A193 · ACCORDION: aligned condition-specific gene representations for multi-sample single-cell analysis

Presenter: Renjie Wu — Massachusetts General Hospital

Authors: Renjie Wu, Rong Ma, Lindsay F. Rizzardi, Luca Pinello

Session: Day 1 · Thu Oct 1, 2026 · 2:15–4:15 PM

Single-cell condition comparisons are often organized around cell-centric representations, with gene-level effects assessed downstream. We present ACCORDION, a condition-aware cell–gene co-embedding framework that learns shared cellular state representations together with separate yet aligned representations of the same named genes across conditions. Cells and genes are represented variationally and jointly trained from expression counts. Same-gene alignment promotes cross-condition comparability, while a structured count decoder separates cell–gene interactions from auxiliary count-scale and nuisance-associated effects. We apply ACCORDION to a multi-donor Alzheimer's disease single-nucleus RNA-seq dataset using all 36,503 genes. The learned cell representation preserves major cell-state structure while showing substantial mixing across donor labels. Moderate gene alignment establishes nearly complete cross-condition comparability of named genes while retaining condition-dependent gene variation. Several genes highlighted in the original study show appreciable representation shifts. Established AD-associated genes show concordant representation changes, including BIN1 and APOE (high displacement and high cell-type-specific log fold change), whereas less detectable neuronal regulator ZEB1 also shows appreciable displacement despite a modest excitatory-neuron log fold change. These results support condition-specific gene representations as objects for studying gene change alongside continuous cellular heterogeneity, with downstream applications to cell–gene, gene–gene, and potentially cross-feature neighborhood queries without requiring predefined cell clusters.

ACCORDION: aligned condition-specific gene representations for multi-sample single-cell analysis

Renjie Wu*[1], Rong Ma[2], Lindsay F. Rizzardi[3,4], Luca Pinello[1,2]

[1] Department of Pathology, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA

[2] Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA

[3] Department of Biochemistry and Molecular Genetics, University of Alabama at Birmingham, Birmingham, AL, USA

[4] HudsonAlpha Institute for Biotechnology, Huntsville, AL, USA

*Presenting author

Background and motivation. Single-cell condition comparisons are often organized around cell-centric representations, with gene-level effects assessed downstream. ACCORDION asks a complementary question: can condition-dependent gene behavior itself be retained as a reusable learned object? The framework extends cell–feature co-embedding by learning shared cellular state representations and separate yet aligned gene representations across conditions. This representation framework provides a direct way to identify genes that vary across conditions and to examine where those changes occur along continuous cell-state variation without requiring predefined cell clusters.

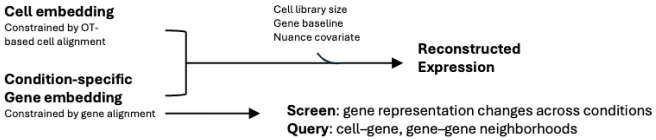
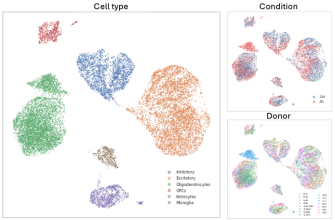


Figure 1. ACCORDION exposes condition-specific named-gene representations alongside cells. Alignment and count-model structure are intended to improve comparability; these roles are evaluated empirically rather than assumed.

Methods. Each cell or gene node is represented variationally by a posterior mean and dispersion. We use regularization including gene alignment that encourages embeddings of the same nominal gene to remain comparable across conditions, and optimal-transport-based cell alignment that promotes comparable cell states. A structured decoder models expression counts from cell–gene interactions while separately accounting for count scale and nuisance-associated variation, helping keep the relational cell–gene geometry interpretable.



Question	Metrics	Value
Cell state preserved?	30-NN cell-type purity	0.996
Cell state embedding comparable?	Condition entropy within cell type	0.741
	Donor entropy within cell type x condition	0.621
Gene embedding comparable?	Same-gene retrieval MRR	0.998
Gene displacement redundant with DE?	Displacement vs max |Welch t| correlation	0.122
Model components carry expected signal?	Cell parameter vs library size correlation	0.977
	Embedding interaction ablation ΔNLL	0.350

Figure 2. Representation diagnostics on the full-gene Alzheimer's disease dataset. *ACCORDION preserves cell-state structure across conditions and donors, establishes comparable condition-specific gene embeddings, and shows distinct but biologically related differential gene representation.*

Analysis. In a multi-donor Alzheimer's disease single-nucleus RNA-seq dataset, ACCORDION preserves coherent cell-state structure with substantial mixing across donor labels. On the gene embedding side, moderate cross-condition gene alignment ensures comparability while retaining condition-dependent gene embedding variation. Established AD-associated genes show concordant representation changes, including BIN1 (displacement 0.723; microglial log2FC = −0.856) and APOE (0.702; microglial log2FC = 0.992), whereas less detectable neuronal regulator ZEB1 also shows appreciable displacement (0.633) despite a modest excitatory-neuron log2FC (0.181). Together, these results support ACCORDION as a gene-aware representation framework for studying condition-dependent changes.