nature communications

Article

https://doi.org/10.1038/s41467-026-70989-w

MIC-Drop-seq: scalable single-cell phenotyping of mutant vertebrate embryos

Received: 9 June 2025

Accepted: 10 March 2026

Published online: 01 April 2026

Check for updates

Clayton M. Carey[1,5], Saba Parvez[2,3,5], Zachary J. Brandt[2], Brent W. Bisgrove[1,2], Christopher J. Yates[2], Randall T. Peterson[2] ✉ & James A. Gagnon[1,4] ✉

Pooled perturbation screens can reveal cellular regulatory networks, yet scaling these techniques for large-scale screens in animals remains challenging. Here we present MIC-Drop-seq, a technique that addresses these challenges by combining high-throughput CRISPR gene disruption in zebrafish embryos with phenotyping by multiplexed single-cell RNAseq. In one MIC-Drop-seq experiment, we simultaneously identified changes in gene expression and cell abundance across 74 cell types resulting from loss of function of 50 transcription factors. These observations recapitulate many known phenotypes, while also uncovering previously uncharacterized roles for transcription factors in brain and mesoderm development. A key advantage of whole-animal screens is that they reveal how changes in one cell type affect the development of other cell types. Surprisingly, such cell-extrinsic phenotypes are abundant, indicating that transcription factors frequently exert effects beyond the cells where they are expressed to adjacent cells. We propose that MIC-Drop-seq will facilitate efforts to dissect the complete gene regulatory networks that guide animal development.

Much of our understanding of biology derives from studying how genetic perturbations manifest in organismal phenotypes[1]. Forward mutagenic screens, for instance, provided a foundational knowledge of genetic networks underlying animal development[2–4]. The recent advent of reverse genetic strategies employing CRISPR has opened new possibilities for targeted screens that enable rapid linking of genotypes to phenotypes[5]. Such targeted approaches, however, have been limited in scope by cumbersome manual delivery of perturbations to individual animals. We recently developed Multiplexed Intermixed CRISPR Droplets (MIC-Drop) as a technique for high-throughput generation of libraries of barcoded mutant zebrafish[6]. In a MIC-Drop screen, a library of intermixed droplets is created and injected dropwise into zebrafish embryos at the single-cell stage to generate a pool of whole-animal mutants. Each droplet contains unique Cas9/guide RNA (gRNA) complexes targeting a single gene alongside a unique DNA barcode to facilitate rapid genotyping. We previously employed MIC-Drop to rapidly identify regulators of cardiac development, demonstrating its power to expand the scale of reverse genetic screens in zebrafish[7].

Genetic screens of laboratory animals are often limited to easily observable phenotypes best suited for detection by simple visual assays. However, many developmental phenotypes are subtle or lack obvious morphological changes, rendering them invisible to conventional screening methods. Furthermore, many experimental measures of phenotype are qualitative or bespoke to specific systems, making them difficult to compare across experiments and laboratories. The advent of high-throughput transcriptional profiling, including single-cell RNA sequencing (scRNAseq), has provided a means to systematically characterize and quantify mutant phenotypes at the cellular and molecular levels. Single-cell methods have been used to map time-resolved cell atlases of zebrafish development, providing a baseline to compare perturbed states in a developing vertebrate[8–10]. Recent

[1]School of Biological Sciences, University of Utah, Salt Lake City, UT, USA. [2]Department of Pharmacology and Toxicology, College of Pharmacy, University of Utah, Salt Lake City, UT, USA. [3]Department of Cell & Developmental Biology, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA. [4]Henry Eyring Center for Cell & Genome Science, University of Utah, Salt Lake City, UT, USA. [5]These authors contributed equally: Clayton M. Carey, Saba Parvez. ✉e-mail: randall.peterson@pharm.utah.edu; james.gagnon@utah.edu

methodological advancements in combining embryo barcoding and combinatorial cell indexing have enabled the generation of genetic perturbation atlases of early development for dozens of mutants in parallel. These datasets have enabled the simultaneous comparison of dozens of mutants and revealed previously uncharacterized developmental phenotypes[11,12]. By pooling samples, these methods circumvent the conventional requirement for separate housing and sequencing of animals with different genotypes. However, challenges persist in scaling up the number of profiled mutants, as the labor and complexity involved in embryo-level indexing impose practical limitations on screen size.

Approaches that couple multiplexed perturbation delivery with pooled scRNAseq offer a scalable platform for single-cell CRISPR screens. Methods such as Perturb-seq and CROP-seq use sequencing of gRNA-linked molecular barcodes, allowing genotypes to be assigned at the cellular level across hundreds of perturbations within a single experiment[13–15]. These tools have allowed the identification of unique cell states and changes in gene regulatory networks caused by perturbing genes in cultured cells[16–18]. Perturb-seq has been extended to organ-targeted in vivo systems, allowing the characterization of many knockout phenotypes within clonal populations in the brains of individual mice using viral vectors[19]. Existing in vivo Perturb-seq approaches, however, fail to capture early developmental phenotypes that would manifest in full knockout animals, and can be confounded by interactions between clones with different genotypes in the same mosaic tissue. A whole-animal Perturb-seq approach, where each individual has a single-gene knockout, is therefore needed to fully characterize the developmental function of genes in vivo.

Here we present MIC-Drop-seq, a technique that couples rapid, multiplexed CRISPR perturbation with scRNAseq from pooled whole-animals. We demonstrate that MIC-Drop-seq enables effective genotype inference by gRNA capture in mutant cells and recapitulates known mutant phenotypes at the cellular level. Using MIC-Drop-seq for a screen of 50 transcription factors, we discover and validate roles for several genes in brain and mesoderm development. We demonstrate that disruptions of transcription factors often cause pervasive yet gene-specific cell-extrinsic effects, indicating that transcription factors frequently alter cellular behaviors beyond the cells in which they are expressed and highlighting the unique advantages that come from performing Perturb-seq experiments in animals. Overall, our findings showcase the utility of MIC-Drop-seq for high-throughput screens of animal development, enabling the mapping of genotypes to phenotypes across dozens of genes in parallel in a single experiment.

Results

MIC-Drop-seq enables multiplexed embryo phenotyping

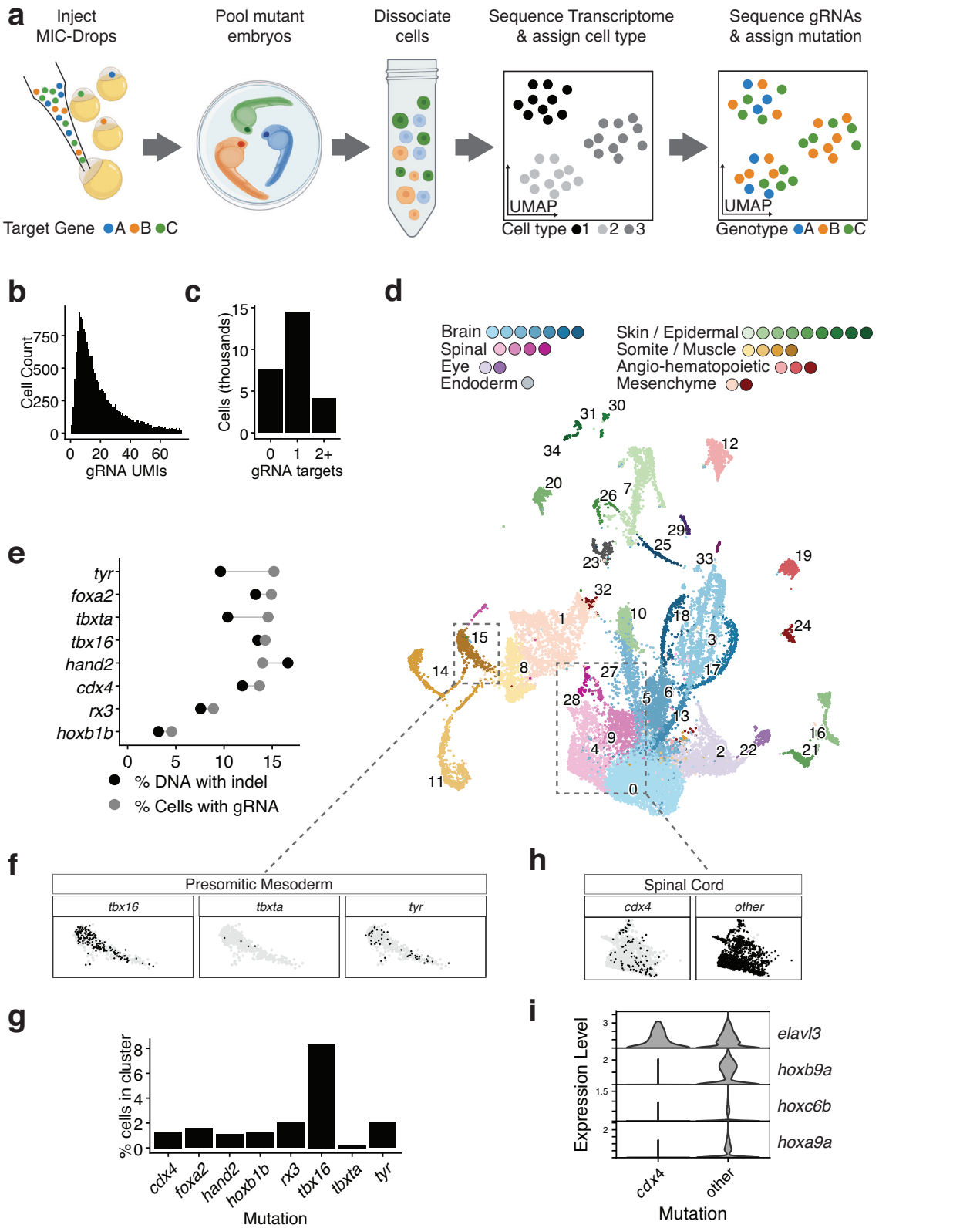
MIC-Drop-seq uses a simple workflow to create and pool animals with different mutant genotypes for multiplexed scRNAseq (Fig. 1a). As in MIC-Drop[6,7], sets of droplets containing Cas9/gRNA complexes targeting different genes are intermixed to form a library. Cohorts of zebrafish embryos are injected at the single-cell stage with a single droplet from the library, so that each embryo has a single gene mutated. Pooled mutant embryos are dissociated into a single-cell suspension of mixed genotypes and prepared for transcriptional profiling. To facilitate demultiplexing of mutant genotypes, MIC-Drop-seq uses direct capture and sequencing of gRNA. We used gRNAs with a modified scaffold that incorporates a capture sequence compatible with 10X Genomics feature barcoding[13] (Supplementary Fig. 1a). We validated that these gRNAs persist for days after injection (Supplementary Fig. 1b) and were highly active for CRISPR mutagenesis in zebrafish embryos (Supplementary Fig. 1c). This approach enables the unified capture of both mRNA and gRNA from each cell in pools derived from zebrafish embryos with mixed genotypes, enabling multiplexed generation and single-cell transcriptomic profiling of many whole-animal mutants in a single experiment.

As a proof-of-concept, we performed a small MIC-Drop-seq screen targeting seven transcription factors with well-characterized mutant phenotypes (*tbx16, tbxta, cdx4, foxa2, hoxb1b, rx3*, and *hand2*), as well as a non-transcription factor control gene (*tyr*). A set of droplets was created for each gene target, with each droplet containing 4 unique gRNAs targeting the gene of interest and Cas9 protein, followed by intermixing to form a MIC-Drop library (Supplementary Data 1). These droplets were used to inject a cohort of embryos at the single-cell stage. A set of 40 individuals were then dissociated in a single pool at 24 hours post-fertilization (hpf) and processed together in a single super-loaded 10X channel. After sequencing the barcoded gRNA library, read thresholds were established to filter out ambient gRNA reads (see Methods). With these criteria, gRNAs were confidently detected in 71% of cells, with a median of 15 gRNA molecules recovered from each classified cell (Fig. 1b). Sequencing of gRNA targeting regions revealed that most cells had gRNAs targeting a single gene, while a smaller subset had gRNAs that targeted multiple genes (Fig. 1c). We reasoned that these cells represent doublets created by super-loading cells during microfluidic library preparation. Indeed, the proportion of cells with guides targeting multiple genes matched the expected doublet rate from super-loading. Computational doublet classification demonstrated that cells with gRNAs targeting multiple genes had characteristics of multiplet artifacts (Supplementary Fig. 2a) and were more likely to be assigned as doublets (Supplementary Fig. 2b)[20]. We therefore considered these cells to be technical doublets and removed them from the dataset. The mutant genotype of the remaining 21,994 cells was inferred based on the sequence of the captured gRNAs, recovering a range of 661 to 2,198 cells per targeted gene (Supplementary Fig. 3a, b).

We determined the efficiency of CRISPR mutagenesis in our MIC-Drop-seq experiment using pooled amplicon sequencing of genomic DNA from the remaining dissociated cells at each of the predicted cut sites. As expected, CRISPR mutagenesis induced frequent deletions in regions adjacent to gRNA cut sites (Supplementary Fig. 3c). For each gene, we calculated the frequency of indels at the target cut sites and in experimental and wild-type cells. We then compared indel frequency at the site with the highest indel rate to the frequency of predicted mutant genotypes among cells in the dataset. The proportion of edited DNA and predicted mutant genotypes were highly correlated, indicating a high efficiency of mutagenesis (Fig. 1e). Furthermore, a conservative estimate of cumulative frameshift probability was high for each target (75-99%), calculated across the amplified cut sites (Supplementary Fig. 3d). Thus, the MIC-Drop-seq protocol induced high rates of CRISPR mutagenesis, similar to previous MIC-Drop experiments[7].

We evaluated whether MIC-Drop-seq could recapitulate known developmental phenotypes. Two of the transcription factors we targeted have well-characterized roles in somitogenesis, each causing distinct changes in the abundance of specific mesodermal cell types. Embryos harboring mutations in *tbx16* have patterning defects in trunk somites and an accumulation of uncommitted presomitic mesoderm[21]. Conversely, mutations in *tbxta* result in the failure to specify presomitic mesoderm and do not make posterior somites[2]. We performed dimensionality reduction and clustering on the dataset, yielding 35 cell clusters (Fig. 1d), which were annotated by label transfer from the Daniocell reference atlas[8] (Supplementary Fig. 4a–c). Cluster annotations were manually verified by examining marker gene expression and represent all tissue types known to be present in the 24 hpf embryo (Supplementary Data 2). We computed the proportion of cells in the presomitic mesoderm cluster for each genotype. Consistent with the known mutant phenotypes, we observed a ~4-fold enrichment of *tbx16* mutant cells and a ~10-fold depletion of *tbxta* mutant cells in the presomitic mesoderm relative to the control *tyr* mutant cells (Fig. 1f,g). Additionally, *rx3* mutant embryos have defects in the specification of retinal neurons from early forebrain progenitors, resulting in an

https://doi.org/10.1038/s41467-026-70989-w



eyeless phenotype[22]. Consistent with this phenotype, we observed a depletion of *rx3* mutant cells among all retinal cell types compared to *tyr* mutant cells (Supplementary Fig. 5a, b). Therefore, MIC-Drop-seq is capable of detecting phenotypes that enrich or deplete particular cell types.

Some mutant phenotypes manifest as a result of altered gene expression without changing the abundance of any cell type. To determine whether MIC-Drop-seq can also detect these phenotypes, we tested for differential gene expression between mutant genotypes in each cell type. *cdx4* has a well characterized role in regulating hox family gene expression domains in the spinal cord[23,24]. We identified *cdx4* mutant cells in the spinal cord and compared their gene expression profile to all other non-*cdx4* mutant cells in these clusters (Fig. 1h). Consistent with the known phenotype, expression of *hoxb9a,*

Fig. 1 | MIC-Drop-seq enables multiplexed single-cell phenotyping of mutant zebrafish embryos. a Overview of MIC-Drop-seq. **b** Histogram of gRNA molecule counts per cell (quantified by Unique Molecular Identifiers, UMIs) for cells with an assigned mutation. **c** Quantification of the number of cells in the dataset with no gRNAs detected (0), gRNA(s) targeting a single gene target (1) or gRNAs targeting multiple genes (2 +). **d** UMAP embedding of 24 hpf zebrafish cells from the MIC-Drop-seq experiment. Cell clusters are classified by color and are further described in Supplementary Data 2. **e** Comparison of observed DNA editing frequency and inferred genotypic composition of sequenced cells. The percentage of cells classified with the indicated genotype among all labeled cells was calculated (black) and compared to the observed indel frequency from bulk genomic DNA at the highest efficiency of the four cut sites for each gene (gray). **f**, **h** Subsets of UMAP embeddings that highlight cells with the predicted genotype indicated (black) vs other cells (gray) in the **f** presomitic mesoderm and **h** spinal cord cluster(s). **g** Quantification of the percent of cells from the indicated genotype in the presomitic mesoderm cluster. **i** Gene expression violin plot comparing selected *hox* gene expression in *cdx4* mutant cells compared to all other cells in the spinal cord cluster. *elavl3* is a common marker of spinal cord cells whose expression is unchanged. Source data are provided as a Source Data file. Created in BioRender. Carey, C. (2026) https://BioRender.com/z1frclu.

hoxc6b, and *hoxa9a* was significantly downregulated among *cdx4* mutant spinal cord cells (Fig. 1i, Supplementary Fig. 5c). Taken together, these experiments demonstrate that MIC-Drop-seq is a reliable and sensitive method for multiplexed mutant phenotyping with cellular resolution.

MIC-Drop-seq screen of 50 developmental genes

We next applied MIC-Drop-seq in a large-scale screen in an effort to discover uncharacterized developmental phenotypes. We selected a set of 50 transcriptional regulator genes that are variable in expression across cell types in the 24 hpf zebrafish embryo and highly expressed in at least one cell type (Supplementary Fig. 6a, see Methods for details). These transcription factors are expressed across a wide variety of cell types and have roles in development that range from well-characterized to mostly uninvestigated (Supplementary Fig. 6b).

A MIC-Drop library containing droplets targeting these 50 genes for mutagenesis was generated and injected into 1,000 zebrafish embryos (Supplementary Data 3). To facilitate downstream statistical analysis, injected embryos were split into pools of 250 embryos, representing four biological replicates (Fig. 2a). Each pool was independently dissociated into single-cells and sequenced at 24 hpf. As in the previous experiment, gRNA was captured and sequenced from overloaded cells, yielding similar rates of cells with single, multiple, or no gRNA types assigned (Supplementary Fig. 6c), and a median of 5 gRNA molecules recovered per cell (Supplementary Fig. 6d). After integration, filtering, and clustering, we recovered 226,492 cells distributed across 74 cell clusters (Fig. 2b, Supplementary Data 4). Cell types were again assigned with reference-based label transfer and broadly classified as either Neural, Mesoderm/Endoderm, or Non-neural ectoderm (Fig. 2b, Supplementary Data 4). Genotypes were confidently assigned in 135,881 cells, with each of the 50 genes represented in each of the 4 replicates with an average of 2,718 cells per genotype (Supplementary Fig. 6e).

To characterize mutant molecular phenotypes, we examined how each transcription factor disruption changed gene expression across all major cell types in the embryo. Differential gene expression testing leveraged the four biological replicates and revealed pervasive cell-type specific gene expression profiles for each mutant (Fig. 2c). Overall, transcription factor disruptions resulted in a median of 12 affected cell types (> 10 Differentially Expressed Genes (DEGs)) per genotype (Supplementary Fig. 7a). Some DEGs appeared repeatedly across all perturbed cell types and mutant genotypes (Supplementary Fig. 7b). Among these frequently observed DEGs, there was an enrichment for genes involved in Notch signaling, neurogenesis, cell cycle regulation, and stem cell differentiation, consistent with the roles of the targeted transcription factors in regulating embryonic development (Supplementary Fig. 7c).

Mutations in developmental regulators can alter cell type differentiation trajectories and change the abundance of different tissues or cell populations. We tested whether each of the 50 gene disruptions produced detectable changes in the abundance of cell types in the 24 hpf embryo. For each perturbation, fold-changes in cell counts were calculated for each replicate, enabling significance testing of enrichment or depletion of each cell cluster (Fig. 2d). Overall, changes in cell type abundance were less common than transcriptional phenotypes, with a median mutant having only a single major cell type with significantly altered abundance (Supplementary Fig. 8a).

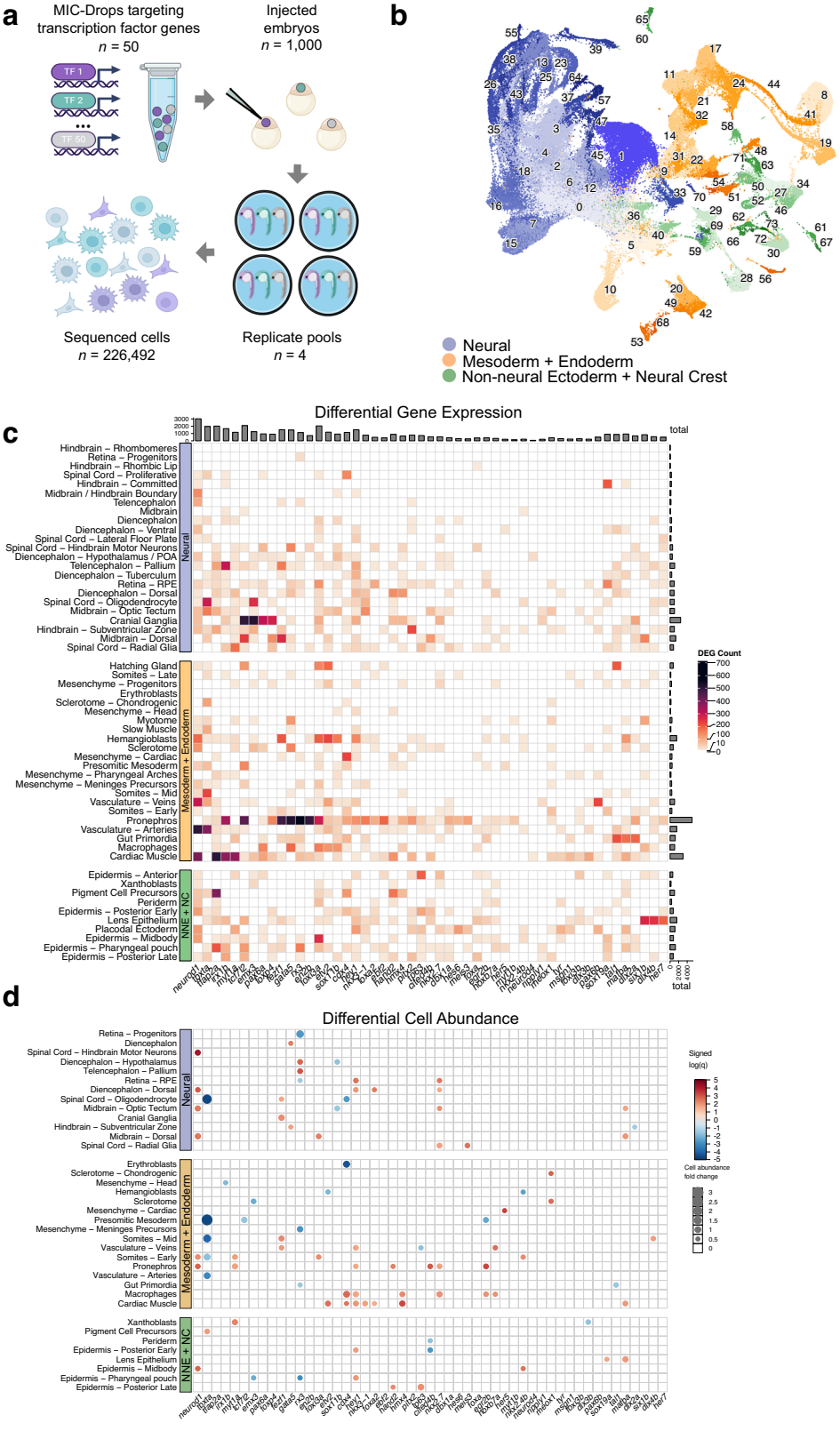
Together, these analyses demonstrate the ability of MIC-Drop-seq to comprehensively profile cellular and molecular phenotypes for hundreds of mutant animals in parallel. This rich dataset contains phenotypic reference information across hundreds of perturbation/cell type combinations. We created a web-based interactive tool to facilitate community exploration of the differential gene expression and cell type abundance data (link).

MIC-Drop-seq reveals transcriptional and cellular phenotypes

The molecular and cellular phenotypes generated and observed from MIC-Drop-seq serve as testable hypotheses for experimental validation, enabling the discovery of biological functions of target genes. Using orthogonal methods, we validated several previously undescribed phenotypes identified in our analyses. Transcriptional regulator *meox1* (mesenchyme homeobox 1) is a critical mesoderm regulator involved in somitogenesis, axial skeleton formation, and mesenchymal cell function[25–27]. Loss-of-function mutations in *meox1* are associated with congenital scoliosis and vertebral fusions, known as Klippel-Feil syndrome, in both humans and in zebrafish[28,29]. The specific set of effector genes that are dysregulated in *meox1* mutants that manifest this phenotype, however, has not been determined. At 24 hpf, *meox1* is highly expressed in paraxial mesoderm cell types. Among these cells, we detected a significant reduction in expression of a set of secreted extracellular matrix proteins in *meox1* mutant cells (Fig. 3a). These factors, including cartilage oligomeric matrix protein (*comp*), are specifically co-expressed with *meox1* in paraxial mesoderm at 24 hpf (Fig. 3b), but have not previously been shown to be altered in *meox1* mutants. To validate these findings, we injected embryos with Cas9/gRNA complexes targeting *meox1* and examined *comp* expression with in situ hybridization. In contrast to control injected embryos (Fig. 3c), *meox1* gRNA-injected embryos showed a loss of *comp* expression in myosepta and tailbud mesoderm, while retaining *comp* expression in the notochord (Fig. 3d). *Comp* has a well-documented role in matrix function, and mutations in this gene are commonly associated with disorders of the joints and skeleton in humans[30]. Thus, our finding of dysregulated *comp* expression in *meox1* mutants demonstrates the ability of MIC-Drop-seq to uncover localized developmental phenotypes for clinically relevant genes.

Developmental phenotypes can manifest as depletion or accumulation of different cell types. From the MIC-Drop-seq dataset, several transcription factor deletions resulted in changes in neural cell type abundance in the developing brain (Fig. 2d). We validated these phenotypes in the optic tectum and the hypothalamus / preoptic area (POA), two cell clusters where multiple mutants had altered cell abundance. In the optic tectum, *mafba, neurod1*, and *nkx2.7* mutant cells appeared enriched while *sox11b* cells were depleted based on cell count data (Fig. 3e and Supplementary Fig. 8b). We identified *gata2a* as a marker gene for the optic tectum, similar to the orthologous Gata2, which marks photorecipient neurons in the mammalian tectum[31]

https://doi.org/10.1038/s41467-026-70989-w



(Fig. 3f). In situ hybridization of *gata2a* confirmed the expansion of the optic tectum in the brains of *neurod1, mafba, nkx2.7* and a decrease in *sox11b* gRNA-injected embryos (Fig. 3g, h). We observed similar dynamic changes in cell abundance in the cell cluster corresponding to a ventral forebrain region encompassing hypothalamus and POA, where eyeless *rx3* mutant cells are enriched and *sox11b* cells are depleted (Fig. 3i, Supplementary Fig. 8b). In situ hybridization of *dlx1a*, a marker of the hypothalamus / POA precursors[32] (Fig. 3j), in *sox11b* gRNA-injected embryos confirmed a depletion of these cells compared to control injected embryos, while *rx3* gRNA-injected embryos showed a marked increase in *dlx1a* expression domain (Fig. 3k, l). Critically, none of the affected mutants showed changes in expression of marker genes *gata2a* and *dlx1a* on a per-cell basis (Supplementary Fig. 9a, b). These mutants also did not cause unexpected changes to other

https://doi.org/10.1038/s41467-026-70989-w

Fig. 2 | A 50 gene MIC-Drop-seq screen for early zebrafish development phenotypes. a Schematic of MIC-Drop-seq experiment. A MIC-Drop library was constructed with droplets targeting a set of 50 highly expressed transcription factors (see Methods) and injected into 1000 embryos. Embryos were divided at 24 hpf into 4 replicate groups of 250 and separately dissociated into single cells for sequencing. **b** UMAP embedding of 226,492 cells derived from MIC-Drop injected zebrafish embryos. Cluster numbers indicate individual cell types; colors indicate broad tissue classifications. Full cluster annotations are detailed in Supplementary Data 4. **c** Heatmap of differential gene expression in prominent cell types (> 800 recovered cells). Color scale indicates the total number of differentially expressed genes (DEGs) in each cell type in cells with the indicated perturbation compared to all other classified cells in the cluster. Total DEG counts (up or downregulated) are shown for each cell type (row) and perturbation (column). Cells are grouped by tissue categories: Neural, Mesoderm + Endoderm, and Non-neural Ectoderm + Neural Crest (NNE + NC). **d** Bubble plot of differentially abundant cell types in prominent cell types (> 800 recovered cells). Cell counts were used to calculate the fold change in cell abundance for each mutant genotype in every cell type (bubble size), while q values were calculated using hooke[59] (color scale), and signed corresponding to a depletion (- sign, blue) or enrichment (+ sign, red) for each cell type. Changes in cell abundance with $q < 0.05$ and absolute $\log_2$FC of > 0.5 are shown. Cell clusters with no detected cell abundance changes in any mutant were omitted. Source data are provided as a Source Data file. Created in BioRender. Carey, C. (2026) https://BioRender.com/7zmkxz0.

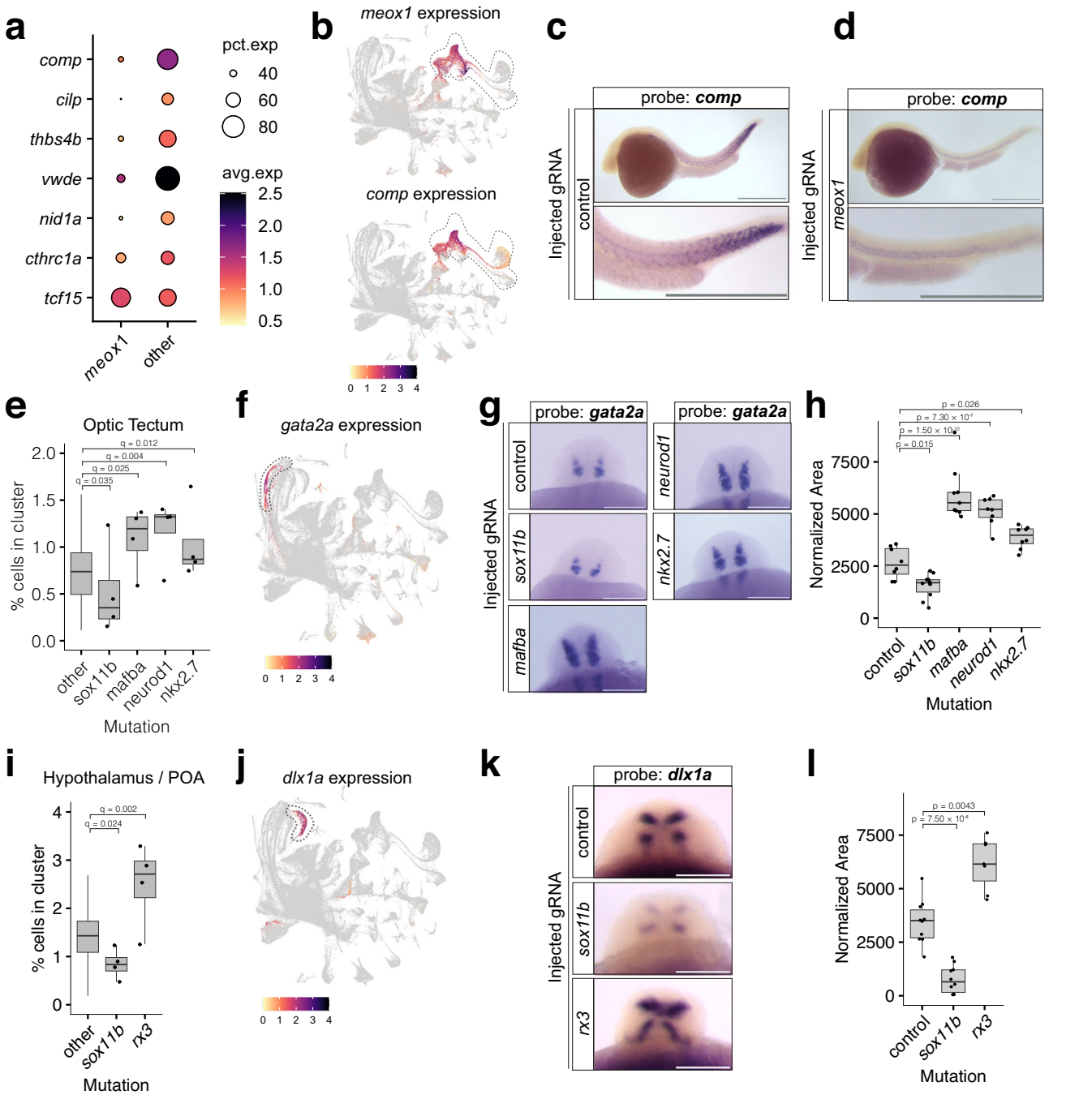


Fig. 3 | MIC-Drop-seq enables discovery of transcriptional and cellular phenotypes. a Gene expression bubble plot comparing expression of a set of secreted extracellular matrix factors in *meox1* mutant vs other cells in the somitic mesoderm (myotome, sclerotome, presomitic mesoderm, and early somite clusters). *tcf15* is a marker of somitic mesoderm and is unchanged in *meox1* mutant cells. **b** Gene expression feature plots for *meox1* and *comp*. Somitic mesoderm clusters are highlighted inside the dashed line. **c**, **d** Colorimetric RNA in situ hybridization of *comp* mRNA in **c** control and **d** *meox1* gRNA injected 24 hpf embryos. Scale bars = 400 μm. **e** Quantification of cell type abundance in the optic tectum cluster. The percentage of cells in the cluster is shown for each indicated mutant genotype, with individual dots representing values from the 4 replicate pools ($n = 4$ biological replicates per genotype), compared to all other genotypes across replicates (Other, $n = 188$ observations). Differential cell abundance was tested using Hooke[59] with Poisson-Lognormal models. Hooke q values for highlighted perturbations: *neurod1* $q = 4.13 \times 10^{-3}$, *nkx2.7* q = 0.012, *mafba* q = 0.025, *sox11b* q = 0.035. **f** Gene expression feature plot of *gata2a*; optic tectum cluster is encircled (dashed line). **g** Representative dorsal-view colorimetric RNA in situ hybridization images for *gata2a* mRNA are shown for 24 hpf zebrafish embryos injected with the indicated gRNA. Scale bars = 100 μm. **h** Quantification of *gata2a* stain area for embryos injected with the indicated gRNA (control $n = 8$, *sox11b* $n = 10$, *mafba* $n = 9$, *neurod1* $n = 8$, *nkx2.7* $n = 8$ embryos; each data point represents one individual embryo). **i** Quantification of cell type abundance in the hypothalamus / preoptic area (POA) cluster. The percentage of cells in the cluster is shown for each mutant genotype, with individual dots representing values from the 4 replicate pools ($n = 4$ biological replicates per genotype), compared to all other genotypes across replicates (Other, $n = 196$ observations). Differential cell abundance was tested using Hooke[59] with Poisson-Lognormal models. Hooke q values for highlighted perturbations: *rx3* $q = 2.02 \times 10^{-3}$, *sox11b* q = 0.024. **j** Gene expression feature plot of *dlx1a* expression in the single-cell dataset, hypothalamus / POA cluster is encircled (dashed line). **k** Representative dorsal-view images of colorimetric RNA in situ hybridization of 24 hpf zebrafish embryos for *dlx1a* mRNA are shown for the indicated genotype. Scale bars = 400 μm. **l** Quantification of *dlx1a* stain area in 24 hpf embryos injected with the indicated gRNA (control $n = 10$, *sox11b* $n = 10$, *rx3* $n = 11$ embryos; each data point represents one individual embryo). Significance testing in **h** and **l** was performed with one-way ANOVA followed by two-sided Dunnett's post-hoc test comparing each mutant to control. For panel **h**: $F(4,38) = 44.65$, $p = 7.24 \times 10^{-14}$; Dunnett's p values: control vs *sox11b* $p = 0.026$, 95% CI [−2069, −108]; control vs *mafba* $p = 1.50 \times 10^{-10}$, 95% CI [2328, 4337]; control vs *neurod1* $p = 7.30 \times 10^{-7}$, 95% CI [1457, 3524]; control vs *nkx2.7* $p = 0.015$, 95% CI [207, 2274]. For panel **l**: $F(2,28) = 43.24$, $p = 2.74 \times 10^{-9}$; Dunnett's p values: control vs *sox11b* $p = 0.0043$, 95% CI [−4533, −823]; control vs *rx3* $p = 7.50 \times 10^{-6}$, 95% CI [2649, 6274]. In all box plots, the center line represents the median, box bounds represent the 25th and 75th percentiles, and whiskers extend to 1.5× the interquartile range (IQR); outliers are shown as individual points. Source data are provided as a Source Data file.

structures including the notochord (Supplementary Fig. 9f, g). In the hindbrain, while *mafba* mutants displayed reduced rhombomere 5 (R5) width (Supplementary Fig. 9e), this phenotype is consistent with the established role of *mafba* in hindbrain patterning[33,34]. Other mutants showed no changes in rhombomere size (Supplementary Fig. 9c–e), demonstrating the specificity of the described mutant phenotypes. Together, these data validate several transcriptional and cellular phenotypes identified by MIC-Drop-seq, demonstrating its use to discover previously uncharacterized regulators of tissue development from large-scale pooled datasets.

MIC-Drop-seq detects indirect developmental effects

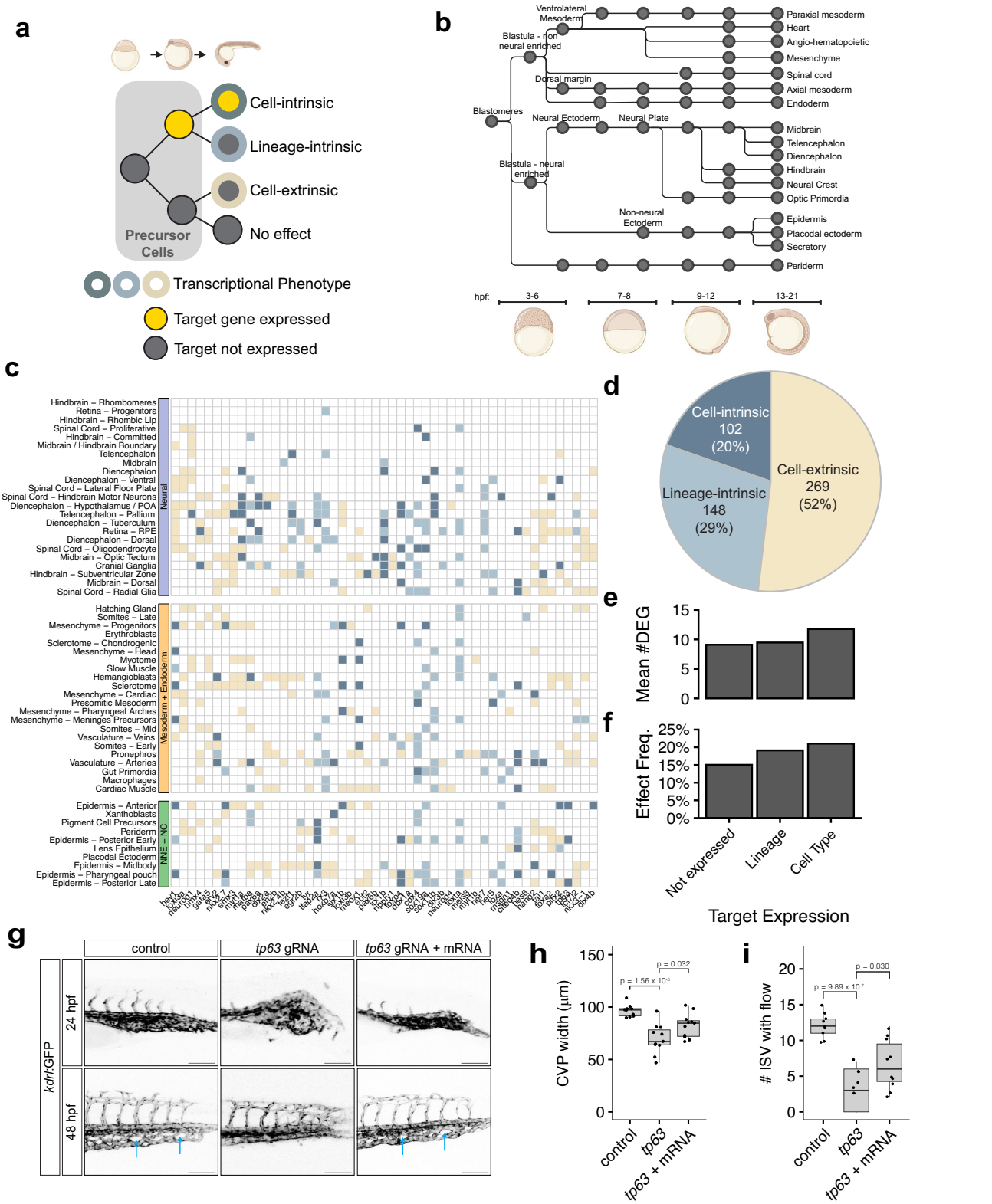
During development, cells frequently influence each other's behavior through direct contact or secreted signals. These cell-extrinsic effects (also called non-cell autonomous or indirect effects) can be triggered by transcription factors acting within one cell type but affecting others. Such interactions are difficult to predict in whole animals and often absent in simplified cell culture systems. We reasoned that our large MIC-Drop-seq dataset, containing transcriptional profiles from 50 mutant genotypes, provides an opportunity to systematically identify and classify these cell-extrinsic phenotypes. To distinguish direct from indirect effects, we established clear definitions (Fig. 4a): phenotypes are "cell-intrinsic" when the disrupted transcription factor is expressed in the affected cell type, "lineage-intrinsic" when expressed in progenitor populations but not the affected 24 hpf cell type, and "cell-extrinsic" when the factor is not expressed in the affected cell type or its developmental lineage. We integrated our MIC-Drop-seq data with Daniocell, a time-course scRNAseq atlas of zebrafish embryogenesis[8], and used published lineage relationships[9,10] to build a simplified developmental tree tracking cell types from early blastula through 24 hpf (Fig. 4b). This allowed us to trace whether each transcription factor was expressed at any point during each cell type's developmental history (Supplementary Fig. 10a).

We next classified transcriptional phenotypes using these definitions across all 50 mutants in each cell type. The number of differentially expressed genes (DEGs) varied widely across cell type/perturbation combinations, with most showing few or no changes. We set a threshold of >10 DEGs to define a meaningful transcriptional phenotype, capturing the top 30% of observations (Supplementary Fig. 10b). For each affected cell type, we determined whether the target transcription factor was expressed in the affected cell type or its lineage using expression data from Daniocell (see Methods), classifying each as cell-intrinsic, lineage-intrinsic, or cell-extrinsic (Fig. 4c).

We sought to address two questions about transcription factor function during development: first, how commonly do transcription factors exert cell-extrinsic effects, and second, does target gene expression predict phenotype occurrence and strength? Strikingly, classification of each transcriptional phenotype based on target gene expression history revealed that cell-extrinsic effects were the most common phenotype class by raw count, accounting for over half of all observed effects (Fig. 4d), a pattern that held across a range of different DEG thresholds (Supplementary Fig. 10c, d). Examining all cell type/perturbation combinations stratified by target gene expression showed that transcriptional responses occurred more frequently when targets were expressed in the affected cell type or its lineage (Fig. 4e), and these cell types showed higher average DEG counts (Fig. 4f). Thus, while cell-extrinsic effects represent the dominant transcriptional phenotype class, transcription factor perturbations still generally produce stronger effects in cells where they are expressed. Taken together, most cell types exhibiting transcriptional responses to perturbations in our dataset do so through cell-extrinsic mechanisms, demonstrating the prevalence of indirect interactions in development.

To verify the veracity of cell-extrinsic phenotypes observed by MIC-Drop-seq, we focused on vascular cell types, which displayed cell-extrinsic phenotypes in multiple mutants. For example, disrupting *tp63*, a transcription factor critical for epidermal cell development[35], caused a decrease in venous cell abundance (Supplementary Fig. 11a) and altered gene expression in all vascular cell types, indicating atypical vascular development (Supplementary Fig. 11b). These phenotypes are unexpected, given that *tp63* is not expressed in vascular cells but is instead restricted to the epidermis (Supplementary Fig. 11c). To verify dysregulated vascular development in *tp63* mutants, a transgenic *kdrl*:GFP zebrafish line was used to compare vascular patterning in embryos injected with either *tp63*-targeting or control gRNAs. At 24 hpf, *tp63* mutants displayed altered caudal vein plexus (CVP) morphology (Fig. 4g). At 48 hpf, vascular development remained disrupted, with failure to form typical vascular fenestrations in the CVP of *tp63* mutants (Fig. 4g) and decreased CVP width (Fig. 4h). Notably, while length of intersegmental vessels (ISVs) remained unchanged (Supplementary Fig. 11d), optical flow analysis revealed defects in erythrocyte transit in *tp63* mutants (Supplementary Fig. 11e), including a significant decrease in the number of ISVs with detectable flow

https://doi.org/10.1038/s41467-026-70989-w



(Fig. 4i). Co-injection of *tp63* mRNA partially rescued these vascular phenotypes, with embryos showing restoration of CVP morphology, width, and blood flow, confirming the specificity of these defects to *tp63* loss-of-function (Fig. 4g–i). Intussusceptive angiogenesis in the CVP is linked to shear force supplied from cardiac output[36], and p63 null mice have defects in heart formation[37]. While *tp63*-injected zebrafish embryos had defects in skin formation (Supplementary Fig. 11f), no changes in heart rate were apparent (Supplementary Fig. 11g). These defects in early vasculogenesis could therefore be the result of altered physical forces stemming from an improperly formed epidermal layer. Together, these findings demonstrate how the unbiased whole-animal datasets generated by MIC-Drop-seq can detect unanticipated emergent phenotypes, even in cell types where the target gene is not expressed.

Fig. 4 | MIC-Drop-seq reveals pervasive cell-extrinsic effects of transcription factor disruption. a Schematic of classification of perturbation types. Perturbed cell types with significant differential gene expression can be a result of direct cellular effects (target gene currently expressed in perturbed cell type), direct lineage effects (target expressed only in precursor cells of the same lineage), or indirect effects (target not expressed in perturbed cell type or its precursors). **b** Map of zebrafish cell lineage differentiation trajectories. Each cell type label is derived from the Daniocell reference dataset, with literature-based lineage connections. **c** Categorical map of perturbation effects. Detected phenotypes were classified as cell-intrinsic when the target gene is expressed in the affected cell type (dark blue) or precursor cells (light blue), while phenotypes in cell types with no history of expression were classified as cell-extrinsic (yellow). Clusters with <10 DEGs are classified as not affected (white). **d** Proportion plot of each perturbation type within the dataset. Totals and percentages for each perturbation type are indicated. **e**, **f** Transcriptional responses were grouped by target gene expression context for each measured cell type: target gene not expressed in the affected cell type (not expressed), target gene expressed only in lineage precursors (lineage), or target gene expressed in the cell type (cell type). For each expression scenario: **e** bars show the frequency of transcriptional phenotypes observed (> 10 DEGs). **f** bars show the mean number of differentially expressed genes. **g** Representative greyscale fluorescence images of Caudal Vein Plexus (CVP) and posterior trunk of *kdrl*:GFP labeled transgenic embryos injected with control gRNA, *tp63* gRNA, or *tp63* gRNA + *tp63* mRNA at 24 and 48 hpf. Blue arrows highlight CVP fenestrations in control embryos. Scale bar = 100 μm. **h**, **i** Quantification of vascular phenotypes at 48 hpf in embryos injected with control gRNA, *tp63* gRNA, or *tp63* gRNA + *tp63* mRNA (control $n = 10$, *tp63* $n = 11$, *tp63* + mRNA $n = 11$ embryos for panel **h**; control $n = 10$, *tp63* $n = 10$, *tp63* + mRNA $n = 10$ embryos for panel **i**; each data point represents one individual embryo) measuring **h** width of the Caudal Vein Plexus (CVP) and **i** number of Intersegmental Vessels (ISVs) with detectable flow. Statistical significance was assessed using one-way ANOVA followed by two-sided Tukey's post-hoc test. For panel **h**: ANOVA $F(2,29) = 15.46$, $p = 2.69 \times 10^{-5}$; Tukey's p values: control vs tp63 $p = 1.56 \times 10^{-5}$, 95% CI [− 38.8, −14.9]; control vs tp63 + mRNA p = 0.017, 95% CI [− 26.2, −2.34]; tp63 vs tp63 + mRNA p = 0.032, 95% CI [0.953, 24.2]. For panel **i**: ANOVA $F(2,25) = 24.02$, $p = 1.51 \times 10^{-6}$; Tukey's p values: control vs tp63 $p = 9.89 \times 10^{-7}$, 95% CI [− 12.3, −5.74]; control vs tp63 + mRNA $p = 6.10 \times 10^{-4}$, 95% CI [− 8.70, −2.34]; tp63 vs tp63 + mRNA $p = 0.030$, 95% CI [0.296, 6.66]. In all box plots, the center line represents the median, box bounds represent the 25th and 75th percentiles, and whiskers extend to 1.5× the interquartile range (IQR); outliers are shown as individual points. Source data are provided as a Source Data file. Created in BioRender. Carey, C. (2026) https://BioRender.com/vxn9jb2, https://BioRender.com/xjdsd2x.

Discussion

Profiling mutant phenotypes with scRNAseq provides a way to comprehensively map the link between genotype and the networks of genes, cells, tissues, and organs underlying animal development. In this report, we present MIC-Drop-seq as a technique for pooled CRISPR screening in whole zebrafish embryos with cellular resolution phenotyping across all embryonic cell types. MIC-Drop-seq is related to existing multiplexed scRNAseq technologies like CROP-seq and Perturb-seq, but offers several differentiating features that enable its use in whole animals. CROP-seq and Perturb-seq similarly rely on sequencing gRNA-linked molecular barcodes to index mutant genotypes, but have been limited to screens of cultured cells, constraining the scope of observable phenotypes to a subset of molecular networks active in a handful of cell types[16,38–40]. In-vivo Perturb-seq uses transduction of CRISPR-containing viral particles to create clones with differing mutant genotypes within individual animals[19,41]. This approach, while useful for studying tissue-specific regulatory networks, fails to capture earlier developmental phenotypes of the targeted genes or to isolate gene-specific cell-extrinsic effects. MIC-Drop-seq uses gene disruption at the zygotic stage, where mutations are rapidly generated and inherited by early progenitors during development, enabling whole-animal developmental phenotypes to be profiled with cellular resolution at scale. Furthermore, MIC-Drop-seq uses injection of purified Cas9 RNP complexes instead of viral vectors, dramatically reducing the possibility of genotyping errors from ambient gRNA or cells infected with multiple viral vectors. These features make MIC-Drop-seq a highly efficient platform that could be scaled to study hundreds to thousands of whole-animal mutants in parallel.

The cellular-level comparisons made possible by MIC-Drop-seq enable detection of phenotypes that would otherwise be difficult to detect by conventional observation methods or would not manifest in vitro. For example, we identified and validated widespread cell abundance changes in the developing brain of *sox11b* mutant embryos, which appear visually indistinguishable from wildtype siblings and have been classified as developmentally normal in previous studies[42,43]. Furthermore, our transcriptomic profiling revealed extensive changes in the makeup of the myosepta extracellular matrix in *meox1* mutants, despite only having a visible phenotype at much later stages of development. While recent studies have also leveraged parallel scRNAseq of multiple mutant embryos to detect such cryptic developmental phenotypes[11], MIC-Drop-seq meaningfully increases throughput by enabling rapid delivery and demultiplexing of perturbations. Indeed, our approach enabled the simultaneous collection of phenotypic information for 50 mutants across 74 embryonic cell types in a single experiment carried out by a small team in a single day. This unbiased approach revealed pervasive cell-extrinsic phenotypes, including an unexpected emergent phenotype in the vasculature of embryos with disrupted *tp63*, a transcription factor that is expressed only in ectoderm in the 24 hpf embryo.

Our MIC-Drop-seq experiments focused on the 24 hpf zebrafish embryo, but further modifications might enable a broader set of applications. MIC-Drop-seq is in principle compatible with any species where embryos are amenable to injection at the zygotic stage. While we demonstrate efficient gRNA sequence recovery at 24 hpf zebrafish embryos, studying later developmental time points will likely result in lower recovery as the gRNAs are diluted and degraded. DNA barcodes or chemically-stabilized gRNAs could enhance genotype recovery at later stages[44]. Another challenge of whole-animal scRNAseq as a phenotyping method is the detection of transcriptional effects in low-abundance cell types. Small sample sizes can reduce statistical power and increase the risk of both false positives and false negatives in differential expression testing. To mitigate these risks, we applied a conservative threshold requiring >800 recovered cells per cluster for inclusion in our analysis. Statistical power for rare cell types could be improved by leveraging emerging scRNAseq technologies that enable profiling of larger cell numbers across additional replicates[45,46], or by using cell sorting to enrich specific populations of interest prior to screening. Combining MIC-Drop-seq with oligonucleotide-based indexing of individual embryos, as implemented in approaches like ZSCAPE[11,12], could further aid in enhancing mutant cell recovery through genotype imputation from cells where gRNAs are not recovered.

MIC-Drop-seq is a flexible phenotyping platform that can be integrated with other perturbation types or single-cell methods in the future. For example, combining MIC-Drop-seq with scATAC or scCUT&TAG[47] would enable deeper probing of developmental gene regulatory networks. MIC-Drop-seq could be used to deliver and characterize the combinatorial effects of chemical and genetic perturbations in pharmacogenetic screens. More precise genome engineering methods, like base or prime editing[48], could enable the simultaneous profiling of phenotypes of panels of specific genetic variants. Finally, MIC-Drop-seq could be coupled with other quantitative and high-throughput measures of phenotype, including body morphology, neural activity, and behavior. Phenotypic data from systematic perturbation screens like these can be used to train predictive models of animal development[49–51]. Large-scale molecular

phenotyping of mutant animals with methods like MIC-Drop-seq will be crucial to providing training data to bridge the gap between molecular and higher-order organismal phenotypes. Such models will aid in connecting, interpreting, and predicting the gene regulatory networks that build a functional animal.

Methods

Ethics statement

All experiments were performed on Zebrafish (*Danio rerio)* strain TuAB. Experiments were approved and performed in accordance with University of Utah Institutional Animal Care and Use Committee protocols #00001701 and #00001487 and Northwestern University Institutional Animal Care and Use Committee protocols #IS00027025 and #IS00026745. Disaggregation by sex was not applicable to the experiments in this manuscript because embryos were profiled prior to sex determination.

Oligonucleotide synthesis

All DNA oligonucleotides (PCR primers, gRNA synthesis templates) were ordered as unmodified custom single-stranded DNA oligonucleotides or ultramers from Integrated DNA Technologies (IDT).

Guide RNA design for feature barcoding

Guide RNAs (gRNAs) targeting the genes-of-interest were designed using CHOPCHOP as previously described[6,7,52]. Briefly, four gRNAs targeting each gene were filtered from CHOPCHOP output based on the following criteria: (1) target sites of 20 bp length that start with GN, (2) non-overlapping target sites targeting the first half of the coding sequence, (3) gRNAs with predicted on-target efficiency score of >40, (4) gRNAs with GC content of 45-75%, (5) gRNAs without poly Ts (TTTT), and (6) gRNAs with no predicted off-targets with 3 bases or fewer mismatches. In cases where gRNAs with no off-targets could not be identified, gRNAs that contain at least 3 mismatches at the off-target sites were used. gRNAs that contain at least one mismatch within the seed region were prioritized. Target sites for all genes are listed in Supplementary Data 1 and 3.

DNA template synthesis and in vitro transcription (IVT)

Gene-specific forward oligos (5′-ATTTAGGTGACACTATA GN_{19} GTTTAAGAGCTAAGCTGGAAACA-3′, where GN_{19} refers to the sgRNA target/spacer sequence) containing an SP6 promoter were ordered at 500 pmol scale from IDT. The sgRNA universal reverse oligo containing the capture sequence (5′AAAAAAAGCACCGACTCGGTGCCACTTGGCCTTGCTAGGACCGGCCTTAAAGCGGCCAAGTTGATAACGGACTAGCCTTATTTAAACTTGCTATGCTGTTTCCAGCTTAGCTCTTAAAC-3′) was ordered from IDT as an ultramer. To generate the DNA template for IVT, a fill-in PCR was performed as described previously with slight modifications[53]. Briefly, a 50 μL mix was assembled containing 2 μM forward gene-specific oligo, 2 μM universal reverse oligo, 1x HF buffer, 200 μM dNTPs (Invitrogen, cat. no. 10297018), 3% DMSO, 1 U of Phusion polymerase (NEB, cat. no. M0535L). The fill-in reaction was performed in a thermal cycler at the following conditions: 98 °C for 2 min, 50 °C for 10 min, 72 °C for 10 min. This DNA template was cleaned up using a PCR cleanup kit (Zymo Research, cat. no. D4024). DNA templates for the four gRNAs targeting each gene were then pooled and used for in vitro transcription (IVT). IVT was performed using a MEGAScript SP6 kit (Invitrogen, cat. No. AM1330) per manufacturer's protocol. Briefly, a 10 μL IVT mix was assembled containing 5 mM of NTPs, 1x reaction buffer, 1 U RNase inhibitor (Thermo Scientific, cat. no. EO0381), 1x SP6 enzyme mix, and 20 ng/μL pooled template DNA. The sample was incubated at 37 °C overnight after which 10 μL nuclease free water and 1 μL of 2 U/μL Turbo DNase I was added, and the sample incubated for an additional 20 minutes. RNA cleanup was performed using a ZR96 RNA Clean and Concentrator-5 (Zymo Research; cat. no. R1080) kit per manufacturer's instruction. RNA concentration was determined using a Nanodrop and RNA integrity verified using gel electrophoresis.

MIC-Drop library design and injection

3% FS-HFE was prepared by dissolving 1 g of 008-fluorosurfactant (Ran Biotechnologies, cat. no. 008-Fluorosurfactant) in 33.3 mL of HFE-7500 (3 M, cat. no. Novec 7500). A 21.5 μL ribonucleoprotein (RNP) complex targeting each gene containing 200 ng/μL gRNAs, 3.36 μM EnGen Cas9 (NEB, cat. no. M0646M) and 1x NEB buffer r3.1 (NEB, cat. no. B6003S) was assembled and incubated at room temperature for 5-10 min. 3.5 μL of 0.5% phenol red dye (Sigma-Aldrich, cat. no. P0290-100 mL) was added to each sample and mixed. The droplets containing RNP were generated using a BioRad QX200 droplet generator (BioRad, cat. no. 1864002) as previously described[6,7]. Briefly, 20 μL of RNP mix was transferred to wells labeled 'sample' of a QX 200 cartridge (BioRad, cat. no. 1864007). 70 μL of 3% FS-HFE was added to the wells labeled 'oil'. The cartridge was sealed using the provided rubber gasket and loaded in the droplet generator. Droplet generation starts as soon as the lid is closed. Droplets for 8 samples were generated simultaneously from a single cartridge. The droplets were carefully transferred to PCR strip tubes containing 50 μL 3% FS-HFE. To generate the MIC-Drop library, equal volume of droplets for each sample were pooled into a separate PCR tube containing 50 μL 3% FS-HFE. The droplets were gently mixed to get an even distribution and stored at 4 °C. 3-5 μL of the intermixed droplets were transferred to a microinjection needle (World Precision Instruments, cat. no. TW100F-3) using a microloader tip (Eppendorf, cat. no. 5242956003). A single droplet per embryo was injected as previously described[6,7]. Injected embryos were transferred to a petri dish (30-50 embryos/petri dish) in E3 medium containing methylene blue and incubated at 28.5 °C.

Embryo dissociation and single-cell library preparation

Dead and severely deformed embryos were removed at 24 hpf. For the pilot experiment, 40 total injected embryos were pooled. Dissociation of embryos was performed similarly to previously published protocols[54]. 1 mL of 1X FACSmax (Fisher Scientific, cat no. T200100) solution was added to the sample and incubated at 28 °C for 20 min. To generate a single-cell suspension, the sample was pipetted up and down 10 times every 5 min using a P1000 pipette. Cells were then passed through a 40 μm strainer and centrifuged at 310 g for 5 minutes, washed once with 1X FACSmax solution, and resuspended in phosphate buffered saline (Thermo Fisher Scientific, cat no. 14190094) with 0.5% bovine serum albumin (Thermo Fisher Scientific, cat no. BP9706100). Cell viability of >90% was determined for each sample using Acridine Orange (Thermo Fisher Scientific, cat no. A1301) and Propidium Iodide (Thermo Fisher Scientific (cat no. P1304MP)) staining. For the screen of 50 transcription factors, after removing any dead and severely deformed embryos at 24 hpf, injected embryos were split into four pools of approximately 250 individuals each. Each pool was dissociated into a single-cell suspension as described earlier. Each pool of dissociated cells was split into four parts, for a total of 16 samples. Each sample was loaded and processed on a 10X Chromium controller using Next Gem 3.1.

Gene expression matrix generation and genotype assignment

Single-cell RNA sequencing transcriptome reads were processed using Cell Ranger software from 10X Genomics and aligned to the GRCz11 v4.3.2 zebrafish reference genome[55]. Gene expression count matrices were generated using the Cell Ranger default settings. A custom reference was used containing the sequence of each gRNA protospacer (Supplementary Data 1 and Supplementary Data 3) to generate a separate CRISPR gRNA count matrix for each sample. To distinguish ambient gRNA reads from injected gRNAs, a Gaussian mixture model

(Cell Ranger software, 10X Genomics) was used to set UMI thresholds for each detection of each gRNA species. Predicted genotypes were assigned to each cell based on the presence of gRNA using these thresholds. Cells with no guide RNA counts exceeding the UMI threshold were assigned "Not Determined". Cells with gRNA counts exceeding UMI thresholds for multiple gene targets were assigned as "Multiple". Cells with detectable gRNAs for a particular gene were assigned as having that mutant genotype. In silico doublet detection was performed with scDblFinder[20].

Dimensionality reduction, clustering and cell cluster labeling

Quality control filtering was implemented to remove cells with >15% of reads mapping to mitochondrial RNA, and cells with fewer than 200 unique detected genes. Cells with gRNA counts for multiple genes were considered doublets and removed from the dataset. Filtered gene expression matrices were log normalized and used to identify variable features for scaling and dimensionality reduction. When multiple sequencing runs were used to generate a dataset, the samples were merged into a single dataset using the data integration functions in Seurat before dimensionality reduction and clustering.

Cell clusters were annotated in two steps. First, a reference atlas was used containing labeled cell types from the Daniocell atlas of zebrafish development, using time points between 21 and 26 hpf[8]. The reference atlas was used to generate transfer anchors using the principal component analysis reference reduction using the TransferData function in Seurat[56], generating a predicted cell type label for each cell in the dataset. Cell cluster labels were then assigned based on a consensus of predicted cell types. Predicted cell types were verified and adjusted using marker gene expression (Supplementary Data 2 and Supplementary Data 4).

Determination of CRISPR editing frequency

A subset of the pool of cells used for single-cell RNA sequencing library preparation were retained for analysis of DNA editing at each target locus. Genomic DNA was extracted by lysing the cells in 1x PCR lysis buffer (10 mM Tris (pH 8), 2 mM EDTA, 0.2% Triton X-100) containing 0.2 mg/mL Proteinase K (MilliporeSigma, cat no. 3115887001). The RNase H2-dependent PCR platform (Integrated DNA Technologies) was used for pooled amplification of each cut site, according to the manufacturer's specifications. Briefly, pooled primers were designed and produced by IDT to amplify each CRISPR cut site using the provided design tool. To avoid overlapping PCR products, the reactions were split into two pools. For each pool, a 20 μl PCR reaction was created containing 2 ng gDNA, 250 μM of each primer, and 1X rhAmpSeq library mix. A second PCR reaction was performed to add sequencing adapters and a sample index, followed by AMPure XP bead purification. Libraries were generated for both experimental MIC-Drop injected embryo cells as well as wild-type control cells. Each library was sequenced using an Illumina MiSeq instrument with a Micro 150 cycle Paired End Sequencing Kit v2. The IDT rhAmpSeq CRISPR Analysis Tool was used to assess DNA editing frequency at each locus. DNA editing was quantified by the presence of nucleotide insertions or deletions (indels) in individual reads compared to each amplicon reference sequence. For each target gene, the gRNA with highest indel frequency was used as a conservative estimate of editing frequency (Fig. 1e). Frameshift rates were calculated using the ampliCan R package for each of the individual gRNA targeting sites[57]. Frameshift frequencies were normalized to the proportion of cells with the corresponding inferred genotype in the single-cell dataset. Because intragenomic frameshift frequency could not be directly measured, we calculated a cumulative frameshift probability for each target gene. This was calculated by determining the complement of the product of wild-type allele frequencies across all four sites: $1 - [(1\text{-}X_1) \times (1\text{-}X_2) \times (1\text{-}X_3) \times (1\text{-}X_4)]$, where X_i represents the frameshift frequency at each gRNA target site.

Differential gene expression analysis

In the pilot dataset (Fig. 1), which did not include biological replicates, differential gene expression was calculated with a Wilcoxon rank sum test as implemented in Seurat[56]. Tests were performed by first subsetting cells from each cluster, then comparing gene expression in cells of each mutant genotype against all other classified cells in the cluster.

For the 50 gene MIC-Drop-seq screen, differential gene expression testing was performed using pseudobulk analysis using edgeR[58] for each genotype in each cell type. First, a subset was created for each cell type with more than 800 total cells, removing any cells with undetermined genotype (ND). Next, aggregate read count libraries were generated for each cell type and genotype, grouped by biological replicate. Each library was filtered by expression and normalized prior to differential gene expression testing with the glmQLFit test in edgeR. Genes with an FDR < 0.05 and an absolute value $\log_2$FC greater than 0.5 were considered differentially expressed.

Differential cell abundance analysis

Cell counts were tabulated for each mutant genotype in each cell cluster across each biological replicate pool, excluding cells with undetermined genotype (ND). Hooke[59] (https://github.com/cole-trapnell-lab/hooke), an extension of the Monocle3 package[60,61], was used to test for differential cell abundance using cell count data. Abundance fold change and significance testing used Poisson-Lognormal models (PLNmodels)[62]. Estimated cell abundances were used to calculate positive or negative $\log_2$FC in cell abundance in mutant cells compared to all other cells in each cluster. Cell type abundance changes were considered significant with an FDR < 0.05 and an absolute $\log_2$FC of > 0.5.

Determination of cell extrinsic and cell intrinsic effects

To determine the lineage history of gene expression for each targeted gene, the Daniocell dataset was downloaded and subsetted to include cell types and time points present from 3-24 hpf[8]. Average expression of each factor was calculated for every cell type and time point. Lineage relationships were manually curated based on prior studies[9,10,63]. Gene expression values for each target in each cluster were scaled to a range of 0-1. Factors were considered expressed if the scaled expression value exceeded 0.2 in the cell type at 24 hpf. Transcriptional phenotypes were classified for each cell type with a threshold of 10 differentially expressed genes (absolute $\log_2$FC > 0.5, FDR < 0.05). Phenotypes were classified as cell-intrinsic when the cell type or its lineage precursors expressed the factor or were otherwise classified as cell-extrinsic.

Zebrafish F0 mutant validation

F0 CRISPR injected embryos for validation experiments were generated using the same gRNAs as were employed in the 50 gene MIC-Drop-seq screen (Supplementary Data 3). Control injections contained equimolar amounts of Cas9 with gRNAs that do not target any locus in the zebrafish genome. 1-1.5nL of the RNP was injected into zebrafish embryos at the single-cell stage. The RNP injection mixture contained 100 ng/μL gRNAs, 2 μM EnGen Cas9 NLS (NEB, cat no. M0646M), 1x Buffer r3.1 (NEB, cat no. B7203S) and 0.1% Phenol Red. At 24 hpf, embryos were manually dechorionated and fixed in 4% Paraformaldehyde Solution in Phosphate Buffered Saline (ChemCruz, cat no. sc-281692) overnight at 4 °C. Fixed embryos were washed with PBS and either taken directly into in situ hybridization staining protocols or washed with successive washing of 50% and 100% methanol for short term storage at 4 °C.

In situ hybridization

Antisense probes for in situ hybridization were generated by first PCR amplifying genomic regions from gDNA extracted from TuAB embryos[64]. Primers were designed incorporating T7 RNA polymerase promoter sequences for amplifying probes for *gata2a* (forward 5′-CACTGACCATGAAGAAGGAG-3′, reverse: 5′-GGATCCTAATAGACT CACTATAGGGAAAGCAAAGGAAGTACTCATGG-3′), *dlx1a* (forward: 5′-AACCTTTGAGCGAG TTTGTG-3′, reverse: 5′GGATCCTAA-TACGACTCACTATAGGGAAAGTCCAACATT CTCGTAGT GC-3′), *comp* (forward: 5′-GCCAAACTCATCACCATCAAAT-3′, reverse: 5′-GGATCC TAATACGACTCACTATAGAGAG CTTTCTTTGGGCCACTT-3′), and *egr2b* (forward: 5′- CCCTCTTGCCGATAGCATCT-3′, reverse 5′-GTTGGAAAAAGCCGGCGTAG-3′). The resulting purified PCR products were used as templates for in vitro transcription using the Roche Digoxigenin labeling kit (Sigma, cat. no. 11175025910). In situ hybridization procedure was performed as described previously[64]. For probe hybridization, 50 ng of purified RNA probe was incubated overnight at 70 °C in 200 μl hybridization mix containing 5% Dextran sulfate.

In situ imaging and quantification

Stained embryos were mounted in glycerol and imaged with a Leica M205 dissecting microscope. For *comp, tbxta*, and *egr2b* in situs, embryos were oriented laterally. For *gata2a* and *dlx1a* in situs, embryos were oriented for dorsal viewing. To quantify the staining area in *gata2a* and *dlx1a* stained embryos, color deconvolution was performed in ImageJ to isolate blue hues. Images were cropped to isolate the head and an 8-bit pixel intensity threshold of 0-230 was applied to measure stained area. Total pixel area within the threshold was quantified and normalized by dividing by the measured width of the head anterior to the eye. For *egr2b* in situs, pixel width measurements were taken at rhombomere 3 and rhombomere 5 at the ventral edge of the embryo. Notochord width was measured above the terminus of the yolk extension for *tbxta* in situs. Notochord, rhombomere 3, and rhombomere 5 width measurements were normalized to the width of the trunk at the beginning of the yolk extension to account for embryo size.

Gene ontology enrichment analysis

The top 50 most frequently occurring DEGs were selected from all genotypes in prominent cell types (>800 recovered cells). Gene ontology term enrichment analysis was performed on the gene list using metascape express analysis via the web portal (https://metascape.org)[65].

Vasculature imaging and quantification

Zebrafish embryos were raised in E3 media until reaching shield stage and then transferred to E3 media with 0.003% N-phenylthiourea (PTU, Sigma Aldrich) to block pigmentation. Embryos from *kdrl*:GFP+ incrosses were screened for fluorescence using a Zeiss fluorescent stereoscope. Confocal microscopy was performed using a Zeiss LSM700 confocal microscope on larvae anesthetized in 3-amino benzoic acid ethyl ester (Tricaine, Sigma Aldrich) in E3 at 0.2 mg/ml and embedded in 1% low-melting agarose in 35 mm glass bottom dishes (MatTek Corporation, Part No. P35G-0.170-14-C). Optical sectioning through the trunk of embryos generated Z-stacks of the developing vasculature with 3.028 μm or 5.400 μm steps between slices and centered around the developing caudal vein plexus just posterior to the yolk extension. Maximum intensity Z-projections were generated using ImageJ (NIH) software. The width of the CVP and ISV length were measured at the location of the first ISV posterior to the yolk extension in imageJ.

Optical flow analysis

To visualize blood flow in caudal vein plexus and intersomitic vessels, embryos were anesthetized in tricaine and embedded in 1% low-melting agarose. Videos capturing blood flow through the tail vasculature were recorded using a Leica Flexacam C5 mounted on a stereoscope. All videos were recorded at 30fps and were at least 10 seconds in duration. The FlowJ plugin integrated into ImageJ was used to measure optical flow. The default central slice of each video was set as the central slice and the Lucas Kanade algorithm with Gaussian derived gradient method, and Gaussian 1D regularization was used to compute optical flow. Parameters were set as follows: Σ-s = 2, Σ-t = 2, τ = 1.75, Σ-w = 1.75.

Embryo editing efficiency with *rx3* capture sequence-modified gRNAs

Zebrafish embryos at the single-cell stage were injected with 1-1.5nL of RNP containing 100 ng/μL of either capture sequence-modified gRNAs or gRNAs without the capture sequence, 2 μM EnGen Cas9 NLS (NEB, cat no. M0646M), 1x Buffer r3.1 (NEB, cat no. B7203S) and 0.1% Phenol Red. Missing eyes phenotype, characteristic of loss of the *rx3* gene, was assessed at 72 hpf and the number of embryos that showed a complete lack of eye development were tallied.

Modified gRNA qPCR

Zebrafish embryos at the single-cell stage were injected with RNP containing capture sequence-modified gRNAs as described earlier. For each sample, 10 embryos at each of the following stages (6 hpf, 12 hpf, 24 hpf, 36 hpf, 48 hpf, 60 hpf, 72 hpf, 84 hpf, 96 hpf, and 120 hpf) were collected, dechorionated and washed three times with PBS. The embryos were flash frozen in liquid nitrogen and stored in −80 °C until ready for RNA extraction. RNA was extracted using a miRNeasy kit (Qiagen, cat no. 217084) following manufacturer's instructions. Homogenization of the tissue was performed using a handheld mechanical tissue grinder (Fisher Scientific, cat no. 12-141-361). A total of 1000 ng of the extracted RNA was reverse transcribed using QuantiTect reverse transcription kit (Qiagen, cat no. 205313) per manufacturer instructions. gRNA barcode-specific primer (5′-TTGCTAGGACCGGCCTTAAAGC-3′) at 500 nM was used for reverse transcription. qRT-PCR was performed on a 7500 Fast Real-Time PCR system (Thermo Fisher Scientific). A 10 μL qRT-PCR reaction mix contained 500 nM forward primer (5′-GACACTATAGATCTGCCAGACG-3′), 500 nM reverse primer (5′-GGACTAGCCTTATTTAAACTTGCTATG-3′), 1.25 ng/μL cDNA, and 1x PowerUp SYBR Green master mix (Fisher, cat no. A25742). PCR cycling conditions were 50 °C for 2 min, 95 °C for 2 min, [95 °C for 15 s, 60 °C for 15 s, 72 °C for 1 min]x 40 cycles and a melt curve analysis at 95 °C for 15 s, 60 °C for 1 min, 95 °C for 30 s and 60 °C for 15 s.

tp63 mRNA rescue

Codon-optimized gene sequence for ΔNp63 was ordered as custom gene synthesis in the pEXP-IDT expression plasmid. The DNA template was amplified using forward primers appending an SP6 promoter site to the 5′ of the template (5′-ATTTAGGTGACACTA-TAGGAGCCACCATGCTGTACCTCGAGAC-3′) and reverse primers specific to the 3′ terminus of the mRNA sequence (5′-TCACTCACCCTCTTCTTTTATACG-3′). mRNA was transcribed and poly(A) tailed using a SP6 mMessage mMachine transcription kit (Thermo Fisher Scientific, cat # AM1340), and a poly(A) tailing kit (Thermo Fisher Scientific, cat # AM1350), following the manufacturer's protocol. 1–1.5 nL of RNP containing 100 ng/μL gRNA, 2 μM Cas9, and 300 ng/μL mRNA was injected in embryos at 1-cell stage. Phenotyping was performed as described in the "Vasculature imaging and quantification" and "Optical flow analysis" sections.

Statistics and reproducibility

Statistical analyses were performed in R. For imaging-based and in situ hybridization experiments, differences between experimental groups were assessed using one-way analysis of variance (ANOVA) followed by Dunnett's post hoc test when comparing multiple mutant conditions

to a single control, or by Tukey's post hoc test when all pairwise comparisons were of interest, such as in *tp63* mRNA rescue experiments. Two-sided unpaired t-tests were used where only two groups were compared. Sample sizes (n) represent individual embryos unless otherwise stated. For rhombomere width analyses, measurements of rhombomeres 3 and 5 were obtained from the same set of embryos but were analyzed separately. Exact test statistics, degrees of freedom, tails, confidence intervals, effect sizes, and P values for all embryo-level comparisons are reported in Supplementary Data 5. All experiments involving embryo imaging in the manuscript main figures were repeated a minimum of 2 times with similar results. The control in situ hybridization experiments in supplementary fig. 9 confirming no unexpected changes in cell abundance were not repeated. No statistical method was used to predetermine sample sizes. No data were excluded from the analyses. The experiments were not randomized. The Investigators were not blinded to allocation during experiments and outcome assessment.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The raw sequencing data and feature-barcode count matrices generated in this study have been deposited in the NCBI Gene Expression Omnibus (GEO) database under accession code GSE315445. The processed single-cell RNA sequencing data, including annotated Seurat objects and metadata, generated in this study have been deposited to Zenodo [https://doi.org/10.5281/zenodo.18602858]. Source data are provided with this paper.

Code availability

Custom code for the MIC-Drop-seq pipeline and downstream single-cell analysis is available at GitHub [https://github.com/clay-carey/MIC-Drop-seq] under an MIT license. The version of the code used to generate the results presented in this study is archived on Zenodo [https://doi.org/10.5281/zenodo.18615355]. Analyses were performed in R (v4.4.0) using Seurat (v5.3.1), edgeR (v4.2.2), monocle3 (v1.3.7), PLNmodels (v1.0.4.200), and hooke (v0.0.1). Raw scRNAseq data was processed using Cell Ranger (v5.0.0) for alignment and gene expression matrix generation. A comprehensive list of software dependencies is provided in the repository's README file.

References

1. Acevedo-Arozena, A. et al. ENU Mutagenesis, A Way Forward To Understand Gene Function. *Annu. Rev. Genom. Hum. Genet.* **9**, 49–69 (2008).
2. Hammerschmidt, M. & Nüsslein-Volhard, C. The expression of a zebrafish gene homologous to Drosophila snail suggests a conserved function in invertebrate and vertebrate gastrulation. *Development* **119**, 1107–1118 (1993).
3. Driever, W. et al. A genetic screen for mutations affecting embryogenesis in zebrafish. *Development* **123**, 37–46 (1996).
4. Haffter, P. et al. The identification of genes with unique and essential functions in the development of the zebrafish, Danio rerio. *Development* **123**, 1–36 (1996).
5. Wang, J. Y. & Doudna, J. A. CRISPR technology: a decade of genome editing is only the beginning. *Science* **379**, eadd8643 (2023).
6. Parvez, S., Brandt, Z. J. & Peterson, R. T. Large-scale F0 CRISPR screens in vivo using MIC-Drop. *Nat. Protoc.* **18**, 1841–1865 (2023).
7. Parvez, S. et al. MIC-Drop: A platform for large-scale in vivo CRISPR screens. *Science* **373**, 1146–1151 (2021).
8. Sur, A. et al. Single-cell analysis of shared signatures and transcriptional diversity during zebrafish development. *Dev. Cell* **58**, 3028–3047.e12 (2023).
9. Wagner, D. E. et al. Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. *Science* **360**, 981–987 (2018).
10. Farrell, J. A. et al. Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. *Science* **360**, eaar3131 (2018).
11. Saunders, L. M. et al. Embryo-scale reverse genetics at single-cell resolution. *Nature* **623**, 782–791 (2023).
12. Dorrity, M. W. et al. Proteostasis governs differential temperature sensitivity across embryonic cell types. *Cell* **186**, 5015–5027.e12 (2023).
13. Przybyla, L. & Gilbert, L. A. A new era in functional genomics screens. *Nat. Rev. Genet* **23**, 89–103 (2022).
14. Replogle, J. M. et al. Combinatorial single-cell CRISPR screens by direct guide RNA capture and targeted sequencing. *Nat. Biotechnol.* **38**, 954 (2020).
15. Datlinger, P. et al. Ultra-high-throughput single-cell RNA sequencing and perturbation screening with combinatorial fluidic indexing. *Nat. Methods* **18**, 635–642 (2021).
16. Dixit, A. et al. Perturb-seq: Dissecting molecular circuits with scalable single cell RNA profiling of pooled genetic screens. *Cell* **167**, 1853–1866.e17 (2016).
17. Yao, D. et al. Scalable genetic screening for regulatory circuits using compressed Perturb-seq. *Nat. Biotechnol.* 1–14 https://doi.org/10.1038/s41587-023-01964-9 (2023).
18. Sunshine, S. et al. Systematic functional interrogation of SARS-CoV-2 host factors using Perturb-seq. *Nat. Commun.* **14**, 6245 (2023).
19. Jin, X. et al. In vivo Perturb-Seq reveals neuronal and glial abnormalities associated with autism risk genes. *Science* **370**, eaaz6063 (2020).
20. Germain, P.-L., Lun, A., Meixide, C. G., Macnair, W. & Robinson, M. D. Doublet identification in single-cell sequencing data using *scDblFinder*. *F1000Res.* **10**, 979 (2022).
21. Griffin, K. J., Amacher, S. L., Kimmel, C. B. & Kimelman, D. Molecular identification of spadetail: regulation of zebrafish trunk and tail mesoderm formation by T-box genes. *Development* **125**, 3379–3388 (1998).
22. Loosli, F. et al. Loss of eyes in zebrafish caused by mutation of chokh/rx3. *EMBO Rep.* **4**, 894–899 (2003).
23. Davidson, A. J. & Zon, L. I. The caudal-related homeobox genes cdx1a and cdx4 act redundantly to regulate hox gene expression and the formation of putative hematopoietic stem cells during zebrafish embryogenesis. *Dev. Biol.* **292**, 506–518 (2006).
24. Hayward II, A. G., Joshi, P. & Skromne, I. Spatiotemporal analysis of zebrafish hox gene regulation by Cdx4. *Dev. Dyn.* **244**, 1564–1573 (2015).
25. Skuntz, S. et al. Lack of the mesodermal homeodomain protein MEOX1 disrupts sclerotome polarity and leads to a remodeling of the cranio-cervical joints of the axial skeleton. *Dev. Biol.* **332**, 383–395 (2009).
26. Kirilenko, P. et al. Transient activation of meox1 is an early component of the gene regulatory network downstream of hoxa2. *Mol. Cell Biol.* **31**, 1301–1308 (2011).
27. Mankoo, B. S. et al. The concerted action of Meox homeobox genes is required upstream of genetic pathways essential for the formation, patterning and differentiation of somites. *Development* **130**, 4655–4664 (2003).
28. Dauer, M. V. P., Currie, P. D. & Berger, J. Skeletal malformations of Meox1-deficient zebrafish resemble human Klippel–Feil syndrome. *J. Anat.* **233**, 687–695 (2018).

29. Mohamed, J. Y. et al. Mutations in MEOX1, Encoding Mesenchyme Homeobox 1, Cause Klippel-Feil Anomaly. *Am. J. Hum. Genet.* **92**, 157–161 (2013).
30. Posey, K. L., Coustry, F. & Hecht, J. T. Cartilage oligomeric matrix protein: COMPopathies and beyond. *Matrix Biol.* **71–72**, 161–173 (2018).
31. Willett, R. T. & Greene, L. A. Gata2 is required for migration and differentiation of retinorecipient neurons in the superior colliculus. *J. Neurosci.* **31**, 4444–4455 (2011).
32. Solek, C. M., Feng, S., Perin, S., Weinschutz Mendes, H. & Ekker, M. Lineage tracing of *dlx1a/2a* and *dlx5a/6a* expressing cells in the developing zebrafish brain. *Dev. Biol.* **427**, 131–147 (2017).
33. Moens, C. B., Yan, Y. L., Appel, B., Force, A. G. & Kimmel, C. B. Valentino: a zebrafish gene required for normal hindbrain segmentation. *Development* **122**, 3981–3990 (1996).
34. Moens, C. B., Cordes, S. P., Giorgianni, M. W., Barsh, G. S. & Kimmel, C. B. Equivalence in the genetic control of hindbrain segmentation in fish and mouse. *Development* **125**, 381–391 (1998).
35. Santos-Pereira, J. M., Gallardo-Fuentes, L., Neto, A., Acemel, R. D. & Tena, J. J. Pioneer and repressive functions of p63 during zebrafish embryonic ectoderm specification. *Nat. Commun.* **10**, 3049 (2019).
36. Karthik, S. et al. Synergistic interaction of sprouting and intussusceptive angiogenesis during zebrafish caudal vein plexus development. *Sci. Rep.* **8**, 9840 (2018).
37. Rouleau, M. et al. TAp63 is important for cardiac differentiation of embryonic stem cells and heart development. *Stem Cells* **29**, 1672–1683 (2011).
38. Liang, J. et al. In-organoid single-cell CRISPR screening reveals determinants of hepatocyte differentiation and maturation. *Genome Biol.* **24**, 251 (2023).
39. Li, C. et al. Single-cell brain organoid screening identifies developmental defects in autism. *Nature* **621**, 373–380 (2023).
40. Datlinger, P. et al. Pooled CRISPR screening with single-cell transcriptome readout. *Nat. Methods* **14**, 297–301 (2017).
41. Zheng, X. et al. Massively parallel in vivo Perturb-seq reveals cell-type-specific transcriptional networks in cortical development. *Cell* **187**, 3236–3248.e21 (2024).
42. Jia, S. et al. Multiple developmental defects in sox11a mutant zebrafish with features of coffin-siris syndrome. *Int. J. Biol. Sci.* **16**, 3039 (2020).
43. Tsurusaki, Y. et al. De novo SOX11 mutations cause Coffin–Siris syndrome. *Nat. Commun.* **5**, 4011 (2014).
44. Takeuchi, S. et al. Pinpoint modification strategy for stabilization of single guide RNA. *J. Chromatogr. B* **1192**, 123149 (2022).
45. Rosenberg, A. B. et al. Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science* **360**, 176–182 (2018).
46. Cao, J. et al. Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science* **357**, 661–667 (2017).
47. Bartosovic, M., Kabbe, M. & Castelo-Branco, G. Single-cell CUT&Tag profiles histone modifications and transcription factors in complex tissues. *Nat. Biotechnol.* **39**, 825–835 (2021).
48. Petri, K. et al. CRISPR prime editing with ribonucleoprotein complexes in zebrafish and primary human cells. *Nat. Biotechnol.* **40**, 189–193 (2022).
49. Fleck, J. S., Camp, J. G. & Treutlein, B. What is a cell type? *Science* **381**, 733–734 (2023).
50. Rood, J. E., Hupalowska, A. & Regev, A. Toward a foundation model of causal cell and tissue biology with a Perturbation Cell and Tissue Atlas. *Cell* **187**, 4520–4545 (2024).
51. Trapnell, C. Revealing gene function with statistical inference at single-cell resolution. *Nat. Rev. Genet* **25**, 623–638 (2024).
52. Labun, K. et al. CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Res.* **47**, W171–W174 (2019).
53. Gagnon, J. A. et al. Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. *PLoS One* **9**, e98186 (2014).
54. Manoli, M. & Driever, W. Fluorescence-activated cell sorting (FACS) of fluorescently tagged cells from zebrafish larvae for RNA isolation. *Cold Spring Harb. Protoc.* **2012**, pdb.prot069633 (2012).
55. Lawson, N. D. et al. An improved zebrafish transcriptome annotation for sensitive and comprehensive detection of cell type-specific genes. *eLife* **9**, e55792 (2020).
56. Stuart, T. et al. Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902.e21 (2019).
57. Labun, K. et al. Accurate analysis of genuine CRISPR editing events with ampliCan. *Genome Res* **29**, 843–847 (2019).
58. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
59. Duran, M. et al. A statistical framework for inferring genetic requirements from embryo-scale single-cell sequencing experiments. 2025.04.03.646654 Preprint at https://doi.org/10.1101/2025.04.03.646654 (2025).
60. Cao, J. et al. The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**, 496–502 (2019).
61. Qiu, X. et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat. Methods* **14**, 979–982 (2017).
62. Chiquet, J., Mariadassou, M. & Robin, S. The Poisson-Lognormal Model as a Versatile Framework for the Joint Analysis of Species Abundances. *Front. Ecol. Evol.* **9**, 588292 (2021).
63. Wagner, D. E. & Klein, A. M. Lineage tracing meets single-cell omics: opportunities and challenges. *Nat. Rev. Genet* **21**, 410–427 (2020).
64. Thisse, C. & Thisse, B. High-resolution in situ hybridization to whole-mount zebrafish embryos. *Nat. Protoc.* **3**, 59–69 (2008).
65. Zhou, Y. et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **10**, 1523 (2019).

Acknowledgements

We thank all members of the Gagnon, Parvez, and Peterson labs for discussions and comments. We thank Opal Allen, Jeffrey Farrell, Austin Green, Nathan Lord, Aaron McKenna, Tajeshree Prakash, Abhinav Sur, Lisandra Vila Ellis, Yogesh Goyal, and Michael Werner for feedback on the manuscript, protocols, reagents, and equipment. We thank CZAR and CBRZ staff for zebrafish care, as well as the University of Utah Cell Imaging and High-Throughput Genomics core facilities. This work was supported by the National Institutes of Health [grants F32HL156644 (C.M.C.), K01HG013682 (C.M.C.), R00HG012593 (S.P.), R01GM134069 (R.T.P.), R35GM142950 (J.A.G.), and R24OD035409 (R.T.P., J.A.G.)]; Northwestern University Startup Funds (S.P.); and the One Utah Data Science Hub (J.A.G.).

Author contributions

S.P., C.M.C., and Z.J.B. led the conceptualization, methodology, visualization, and writing of the original draft. These three authors, along with B.W.B. and C.J.Y., conducted the investigation. R.T.P. and J.A.G. provided supervision, project administration and contributed to the conceptualization. Funding acquisition was a joint effort by R.T.P., J.A.G., S.P., C.M.C., and Z.J.B.; R.T.P., J.A.G., C.M.C., S.P., Z.J.B., and B.W.B. reviewed and edited the manuscript.

Competing interests

S.P., R.T.P. have submitted a patent application to the USPTO pertaining to the core MIC-Drop technology used in this work (application number US18/568,690). The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-026-70989-w.

Correspondence and requests for materials should be addressed to Randall T. Peterson or James A. Gagnon.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.