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

Authors: Clayton M. Carey[1]†, Saba Parvez[2,3]†, Zachary J. Brandt[2], Brent W. Bisgrove[1,2], Christopher J. Yates[2], Randall T. Peterson[2]*, James A. Gagnon[1,4]*

Affiliations:

[1] School of Biological Sciences, University of Utah; Salt Lake City, UT, 84112, USA.
[2] Department of Pharmacology and Toxicology, College of Pharmacy, University of Utah; Salt Lake City, UT, 84112, USA.
[3] Department of Cell & Developmental Biology, Feinberg School of Medicine, Northwestern University; Chicago, IL, 60611, USA.
[4] Henry Eyring Center for Cell & Genome Science, University of Utah; Salt Lake City, UT, 84112, USA.

† Equal contribution
* Corresponding authors: randall.peterson@pharm.utah.edu, james.gagnon@utah.edu

Abstract: Advances in genome engineering and single-cell RNA sequencing (scRNAseq) have revolutionized the ability to precisely map gene functions, yet scaling these techniques for large-scale genetic screens in animals remains challenging. We combined high-throughput gene disruption in zebrafish embryos via Multiplexed Intermixed CRISPR Droplets with phenotyping by multiplexed scRNAseq (MIC-Drop-seq). In one MIC-Drop-seq experiment, we intermixed and injected droplets targeting 50 transcriptional regulators into 1,000 zebrafish embryos, followed by pooled scRNAseq. Tissue-specific gene expression and cell abundance analysis of demultiplexed mutant cells recapitulated many known phenotypes, while also uncovering novel functions in brain and mesoderm development. We observed pervasive cell-extrinsic effects among these phenotypes, highlighting how whole-embryo sequencing captures complex developmental interactions. Thus, MIC-Drop-seq provides a powerful and scalable platform for mapping gene functions in vertebrate development with cellular resolution.

Main Text:

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 up 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 DNA barcode to facilitate rapid genotyping. We previously employed MIC-Drop to rapidly identify new 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 high-resolution scRNAseq-based comparisons of mutant zebrafish embryos revealed previously unappreciated developmental phenotypes, even for well-characterized transcriptional regulators (*11*). Scaling these approaches for screening large numbers of mutants, however, remains logistically challenging and cost-prohibitive due to the individual engineering, housing, dissociation and separate sequencing of animals with different genotypes.

Approaches that employ multiplexed perturbation and pooled scRNAseq offer a path forward. Such methods, including Perturb-seq and CROP-seq, use indexing of CRISPR gRNAs from pooled cells to assign genotypes across hundreds of gene knockouts in a single experiment (*12–14*). These tools have allowed the identification of unique cell states and changes in gene regulatory networks caused by perturbing genes in cultured cells (*15–17*). 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 (*18*). 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 tissue. A whole-animal Perturb-seq approach, where each individual receives a single perturbation, 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 in 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 transcription factors 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, accelerating the mapping of genotypes to phenotypes across dozens of genes in parallel.

MIC-Drop-seq enables multiplexed single-cell phenotyping of mutant zebrafish embryos

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*) (**Fig. S1A)**. We validated that these gRNAs persist for days after injection (**Fig. S1B**) and were highly active for CRISPR mutagenesis in zebrafish embryos (**Fig. S1C**). This approach enables the unified capture of both mRNA and gRNA from each cell in pools derived from zebrafish embryos with mixed genotypes. MIC-Drop-seq therefore enables the multiplexed generation and single-cell transcriptomic profiling of many whole-animal mutants in a single experiment.

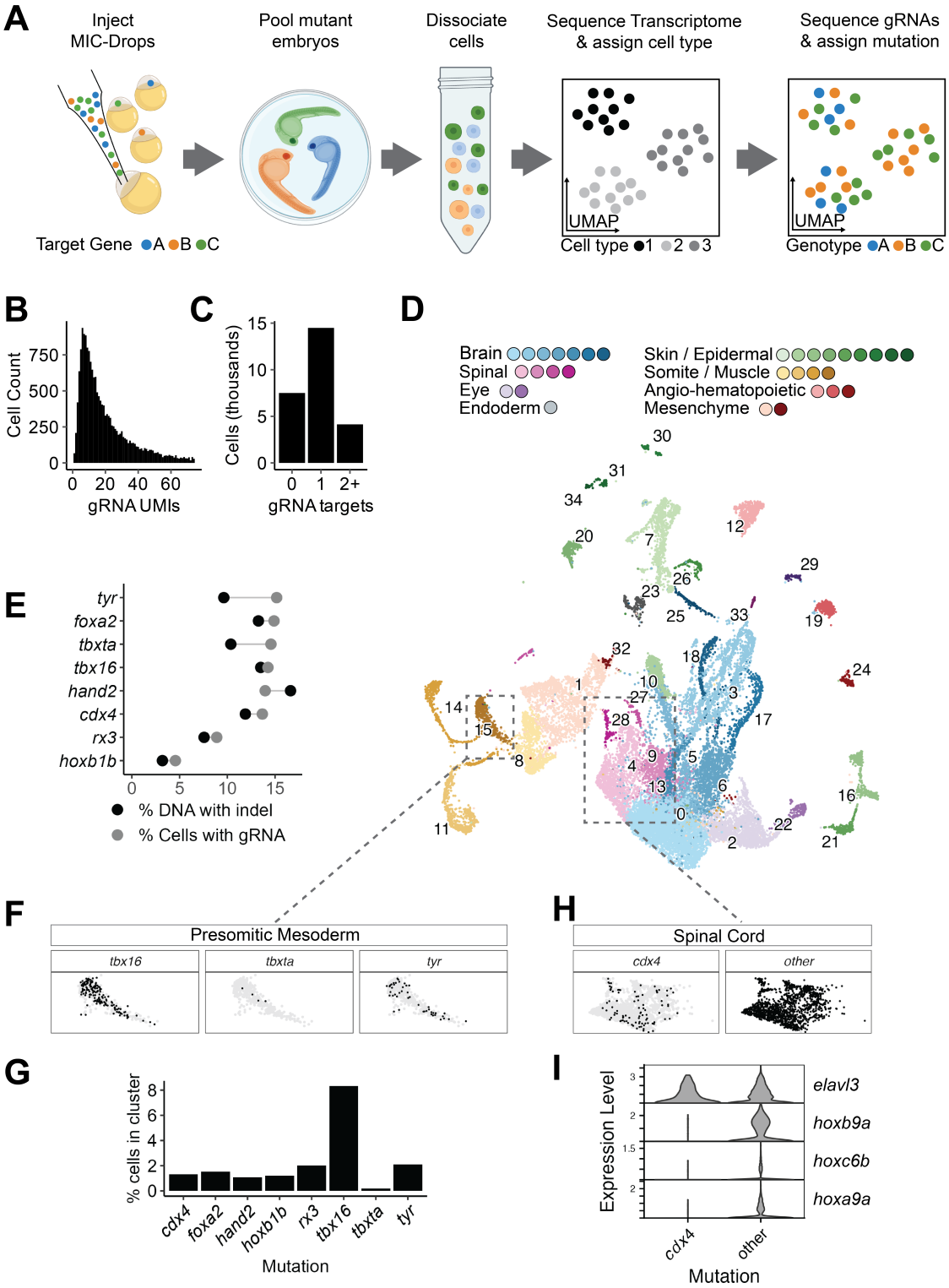
As a proof-of-concept, we performed a small MIC-Drop-seq screen targeting seven transcription factors with well-characterized mutant phenotypes (*tbx16, tbxta, cdx4, foxa2, hoxb1b, rx3,* and *hand2*), as well as a non-transcription factor control gene (*tyr*). A set of droplets was created for each gene 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 (**Table S1**). 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 (**Fig. S2A**), and were more likely to be assigned as doublets (**Fig. S2B**)(*19*). We therefore considered these cells to be technical doublets and removed them from the dataset. The mutant genotype of the remaining 21,994 cells were inferred based on the sequence of the captured gRNAs, recovering a range of 661 to 2,198 cells per targeted gene (**Fig. S2A,B**). Dimensionality reduction and clustering yielded 35 cell clusters (**Fig. 1D**), which were annotated by label transfer from the Daniocell reference atlas (*8*) (**Fig. S4A-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 (**Table S2**).

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. For each gene, we calculated the frequency of indels at the target cut sites and in experimental and wild type cells. We then compared editing frequency 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**), 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 (*20*). Conversely, mutations in *tbxta,* result in the failure to specify presomitic mesoderm and do not make posterior somites (*2*). We computed the proportion of cells in the presomitic mesoderm cluster for each genotype. Consistent with the known 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 eyeless phenotype (*21*). Consistent with this phenotype, we observed a depletion of *rx3* mutant cells among all retinal cell types compared to *tyr* mutant cells (**Fig. S5A,B**). Therefore, MIC-Drop-seq is capable of detecting phenotypes that enrich or deplete particular cell types.

Some mutant phenotypes manifest as altered gene expression without changing the abundance of a given 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 (*22*, *23*). We identified *cdx4* mutant cells in the spinal cord and compared their gene expression profile to all other cells in these clusters (**Fig. 1H**). Expression of *hoxb9a, hoxc6b,* and *hoxa9a* was significantly downregulated among *cdx4* mutant spinal cord cells (**Fig. 1I**). Taken together, these experiments demonstrate that MIC-Drop-seq is a reliable and sensitive method for multiplexed mutant phenotyping with cellular resolution.


A
Inject MIC-Drops
Pool mutant embryos
Dissociate cells
Sequence Transcriptome & assign cell type
Sequence gRNAs & assign mutation
Target Gene A B C
UMAP
Cell type 1 2 3
UMAP
Genotype A B C
B
Cell Count
750
500
250
0
0 20 40 60
gRNA UMIs
C
Cells (thousands)
15
10
5
0
0 1 2+
gRNA targets
D
Brain
Spinal
Eye
Endoderm
Skin / Epidermal
Somite / Muscle
Angio-hematopoietic
Mesenchyme
E
tyr
foxa2
tbxta
tbx16
hand2
cdx4
rx3
hoxb1b
0 5 10 15
% DNA with indel
% Cells with gRNA
F
Presomitic Mesoderm
tbx16
tbxta
tyr
H
Spinal Cord
cdx4
other
G
% cells in cluster
8
6
4
2
0
cdx4 foxa2 hand2 hoxb1b rx3 tbx16 tbxta tyr
Mutation
I
Expression Level
elavl3
hoxb9a
hoxc6b
hoxa9a
cdx4 other
Mutation

165

Fig. 1: MIC-Drop-seq enables multiplexed single-cell phenotyping of mutant zebrafish embryos.

(A) An overview of MIC-Drop-seq. (B) Histogram of unique guide RNA molecule counts sequenced in each cell 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 a MIC-Drop-seq experiment targeting eight genes. (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 each locus (grey) (F,H). Subsets of UMAP embeddings that highlight cells with the predicted genotype indicated (black) vs other cells (grey) in the (F) presomitic mesoderm and (H) spinal cord cluster(s). (G) Quantification of the percent of cells with the indicated genotype in the presomitic mesoderm cluster. (I) Gene expression violin plot comparing 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.

A 50 gene MIC-Drop-seq screen for early zebrafish development phenotypes

We next applied MIC-Drop-seq in a large-scale screen in an effort to discover novel developmental phenotypes. We selected a set of 50 transcriptional regulator genes that are variable in expression across cell types in the zebrafish embryo and highly expressed in at least one cell type (**Fig. S6A,** 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 (**Fig. S6B**).

A MIC-Drop library containing droplets targeting these 50 genes for mutagenesis was generated and injected into 1,000 zebrafish embryos (**Table S3**). 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. After integration, filtering, and clustering, we recovered 226,492 cells distributed across 74 cell clusters (**Fig. 2B, Table S4**). Cell types were assigned with reference-based label transfer and broadly classified as either Neural, Mesoderm/Endoderm, or Non-neural ectoderm (**Fig. 2B, Table S4**). 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 (**Fig. S6C**).

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 (**Fig. S7A**). Some DEGs appeared repeatedly across all perturbed cell types and mutant genotypes (**Fig. S7B**). 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 (**Fig. S7C**).

Mutations in developmental regulators can alter cell type differentiation trajectories and change the abundance of different tissues. 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 type 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 (**Fig. S8**).

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**).

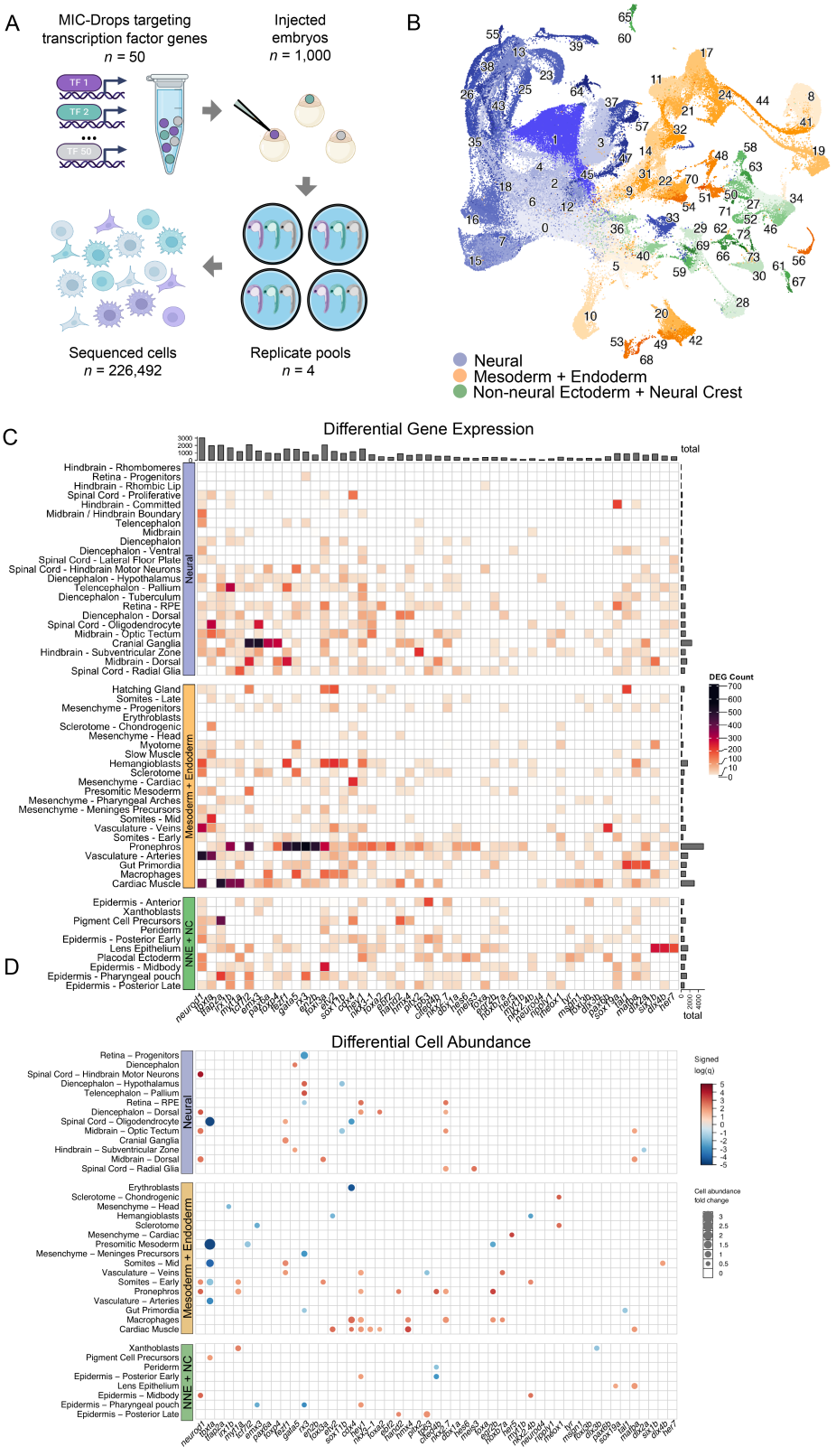
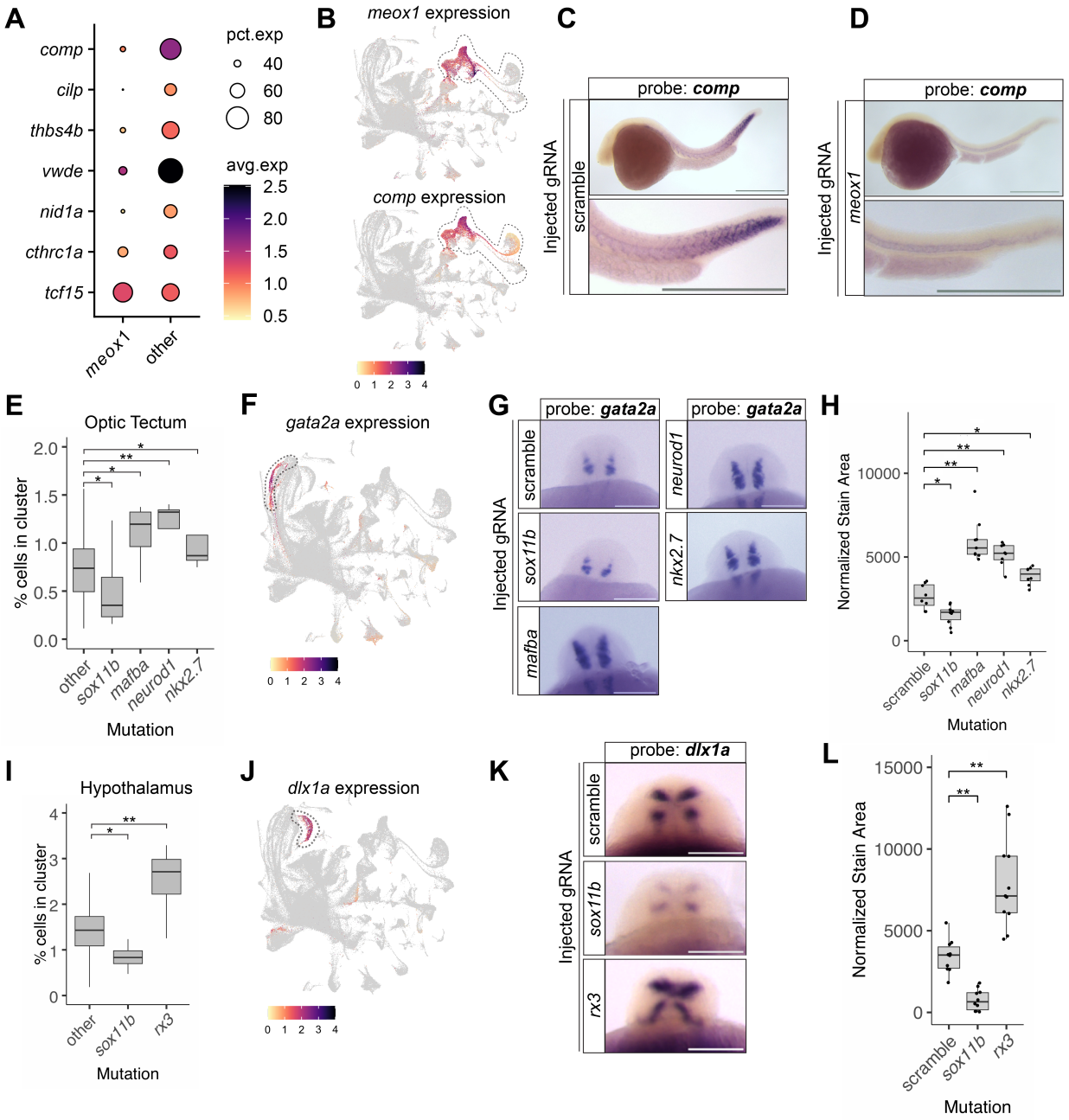

A
MIC-Drops targeting transcription factor genes
n = 50
TF 1
TF 2
TF 50
Injected embryos
n = 1,000
Sequenced cells
n = 226,492
Replicate pools
n = 4
B
Neural
Mesoderm + Endoderm
Non-neural Ectoderm + Neural Crest
C
Differential Gene Expression
total
Neural
Mesoderm + Endoderm
NNE + NC
DEG Count
D
Differential Cell Abundance
Signed log(q)
Cell abundance fold change

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 1,000 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. (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 cells in the cluster. Total DEG counts (up or downregulated) are shown for each cell type (row) and perturbation (column) (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 monocole3 (*24*) (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 log2FC of > 0.5 are shown. Cell clusters with no detected cell abundance changes in any mutant were omitted.

MIC-Drop-seq enables discovery of transcriptional and cellular phenotypes

The molecular and cellular phenotypes generated from MIC-Drop-seq serve as testable hypotheses for experimental validation, enabling discovery of new biological functions of the targeted genes. Using orthogonal methods, we validated several novel 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. We injected embryos with Cas9/gRNA complexes targeting *meox1* and examined *comp* expression with in situ hybridization. In contrast to control embryos, *meox1* gRNA-injected embryos showed a loss of *comp* expression in myosepta and tailbud mesoderm, while retaining *comp* expression in the notochord (**Fig. 3C**). *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 in 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, two cell types 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 (**Fig. 3E**). We identified *gata2a* as a specific marker for optic tectum, similar to the orthologous Gata2, which marks photorecipient neurons in the mammalian tectum (*31*) (**Fig. 3F**). In situ hybridization of *gata2a* confirmed the expansion of the optic tectum in the brains of the *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 hypothalamus cell cluster, where eyeless *rx3* mutant cells are enriched and *sox11b* cells are depleted (**Fig. 3I**). In situ hybridization of *dlx1a*, a marker of the hypothalamus (*32*) (**Fig. 3J**), in *sox11b* gRNA-injected embryos confirmed a depletion of hypothalamic cells compared to control injected embryos, while *rx3* gRNA-injected embryos showed a marked increase (**Fig. 3K, L**). Critically, none of the affected mutants showed changes in expression of marker genes *gata2a* and *dlx1a* on a per-cell basis (**Fig. S9).** Together, these data validate several transcriptional and cellular phenotypes elevated by MIC-Drop-seq, demonstrating its use to discover previously uncharacterized regulators of tissue development from large-scale pooled datasets.


A
comp
cilp
thbs4b
vwde
nid1a
cthrc1a
tcf15
meox1
other
pct.exp
40
60
80
avg.exp
2.5
2.0
1.5
1.0
0.5
B
meox1 expression
comp expression
0 1 2 3 4
C
probe: comp
Injected gRNA
scramble
D
probe: comp
Injected gRNA
meox1
E
Optic Tectum
% cells in cluster
2.0
1.5
1.0
0.5
0.0
other
sox11b
mafba
neurod1
nkx2.7
Mutation
F
gata2a expression
0 1 2 3 4
G
probe: gata2a
probe: gata2a
Injected gRNA
scramble
sox11b
mafba
neurod1
nkx2.7
H
Normalized Stain Area
10000
5000
0
scramble
sox11b
mafba
neurod1
nkx2.7
Mutation
I
Hypothalamus
% cells in cluster
4
3
2
1
0
other
sox11b
rx3
Mutation
J
dlx1a expression
0 1 2 3 4
K
probe: dlx1a
Injected gRNA
scramble
sox11b
rx3
L
Normalized Stain Area
15000
10000
5000
0
scramble
sox11b
rx3
Mutation

294
295

296 **Fig. 3: MIC-Drop-seq enables discovery of transcriptional and cellular phenotypes.**
297 (A) Gene expression bubble plot comparing expression of a set of secreted extracellular matrix
298 factors in *meox1* mutant vs other cells in the somitic mesoderm (myotome, sclerotome,
299 presomitic mesoderm, and early somite clusters). *tcf15* is a marker of somitic mesoderm and is
300 unchanged in *meox1* mutant cells. Percent of cells expressing the indicated gene is indicated by
301 bubble size and average expression level per cell is indicated with color scale. (B) Gene
302 expression feature plots for *meox1* and *comp*. Somitic mesoderm clusters are highlighted inside
303 the dashed line. (C-D) Colorimetric RNA in situ hybridization with DIG-labeled *comp* probe in (C)
304 wild type and (D) *meox1* mutant 24 hpf embryos. Top: whole embryo. Bottom: zoomed in view
305 of embryo tail. Scale bars = 400 μm. (E) Quantification of percentage of total cells of the
306 indicated mutant in the optic tectum cluster. (F) Gene expression feature plot of *gata2a*
307 expression in the single-cell dataset; optic tectum cluster is encircled (dashed line). (G)
308 Representative dorsal-view images of colorimetric RNA in situ hybridization using *gata2a* probe
309 are shown for the indicated genotype. Scale bars = 100 μm. (H) Plot quantifying *gata2a* stain
310 area for the indicated mutants. (I) Quantification of percentage of total cells of the indicated
311 mutant in the hypothalamus cluster. (J) Gene expression feature plot of *dlx1a* expression in the
312 single-cell dataset, hypothalamus cluster is encircled (dashed line). (K) Representative dorsal-
313 view images of colorimetric RNA in situ hybridization using *gata2a* probe are shown for the
314 indicated genotype. Scale bars = 400 μm. (L) Plot quantifying *dlx1a* stain area for the indicated
315 mutants. Significance testing in (H) and (L) was performed using unpaired t-tests comparing
316 each mutant to scramble injected embryos. ** indicates $p < 0.0005$, * $p < 0.005$. Significance
317 testing in (E) and (I) was performed using Hooke (part of the Monocle 3 package (*24*)). *
318 indicates a q value of < 0.05, and ** indicates a q value of < 0.005.

319

MIC-Drop-seq reveals pervasive cell-extrinsic effects of transcription factor disruption

Cells in developing embryos can influence the behaviors of other cell types, often via direct contact or through secreted signals. These cell-extrinsic effects are often termed non-cell autonomous or indirect and can be regulated by the cell-intrinsic activities of transcription factors. These emergent phenotypes can be absent in simplified cell culture systems and are difficult to predict in complex whole-animal systems. We reasoned that the large number of profiled genotypes in the MIC-Drop-seq dataset represents an avenue for unbiased identification and classification of these cell-extrinsic phenotypes. We define perturbation phenotypes as cell-intrinsic when the target transcription factor is expressed in the affected cell type or lineage-extrinsic when expressed in progenitors, while effects on cell types without current or prior expression of the perturbed gene are considered cell-extrinsic (**Fig. 4A**). To classify each transcriptional phenotype as cell-intrinsic or cell-extrinsic, we integrated the MIC-Drop-seq dataset with the Daniocell dataset, a time-course scRNAseq atlas of zebrafish embryogenesis (*8*). We constructed a simplified tree of cell type differentiation, using prior studies (*9*, *10*), to establish the identities and trajectories from early blastula to cell types present at 24 hpf (**Fig. 4B**). This representation allows quantification of the expression of each transcription factor throughout the differentiation history of each cell type (**Fig. S10**).

We next quantified the cell-intrinsic and cell-extrinsic impacts of each transcription factor disruption using lineage and target gene expression relationships. We set a simple threshold of > 10 DEGs to define a transcriptional phenotype and computed the scaled expression of each target gene across each cell type. Phenotypes were classified as cell-intrinsic when the affected cell types or their precursors expressed the target gene or were otherwise classified as cell-extrinsic (**Fig. 4C**). Overall, we found that cell-extrinsic effects are pervasive, representing the most commonly observed phenotype class among all affected cell types (**Fig. 4D**). Despite the high frequency of cell-extrinsic effects, phenotypes were more likely to be observed when the target gene is expressed in the affected cell or its lineage (**Fig. 4E**), and these cell-intrinsic phenotypes had a greater average number of DEGs (**Fig. 4F**).

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 (*33*), caused a decrease in venous cell abundance (**Fig. S11A**) and altered gene expression in all vascular cell types, indicating atypical vascular development (**Fig. S11B**). These phenotypes are unexpected, given that *tp63* is not expressed in vascular cells but is instead restricted to the epidermis (**Fig. S11C**). 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 scrambled 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 displayed decreased CVP width (**Fig. 4F**). Notably, length of intersegmental vessels (ISVs) remained unchanged (**Fig. S11D**). Optical flow analysis revealed defects in erythrocyte transit in *tp63* mutants (**Fig. S11E**), including a significant decrease in the number of ISVs with detectable flow (**Fig. 4I**), and overall decreased flow area (**Fig. S11F**). Intussusceptive angiogenesis in the CVP is linked to shear force supplied from cardiac output (*34*), and p63 null mice have defects in heart formation (*35*). While *tp63* injected zebrafish embryos had defects in skin formation (**Fig. S11G**), no changes in heart rate were apparent (**Fig S11H**). We hypothesize that these defects in early vasculogenesis could be the result of altered physical forces stemming from 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.

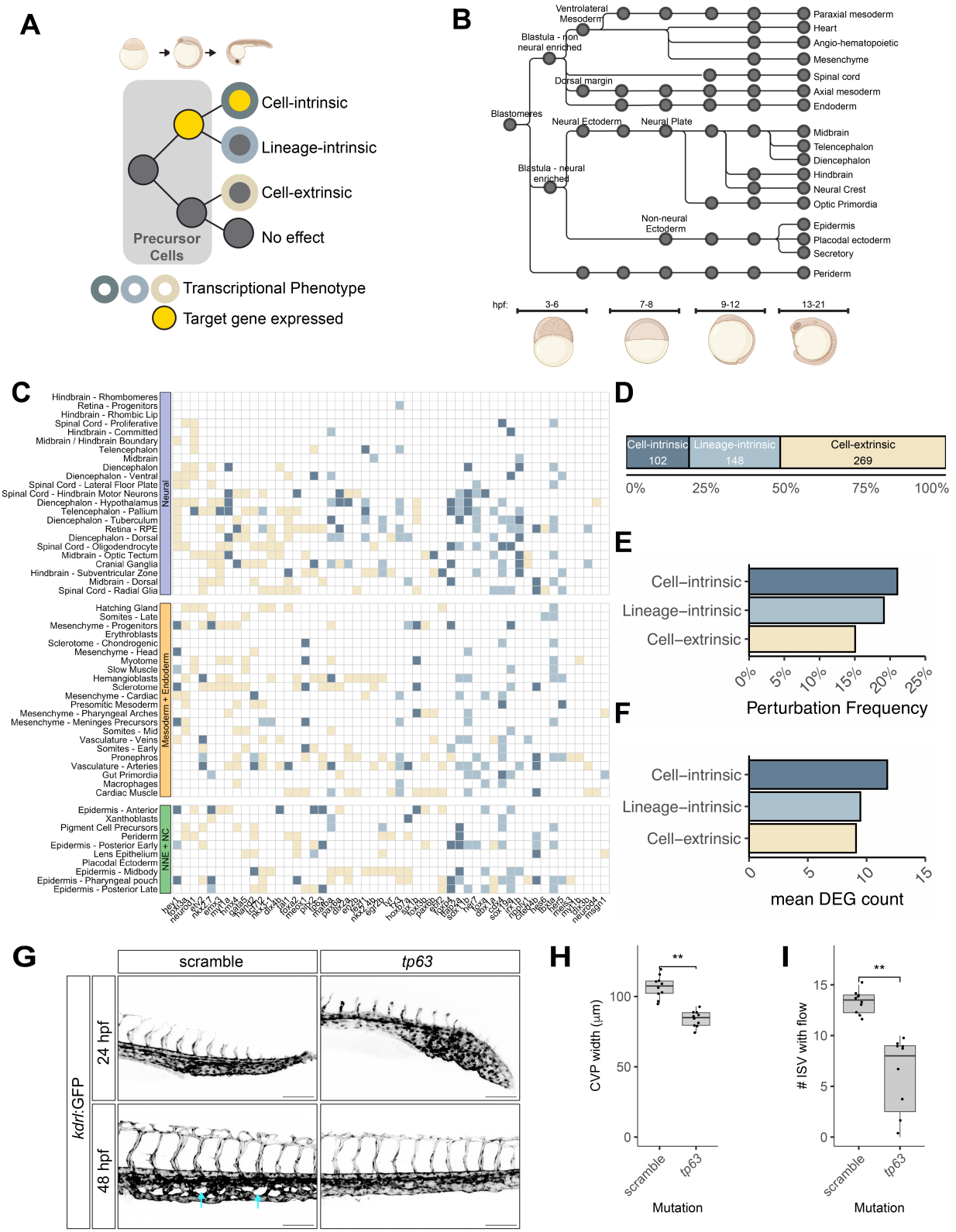

A
Cell-intrinsic
Lineage-intrinsic
Cell-extrinsic
No effect
Precursor Cells
Transcriptional Phenotype
Target gene expressed
B
Ventrolateral Mesoderm
Blastula - non neural enriched
Dorsal margin
Blastomeres
Neural Ectoderm
Neural Plate
Blastula - neural enriched
Non-neural Ectoderm
Paraxial mesoderm
Heart
Angio-hematopoietic
Mesenchyme
Spinal cord
Axial mesoderm
Endoderm
Midbrain
Telencephalon
Diencephalon
Hindbrain
Neural Crest
Optic Primordia
Epidermis
Placodal ectoderm
Secretory
Periderm
hpf:
3-6
7-8
9-12
13-21
C
Neural
Mesoderm + Endoderm
NNE + NC
D
Cell-intrinsic 102
Lineage-intrinsic 148
Cell-extrinsic 269
0%
25%
50%
75%
100%
E
Cell–intrinsic
Lineage–intrinsic
Cell–extrinsic
0%
5%
10%
15%
20%
25%
Perturbation Frequency
F
Cell–intrinsic
Lineage–intrinsic
Cell–extrinsic
0
5
10
15
mean DEG count
G
scramble
tp63
kdrl:GFP
24 hpf
48 hpf
H
CVP width (μm)
100
50
0
**
scramble
tp63
Mutation
I
ISV with flow
15
10
5
0
**
scramble
tp63
Mutation

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) Schematic of zebrafish cell 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 (see Methods for details). (D) Proportion plot of each perturbation type within the dataset. Totals for each perturbation type are indicated. (E) Quantification of perturbation frequency for each category. For each target, expression history was determined for every cell type (cell, lineage, or indirect). The frequency of cell type perturbation (>10 DEGs) for each category is plotted from all perturbations. (F) The mean number of DEGs for perturbed cell types (>10 DEGs) for each category is plotted. (G) Representative greyscale fluorescence images of CVP and posterior trunk of *kdrl*:GFP labeled transgenic embryos injected with scrambled or *tp63* targeting gRNA at 24 and 48 hpf. Blue arrows highlight CVP fenestrations in control embryos. Scale bar = 100 μm. (H) Measured CVP pixel width at 48 hpf in injected embryos. (I) Quantification of ISVs with detectable erythrocyte flow.

401 **Discussion:**
402
403 Profiling mutant phenotypes with scRNAseq provides a way to comprehensively map the link
404 between genotype and the networks of genes, cells, tissues, and organs underlying animal
405 development. In this report, we present MIC-Drop-seq as a technique for pooled CRISPR
406 screening in whole zebrafish embryos with cellular resolution phenotyping across all embryonic
407 cell types. MIC-Drop-seq is related to existing multiplexed scRNA-seq technologies like CROP-
408 seq and Perturb-seq, but offers several differentiating features that enable its use in whole
409 animals. CROP-seq and Perturb-seq similarly rely on gRNA sequencing to index mutant
410 genotypes, but have been limited to screens of cultured cells, limiting the scope of observable
411 phenotypes to a subset of molecular networks active in a handful of cell types (*15*, *36–38*). *In-*
412 *vivo* Perturb-seq uses transduction of CRISPR-containing viral particles to create clones with
413 differing mutant genotypes within individual animals (*18*, *39*). This approach, while useful for
414 studying tissue-specific regulatory networks, fails to capture earlier developmental phenotypes
415 of the targeted genes or to isolate gene-specific cell-extrinsic effects. MIC-Drop-seq uses gene
416 disruption at the zygotic stage, where mutations are rapidly generated and inherited by early
417 progenitors during development, enabling whole-animal developmental phenotypes to be
418 profiled with cellular resolution at scale. Furthermore, MIC-Drop-seq uses injection of purified
419 Cas9 RNP complexes instead of viral vectors, dramatically reducing the possibility of
420 genotyping errors from ambient gRNA or cells infected with multiple viral vectors. These
421 features make MIC-Drop-seq a highly efficient platform that could be scaled to study hundreds
422 to thousands of whole-animal mutants in parallel.
423
424 The cellular-level comparisons made possible by MIC-Drop-seq enable detection of phenotypes
425 that would otherwise be difficult to detect by conventional observation methods or would not
426 manifest *in vitro*. For example, we identified and validated widespread cell abundance changes
427 in the developing brain of *sox11b* mutant embryos, which appear visually indistinguishable from
428 wildtype siblings and have been classified as developmentally normal in previous studies (*40*,
429 *41*). Furthermore, our transcriptomic profiling revealed extensive changes in the makeup of the
430 myosepta extracellular matrix in *meox1* mutants, despite only having a visible phenotype at
431 much later stages of development. While recent studies have also leveraged parallel scRNAseq
432 of multiple mutant embryos to detect such cryptic developmental phenotypes (*42*), MIC-Drop-
433 seq drastically increases throughput by enabling rapid delivery and demultiplexing of
434 perturbations. Indeed, our approach enabled the simultaneous collection of phenotypic
435 information for 50 mutants across 74 embryonic cell types in a single experiment carried out in a
436 single day. This unbiased approach revealed pervasive cell-extrinsic phenotypes, including an
437 unexpected emergent phenotype in the vasculature of embryos with disrupted *tp63*, a
438 transcription factor that is expressed only in ectoderm in the 24 hpf embryo.
439
440 Our MIC-Drop-seq experiments focused on the 24 hpf zebrafish embryo, but further
441 modifications might enable a broader set of applications. MIC-Drop-seq is in principle
442 compatible with any species where embryos are amenable to injection at the zygotic stage.
443 While we demonstrate efficient gRNA sequence recovery at 24 hpf zebrafish embryos, studying
444 later developmental time points will likely result in lower recovery as the gRNAs are diluted and
445 degraded. DNA barcodes or chemically-stabilized gRNAs could enhance genotype recovery at
446 later stages (*43*). Supplementary oligonucleotide-based indexing of individual embryos could
447 further aid in profiling more advanced developmental stages through genotype imputation from
448 cells where gRNAs are not recovered (*42*). Another challenge of whole-animal scRNAseq as a
449 phenotyping method is the detection of transcriptional effects in low-abundance cell types. More
450 statistical power could be acquired by leveraging emerging scRNAseq technologies to
451 sequence far more cells across more replicates (*44*, *45*). Alternatively, cell sorting could enrich
452 rare cell types of interest for focused screens.

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 (*46*) 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 (*47*), 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 (*48–50*). 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.

Materials and Methods:

Guide RNA design for feature barcoding

Guide RNAs (gRNAs) targeting the genes-of-interest were designed using CHOPCHOP as previously described (*6*, *7*, *51*). 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 less mismatches, In case when 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 **Supplemental Table S1 and S3.**

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'AAAAAAAGCACCGACTCGGTGCCACTTGGCCTTGCTAGGACCGGCCTTAAAGCGGCCA AGTTGATAACGGACTAGCCTTATTTAAACTTGCTATGCTGTTTCCAGCTTAGCTCTTAAAC-3') was ordered from IDT as an ultramer. To generate the DNA template for IVT, a fill-in PCR was performed as described previously. Briefly, a 50 µL PCR 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

500 DNase I was added, and the sample incubated for an additional 20 minutes. RNA cleanup was
501 performed using a ZR96 RNA Clean and Concentrator-5 (Zymo Research; cat. no. R1080) kit
502 per manufacturer's instruction. RNA concentration was determined using a Nanodrop and RNA
503 integrity verified using gel electrophoresis.
504

505 **MIC-Drop library design and injection**

506 3% FS-HFE was prepared by dissolving 1 g of 008-fluorosurfactant (Ran Biotechnologies, cat.
507 no. 008-Fluorosurfactant) in 33.3 mL of HFE-7500 (3M, cat. no. Novec 7500). A 21.5 µL
508 ribonucleoprotein (RNP) complex targeting each gene containing 200 ng/µL gRNAs, 3.36 µM
509 EnGen Cas9 (NEB, cat. no. M0646M) and 1x NEB buffer r3.1 (NEB, cat. no. B6003S) was
510 assembled and incubated at room temperature for 5-10 min. 3.5 µL of 0.5% phenol red dye
511 (Sigma-Aldrich, cat. no. P0290-100 mL) was added to each sample and mixed. The droplets
512 containing RNP were generated using a BioRad QX200 droplet generator (BioRad, cat. no.
513 1864002) as previously described (*6*, *7*). Briefly, 20 µL of RNP mix was transferred to wells
514 labeled 'sample' of a QX 200 cartridge (BioRad, cat. no. 1864007). 70 µL of 3% FS-HFE was
515 added to the wells labeled 'oil'. The cartridge was sealed using the provided rubber gasket and
516 loaded in the droplet generator. Droplet generation starts as soon as the lid is closed. Droplets
517 for 8 samples were generated simultaneously from a single cartridge. The droplets were
518 carefully transferred to PCR strip tubes containing 50 µL 3% FS-HFE. To generate the MIC-
519 Drop library, equal volume of droplets for each sample were pooled into a separate PCR tube
520 containing 50 µL 3% FS-HFE. The droplets were gently mixed to get an even distribution and
521 stored at 4°C. 3-5 µL of the intermixed droplets were transferred to a microinjection needle
522 (World Precision Instruments, cat. no. TW100F-3) using a microloader tip (Eppendorf, cat. no.
523 5242956003). A single droplet per embryo was injected as previously described (*6*, *7*). Injected
524 embryos were transferred to a petri dish (30-50 embryos/petri dish) in E3 medium containing
525 methylene blue and incubated at 28.5 °C.

526 **Embryo dissociation and single-cell library preparation**

527 Dead and severely deformed embryos were removed at 24 hpf. For the pilot experiment, 40
528 total injected embryos were pooled. Dissociation of embryos was performed similarly to
529 previously published protocols (*52*). 1 mL of 1X FACSmax (Fisher Scientific, cat no. T200100)
530 solution was added to the sample and incubated at 28°C for 20 min. To generate a single cell
531 suspension, the sample was pipetted up and down 10 times every 5 min using a P1000 pipette.
532 Cells were then passed through a 40 µm strainer and centrifuged at 310 g for 5 minutes,
533 washed once with 1X FACSmax solution, and resuspended in phosphate buffered saline
534 (ThermoFisher Scientific, cat no. 14190094) with 0.5% bovine serum albumin (FisherScientific,
535 cat no. BP9706100). Cell viability of >90% was determined for each sample using Acridine
536 Orange (ThermoFisher Scientific, cat no. A1301) and Propidium Iodide (ThermoFisher Scientific
537 (cat no. P1304MP)) staining. For the screen of 50 transcription factors, after removing any dead
538 and severely deformed embryos at 24 hpf, injected embryos were split into four pools of
539 approximately 250 individuals each. Each pool was dissociated into a single-cell suspension as
540 described earlier. Each pool of dissociated cells was split into four parts, for a total of 16
541 samples. Each sample was loaded and processed on a 10X Chromium controller using Next
542 Gem 3.1.

543 **Gene expression matrix generation and genotype assignment**

544 Single-cell RNA sequencing transcriptome reads were processed using Cell Ranger software
545 from 10X Genomics and aligned to the GRCz11 v4.3.2 zebrafish reference genome (*53*). Gene
546 expression count matrices were generated using the default settings. A custom reference was
547 used containing the sequence of each gRNA protospacer (Table S1, Table S3) to generate a
548 separate CRISPR gRNA count matrix for each sample. To distinguish ambient gRNA reads
549 from injected gRNAs, a gaussian mixture model (Cell Ranger software, 10X Genomics) was
550 used to set UMI thresholds for each detection of each gRNA species. Predicted genotypes were
551 assigned to each cell based on the presence of gRNA using these thresholds. Cells with no
552 guide RNA counts exceeding the UMI threshold were assigned "Not Determined". Cells with
553 gRNA counts exceeding UMI thresholds for multiple gene targets were assigned as "Multiple".
554 Cells with detectable gRNAs for a particular gene were assigned as having that mutant
555 genotype. In silico doublet detection was performed with scDblFinder (*19*).

556 **Dimensionality reduction, clustering and cell cluster labeling**

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

564 Cell clusters were annotated in two steps. First, a reference atlas was used containing labeled
565 cell types from the Daniocell atlas of zebrafish development, using time points between 21 and
566 26 hpf (*8*). The reference atlas was used to generate transfer anchors using the principle
567 component analysis reference reduction using the TransferData function in Seurat (*54*),
568 generating a predicted cell type label for each cell in the dataset. Cell cluster labels were then
569 assigned based on a consensus of predicted cell types. Predicted cell types were verified and
570 adjusted using marker gene expression (**Table S2 and S4**).

571 **Determination of CRISPR editing frequency**

572 A subset of the pool of cells used for single-cell RNA sequencing library preparation were
573 retained for analysis of DNA editing at each target loci. Genomic DNA was extracted by lysing
574 the cells in 1x PCR lysis buffer (10 mM Tris (pH 8), 2mM EDTA, 0.2% Triton X-100) containing
575 0.2 mg/mL Proteinase K (MilliporeSigma, cat no. 3115887001). The RNase H2-dependent PCR
576 platform (Integrated DNA Technologies) was used for pooled amplification of each cut site,
577 according to the manufacturer's specifications. Briefly, pooled primers were designed and
578 produced by IDT to amplify each CRISPR cut site using the provided design tool. To avoid
579 overlapping PCR products, the reactions were split into two pools. For each pool, a 20 µl PCR
580 reaction was created containing 2 ng gDNA, 250 µM of each primer, and 1X rhAmpSeq library
581 mix. A second PCR reaction was performed to add sequencing adapters and a sample index,
582 followed by AMPure XP bead purification. Libraries were generated for both experimental MIC-
583 Drop injected embryo cells as well as wild type control cells. Each library was sequenced using
584 an Illumina MiSeq instrument with a Micro 150 cycle Paired End Sequencing Kit v2. The IDT
585 rhAmpSeq CRISPR Analysis Tool was used to assess DNA editing frequency at each loci. DNA
586 editing was quantified by the presence of nucleotide insertions or deletions (indels) in individual
587 reads compared to each amplicon reference sequence.

588 **Differential gene expression analysis**

589 In the pilot dataset (**Figure 1**), which did not include biological replicates, differential gene
590 expression was calculated with a Wilcoxon rank sum test as implemented in Seurat V4 (*54*).
591 Tests were performed by first subsetting cells from each cluster, then comparing gene
592 expression in cells of each mutant genotype against all other classified cells.

593 For the 50 gene MIC-Drop-seq screen, differential gene expression testing was performed using
594 pseudobulk analysis using edgeR (*55*) for each genotype in each cell type. First, a subset was
595 created for each cell type with more than 800 total cells. Next, aggregate read count libraries
596 were generated for each cell type and genotype, grouped by biological replicate. Each library
597 was filtered by expression and normalized prior to differential gene expression testing with the
598 glmQLFit test in edgeR. Genes with an FDR < 0.05 and an absolute value of log2-fold-change
599 greater than 0.5 were considered differentially expressed.

600 **Differential cell abundance analysis**

601 Cells counts were tabulated for each mutant genotype in each cell cluster across each biological
602 replicate pool. Hooke (https://github.com/cole-trapnell-lab/hooke), an extension of the Monocle3
603 package(*24*, *56*), was used to test for differential cell abundance using cell count data.
604 Abundance fold change and significance testing used Poisson-Lognormal models (PLNmodels)
605 version 1.0.4.200 (*57*). Estimated cell abundances were used to calculate positive or negative
606 log2 fold changes in cell abundance in mutant cells compared to all other cells. Cell type
607 abundance changes were considered significant with a FDR < 0.05 and an absolute log_2-fold
608 change of > 0.5.

609 **Determination of cell extrinsic and cell intrinsic effects**

610 To determine the lineage history of gene expression for each targeted gene, the Daniocell
611 dataset was downloaded and subsetted to include cell types and time points present from 3-24
612 hpf (*8*). Average expression of each factor was calculated for every cell type and time point.
613 Lineage relationships were manually curated based on prior studies (*10*, *58*). Gene expression
614 values for each target in each cluster were scaled to a range of 0-1. Factors were considered
615 expressed if the scaled expression value exceeded 0.2 in the cell type at 24 hpf. Transcriptional
616 phenotypes were classified for each cell type with a threshold of 10 differentially expressed
617 genes (absolute log_2FC > 0.5, FDR < 0.05). Phenotypes were classified as cell-intrinsic when
618 the cell type or its lineage precursors expressed the factor or were otherwise classified as cell-
619 extrinsic.

620 **Zebrafish F0 mutant validation**

621 F0 CRISPR injected embryos for validation experiments were generated using the same gRNAs
622 as was employed in the 50 gene MIC-Drop-seq screen. 1-1.5nL of the RNP was injected into
623 zebrafish embryos at the single cell stage. The RNP injection mixture contained 100 ng/µL
624 gRNAs, 2 µM EnGen Cas9 NLS (NEB, cat no. M0646M), 1x Buffer r3.1 (NEB, cat no. B7203S)
625 and 0.1% Phenol Red. At 24 hpf, embryos were manually dechorionated and fixed in 4%
626 Paraformaldehyde Solution in Phosphate Buffered Saline (ChemCruz, cat no. sc-281692)
627 overnight at 4ºC. Fixed embryos were washed with PBS and either taken directly into *in situ*
628 hybridization staining protocols or washed with successive washing of 50% and 100% methanol
629 for short term storage at 4ºC.

630 **In situ hybridization**

631 Antisense probes for in situ hybridization were generated by first PCR amplifying genomic
632 regions from gDNA extracted from TuAB embryos (*59*). Primers were designed incorporating T7
633 RNA polymerase promoter sequences for amplifying probes for *gata2a* (forward 5'-
634 CACTGACCATGAAGAAGGAG-3', reverse: 5'-GGATCCTAATAGACT
635 CACTATAGGGAAAGCAAAGGAAGTACTCATGG-3'), *dlx1a* (forward: 5'-AACCTTTGAGCGAG
636 TTTGTG-3' , reverse: 5'GGATCCTAATACGACTCACTATAGGGAAAGTCCAACATT
637 CTCGTAGT GC-3'), and *comp* (forward: 5'-GCCAAACTCATCACCATCAAAT-3', reverse: 5'-
638 GGATCC TAATACGACTCACTATAGAGAG CTTTCTTTGGGCCACTT-3'). The resulting
639 purified PCR products were used as templates for in vitro transcription using the Roche
640 Digoxigenin labeling kit (Sigma, cat. no 11175025910). In situ hybridization procedure was
641 performed as described previously (*59*). For probe hybridization, 50 ng of purified RNA probe
642 was incubated overnight at 70ºC in 200 µl hybridization mix containing 5% Dextran sulfate.

643 **In situ imaging and quantification**

644 Stained embryos were mounted in glycerol and imaged with a Leica M205 dissecting
645 microscope. For *comp* in situs, embryos were oriented laterally. For *gata2a* and *dlx1a* in insitus,
646 embryos were oriented for dorsal viewing. To quantify the staining area in *gata2a* and *dlx1a*
647 stained embryos, color deconvolution was performed in ImageJ to isolate blue hues. Images
648 were cropped to isolate the head and an 8-bit pixel intensity threshold of 0-230 was applied to
649 measure stained area. Total pixel area within the threshold was quantified and normalized by
650 dividing by the measured width of the head anterior to the eye.

651 **Gene ontology enrichment analysis**

652 The top 50 most frequently occurring DEGs were selected from all genotypes in prominent cell
653 types (>800 recovered cells). Gene ontology term enrichment analysis was performed on the
654 gene list using metascape express analysis via the web portal (https://metascape.org)(*60*).

655 **Vasculature imaging and quantification**

656 Zebrafish embryos were raised in E3 media until reaching shield stage and then transferred to
657 E3 media with 0.003% N-phenylthiourea (PTU, Sigma Aldrich) to block pigmentation. Embryos
658 from *kdrl*:GFP+ in-crosses were screened for fluorescence using a Zeiss fluorescent
659 stereoscope. Confocal microscopy was performed using a Zeiss LSM700 confocal microscope
660 on larvae anesthetized in 3-amino benzoic acid ethyl ester (Tricaine, Sigma Aldrich) in E3 at
661 0.2mg/ml and embedded in 1% low-melting agarose in 35mm glass bottom dishes (MatTek
662 Corporation, Part No. P35G-0.170-14-C). Optical sectioning through the trunk of embryos
663 generated Z-stacks of the developing vasculature with 3.028µm or 5.400µm steps between
664 slices and centered around the developing caudal vein plexus just posterior to the yolk
665 extension. Maximum intensity Z-projections were generated using ImageJ (NIH) software. The
666 width of the CVP and ISV width was measured at the location of the first ISV posterior to the
667 yolk extension in imageJ.

668 **Optical flow analysis**

669 To visualize blood flow in caudal vein plexus and intersomitic vessels, embryos were
670 anesthetized in tricaine and embedded in 1% low-melting agarose as previously described.
671 Videos capturing blood flow through the tail vasculature were recorded using a Leica Flexacam
672 C5 mounted on a stereoscope. All videos were recorded at 30fps and were at least 10 seconds
673 in duration. The FlowJ plugin integrated into ImageJ was used to measure optic flow. The

674 default central slice of each video was set as the central slice and the Lucas Kanade algorithm
675 with Gaussian derived gradient method, and Gaussian 1D regularization was used to compute
676 optic flow. Parameters were set as follows: **Σ**-s=2, **Σ**-t=2, **τ**=1.75, **Σ**-w=1.75.

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

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

683 **Modified gRNA qPCR**

684 Zebrafish embryos at the single cell stage were injected with RNP containing capture sequence-
685 modified gRNAs as described earlier. For each sample, 10 embryos at each of the the following
686 stages (6 hpf, 12 hpf, 24 hpf, 36 hpf, 48 hpf, 60 hpf, 72 hpf, 84 hpf, 96 hpf, and 120 hpf) were
687 collected, dechorionated and washed three times with PBS. The embryos were flash frozen in
688 liquid nitrogen and stored in -80ºC until ready for RNA extraction. RNA was extracted using a
689 miRNeasy kit (Qiagen, cat no. 217084) following manufacturer's instructions. Homogenization of
690 the tissue was performed using a handheld mechanical tissue grinder (Fisher Scientific, cat no.
691 12-141-361). A total of 1000ng of the extracted RNA was reverse transcribed using QuantiTect
692 reverse transcription kit (Qiagen, cat no. 205313) per manufacturer instructions. gRNA
693 barcode-specific primer (5'-TTGCTAGGACCGGCCTTAAAGC-3') at 500 nM was used for
694 reverse transcription. (qRT-PCR was performed on a 7500 Fast Real-Time PCR system
695 (Thermo Fisher) . A 10 µL qRT-PCR reaction mix contained 500 nM forward primer (5'-
696 GACACTATAGATCTGCCAGACG-3'), 500 nM reverse primer (5'-
697 GGACTAGCCTTATTTAAACTTGCTATG-3'), 1.25 ng/µL cDNA, and 1x PowerUp SYBR Green
698 master mix (Thermo Fisher, cat no. A25742). PCR cycling conditions were 50ºC for 2min, 95ºC
699 for 2min, [95ºC for 15s, 60ºC for 15s, 72ºC for 1min]x 40 cycles and a melt curve analysis at
700 95ºC for 15s, 60ºC for 1min, 95ºC for 30s and 60ºC for 15s.

701

702 ## References and Notes

703 1. A. Acevedo-Arozena, S. Wells, P. Potter, M. Kelly, R. D. Cox, S. D. M. Brown, ENU
704 Mutagenesis, a Way Forward to Understand Gene Function. *Annual Review of Genomics*
705 *and Human Genetics* **9**, 49–69 (2008).

706 2. M. Hammerschmidt, C. Nüsslein-Volhard, The expression of a zebrafish gene homologous
707 to Drosophila snail suggests a conserved function in invertebrate and vertebrate
708 gastrulation. *Development* **119**, 1107–1118 (1993).

709 3. W. Driever, L. Solnica-Krezel, A. F. Schier, S. C. Neuhauss, J. Malicki, D. L. Stemple, D. Y.
710 Stainier, F. Zwartkruis, S. Abdelilah, Z. Rangini, J. Belak, C. Boggs, A genetic screen for
711 mutations affecting embryogenesis in zebrafish. *Development* **123**, 37–46 (1996).

712 4. P. Haffter, M. Granato, M. Brand, M. C. Mullins, M. Hammerschmidt, D. A. Kane, J.
713 Odenthal, F. J. van Eeden, Y. J. Jiang, C. P. Heisenberg, R. N. Kelsh, M. Furutani-Seiki, E.
714 Vogelsang, D. Beuchle, U. Schach, C. Fabian, C. Nüsslein-Volhard, The identification of
715 genes with unique and essential functions in the development of the zebrafish, Danio rerio.
716 *Development* **123**, 1–36 (1996).

5. J. Y. Wang, J. A. Doudna, CRISPR technology: A decade of genome editing is only the beginning. *Science* **379**, eadd8643 (2023).

6. S. Parvez, Z. J. Brandt, R. T. Peterson, Large-scale F0 CRISPR screens in vivo using MIC-Drop. *Nat Protoc* **18**, 1841–1865 (2023).

7. S. Parvez, C. Herdman, M. Beerens, K. Chakraborti, Z. P. Harmer, J.-R. J. Yeh, C. A. MacRae, H. J. Yost, R. T. Peterson, MIC-Drop: A platform for large-scale in vivo CRISPR screens. *Science* **373**, 1146–1151 (2021).

8. A. Sur, Y. Wang, P. Capar, G. Margolin, M. K. Prochaska, J. A. Farrell, Single-cell analysis of shared signatures and transcriptional diversity during zebrafish development. *Dev Cell* **58**, 3028-3047.e12 (2023).

9. D. E. Wagner, C. Weinreb, Z. M. Collins, J. A. Briggs, S. G. Megason, A. M. Klein, Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. *Science* **360**, 981–987 (2018).

10. J. A. Farrell, Y. Wang, S. J. Riesenfeld, K. Shekhar, A. Regev, A. F. Schier, Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. *Science* **360**, eaar3131 (2018).

11. L. M. Saunders, S. R. Srivatsan, M. Duran, M. W. Dorrity, B. Ewing, T. Linbo, J. Shendure, D. W. Raible, C. B. Moens, D. Kimelman, C. Trapnell, Deep molecular, cellular and temporal phenotyping of developmental perturbations at whole organism scale. bioRxiv [Preprint] (2022). https://doi.org/10.1101/2022.08.04.502764.

12. L. Przybyla, L. A. Gilbert, A new era in functional genomics screens. *Nat Rev Genet* **23**, 89–103 (2022).

13. J. M. Replogle, T. M. Norman, A. Xu, J. A. Hussmann, J. Chen, J. Z. Cogan, E. J. Meer, J. M. Terry, D. P. Riordan, N. Srinivas, I. T. Fiddes, J. G. Arthur, L. J. Alvarado, K. A. Pfeiffer, T. S. Mikkelsen, J. S. Weissman, B. Adamson, Combinatorial single-cell CRISPR screens by direct guide RNA capture and targeted sequencing. *Nature biotechnology* **38**, 954 (2020).

14. P. Datlinger, A. F. Rendeiro, T. Boenke, M. Senekowitsch, T. Krausgruber, D. Barreca, C. Bock, Ultra-high-throughput single-cell RNA sequencing and perturbation screening with combinatorial fluidic indexing. *Nat Methods* **18**, 635–642 (2021).

15. A. Dixit, O. Parnas, B. Li, J. Chen, C. P. Fulco, L. Jerby-Arnon, N. D. Marjanovic, D. Dionne, T. Burks, R. Raychndhury, B. Adamson, T. M. Norman, E. S. Lander, J. S. Weissman, N. Friedman, A. Regev, Perturb-seq: Dissecting molecular circuits with scalable single cell RNA profiling of pooled genetic screens. *Cell* **167**, 1853-1866.e17 (2016).

16. D. Yao, L. Binan, J. Bezney, B. Simonton, J. Freedman, C. J. Frangieh, K. Dey, K. Geiger-Schuller, B. Eraslan, A. Gusev, A. Regev, B. Cleary, Scalable genetic screening for regulatory circuits using compressed Perturb-seq. *Nat Biotechnol*, 1–14 (2023).

17. S. Sunshine, A. S. Puschnik, J. M. Replogle, M. T. Laurie, J. Liu, B. S. Zha, J. K. Nuñez, J. R. Byrum, A. H. McMorrow, M. B. Frieman, J. Winkler, X. Qiu, O. S. Rosenberg, M. D.

757 Leonetti, C. J. Ye, J. S. Weissman, J. L. DeRisi, M. Y. Hein, Systematic functional
758 interrogation of SARS-CoV-2 host factors using Perturb-seq. *Nat Commun* **14**, 6245
759 (2023).

760 18. X. Jin, S. K. Simmons, A. Guo, A. S. Shetty, M. Ko, L. Nguyen, V. Jokhi, E. Robinson, P.
761 Oyler, N. Curry, G. Deangeli, S. Lodato, J. Z. Levin, A. Regev, F. Zhang, P. Arlotta, In vivo
762 Perturb-Seq reveals neuronal and glial abnormalities associated with autism risk genes.
763 *Science* **370**, eaaz6063 (2020).

764 19. P.-L. Germain, A. Lun, C. G. Meixide, W. Macnair, M. D. Robinson, Doublet identification in
765 single-cell sequencing data using *scDblFinder*. F1000Research 10:979 [Preprint] (2022).
766 https://doi.org/10.12688/f1000research.73600.2.

767 20. K. J. Griffin, S. L. Amacher, C. B. Kimmel, D. Kimelman, Molecular identification of
768 spadetail: regulation of zebrafish trunk and tail mesoderm formation by T-box genes.
769 *Development* **125**, 3379–3388 (1998).

770 21. F. Loosli, W. Staub, K. C. Finger-Baier, E. A. Ober, H. Verkade, J. Wittbrodt, H. Baier, Loss
771 of eyes in zebrafish caused by mutation of chokh/rx3. *EMBO Rep* **4**, 894–899 (2003).

772 22. A. J. Davidson, L. I. Zon, The caudal-related homeobox genes cdx1a and cdx4 act
773 redundantly to regulate hox gene expression and the formation of putative hematopoietic
774 stem cells during zebrafish embryogenesis. *Developmental Biology* **292**, 506–518 (2006).

775 23. A. G. Hayward II, P. Joshi, I. Skromne, Spatiotemporal analysis of zebrafish hox gene
776 regulation by Cdx4. *Developmental Dynamics* **244**, 1564–1573 (2015).

777 24. J. Cao, M. Spielmann, X. Qiu, X. Huang, D. M. Ibrahim, A. J. Hill, F. Zhang, S. Mundlos, L.
778 Christiansen, F. J. Steemers, C. Trapnell, J. Shendure, The single-cell transcriptional
779 landscape of mammalian organogenesis. *Nature* **566**, 496–502 (2019).

780 25. S. Skuntz, B. Mankoo, M.-T. T. Nguyen, E. Hustert, A. Nakayama, E. Tournier-Lasserve,
781 C. V. E. Wright, V. Pachnis, K. Bharti, H. Arnheiter, Lack of the mesodermal homeodomain
782 protein MEOX1 disrupts sclerotome polarity and leads to a remodeling of the cranio-
783 cervical joints of the axial skeleton. *Developmental Biology* **332**, 383–395 (2009).

784 26. P. Kirilenko, G. He, B. S. Mankoo, M. Mallo, R. Jones, N. Bobola, Transient Activation of
785 Meox1 Is an Early Component of the Gene Regulatory Network Downstream of Hoxa2.
786 *Molecular and Cellular Biology* **31**, 1301–1308 (2011).

787 27. B. S. Mankoo, S. Skuntz, I. Harrigan, E. Grigorieva, A. Candia, C. V. E. Wright, H.
788 Arnheiter, V. Pachnis, The concerted action of Meox homeobox genes is required
789 upstream of genetic pathways essential for the formation, patterning and differentiation of
790 somites. *Development* **130**, 4655–4664 (2003).

791 28. J. Y. Mohamed, E. Faqeih, A. Alsiddiky, M. J. Alshammari, N. A. Ibrahim, F. S. Alkuraya,
792 Mutations in MEOX1, Encoding Mesenchyme Homeobox 1, Cause Klippel-Feil Anomaly.
793 *The American Journal of Human Genetics* **92**, 157–161 (2013).

794 29. M. V. P. Dauer, P. D. Currie, J. Berger, Skeletal malformations of Meox1-deficient
795 zebrafish resemble human Klippel–Feil syndrome. *Journal of Anatomy* **233**, 687–695
796 (2018).

30. K. L. Posey, F. Coustry, J. T. Hecht, Cartilage oligomeric matrix protein: COMPopathies and beyond. *Matrix Biology* **71–72**, 161–173 (2018).

31. R. T. Willett, L. A. Greene, Gata2 Is Required for Migration and Differentiation of Retinorecipient Neurons in the Superior Colliculus. *J. Neurosci.* **31**, 4444–4455 (2011).

32. C. M. Solek, S. Feng, S. Perin, H. Weinschutz Mendes, M. Ekker, Lineage tracing of *dlx1a/2a* and *dlx5a/6a* expressing cells in the developing zebrafish brain. *Developmental Biology* **427**, 131–147 (2017).

33. J. M. Santos-Pereira, L. Gallardo-Fuentes, A. Neto, R. D. Acemel, J. J. Tena, Pioneer and repressive functions of p63 during zebrafish embryonic ectoderm specification. *Nat Commun* **10**, 3049 (2019).

34. S. Karthik, T. Djukic, J.-D. Kim, B. Zuber, A. Makanya, A. Odriozola, R. Hlushchuk, N. Filipovic, S. W. Jin, V. Djonov, Synergistic interaction of sprouting and intussusceptive angiogenesis during zebrafish caudal vein plexus development. *Sci Rep* **8**, 9840 (2018).

35. M. Rouleau, A. Medawar, L. Hamon, S. Shivtiel, Z. Wolchinsky, H. Zhou, L. De Rosa, E. Candi, S. de la Forest Divonne, M. L. Mikkola, H. van Bokhoven, C. Missero, G. Melino, M. Pucéat, D. Aberdam, TAp63 Is Important for Cardiac Differentiation of Embryonic Stem Cells and Heart Development. *Stem Cells* **29**, 1672–1683 (2011).

36. J. Liang, J. Wei, J. Cao, J. Qian, R. Gao, X. Li, D. Wang, Y. Gu, L. Dong, J. Yu, B. Zhao, X. Wang, In-organoid single-cell CRISPR screening reveals determinants of hepatocyte differentiation and maturation. *Genome Biol* **24**, 251 (2023).

37. C. Li, J. S. Fleck, C. Martins-Costa, T. R. Burkard, J. Themann, M. Stuempflen, A. M. Peer, Á. Vertesy, J. B. Littleboy, C. Esk, U. Elling, G. Kasprian, N. S. Corsini, B. Treutlein, J. A. Knoblich, Single-cell brain organoid screening identifies developmental defects in autism. *Nature* **621**, 373–380 (2023).

38. P. Datlinger, A. F. Rendeiro, C. Schmidl, T. Krausgruber, P. Traxler, J. Klughammer, L. C. Schuster, A. Kuchler, D. Alpar, C. Bock, Pooled CRISPR screening with single-cell transcriptome readout. *Nat Methods* **14**, 297–301 (2017).

39. X. Zheng, B. Wu, Y. Liu, S. K. Simmons, K. Kim, G. S. Clarke, A. Ashiq, J. Park, J. Li, Z. Wang, L. Tong, Q. Wang, K. T. Rajamani, R. Muñoz-Castañeda, S. Mu, T. Qi, Y. Zhang, Z. C. Ngiam, N. Ohte, C. Hanashima, Z. Wu, X. Xu, J. Z. Levin, X. Jin, Massively parallel in vivo Perturb-seq reveals cell-type-specific transcriptional networks in cortical development. *Cell* **187**, 3236-3248.e21 (2024).

40. S. Jia, X. Wu, Y. Wu, X. Cui, B. Tao, Z. Zhu, W. Hu, Multiple Developmental Defects in sox11a Mutant Zebrafish with Features of Coffin-Siris Syndrome. *International Journal of Biological Sciences* **16**, 3039 (2020).

41. Y. Tsurusaki, E. Koshimizu, H. Ohashi, S. Phadke, I. Kou, M. Shiina, T. Suzuki, N. Okamoto, S. Imamura, M. Yamashita, S. Watanabe, K. Yoshiura, H. Kodera, S. Miyatake, M. Nakashima, H. Saitsu, K. Ogata, S. Ikegawa, N. Miyake, N. Matsumoto, De novo SOX11 mutations cause Coffin–Siris syndrome. *Nat Commun* **5**, 4011 (2014).

836 42. L. M. Saunders, S. R. Srivatsan, M. Duran, M. W. Dorrity, B. Ewing, T. H. Linbo, J.
837 Shendure, D. W. Raible, C. B. Moens, D. Kimelman, C. Trapnell, Embryo-scale reverse
838 genetics at single-cell resolution. *Nature* **623**, 782–791 (2023).

839 43. S. Takeuchi, M. Yamamoto, S. Matsumoto, E. Kenjo, M. Karashima, Y. Ikeda, Pinpoint
840 modification strategy for stabilization of single guide RNA. *Journal of Chromatography B*
841 **1192**, 123149 (2022).

842 44. A. B. Rosenberg, C. M. Roco, R. A. Muscat, A. Kuchina, P. Sample, Z. Yao, L. T.
843 Graybuck, D. J. Peeler, S. Mukherjee, W. Chen, S. H. Pun, D. L. Sellers, B. Tasic, G.
844 Seelig, Single-cell profiling of the developing mouse brain and spinal cord with split-pool
845 barcoding. *Science* **360**, 176–182 (2018).

846 45. J. Cao, J. S. Packer, V. Ramani, D. A. Cusanovich, C. Huynh, R. Daza, X. Qiu, C. Lee, S.
847 N. Furlan, F. J. Steemers, A. Adey, R. H. Waterston, C. Trapnell, J. Shendure,
848 Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science* **357**,
849 661–667 (2017).

850 46. M. Bartosovic, M. Kabbe, G. Castelo-Branco, Single-cell CUT&Tag profiles histone
851 modifications and transcription factors in complex tissues. *Nat Biotechnol* **39**, 825–835
852 (2021).

853 47. K. Petri, W. Zhang, J. Ma, A. Schmidts, H. Lee, J. E. Horng, D. Y. Kim, I. C. Kurt, K.
854 Clement, J. Y. Hsu, L. Pinello, M. V. Maus, J. K. Joung, J.-R. J. Yeh, CRISPR prime
855 editing with ribonucleoprotein complexes in zebrafish and primary human cells. *Nat*
856 *Biotechnol* **40**, 189–193 (2022).

857 48. J. S. Fleck, J. G. Camp, B. Treutlein, What is a cell type? *Science* **381**, 733–734 (2023).

858 49. J. E. Rood, A. Hupalowska, A. Regev, Toward a foundation model of causal cell and tissue
859 biology with a Perturbation Cell and Tissue Atlas. *Cell* **187**, 4520–4545 (2024).

860 50. C. Trapnell, Revealing gene function with statistical inference at single-cell resolution. *Nat*
861 *Rev Genet* **25**, 623–638 (2024).

862 51. K. Labun, T. G. Montague, M. Krause, Y. N. Torres Cleuren, H. Tjeldnes, E. Valen,
863 CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic*
864 *Acids Research* **47**, W171–W174 (2019).

865 52. M. Manoli, W. Driever, Fluorescence-activated cell sorting (FACS) of fluorescently tagged
866 cells from zebrafish larvae for RNA isolation. *Cold Spring Harb Protoc* **2012**,
867 pdb.prot069633 (2012).

868 53. N. D. Lawson, R. Li, M. Shin, A. Grosse, O. Yukselen, O. A. Stone, A. Kucukural, L. Zhu,
869 An improved zebrafish transcriptome annotation for sensitive and comprehensive detection
870 of cell type-specific genes. *eLife* **9**, e55792 (2020).

871 54. T. Stuart, A. Butler, P. Hoffman, C. Hafemeister, E. Papalexi, W. M. Mauck, Y. Hao, M.
872 Stoeckius, P. Smibert, R. Satija, Comprehensive Integration of Single-Cell Data. *Cell* **177**,
873 1888-1902.e21 (2019).

874 55. M. D. Robinson, D. J. McCarthy, G. K. Smyth, edgeR: a Bioconductor package for
875 differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140
876 (2010).

877 56. X. Qiu, Q. Mao, Y. Tang, L. Wang, R. Chawla, H. A. Pliner, C. Trapnell, Reversed graph
878 embedding resolves complex single-cell trajectories. *Nat. Methods* **14**, 979–982 (2017).

879 57. J. Chiquet, M. Mariadassou, S. Robin, The Poisson-Lognormal Model as a Versatile
880 Framework for the Joint Analysis of Species Abundances. *Front. Ecol. Evol.* **9** (2021).

881 58. D. E. Wagner, A. M. Klein, Lineage tracing meets single-cell omics: opportunities and
882 challenges. *Nature Reviews Genetics* **21**, 410–427 (2020).

883 59. C. Thisse, B. Thisse, High-resolution in situ hybridization to whole-mount zebrafish
884 embryos. *Nat Protoc* **3**, 59–69 (2008).

885 60. Y. Zhou, B. Zhou, L. Pache, M. Chang, A. H. Khodabakhshi, O. Tanaseichuk, C. Benner,
886 S. K. Chanda, Metascape provides a biologist-oriented resource for the analysis of
887 systems-level datasets. *Nat Commun* 10, 1523 (2019).

888
889

Acknowledgments: 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. We thank the Cell Imaging and High-Throughput Genomics core facilities.

Funding:

National Institutes of Health grant F32HL156644 (CMC)

National Institutes of Health grant K01HG013682 (CMC)

National Institutes of Health grant R00HG012593 (SP)

National Institutes of Health grant R01GM134069 (RTP)

National Institutes of Health grant R35GM142950 (JAG)

National Institutes of Health grant R24OD035409 (RTP, JAG)

Northwestern University Startup Funds (SP)

One Utah Data Science Hub (JAG)

Author contributions:

Conceptualization: SP, RTP, JAG, CMC, ZJB

Methodology: SP, CMC, ZJB

Investigation: SP, CMC, ZJB, BWB, CJY

Visualization: CMC, SP, ZJB

Funding acquisition: RTP, JAG, SP, CMC, ZJB

Project administration: RTP, JAG

Supervision: RTP, JAG

Writing – original draft: CMC, SP, ZJB

Writing – review & editing: RTP, JAG, CMC, SP, ZJB, BWB

Competing interests: Authors declare that they have no competing interests.

Data and materials availability: All raw sequencing data, cell by gene expression matrices, and processed Seurat objects will be available on SRA.

All code and supporting data used to process data and generate visualizations in the manuscript are available at: https://github.com/clay-carey/MIC-Drop-seq