Available Online: 25 November 2026 DOI: 10.54254/3049-5458/2026.39541

Local glycolysis-hypoxia gene-expression relationships recur across differently organized human brain organoids

Quan Gui

Department of Biological Sciences, Purdue University, West Lafayette, USA

qgui@purdue.edu

Abstract. Brain organoids can contain similar cell populations yet differ in how those cells are arranged. Spatial comparisons reveal differences in internal organization alongside cellular composition and overall expression. This study compared spatial gene-expression data from three two-month cortical organoids and 24 midbrain-like organoids from two donors, linking nearby expression relationships to regional distributions and population composition within each organoid. Glycolysis-associated expression at one location was positively related to hypoxia-response expression nearby in all three cortical organoids and 23 of 24 midbrain-like organoids after accounting for recorded identity, tissue position and sampling-related variation. Within the midbrain-like cohort, both programs were enriched inward in fifteen organoids, outward in five, and in opposing directions in four. Positive local relationships occurred across all three configurations. In this same cohort, a greater proportion of spots labeled as immature neurons accompanied stronger hypoxia-response expression in inner relative to outer tissue after accounting for culture stage, donor and total spot count. Linking local expression, regional enrichment and population proportions reveals spatial differences among organoids that share a positive nearby metabolic relationship.

Keywords: brain organoids, spatial transcriptomics, spatial organization, glycolysis, hypoxia response

1. Introduction

Brain tissue organization places cells in distinct environments as they proliferate, differentiate and establish connections. Cellular activity requires oxygen and nutrients, whose availability depends on delivery, transport through tissue and local consumption. Cell position and the arrangement of neighboring populations can therefore matter both for access to these resources and for the demands placed on them. Human brain organoids reproduce progenitor zones and neuronal territories and allow their organization to be examined alongside molecular state [1, 2]. Relating metabolic gene expression to these arrangements asks whether organoids with similar cellular contents also share local expression relationships and regional patterns.

The supply environment of conventional nonvascularized organoids differs from that of the native brain. Native tissue receives oxygen and nutrients through a perfused vascular network connected to the circulation. Conventional nonvascularized organoid cultures receive nutrients from the surrounding medium, with transport into the tissue shaped by geometry and culture conditions [3-5]. Endothelial engineering can produce

Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

vascular-like networks with experimentally demonstrated transport features: Cakir et al. reported luminal tracer entry in vitro and perfused vessels after transplantation [5].

Experiments that alter tissue access link culture conditions to metabolic and developmental state. Slicing cortical organoids reduced interior hypoxia and cell death while supporting prolonged neurogenesis and cortical layer formation [3]. Cai et al. introduced artificial diffusible channels through vascular-network-inspired scaffolds, facilitating medium transport into midbrain organoids, reducing hypoxia and necrosis, and improving neuronal development and activity [4]. Stress-associated changes in cortical subtype specification [6] and protocol-dependent differences in cellular and metabolic profiles [7, 8] further connect culture environment with developmental state. These findings motivate comparing metabolic gene expression and cellular arrangement across independently generated organoids.

Existing spatial studies establish important source-specific observations. Uzquiano et al. described inward glycolysis and hypoxia-related transcription within assigned cortical identities and corroborated core hypoxia [9]. Budinger et al. profiled neuronal trajectories, mitochondrial programs and spatial neighborhoods in midbrain-like organoids and human tissue [10]. Lin et al. characterized control cortical architecture while studying structural disarray in autism models [11]. Patterned neural organoids and assembloids extend the range of experimentally generated organizations [12-14].

Local neighborhoods, regional enrichment and population position provide complementary descriptions of organoid organization. A glycolysis-hypoxia relationship between nearby locations may occur whether both programs are enriched inward, outward or in opposite directions. Similarly, a population can be prevalent in an outer region while expressing more of a program per observation inward. Connecting these measurements within identified organoids makes it possible to distinguish recurring relationships from configurations that separate individual cultures. Geometric inner-outer position describes sampled tissue organization, and gene-expression scores describe metabolic and hypoxia-response transcription.

This study compares three month-2 cortical organoids from Uzquiano with 24 Midbrain-Like Organoids (MLOs) from Budinger. Lin forebrain organoids supply an independent cortical comparison of neuronal organization; the broader eight-study WT/control resource supplies source-specific context (Table 1). The central question is whether nearby glycolysis-hypoxia relationships recur across organoids with different regional expression distributions. The primary glycolysis and hypoxia programs use mutually exclusive gene sets, termed disjoint programs. The analysis examines whether the relationship is stronger nearby than at comparable distant locations and estimates how much remains after accounting for the fitted linear relationship at the same locations. Energetic pairings and population proportions, within-population expression and relative position connect these local relationships to regional organization in identified organoids. The parallel cortical comparison asks whether shared broad neuronal arrangements also accompany shared local interfaces and molecular relationships.

2. Methods

2.1. Primary cohorts, biological units and recorded populations

The primary metabolic comparison comprises three month-2 Uzquiano cortical organoids and 24 Budinger MLOs at days 40, 70 and 120 [9, 10]. MLO hierarchy includes two donors, two differentiations per donor and four captures. The primary cohorts differ in source, developmental route and assay: cortical observations are Slide-seq beads, whereas MLO observations are multicellular Visium spots. Eight WT/control sources contribute the broader 92-map resource [9-12, 15-18]; contextual comparisons use source-specific units, including three serial iTracer sections from one organoid and five patterned model/stage maps. Table 1 and

Appendix A1 summarize the source-specific units, primary cohort hierarchy and annotation provenance. Sections are averaged equally within resolved organoids; line-level summaries average organoids equally. Fields, arrays and serial sections are linked through the available source specimen identifiers. Disease, tumor, transplant and retinal observations are excluded. GSE271118 is excluded for its coordinate defect; GSE294759 is excluded under the WT/control eligibility rule (Appendix A1).

Table 1. Study cohorts and source-specific measurement units

Source	Dataset	Maps	Observations	Interpretation
Lin	GSE197150	32	628,953	Panel-separated organoids; primary 14 organoids / 10 Control lines
Uzquiano	UZQUIANO_CELL_2022	10	23,290	Ten source organoids; primary month2 $n = 3$
Budinger	GSE277032	24	3,335	24 organoids, two donors, four captures
Salas-Lucia	GSE305978	2	1,525	Two Control FOVs
Lozachmeur	GSE223020	15	6,577	15 acquisition arrays
He	E-MTAB-10972	3	2,038	Three serial sections of one organoid
Yang	GSE289229	5	1,583	Five patterned model/stage maps
Ishahak	GSE283497	1	43,428	One in-vitro WT spatial map

Notes: Eight source studies with their map and observation counts. Biological units and endpoint eligibility remain source-specific; disease, transplant and retinal observations are excluded.

The source cortical cultures used agitation: Uzquiano organoids entered orbital culture and subsequently spinner flasks, whereas Lin organoids were cultured with shaking after release from Matrigel. Budinger MLOs were maintained under static conditions with medium changes every other day [9-11]. These cohorts were cultured under conventional nonvascularized conditions. Culture mixing describes medium movement around organoids.

The neuronal comparison comprises sixteen Lin sections from fourteen organoids across ten Control lines and the three month-2 Uzquiano organoids [9, 11]. Lin provides cell-resolved MERFISH observations; Uzquiano provides Slide-seq beads with transferred assignments. Fine identities were mapped to progenitor, immature-neuron and differentiated-neuron classes, with other named, mixed and unresolved categories retained in all-observation fraction denominators (Appendix A5). Section fractions were computed before equal-section organoid aggregation; Control-line means then weighted organoids equally [19]. Pooled-observation weighting is a separate sensitivity.

Lin project labels used the maximum unweighted mean log-transformed counts per 10,000 (main Equation (3)) marker-program score, with insertion-order tie handling and an undetected-marker label for nonpositive maxima. The immature program comprised *DCX*, TUBB3, TUBB, NHLH1, SOX11 and *STMN2*. The four settings used all members, omitted *DCX*, omitted *STMN2* or omitted both from the affected labeling programs, on the same observations and expression values to assess dependence on the expression-derived assignments [20]. These project classes are distinct from the source paper's Seurat identities and BANKSY domains [11]. The *SYT1*/*MAP2* distance analysis instead omitted both of those outcomes jointly from every Lin labeling program. Uzquiano annotations are source-transferred bead assignments (Appendix A1). Appendices A1 and A5 specify score membership, annotation provenance and neuronal-class definitions.

2.2. Neuronal position, adjacency and fetal relative order

Lin inner-positive depth is one minus radius relative to local angular boundary reach from median XY, using 360 angular bins and the 99.5th radial percentile. Uzquiano uses normalized distance to its 0.20 concave-hull boundary. Immature-neuron fraction contrasts are inner minus outer. In Lin, outer observations have depth < 0.33 and inner observations have depth ≥ 0.66; Uzquiano uses the source-defined outer and inner depth tertiles. Every accepted observation in a zone enters its fraction denominator. The early-neuronal expression endpoint correlates the mean of within-map *DCX* and *STMN2 z* scores with depth. Equal-section summaries produce fourteen Lin and three Uzquiano organoid effects. Marker-sensitivity analyses intersect supported endpoints before comparing each organoid with its baseline (Appendix A5).

Progenitor-immature adjacency uses directed k10 edges, excluding self and capping length at four times the median positive first-neighbor distance. Cross-population counts combine both edge directions. Component-conditioned label permutations preserve class frequencies within coordinate components; depth-conditioned permutations additionally preserve frequencies in the existing [0,0.2,0.4,0.6,0.8,1] depth strata. The latter requires twenty observations per stratum and 80% map coverage; comparisons also require thirty observations per class and ten expected edges. The 114 planned map/pair/null slots use 999 permutations, two-sided absolute deviations and a +1 correction. Twelve unavailable slots enter BH adjustment with temporary $p = 1$ while their displayed p/q remain missing. The twelve separate hierarchy tests use exact two-sided sign tests on ten Lin Control-line means or three Uzquiano organoids. Fixed-graph label accounting and marker-omission changes are descriptive comparisons under the component-conditioned definition (Appendix A5).

Fetal relative order compares source-annotated cortical states along the fetal ventricular-to-marginal-zone ordering with the direction of corresponding organoid state-position differences [21]. Each state pair uses its own supported observations, distinguishing radial glia, intermediate progenitors, immature/newborn neurons and differentiated neurons. Correspondence is summarized as the number of supported organoids matching the fetal order for each state pair. Appendix A5 gives the state mappings and pair-specific relative-order comparisons.

2.3. Crowding and progenitor-relative neuronal expression

Relative crowding is the relative spacing statistic in Equation (1). Higher values indicate closer spacing relative to the reference distance. Whole-map *DCX* associations use all-observation k10 crowding and Spearman correlation. Label centering subtracts each recorded label's mean expression before correlation with the same crowding values. *STMN2* analyses restrict expression and neighbor identity to the recorded immature population. Uncapped and component/gap-restricted k5/k10/k20 neighborhoods use estimator-specific finite masks, which can differ even among observations with shared annotation membership (Appendix A5). Count, depth, joint count/depth and quadratic-XY rank adjustments are evaluated separately. The four-gene/five-adjustment crowding family contains twenty exact two-sided line-level sign tests with BH20; the separate radial-expression/spacing family has 33 registered endpoints and 27 finite tests. Median-line intervals use 5,000 percentile-bootstrap draws (seed 9222026); sign-test ties use an absolute deviation at most 10^{-12}. In Appendix A5, each omission is paired with a baseline correlation evaluated on that omission's shared annotation membership, followed by equal-section organoid and equal-organoid Control-line aggregation.

$$c_{ik} = -\log\left(\frac{r_{ik}}{\mathrm{median}_{j \in V_k}(r_{jk})}\right) \tag{1}$$

In Equation (1), the distance is measured from each focal observation to its kth neighbor, excluding self. The reference set contains finite positive kth-neighbor distances for the specified estimator in the map or stated

population, after self exclusion and any stated population, component and gap restrictions; its median normalizes distance. Higher crowding values denote closer relative spacing.

For progenitor-relative expression, the neuronal population combines immature and differentiated classes. A symmetric, positive-edge k10 graph with the same distance cap defines multi-source shortest paths to progenitors; paths remain within connected components. The exposure is the graph-distance transform in Equation (2). Support requires at least twenty progenitors, at least one hundred reachable finite neuronal observations and at least five observations with detectable expression of the outcome gene. Conditional rank correlations adjust for fine-label indicators, normalized neuronal depth ranks and their squares, and ranks of log(1 + inherited total counts) and log(1 + relative tenth-neighbor spacing). Within each map or section, depth ranks are calculated across the finite neuronal observations in each gene-specific fit, with average ranks for ties. Each rank is divided by the number of observations in that fit; the normalized rank is then squared for the quadratic term. The inherited depth variable derives from the full source-map geometry before neuronal subsetting. Fit eligibility requires finite gene expression, inherited depth and progenitor graph distance, and positive inherited total counts; zero gene-expression values are eligible. Relative spacing divides the nonself tenth-neighbor distance by the median positive first-neighbor distance across the full native map. The separately fitted within-source expression models instead use fine-label indicators, raw depth, depth squared, log(1 + inherited total counts), and log(1 + relative tenth-neighbor spacing). The augmented model adds log(1 + relative progenitor graph distance); both distances use the full native map's median positive first-neighbor distance as denominator. These prediction models and rank correlations have separate estimands.

$$t_i = log(1 + \frac{d_i}{m_1}) \tag{2}$$

In Equation (2), shortest-path distance to the nearest progenitor on the capped coordinate graph is normalized by the median positive first-neighbor distance. Unreachable observations are unavailable. This reference differs from the median kth-neighbor distance in Equation (1).

Prediction holds out entire Control lines in Lin and entire organoids in Uzquiano, fitting models separately within each source. Training assigns equal weight to lines, to organoids within each line and to sections within each organoid; both models are evaluated on identical held-out IDs. Section results are averaged within organoid, then organoids within line. Error reduction compares the augmented and baseline prediction errors; negative values denote poorer prediction. The four planned neuronal sign tests (two genes by two sources) form a BH4 family; unavailable Uzquiano *MAP2* occupies a multiplicity slot with its reported result marked unavailable. Pointwise 95% percentile intervals resample the stated biological units in 10,000 draws. Appendix A5 summarizes neuronal population definitions and source-specific prediction outcomes.

2.4. Score memberships, expression and adjustment

Uzquiano expression fractions are used as supplied. Count-derived vectors use the inherited full-assay totals in Equation (3). Scores are equal-weight member means. Original-common glycolysis/hypoxia scores contain fifteen/thirteen genes, including *ENO1*, *LDHA*, *PDK1*, *PGK1* and *SLC2A1* in both. Removing these five defines the primary disjoint ten/eight-member scores. Table A1 gives the exact primary and complementary memberships; Appendix A1 distinguishes source-available scores and source normalization. Biological

annotations [22-24] and curated Reactome records[1] provide functional context for the program definitions.

$$x_{ig} = log(1 + \frac{10^4 c_{ig}}{T_i}) \tag{3}$$

For count inputs, c is the measured count of gene g in observation i; T is its inherited full-assay count total. For Uzquiano, x is instead the source-supplied expression fraction.

$$s_{iP} = \frac{1}{|G_P|} \sum_{g \in G_P} [x_{ig}] \tag{4}$$

Equation (4) averages expression equally over the measured members of the stated program and version. Appendix A1 specifies eligibility rules and the disjoint ten/eight-member definitions. Program names and analytical keys for the broad-source comparison are listed in Table 2.

Table 2. Program key for the broad-source comparison

Code	Analytical key	Display name
P01	developmental_stress_hypoxia	Developmental stress / hypoxia
P02	glycolysis	Glycolysis
P03	hypoxia	Hypoxia response
P04	oxidative_phosphorylation_etc	Oxidative phosphorylation / ETC
P05	mitochondrial_respiratory_chain	Respiratory chain
P06	atp_synthesis	ATP-related transcription
P07	tca_cycle	TCA cycle
P08	fatty_acid_oxidation	Fatty-acid oxidation
P09	oxidative_stress	Oxidative stress
P10	integrated_stress_response	Integrated stress response
P11	apoptosis_cell_death_stress	Apoptosis / cell-death stress
P12	proliferation_associated_metabolism	Proliferation-associated program
P13	glutamate_glutamine_metabolism	Glutamate / glutamine metabolism
P14	glycolysis_nonoverlap	Glycolysis: shared genes removed
P15	hypoxia_nonoverlap	Hypoxia response: shared genes removed

Notes: Program names and analytical keys used in the expression comparisons. P12 denotes the Proliferation-associated program, defined by eight requested genes; the C002 three-member subset is *CENPF*/*MKI67*/*TOP2A* (Appendix A1). Membership and source-specific score definitions are described in Appendix A1 and Table A1.

For the primary local analysis, each program score is regressed on an intercept, recorded fine-label indicators, raw depth, depth squared, log(1 + inherited full-assay total counts), and log(1 + relative tenth-neighbor spacing). Relative spacing divides the nonself tenth-neighbor distance by the median positive first-neighbor distance across the full native map. Residuals are centered and divided by their population SD over the finite fitted observations, which require finite depth and positive finite counts. Self is excluded from

[1] Reactome. (n.d.). Curated pathway and complex records: Glycolysis (R-HSA-70171); respiratory electron transport (R-HSA-611105); ATPase:ATP (R-HSA-164835); citric acid cycle (TCA cycle) (R-HSA-71403); mitochondrial fatty acid beta-oxidation (R-HSA-77289); detoxification of reactive oxygen species (R-HSA-3299685); apoptosis (R-HSA-109581); and cell cycle (R-HSA-1640170). Retrieved September 24, 2026, from https://reactome.org

spacing calculations. These scores are distinct from the uncentered whole-map scores and from separately label-centered or count/depth rank-adjusted comparisons. Appendices A1 and A3 specify source-available energetic definitions, coverage and scaling. Figure 5B,C uses disjoint 10/8 population contributions; Appendix A4 distinguishes the complementary original-common 15/13 cortical decomposition and source-available MLO contributions.

2.5. Cross-location spatial estimators

Bivariate spatial association measures incorporate the spatial arrangement of two variables [25]. Coordinate components use the symmetric ten-neighbor graph capped at four times median positive first-neighbor distance. For spatial scales, each exact score is map-standardized before analysis or residualized on an intercept, fine identity, raw depth, depth squared, log(1 + inherited full-assay total counts), and log(1 + relative tenth-neighbor spacing), then centered and standardized by its residual population SD. Relative spacing uses the median positive first-neighbor distance across the full native map as its denominator. Positive-distance pairs occupy (0,2], (2,4] and (4,8] multiples of median first-neighbor spacing, share a capped-graph component, and have a midpoint within two spacing units of observed support. Self-pairs are excluded. For focal i, each band's association contribution is (Equation (5)):

$$\frac{z^A{}_i(W_b z^B)_i + z^B{}_i(W_b z^A)_i}{2} \tag{5}$$

the operator averages eligible partners in band b. Map estimates give each supported focal observation equal weight. The same-observation cross-product uses the identical common focal mask, with at least 30 supported focal observations. The standardized cross-product is dimensionless; distance bands are defined relative to each map's median positive first-neighbor spacing.

The original 4-8 spacing band is a distance-resolved field statistic on full common focal support. Fixed native 4-by-4 bins within components define regional means, requiring ten observations per region and four regions. Appendix A2 describes regional-deletion sensitivity and the contextual source units. The actual-partner matched comparison beyond eight spacing units and the same-location decomposition are defined below. These comparisons are descriptive.

Source-specific coordinate calibration is described in Appendix A1, using the recorded Space Ranger registration for MLOs[2].

2.6. Whole-map effects and same-organoid pairings

Whole-map effects use source expression and the original-map population SD defined in Equation (6), with outer observations at the lower-tercile rule in Equation (7) and inner observations at the strict upper-tercile rule in Equation (7) among finite expression/depth values. Support requires thirty observations overall and ten per zone. Uzquiano depth is normalized distance to its 0.20 concave-hull boundary; MLO depth is one minus median-center radius divided by maximum radius. Both are inner-positive. An inclusive upper-tercile rule (depth $\geq q_2$) was evaluated separately as a boundary-sensitivity analysis (Appendix A2). Constant scores and unsupported effects are recorded as unavailable.

$$\sigma_s = \sqrt{\frac{1}{|V|}\sum_{i\in V}[(s_i - \mu_s)^2]};\ \mu_s = \frac{1}{|V|}\sum_{i\in V}[s_i] \tag{6}$$

[2] 10x Genomics. (n.d.). Space Ranger spatial outputs (version 2.1). Retrieved October 9, 2026, from https://www.10xgenomics.com/support/software/space-ranger/2.1/analysis/outputs/spatial-outputs

V is the finite score/depth support for this positional endpoint. The complete-map mean and SD use the population denominator, the number of observations in V.

$$O = \{i \in V : h_i \leq q_1\}, I = \{i \in V : h_i > q_2\} \tag{7}$$

Equation (7) defines the outer and inner sets from the one-third and two-thirds quantiles of inner-positive depth on finite score/depth support. The outer set includes its quantile boundary; the primary inner set lies strictly above its quantile boundary.

$$\Delta = \frac{\mu_I - \mu_O}{\sigma_s} \tag{8}$$

The regional means average s within I and O. Delta is their difference in original-map SD units. Label-centered and fixed-denominator variants are defined in Appendix Equations (A2)-(A3).

The same-organoid comparison joins adjusted disjoint local effects to strict disjoint uncentered regional effects by exact identity. The original-common regional comparison and source-available energetic pairs use their respective memberships. Directional configurations use zero as the sign boundary, with exact zeros and unavailable values recorded separately. Figure 2 examples minimize distance from each directional group's bivariate median, with ties resolved by identity; selection depends only on the two regional effects. Energetic examples are the first member in the recorded source order of each nonempty uncentered quadrant. The source-unit displays and native gallery show all individuals. Native figures use original coordinates and equal aspect; metabolic panels display Cartesian upward y. Calculations use the original analytical coordinates.

2.7. Actual-partner matching and same-location decomposition

The descriptive specification was fixed retrospectively, after inspecting the spatial fields and before evaluating these comparisons. It uses the scores, annotations, coordinates and graph components defined above, with focal observations supported across all three distance bands. The primary pair is disjoint glycolysis/hypoxia after fine-identity, depth/depth-squared, inherited-count and spacing adjustment. Adjusted scores and same-location residuals use the original map residual-SD scale.

For every directed admissible near pair (i, j) within two median-first-neighbor-spacing units, the distant pool comprises every finite observation j-prime with j's fine label, existing equal-width normalized-depth band and tissue component, at Euclidean distance greater than eight spacing units from i. The pool is matched to the near partner j. A near partner with an empty pool is removed from both sides. Remaining partners receive equal weight within each focal, and retained focals equal weight within map. The distant estimate averages every eligible partner. At least thirty retained focals support a map estimate. Focal and directed-partner retention are separately reported. The 4-8 distance band uses the original three-band support; the matched distant control lies beyond eight. Equal section means summarize the single iTracer organoid; counts are also supplied by section.

Write A for the existing adjusted disjoint glycolysis score and B for hypoxia. On all finite map observations, ordinary least squares fits Equation (9). Equation (10) defines the focal-average statistic for a row-normalized near or matched-distant operator W on the common retained focal set. The fitted and remaining components in Equations (11)-(12) reconstruct the original statistic. Equation (13) retains the intercept contribution on restricted focal support. Each organoid's near-minus-distant values are computed before source summaries.

$$B = \alpha + \beta A + e, f = \alpha + \beta A, e = B - f \tag{9}$$

A and B are the existing adjusted disjoint glycolysis and hypoxia vectors. Ordinary least squares over all finite map observations estimates the intercept and slope; f and e are fitted values and residuals. Residuals e

use the response-score scale.

$$C(W) = \frac{\sum_{i \in F}[A_i(WB)_i + B_i(WA)_i]}{2|F|} \tag{10}$$

W is a row-normalized near or actual-partner-matched distant operator; Equation (10) averages over the common retained focal set with equal focal weights. Near partners have equal weight within each focal. Distant weights reproduce the near partners' label, depth-band and component composition. Each operator's weights are used for all three of its decomposition terms.

$$F(W) = \frac{\sum_{i \in F}[A_i(Wf)_i + f_i(WA)_i]}{2|F|} \tag{11}$$

Equation (11) evaluates the fitted-score component with the same operator and focal set as the original statistic in Equation (10).

$$D(W) = \frac{\sum_{i \in F}[A_i(We)_i + e_i(WA)_i]}{2|F|}, C(W) = F(W) + D(W) \tag{12}$$

Equation (12) gives the remaining component. This decomposition is algebraic on the fixed score scales and retained support.

$$F(W) = \frac{\alpha}{2|F|} \sum_{i \in F}[A_i + (WA)_i] + \frac{\beta}{|F|} \sum_{i \in F}[A_i(WA)_i] \tag{13}$$

The intercept term is retained on the restricted focal set after full-map centering. The fitted and remaining components may each be positive or negative.

$$\Delta_T = T(W_{near}) - T(W_{distant}), T \in \{C, F, D\} \tag{14}$$

The near-minus-distant difference in Equation (14) is formed within each organoid for each component before source summaries. The matched distant operator uses controls beyond eight spacing units; the 4-8-spacing band is evaluated separately.

The fixed model-form check uses a natural cubic spline with three predictor degrees of freedom plus an intercept, with knots at the predictor's minimum, one-third and two-thirds quantiles, and maximum. Natural end constraints define a four-dimensional interpolation basis; an intercept plus three basis columns provides the fitted design. Fits with non-distinct knots or rank deficiency are unavailable. The reverse A-on-B mapping is a separate predefined sensitivity. All four fits use the same finite map observations, and their evaluations use identical paired support. Residuals are in response-score units. Linear, spline and reverse fits provide descriptive conditional decompositions.

Each sufficiently populated existing component-specific 4-by-4 native region (at least ten finite observations) is deleted from focal, near and distant pools. Original map-wide fitted values and score scales remain fixed; actual-partner controls and weights are reconstructed within the surviving pools, retaining original focal eligibility and the thirty-focal threshold. Deletion ranges describe sensitivity to regional support. Donor, capture and unit omissions summarize equal-organoid effects in overlapping subsets. Exact matrices are stored as near weights plus a factorized all-control operator, with point identities and group exclusions. Independent explicit partner enumeration and alternate fitted-basis reconstruction verify means and decomposition identities.

2.8. Regional support and population accounting

Population accounting uses each program's source-available, original-common or disjoint definition. Among finite score/depth observations, outer and inner regions use the original first and upper depth terciles. A fine population with at least five observations in each region supports symmetric composition and within-group

terms, given in Equation (16). Fractions include every finite regional observation, including other/unassigned groups. Unsupported populations contribute their observed region sums divided by full-region counts. All terms share the original complete-map population SD. Their sum reconstructs the complete-map difference. Individual partial terms can depend on score origin and offset one another; they are algebraic components of the regional contrast. Sections are averaged equally within their documented organoid.

$$p_{kR} = \frac{n_{kR}}{N_R}, \mu_{kR} = \frac{\sum_{i \in R_k}[s_i]}{n_{kR}} \tag{15}$$

In Equation (15), the region is either inner or outer. The population count includes observations of the specified population in that region, whereas the regional denominator includes all finite observations regardless of label. Zero-count population means are recorded as NA.

$$C_k = \frac{(p_{kI} - p_{kO})(\mu_{kI} + \mu_{kO})}{2\sigma_s}, H_k = \frac{(p_{kI} + p_{kO})(\mu_{kI} - \mu_{kO})}{2\sigma_s} \tag{16}$$

C and H denote composition and within-population terms for populations supported by at least five observations per region. All terms share the original complete-map SD.

$$U = \frac{\frac{\sum_{i \in I_{uns}}[s_i]}{N_I} - \frac{\sum_{i \in O_{uns}}[s_i]}{N_O}}{\sigma_s} \tag{17}$$

Equation (17) sums the inner and outer observations from unsupported populations. The unresolved term uses observed sums, including when a population is absent from the other region.

$$\Delta = \sum_{k \in K}[C_k + H_k] + U \tag{18}$$

K is the supported population set. These terms reconstruct the complete-map effect as shown in Equation (18). In the matched-support comparison, region fractions are renormalized over the retained IDs but the original complete-map SD is kept.

The cortical matched-support comparison includes fine labels meeting the same five-per-region criterion. Each fine label maps unambiguously to one inherited coarse category. Both resolutions use exactly the same retained inner/outer observation IDs, fixed original region assignments and original complete-map SD. Regional fractions are renormalized within this subset, with at least ten observations required per region. Both complete decompositions reconstruct the same restricted contrast; the original full-map contrast and retained regional fractions are reported separately. For a complete partition on fixed matched support, a common additive score offset leaves the summed composition and within-group terms invariant. Appendix A4 summarizes population-level contributions and matched-support comparisons. Six additional descriptive paired contrasts compare the disjoint within-population contributions: glycolysis minus hypoxia within each of four populations, and mixed neural minus stress/necrosis-associated for each program on their common organoids. These comparisons were selected after the earlier population results were known and specified before evaluating their paired differences. They use the five-per-region rule, finite intersection support and program-specific original-map SD. Equal-organoid paired summaries and omission ranges are reported in Table A3, with individual differences in Figure A2. The unresolved population contributes one descriptive organoid pair.

2.9. Inferential scope and supporting comparisons

Local signs, same-organoid configurations, partner-matched contrasts and fitted/remaining decompositions are descriptive source-unit comparisons.

The exploratory MLO donor-preparation associations describe overlapping strata within the sampled cohort (Appendix A4).

Paired *DCX* centering changes were evaluated separately within each cortical source. Equal-section organoid coefficients were differenced before averaging organoids equally within each Lin Control line; the Uzquiano inference units were its three organoids. A retrospective two-source family used exact two-sided sign tests against zero paired change, excluding ties within 10^{-12}, with BH2 and pointwise 95% percentile intervals from 10,000 resamples of the stated units (seed 20261013). Donor, donor-by-differentiation and capture omissions quantify sensitivity of MLO paired program and near-minus-distant summaries across overlapping cohort subsets (Appendices A2 and A3).

Baseline immature-fraction effects were also summarized at the hierarchy used for cortical inference. Equal-section organoid effects were averaged equally within each Lin Control line; the Uzquiano inference units were its three organoids. A separate retrospective family of two source-specific exact sign tests against zero used BH2, with absolute effects at most 10^{-12} treated as ties. The effect estimator is the median inference-unit effect; pointwise 95% percentile intervals use 10,000 unit resamples (seed 20261013, reset per source). The historical fourteen-organoid signed-rank family remains separate. Adjacency and within-state concentration use 5,000-resample effect intervals and complete twelve- and twenty-endpoint families. Percentile intervals estimate effect uncertainty by resampling; exact sign tests evaluate direction separately (Appendix A5).

The original-common CFuPN glycolysis-minus-hypoxia comparison uses the arithmetic mean of three within-organoid differences in original-map-SD contributions. A separate retrospective one-endpoint family was specified before evaluating the comparison. The two-sided exact paired sign-flip test enumerates all eight sign assignments, including the observed one, and counts absolute mean statistics at least as large as observed (tolerance 10^{-12}); q equals p in this BH1 family. This requires independent, sign-exchangeable paired organoid differences under the null. The original-common CFuPN contributions are summarized by arithmetic means with one sample SD (the square root of the sum of squared deviations from the sample mean divided by $n - 1$). Boxplots elsewhere show medians, quartiles and 1.5-IQR whiskers, with every observation displayed once. Joint-support MLO program contrasts and donor/capture omissions remain descriptive. The paired effect and its exact probability are reported in Appendix A4.

2.10. Population-specific regional and SYT1 models

Regional population accounting used 24 MLOs and three month-2 cortical organoids, with sources analyzed separately. Original-common 15/13 and disjoint 10/8 glycolysis/hypoxia definitions share observation identities and full-map depth-tertile boundaries. Each score uses its own complete-map population SD, fixed under restriction. Fine populations require five observations in each region, and matched retained tissue requires ten observations per region. Coarse and fine symmetric composition/within terms use identical retained IDs; unsupported full-map contributions are recorded separately (Appendix A4).

Organoid-level composition models use ordinary least squares for coefficients and R^2 (Appendix A4). With H denoting the disjoint hypoxia inner-minus-outer effect in original-map SD units, the baseline model is given by Equation (19)

$$H = intercept + stage\ indicators + donor\ indicator + log(N) \tag{19}$$

and the expansion adds immature fraction f. Stage is categorical, with day 120 as reference and indicators for days 40 and 70; Donor1 is the reference. N counts all sampled spots in the organoid. The full model has 24 organoids, rank six and 18 residual degrees of freedom; slopes per ten percentage points are 0.1 times the fraction coefficient. The same baseline/expansion predicts regional glycolysis, source-available OXPHOS or adjusted disjoint full-near correspondence in B1, B3 and B4. B5 instead predicts mean immature-population

SYT1 from stage, donor, $log(N)$ and f, adding full-near correspondence. OLS requires twelve finite organoids, full rank after constant-column removal and five residual degrees of freedom. Ten omission sensitivities remove each of two donors, four donor-specific differentiations or four captures; constant donor indicators are omitted in single-donor fits. The two pooled differentiation-index diagnostics are separate because those indices span distinct donors.

Relative radial position was defined in Equation (20) as

$$S = mean(d|immature - neuron\ label) - mean(d|all\ sampled\ spots) \tag{20}$$

using inward-positive MLO depth d. Positive S places the population inward of the sampled map mean. Population position and regional expression use the same map geometry. Eligibility required at least thirty total spots, five immature-labeled and five other spots, and finite depth for every sampled spot; all 24 organoids qualified. The primary outcome was disjoint H8 regional enrichment and the prespecified secondary outcome disjoint G10 enrichment. Outcome values, residuals, MAE and RMSE use original-map SD units; MSE uses (original-map SD)². Percentage error changes are dimensionless and use the stated baseline-error denominators. The baseline contained categorical stage, donor, log total spot count and whole-organoid immature fraction; the extension added S. Descriptive OLS fits and alpha-one ridge organoid/capture holdouts used identical units. Category levels and scaling were learned within training folds. The spatial descriptor and secondary outcome were specified after reviewing prior composition results but before evaluating the spatial association. Annotation sensitivity recomputed fraction and S from the saved overlap-removed labels. Appendix A4 specifies model ranks, error denominators, omission sensitivity and prediction summaries. These analyses describe associations and prediction within the sampled two-donor cohort.

The annotation-module overlap-removal sensitivity removes every primary disjoint glycolysis/hypoxia member from every competing MLO marker module. The actual overlap comprises *BNIP3* and *VEGFA* in the nine-member stress/necrosis module; seven members remain. Scores are unweighted means of measured exact-symbol members in full-assay log1p(CP10000) expression. Eligibility requires at least three measured retained members and at least half the original requested membership, specified before evaluating the sensitivity. All seven modules satisfy both rules in all 24 organoids. Ineligible modules are excluded from competition. Assignment uses the maximum score with source insertion-order tie resolution; Unresolved is assigned when the maximum is nonpositive or zero modules are eligible. The denominator of f includes every spot, including newly unresolved assignments. This sensitivity varies annotation scores, labels and derived population predictors with a fixed hypoxia outcome, regional boundaries, SDs, covariates and folds. The local label-adjusted correspondence endpoint is a separate analysis. MLO population labels were assigned with the project marker modules; source spatial-region keys identified tissue regions. Appendix A4 summarizes assignment changes and their effects on conditional associations and prediction.

Observation-level *SYT1* models are fitted within immature-labeled observations of each source. The baseline includes depth, depth squared, log(1 + inherited full-assay counts), log(1 + d10/NN), local immature fraction, focal disjoint glycolysis and hypoxia scores, and applicable fine-label indicators. d10 is the nonself tenth-neighbor distance and NN is the full map's median positive first-neighbor distance. The expansion adds the two mean nonself-neighbor metabolic scores. Neighbors are in the same capped-k10 coordinate component, at positive distance at most 2 NN, with a pair midpoint within 2 NN of a sampled location. The separate far-band sensitivity uses 2-4 NN. Focal metabolic scores use complete-map means and population SDs; neighbor means average those standardized values. In per-map OLS coefficient displays, predictors and *SYT1* are standardized with their population SDs among the finite eligible immature observations. Prediction uses source-scale *SYT1*: log1p(CP10000) in MLOs and supplied expression fractions in cortex. Coefficient support requires thirty observations, five *SYT1* detections, variable *SYT1*, at least three observations per fitted

rank and ten residual degrees of freedom. Prediction support requires ten finite immature observations and five detections. These rules support 16 MLO coefficient estimates, 19 MLO prediction units and three cortical organoids for each analysis. Appendix A4 reports prediction support and source-level errors. The MLO assigning modules exclude *SYT1*; cortical assignments are the source-supplied transferred labels.

Held-unit prediction uses weighted ridge regression, separately within each source, with objective

$$\sum_i[w_i(y_i - b_0 - x_i{}^T b)^2] + \alpha \sum_j[b_j{}^2] \tag{21}$$

The intercept in Equation (21) is unpenalized and α = 1 on this weighted-sum scale. For K training organoids containing N training observations, each row in organoid u receives

$$w_i = \frac{N}{K n_u} \tag{22}$$

where the subscripted n is that organoid's training-row count; the weights in Equation (22) sum to N and each organoid has equal total weight. In the one-row-per-organoid models, every weight is one. Nonintercept predictors are centered and scaled by weighted training means and population SDs; near-constant SDs are set to one. Category levels and scaling are derived from each training fold. Entire organoids or captures are held out; baseline and expanded models use identical evaluation observations. Source-level MSE averages within each held organoid and then equally across organoids, including under capture holdout. Metabolic score standardization uses values from the complete test map, which is assumed available at prediction time; SYT1 is used only as the outcome. This measurement transform is distinct from training-only predictor scaling in the ridge model. Buffered component-by-4x4-region deletions remove focal rows whose neighborhoods touch the deleted region, using full eligible-population coefficient units. Their ranges describe sensitivity to regional support. Results are reported as descriptive coefficients and held-out prediction errors.

3. Results

3.1. Local glycolysis-hypoxia correspondence recurs across organoid sources

Glycolysis-associated expression at one location corresponded positively to hypoxia-response expression nearby across the cortical and MLO cohorts (Figure 1). The comparison includes three month-2 cortical organoids and 24 MLOs at days 40, 70 and 120. The local statistic averages products of adjusted, residual-standardized program scores at different nearby locations. Positive products can arise when both scores deviate above their adjusted means or when both deviate below them. Under the primary disjoint definition, adjusted near cross-products were positive in 3/3 cortical organoids and 23/24 MLOs, with medians +0.0301 and +0.1112. The original-common comparison, which includes five shared genes, was positive in 3/3 and 22/24, with medians +0.0787 and +0.2054.

The local relationship links adjusted glycolysis at one location with hypoxia-response transcription at distinct locations within two map-relative neighbor-spacing units. Each eligible focal observation receives equal weight, and the statistic averages the two program directions. Adjustment uses recorded fine identity, depth and depth squared, inherited full-assay counts and local spacing. The observations are Slide-seq beads in cortex and Visium spots in MLOs.

The distance profile varied between sources. Each distance comparison used the same set of eligible focal observations with finite scores, termed common focal support. Across three distance bands, the MLO median disjoint cross-product approached zero at 4-8 spacing units (+0.000497), whereas the cortical median remained +0.0301. Distance is expressed relative to each source map's neighbor spacing.

Pairing with ATP-related transcription supplied a contrasting endpoint. Adjusted glycolysis-ATP and hypoxia-ATP near effects were negative in all three cortical organoids, while MLO directions were mixed, with 13/24 and 8/24 positive, respectively (Appendix A2).

Nearby correspondence was also positive in the iTracer organoid and patterned DV/RC maps. The day-60 iTracer organoid, represented by three serial sections, had a small positive disjoint near effect (+0.005455); its matched near-minus-distant effect was also positive (+0.006943). All five patterned DV/RC maps had positive adjusted disjoint near effects, spanning +0.110996 to +0.239861, and positive matched contrasts. The contextual comparisons summarize one iTracer organoid and five patterned model/stage maps (Appendix A2).

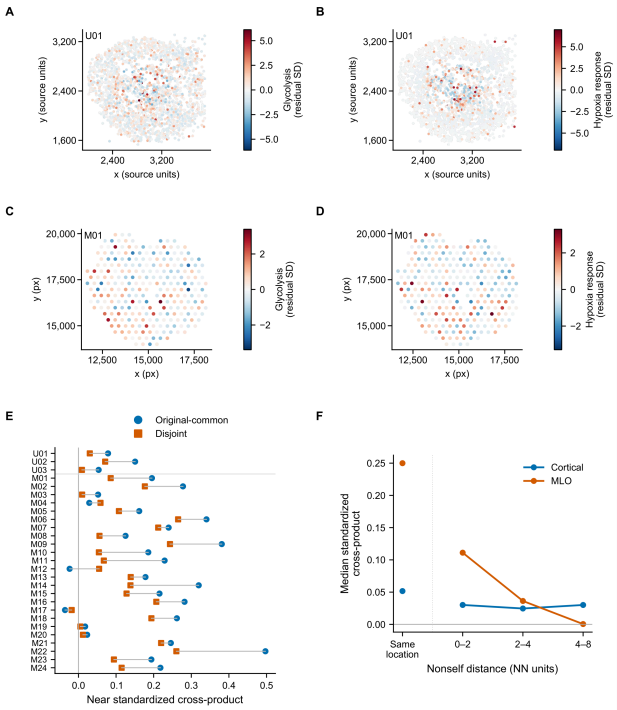


Figure 1. Glycolysis-hypoxia correspondence recurs between nearby observations. (A) Adjusted glycolysis scores across Slide-seq beads in cortical organoid U01. (B) Adjusted hypoxia-response scores in the same beads. (C) Adjusted glycolysis scores across Visium spots in midbrain-like organoid M01. (D) Adjusted hypoxia-response scores in the same spots. (E) Adjusted near cross-products for the original-common and disjoint gene sets, paired within each organoid. (F) Median disjoint cross-products in each cohort at the same location and across three distance bands

Notes: Nearby correspondence compares glycolysis at one sampled location with hypoxia-response transcription at other locations, and averages the two program directions with equal focal weighting. The four native maps use disjoint scores and show the first organoid in recorded source-ID order from each primary cohort. Colors show glycolysis residual score (SD) in A,C and hypoxia-response residual score (SD) in B,D, each over its full symmetric range. U01 uses native source units; for M01, px denotes full-resolution image pixels. All four maps use equal aspect and upward Y. Panel E includes all three month-2 cortical organoids and 24 MLOs. F, Equal-organoid medians of adjusted disjoint cross-products on common three-band focal support within each map. E,F show means of products of residual-standardized program scores; the result is a dimensionless cross-location statistic. The isolated same-location point compares both scores at one observation; lines connect the nonself distance bands. NN denotes the full native map's median positive first-neighbor spacing. Covariate adjustment and equal-focal averaging are defined in main Methods, "Cross-location spatial estimators". Original-common scores use 15 glycolysis/13 hypoxia

members; disjoint scores use ten/eight after five shared genes are removed. Table A1 gives exact members; Appendix A2 describes support and partner-matched controls.

3.2. The same local relationship accompanies different regional configurations

Positive local correspondence accompanied jointly inward, jointly outward and opposing whole-map distributions in identified MLOs (Figure 2). Using the primary disjoint glycolysis and hypoxia scores, fifteen organoids were jointly inward, five jointly outward and four opposed: one glycolysis-inward/hypoxia-outward and three glycolysis-outward/hypoxia-inward. The adjusted disjoint full-support near effect was positive in all fifteen inward pairs, all five outward pairs and three of the four opposing cases.

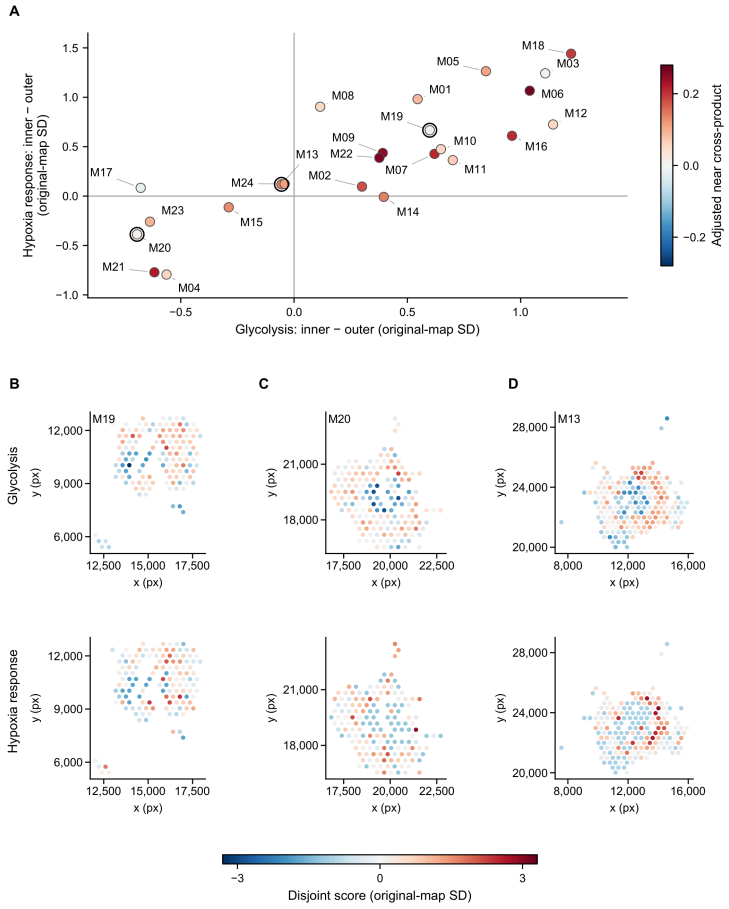


Figure 2. Local correspondence and global arrangement in the same MLOs. (A) Inner-minus-outer glycolysis and hypoxia-response effects in 24 MLOs, colored by each organoid's adjusted near cross-product. (B) Glycolysis (top) and hypoxia-response (bottom) maps in M19, with both programs enriched inward. (C) Corresponding maps in M20, with both programs enriched outward. (D) Corresponding maps in M13, with glycolysis enriched outward and hypoxia-response expression enriched inward

Notes: A, All 24 MLOs: uncentered disjoint glycolysis (10 genes) and hypoxia (8 genes) inner-minus-outer effects in original-map SD units, colored by the same organoid's adjusted full-support near cross-product. The strict inner zone is depth > q_2 and the outer zone is depth ≤ q_1. Fifteen organoids are jointly inward, five jointly outward and four opposed. Configurations classify the directions of the two regional expression contrasts. The original-common 17-inward/7-outward definition is a separate comparison. Open rings in A identify the organoids shown in B-D. Each mapped organoid is the one nearest its directional group's bivariate median global effect, with display-ID tie-breaking. For B-D, the global glycolysis/hypoxia effects (original-map

SD), adjusted full-support local near cross-products and sampling support are, respectively: M19, +0.5987/+0.6666, +0.0051, 130 total spots and 125 focals; M20, -0.6930/-0.3873, +0.0119, 153 total spots and 152 focals; M13, -0.0557/+0.1192, +0.1389, 187 total spots and 183 focals. All observations use native coordinates in px (full-resolution image pixels), upward Y and equal aspect. Colors show whole-map standardized glycolysis or hypoxia-response scores, with the same symmetric, unclipped range across these panels and Figure 4. Figure 3 shows the partner-matched estimator on different retained focals and partner weights: M19 and M20 have small positive full-support values here but negative original matched-near values there. Global uncentered contrasts and adjusted local cross-products are distinct estimators. Appendix A2 distinguishes the full-support and matched estimators and the strict and inclusive boundary definitions.

The local and global endpoints use the same gene memberships to measure cross-location correspondence and regional enrichment, respectively. Global effects are uncentered inner-minus-outer differences divided by each program's original-map SD; local cross-products use the adjusted score deviations. Native paired maps show the organoid nearest each directional group's bivariate median global effect. Each example links its actual full-support local cross-product to both global effects, with total map observations distinguished from supported focal observations (Figure 2; Appendix A2). The small positive full-support values for M19 and M20 become negative under the original matched-near estimator in Figure 3, which uses different partners and focals.

The original-common positional comparison gives seventeen jointly inward and seven jointly outward organoids, paired with positive disjoint near effects in seventeen and six, respectively (Appendix A2). Its five shared genes and resulting score values distinguish this comparison from the primary 15/5/4 configuration. Both definitions use the same observations and strict regional thresholds.

3.3. Nearby correspondence exceeds matched distant relationships in most organoids

Nearby correspondence exceeded the matched distant comparison in most eligible organoids (Figure 3). This comparison asks whether the relationship is stronger locally than across more distant tissue with comparable partner identities and positions. Adjusted disjoint scores and matched partner composition were available in all three cortical organoids and nineteen MLOs. On the same eligible focal observations, near/distant medians were +0.0283/+0.0037 in cortex and +0.1637/-0.1085 in MLOs. Paired near-minus-distant effects were positive in 2/3 and 16/19, with medians +0.0277 and +0.2728. The MLO contrast combines positive nearby correspondence with a negative distant relationship.

Positive nearby correspondence also persisted after accounting for the fitted linear relationship between the two programs at the same locations. Near remainders on the common matched focal set were positive in 3/3 cortical organoids and 16/19 MLOs, with medians +0.0258 and +0.1179. Distant remainders had medians +0.0027 and -0.0841; paired near-minus-distant remainders were positive in 2/3 and 16/19, with medians +0.0250 and +0.2062. The fixed nonlinear mapping retained these direction counts. Figure 3C summarizes the within-organoid contrasts; each statistic uses the same focal set, partner weights and score units for its original, fitted and remaining components (Appendix A2).

A positive relationship at both distances can yield a small or negative near-minus-distant contrast; this occurred in one cortical organoid. Prediction direction and regional deletion affected some individual contrasts, while positive MLO median near remainders persisted after either donor and every capture omission. Appendix A2 and Figure A1 report changes in support and model-form and regional-deletion sensitivities.

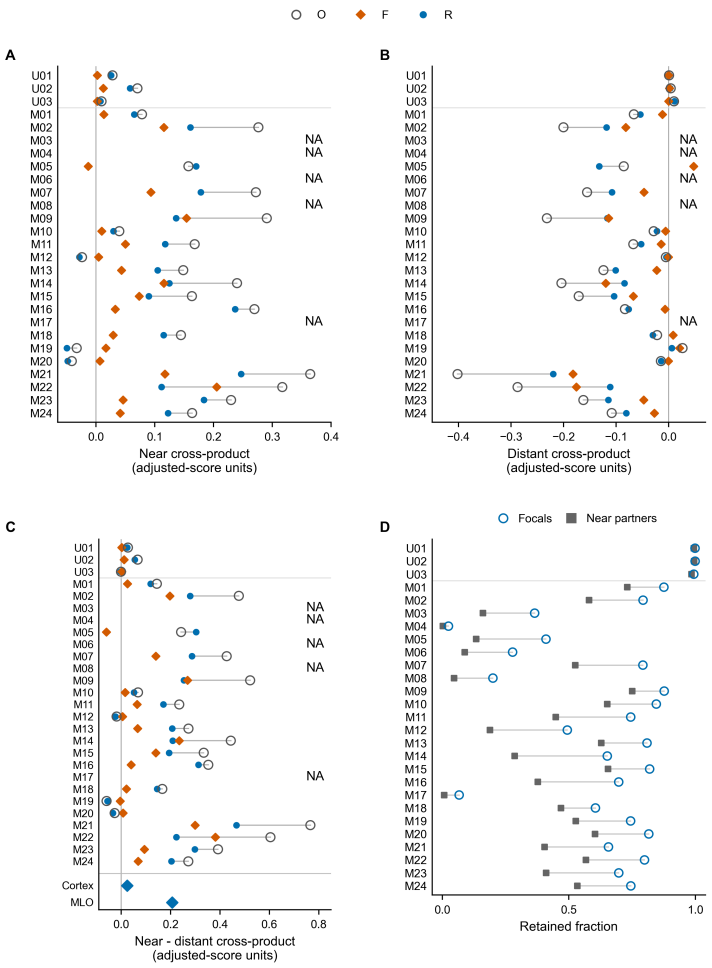


Figure 3. Partner-matched correspondence and its within-organoid distance contrast. (A) Near glycolysis–hypoxia cross-products by organoid, showing the original estimate, its fitted same-location component and the remainder. (B) The same quantities for matched distant partners. (C) Near-minus-distant contrasts within each organoid and median remaining contrasts by cohort. (D) Fractions of focal observations and directed near partners retained after matching

Notes: A, Adjusted glycolysis at one focal location paired with hypoxia-response expression at distinct near locations; the two program directions are averaged. B, Distant partners matched to each actual near partner's recorded fine identity, depth band and component. Near distances are at most 2 NN and distant distances exceed 8 NN, where NN is the full native map's median positive first-neighbor spacing. C, Within-organoid near-minus-distant contrasts: positive values indicate stronger near correspondence, zero equal estimates and negative values stronger distant correspondence. Each row in A-C is one identified organoid; the cohort contains three cortical organoids and 24 MLOs, of which three and nineteen support the matched comparison. O (open circle) is the original matched statistic on covariate-adjusted disjoint scores; F (orange diamond) is its fitted same-location component; R (blue circle) is the remainder, with O = F + R. Gray segments connect O and R. Products of matching positive or negative residual-standardized deviations contribute to a positive cross-product. The statistic is dimensionless. The linear hypoxia-on-glycolysis relationship is fitted across all finite map observations. Near and distant statistics are then evaluated on their common matched focal set; within each statistic, its own partner weights are used for the original, fitted and remaining components. Blue diamonds below C show descriptive median remaining contrasts: +0.0250 for the three cortical organoids (2/3 positive) and +0.2062 for nineteen MLOs (16/19 positive). In D, open circles divide matched focal counts by original common-three-band focal counts; squares divide retained directed near-partner counts by original

directed near-partner counts. NA in A-C denotes fewer than thirty matched focals: M03 = 27, M04 = 1, M06 = 20, M08 = 11 and M17 = 3. Regional and hierarchy omissions characterize subset sensitivity (Appendix A2; Figure A1).

3.4. Concordant energetic pairings accompany heterogeneous cohort distributions

Within-organoid energetic pairings differed from the impression given by separate cohort summaries: glycolysis and OXPHOS were enriched in the same region in 18 of 24 MLOs, although their cohort medians pointed in opposite directions (Figure 4). Eleven pairs were jointly inward, seven jointly outward and six opposed. Seventeen inward glycolysis effects and thirteen outward OXPHOS effects produced the opposite cohort medians. ATP-related and glutamate/glutamine pairs were concordant in 17/24 and 16/24, respectively, with different same-organoid combinations (Appendix A3). Each energetic comparison uses its source-available program definition.

Positive glycolysis-OXPHOS relationships predominated within recorded MLO populations (Figure 4B). Spearman correlations were positive in 52/57 supported population-organoid records: 15/18 dopaminergic-neuroblast, 20/22 mixed-neural, 14/14 stress/necrosis-associated and 3/3 unresolved records. Each coefficient compares paired source-available scores across at least twenty spots with variation in both programs; the 57 records are nested within 23 of the 24 MLOs (Appendix A3).

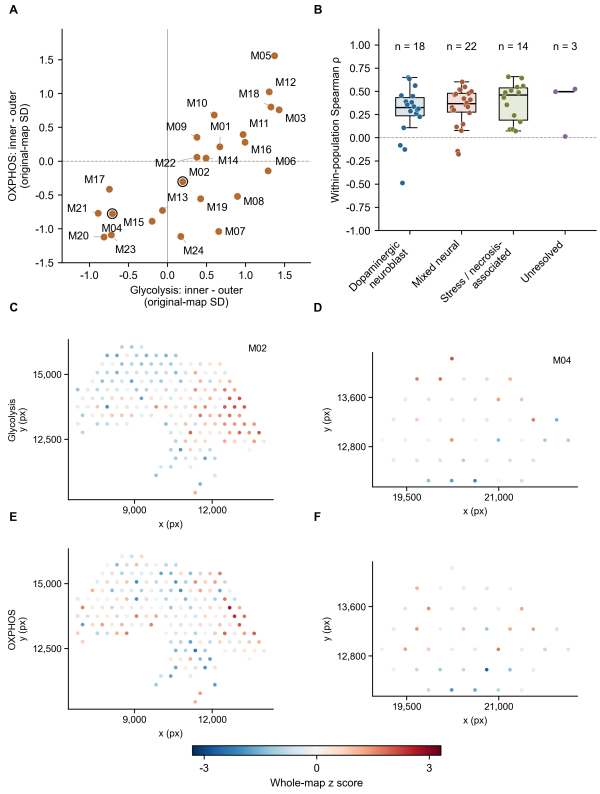


Figure 4. Energetic regional pairings coexist with positive within-population correlations. (A) Paired inner-minus-outer glycolysis and OXPHOS effects in 24 MLOs. (B) Within-population correlations between glycolysis and OXPHOS scores across spots, grouped by recorded population. (C) Glycolysis scores in M02. (D) Glycolysis scores in M04. (E) OXPHOS scores at the same spots in M02. (F) OXPHOS scores at the same spots in M04

Notes: A, All 24 MLO source-available glycolysis/OXPHOS inner-minus-outer effect pairs, in original-map SD: 11 jointly inward, seven jointly outward, six glycolysis-inward/OXPHOS-outward and zero glycolysis-outward/OXPHOS-inward. Guides identify organoids; rings mark M02 and M04. B, All 57 finite within-population, same-observation glycolysis-OXPHOS Spearman coefficients are nested within 23 of the 24 MLOs and colored and grouped by recorded spot population; jitter separates observations along the categorical axis. Boxes with black outlines and translucent population-colored fills for 18 dopaminergic-neuroblast, 22 mixed-neural and 14 stress/necrosis-associated records show medians and interquartile ranges; whiskers reach the most extreme observations within 1.5 interquartile ranges. Each point, including outliers, represents one supported population-organoid coefficient; n counts those records. The three-record unresolved group is displayed as individual points and a median bar. Support requires at least twenty paired spots and variation in both scores. Positive counts are 15/18, 20/22, 14/14 and 3/3, respectively; these counts summarize coefficient directions. The reference line denotes Spearman rho = 0. This same-observation source-available endpoint differs from the adjusted cross-location disjoint glycolysis-hypoxia statistic in Figure 3. The two source-available programs use disjoint gene sets. The maps include 174 spots in M02 and all 41 in M04. Colors show whole-map standardized glycolysis scores in C,D and OXPHOS scores in E,F, using the shared symmetric unclipped range of Figure 2. Coordinates use px (full-resolution image pixels), with upward Y and equal aspect. M02 is the first opposing example in recorded source order; M04 is the jointly outward example. Appendix A3 and Table A2 report the energetic configurations, population support and hierarchy sensitivities; Table A1 specifies the source-available glycolysis/OXPHOS memberships.

Paired concordance persisted across donor, differentiation and capture omissions, whereas the opposite cohort medians depended on the subset (Appendix A3). Uncentered glycolysis/OXPHOS effects were concordant in 10/12 organoids after omitting Donor1 and 8/12 after omitting Donor2, in 12-14/18 after donor-by-differentiation omissions, and in 13-14/18 after capture omissions. The OXPHOS median changed from -0.2219 in the full cohort to +0.2830 after removing Donor2, +0.0541 after removing Donor2_dif2, and +0.0359 after removing capture GSM8513021, while glycolysis remained positive in these subsets. Table A2 reports direction counts for all six pair/version combinations; Appendix A3 summarizes hierarchy sensitivity.

3.5. Population representation and position relate to regional metabolic organization

Peripheral population prevalence coexisted with inward expression contributions in the cortical organoids (Figure 5A,B). Regional fraction measures a population's share of all observations in a region; its within-population contribution weights the inner-outer expression difference by the mean regional fraction and divides by the program's original-map SD. The Corticofugal Projection-Neuron (CFuPN) population was more prevalent outward in all three month-2 organoids, yet its disjoint glycolysis and hypoxia contributions were positive. The U02 hypoxia contribution was +0.0023 SD. Population abundance and expression within that population therefore contributed differently to the regional pattern.

Mixed-neural and dopaminergic-neuroblast populations supplied predominantly inward disjoint contributions in MLOs, whereas stress/necrosis-associated contributions had mixed directions (Figure 5C). Within the same organoids, mixed-neural contributions exceeded stress/necrosis-associated contributions in 10/11 glycolysis pairs and 8/11 hypoxia pairs, with median differences of +0.1253 and +0.1713 program-specific original-map SD. These descriptive pairs comprised one Donor1 and ten Donor2 organoids (Appendix A4). Table A3 and Figure A2 report paired population support, the singleton estimate and within-population program differences.

Regional accounting was sensitive to both sampled tissue and annotation resolution. Restricting cortical comparisons to fine populations present in both regions removed much of the inner tissue in two organoids. In the third, finer labeling changed the sign of the within-group term on identical observations. Both decompositions reconstructed the same restricted regional contrast. The matched-support and alternative-score comparisons distinguish these sources of variation (Appendix A4).

Across the 24 MLOs, a greater fraction of immature-neuron-labeled spots was associated with stronger regional hypoxia-response enrichment inward (Figure 5D). After adjustment for categorical stage, donor and

log total sampled-spot count, a ten-percentage-point greater fraction was associated with an increase of 0.154 original-map SD in the disjoint inner-minus-outer hypoxia effect. The fraction denominator includes every sampled spot (Appendix A4).

Relative population position added a distinct conditional association (Figure 5E). Position is the mean inward-positive depth of immature-labeled spots minus the mean depth of all sampled spots. After adjustment for stage, donor, log spot count and immature fraction, a 0.1-unit more inward relative position was associated with a 0.306-SD less inward disjoint hypoxia contrast (Appendix A4).

The fraction coefficient remained positive and the position coefficient negative after removing the shared *BNIP3* and *VEGFA* genes from the competing stress/necrosis annotation module and across all ten donor, donor-specific differentiation and capture omissions. The hypoxia score was the same in these annotation comparisons (Appendix A4).

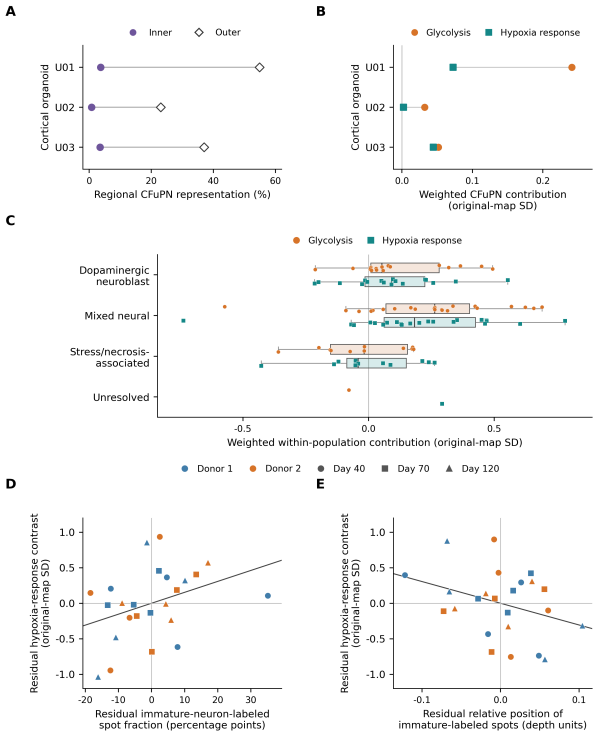


Figure 5. Regional metabolic expression in relation to population representation, within-population expression and relative position. (A) Corticofugal projection-neuron fractions in the inner and outer regions of three cortical organoids. (B) Weighted contributions of expression within this population to the regional glycolysis and hypoxia-response contrasts. (C) Corresponding contributions by recorded MLO spot population. (D) Adjusted association between the immature-neuron-labeled spot fraction and the regional hypoxia-response contrast across 24 MLOs. (E) Association between the relative depth of these spots and the same contrast after additional adjustment for their fraction

Notes: A,B, Three month-2 cortical organoids (U01-U03) with source-recorded corticofugal projection-neuron (CFuPN) assignments. A, CFuPN percentages among all original regional observations, including unresolved assignments. Purple circles denote inner regions and open diamonds outer regions; segments join each organoid. Inner/outer CFuPN counts are 26/387, 14/406 and 14/147, with respective regional denominators of 705/705, 1,759/1,759 and 397/397 observations. B, Disjoint within-CFuPN contributions: orange circles, glycolysis; teal squares, hypoxia response (also in C). Segments pair programs within each organoid. Each term weights the within-population inner-minus-outer expression difference by the mean regional population

fraction and divides by the program's SD across the original complete map (divisor N; Equation (16)). Glycolysis contributions are +0.241, +0.032 and +0.052 SD; hypoxia contributions are +0.073, +0.0023 and +0.045 SD. Zero denotes a weighted within-population contrast of zero. C, All 102 supported population-program-organoid terms, nested within 24 MLOs. Each program has 17 dopaminergic-neuroblast, 22 mixed-neural, 11 stress/necrosis-associated and one unresolved organoid contribution. These are recorded multicellular spot assignments. Support requires at least five observations per population in each region; unsupported terms are missing. Points show every supported term with categorical jitter. Boxes show quartiles and medians; whiskers reach observations within 1.5 interquartile ranges. Outliers are included; singleton groups are displayed as individual points. The same organoid may contribute several population and program records. Figure 4B instead requires twenty paired spots within a population over the map and variation in both scores; that distinct same-observation endpoint has different support. Figure A2 and Table A3 show the same-organoid program and population differences and descriptive hierarchy sensitivities. B,C use disjoint ten-gene glycolysis and eight-gene hypoxia-response scores with program-specific original-map SDs. Positive terms favor inner regions. Appendix A4 describes complete accounting, matched-tissue comparisons and the complementary original-common CFuPN comparison; Table A1 specifies primary program membership. The 24 MLOs in D and E came from two donors, two differentiations per donor and four captures, at days 40, 70 and 120. The outcome is the disjoint hypoxia inner-minus-outer effect across all eligible regional spots in original-map SD. Both variables are residualized against an intercept, categorical stage, donor and log total sampled-spot count. Horizontal residuals are fraction percentage points, with all spots in the fraction denominator; vertical residuals use original-map SD units. The adjusted OLS line represents +0.154 SD per ten percentage points. In D,E, blue/orange identify Donor1/Donor2 and circles/squares/triangles identify days 40/70/120; gray guides mark zero residuals. Annotation sensitivity, hierarchy omissions and held-unit predictions are in Appendix A4 and Figure A3. Panel E uses the original labels. Relative position S is mean depth among immature-labeled spots minus mean depth among all spots, with positive depth directed inward. S and the disjoint regional hypoxia effect are residualized against an intercept, categorical stage, donor, log total sampled-spot count and immature fraction. The horizontal axis uses residual depth units; vertical residuals use the outcome's original-map SD units. The OLS line represents −0.306 SD per 0.1-unit greater S. Each point represents one organoid. Position fits, sensitivities and predictions are reported in Appendix A4, Figure A4 and Table A4. Annotation and hierarchy omissions assess coefficient sensitivity; organoid and capture holdouts evaluate prediction within the two-donor cohort.

Prediction performance varied with the holdout scheme and annotation definition. Adding fraction reduced mean organoid-held hypoxia prediction error. Under capture holdout, error decreased with the original labels and increased after annotation-module overlap removal. Adding relative position to the fraction-adjusted model reduced mean prediction error under both holdouts and annotation definitions; error increased in one capture under both definitions. Omitting M04 from the overlap-removed organoid-held error summary yielded a 2.57% error increase. The prespecified glycolysis comparison increased mean prediction error under both holdouts and annotation versions (Appendix A4). Appendix A4 and Table A4 report fit and prediction summaries.

The exploratory relationship between immature-labeled spot fraction and full-support local correspondence varied across donor preparations, with a positive pooled association and opposite within-donor directions. This comparison uses a different outcome from the regional hypoxia model (Appendix A4).

Adding local correspondence increased mean held-out prediction error for immature-population *SYT1* across 19 supported MLOs. In the observation-level models, adding neighboring program scores increased equal-organoid mean error in both MLOs and cortex and increased MLO capture-held error (Appendix A4).

3.6. Cortical sources share peripheral neuronal positioning and annotation-sensitive local interfaces

The parallel cortical comparison examined whether shared broad neuronal positions accompany shared local interfaces and molecular relationships. Immature-neuron peripheral representation recurred across the two cortical sources (Figure 6A-C; Appendix A5). Inner-minus-outer population-fraction effects were negative in 12/14 Lin organoids and all three month-2 Uzquiano organoids, with medians -0.2549 and -0.2440. Fractions

include all observations in each region, including other and unresolved assignments. The separate DCX/STMN2 expression-position endpoint also showed negative median correlations with inner-positive depth (-0.1083 and -0.1573). Both the fraction and expression measurements favor peripheral early-neuronal positioning.

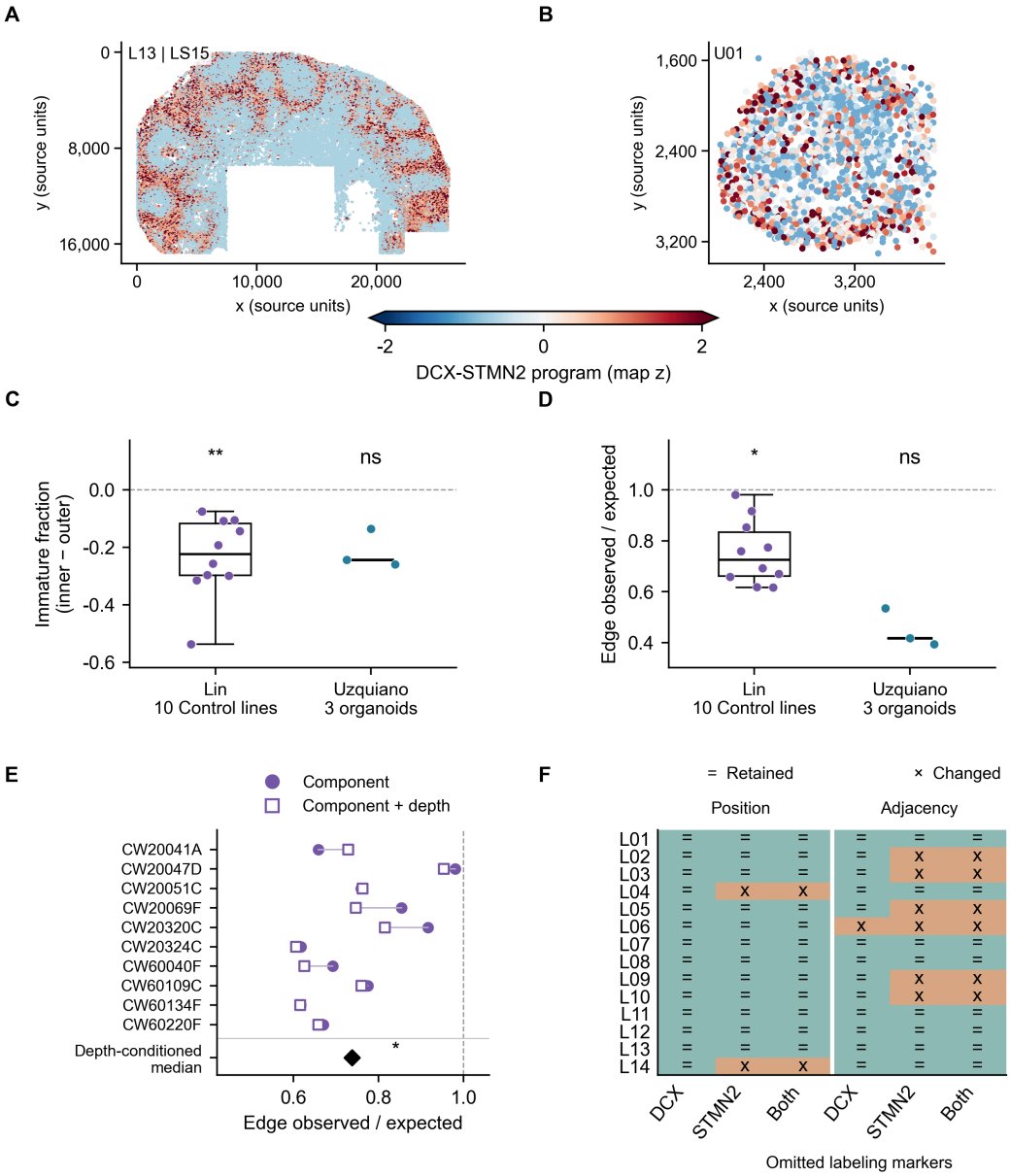


Figure 6. Peripheral neuronal representation recurs while local interfaces depend on population definition. (A) DCX–STMN2 program scores in Lin section L13/LS15. (B) The same program in Uzquiano organoid U01. (C) Inner-minus-outer immature-neuron fractions in both cohorts. (D) Observed-to-expected ratios of progenitor–immature-neuron edges under component-conditioned expectations. (E) Edge ratios for each Lin Control line under component-only and component-by-depth expectations. (F) Direction of position and adjacency effects relative to baseline after omitting DCX, STMN2 or both from labeling in 14 Lin organoids

Notes: The maps show the mean of within-map *DCX* and *STMN2* z scores, standardized again for display. Colors show expression and saturate beyond ± 2 map SD; native coordinates use equal aspect and downward Y in each source's own coordinate units. C, Inner-minus-outer immature-neuron fractions in ten Lin Control-line summaries and three month-2 Uzquiano organoids. Sixteen Lin sections contribute equally to fourteen organoid means, which contribute equally within each Control line. Regional fractions include all recorded observations. Lin and Uzquiano median effects are -0.2245 (95% percentile CI -0.3056 to -0.1077; p = 0.0020, q = 0.0039) and -0.2440 (95% percentile CI -0.2594 to -0.1359; p = 0.25, q = 0.25), respectively. Exact source-specific sign tests compare each effect with zero in the baseline-fraction BH2 family. D, Component-conditioned progenitor-immature edge observed/expected ratios using the same biological inference units: Lin median 0.7257 (95% percentile CI 0.6436 to 0.8535; p = 0.0020, q = 0.0117); Uzquiano 0.4174 (95% percentile CI 0.3932 to 0.5352; p = 0.25, q = 0.3). These tests compare each source with one in the twelve-endpoint adjacency family. C,D, Lin boxes summarize ten Control-line values; the three Uzquiano organoids are shown individually with a median bar. E, Each row connects the same Control line under component (filled circle) and component-by-depth (open square) expectations. The black diamond marks the median depth-conditioned ratio across ten Control lines (0.7375); the symbol denotes an exact sign test of these ratios against one (p = 0.0020; BH12 q = 0.0117). The capped directed k10 graph excludes self and combines both cross-population edge directions. F, All fourteen Lin organoids under marker omissions: teal (=) indicates retained baseline direction and tan (×) changed direction, relative to zero for position and one for adjacency. Identical contributing sections are used across definitions; six organoids cross the adjacency reference after *STMN2* or double omission. Boxes span the first and third quartiles, with a black median line; black whiskers extend to the most extreme observations within 1.5 interquartile ranges. All observations, including outliers, are shown. Percentile confidence intervals resample biological inference units (main Methods, 'Inferential scope and supporting comparisons'). Three-organoid bootstrap intervals are discrete; the permutation-null distributions and exact sign tests are defined separately in Methods. Appendix A5 gives the individual marker-omission transitions and distinguishes organoid- and Control-line-level summaries. Symbols denote adjusted q values: ns, $q \geq 0.05$; *, $q < 0.05$; **, $q < 0.01$; ***, $q < 0.001$; ****, $q < 0.0001$.

Progenitors and immature neurons were less adjacent than expected from their component-specific frequencies in both sources (Figure 6D,E; Appendix A5). Observed-to-expected edge ratios on the capped ten-neighbor graph were below one in 14/14 Lin and 3/3 Uzquiano organoids, with component-conditioned medians of 0.7554 and 0.4174. Conditioning additionally on depth showed the deficit in all ten Lin Control-line means (median 0.7375; exact p = 0.00195, q = 0.0117 in the twelve-endpoint family). All three Uzquiano organoids had the same direction (exact sign-test p = 0.25; Appendix A5).

Position and adjacency responded differently to population redefinition on identical section/organoid support (Figure 6F). Each of the four Lin marker settings produced twelve negative and two positive fraction contrasts. Eleven organoids remained negative across all settings, one remained positive and two exchanged signs (Appendix A5). *DCX* omission retained the baseline fraction direction in 14/14 organoids and an adjacency deficit in 13/14; *STMN2* and double omission retained 12/14 positional directions and adjacency deficits in 8/14 (Appendix A5). The same six organoids crossed the adjacency reference of one under *STMN2* and double omission. Fixed-graph accounting showed changes in both observed edge assignments and their label-frequency expectation (Appendix A5). Broad peripheral distribution is therefore more stable under these marker omissions than the finer population interfaces.

Fetal relative-order comparisons placed this recurring architecture in a developmental context [21]. Radial-glia/intermediate-progenitor, radial-glia/immature-neuron and intermediate-progenitor/immature-neuron order corresponded in 9/14, 10/14 and 8/14 Lin organoids, respectively, and in all three month-2 Uzquiano organoids for each comparison. Immature-to-differentiated order corresponded in 1/14 Lin and 0/3 Uzquiano organoids (Appendix A5). Each state pair uses its own supported organoids and the source laminar annotations as the fetal reference.

3.7. Shared tissue organization contains source-specific neuronal molecular relationships

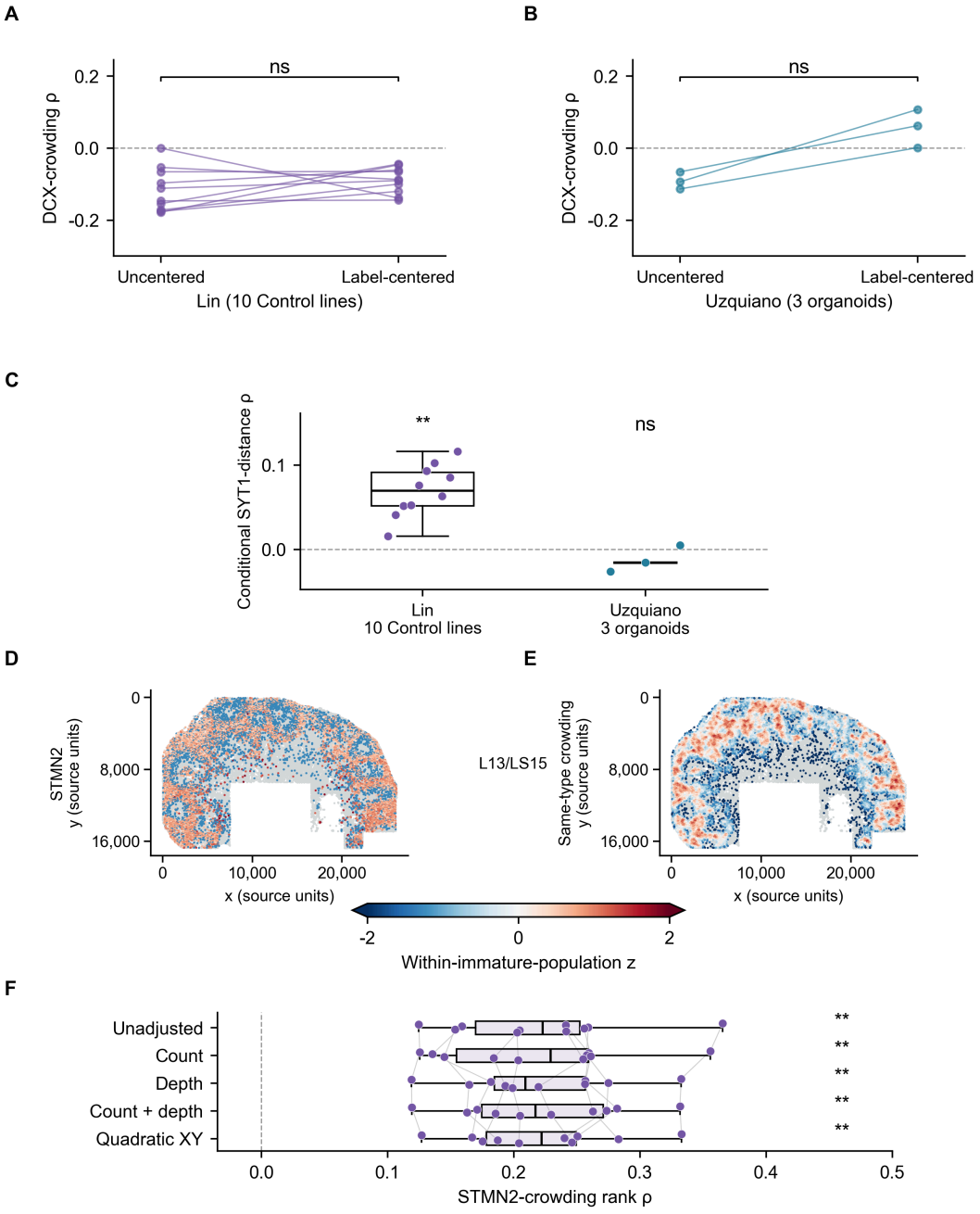


Figure 7. Shared cortical organization contains source-specific molecular-spatial relationships. (A) DCX–crowding correlations before and after label centering in ten Lin Control lines. (B) The same comparison in three Uzquiano organoids. (C) Conditional correlations of neuronal SYT1 with graph distance to progenitors in both cohorts. (D) STMN2 expression in immature neurons in L13/LS15. (E) Same-type crowding in those neurons. (F) STMN2–crowding correlations in immature neurons under five adjustments in Lin

Notes: Panels A and B use k10 crowding calculated across all observations. Thin lines connect the same inference unit. Label centering subtracts each recorded-label expression mean. Lin first averages sections into organoids and then organoids equally within Control lines. Brackets show source-specific exact sign tests of paired centered-minus-uncentered changes against zero:

Lin median +0.0287, 95% percentile CI [-0.0156, +0.0823], q = 0.2188; Uzquiano +0.1282, CI [+0.1146, +0.2008], q = 0.25. The tests form the paired-change BH2 family. Bootstrap procedures are in main Methods, "Inferential scope and supporting comparisons"; the three-unit percentile interval is discrete and calculated separately from the exact sign test. C, Conditional *SYT1* correlations with shortest-path progenitor distance in ten Lin Control-line summaries and three Uzquiano organoids: Lin +0.0694 (95% percentile CI +0.0515 to +0.0940; p = 0.0020, q = 0.0039); Uzquiano -0.0156 (95% percentile CI -0.0262 to +0.0051; p = 1, q = 1). C includes immature and differentiated neuronal observations and conditions on fine identity, depth, inherited counts and spacing. Rank construction, quadratic depth adjustment and finite-fit eligibility are specified in main Methods section 2.3. Lin labeling jointly omits *SYT1* and *MAP2*; Uzquiano uses source labels. The reference line is zero. Symbols denote source-specific exact sign tests against zero in the four-endpoint SYT1/MAP2 family. Panels D and E show the same 12,313 immature profiles in color. The remaining profiles among 38,263 section observations are gray. Each measure is standardized within the immature population, and the color scale saturates beyond ± 2 SD. Both maps use the same native source-coordinate units, equal aspect and downward Y. Same-type crowding is negative log relative tenth-neighbor spacing within that population, excluding self (Equation (1)). In F, thin gray segments connect each Control line across adjustment settings. For each adjustment, an exact sign test assesses the ten Control-line coefficients against zero. All five q values are 0.00355 in the original complete twenty-endpoint family. Medians and 95% percentile intervals are Unadjusted: 0.2229 [0.1588, 0.2557]; Count: 0.2291 [0.1448, 0.2596]; Depth: 0.2092 [0.1788, 0.2657]; Count + depth: 0.2173 [0.1709, 0.2736]; Quadratic XY: 0.2220 [0.1752, 0.2644]; resampling procedures are in main Methods section 2.3. Boxes show medians and interquartile ranges; black whiskers extend to the most extreme observed values within 1.5 IQR. All observations, including outliers, appear once. The three-organoid group in C has individual points and a median bar. Appendix A5 distinguishes the DCX, SYT1 and STMN2 analysis units and testing families. Figure A5 and Table A5 show omission-specific STMN2 baselines and restricted finite support. Symbols denote adjusted q values: ns, $q \geq 0.05$; *, $q < 0.05$; **, $q < 0.01$; ***, $q < 0.001$; ****, $q < 0.0001$.

DCX expression and all-observation crowding shared a predominantly negative uncentered association across the cortical sources but diverged after label centering (Figure 7A,B; Appendix A5). Uncentered coefficients were negative in 13/14 Lin and 3/3 Uzquiano organoids, with medians -0.0885 and -0.0934. After subtracting each recorded label's mean expression, coefficients were negative in all fourteen Lin organoids and positive in all three Uzquiano organoids; medians were -0.0980 and +0.0624. Centering compares expression deviations around label means against the same crowding field. Median paired changes within the biological inference units in Figure 7A,B were +0.0287 across ten Lin Control lines and +0.1282 across three Uzquiano organoids (BH-adjusted q = 0.21875 and 0.25, respectively).

A small positive conditional *SYT1*-distance association recurred across the Lin organoids, whereas the Uzquiano coefficients had mixed signs (Figure 7C). The analysis measured shortest-path distance to progenitor observations on the capped coordinate graph, within immature and differentiated neuronal observations; *SYT1* and *MAP2* were jointly omitted from every Lin labeling program. After fine-label, depth, library-support and spacing adjustment, *SYT1* increased with progenitor distance in all fourteen Lin organoids (median conditional rank correlation +0.0755). All ten Control-line means were positive (median +0.0694; 95% bootstrap interval +0.0515 to +0.0940; q = 0.003906 in the four-endpoint family). Adding progenitor distance to the corresponding depth/support prediction model reduced held-out Control-line prediction error in 10/14 organoid evaluations, with median reduction 0.5821% and four nonpositive evaluations (Appendix A5).

In Uzquiano, conditional *SYT1* coefficients were -0.0262, -0.0156 and +0.0051. Adding progenitor distance increased held-out-organoid prediction error in all three organoids (median error reduction -0.2246%). *MAP2* showed a weaker positive relationship in 13/14 Lin organoids.

Within the Lin immature-neuron population, higher *STMN2* expression was associated with closer relative same-type spacing across all ten Control lines (Figure 7D-F; Appendix A5). Median line coefficients were +0.2229 unadjusted and +0.2291 after count adjustment. Across the same ten lines, the association remained positive after depth, joint count/depth and quadratic-coordinate adjustment, with q approximately 0.00355 for each of the five endpoints in the twenty-endpoint, four-gene/five-adjustment family. The radial-expression endpoint was negative, with q = 0.0527 in its separate radial/spacing family.

The *STMN2* association attenuated with changes to the included population and its neighborhoods. Using the observations assigned to the immature population under each labeling definition, component/gap restriction reduced the baseline median from +0.2229 to +0.1627, with ten positive line means. *STMN2* omission gave uncapped/restricted medians +0.1592/+0.1373, also with ten positive means; double omission gave +0.1002/+0.0443, with eight positive means in each version. Separate paired comparisons used observations assigned to the immature population under both baseline and the particular omission. These showed decreases in every line after *STMN2* or double omission, using a separately calculated baseline for each omission (Appendix A5). Within those shared observations, neighborhood restrictions can leave different subsets with finite correlation inputs in the two estimators. More broadly, 49 of 213 comparable label-by-map/panel records had opposing regional fraction and same-type crowding effects (164 concordant; 34 of 247 evaluated records lacked crowding support; Appendix A5). A population can therefore be more prevalent in a region while its same-type observations are more widely spaced.

4. Discussion

Cross-source comparison reveals a shared nearby glycolysis-hypoxia relationship within organoids that differ in regional expression patterns. The source studies established cortical metabolic gradients and midbrain cellular, mitochondrial and spatial organization [9, 10]. Connecting their point-resolved measurements shows that positive nearby relationships occur in MLOs with jointly inward, jointly outward and opposing program enrichment. Local effects also vary in magnitude and distance profile between sources.

In many MLOs, positive nearby terms coincide with negative distant terms; in cortex, positive associations can extend to matched distant locations. Positive near remainders persist after partitioning out the fitted linear relationship at the same locations. Regional deletion and prediction direction affect some individual remainders and contrasts, whereas MLO median near remainders remain positive across donor and capture omissions. These measurements distinguish the relationship observed nearby, its extension through tissue and its relation to the fitted same-location component.

Within-organoid pairing identifies the joint regional arrangement of the two programs in each organoid. Glycolysis and OXPHOS favor the same region in eighteen of 24 MLOs despite their opposite cohort medians. Paired concordance persists across hierarchy omissions, while the OXPHOS median can reverse with the subset. ATP-related and glutamate/glutamine transcription supply further combinations, including near-zero effects. Positive glycolysis-OXPHOS correlations among similarly labeled spots also occur when the two programs' regional enrichments point in opposite directions.

Population fraction, within-population expression and relative position distinguish components of regional hypoxia-response enrichment. Peripheral CFuPN prevalence coexists with inward weighted expression contributions because fraction and expression enter separate terms of the regional contrast. Matched-tissue comparisons distinguish changes caused by restricting regional coverage from changes caused by subdividing labels on the same observations. Across MLOs, greater immature-labeled fraction is associated with stronger inward hypoxia enrichment, whereas more inward relative position is associated with weaker enrichment after fraction adjustment.

Organoid slicing, artificial diffusion channels and vascular engineering link tissue access to developmental outcomes [3-5], while culture protocol and morphology also relate to developmental and metabolic differences [7, 8, 26]. Mil et al. altered cortical-organoid composition by manipulating glucose availability or pentose phosphate pathway enzymes, with ribose rescuing key phenotypes [27]. The present spatial comparisons show recurring local glycolysis-hypoxia transcriptional relationships across different regional expression

distributions and relate regional hypoxia-response expression to population composition and position in MLOs.

The parallel neuronal comparison separates shared broad positioning from more variable local interfaces. Peripheral early-neuronal representation is more stable under marker omission than progenitor-neuron adjacency. Position summarizes distribution over the organoid, whereas adjacency depends on identities at individual graph edges and their expected frequencies. Reclassification can consequently alter inferred interfaces more strongly than the broad positional tendency. Selected early fetal state orders correspond, while immature-to-differentiated ordering shows little agreement [21].

Shared peripheral positioning coexists with source-specific neuronal molecular relationships. *DCX* centering reveals different within-label associations beneath the common uncentered tendency. Progenitor distance is positively associated with *SYT1* in Lin, whereas Uzquiano shows mixed associations and increased prediction error. Within Lin immature neurons, higher *STMN2* expression accompanies closer same-type spacing, with attenuation under changes to population membership and neighborhood support. Opposing regional fraction and crowding effects further distinguish population prevalence from local spacing.

5. Conclusion

The present comparison establishes that a shared local expression relationship coexists with diverse regional arrangements, and that population fraction, position and within-population expression distinguish components of that diversity. Together, these comparisons support evaluating local expression relationships alongside regional enrichment and population organization when characterizing brain organoids.

Author contributions

Quan Gui: Conceptualization, methodology, software, data curation, formal analysis, visualization, writing original draft, and writing review and editing.

Acknowledgments

The author thanks Dr. Guang-Wei Zhang for his guidance and for taking the time to read and comment on the manuscript.

Data and code availability

Original source data can be obtained through GEO (GSE197150, GSE277032, GSE289229, GSE305978, GSE283497 and GSE223020), BioStudies/ArrayExpress (E-MTAB-10972) and the Broad Single Cell Portal (SCP1756). Source-specific processing and annotation provenance are described in Appendix A1.

Ethics statement

This study reanalyzed publicly available spatial transcriptomic datasets from the cited source studies.

References

[1] Lancaster, M. A., Renner, M., Martin, C.-A., Wenzel, D., Bicknell, L. S., Hurles, M. E., Homfray, T., Penninger, J. M., Jackson, A. P., & Knoblich, J. A. (2013). Cerebral organoids model human brain development and microcephaly. *Nature, 501*(7467), 373-379. https://doi.org/10.1038/nature12517

[2] Lancaster, M. A., Corsini, N. S., Wolfinger, S., Gustafson, E. H., Phillips, A. W., Burkard, T. R., Otani, T., Livesey, F. J., & Knoblich, J. A. (2017). Guided self-organization and cortical plate formation in human brain organoids. *Nature Biotechnology, 35*(7), 659-666. https://doi.org/10.1038/nbt.3906

[3] Qian, X., Su, Y., Adam, C. D., Deutschmann, A. U., Pather, S. R., Goldberg, E. M., Su, K., Li, S., Lu, L., Jacob, F., Nguyen, P. T. T., Huh, S., Hoke, A., Swinford-Jackson, S. E., Wen, Z., Gu, X., Pierce, R. C., Wu, H., Briand, L. A., … Ming, G.-L. (2020). Sliced human cortical organoids for modeling distinct cortical layer formation. *Cell Stem Cell, 26*(5), 766-781.e9. https://doi.org/10.1016/j.stem.2020.02.002

[4] Cai, H., Tian, C., Chen, L., Yang, Y., Sun, A. X., McCracken, K., Tchieu, J., Gu, M., Mackie, K., & Guo, F. (2025). Vascular network-inspired diffusible scaffolds for engineering functional midbrain organoids. *Cell Stem Cell, 32*(5), 824–837.e5. https://doi.org/10.1016/j.stem.2025.02.010

[5] Cakir, B., Xiang, Y., Tanaka, Y., Kural, M. H., Parent, M., Kang, Y.-J., Chapeton, K., Patterson, B., Yuan, Y., He, C.-S., Raredon, M. S. B., Dengelegi, J., Kim, K.-Y., Sun, P., Zhong, M., Lee, S., Patra, P., Hyder, F., Niklason, L. E., … Park, I.-H. (2019). Engineering of human brain organoids with a functional vascular-like system. *Nature Methods, 16*(11), 1169–1175. https://doi.org/10.1038/s41592-019-0586-5

[6] Bhaduri, A., Andrews, M. G., Mancia Leon, W., Jung, D., Shin, D., Allen, D., Jung, D., Schmunk, G., Haeussler, M., Salma, J., Pollen, A. A., Nowakowski, T. J., & Kriegstein, A. R. (2020). Cell stress in cortical organoids impairs molecular subtype specification. *Nature, 578*(7793), 142-148. https://doi.org/10.1038/s41586-020-1962-0

[7] He, Z., Dony, L., Fleck, J. S., Szałata, A., Li, K. X., Slišković, I., Lin, H.-C., Santel, M., Atamian, A., Quadrato, G., Sun, J., Pașca, S. P., Human Cell Atlas Organoid Biological Network, Amin, N. D., Kelley, K. W., Bertucci, T., Temple, S., Bowles, K. R., Caporale, N., … Treutlein, B. (2024). An integrated transcriptomic cell atlas of human neural organoids. *Nature, 635*(8039), 690-698. https://doi.org/10.1038/s41586-024-08172-8

[8] Øhlenschlæger, M. S., Jensen, P., Havelund, J. F., Schmidt, S. I., Mohamed, F. A., Sutcliffe, M., Elmkvist, S. B., Criscuolo, L., Wingett, S. W., Chiaradia, I., Bayram, E., Nicolaisen, J. A. A., Jakobsen, L. A., Brewer, J., Benros, M. E., Freude, K., Færgeman, N. J., Lancaster, M. A., Larsen, M. R., & Bogetofte, H. (2026). Multi-omic analysis of guided and unguided forebrain organoids reveals differences in cellular composition and metabolic profiles. *Cell Reports Methods, 6*(2), Article 101295. https://doi.org/10.1016/j.crmeth.2025.101295

[9] Uzquiano, A., Kedaigle, A. J., Pigoni, M., Paulsen, B., Adiconis, X., Kim, K., Faits, T., Nagaraja, S., Antón-Bolaños, N., Gerhardinger, C., Tucewicz, A., Murray, E., Jin, X., Buenrostro, J., Chen, F., Velasco, S., Regev, A., Levin, J. Z., & Arlotta, P. (2022). Proper acquisition of cell class identity in organoids allows definition of fate specification programs of the human cerebral cortex. *Cell, 185*(20), 3770-3788.e27. https://doi.org/10.1016/j.cell.2022.09.010

[10] Budinger, D., Puigdevall, P., Hall, G. T., Roth, C., Xenakis, T., Marrosu, E., Jerber, J., Di Domenico, A., Picco, F., Kilpinen, H., Castellano, S., Kurian, M. A., & Barral, S. (2026). An in vivo and in vitro spatiotemporal profile of human midbrain development. *Nature Communications, 17*(1), 1354. https://doi.org/10.1038/s41467-025-67779-1

[11] Lin, L., Saw, T. Y., Chou, N., Goh, J. L. J., Kwa, J. E., Chock, W. K., Singhal, V., Li, Z., Huang, M. J., Ng, H. H., Khor, C. C., Kuan, H. L., Chen, K. H., Prabhakar, S., & Liu, J. (2026). Spatial architecture of autism pathogenesis reveals mosaic structural disarray during early development. *Nature Communications, 17*(1), 7522. https://doi.org/10.1038/s41467-026-74320-5

[12] Yang, F., Pavon, N., Matsubara, T., Sebastian, R., Martinez-Martin, B., Pak, C., Sun, Y., & Kim, D.-H. (2026). Distinct spatial patterning and transcriptomic landscapes of human neural organoids by localized delivery of morphogens. *Cell Stem Cell, 33*(3), 438-453.e7. https://doi.org/10.1016/j.stem.2026.01.008

[13] Tsai, Y.-C., Ozaki, H., Almet, A. A., Fleming, I., Nie, C., Luo, S., Wang, X., Shiraiwa, K., Noh, M. J. M., Trejo, S., Sugita, B. K., Dalal, J., Gonzalez, R. A., De Jesus, B., Branco, N., Chen, G. L.-M., Gandal, M. J., Nie, Q., & Watanabe, M. (2026). Morphogen-guided neocortical organoids with anteroposterior areal identity. *Cell Stem Cell, 33*(9), 1540–1552.e9. https://doi.org/10.1016/j.stem.2026.07.014

[14] Amin, N. D., Kelley, K. W., Kaganovsky, K., Onesto, M., Hao, J., Miura, Y., McQueen, J. P., Reis, N., Narazaki, G., Li, T., Kulkarni, S., Pavlov, S., & Pașca, S. P. (2024). Generating human neural diversity with a multiplexed morphogen screen in organoids. *Cell Stem Cell, 31*(12), 1831–1846.e9. https://doi.org/10.1016/j.stem.2024.10.016

[15] Ishahak, M., Han, R. H., Annamalai, D., Woodiwiss, T., McCornack, C., Cleary, R. T., DeSouza, P. A., Qu, X., Dahiya, S., Kim, A. H., & Millman, J. R. (2025). Genetically engineered brain organoids recapitulate spatial and developmental states of glioblastoma progression. *Advanced Science, 12*(10), 2410110. https://doi.org/10.1002/advs.202410110

[16] Salas-Lucia, F., Escamilla, S., Charest, A., Jiang, H., Stout, R., & Bianco, A. C. (2025). Thyroid hormone promotes fetal neurogenesis. *JCI Insight, 10*(19), e194445. https://doi.org/10.1172/jci.insight.194445

[17] He, Z., Maynard, A., Jain, A., Gerber, T., Petri, R., Lin, H.-C., Santel, M., Ly, K., Dupré, J.-S., Sidow, L., Sanchis Calleja, F., Jansen, S. M. J., Riesenberg, S., Camp, J. G., & Treutlein, B. (2022). Lineage recording in human cerebral organoids. *Nature Methods, 19*(1), 90–99. https://doi.org/10.1038/s41592-021-01344-8

[18] Lozachmeur, G., Bramoulle, A., Aubert, A., Stüder, F., Moehlin, J., Madrange, L., Yates, F., Deslys, J.-P., & Mendoza-Parra, M. A. (2023). Three-dimensional molecular cartography of human cerebral organoids revealed by double-barcoded spatial transcriptomics. *Cell Reports Methods, 3*(9), 100573. https://doi.org/10.1016/j.crmeth.2023.100573

[19] Zimmerman, K. D., Espeland, M. A., & Langefeld, C. D. (2021). A practical solution to pseudoreplication bias in single-cell studies. *Nature Communications, 12*(1), 738. https://doi.org/10.1038/s41467-021-21038-1

[20] Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S. F., & Baker, C. I. (2009). Circular analysis in systems neuroscience: The dangers of double dipping. *Nature Neuroscience, 12*(5), 535–540. https://doi.org/10.1038/nn.2303

[21] Qian, X., Coleman, K., Jiang, S., Kriz, A. J., Marciano, J. H., Luo, C., Cai, C., Manam, M. D., Caglayan, E., Lai, A., Exposito-Alonso, D., Otani, A., Ghosh, U., Shao, D. D., Andersen, R. E., Neil, J. E., Johnson, R., LeFevre, A., Hecht, J. L., … Walsh, C. A. (2025). Spatial transcriptomics reveals human cortical layer and area specification. *Nature, 644*(8075), 153–163. https://doi.org/10.1038/s41586-025-09010-1

[22] Semenza, G. L., Roth, P. H., Fang, H. M., & Wang, G. L. (1994). Transcriptional regulation of genes encoding glycolytic enzymes by hypoxia-inducible factor 1. *Journal of Biological Chemistry, 269*(38), 23757–23763. https://doi.org/10.1016/s0021-9258(17)31580-6

[23] Harding, H. P., Zhang, Y., Zeng, H., Novoa, I., Lu, P. D., Calfon, M., Sadri, N., Yun, C., Popko, B., Paules, R., Stojdl, D. F., Bell, J. C., Hettmann, T., Leiden, J. M., & Ron, D. (2003). An integrated stress response regulates amino acid metabolism and resistance to oxidative stress. *Molecular Cell, 11*(3), 619–633. https://doi.org/10.1016/s1097-2765(03)00105-9

[24] Balsa, E., Marco, R., Perales-Clemente, E., Szklarczyk, R., Calvo, E., Landázuri, M. O., & Enríquez, J. A. (2012). NDUFA4 is a subunit of complex IV of the mammalian electron transport chain. *Cell Metabolism, 16*(3), 378–386. https://doi.org/10.1016/j.cmet.2012.07.015

[25] Lee, S.-I. (2001). Developing a bivariate spatial association measure: An integration of Pearson's r and Moran's I. *Journal of Geographical Systems, 3*(4), 369–385. https://doi.org/10.1007/s101090100064

[26] Chiaradia, I., Imaz-Rosshandler, I., Nilges, B. S., Boulanger, J., Pellegrini, L., Das, R., Kashikar, N. D., & Lancaster, M. A. (2023). Tissue morphology influences the temporal program of human brain organoid

development. *Cell Stem Cell, 30*(10), 1351–1367.e10. https://doi.org/10.1016/j.stem.2023.09.003

[27] Mil, J., Soto, J. A., Krall, A. S., Peloggia, J., Frigui, S., Fong, O. S., Sathe, L., Matulionis, N., Day, F., Stiles, L., Montales, K. P., Azizad, D. J., Gonzalez, C. E., Fazzari, E., Li, M. X., Nano, P. R., Martija, A. A., Perez-Ramirez, C. A., Nguyen, C. V., … Bhaduri, A. (2026). Metabolic atlas of early human cortex reveals glycolytic remodeling and pentose phosphate pathway control of cell fate transitions. *Cell, 189*(21), 6810–6828.e15. https://doi.org/10.1016/j.cell.2026.07.023

Appendix

Appendix A1. Cohorts, expression layers and program definitions

The primary metabolic comparison uses three month-2 Uzquiano cortical organoids and 24 Budinger Midbrain-Like Organoids (MLOs). The MLOs comprise two donors, two differentiations within each donor and four captures, with six organoids per capture. The cortical neuronal comparison uses sixteen Lin sections from fourteen organoids and ten Control lines, alongside the same three Uzquiano organoids. Sections are averaged equally within organoid, then organoids equally within Control line. Table 1 identifies the eight contributing studies and their observation units.

He contributes three serial day-60 sections of one iTracer organoid; Yang contributes five patterned model/stage maps. These counts describe sections and maps at their documented source hierarchy. CosMx contributes two Control fields, DBiT fifteen acquisition arrays and Xenium one in-vitro WT map. Disease, tumor, transplant and retinal material is outside the comparison. GSE271118 is excluded because its coordinates do not support the required geometry; GSE294759 is excluded because the released spatial regions cannot be assigned unambiguously to WT/control versus edited material.

Lin expression is log1p(CP10000) from deposited counts using supplied all-feature totals, with project maximum-mean marker assignments. Uzquiano uses supplied expression fractions and source-transferred bead labels. Budinger uses count-derived expression normalized by inherited full-assay totals, with project marker-module assignments for multicellular spots and separate source spatial-region keys. CosMx uses endogenous-gene counts and source Seurat identities; DBiT uses project mixed capture-position identities. iTracer uses full raw-feature totals rather than the named-feature normalization of processed X. Yang project spot-enriched labels and author annotations are separate. Xenium expression uses total_gene_counts; its all-feature total is a distinct count covariate.

MLO coordinates are source Space Ranger full-resolution image pixels. The recorded conversion (Equation (A1)):

$$\frac{65}{spot_diameter_fullres} \tag{A1}$$

micrometers per pixel, is approximately 0.26347-0.26377; near bounds are 199.71-199.94 micrometers and matched distant controls lie beyond 798.86-799.76 micrometers. Uzquiano coordinates are source puck units. Cross-source distance comparisons use the median positive first-neighbor spacing of each map, with no common physical range assumed.

Table A1. Program membership for the primary and complementary metabolic comparisons

Definition	Members
Disjoint glycolysis (10)	ALDOA, ALDOC, ENO2, GAPDH, GPI, HK1, PFKL, PFKM, PFKP, PKM
Disjoint hypoxia response (8)	ADM, ANGPTL4, BNIP3, CA9, DDIT4, EGLN3, NDRG1, VEGFA
Original-common glycolysis (15)	ALDOA, ALDOC, ENO1, ENO2, GAPDH, GPI, HK1, LDHA, PDK1, PFKL, PFKM, PFKP, PGK1, PKM, SLC2A1
Original-common hypoxia response (13)	ADM, ANGPTL4, BNIP3, CA9, DDIT4, EGLN3, ENO1, LDHA, NDRG1, PDK1, PGK1, SLC2A1, VEGFA
Common ATP-related (5)	ATP5F1D, ATP5F1E, ATP5MC3, ATP5ME, ATP5MG

Table A1. Continued

MLO source-available glycolysis (18)	HK1, HK2, GPI, PFKL, PFKM, PFKP, ALDOA, ALDOC, TPI1, GAPDH, PGK1, PGAM1, ENO1, ENO2, PKM, LDHA, SLC2A1, PDK1
MLO source-available hypoxia (14)	VEGFA, CA9, BNIP3, EGLN1, EGLN3, LDHA, PDK1, SLC2A1, ADM, NDRG1, DDIT4, ENO1, PGK1, ANGPTL4
MLO source-available OXPHOS (24)	NDUFA1, NDUFA2, NDUFA3, NDUFA4, NDUFA5, NDUFB5, NDUFB8, NDUFS1, NDUFS2, SDHA, SDHB, UQCRC1, UQCRC2, UQCRFS1, COX4I1, COX5A, COX6A1, COX6C, ATP5F1A, ATP5F1B, ATP5F1C, ATP5MC1, ATP5MC2, ATP5ME
MLO source-available ATP-related (10)	ATP5F1A, ATP5F1B, ATP5F1C, ATP5F1D, ATP5F1E, ATP5MC1, ATP5MC2, ATP5MC3, ATP5ME, ATP5MG
MLO source-available Glu/Gln (14)	ALDH18A1, OAT, PYCR3, GLS, GOT2, GLS2, GLUL, PYCR2, GLUD1, RIMKLB, KYAT1, RIMKLA, GLUD2, PYCR1

Notes: Primary disjoint scores remove all five shared genes from both original-common programs. Source-available MLO energetic scores use their own measured members. Scores average members equally; main Equations (3)-(4) and (6)-(8) specify expression normalization and regional standardization.

The five shared original-common members are ENO1, LDHA, PDK1, PGK1 and SLC2A1. General source-available coverage requires at least five measured members or 30% of intended members; glutamate/glutamine requires five measured members. The Proliferation-associated program requests MKI67, TOP2A, PCNA, CENPF, TYMS, RRM2, MCM2 and MCM5; its C002 comparison uses CENPF, MKI67 and TOP2A. Measurement and detection are distinct. Missing members and constant scores are unavailable, rather than zero effects.

Label centering subtracts the map-specific mean of the recorded label. The centered effect divides the centered inner-minus-outer difference by residual-map population SD. The fixed-denominator variant uses the original score SD on the same finite observations and zones. These are different denominators for the same centered numerator (Equation (A2)):

$$e_i = s_i - \mu_{l(i)}, \sigma_e = \sqrt{\frac{1}{|V|}\sum_{i\in V}[(e_i - \mu_e)^2]} \tag{A2}$$

The label-centered residual e subtracts the map-specific mean for the recorded label of i. Its SD is computed on the relevant finite map support V with population denominator; the residual map mean is retained in the SD calculation.

$$\Delta_{centered} = \frac{\mu_{e,I} - \mu_{e,O}}{\sigma_e}, \Delta_{fixed} = \frac{\mu_{e,I} - \mu_{e,O}}{\sigma_s} \tag{A3}$$

Both effects use the same centered numerator on identical IDs and fixed zones, with different SD denominators. The fixed-denominator version expresses the centered contrast relative to the original score variation.

Appendix A2. Local correspondence, matched support and regional sensitivity

The primary disjoint full-support near statistic and the actual-partner-matched comparison use different eligible focal sets and weights. Matching supports all three month-2 cortical organoids and nineteen MLOs. M03, M04, M06, M08 and M17 have only 27, 1, 20, 11 and 3 matched focals, respectively, below the thirty-focal threshold. M12, M19 and M20 have negative original matched-near values despite positive full-near values; M17 has a negative full-near value but insufficient matched support. The decomposition in Figure 3

uses each operator's own partner weights and its matched focal set; the linear fit uses all finite map observations.

For U01-U03, primary linear near remainders are +0.025827, +0.058100 and +0.007451, and distant remainders +0.000864, +0.002683 and +0.010393. U03 therefore has a positive near remainder but a negative near-minus-distant remainder (-0.002941). The MLO median near remainder is +0.117907 and median paired contrast +0.206223, each positive in sixteen of nineteen organoids. M12, M19 and M20 have negative near remainders.

The fixed natural-cubic model gives near-remainder medians +0.025432 in cortex and +0.119741 in MLOs, positive in 3/3 and 16/19. Paired contrast medians are +0.025097 and +0.201981, positive in 2/3 and 16/19. Reversing predictor and response makes all three cortical contrasts positive. These directional decompositions are descriptive. Regional deletions reconstruct near and distant partner pools while keeping the original fitted mapping and scale (Figure A1). U03 has positive near remainders in 15/16 deletions but positive contrasts in 3/16; M18 also crosses zero. M12 has forty matched focals but no occupied deletion region meeting the ten-observation threshold. MLO median near remainders remain positive after each donor and capture omission.

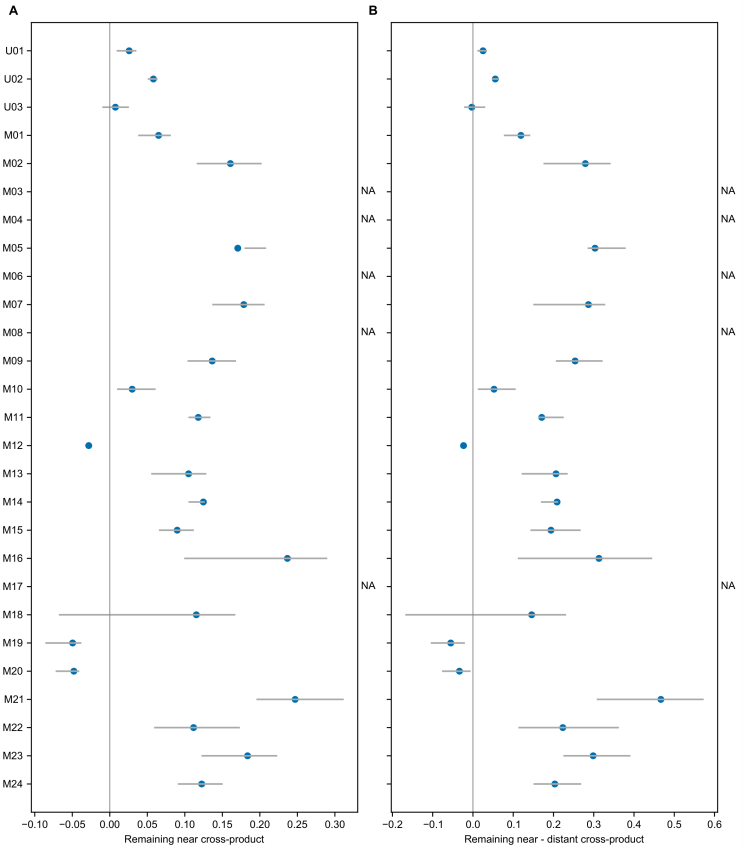


Figure A1. Regional deletions assess sensitivity of nearby correspondence to sampled tissue support. (A) Remaining near correspondence in each organoid under the primary linear model. (B) The corresponding near-minus-distant contrast. In both panels, points show the estimate for the full matched sample and horizontal lines span estimates after individual tissue regions are omitted

Notes: Each successful deletion removes the selected native grid region from the focal, near-partner and distant-control pools. The map-wide fitted mapping, original residual scaling and original three-band focal eligibility remain fixed; matched weights

are reconstructed within each surviving pool. Each deletion still requires 30 retained focals. These ranges quantify sensitivity to the retained spatial support. NA denotes fewer than 30 matched focal observations for the baseline comparison (M03, M04, M06, M08 and M17). M12 contributes a baseline estimate from 40 matched focals; each occupied grid region contains fewer than ten finite observations. All other finite baselines have nonzero deletion ranges. The outcome-independent regions are the existing component-specific 4-by-4 native grid, with at least ten finite observations in a deleted region. The primary analysis and model-form checks are defined in main Methods section 2.7; this appendix describes the organoid and hierarchy sensitivities.

The contextual iTracer comparison averages three serial sections equally within one organoid: disjoint adjusted near correspondence is +0.005455 and its actual-partner-matched contrast +0.006943. All five patterned DV/RC maps have positive adjusted disjoint near effects (+0.110996 to +0.239861) and positive matched contrasts. Their units are one iTracer organoid and five patterned model/stage maps. Glycolysis-ATP and hypoxia-ATP near comparisons are negative in all three primary cortical organoids; MLO positive counts are 13/24 and 8/24, respectively.

Joining strict disjoint regional effects to the same organoid's adjusted full-near value yields fifteen jointly inward, five jointly outward and four opposing MLO configurations. The original-common comparison yields seventeen jointly inward and seven jointly outward. Under inclusive upper-tercile membership, these counts are sixteen/five/three and seventeen/six/one, respectively; M13 and M14 contribute hypoxia sign changes near zero. The score definitions, strict/inclusive boundaries and local estimators remain separate.

Appendix A3. Same-organoid energetic configurations

Table A2. Energetic configurations before and after label centering

Partner	Version	*n*	Both inward	Both outward	Gly in / B out	Gly out / B in	Concordant	Marginal bounds
OXPHOS	Uncentered	24	11	7	6	0	18	4-18
OXPHOS	Centered	24	10	7	7	0	17	3-17
ATP-related	Uncentered	24	10	7	7	0	17	3-17
ATP-related	Centered	24	11	7	6	0	18	4-18
Glu/Gln	Uncentered	24	9	7	8	0	16	2-16
Glu/Gln	Centered	24	10	6	7	1	16	4-18

Notes: All rows contain 24 MLOs with finite, nonzero source-available effects. Gly denotes glycolysis and B the partner; Glu/Gln denotes glutamate/glutamine. Uncentered effects use original-map SD, centered effects residual-map SD (Equations (A2)-(A3)). Concordance counts pairs with the same sign. Marginal bounds enumerate possible concordant counts given the two fixed marginal direction counts.

Paired concordance remains a majority across every donor, donor-specific differentiation and capture omission, although the partner's cohort median can change sign. For uncentered glycolysis/OXPHOS, concordance is 10/12 after omitting Donor1 and 8/12 after omitting Donor2, 12-14/18 across donor-specific differentiation omissions and 13-14/18 across capture omissions. The OXPHOS median changes from -0.2219 in all 24 organoids to +0.2830 after omitting Donor2, +0.0541 after Donor2_dif2 and +0.0359 after capture GSM8513021. Glycolysis remains positive in these subsets. Omissions describe overlapping subsets of the same two-donor cohort.

The within-population glycolysis-OXPHOS comparison has 57 finite correlations across 23 MLOs: 18 dopaminergic-neuroblast, 22 mixed-neural, fourteen stress/necrosis-associated and three unresolved population-organoid records. Positive counts are 15/18, 20/22, 14/14 and 3/3. A population needs twenty

paired spots and variation in both scores. These same-observation correlations use source-available programs and differ from the adjusted disjoint cross-location endpoint.

Appendix A4. Population accounting, conditional associations and prediction

Population contribution terms weight the within-population regional expression difference by mean regional fraction and divide by the original complete-map SD (main Equations (15)-(18)). Each population requires five observations per region. Fine and coarse accounting on matched tissue uses identical retained IDs, original region assignments and SD, with ten observations required per retained region. Both decompositions reconstruct the same restricted contrast.

Restricting the original-common cortical comparison retains 43.1%/94.8%, 99.9%/69.9% and 40.6%/97.7% of inner/outer observations in U01-U03. U01/U03 lose much of their inner support; U02 instead shows an annotation-resolution effect on the same retained observations. In the disjoint comparison, U02 coarse/fine within-group terms are +0.677637/-0.258900 for glycolysis and +0.420253/-0.208472 for hypoxia (original-map SD). Among 21 MLOs with matched regional support, fine-group terms are positive in fifteen for glycolysis and sixteen for hypoxia.

Under the complementary original-common 15/13 definition, CFuPN contributions have means ± sample SD of 0.131765 ± 0.114178 for glycolysis and 0.100248 ± 0.063775 for hypoxia across three cortical organoids. Their mean paired difference is +0.031517; exact two-sided sign-flip $p = q = 0.5$ in its separate one-endpoint family. These values use different memberships from Figure 5B,C. The six disjoint same-organoid comparisons in Table A3 and Figure A2 preserve both signs and the unresolved singleton.

Table A3. Paired disjoint-program and population contributions

Contrast	Organoids	Positive / negative	Median	Mean	Omission median range
P1	17	10 / 7	+0.002171	+0.025248	-0.012859 to +0.012822
P2	22	13 / 9	+0.028691	+0.036359	-0.015614 to +0.094200
P3	11	3 / 8	-0.038809	-0.030107	-0.077742 to +0.035127
P4	1	0 / 1	-0.370843	-0.370843	-0.370843 to -0.370843
P5	11	10 / 1	+0.125304	+0.162394	+0.115520 to +0.280582
P6	11	8 / 3	+0.171315	+0.157251	+0.068128 to +0.288808

Notes: P1-P4 are glycolysis minus hypoxia within dopaminergic-neuroblast, mixed-neural, stress/necrosis-associated and unresolved populations. P5-P6 are mixed-neural minus stress/necrosis-associated contributions for glycolysis and hypoxia. Each program uses its own original-map SD. P5/P6 comprise one Donor1 and ten Donor2 organoids. Ranges are finite subset medians across donor, donor-specific differentiation and capture omissions; P4 includes only subsets retaining M17. These six comparisons are descriptive, selected after earlier population results and specified before evaluating paired differences.

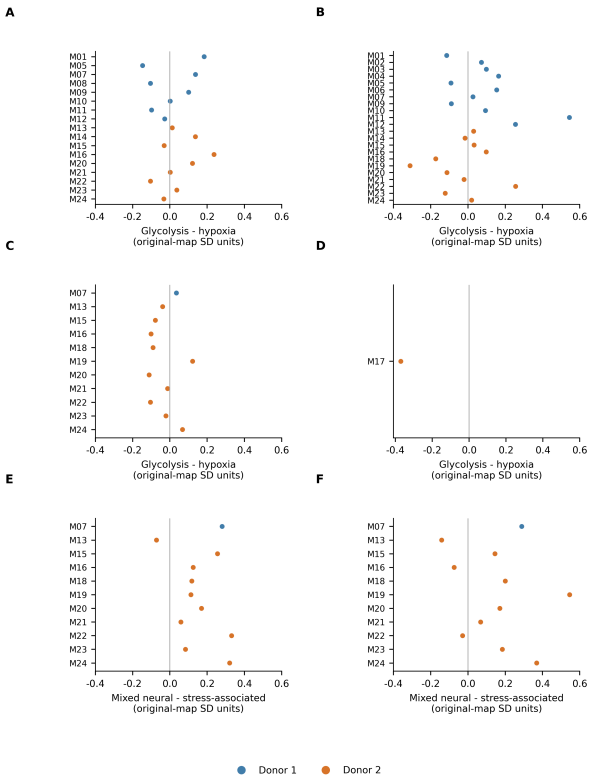


Figure A2. Paired disjoint-program and population contributions within identified MLOs. (A) Glycolysis minus hypoxia-response contribution within the dopaminergic neuroblast population (17 organoids). (B) The same program contrast within the mixed neural population (22 organoids). (C) Program contrast within the stress/necrosis-associated population (11 organoids). (D) Program contrast among unresolved spots in the one eligible organoid, M17. (E) Mixed neural minus stress/necrosis-associated contribution to regional glycolysis in 11 organoids. (F) The corresponding population contrast for the hypoxia-response program in the same 11 organoids

Notes: The comparisons use weighted within-population contributions. Panels E and F include one organoid from Donor1 and ten from Donor2. Each point is a within-organoid difference, labeled by organoid identity and colored by donor. Zero denotes equal contributions; positive values favor the first named program or population. Each contribution uses its program's original-map SD. Eligibility requires at least five population observations in each region and finite values on both sides of the pair. Organoids recur across the six descriptive comparisons. Table A3 reports all six comparisons and their finite omission-median ranges; the singleton remains descriptive. Omission ranges characterize sensitivity across the supported sample subsets.

Appendix A4. Population models and prediction (continued)

Removing BNIP3 and VEGFA from the competing stress/necrosis annotation module changes 263 of 3,335 assignments, increasing immature-labeled spots from 1,079 to 1,142 and unresolved spots from 143 to 163. All seven modules pass the member-coverage rules. Every spot remains in the fraction denominator. The fraction coefficient is +0.153973 original-map SD per ten percentage points under original labels and +0.149566 after overlap removal; both remain positive across all ten hierarchy omissions (Figure A3). The hypoxia outcome and original-map SD remain fixed.

Adding fraction to stage, donor and log spot count reduces mean organoid-held hypoxia MSE from 0.421226 to 0.403807 under original labels (4.14%) and to 0.405086 after overlap removal (3.83%). Capture-

held MSE changes from 0.488039 to 0.484090 under original labels (0.81% reduction) but to 0.496446 after overlap removal (1.72% increase). Thirteen of 24 organoids improve in each label/holdout combination. MSE units are (original-map SD)²; prediction holds out complete organoids or captures.

Relative position S equals mean depth in immature-labeled spots minus mean depth in all spots. Its identity (Equation (A4))

$$S = (1 - f)[mean(d|immature) - mean(d|other)] \tag{A4}$$

makes its dependence on fraction explicit. With stage, donor, log spot count and fraction adjustment, original/overlap-removed hypoxia slopes are -0.305640/-0.334673 original-map SD per 0.1 depth unit. Baseline/extended R^2 values are 0.400388/0.468776 and 0.396166/0.458776. Full-cohort ranks are six/seven, with 18/17 residual degrees of freedom. All individual and hierarchy omissions retain negative slopes. Original-label hierarchy ranges are -0.505179 to -0.119925; overlap-removed ranges -0.624804 to -0.102172.

Table A4. Held-unit prediction after adding relative population position

Outcome	Labels	Holdout	Baseline MSE	Extended MSE	Reduction (%)	Improved organoids
Glycolysis	Original	Capture	0.658692	0.733964	-11.427	7/24
Glycolysis	Original	Organoid	0.558105	0.583368	-4.526	13/24
Hypoxia	Original	Capture	0.484090	0.456089	+5.784	11/24
Hypoxia	Original	Organoid	0.403807	0.378040	+6.381	15/24
Glycolysis	Overlap removed	Capture	0.682445	0.752243	-10.228	8/24
Glycolysis	Overlap removed	Organoid	0.554910	0.591231	-6.545	13/24
Hypoxia	Overlap removed	Capture	0.496446	0.460817	+7.177	12/24
Hypoxia	Overlap removed	Organoid	0.405086	0.389373	+3.879	13/24

Notes: Each row evaluates all 24 MLOs with alpha-one ridge regression and training-only scaling. The baseline includes stage, donor, log spot count and immature fraction; the extension adds S. Errors are equal-organoid MSE in (original-map SD)². Reduction is $100 \times \frac{(baselineminusextendedmeanMSE)}{baselinemeanMSE}$; negative values denote increased error. Improved counts describe individual organoids. Original and overlap-removed labels have the definitions above.

Hypoxia prediction improves in three captures and worsens in GSM8513022 under both annotation definitions. M04 has six immature-labeled spots among 41 total spots. Excluding M04 from the overlap-removed organoid-held error summary gives a 2.57% increase in error. The glycolysis secondary comparison increases mean error under both annotation versions and holdout schemes (Table A4). Figure A4 shows every unit's hypoxia prediction change.

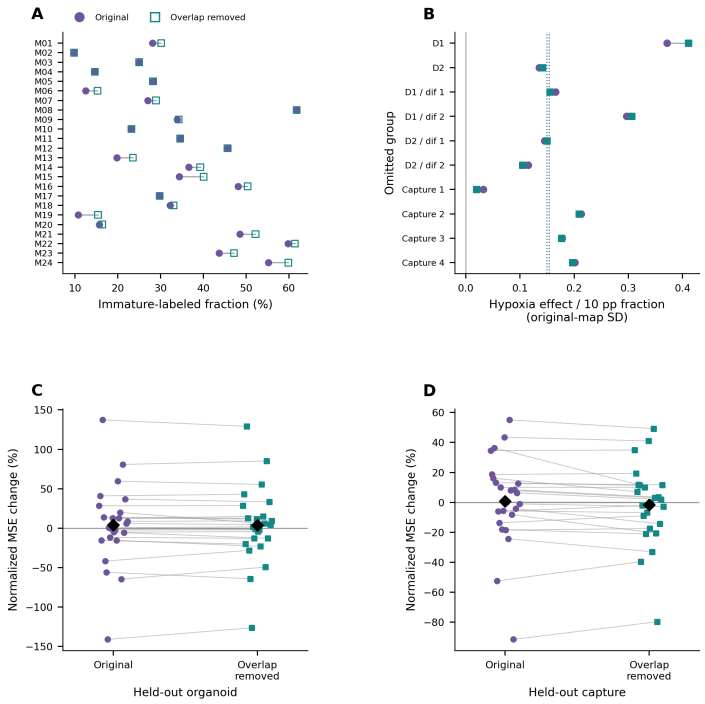


Figure A3. Composition-hypoxia association persists after removing metabolic genes from annotation, while capture-held prediction worsens. (A) Immature-labeled spot fractions in each of 24 MLOs before and after *BNIP3* and *VEGFA* were removed from the stress/necrosis annotation module. (B) Adjusted hypoxia-response coefficients per ten-percentage-point increase in immature fraction after each donor, differentiation or capture omission. (C) Change in each organoid's hypoxia prediction error when immature fraction is added to the model, with whole organoids held out. (D) The same error comparison with whole captures held out

Notes: In A, each line joins the same organoid across annotation definitions, including unchanged values. All 3,335 original spots remain in the denominators. Panel B includes all ten donor, donor-specific differentiation or capture omissions. Purple circles use original labels; teal squares remove BNIP3 and VEGFA from the competing stress module. D1/D2 denote Donor1/Donor2; dif identifies the differentiation within that donor. Captures 1-4 are GSM8513021-GSM8513024. Colored dotted lines show full-cohort coefficients. Connecting lines pair the same omitted group across definitions; omission estimates quantify sensitivity to group support. Models adjust categorical stage, donor and log total sampled-spot count. Panels C and D report normalized MSE changes under both annotation definitions. Each point is $100x\frac{(organoidbaselineMSE-organoidexpandedMSE)}{meanbaselineMSEforthatholdoutscheme}$, using the mean baseline-error denominator for that holdout. Black diamonds are the equal-organoid means. Lines connect the same evaluation organoid across annotation definitions; the ordering and validation folds are identical. Positive values favor adding immature fraction. The annotation modules require at least three measured retained markers and half their original requested membership. All seven modules qualify; maximum-mean assignment resolves ties in source order and assigns Unresolved for nonpositive maxima. The original disjoint hypoxia outcome and original-map SD define both comparisons. Appendix A4 reports full-cohort changes, absolute prediction errors and the corresponding sampling hierarchy.

Metabolic predictors do not improve the mean held-out SYT1 outcomes examined here. Adding full-near correspondence to stage, donor, log spot count and fraction increases mean immature-population SYT1 MSE from 0.025641 to 0.027355 across nineteen MLOs; eight improve. In the separate observation-level model, neighboring glycolysis/hypoxia scores increase equal-organoid MSE from 0.530877 to 0.534133 in nineteen MLOs and from 0.020128 to 0.020156 in three cortical organoids; seven and one improve, respectively. MLO capture-held error increases from 0.531872 to 0.537812. These errors use each source's SYT1 expression

scale. The progenitor-immature adjacency outcome has zero eligible MLOs under its support criteria. The exploratory fraction/full-near association has opposite within-donor directions in two sampled donor preparations, distinct from the regional hypoxia outcome.

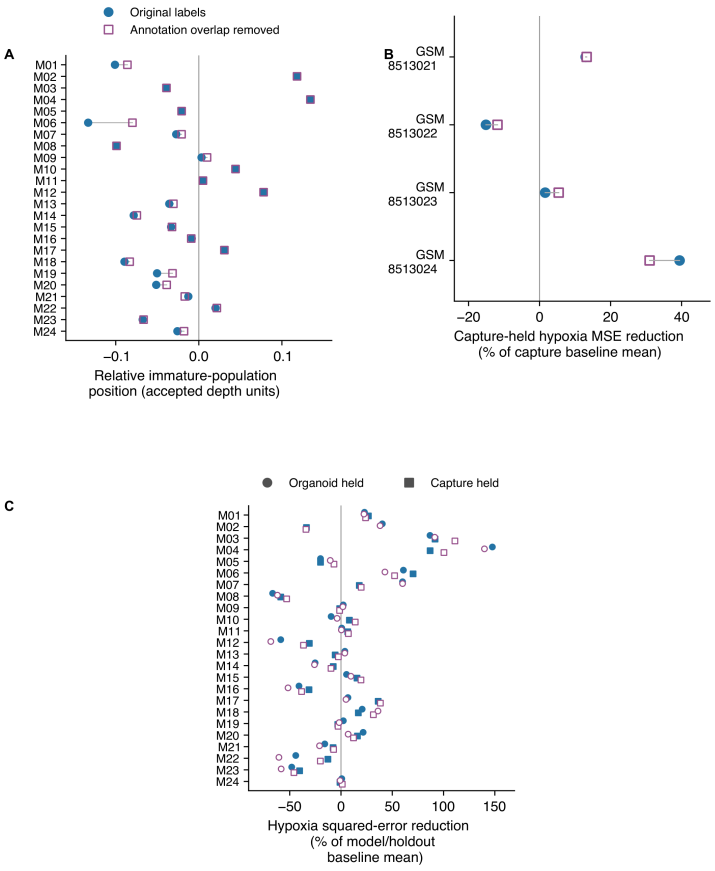


Figure A4. Annotation and held-unit sensitivity of relative-position prediction. (A) Mean depth of immature-labeled spots relative to mean depth of all spots in each MLO, under the original and overlap-removed annotations. (B) Percentage reduction in hypoxia prediction error within each held-out capture after adding relative position. (C) Prediction-error reduction for each organoid under both annotation definitions and both holdout schemes

Notes: In A, lines join original (blue filled circles) and overlap-removed (purple open squares) estimates for each of the 24 MLOs. Positive relative position (S) places the immature-labeled population inward of the sampled-map mean. All spots enter fraction denominators. Figure 5E shows the original-label partial association. Each capture in B has six organoids. Gray lines join the annotation versions, using that capture's own baseline mean squared error as denominator. In C, circles/squares denote organoid/capture holdout; blue filled/purple open markers denote original/overlap-removed annotations. Each value is 100 times baseline minus extended squared error divided by the model/version/holdout mean baseline squared error over all 24 organoids. Positive values indicate lower error with position; vertical offsets separate the four evaluations per organoid. The baseline contains stage, donor, log spot count and immature fraction; the extension adds S. Alpha-one ridge models use training-only scaling and complete-unit holdout. Whole-cohort and capture-specific percentages have the distinct denominators stated above. Table A4 includes the hypoxia and glycolysis prediction summaries; Appendix A4 reports hierarchy and M04 sensitivity.

Appendix A5. Neuronal annotation, neighborhood and molecular sensitivities

Lin progenitors combine radial_glia_progenitor, intermediate_progenitor and cycling_progenitor; immature_neuron is the immature class; differentiated_neuron and gabaergic_interneuron form the differentiated class. Uzquiano progenitors are aRG, oRG and IP; the immature class combines Newborn CFuPN and Newborn CPN, while the differentiated class includes CFuPN and Unspecified PN. Other named, mixed, unresolved and marker-not-detected assignments remain in regional fraction denominators. Lin marker programs and source-transferred Uzquiano classes have different provenance.

All sixteen Lin sections and fourteen organoids support the positional and component-conditioned adjacency comparisons in every marker setting. Eleven organoids have negative fraction effects throughout, one positive throughout, and two exchange direction. CW20051C-1 changes from -0.0757 to +0.0165/+0.0567 after STMN2/double omission; CW60220F-2 changes from +0.0419 to -0.0434/-0.0581. The six adjacency-reference crossings after STMN2 and double omission occur in CW20047D-1, CW20047D-2, CW20069F-1, CW20320C-1, CW60040F-2 and CW60109C-1; CW20320C-1 also crosses after DCX omission. Identical section support is used across definitions. These descriptive transitions are distinct from the 114-slot conditional-null family and twelve hierarchy-level tests.

Whole-map DCX is negative before centering in 13/14 Lin and 3/3 Uzquiano organoids; after centering, all fourteen Lin coefficients are negative and all three Uzquiano coefficients positive. CW60109C-1 changes from +0.0314 to -0.1672. Paired changes are tested at ten Lin Control lines and three Uzquiano organoids in BH2. The separate SYT1/MAP2 progenitor-distance family uses BH4 and the neuronal observation sets in main Methods section 2.3. All fourteen Lin SYT1 coefficients are positive after jointly omitting SYT1/MAP2 from labeling; Uzquiano values are -0.0262, -0.0156 and +0.0051. The Lin prediction improvement is small and occurs in 10/14 organoid evaluations under Control-line holdout; all three cortical held-organoid errors increase.

Table A5. STMN2 correlations on omission-specific shared membership

Omission	Neighborhood	Baseline median	Omission median	Median paired change	Increase	Decrease	Zero
					n	*n*	*n*
DCX	Restricted	0.1595	0.1280	-0.0105	3	7	0
DCX	Uncapped	0.2213	0.1467	-0.0422	1	9	0
Both	Restricted	0.1768	0.0248	-0.1160	0	10	0
Both	Uncapped	0.2450	0.0692	-0.1765	0	10	0
STMN2	Restricted	0.1789	0.0943	-0.0546	0	10	0
STMN2	Uncapped	0.2500	0.1382	-0.1030	0	10	0

Notes: Each row summarizes ten Control-line pairs from fourteen organoids and sixteen sections. Each omission has its own shared-membership baseline. Paired change is calculated within line before summarization and differs from a difference of marginal medians. All 48 uncapped section pairs have identical finite IDs; all 48 restricted pairs have different finite supports despite common annotation membership. Increase/decrease/zero columns count Control lines. Figure A5 distinguishes own-membership and shared-membership comparisons.

Original correlation masks require finite expression, crowding, counts and depth, including unadjusted rows. Restricted finite focal support ranges from 33.9% to 94.7% of original section profiles. The within-immature crowding family has twenty exact sign tests for four genes and five adjustments; the separate radial-expression/spacing family has 33 registered endpoints and 27 finite tests. Median-line intervals use 5,000 percentile draws (seed 9222026), and the radial intervals use seed 412. The original immature-fraction signed-

rank p = 0.001709 and q = 0.027344 belong to their own source family, distinct from the baseline-fraction BH2 used in Figure 6.

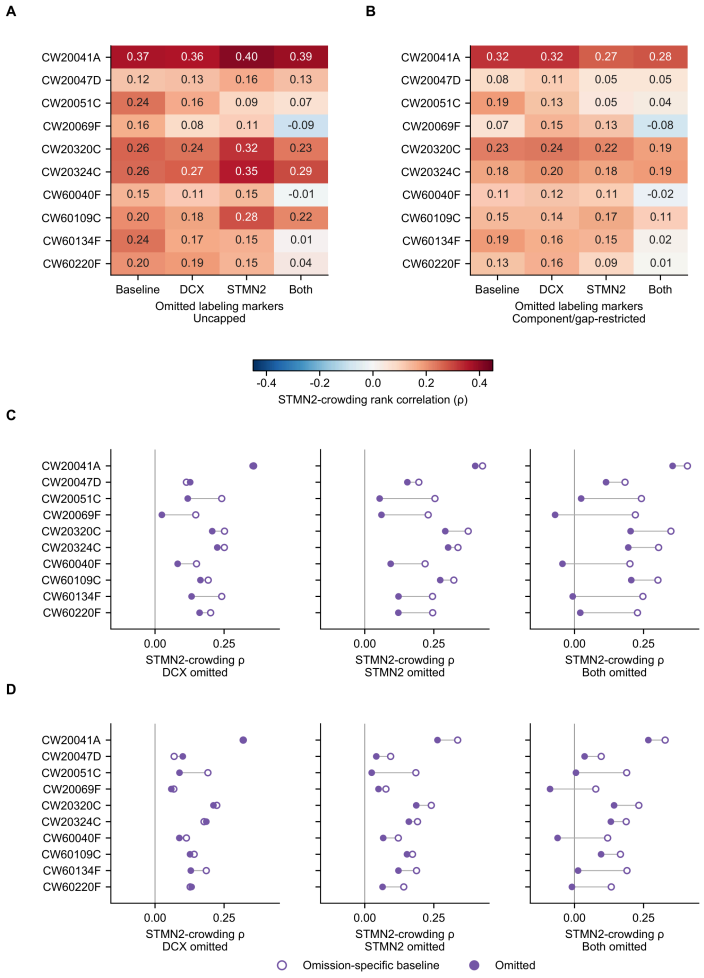


Figure A5. STMN2–crowding correlations depend on labeling and neighborhood definition. (A) *STMN2*–crowding correlations for each labeling setting, using its own membership and uncapped neighborhoods. (B) The same correlations with component/gap restrictions. (C) Baseline and marker-omitted correlations on shared membership using uncapped neighborhoods. (D) The same paired comparison with component/gap restrictions

Notes: All panels show unadjusted rank correlations with tenth-neighbor (k10) crowding in ten Control lines. The heatmaps cover four labeling settings. In C and D, lines pair baseline (open circles) and omitted (filled circles) coefficients within each Control line for *DCX*, *STMN2* or joint omission. Each omission defines its own shared population and baseline. Sixteen section pairs are averaged equally into fourteen organoids and then ten lines. All 48 uncapped section pairs have identical finite IDs, whereas all 48 restricted pairs have unequal finite support despite shared annotation membership. The current-membership baseline medians are 0.2229/0.1627, STMN2-omission medians 0.1592/0.1373 and both-marker-omission medians 0.1002/0.0443; both-marker omissions have 8/10 positive lines. Table A5 distinguishes median within-line changes from differences of medians. Main Methods section 2.3 and Appendix A5 define neighborhood restrictions, finite support and aggregation.

Across the broader label-by-map/panel inventory, 247 records were evaluated: 213 have finite regional fraction and same-type crowding effects, of which 164 are concordant and 49 opposed, with no zero ties; 34

lack crowding support. Records repeat labels within maps and source units. Twelve CosMx field-axis records are separate: seven same-direction, one tie and four unavailable. Relative crowding measures neighbor spacing and is not a physical-area cell density.

The fetal reference uses supplied cortical-state annotations [21]. Radial-glia/intermediate-progenitor, radial-glia/immature-neuron and intermediate-progenitor/immature-neuron order agrees in 9/14, 10/14 and 8/14 Lin organoids, and in all three month-2 Uzquiano organoids for each pair. Immature-to-differentiated order agrees in 1/14 and 0/3. Every pair uses its own supported organoids; these positional comparisons are separate from expression-contrast comparisons. For the latter, fetal H1 RG/IPC and EN-Mig/EN-Newborn are the precursor and early-neuronal classes, DP is excluded and GW34 lacks the required early classes.