and organoid biology. Meta-analytic approaches across many primary tissue and organoid datasets offer a route around these confounds, enabling the discovery of replicable primary tissue and organoid signatures independent of batch, which can then be interrogated for how well organoids recapitulate primary tissue biology[16–18]. An important biological signature for this purpose is gene co-expression[19]. Genes that are functionally related tend to be expressed together, resulting in correlated gene expression dynamics that can define functionally relevant gene modules[19]. Gene co-expression relationships represent a shared genomic space that can be aggregated across experiments (e.g.,[20]) in either *in vivo* or *in vitro* systems, thus providing a useful framework for quantifying functional similarities and differences. Excitingly, coupling meta-analytic comparisons of primary tissue and organoid co-expression with single-cell RNA-sequencing data (scRNA-seq) stands to deliver cell-type specific quantifications of organoids' current capacity for producing functionally equivalent cell-types to primary tissues[21,22].

Among organoid systems, human neural organoids are particularly well suited for meta-analytic evaluation due to well-described broad cell-type annotations and their known lineage relationships[23], the wide variety of differentiation protocols in use[24], and the increasing amount of single-cell primary brain tissue and neural organoid data publicly available. In particular, the diversity of differentiation protocols for human neural organoids poses a unique challenge for organoid quality control that can be met by meta-analytic approaches. Neural organoids can either be undirected[25] (multiple brain region identities) or directed (specific brain region identity) with an increasing number of protocols striving to produce a wider variety of region-specific organoids[11,26–37]. Meta-analytic primary tissue/organoid comparisons across differentiation protocols stand to derive generalizable quality control metrics applicable to any differentiation protocol, fulfilling a currently unmet need for unified quality control metrics across heterogeneous neural organoids.

Prior comparisons between primary brain tissues and neural organoids demonstrated that organoids have the capacity to produce diverse cell-types that capture both regional and temporal variation similar to primary tissue data as assayed through transcriptomic[10,11,13,16,17,38], epigenomic[39,40], electrophysiologic[41], and proteomic studies[42]. At the morphological level,

neural organoids can produce cellular organizations structurally similar to various *in vivo* brain regions, including cortical layers[43] and hippocampus[27], as well as modeling known inter-regional interactions like neuromuscular junctions[34] and interneuron migration[29]. Additionally, several prior studies have compared primary tissue/organoid co-expression and concluded that neural organoids recapitulate primary brain tissue co-expression[5,13,39], but these assessments are highly targeted to study-specific properties, limiting potential generalization or potential assessment across the field. Typically, only a single organoid differentiation protocol is used in these assessments and it remains unclear whether organoids across different protocols will produce similar results. This lack of breadth also affects the use of primary tissue data used as a reference, with the primary tissue datasets utilized being treated as gold-standard datasets with little consideration for the extent one primary tissue reference may generalize to another. While prior meta-analytic comparisons of primary tissue/organoid co-expression have been performed[17], these were done at the bulk level (lack cell-type resolution) and included a small number of cortical organoid protocols, limiting the biological resolution and generalizability of these findings.

In this study, we perform a meta-analytic assessment of primary brain tissue (2.95 million cells, 50 datasets, Fig. 1A) and neural organoid (1.63 million cells, 130 datasets, 12 protocols, Fig. 1B) scRNA-seq datasets, constructing robust primary tissue cell-type specific markers and co-expression to query how well neural organoids recapitulate primary tissue cell-type specific biology. We sample primary brain tissue data over the first and second trimesters and across 15 different developmentally defined brain regions, extracting lists of cell-type markers that define broad primary tissue cell-type identity regardless of temporal, regional, or technical variation (Fig. 1A). We derive co-expression networks from individual primary tissue and organoid datasets as well as aggregate co-expression networks across datasets (Fig. 1C). From these networks, we assess the strength of co-expression within primary tissue cell-type marker sets as well as the preservation of co-expression patterns between primary tissue and organoid data (Fig. 1D-E). We also provide an R package to download our primary tissue reference co-expression network and assay new neural organoid data using simple, meaningful, and fast statistics (Fig. 1F). By constructing robust primary tissue cell-type representations through meta-analytic approaches, we demonstrate the preservation of primary tissue cell-type co-expression provides both specific and generalizable characterization of the primary tissue fidelity of human neural organoids.

Results:

Meta-analytic framework for primary tissue/organoid comparisons

We reason that, if they exist, primary tissue cell-type specific signals robust to temporal, regional, and technical variation will constitute *in vivo* standards applicable to any organoid dataset regardless of time in culture or differentiation protocol. We first show it is possible to learn sets of marker genes that define broad primary tissue cell-types (Fig. 1A, Supp. Table 1) across timepoints (gestational weeks GW5-GW25) and brain regions (15 developmentally defined brain regions) through a meta-analytic differential expression framework (Fig. 1A, Fig. 2A-B). We then compare co-expression within these marker sets between primary tissue and organoid data to quantify the degree organoids preserve primary tissue cell-type specific co-expression. An important aspect of our analysis is our cross-validation of primary tissue differential expression and co-expression. We employ a leave-one-out cross-validation approach when learning robust differentially expressed marker genes from our annotated primary tissue datasets (2,174,934 cells, 37 datasets) and we interrogate co-expression of our primary tissue marker genes within a large cohort of unannotated primary tissue datasets (776,343 cells, 14 datasets). This approach ensures we are extracting primary tissue markers and co-expression relationships independent of temporal, regional, and technical variation, a powerful approach for deriving broad primary tissue signatures appropriate for comparison to a wide range of organoid datasets.

Cross-temporal and -regional primary tissue cell-type markers

To learn markers that define broad primary tissue cell-types, we apply the MetaMarkers[44] framework to our cross-temporal and -regional annotated primary tissue datasets (Fig. 2A-B). MetaMarkers uses robust differential expression statistic thresholds (log2 fold-change >= 4 and FDR-adjusted p-value <= 0.05) for determining whether a gene is differentially expressed (DE) within individual datasets, then ranks all genes via the strength of their recurrent DE across datasets (see methods). We test the generalizability of our primary tissue MetaMarker gene sets in predicting primary cell-types by employing a leave-one-out primary tissue cross-validation (Fig. 2A-B). We construct an aggregate expression predictor in the left-out dataset using MetaMarkers learned from the remaining datasets (see methods), quantifying how well the MetaMarker gene sets predict the left-out cell-type annotations with the area-under-the-receiver-operating-characteristic curve statistic (AUROC, Fig. 2B-C). The AUROC is the probability of correctly prioritizing a true positive (e.g., cell of the right type) above a negative, (e.g., cell of the wrong type), given some predictor of the positive class, in this case, aggregate cell-type marker expression.

Starting with just the top 10 primary tissue MetaMarkers per cell-type, we achieve a mean AUROC across all primary tissue datasets of 0.944 ± 0.0280 SD, 0.865 ± 0.0653 SD, 0.873 ± 0.0676 SD, 0.937 ± 0.0669 SD, 0.879 ± 0.0535 SD, and 0.863 ± 0.0768 SD, for dividing progenitors, neural progenitors, intermediate progenitors, GABAergic neurons, glutamatergic neurons, and non-neuronal cell-types respectively (Fig. 2C). These extremely high performances demonstrate that even a small number of meta-analytically derived primary tissue cell-type markers have high utility in predicting primary tissue cell-type annotations regardless of temporal and regional variability. For all following analysis, we take the top 100 MetaMarkers per cell-type as robust representations of our 6 broad primary tissue cell-type annotations (average AUROC >= 0.90 except for intermediate progenitors: 0.897 ± 0.0777

SD), with the 100 MetaMarkers achieving modest increases in performance over the top 10 MetaMarkers for all cell-types except GABAergic cells (Fig. 2C, mean AUROC for 100 GABAergic MetaMarkers: 0.922 ± 0.0777 SD). When comparing MetaMarkers to markers derived from individual primary tissue datasets, we find the MetaMarkers are consistently top performers in predicting primary tissue annotations (Fig. 2D), with MetaMarkers producing the top results for intermediate progenitors, glutamatergic neurons, and GABAergic neurons (Supp. Fig. 1), as well as comparable performance to top individual datasets for dividing progenitors, neural progenitors, and non-neuronal cell-types (Supp. Fig. 1).

We explore the primary tissue MetaMarker sets further by computing the average expression of the top 100 MetaMarkers for our 6 annotated cell-types across all cells within our 37 annotated primary tissue datasets (Fig. 2E), continuing our leave-one-out approach. Each annotated primary tissue cell-type expresses the corresponding matched MetaMarker set over all other MetaMarker sets, with the exception of some off-target expression for the neural progenitor MetaMarkers in dividing progenitors and non-neurons (aggregated over all datasets Fig. 2E, individual datasets Supp. Fig. 1B). This demonstrates our MetaMarker gene sets act as robust cell-type markers in aggregate across all first and second trimester timepoints (Fig. 2E, Supp. Fig. 1B). Additionally, we investigate the expression of the top 100 MetaMarker gene sets across annotated primary brain regions, demonstrating each primary tissue cell-type maximally expresses the corresponding primary tissue MetaMarker set across all annotated brain regions (Supp. Fig. 2A-B). Overall, we are able to meta-analytically extract cell-type markers that define broad primary tissue cell-types independent of temporal and regional variation.

Broad primary tissue cell-type markers capture organoid temporal variation

After extracting meta-analytic cell-type markers that capture broad primary tissue temporal and regional variation, we can test how well these markers also capture organoid temporal and regional (protocol) variation. We start with a large-scale temporal organoid atlas[38] derived from a forebrain differentiation protocol containing timepoints ranging from 23 days to 6 months in culture. When comparing primary tissue and organoid data along a temporal axis, one might expect younger primary tissue expression data to be a better reference for younger organoid cell-types (better able to predict cell-types) and vice-versa for older primary and organoid data (Supp. Fig. 3A). We test this relationship using the same AUROC quantification as in Figure 1C, but now using the top 100 primary tissue cell-type markers per primary tissue dataset to predict organoid cell-type annotations across all organoid timepoints (Supp. Fig. 3B, see methods).

We observe highly consistent performance across all primary tissue datasets (GW5 – GW25) when predicting organoid cell-types regardless of the organoid timepoint (Supp. Fig. 3B). The average difference in AUROC scores when predicting organoid cell-types using either our youngest (GW5) or oldest (GW25) primary data is 0.000382 ± 0.0357 SD, 0.132 ± 0.188 SD, 0.141 ± 0.0980 SD, 0.0379 ± 0.130 SD and 0.0845 ± 0.209 SD for dividing progenitors, neural progenitors, glutamatergic neurons GABAergic neurons, and non-neuronal cells respectively (No annotated intermediate progenitors in the GW25 primary tissue dataset). This demonstrates strikingly consistent performance across primary tissue timepoints, highlighting that broad primary tissue cell-type signatures are applicable as reference for organoid cell-types regardless of the primary tissue or organoid timepoint. The one exception is for neural progenitors, where there seemingly is a temporal shift in performance with younger primary

tissue datasets predicting younger organoid annotations over older organoid annotations and vice-versa for older primary tissue/organoid data (Supp. Fig. 3B). However, a subset of the young GW6-8 primary tissue datasets report sharp increases in performance predicting older organoid timepoints in opposition to other GW6-8 primary tissue datasets, suggesting variance in performance is driven by intersections between the quality of individual organoid and primary tissue datasets rather than overarching temporal variability. Importantly, our lists of top 100 primary tissue MetaMarkers perform comparably to marker sets from individual primary tissue datasets, with less variance in performance across the organoid timepoints for the differentiated cell-types (mean AUROC variance across organoid timepoints for individual primary tissue datasets vs. primary MetaMarker variance; glutamatergic: 0.0142, 0.00477, GABAergic: 0.00921, 0.00199, non-neuronal: 0.00901, 0.00670, Supp. Fig. 3B). This demonstrates our meta-analytic primary tissue cell-type markers robustly capture organoid temporal variation.

Broad primary tissue cell-type markers capture organoid protocol variation

We assess whether our primary tissue MetaMarker gene sets capture organoid variation outside the annotated forebrain temporal organoid atlas by performing principal-component analysis (PCA) across all organoid datasets, representing data from 12 different differentiation protocols. Our lists of 100 primary tissue MetaMarkers are consistently heavily weighted in the first PC across organoid datasets (Supp. Fig. 3C-D). While a large portion of PC1-weighted genes are dividing progenitor MetaMarkers (representing cell-cycle signal), markers for non-dividing fetal cell-types also comprise those genes consistently heavily weighted in PC1 across organoid datasets (Supp. Fig. 3C-D).

Aggregate organoid co-expression preserves primary tissue cell-type co-expression

Our primary tissue MetaMarkers that capture both primary tissue and organoid temporal/regional variation enable assessments of cell-type specific co-expression between arbitrary primary tissue and organoid datasets. One normally would need matched cell-type annotations across datasets to compare cell-type specific biology, but here we couple our meta-analytically derived cell-type markers with gene co-expression quantifications, which do not rely on cell-type annotations, to extract cell-type specific co-expression from any given scRNA-seq dataset. Practically, if organoids are producing cell-types functionally identical to primary tissue cell-types, we would expect near identical co-expression relationships within our primary tissue MetaMarker gene sets across primary tissue and organoid datasets.

Deriving co-expression relationships from single-cell data is challenging due to inherent sparsity of the expression data (Fig. 3A). We overcome this sparsity with straightforward standardization and aggregation approaches (Fig. 3A, see methods), which prioritize replicable signal across datasets. We first explore marker set co-expression within our unannotated primary tissue datasets, which were not included in deriving our primary tissue MetaMarker sets. The aggregate unannotated primary tissue co-expression network nearly perfectly constructs cell-type specific co-expression modules when hierarchically clustering the co-expression of our top 100 primary tissue MetaMarker gene sets (Fig. 3B). Turning to the aggregated organoid co-expression network, while the intra- and inter-MetaMarker gene set co-expression appears dysregulated compared to the unannotated primary tissue co-expression network, the overall clustering of MetaMarker genes by co-expression still largely captures cell-type specific clustering (Fig. 3B). We quantify this through the Adjusted Rands

Index (ARI) metric, comparing the MetaMarker clustering through co-expression in any given network to the perfect clustering of MetaMarker gene sets by cell-type. We perform this quantification for both the aggregated co-expression networks (diamond, triangle, and square special characters, Fig. 3C) and for all individual primary tissue and organoid co-expression networks (boxplots, Fig. 3C). While individual organoid networks perform worse than individual primary tissue networks on average, the aggregated organoid network is largely comparable to individual primary tissue networks (Fig. 3C, median annotated and unannotated primary tissue network ARI: 0.403, 0.437, aggregated organoid network ARI: 0.381). In aggregate, organoid co-expression largely captures broad primary tissue cell-type specific co-expression.

Organoid datasets vary in primary tissue cell-type marker set co-expression

Having broadly assessed co-expression across our MetaMarker gene sets, we then asked how well do organoids recapitulate primary tissue co-expression within each cell-type specific MetaMarker gene set. We score intra-gene set co-expression strength through a simple machine learning framework[45,46], which quantifies whether genes in a given set are more strongly co-expressed with each other compared to the rest of the genome (Fig. 3D).

Co-expression module scores across the annotated and unannotated primary tissue datasets are largely comparable with the exception of a sharp decrease in intermediate progenitor performance for the unannotated primary tissue datasets (Fig. 3E). Six out of the fourteen unannotated datasets are sampled from either the ganglionic eminences or the hypothalamus, potentially explaining this decrease in performance and suggesting our intermediate progenitor MetaMarkers are enriched for signal from cortical areas. In contrast, performance is much more variable across the individual organoid datasets for all cell-types except the dividing progenitors, ranging from no signal (AUROC <= 0.50) to comparable results with primary tissue networks (Fig. 3E). We visualize the top and bottom performing organoid co-expression networks for glutamatergic co-expression to highlight the extreme variability across organoid datasets in recapitulating primary tissue co-expression (Fig. 3F). In the top performing organoid network, we find cell-type specific co-expression modules with a clear glutamatergic module (Fig. 3F). While co-expression for dividing progenitor markers constructs a clear module in the bottom performing organoid network, non-dividing primary tissue cell-type co-expression is clearly dysregulated with clusters composed of all primary tissue cell-type markers (Fig. 3F). By quantifying the intra-gene set co-expression of our primary tissue MetaMarkers, we are able to place organoid datasets on a spectrum ranging from complete failure to primary tissue-level recapitulation of primary tissue co-expression.

Importantly, organoid datasets vary strongly by protocol type in recapitulating primary tissue cell-type specific co-expression (Supp. Fig. 4). The undirected differentiation protocols (cerebral and cortical, Supp. Fig. 4) produce highly variable results across the primary tissue cell-types, in line with previous reports of high variability for undirected organoids. Intriguingly, the vascularized cortical organoid protocol produces consistently high performance across all primary tissue cell-types (Supp. Fig. 4), suggesting vascularized models increase organoids' capacity to produce comparable primary tissue cell-types *in vitro*. We also find the vascularized cortical, dorsal patterned forebrain, and undirected cortical protocols produce some of the highest co-expression module scores for our GABAergic primary tissue MetaMarkers (Supp. Fig. 4). These results agree with previous observations of *in vitro* production of inhibitory cell-types within these cortical models previously expected to produce exclusively excitatory lineages.

Organoid datasets vary in preserving gene-level primary tissue co-expression

We take our primary tissue/organoid co-expression comparisons a step further and ask how well individual organoid datasets preserve gene-level primary tissue co-expression relationships. For any given individual gene, we can quantify whether that gene's top co-expressed partners are preserved in one co-expression network compared to another (Fig. 4A). We use the aggregated co-expression network from the annotated primary tissue datasets as our reference co-expression network and test how well individual co-expression networks, either primary tissue or organoid, perform in preserving primary tissue gene-level co-expression patterns (Fig. 4A, top 10 co-expressed neighbors). We start by quantifying the preserved co-expression of genes within our primary tissue MetaMarker gene sets, using the average preserved co-expression AUROC as a measure of preserved co-expression for any given gene set (Fig. 4A). Across our 6 annotated primary tissue cell-types, primary tissue co-expression networks deliver consistently high performance for preserved co-expression scores of our primary tissue MetaMarker gene sets (Fig. 4B, mean preserved co-expression score across cell-types and primary tissue datasets: annotated 0.971 ± 0.0227 SD, unannotated 0.963 ± 0.00957 SD). This indicates the top 10 co-expressed partners are highly preserved for the vast majority of genes within each MetaMarker gene set across all primary tissue datasets.

In contrast, individual organoid datasets vary substantially in preserved co-expression scores across our primary tissue MetaMarker gene sets (Fig. 4B). As before with our quantification of intra-gene set co-expression, quantifying preserved gene-level co-expression places organoid datasets on a spectrum of near zero to indistinguishable preserved co-expression to primary tissue data. Organoid datasets vary substantially by protocol in preserving primary tissue cell-type specific co-expression, echoing similar trends as observed from our co-expression module analysis (Supp. Fig. 5). Since the majority of our organoid protocols are designed for producing excitatory lineages, it is encouraging we report a higher average preservation of glutamatergic primary tissue co-expression over non-neuronal or GABAergic primary tissue co-expression across our organoid datasets (Fig. 4B). Unsurprisingly, preservation of dividing progenitor co-expression is universally high with a preserved co-expression score of approximately 1 in nearly every primary tissue and organoid dataset, representing consistent co-expression of cell-cycle marker genes across systems (Fig. 4B, Supp. Fig. 5). A subset of organoid datasets are clear outliers to this trend (Fig. 4B, Supp. Fig. 5), suggesting that cell-cycle co-expression is not preserved, indicating basic cellular functions may be dysregulated in these datasets. One intriguing observation came from a study that compared organoids grown in a vertical shaker versus an orbital shaker[47]. We show that organoids grown in an orbital shaker produce higher preserved primary tissue co-expression scores for intermediate progenitors and glutamatergic cell-types whereas organoids grown in a vertical shaker produce higher scores for GABAergic cell-types (3 replicates each, glutamatergic, intermediate progenitor, GABAergic; Orbital: 0.896 ± 0.00102 SD, 0.795 ± 0.00148 SD, 0.665 ± 0.0308 SD. Vertical: 0.644 ± 0.0125 SD, 0.686 ± 0.0161 SD, 0.763 ± 0.00731 SD). This suggests the mechanical conditions of organoid growth can distinctly impact lineage and cell-type production in organoids.

With measures of preserved primary tissue co-expression for multiple cell-types within organoids, we can additionally assess variation in preserved co-expression across cell-types within individual organoid datasets. We compute correlations of preserved co-expression scores between the 6 MetaMarker sets across all organoid datasets and find significantly

positive correlations (FDR-adjusted p-value < .001) across all comparisons with the exception of the non-neuronal cell-type (Fig. 4C, non-neuronal FDR-adjusted p-values range from < 0.001 to 0.745). This indicates preserved primary tissue co-expression is a global feature of organoid datasets. For example, if an organoid is producing neural progenitors that preserve primary tissue co-expression, that organoid is likely producing other cell-types that preserve primary tissue co-expression. Similarly, we asked if preserved co-expression varies across normal or perturbed organoids. A subset of our organoid datasets come from studies that performed various perturbations (22q11.2 deletion, SMARCB1 knockdown, exposure to Alzheimer's serum, SETBP1 point mutations, amyotrophic lateral sclerosis patient-derived organoids). We compare the MetaMarker preserved co-expression scores between normal and perturbed organoids and find only a single significant difference across all cell-type MetaMarker sets (intermediate progenitor normal vs. mutant preserved co-expression score FDR-adjusted p-value: 0.0287, Supp. Fig. 6). This demonstrates our broad primary tissue cell-type co-expression signatures are also applicable for comparison with organoids in perturbation experiments.

After revealing cell-type specific variation for preserving primary tissue co-expression within organoids, our co-expression networks additionally allow genome-wide assessments of preserved co-expression. We extend our analysis via GO terms to quantify preserved primary tissue co-expression within organoids across the whole genome. GO terms with significantly preserved primary tissue co-expression (see methods) in organoids are mostly related to basic cellular functions like response to DNA damage and protein translation, as well as GO terms related to neurodevelopment (Fig. 4D). GO terms that significantly lack preservation of primary tissue co-expression are almost exclusively related to angiogenesis or immune function (Fig. 4D), concordant with the fact that organoids lack vasculature and an immune system.

While GO terms are useful for partitioning the genome into functional units for comparison, our co-expression networks also enable assessments of preserved co-expression for individual genes. As a particular use-case, we search for genes with exceptionally high preserved primary tissue co-expression across primary tissue datasets that also have poor preserved primary tissue co-expression across organoid datasets. We only consider genes that have some measurable expression in every organoid and primary tissue dataset and compute the average preserved co-expression AUROC for each gene across the organoid and primary tissue datasets (Supp. Fig. 6). The top 10 enriched GO terms for genes (76 in total) with high primary tissue (average AUROC >= 0.99) and low organoid (average AUROC < 0.70) preserved co-expression are related to extra-cellular matrix (ECM) and vascular characterizations (Fig. 4E). The poor conservation of genes related to vasculature can be explained by the absence of vascularization in the vast majority of our organoid datasets. The subset of these 76 genes in the ECM GO terms are CAV1, CAV2, COL4A1, CTSK, ENG, LAMB1, LAMC1, NID1, NID2, DDR2, and VWA1. Notably, these genes produce collagen and laminins, components of Matrigel, the artificial ECM typically included in organoid cultures. These results highlight preserved primary tissue co-expression of ECM-related genes as a particularly consistent deficit across neural organoids, suggesting that investigations into the signaling between artificial ECM and cells in organoid cultures may be a route forward for general improvements of organoid fidelity.

In summary, we interrogate co-expression in organoids at multiple levels, revealing organoids vary in preserving primary tissue co-expression at gene-, cell-type, and whole

genome resolutions through the use of a robust aggregate primary tissue co-expression network.

Temporal variation in organoid preservation of primary tissue co-expression

We score preserved co-expression in organoids using the aggregate primary tissue co-expression network (GW5-25), which by design aims to capture signal robust to temporal variation. To investigate temporal trends in organoid co-expression, we employ a similar approach as when predicting organoid cell-type annotations (Supp. Fig. 3), this time quantifying the preservation of primary tissue co-expression for the top 100 cell-type markers per individual primary tissue dataset across all organoid timepoints (Fig. 5A-B). We uncover a broad temporal shift in the preservation of primary tissue co-expression within organoids across all cell-types, with younger organoids (23 days – 1.5 months) as the top performers for mostly first trimester primary tissue co-expression transitioning to older organoids (2 – 6 months) as top performers for mostly second trimester primary tissue co-expression (Fig. 5B). This temporal shift is broadly consistent across the cell-types, beginning around GW9-10 (Fig. 5B). Our approach in predicting organoid annotations in Figure 2 is based on aggregate marker expression and did not produce temporally variable results, whereas our approach here comparing preserved co-expression of the same marker genes does produce temporally variable results. This indicates that the co-expression relationships of genes rather than their expression levels better capture temporal variation in developing systems.

Organoids preserve developing brain co-expression over adult brain co-expression

We demonstrate temporal variation in developing brain co-expression relationships is captured by organoids, but only from the single forebrain organoid protocol used in the temporal organoid atlas. In order to extend analysis across all our organoid datasets and assess broad temporal variation in co-expression, we next investigate the preserved co-expression within organoids of both developing and adult brain co-expression relationships.

We construct an aggregate adult co-expression network from a medial temporal gyrus scRNA-seq dataset[48] (157,508 cells) sampling 7 adult individuals. We compare the preserved co-expression scores of organoids for either developing or adult glutamatergic, GABAergic, and non-neuronal cell-types. Organoids unanimously preserve developing brain co-expression over adult co-expression (Supp. Fig. 6) for glutamatergic and GABAergic cell-types with equally poor performance for the non-neuronal cell-type, suggesting organoids generally fail to produce non-neuronal cell-types. We extend this analysis genome-wide and place organoids in context between developing and adult data by computing the average preservation of co-expression AUROC across all genes for organoid, developing, and adult co-expression using the annotated primary developing brain tissue network as the reference. The adult co-expression network produces a global preserved developing brain co-expression score of 0.591, indicating very poor performance across the genome in preserving developing co-expression relationships (Supp. Fig. 6). Organoids vary substantially in their global preservation of developing brain co-expression with some organoid datasets performing comparably to the adult data. This result is largely influenced by the number of cells present within individual organoid datasets (Supp. Fig. 6, corr 0.647, p-value < .001), suggesting a cell-sampling limitation for uncovering developing brain co-expression within organoids. However, organoid datasets report more variable global preserved co-expression scores compared to down-sampled developing brain data (Supp. Fig. 6), indicating a remaining biological gap

between primary developing brain tissue and organoid data not explained through technical means.

An intriguing study generated data from human cortical organoids either transplanted or not into developing rat brains to test the limits of maturation organoids can achieve *in vitro*[49]. We compare the preservation of developing and adult co-expression between these age-matched non-transplanted and transplanted human cortical organoids. We report that while the non-transplanted organoids preserve developing co-expression over adult for glutamatergic and GABAergic markers (Supp. Fig. 6, non-transplanted glutamatergic and GABAergic mean developing brain AUROCs: 0.797 ± 0.0281 SD, 0.697 ± 0.0212 SD. Non-transplanted glutamatergic and GABAergic mean adult AUROCs: 0.672 ± 0.0234 SD, 0.586 ± 0.0309 SD), the transplanted organoids have increased preservation of adult co-expression for glutamatergic and non-neuronal cell-types (Supp. Fig. 6, transplanted glutamatergic and non-neuronal mean developing brain AUROCs: 0.759 ± 0.00908 SD, 0.501 ± 0.0127 SD. Transplanted glutamatergic and non-neuronal mean adult AUROCs: 0.849 ± 0.0332 SD, 0.738 ± 0.00779 SD). This indicates the transplanted human organoids are adopting adult human glutamatergic and non-neuronal co-expression, concordant with the original authors' conclusions of increased maturation in transplanted organoids.

Variability in organoid co-expression is driven by marker gene expression

We investigate the impact of various technical features in our analysis on our co-expression results by assessing their correlation with our co-expression module scores and preserved co-expression AUROCs, focusing on technical features like sequencing depth, number of cells, etc. An important technical consideration for our analysis is ensuring all datasets have an identical gene namespace for meaningful comparisons of expression data. We fit all datasets to the GO gene universe, dropping gene annotations not in GO or zero-padding missing GO annotations in individual datasets. Excessive zero-padding of genes within our MetaMarker gene sets may artificially lower co-expression module scores or preserved co-expression scores, though we find this relationship to be relatively weak with little impact on score variance (Supp. Fig. 7, R^2 for co-expression module scores and zero-padding: 0.00123, 0.0165, 0.103, 0.0284, 0.0337, 0.052, R^2 for preserved co-expression and zero-padding: 0.0825, 0.321, 0.149, 0.0302, 0.0409, 0.000769 for neural prog., dividing prog., intermediate prog., glutamatergic, GABAergic, and non-neuronal cell-types respectively). Sequencing depth is also similarly found to have little impact on our co-expression module scores or preserved co-expression scores (Supp. Fig. 7). Rather, the features strongly related to performance are the number of cells in a dataset and the strength of marker set expression, all significantly strongly correlated with co-expression module scores and preserved co-expression scores (Supp. Fig. 7, range of significant (p-value < .001) correlations between marker set expression or cell number and co-expression module scores or preserved co-expression scores: 0.351 – 0.808, with the exception of dividing progenitors having a significant negative correlation of -0.453 between marker set expression and co-expression module scores).

Preservation of primary tissue co-expression as a generalizable quality control metric

As a general summary, our approach for quantifying preserved primary tissue co-expression across numerous organoid protocols revealed the axes on which organoids lie for recapitulating primary tissue co-expression relationships at gene, cell-type, and whole-genome

resolutions. These assessments provide powerful quality control information, identifying which genes and/or cell-types organoids can or cannot currently model on par with primary tissue data. We make our methods accessible through an R package to aid in future organoid studies and protocol development, providing means for rapidly constructing co-expression networks from scRNA-seq data (Fig. 6A) as well as querying preserved co-expression of users' data with our aggregate primary tissue brain co-expression network (Fig. 6A). Additionally, we make the results of our meta-analysis across primary tissue and organoid datasets available for users to place their data in reference to a field-wide collection (Fig. 6B).

Discussion

Through the use of meta-analytic differential expression and co-expression, we are able to provide cell-type specific measurements of human neural organoids' current capacity to replicate primary tissue biology. We extracted broad cell-type markers that define primary brain tissue cell-types across a large temporal axis (GW5 – 25) and across numerous heterogenous brain regions to act as a generalizable primary tissue reference for organoids that also vary temporally and regionally (by protocol). By quantifying intra-marker set co-expression and the preservation of co-expression across networks, we revealed human neural organoids lie on a spectrum of near-zero to near-identical recapitulation of primary tissue cell-type specific co-expression in comparison to primary tissue data. We made our aggregate primary tissue reference data and methods for measuring preserved co-expression publicly available as an R package to aid in the quality control and protocol development of future human neural organoids.

Prior work comparing primary brain tissue and neural organoid systems demonstrated organoids can produce cell-types[11,12] and morphological structures[27,43] similar to primary tissues and are capable of modeling temporal[13,38,40] and regional[3,12,28,29] primary tissue variation. Multiple lines of evidence support these findings such as assessments of cytoarchitecture and cell-type proportions[3,11,16,23], whole transcriptome and marker gene expression correlations[10,12], and comparisons of co-expression modules[5,13,17,39]. Our meta-analytic approach is able to quantify these field-wide observations within a generalizable framework, recapitulating that organoids model broad primary tissue biology with our specific approach offering several key advancements for primary tissue/organoid comparisons. First, we derive quantifications of preserved primary tissue co-expression that can be extended from individual genes to the entire genome and, second, we place organoid co-expression in reference to robust meta-analytic primary tissue performance providing a general benchmark for protocol development and quality control across heterogeneous organoid systems.

A key aspect of our study design is our cross-validation of primary tissue differential expression and co-expression. We demonstrated that temporally and regionally heterogenous primary tissue data are able to strongly recapitulate our meta-analytic primary tissue marker gene expression and co-expression relationships. This meta-analytic primary tissue performance defines a clear benchmark for gauging the fidelity of organoid models, where organoids that produce functionally equivalent primary tissue cell-types are expected to perform comparably to primary tissue data. In our assessment across 12 different organoid differentiation protocols, we showed a subset of protocols produce organoids with comparable cell-type specific co-expression to primary tissue data, demonstrating high primary tissue fidelity is possible with current methods. While we employ a broad approach sampling across temporal and regional variation to optimize for generalizability, more precisely matched primary tissue data for specific organoid timepoints or protocols is better suited for comparisons studying more subtle variation.

Certainly, while comparisons between primary tissue and organoid systems at a high-resolution of cell-type annotation are of interest, our results centered on broad cell-types at the cell-class level constitute a critical foundation for these more fine-tuned investigations of organoids. Cell-type specification within the brain involves complex spatial and temporal mechanisms[50] to produce the high cellular heterogeneity we observe, with the exact resolution of meaningful cell-type annotations still being actively debated and posing a general

conceptual challenge within the field of single-cell genomics[51]. We focus here on establishing methods for assessing consistent and accurate production of primary tissue cell-types at the class-level within organoids as a critical actionable first step towards increasing primary tissue fidelity across variable organoid differentiation protocols, with meta-analysis at a higher resolution of cell-type annotations (e.g., MGE and CGE interneurons, layer-specific excitatory neurons, progenitor subtypes) as an exciting future venture once class-level fidelity in organoids is consistently achieved.

One exciting application for the use of neural organoid systems is the study of a wide-range of human neurological diseases using human *in vitro* models[52,53], which critically depends on the *in vivo* fidelity of cell-types produced in organoids. Neural organoids have been used to model and investigate human disorders of neurodevelopmental[3,54], neuropsychiatric[55–57], and neurodegenerative[58–60] nature, as well as infectious diseases[28,61,62]. It is essential that organoid systems model *in vivo* cell-types with extreme fidelity to fully realize the therapeutic potential of human organoids and ensure findings in these *in vitro* models are not specific to potential artifactual or inaccurate *in vitro* biology. While our results demonstrate that high primary tissue fidelity in organoids is currently methodologically possible, we also report a high degree of variability across organoids and studies/protocols indicating a remaining methodological gap. The broad applicability of our meta-analytic approach offers the potential for benchmarking primary tissue fidelity across numerous organoid protocols, aiding in increasing the quality of neural organoids for use in a wide-range of human health-related translational investigations.

The generalizable and flexible nature of our analysis is well suited to aid in the development of organoid differentiation protocols and the general quality control of neural organoids. Our results demonstrate the type of experiments possible through comparing preserved co-expression across organoid experimental variables, such as the differences in preserved co-expression between organoids grown in vertical or orbital shakers, as well as between transplanted or non-transplanted organoids. Importantly, our broad sampling across organoid protocols enabled clear identification of promising avenues for increasing organoid primary tissue fidelity. The strong performance across cell-types for the vascularized protocol we assessed suggests vascularized protocols as a route forward for global increases in primary fidelity. Additionally, our findings of specific ECM-related genes with consistent poorly preserved primary tissue co-expression in organoids suggests investigations into the interactions between Matrigel or other ECM-substrates and organoids may lead to general protocol adjustments for increasing primary tissue fidelity[63]. Looking beyond neural organoids, our framework for quantifying preserved co-expression can be applied to other organoid systems granted there is sufficient annotated primary tissue data to act as a reference.

Methods

Dataset download and scRNA-seq pre-processing

Links for all downloaded data (GEO accession numbers, data repositories, etc.) are provided in Supp. Table 1. All scRNA-seq data was processed using the Seurat v4.2.0 R package. Data made available in 10XGenomics format (barcodes.tsv.gz, features.tsv.gz, matrix.mtx.gz) were converted into Seurat objects using the Read10X() and CreateSeuratObject() Seurat functions. Data made available as expression matrices were converted into sparse matrices and then converted into Seurat objects using the CreateSeuratObject() function. Ensembl gene IDs were converted into gene names using the biomaRt v2.52.0[64] package.

Where metadata was made available, we separated data by batch (Age, Donor, Cell line, etc.) for our final total of 130 organoid and 51 primary tissue datasets (Supp. Table 1). We processed and analyzed each batch independently without integration. We used consistent thresholds for filtering cells across all datasets, keeping cells that had less than 50% of reads mapping to mitochondrial genes and had between 200 and 6000 detected genes. Several datasets provided annotations for potential doublets; we excluded all cells labeled as doublets when annotations were made available. All data made available with raw expression counts were CPM normalized with NormalizeData(normalization.method = 'RC', scale.factor = 1e6), otherwise normalizations were kept as author supplied.

For primary tissue and organoid data made available with cell-type annotations, we provide our mapping between author provided annotations and our broad cell-type annotations in Supp. Table 2.

Primary tissue MetaMarker generation and cross-validation

MetaMarkers were computed using the MetaMarkers v0.0.1[44] R package, which requires shared cell-type and gene annotations across datasets to derive a ranked list of MetaMarkers. Gene markers for individual datasets were first computed using the compute_markers() function on the CPM normalized expression data for our annotated primary tissue datasets (Supp. Table. 1). A ranked list of MetaMarkers was then computed using the make_meta_markers() function using all 37 individual annotated primary tissue dataset marker lists. Genes are first ranked through their recurrent differential expression (the number of datasets that gene was called as DE using a threshold of log2 FC >= 4 and FDR p-value <= .05) and then through the averaged differential expression statistics of each gene across individual datasets. When we take the top 100 markers per individual dataset as in Fig. 2D, Fig. 5, Supp. Fig. 1A, and Supp. Fig. 3B, we rank markers for each dataset by their AUROC statistic as computed with the compute_markers() MetaMarkers function.

For the cross-validation of our primary tissue MetaMarkers, we excluded a single annotated primary tissue dataset, computed MetaMarkers from the remaining 36 annotated primary tissue datasets, and then used those MetaMarkers to predict the cell-type annotations of the left-out dataset. We construct an aggregate expression predictor to quantify the predictive strength a list of genes has, in this case our MetaMarker lists, in predicting cell-type annotations. Taking any arbitrary number of genes (10, 20, 50, 100, 250, or 500 MetaMarkers), we sum the expression counts for those genes within each cell and then rank all cells by this

aggregate expression vector. We compute an AUROC using this ranking and the cell-type annotations for a particular cell-type through the Mann-Whitney U test. Formally:

$$AUROC = \frac{U}{n_0 * n_1}$$

where U is the Mann-Whitney U test statistic, n_0 is the number of positives (cells with a given cell-type annotation) and n_1 is the number of negatives (cells without that cell-type annotation).

$$U = R_0 - \frac{n_0(n_0 + 1)}{2}$$

where R_0 is the sum of the positive ranks.

As an example, if there are 10 genes that are perfect glutamatergic markers (only glutamatergic cells express these genes), then ranking cells by the summed expression of these genes will place all glutamatergic cells (positives) in front of all other cells (negatives), producing an AUROC of 1. The violin plots in Supp. Fig. 1B and in Figure 2E visualize our aggregate expression approach, where datapoints per cell-type are the aggregated expression counts for the given top 100 MetaMarkers across all cells per dataset (Supp. Fig. 1B) or aggregated across all datasets (Fig. 2E).

For Supp. Fig. 1A, we took the top 100 cell-type markers per individual primary tissue dataset (x-axis) and used those genes to predict cell-type annotations as described above for all other annotated primary tissue datasets, reported as the AUROC boxplot distributions. The MetaMarker distribution was computed using a leave-one-out approach as described above. We ranked the individual primary tissue datasets by their median AUROC performance per cell-type to derive the distributions of ranks presented in Figure 2D, excluding the dividing progenitor data as performance was highly consistent across all primary tissue datasets.

Cross-regional primary tissue MetaMarker expression

We investigated the aggregate expression of our top 100 MetaMarkers per cell-type across annotated brain regions separately for the annotated first-trimester and second-trimester primary tissue atlases due to differing regional annotations. MetaMarkers were computed with a leave-one-out approach as described above using all 37 of the annotated primary tissue datasets. For the heatmaps in Supp. Fig. 2, rows represent the annotated cells present within the given dataset, columns represent the aggregated expression for the top 100 given cell-type MetaMarkers and each annotated region present. We average the aggregated expression for each cell-type per region and then normalize each region (column) by the maximum average expression value across the cell-types. A value of 1 indicates that cell-type is the one maximally expressing the given MetaMarker set for that brain region. The heatmaps are ordered by cell-type and region and are not clustered.

Organoid PCA

PCA analysis was performed using the Seurat function RunPCA() with the top 2000 variable features, determined using the Seurat function

FindVariableFeatures(selection.method = 'vst', nfeatures = 2000). For each organoid dataset, we took the eigenvector for the first principal component, computed the absolute value, and then divided by the maximum value to compute a normalized vector between 0 and 1. We visualized the normalized eigenvectors for each organoid dataset in Supp. Fig. 3C, keeping primary tissue MetaMarker genes that were detected in the top 2000 variable genes of at least 10 organoid datasets. Genes missing from any given dataset's top 2000 variable genes were given a value of 0. The heatmap was produced using the ComplexHeatmap v2.12.1[65] package and was hierarchically clustered using the ward.D2 method for both rows and columns.

Generating co-expression networks from scRNA-seq data

To generate a shared gene annotation space across all datasets, we fit each dataset to the GO gene universe before computing co-expression matrices. Using human GO annotations (sourced 2022-03-10 using the org.Hs.eg.db v3.15.0[66] and AnnotationDbi v1.58.0[67] R packages), we excluded gene expression from a dataset if the gene annotation was not present in GO and we zero-padded missing GO genes for each dataset.

We compute a gene-by-gene co-expression matrix per dataset using the spearman correlation coefficient computed across all cells in a given dataset. We then rank the correlation coefficients in the gene-by-gene matrix and divide by the maximum rank to obtain a rank-standardized co-expression matrix. All results reported using individual dataset co-expression networks (Fig. 3C-F, Fig. 4B, Figs. 5-6, Supp. Figs. 4-7) were obtained using the rank-standardized co-expression networks.

We compute the aggregated co-expression networks by taking the average of the rank standardized co-expression networks for each gene-gene index.

Hierarchical clustering of primary tissue MetaMarkers by co-expression

We visualize the co-expression of primary tissue MetaMarker genes using the ComplexHeatmap package and the ward.D2 algorithm for hierarchical clustering. We use the fossil v0.4.0 package[68] to compute the adjusted Rands Index with the adj.rand.index() function. To compute the adjusted Rands Index, we calculate a consensus clustering of MetaMarkers per co-expression network across 100 k-means clusterings (using the arguments row_km = 6, column_km = 6, row_km_repeats = 100, column_km_repeats = 100 within the Heatmap function) to compare to the perfect grouping of MetaMarkers by cell-type.

Co-expression module learning analysis

EGAD v1.24.0[45] is a machine learning framework that quantifies the strength of co-expression within an arbitrary gene-set compared to the rest of the genome with an AUROC quantification (Fig. 3D). We compute co-expression module AUROCs for all GO gene-sets (between 10 and 1000 genes per GO term) and our top 100 primary tissue MetaMarker gene-sets for each individual primary tissue and organoid co-expression network as well as the aggregated annotated, unannotated and organoid networks. For the annotated primary tissue co-expression networks, we employ a leave-one-out approach, learning MetaMarkers from 36 of the annotated datasets and computing co-expression module AUROCs for these MetaMarkers in the left-out dataset's co-expression network. We compute co-expression module AUROCs using the EGAD run_GBA() function with default parameters. In Figure 3E,

the ‘All GO terms’ distributions report the average co-expression module AUROC across all GO terms for each individual network.

Preservation of co-expression

To compute our preservation of co-expression AUROC, we take the top 10 co-expressed partners for gene A in a reference co-expression network as our positive gene annotations. In a test co-expression network, we rank all genes through their co-expression with gene A and compute an AUROC using this ranking and the positive annotations derived from the reference network. If gene A in the test network has the exact same top 10 co-expressed partners as in the reference network, that would result in an AUROC of 1. To summarize a given gene-set’s preserved co-expression, we take the average preserved co-expression AUROC across all genes in that gene set as the preservation of co-expression score for that gene set. We use the aggregated annotated primary tissue co-expression matrix as our reference network.

The preserved co-expression scores for the annotated primary tissue data in Figure 4B were computed with a leave-one-out approach. MetaMarkers and an aggregated co-expression matrix were computed from 36 of the annotated primary tissue datasets and then preserved co-expression scores were computed using the co-expression network of the left-out annotated primary tissue dataset.

Preservation of GO term co-expression

We compute p-values for the preservation of co-expression of GO terms using a mean sample error approach. Using the aggregated annotated primary tissue co-expression network as the reference and the aggregated organoid network as the test network, we first compute the preserved co-expression AUROCs for all individual genes, taking the mean and standard deviation value as the population mean and population standard deviation. For any given GO term, we first compute the preserved co-expression score for the term (the average of the preserved co-expression AUROCs for the genes in the term) and then compute the sample error for that score with:

$$SE = \frac{SD_{pop}}{\sqrt{n_{GO}}}$$

where $SD_{pop}$is the population standard deviation and n_{GO} is the number of genes in the GO term. We then compute a z-score through:

$$Z_{GO} = \frac{mu_{GO} - mu_{pop}}{SE}$$

where mu_{go} is the preserved co-expression score for the GO term and mu_{pop} is the population mean preserved co-expression AUROC. We compute left-sided p-values using the standard normal distribution:

$$p_L = P(X \leq Z_{GO})$$

Where X is a normal distribution with mean = 0 and standard deviation = 1. We use the R function pnorm(Z_{GO}) to compute this p-value.

We then compute the right-sided p-value as:

$$p_R = 1 - p_L$$

We adjust p-values using the R function p.adjust(method = 'BH'). We filter for GO terms that have between 20 and 250 genes per term and use a threshold of FDR-corrected p-value <= 0.0001 to call significance. Significant left-sided p-values are interpreted as GO terms with significantly smaller preserved co-expression scores (significantly not preserved) than expected through sampling error and right-sided p-values are interpreted as GO terms with significantly larger preserved co-expression scores (significantly preserved) than expected through sampling error. We use the R package rrvgo to visualize the significant GO terms in Fig. 4D.

Computing correlation significance

We employ a permutation test to compute p-values for any given correlation coefficient. We permute data-pairs and compute a correlation coefficient, repeating for 10,000 random permutations to generate a distribution of correlation coefficients under the null hypothesis of independence. We calculate a two-sided p-value for the original correlation coefficient as the number of permuted correlation coefficients whose absolute value is greater than or equal to the absolute value of the original correlation coefficient, divided by 10,000. We adjust p-values using the R function p.adjust(method = 'BH') and use a FDR-corrected p-value threshold of <= .05 to call significance.

Comparing co-expression of normal vs. perturbed organoids

For both the co-expression module AUROCs and the preserved co-expression scores of normal and perturbed organoids, we test for significant differences per cell-type using the Mann Whitney U test, adjusting p-values with the R function p.adjust(method = 'BH') and using a FDR-corrected p-value threshold of <= .05 to call significance.

Organoid temporal analysis

The organoid temporal analysis for both predicting organoid annotations with primary tissue markers (Supp. Fig. 3B) and scoring the preserved co-expression of organoid co-expression using primary tissue networks as reference (Fig. 5) were performed for all pair-wise combinations of the 37 annotated primary tissue datasets and the 26 temporally annotated forebrain organoid datasets. We excluded the GW7-28 annotated primary tissue dataset from the temporal preserved co-expression analysis (Fig. 5) due to the wide temporal range sampled. For predicting organoid annotations with primary tissue markers, we used the top 100 markers per primary tissue dataset to construct aggregate expression predictors in the organoid datasets as described above. The MetaMarkers performance was calculated using MetaMarkers derived from all 37 annotated primary tissue datasets. For scoring preserved co-expression, individual primary tissue networks were used as the reference with individual organoid networks as the test networks. We computed the preserved co-expression scores of the top 100 primary tissue cell-type markers per individual primary dataset for each individual organoid network.

GO enrichment analysis

We compute enrichment for GO terms using Fisher's Exact Test as implemented through the hypergeometric test. We compute raw p-values for GO terms with between 10-1000 genes and compute FDR-adjusted p-values using p.adjust(method = 'BH'). We only consider GO sets with between 20 and 500 when choosing the top 10 GO sets in Figure 4E, ranked by FDR-adjusted p-value.

R and R packages

All analysis was carried out in R v4.2.2. Colors with selected using the MetBrewer v0.2.0 R library. Plots were generated using ggplot2 v3.3.6[69]. Spearman correlation matrices for co-expression networks were computed using a python v3.6.8 script, implemented in R with the reticulate v1.26 R package, as well as using functions from the matrixStats v0.62.0 R library. All code used in generating results and visualizations will be made public at the time of publication. The preservedCoexp R library is made available at https://github.com/JonathanMWerner/preservedCoexp.

Author Contributions

JMW and JG conceived the project. JMW and JG designed analyses. JMW performed analyses. JMW and JG wrote the manuscript.

Acknowledgments

JMW was supported by NSF award no. DGE-1938105. JG and JMW were supported by NIH grants R01MH113005 and R01LM012736. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under grant no. DGE-1938105. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. We thank Tomasz Nowakowski and members of the Gillis lab for helpful comments on the manuscript.

976 **References**

977

978 1. Eiraku, M. *et al.* Self-Organized Formation of Polarized Cortical Tissues from ESCs and Its Active

979 Manipulation by Extrinsic Signals. *Cell Stem Cell* **3**, 519–532 (2008).

980 2. Sato, T. *et al.* Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal

981 niche. *Nature* **459**, 262–265 (2009).

982 3. Lancaster, M. A. *et al.* Cerebral organoids model human brain development and microcephaly.

983 *Nature* **501**, 373–379 (2013).

984 4. Corrò, C., Novellasdemunt, L. & Li, V. S. W. A brief history of organoids. *Am. J. Physiol.-Cell Physiol.*

985 **319**, C151–C165 (2020).

986 5. Pollen, A. A. *et al.* Establishing Cerebral Organoids as Models of Human-Specific Brain Evolution. *Cell*

987 **176**, 743-756.e17 (2019).

988 6. Kanton, S. *et al.* Organoid single-cell genomic atlas uncovers human-specific features of brain

989 development. *Nature* **574**, 418–422 (2019).

990 7. Benito-Kwiecinski, S. *et al.* An early cell shape transition drives evolutionary expansion of the human

991 forebrain. *Cell* **184**, 2084-2102.e19 (2021).

992 8. Fleck, J. S. *et al.* Inferring and perturbing cell fate regulomes in human brain organoids. *Nature* 1–8

993 (2022) doi:10.1038/s41586-022-05279-8.

994 9. Sarieva, K. & Mayer, S. The Effects of Environmental Adversities on Human Neocortical

995 Neurogenesis Modeled in Brain Organoids. *Front. Mol. Biosci.* **8**, (2021).

996 10. Camp, J. G. *et al.* Human cerebral organoids recapitulate gene expression programs of fetal

997 neocortex development. *Proc. Natl. Acad. Sci.* **112**, 15672–15677 (2015).

998 11. Velasco, S. *et al.* Individual brain organoids reproducibly form cell diversity of the human

999 cerebral cortex. *Nature* **570**, 523–527 (2019).

1000 12. Bhaduri, A. *et al.* Cell stress in cortical organoids impairs molecular subtype specification.
1001 *Nature* **578**, 142–148 (2020).
1002 13. Gordon, A. *et al.* Long-term maturation of human cortical organoids matches key early
1003 postnatal transitions. *Nat. Neurosci.* **24**, 331–342 (2021).
1004 14. Feng, W. *et al.* Computational profiling of hiPSC-derived heart organoids reveals chamber
1005 defects associated with NKX2-5 deficiency. *Commun. Biol.* **5**, 1–18 (2022).
1006 15. Leek, J. T. *et al.* Tackling the widespread and critical impact of batch effects in high-throughput
1007 data. *Nat. Rev. Genet.* **11**, 733–739 (2010).
1008 16. Tanaka, Y., Cakir, B., Xiang, Y., Sullivan, G. J. & Park, I.-H. Synthetic Analyses of Single-Cell
1009 Transcriptomes from Multiple Brain Organoids and Fetal Brain. *Cell Rep.* **30**, 1682-1689.e3 (2020).
1010 17. Cheroni, C. *et al.* Benchmarking brain organoid recapitulation of fetal corticogenesis. *Transl.*
1011 *Psychiatry* **12**, 1–16 (2022).
1012 18. Kim, H. J. *et al.* Comprehensive characterization of fetal and mature retinal cell identity to
1013 assess the fidelity of retinal organoids. *Stem Cell Rep.* **18**, 175–189 (2023).
1014 19. Zhang, B. & Horvath, S. A General Framework for Weighted Gene Co-Expression Network
1015 Analysis. *Stat. Appl. Genet. Mol. Biol.* **4**, (2005).
1016 20. Lee, J., Shah, M., Ballouz, S., Crow, M. & Gillis, J. CoCoCoNet: conserved and comparative co-
1017 expression across a diverse set of species. *Nucleic Acids Res.* **48**, W566–W571 (2020).
1018 21. Crow, M., Paul, A., Ballouz, S., Huang, Z. J. & Gillis, J. Exploiting single-cell expression to
1019 characterize co-expression replicability. *Genome Biol.* **17**, 101 (2016).
1020 22. Mead, B. E. *et al.* Harnessing single-cell genomics to improve the physiological fidelity of
1021 organoid-derived cell types. *BMC Biol.* **16**, 62 (2018).

1022 23. Agboola, O. S., Hu, X., Shan, Z., Wu, Y. & Lei, L. Brain organoid: a 3D technology for investigating
1023 cellular composition and interactions in human neurological development and disease models in
1024 vitro. *Stem Cell Res. Ther.* **12**, 430 (2021).
1025 24. Mayhew, C. N. & Singhania, R. A review of protocols for brain organoids and applications for
1026 disease modeling. *STAR Protoc.* **4**, 101860 (2023).
1027 25. Lancaster, M. A. & Knoblich, J. A. Generation of cerebral organoids from human pluripotent
1028 stem cells. *Nat. Protoc.* **9**, 2329–2340 (2014).
1029 26. Muguruma, K., Nishiyama, A., Kawakami, H., Hashimoto, K. & Sasai, Y. Self-Organization of
1030 Polarized Cerebellar Tissue in 3D Culture of Human Pluripotent Stem Cells. *Cell Rep.* **10**, 537–550
1031 (2015).
1032 27. Sakaguchi, H. *et al.* Generation of functional hippocampal neurons from self-organizing human
1033 embryonic stem cell-derived dorsomedial telencephalic tissue. *Nat. Commun.* **6**, 8896 (2015).
1034 28. Qian, X. *et al.* Brain-Region-Specific Organoids Using Mini-bioreactors for Modeling ZIKV
1035 Exposure. *Cell* **165**, 1238–1254 (2016).
1036 29. Xiang, Y. *et al.* Fusion of Regionally Specified hPSC-Derived Organoids Models Human Brain
1037 Development and Interneuron Migration. *Cell Stem Cell* **21**, 383-398.e7 (2017).
1038 30. Birey, F. *et al.* Assembly of functionally integrated human forebrain spheroids. *Nature* **545**, 54–
1039 59 (2017).
1040 31. Xiang, Y. *et al.* hESC-Derived Thalamic Organoids Form Reciprocal Projections When Fused with
1041 Cortical Organoids. *Cell Stem Cell* **24**, 487-497.e7 (2019).
1042 32. Miura, Y. *et al.* Generation of human striatal organoids and cortico-striatal assembloids from
1043 human pluripotent stem cells. *Nat. Biotechnol.* **38**, 1421–1430 (2020).

1044 33. Eura, N. *et al.* Brainstem Organoids From Human Pluripotent Stem Cells. *Front. Neurosci.* **14**,
1045 (2020).
1046 34. Andersen, J. *et al.* Generation of Functional Human 3D Cortico-Motor Assembloids. *Cell* **183**,
1047 1913-1929.e26 (2020).
1048 35. Huang, W.-K. *et al.* Generation of hypothalamic arcuate organoids from human induced
1049 pluripotent stem cells. *Cell Stem Cell* **28**, 1657-1670.e10 (2021).
1050 36. Nayler, S., Agarwal, D., Curion, F., Bowden, R. & Becker, E. B. E. High-resolution transcriptional
1051 landscape of xeno-free human induced pluripotent stem cell-derived cerebellar organoids. *Sci. Rep.*
1052 **11**, 12959 (2021).
1053 37. Sozzi, E., Nilsson, F., Kajtez, J., Parmar, M. & Fiorenzano, A. Generation of Human Ventral
1054 Midbrain Organoids Derived from Pluripotent Stem Cells. *Curr. Protoc.* **2**, e555 (2022).
1055 38. Uzquiano, A. *et al.* Proper acquisition of cell class identity in organoids allows definition of fate
1056 specification programs of the human cerebral cortex. *Cell* **185**, 3770-3788.e27 (2022).
1057 39. Luo, C. *et al.* Cerebral Organoids Recapitulate Epigenomic Signatures of the Human Fetal Brain.
1058 *Cell Rep.* **17**, 3369–3384 (2016).
1059 40. Amiri, A. *et al.* Transcriptome and epigenome landscape of human cortical development
1060 modeled in brain organoids. *Science* **362**, eaat6720 (2018).
1061 41. Fair, S. R. *et al.* Electrophysiological Maturation of Cerebral Organoids Correlates with Dynamic
1062 Morphological and Cellular Development. *Stem Cell Rep.* **15**, 855–868 (2020).
1063 42. Nascimento, J. M. *et al.* Human Cerebral Organoids and Fetal Brain Tissue Share Proteomic
1064 Similarities. *Front. Cell Dev. Biol.* **7**, 303 (2019).
1065 43. Qian, X. *et al.* Sliced Human Cortical Organoids for Modeling Distinct Cortical Layer Formation.
1066 *Cell Stem Cell* **26**, 766-781.e9 (2020).

1067 44. Fischer, S. & Gillis, J. How many markers are needed to robustly determine a cell's type?
1068 *iScience* **24**, 103292 (2021).
1069 45. Ballouz, S., Weber, M., Pavlidis, P. & Gillis, J. EGAD: ultra-fast functional analysis of gene
1070 networks. *Bioinformatics* **33**, 612–614 (2017).
1071 46. Skinnider, M. A., Squair, J. W. & Foster, L. J. Evaluating measures of association for single-cell
1072 transcriptomics. *Nat. Methods* **16**, 381–386 (2019).
1073 47. Suong, D. N. A. *et al.* Induction of inverted morphology in brain organoids by vertical-mixing
1074 bioreactors. *Commun. Biol.* **4**, 1–13 (2021).
1075 48. Jorstad, N. L. *et al. Comparative transcriptomics reveals human-specific cortical features*.
1076 http://biorxiv.org/lookup/doi/10.1101/2022.09.19.508480 (2022) doi:10.1101/2022.09.19.508480.
1077 49. Revah, O. *et al.* Maturation and circuit integration of transplanted human cortical organoids.
1078 *Nature* **610**, 319–326 (2022).
1079 50. Nowakowski, T. J. *et al.* Spatiotemporal gene expression trajectories reveal developmental
1080 hierarchies of the human cortex. *Science* **358**, 1318–1323 (2017).
1081 51. Zeng, H. What is a cell type and how to define it? *Cell* **185**, 2739–2755 (2022).
1082 52. Chen, H. I., Song, H. & Ming, G. Applications of Human Brain Organoids to Clinical Problems.
1083 *Dev. Dyn.* **248**, 53–64 (2019).
1084 53. Eichmüller, O. L. & Knoblich, J. A. Human cerebral organoids — a new tool for clinical neurology
1085 research. *Nat. Rev. Neurol.* **18**, 661–680 (2022).
1086 54. Mariani, J. *et al.* FOXG1-Dependent Dysregulation of GABA/Glutamate Neuron Differentiation in
1087 Autism Spectrum Disorders. *Cell* **162**, 375–390 (2015).

1088 55. Notaras, M. *et al.* Schizophrenia is defined by cell-specific neuropathology and multiple
1089 neurodevelopmental mechanisms in patient-derived cerebral organoids. *Mol. Psychiatry* **27**, 1416–
1090 1434 (2022).
1091 56. Stachowiak, E. K. *et al.* Cerebral organoids reveal early cortical maldevelopment in
1092 schizophrenia—computational anatomy and genomics, role of FGFR1. *Transl. Psychiatry* **7**, 1–24
1093 (2017).
1094 57. Dixon, T. A. & Muotri, A. R. Advancing preclinical models of psychiatric disorders with human
1095 brain organoid cultures. *Mol. Psychiatry* **28**, 83–95 (2023).
1096 58. Smits, L. M. *et al.* Modeling Parkinson's disease in midbrain-like organoids. *Npj Park. Dis.* **5**, 1–8
1097 (2019).
1098 59. Chen, X. *et al.* Modeling Sporadic Alzheimer's Disease in Human Brain Organoids under Serum
1099 Exposure. *Adv. Sci.* **8**, 2101462 (2021).
1100 60. Szebényi, K. *et al.* Human ALS/FTD brain organoid slice cultures display distinct early astrocyte
1101 and targetable neuronal pathology. *Nat. Neurosci.* **24**, 1542–1554 (2021).
1102 61. Garcez, P. P. *et al.* Zika virus impairs growth in human neurospheres and brain organoids.
1103 *Science* **352**, 816–818 (2016).
1104 62. Pellegrini, L. *et al.* SARS-CoV-2 Infects the Brain Choroid Plexus and Disrupts the Blood-CSF
1105 Barrier in Human Brain Organoids. *Cell Stem Cell* **27**, 951-961.e5 (2020).
1106 63. Kozlowski, M. T., Crook, C. J. & Ku, H. T. Towards organoid culture without Matrigel. *Commun.*
1107 *Biol.* **4**, 1–15 (2021).
1108 64. Durinck, S., Spellman, P. T., Birney, E. & Huber, W. Mapping identifiers for the integration of
1109 genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc.* **4**, 1184–1191 (2009).

65. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* (2016).

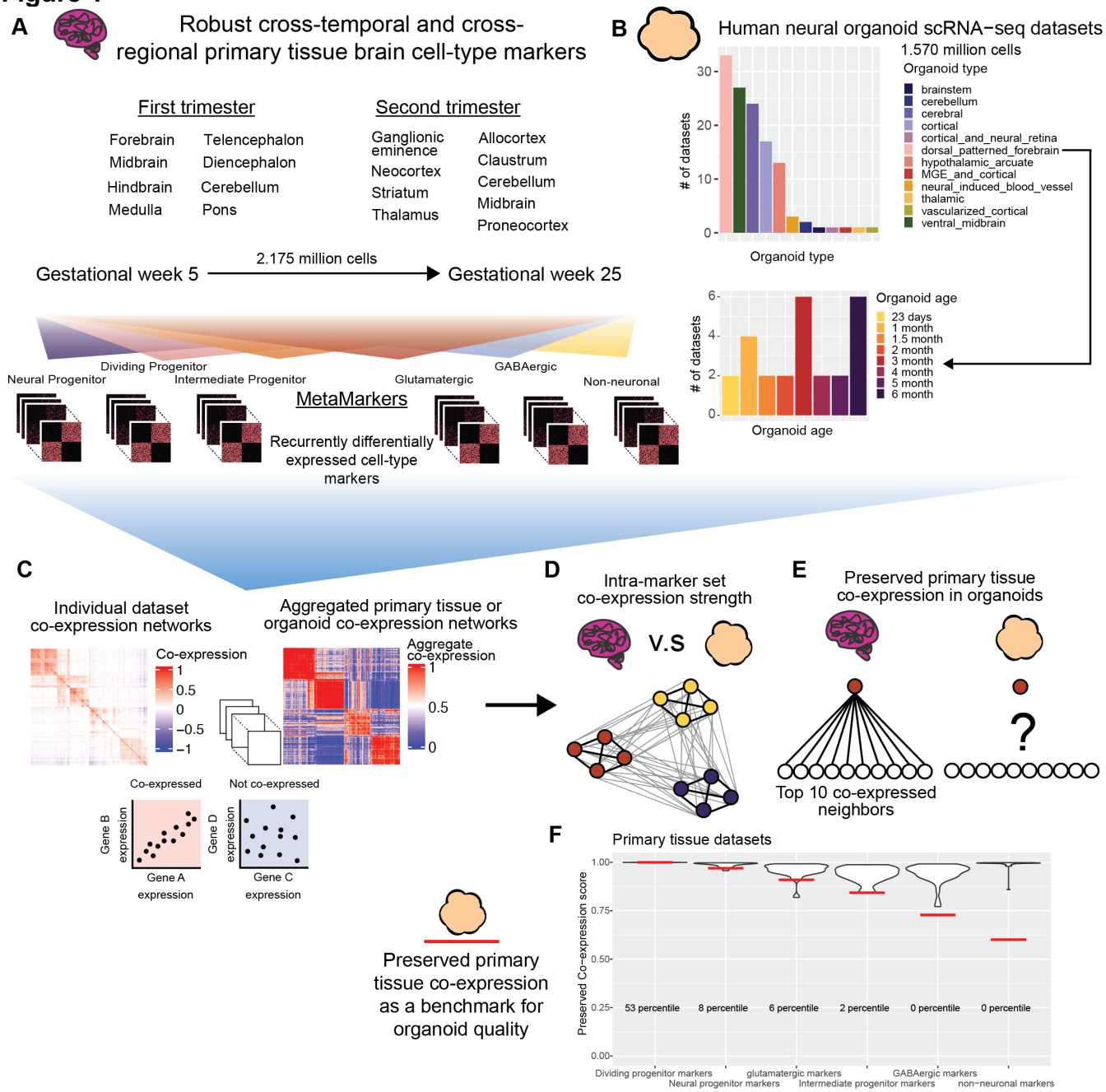
66. Carlson, M. org.Hs.eg.db: Genome wide annotation for Human. (2019).

67. Pagès, H., Carlson, M., Falcon, S. & Li, N. AnnotationDbi: Manipulation of SQLite-based annotations in Bioconductor. (2022).

68. Vavrek, M. J. fossil: Palaeoecological and Palaeogeographical Analysis Tools. (2020).

69. Wickham, H. ggplot2: Elegant Graphics for Data Analysis. *Springer-Verl. N. Y.* (2016).

Figure 1



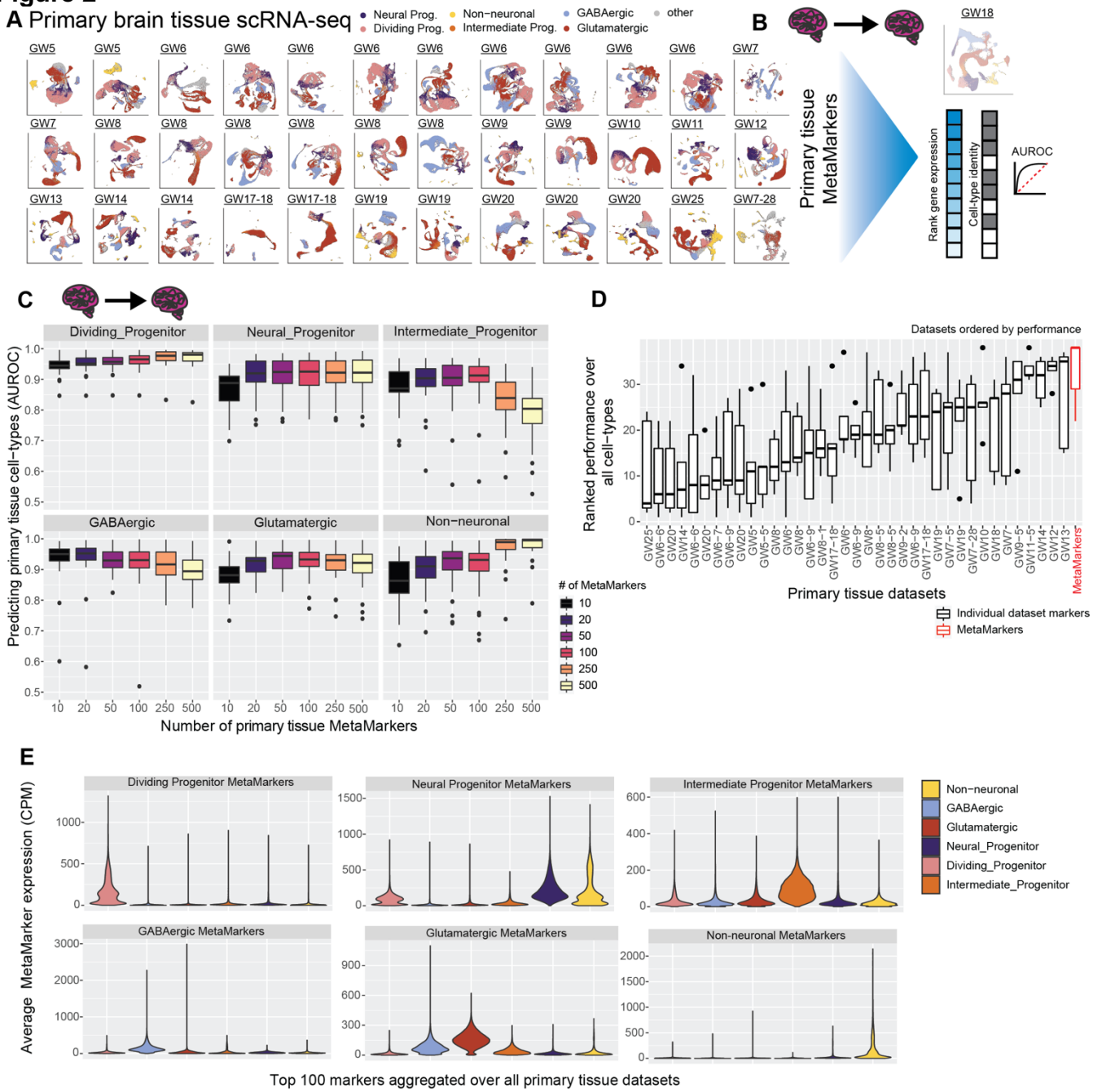
Using meta-analysis to quantify preserved primary tissue co-expression in organoids

A Collection of annotated primary tissue brain scRNA-seq datasets, ranging from gestational week (GW) 5 to 25 and sampling from 15 developmentally defined brain regions. The primary tissue datasets are annotated at broad cell-type levels (Neural Progenitor, Dividing Progenitor, Intermediate Progenitor, Glutamatergic, GABAergic, and Non-neuronal) and these annotations are used to compute MetaMarkers, cell-type markers identified through recurrent differential expression.

B Collection of human neural organoid scRNA-seq datasets, sampling from 12 different differentiation protocols. Included is an annotated temporal forebrain organoid dataset.

C Example of a sparse co-expression network derived from a scRNA-seq data and an example of an aggregate co-expression network averaged over many scRNA-seq datasets. The aggregate network enhances the sparse signal from the individual network.
D Schematic showing a quantification of intra-marker set co-expression
E Schematic showing a quantification for the strength of preserved co-expression between two co-expression networks, measuring the replication of the top 10 co-expressed partners of an individual gene across the networks.
F Example plot from the preservedCoexp R library, placing cell-type specific preserved co-expression scores of an example forebrain organoid dataset in reference to scores derived from primary tissue datasets. Red lines denote the percentile of the organoid cell-type scores within the primary tissue distributions.

Figure 2



Meta-analytic primary tissue cell-type markers

A Annotated UMAPs of the annotated primary tissue brain scRNA-seq datasets.

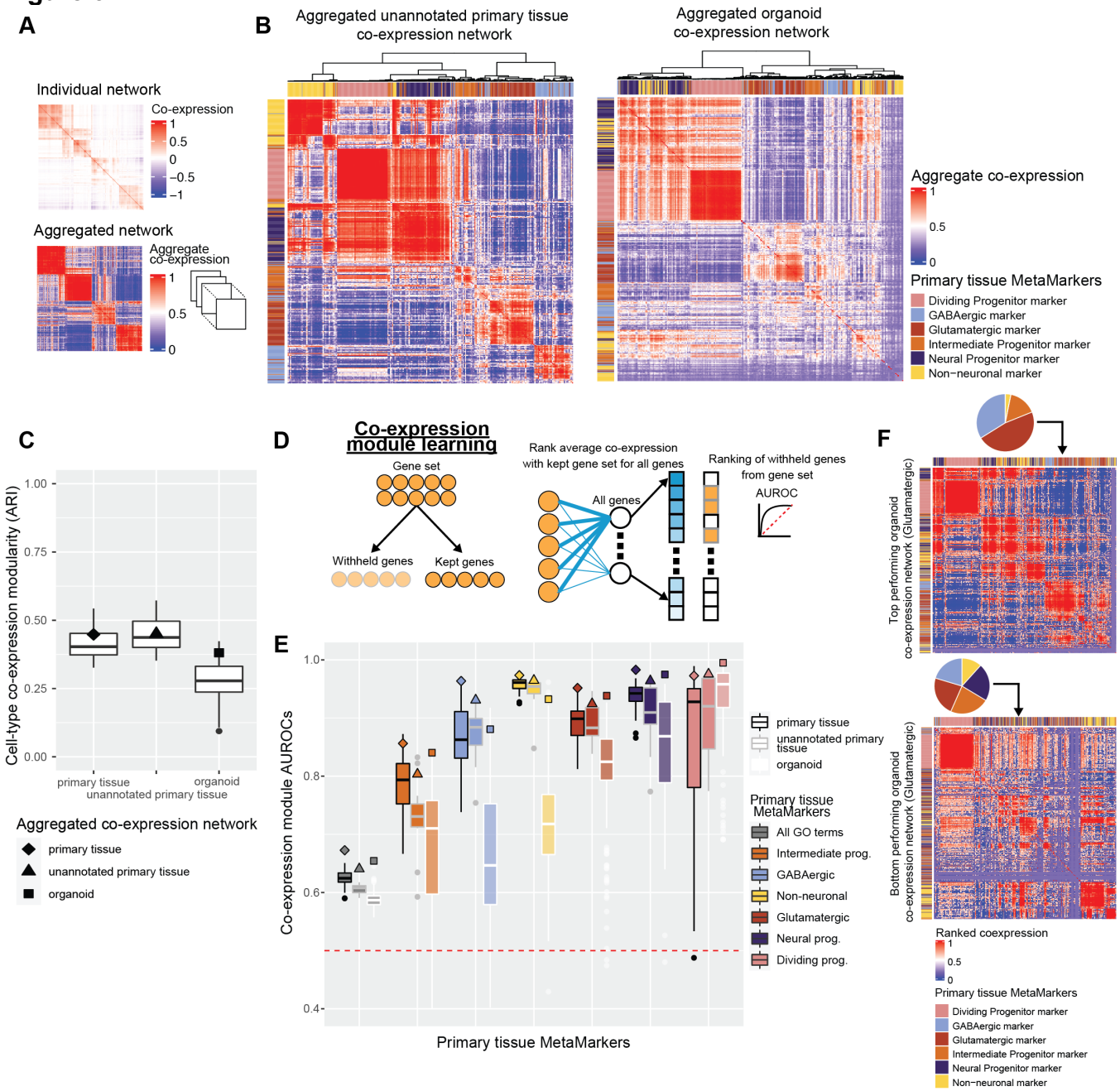
B Example of our leave-one-out cross-validation approach for learning primary tissue MetaMarkers and testing the markers' capacity for predicting annotations in the left-out dataset, quantified with the AUROC statistic.

C Meta-analytic primary tissue markers have high performance in predicting primary tissue cell-type annotations. Boxplot distributions of the AUROC statistic for predicting cell-type annotations across all leave-one-out combinations of our annotated primary tissue datasets, with an increasing number of MetaMarkers used for predicting cell-type annotations on the x-axis.

D MetaMarkers have the highest performance in predicting primary tissue cell-type annotations. Boxplots of marker gene-set performances. Gene-sets are the top 100 cell-type markers from individual primary tissue datasets compared to the MetaMarker performance. Performances for each cell-type in individual primary tissue datasets are presented in Supp. Fig. 1A. Datasets are ordered by their median performance.

E Averaged distributions of gene expression for the top 100 MetaMarkers demonstrating clear cell-type specificity. This is performed with a leave-one-out cross-validation, with individual dataset distributions reported in Supp. Fig. 1B.

Figure 3



Neural organoids vary in recapitulating primary tissue cell-type marker set co-expression

A Example of a sparse co-expression network derived from a scRNA-seq data and an example of an aggregate co-expression network averaged over many scRNA-seq datasets. The aggregate network enhances the sparse signal from the individual network.

B Marker gene-sets show clear cell-type clusters via their co-expression relationships in primary tissue and organoid networks. The aggregated co-expression networks for the unannotated primary tissue datasets and organoid datasets, showing the hierarchically clustered co-expression of the primary tissue MetaMarkers for the 6 cell-types.

C Organoid cell-type clustering via co-expression is notable lower compared to all primary tissue datasets. Distributions of the Adjusted Rands Index (ARI) for individual annotated

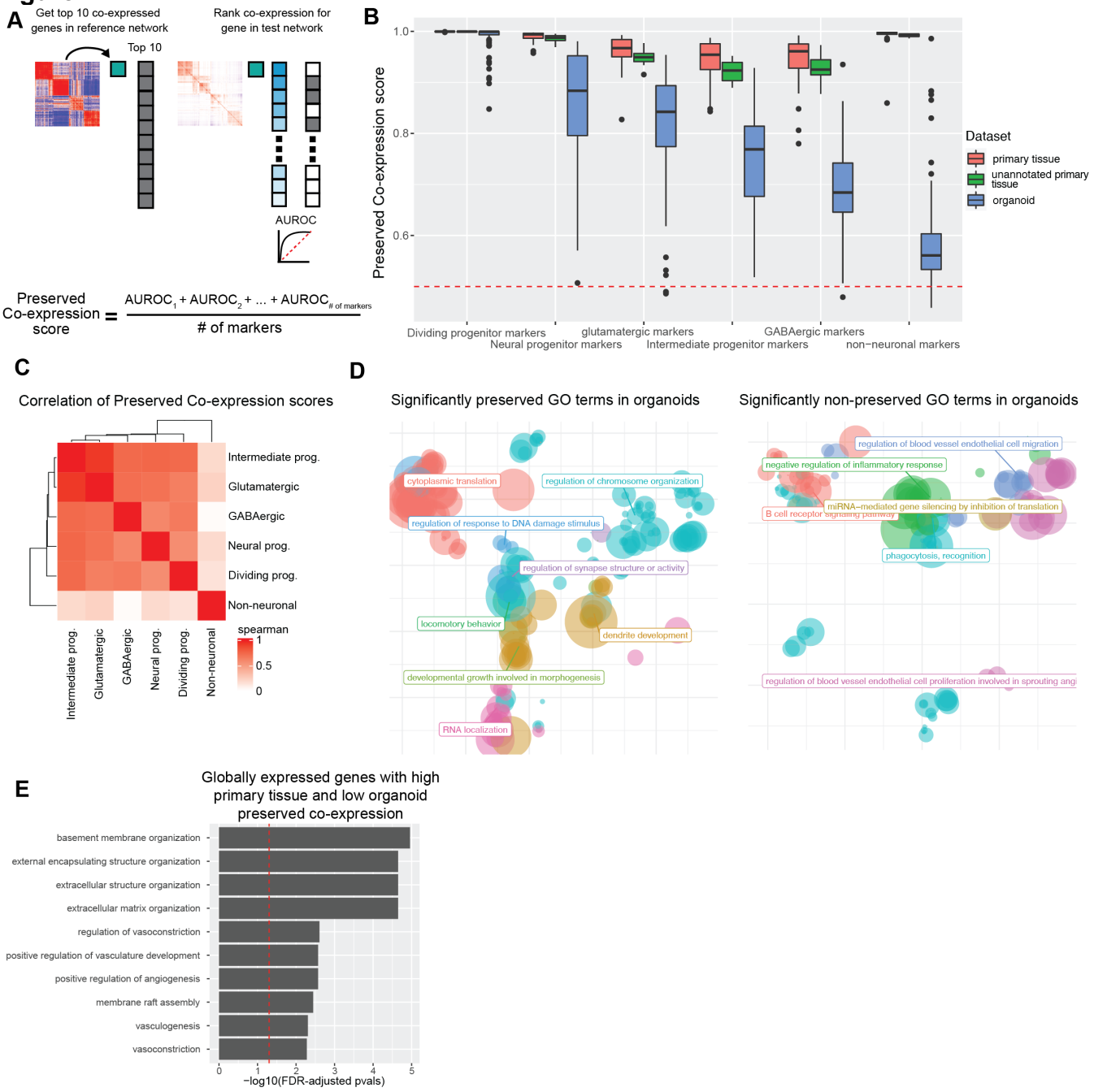
primary tissue, unannotated primary tissue, and organoid datasets. The ARI scores for the aggregate networks are denoted with the special characters.

D Schematic for the co-expression module learning framework, measuring the co-expression strength within an arbitrary gene-set compared to the rest of the genome, quantified with the AUROC statistic.

E Distributions of co-expression module AUROCs for individual annotated primary tissue, unannotated primary tissue, and organoid datasets for the co-expression strength of the MetaMarker gene-sets for the 6 cell-types. The grey 'All GO terms' distributions report the average co-expression module AUROC across all GO terms for each individual dataset. Co-expression module AUROCs for the aggregate co-expression networks are denoted with the special characters.

F Top and bottom organoid co-expression networks based on Glutamatergic performance. Heatmaps depict the hierarchically clustered co-expression of the primary tissue MetaMarker gene-sets for the 6 cell-types. Cell-type specific clusters are apparent in the top network, but are more mixed in the bottom network. Pie-charts depict the percentage of MetaMarker gene-sets that make up an example cluster in each network determined via the hierarchical clustering.

Figure 4



Neural organoids vary in their preservation of primary tissue gene-level co-expression

A Schematic showing the quantification for gene-level preserved co-expression. The preserved co-expression score for any given gene-set is the average preserved co-expression AUROC across all genes within that gene set.

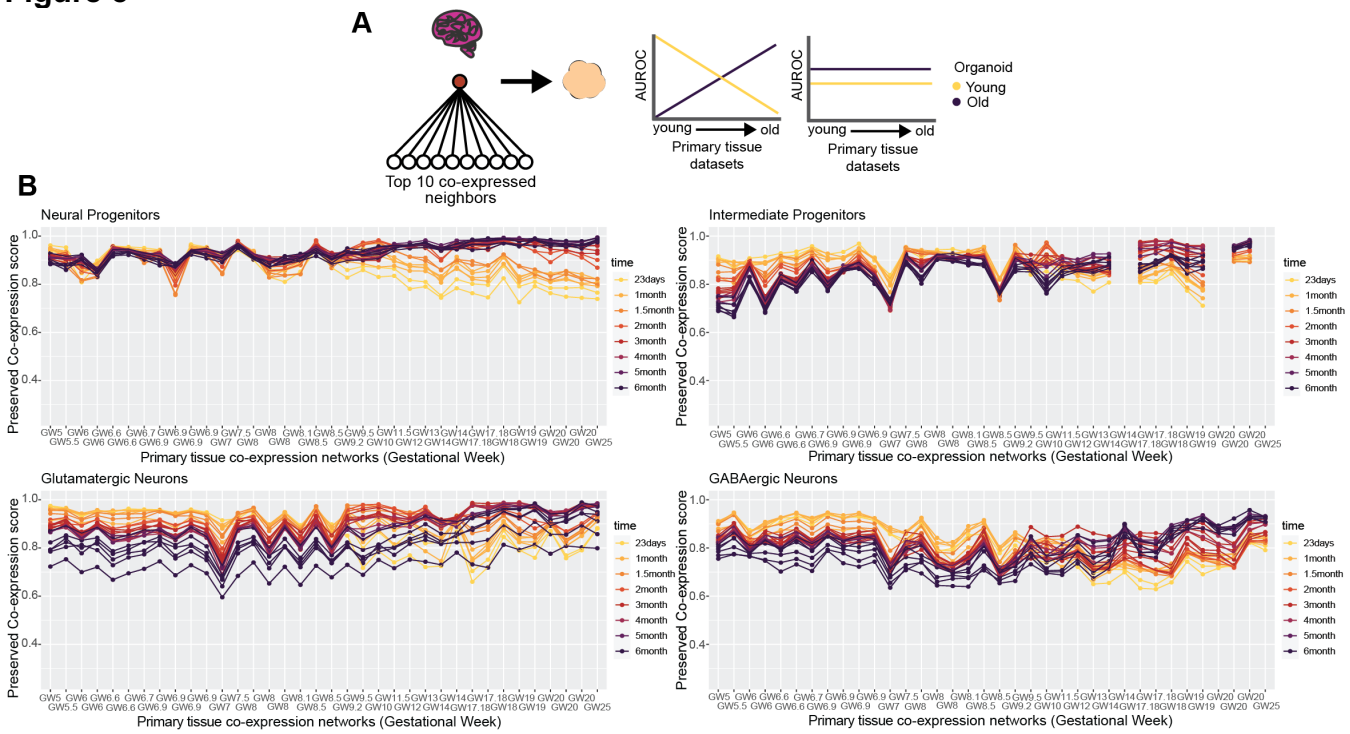
B Organoids strongly vary in preserved primary tissue cell-type specific co-expression in comparison to fetal data. Boxplot distributions show the preserved co-expression scores for the primary tissue MetaMarker gene-sets of the 6 cell-type annotations across all individual networks.

C The majority of cell-types are significantly correlated in preserved co-expression within organoid networks. Spearman correlation matrix for the preserved co-expression scores for all 6 cell-type annotations across all individual organoid datasets.

D Scatter plots summarizing the semantic distances of GO terms that are significantly preserved or non-preserved between the aggregate annotated primary tissue and organoid co-expression networks.

E Organoids globally fail to preserve primary tissue co-expression of ECM and vascular related genes. Bar plot detailing the top 10 GO terms from a GO enrichment test of the 76 genes with high and low preserved co-expression AUROCs within primary tissue networks and organoid networks respectively. The preserved co-expression for each individual gene from primary tissue networks and organoid networks is reported in Supp. Fig. 6E.

Figure 5

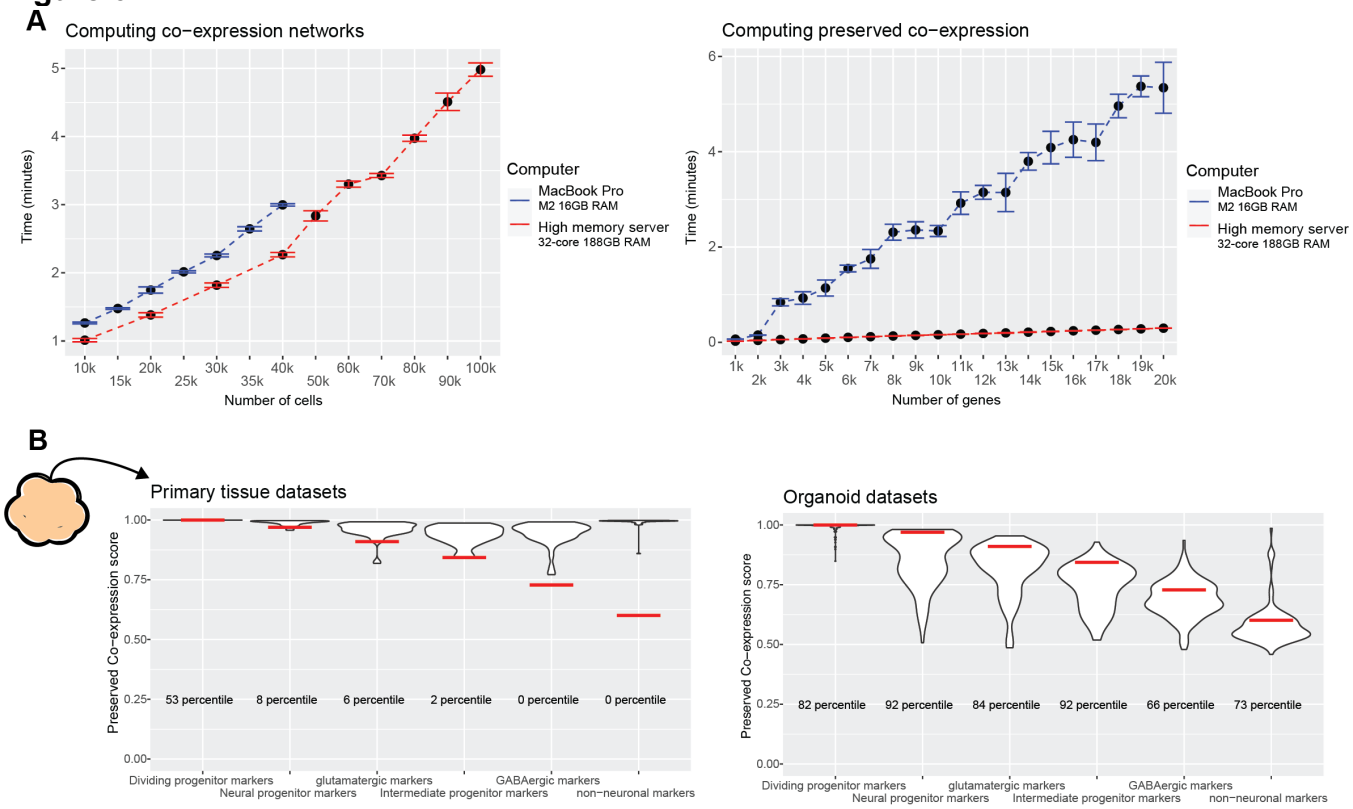


Neural organoids capture temporal dynamics in primary tissue co-expression

A Schematic showing two potential outcomes when comparing the preserved co-expression between primary tissue and organoid data on a temporal axis. There may be a temporal relationship, with younger organoids recapitulating younger primary tissue co-expression over older primary tissue co-expression and vice versa for older organoids, or there may be no temporal relationship.

B Organoid co-expression models temporal trends in primary tissue co-expression. Line plots showing the preserved co-expression scores computed from individual organoid co-expression networks for cell-type markers of individual primary tissue datasets. Primary tissue datasets on the x-axis are ordered from youngest to oldest.

Figure 6



The preservedCoexp R package enables fast computation of preserved co-expression
A The preservedCoexp R package can compute co-expression networks and genome-wide preservation of co-expression in a few minutes even for low-memory computers. Line plots showing the computational time to either compute co-expression networks or preserved co-expression as the number of cells or genes increases. Points are the mean value from 10 replicates, with error bars depicting ± 1 standard deviation.
B Example plot from the preservedCoexp R package, placing cell-type specific preserved co-expression scores of an example forebrain organoid dataset in reference to scores derived from primary tissue datasets or organoid datasets. Red lines denote the percentile of the forebrain organoid cell-type scores within either the primary tissue distributions or organoid distributions.