Single-cell sequencing of the *Drosophila* embryonic heart and muscle cells during differentiation and maturation

Georg Vogler[1,*], Bill Hum[1], Marco Tamayo[1], Yoav Altman[2], Rolf Bodmer[1,*]

[1] Development, Aging and Regeneration Program, [2] NCI-designated Cancer Center, Sanford Burnham Prebys Medical Discovery Institute, La Jolla, CA 92037 USA

* Co-corresponding authors

ORCID

Georg Vogler: 0000-0002-8303-3531

Bill Hum: 0000-0001-5591-8060

Marco Tamayo: 0000-0001-9891-0755

Yoav Altman: 0000-0001-6075-1111

Rolf Bodmer: 0000-0001-9087-1210

Abstract

The developing *Drosophila* heart consists of cardioblasts that differentiate into different types of cardiomyocytes and pericardial cells. A large body of work has identified numerous genes and pathways involved in heart specification and differentiation, downstream of cardiac transcription factors, such as Tinman (NKX2-5) and Dorsocross1/2/3 (TBX5). The advent of single-cell RNA sequencing (scRNAseq) technology allowed us for the first time to describe the transcriptome of different cardiac cell types in the *Drosophila* model at high resolution. Here, we applied scRNAseq on sorted cells of late-stage *Drosophila* embryos expressing a cardiac GFP reporter. We find distinct expression profiles of cardioblasts as they mature to cardiomyocytes, as well as discretely clustering pericardial cells, including a set expressing Tinman that potentially assist in heart morphogenesis. In addition, we describe other cell types that were sequenced as by-catch due to low but distinct extracardiac expression of the GFP reporter. Our studies on wildtype cardioblasts will be the foundation for investigating developmental profiles in mutant backgrounds and for generating gene regulatory networks at single-cell resolution during cardiogenesis.

Introduction

Congenital heart disease (CHD) is the most frequent birth defect among newborns and the leading cause of infant death or illness. While massive genome sequencing efforts have exponentially expanded the number of potential genes involved in human disease, including CHD (Claussnitzer et al., 2020; Jin et al., 2017; Sifrim et al., 2016) we still lack a fundamental mechanistic understanding how most of these genes and variants contribute to a disease (Blue et al., 2017). In addition, the majority of candidate genes are likely to act in concert with other genes in a patient-specific genetic background and the study of such gene-gene interactions has not been established yet at a scale required for the vast majority of diseases and conditions (Gifford et al., 2019; Kathiriya et al., 2020). Model systems that allow to systematically identify genetic interactions with high throughput are typically based on cell assays of organisms with a simple genetic architecture such as yeast or *Drosophila* S2 cell cultures (Billmann et al., 2016; Leeuwen et al., 2017; Tong et al., 2004). For genetic interactions within a multi-cellular organism, the fruit fly *Drosophila* has always been an important model organism (for example *NOTCH*, see Salazar and Yamamoto, 2018). The high degree of genetic conservation between flies and humans, including the presence of orthologous organ systems (including the heart) and conservation of many developmental and molecular-genetic mechanisms make *Drosophila* highly relevant for the study of human disease. This sets the stage for new approaches that integrate patient-specific genomics with an amenable *Drosophila* model to understand fundamental genetic pathways likely at play in the disease process (e.g. Harst et al., 2016; Theis et al., 2020). The prospect to study genetic interactions at the single-cell level in a multicellular organism is a logical next step in this important endeavor.

The formation of the heart is a finely controlled process of specification and differentiation events thought to be under tight control of cardiac transcription factors and signaling pathways, in vertebrates as well as in the fruit fly (Ahmad, 2017; Bodmer and Frasch, 2010; Cripps and Olson, 2002). The segmental *Drosophila* embryonic heart consists of 104 cardioblasts (CBs), distributed over eight segments (T3, A1-A7), as well as several types of pericardial cells (Panta et al., 2020; Reim and Frasch, 2009; Rotstein and Paululat, 2016). Post-specification the CBs express a common set of transcription factors (TFs), including Tail-up, H15 and Mef2 (Gajewski et al., 1997; Griffin et al., 2000; Mann et al., 2009). However, within each segment there are defined subsets of CBs that in addition express a combination of other cardiac TFs (Tinman and/or Ladybird; Doc1/2/3/Svp; Azpiazu and Frasch, 1993; Bodmer, 1993; Jagla et al., 1997b; Lo and Frasch, 2001; Reim and Frasch, 2005). CBs are flanked by

~140 pericardial cells (PCs) that are also expressing a common set of TFs (Hand, Zfh1), as well as other TF combinations that create different types of PCs, including Tin, Odd and Eve (Han et al., 2006; Han and Bodmer, 2003; Han and Olson, 2005; Lai et al., 1991; Perrin et al., 2004).

Of these transcription factors, *tinman* is already expressed and required in the cardiac mesoderm prior to CB subtype diversification (Azpiazu and Frasch, 1993; Bodmer, 1993). The cardiac specificity and DNA sequence-specific binding of the homeodomain-containing Tinman protein has allowed for detailed ChIP-on chip characterization of cardiac Tinman target genes (Jin et al., 2013; Junion et al., 2012; Zinzen et al., 2009), in parallel to several groups defining Tinman-dependent regulatory networks by bioinformatical FACS-assisted characterization of chromatin states of cardiac cells and machine learning approaches (Ahmad et al., 2014; Seyres et al., 2016). Nevertheless, the regulatory gene networks involved are still incompletely understood, in part because these could not be done in a cell specific manner.

In order to obtain a comprehensive picture of the logic and landscape of the gene networks within cardiac subtypes, it is now possible to perform genome wide RNA sequencing at the single cell level (scRNAseq). We therefore sought to create a transcriptional atlas of the embryonic fly heart to identify the full set of CB- and PC-expressed genes potentially involved in cardiac subtype differentiation and heart-tube morphogenesis and maturation. This will allow us to examine co-expression data of transcription factors and target genes at the single cell level in order to establish comprehensively built gene regulatory networks.

Here we present an atlas of the embryonic *Drosophila* heart during the last stages of embryonic development, based on FACS-sorted scRNAseq of heart cells. Firstly, we optimized single cell isolation protocols to maximize cell survival after FACS sorting for droplet-based RNA sequencing. We present a high-resolution map of cardiac cell types as well as other systematically recovered cell types of the late developing fly embryo, and our analysis uncovers distinct biological processes active during tissue differentiation and maturation.

Results and Discussion

FACS analysis / optimization for low cell number

Sample and collection buffer composition (Schneider's insect medium vs. 1x Accumax vs. 1% BSA/1xDPBS), droplet deflection direction (left vs. right), collection volume and post-sort centrifugation (20µl collection buffer volume without centrifugation vs. 500µl collection volume followed by centrifugation and resuspension in 50µl final volume) were varied to optimize the recovery of live cells. For each combination of conditions, 20,000 live (DAPI-negative) cells were sorted into cooled 1.5ml Eppendorf tubes. Recovery and viability were assessed immediately after sorting by adding 300ng/ml DAPI and then running the entire collection volume back through the cytometer to count live cells. We did not observe large differences in cell survival between sample buffers (Accumax or Schneider's), however, maximum cell survival after sorting was observed with 1%BSA/1xDPBS with ~32% surviving cells.

Highly-autofluorescent GFP-negative cells were identified and excluded from the sort gate by plotting a second fluorescence channel (561nm Ex., 610nm Em.) against GFP and avoiding cells along the diagonal (Fig. 1B'). The GFP-high sort gate position was further optimized using Imaging Flow Cytometry (IFC). Briefly, FACS-enriched GFP+ cells were imaged in flow at 40X magnification and target cells identified based on brightfield and GFP imagery. These handpicked events were plotted for GFP fluorescence and then circumscribed by a provisional gate. This gate was then projected onto the unsorted sample to determine the baseline frequency of GFP-high cells. The final GFP-high sort gate was adjusted so that the frequency of events matched the result from imaging flow cytometry.

Heart-specific GFP reporter line

To sort cardiac cells we used *mid*^{E19}::GFP (Jin, 2013) as fluorescent reporter, which we combined with *tinD*-Gal4 to express RNAi in cardiac precursors. Homozygous *mid*^{E19}::GFP has high GFP expression in Tinman-positive CBs (Fig. 1A), but also some low but detectable levels of GFP in other cells. Since we used heterozygous *mid*^{E19}::GFP throughout the experiments, this extra-cardiac GFP signal dropped significantly compared to the homozygous condition, but could still be visualized when using high image contrast (Fig. 1A', B'). The majority of these weakly GFP-positive cells appear to

belong to nuclei of dorsal muscles (Fig.1A'), as well as cells in the nervous system, head region and visceral mesodermal derivatives. In addition, we founnd on occasion GFP-positive ostia cells as well as plasmatocytes even in heterozygous *mid*^{E19}::GFP embryos (Fig. 1 and Supplemental Movie 1). Since we aimed to limit FAC-sorting time to 1h to maximize cell viability, but also aimed at a high cell yield for subsequent Chromium steps (10k cells input), we decided to collect cells first at high stringency settings (Fig. 1B', high GFP) for 50min and added cells with lower GFP (Fig. 1B', low GFP) for the remaining time to reach a target cell number of 30k sorted cells prior to loading on the Chromium chip (32% survival rate → ~10k cells loaded on chip).

Single-cell RNA-sequencing of Drosophila embryonic heart cells

Following cell isolation and fluorescent activated cell sorting, cells were processed using 10xGenomic's Chromium to recover single cell transcriptomic data. Cell recovery following FACS optimizations (see above) improved capture rate up to ten-fold (before/after: sample pair 1 472/4385, sample pair 2 914/5489). Seurat data integration resulted in a rich dataset with a total of about 12,000 cells from 6 samples and consistently identified 10k genes, all of which were used for subsequent clustering and marker gene identification (Fig. 2).

As indicated by the analysis of *mid*^{E19}::GFP reporter gene expression (Fig. 1A), we expected to identify high GFP-expressing CBs as well as low GFP-expressing cells from a number of different tissues. In concordance with fluorescent *in situ* staining of *eGFP* (Supplemental Figure 3), we find that most recovered cells also express this reporter gene, while some clusters already have shut down marker gene expression, but still contain eGFP protein (Supplemental Figure 3). Our single-cell RNA data clustered very distinctly, and we were able to assign the identity of these cell types based on marker gene expression (Fig. 2A, Supplemental Figure 1). As expected, *tin* and *Hand* co-expression let us identify unequivocally the CBs as well as the PCs (see next section). In addition, we identified clusters with gene expression characteristic of muscle cells of different stages identified by the expression of *lame-duck* and *Mhc*, gonadal cells by their expression of *traffic jam* and *Six4*, neurons (by *elav)* as well as fat body (*apolpp*; Fig. 2C, Supplemental Figure 1).

Cardiac and pericardial cell gene expression differences at high resolution

Importantly, we captured two clusters of cells that express the CB and PC marker gene *Hand* (Figure 2, 3, see dashed lines in Figure 2A). To test whether these cells and the computed clustering include and reflect current knowledge of *Drosophila* heart development, we analyzed select genes known to be expressed in the developing heart for their respective expression pattern (Fig. 3A). Indeed, the cardiac expression of these genes correlated with their predicted clusters within the heart, confirming the high quality of the single-cell transcriptome data. For example, the muscle transcription factor *Mef2* was found in presumptive CBs and was absent in PCs, while the cardiac collagen *pericardin* was strongly expressed in PCs and at low levels in CBs.

Reclustering of these cardiac cells further subdivided the CBs into *Mhc*-negative and *Mhc*-positive CBs: *Mhc*-negative 'CBs' are non-differentiated cardioblasts, and *Mhc*-positive 'CMs' as differentiating cardiomyocytes. PCs distinctly as *tinman*-positive 'tin-PCs' and *odd-skipped*-positive 'odd-PCs' (Figure 3B, C). In our data set of 12,000 cells, we recovered a total of 520 CBs, 439 CMs, 236 tin-PCs and 270 odd-PCs.

We then tested for gene expression differences between CBs and CMs using GO term analysis (using g:Profiler, Raudvere et al., 2019) to gain insight into the underlying biological processes in CBs vs CMs. When analyzing genes overrepresented in CMs (86 genes, avg_log2FC < -1), the GO terms with highest significance included all myofibril development and sarcomere (Supplemental Figure 4). In contrast, genes overrepresented in CBs (82 genes, avg_log2FC > 0.6) show GO terms mostly of cellular biogenesis, including ribosome biogenesis (Supplemental Figure 5). Interestingly, the ribosomal gene *RpL13* is required in the embryonic heart for later growth and remodeling of the larval and pupal heart, and its knockdown in the embryo causes a 'no adult heart' phenotype (Schroeder et al., 2019). The GO term analysis suggests that CBs increase the expression of genes necessary later for the translation of genes such as *Mhc*, *Mlc2* and *bent* into muscle structural proteins during differentiation. This is also reflected by the anti-correlation of the expression of *Myc* and *Mhc* in CBs vs. CMs (Figure 3D): *Myc* is a master regulator of ribosome biogenesis and cell growth (Destefanis et al., 2020; Gallant, 2013; Grewal et al., 2005) and it is highly expressed prior to CM differentiation. We conclude that CBs during their final developmental steps of morphogenesis initiate a CM differentiation program, that enables robust

expression (and translation) of sarcomeric genes, as well as other components necessary for contraction, such as ion channels (see below).

Expression of ion channels in fly heart cells

We then compared a comprehensive list of *Drosophila* ion channels (FlyBase; Larkin et al., 2020) with the genes expressed in the heart (Table 1). Similar to sarcomeric gene expression, CMs express many ion channels that are important for muscle contraction, including calcium channels (e.g. *Ca-alpha1D*), glutamate receptor proteins (e.g. *GluRIIA-D*) and potassium channels (*SK*, *ShaI*). Interestingly, while Odd-PCs do not show discrete expression of ion channels, Tin-PCs express channels involved in mechano-sensation (*piezo*, *pain*), as well as some channels shared with CMs (Table 1).

Identification of new markers of Tinman-positive pericardial cells

While cardioblasts and Odd-skipped-positive pericardial cells develop into critical functional organs (CBs into the cardiomyocytes of the larval and adult heart, Odd-PCs into nephrocytes), the developmental function of Tinman-positive pericardial cells (Tin-PCs) is unclear. Tin-PCs are created via asymmetric cell division from a Tin-positive precursor that gives rise to a Tin/Lb-positive CB and Tin/Lb-positive PC (Han and Bodmer, 2003; Jagla et al., 1997a; Ward and Skeath, 2000). At the end of heart morphogenesis, these cells position ventral to the closing heart tube (Fig. 4A-C), with high expression levels of the cell adhesion receptor Neurotactin (Nrt), as well as the cardiac transcription factors Zfh1 and Tin. Our single-cell data allowed for the first time to identify new genes to describe this cell in greater detail (Table 2). We identified several new marker genes that have not been described in this cell context and report here for the first time their detailed expression (Figure 4C-H). First, we found that the homeobox transcription factor *cut* is expressed in Tin-PCs throughout heart morphogenesis and closure (Fig. 4C). Furthermore, we find that the GPCR *Corazonine receptor* (*CrzR*) is also expressed in these cells, already during dorsal closure of the heart and again throughout morphogenesis (Fig. 4G). Lastly, we tested for the expression of the cell adhesion molecule and septate junction protein *Lachesin* (*Lac*), which was previously identified to be necessary for the integrity of the embryonic heart (Yi et al., 2008). *Lac* is expressed in the heart already at stage 15 (as well as in tissues outside the heart such as the developing trachea, Fig. 4H). However, especially during the final stages of heart morphogenesis *Lac* expression is increased especially in the Tin-PCs of the heart proper (Fig. H'''), indicating a region-specific function for *Lac*.

These gene expression patterns let us hypothesize that Tin-PCs may be required for heart-tube closure, lumen formation and proper cardiac morphogenesis, which can now be tested with cell type-specific genetic manipulations using our high-resolution data.

Cell removal during muscle differentiation

Interestingly, we also observed one cluster with a 'dual identity' – plasmatocytes and muscle cell, and this cluster is also closest to the actual plasmatocytes cluster (Figure 2, Supplemental Figure 1). These dual-identity cells express markers genes that identify these cells as plasmatocytes (i.e., expression of *SPARC, Pvr*, Fig. 5A), but they also express muscle markers such as *Mhc* and *wupA* (Fig. 5A). Since these cells are identified throughout our replicates and are the only occurrence of such dual-identity cells, we hypothesize that they are a product of macrophage-like plasmatocytes fusing with muscle cells or myoblasts likely to remove them. Interestingly, time-lapse imaging of the dorsal muscle area shows high activity of migrating plasmatocytes, and we were able to capture an event that shows the uptake of cells by a plasmatocytes during embryonic muscle morphogenesis (Figure 5B and Supplemental Video 1). This suggests that plasmatocytes are not only 'sculpting' surrounding tissue by secretion of ECM/collagen (*Col4a1, vkg*), but also by scavenging superfluous cells such as fusion competent myoblasts similar to their role in clearing apoptotic cells (Franc et al., 1999).

Muscle differentiation at the single-cell level

In addition to these ectopic *Mhc*-positive cells, we recover a large number of cells that co-express the myoblast marker *lame-duck* (*lmd*, Fig. 6A) and the somatic muscle transcription factor *Mef2* (Supplemental Figure 2). These cell clusters show a trajectory within the UMAP plot that have a gradient of increased expression of muscle differentiation markers such as sarcomeric components *Mhc* and *wupA* (Figure 6A, Supplemental Figure 2), as well as a complementary expression pattern of early muscle precursor markers such as *twist* and – similar to the developing heart – strong early presence of *Myc* transcripts that is diminished as cells differentiate and express sarcomeric genes.

Interestingly, there is also an inverse gradient of the signaling protein Notch and most of its direct target genes of the Bearded- and E(spl)-complex (Fig. 6A and Supplemental Figure 2). This expression parallels their known role for adult muscle precursor (AMP) and adult muscle development

(Figeac et al., 2010; Lai et al., 2000). However, with regard to other instructive muscle transcription factors such as *homothorax (Hth), extradenticle (exd)* and *spalt major* (*salm*) (Bryantsev et al., 2012; Schönbauer et al., 2011), *hth* and *exd* are broadly expressed, while *salm* is only expressed in those muscle cells with high Notch signalling (Supplemental Figure 2). We hypothesize that Notch signaling is likely antagonizing muscle differentiation, which is an additional role that is likely different from its early function during cell specification in the developing mesoderm. Because the cell layout within the UMAP plot could be due to temporal as well as spatial differences between cells we sought to analyze their relationship using RNA velocity (using velocytoR; Manno et al., 2018). Since most single cell RNA sequencing data in addition to mature transcripts also contains reads from unspliced mRNA, these can be separately analyzed to determine their expression patterns. If the gene expression gradients within UMAP plots reflect a gradual progression between states, the relative fractions of spliced and unspliced mRNA molecules for each gene within each cell should allow to predict their future state and would result in a velocity vector showing this progression. When we ran this analysis, we indeed found a robust trajectory from early (less differentiated) myoblasts/muscle cells to differentiated muscles cells with high expression of sarcomeric genes (Figs. 6C). For example, the onset of expression of the main muscle structural gene *Mhc* (by relative increased amounts of unspliced *Mhc*) occurs before the we observe the full expression level of the spliced *Mhc* transcript. Since our biological samples all covered a developmental time window of 2 hours, the RNA velocity analysis seems to reflect this at the molecular level.

These are just two examples of how one can leverage single cell sequencing of sorted groups of cells from a small temporal window of development to gain insight into biological processes and to motivate their detailed analysis. Using this scRNAseq technology of embryonic cardioblasts with mutant or knockdown hearts, e.g for cardiogenic TFs (*tin, Doc, pnr, etc.*), we will be able to generate detailed gene regulatory networks at single heart cell level during cardiogenesis.

Material and Methods

Fly stocks and reagents

For embryo collections of *Drosophila melanogaster* we combined *tinD*::Gal4 (gift from M. Frasch) with mid^{E19}::GFP (gift from I. Reim). Virgins of *tinD*::Gal4; mid^{E19}::GFP were collected *en masse* (~7000

females) and crossed to ~ 4000 w^{1118} males (VDRC stock #60000). To overexpress membranous RFP, virgins were crossed to UAS-*CD4::tdTomato* (RRID: BDSC_35837). For RNAi knockdown of *tin*, we used VDRC line v12656. Flies were maintained on standard cornmeal/molasses fly food at 25°C. Egg collections were done in cages (10cm diameter, Genesee Scientific) with grape juice agar plates supplemented with live yeast paste.

Timing of egg collections

Cages were kept at 25°C throughout the experiment, and yeast-paste supplemented grape juice agar plates were changed 3x per day. On the day before cell sorting, plates were replaced every 2h from 10:00-18:00h; the collections from 14-16h and 16-18h were directly transferred into a programmable incubator with the following temperature cycle: 9:00-00:00 (18°C), 00:00-3:00 (18°C), 3:00-9:00 (29°C). Staged embryos (st.16/17) were directly processed at 9:00 the next morning (24h time format).

Preparation of *Drosophila* embryo cell suspension

We adapted the protocols to prepare a cell suspension from primary from Bryantsev, 2012 and Salmand, 2011. Embryos were dechorionated using 4% bleach for 3min and transferred into a 1ml Dounce homogenizer containing 500ml Accumax (Innovative Cell Technologies). Embryos were homogenized on ice using 10 strokes ('tight' pestle) and incubated at room temperature for 15min. The suspension was vortexed briefly every 5min to resuspend any cells that have settled. The suspension was then poured over a 100µm cell strainer and the flow-through collected in a 50ml Falcon tube. The Dounce homogenizer was rinsed with another 0.5ml of Accumax and added to the 100µm filter. Next, we poured the filtered suspension through a 30µm MACS filter (Miltenyi Biotech) and the flow-through collected in a 15ml Falcon tube. Cells were centrifuged for 3min at 300g and the cell pellet was washed once with 1.5ml of Accumax (or Schneider's medium or 1% BSA/1xDPBS during optimization steps) and stored on ice.

FACS of Drosophila GFP-positive cells

Cells were sorted on a SORP FACSAriaIIu (BD Biosciences, San Jose) with 100 µm nozzle, 23 psi sheath pressure, and 32.7 kHz drop frequency and FACSDiVa v8.0.3 software. Standard excitation and filters were used for both GFP and DAPI. Autofluorescence was assessed with 561 nm excitation and

610/20BP emission filter. Both sample chamber and collection tubes were kept at 4°C for the duration of the sort. FACSDiVa and FlowJo v10 were used for data analysis.

Imaging Flow Cytometry

Cells were imaged both before and after sorting at 40X magnification on a 2-camera Amnis ImageStream MkII (Luminex, Seattle) set to low speed/high resolution. IFC files were acquired with Amnis INSPIRE v6.2.187 and analyzed with Amnis IDEAS v6.2.187.0 software (Luminex, Seattle).

Chromium 10x Library preparation and sequencing

An estimated input of up to 10,000 sorted cells were directly processed following the manufacturer's protocol for Chromium Single Cell 3 Reagent Kits v3. Libraries were sequenced at a target depth of 30k reads per cell at the La Jolla Institute for Immunology Sequencing Core.

Data analysis

All scSeq data was processed using R (version 4.0.3) and Seurat (version 3.9.9.9010, REF). The full R script source code will be available on Github.

HCR RNA in situ

HCR RNA in situ hybridization was done according to the manufacturer's protocol for *Drosophila* embryos (Molecular Instruments). Probes were synthesized for the following genes (using the provided GenBank accession numbers): *CG6415* (NM_135597.3); *H15* (NM_135082); *CrzR* (NM_140314.4); *Lac* (NM_078989.3). YFP/GFP probe *d2eGFP* was obtained from the manufacturer. Images were acquired using a C-Apochromat 40×/1.2 NA water immersion objective lens on an AxioImager Z1 equipped with an Apotome.2 (all Carl Zeiss) and a OrcaFlash4.0LT camera (Hamamatsu), with ZEN software 2.3 software (Carl Zeiss). Image processing was done using ImageJ/FIJI (REF).

Acknowledgements

We appreciate the support of the Sanford Burnham Prebys Shared Resources, including Amy Cortez, and Ziyi Sang (Flow Cytometry), and Brian James (Genomics).

References

Ahmad SM. 2017. Conserved signaling mechanisms in Drosophila heart development. *Dev Dynam* **246**:641–656. doi:10.1002/dvdy.24530

Ahmad SM, Busser BW, Huang D, Cozart EJ, Michaud S, Zhu X, Jeffries N, Aboukhalil A, Bulyk ML, Ovcharenko I, Michelson AM. 2014. Machine learning classification of cell-specific cardiac enhancers uncovers developmental subnetworks regulating progenitor cell division and cell fate specification. *Development (Cambridge, England)* **141**:878–888. doi:10.1242/dev.101709

Azpiazu N, Frasch M. 1993. tinman and bagpipe: two homeo box genes that determine cell fates in the dorsal mesoderm of Drosophila. **7**:1325–1340.

Billmann M, Horn T, Fischer B, Sandmann T, Huber W, Boutros M. 2016. A genetic interaction map of cell cycle regulators. *Mol Biol Cell* **27**:1397–1407. doi:10.1091/mbc.e15-07-0467

Blue GM, Kirk EP, Giannoulatou E, Sholler GF, Dunwoodie SL, Harvey RP, Winlaw DS. 2017. Advances in the Genetics of Congenital Heart Disease: A Clinician's Guide. *Journal of the American College of Cardiology* **69**:859–870. doi:10.1016/j.jacc.2016.11.060

Bodmer R. 1993. The gene tinman is required for specification of the heart and visceral muscles in Drosophila. *Development (Cambridge, England)* **118**:719–729.

Bodmer R, Frasch M. 2010. Heart Development and Regeneration, Academic Press; 1 edition (June 18, 2010). Academic Press; 1 edition (June 18, 2010).

Bryantsev AL, Duong S, Brunetti TM, Chechenova MB, Lovato TL, Nelson C, Shaw E, Uhl JD, Gebelein B, Cripps RM. 2012. Extradenticle and Homothorax Control Adult Muscle Fiber Identity in Drosophila. *Developmental cell* **23**:664–673. doi:10.1016/j.devcel.2012.08.004

Claussnitzer M, Cho JH, Collins R, Cox NJ, Dermitzakis ET, Hurles ME, Kathiresan S, Kenny EE, Lindgren CM, MacArthur DG, North KN, Plon SE, Rehm HL, Risch N, Rotimi CN, Shendure J, Soranzo N, McCarthy MI. 2020. A brief history of human disease genetics. *Nature* **577**:179–189. doi:10.1038/s41586-019-1879-7

Cripps RM, Olson EN. 2002. Control of cardiac development by an evolutionarily conserved transcriptional network. *Developmental biology* **246**:14–28. doi:10.1006/dbio.2002.0666

Destefanis F, Manara V, Bellosta P. 2020. Myc as a Regulator of Ribosome Biogenesis and Cell Competition: A Link to Cancer. *Int J Mol Sci* **21**:4037. doi:10.3390/ijms21114037

Figeac N, Jagla T, Aradhya R, Ponte JPD, Jagla K. 2010. Drosophila adult muscle precursors form a network of interconnected cells and are specified by the rhomboid-triggered EGF pathway. *Development* **137**:1965–1973. doi:10.1242/dev.049080

Franc NC, Heitzler P, B. RA, Ezekowitz, White K. 1999. Requirement for Croquemort in Phagocytosis of Apoptotic Cells in Drosophila. *Science* **284**:1991–1994. doi:10.1126/science.284.5422.1991

Gajewski K, Kim Y, Lee YM, Olson EN, Schulz RA. 1997. D-mef2 is a target for Tinman activation during Drosophila heart development. *The EMBO Journal* **16**:515–522. doi:10.1093/emboj/16.3.515

Gallant P. 2013. Myc Function in Drosophila. *Csh Perspect Med* **3**:a014324. doi:10.1101/cshperspect.a014324

Gifford CA, Ranade SS, Samarakoon R, Salunga HT, Soysa TY de, Huang Y, Zhou P, Elfenbein A, Wyman SK, Bui YK, Metzler KRC, Ursell P, Ivey KN, Srivastava D. 2019. Oligogenic inheritance of a human heart disease involving a genetic modifier. *Science (New York, NY)* **364**:865–870. doi:10.1126/science.aat5056

Grewal SS, Li L, Orian A, Eisenman RN, Edgar BA. 2005. Myc-dependent regulation of ribosomal RNA synthesis during Drosophila development. *Nat Cell Biol* **7**:295–302. doi:10.1038/ncb1223

Griffin KJ, Stoller J, Gibson M, Chen S, Yelon D, Stainier DY, Kimelman D. 2000. A conserved role for H15-related T-box transcription factors in zebrafish and Drosophila heart formation. *Developmental biology* **218**:235–247. doi:10.1006/dbio.1999.9571

Han Z, Bodmer R. 2003. Myogenic cells fates are antagonized by Notch only in asymmetric lineages of the Drosophila heart, with or without cell division. *Development (Cambridge, England)* **130**:3039–3051.

Han Z, Olson EN. 2005. Hand is a direct target of Tinman and GATA factors during Drosophila cardiogenesis and hematopoiesis. *Development (Cambridge, England)* **132**:3525–3536. doi:10.1242/dev.01899

Han Z, Yi P, Li X, Olson EN. 2006. Hand, an evolutionarily conserved bHLH transcription factor required for Drosophila cardiogenesis and hematopoiesis. *Development (Cambridge, England)* **133**:1175–1182. doi:10.1242/dev.02285

Harst P van der, Setten J van, Verweij N, Vogler G, Franke L, Maurano MT, Wang X, Leach IM, Eijgelsheim M, Sotoodehnia N, Hayward C, Sorice R, Meirelles O, Lyytikäinen L-P, Polasek O, Tanaka T, Arking DE, Ulivi S, Trompet S, Müller-Nurasyid M, Smith AV, Dörr M, Kerr KF, Magnani JW, M FDG, Zhang W, Nolte IM, Silva CT, Padmanabhan S, Tragante V, Esko T, Abecasis GR, Adriaens ME, Andersen K, Barnett P, Bis JC, Bodmer R, Buckley BM, Campbell H, Cannon MV, Chakravarti A, Chen LY, Delitala A, Devereux RB, Doevendans PA, Dominiczak AF, Ferrucci L, Ford I, Gieger C, Harris TB, Haugen E, Heinig M, Hernandez DG, Hillege HL, Hirschhorn JN, Hofman A, Hubner N, Hwang S-J, Iorio A, Kähönen M, Kellis M, Kolcic I, Kooner IK, Kooner JS, Kors JA, Lakatta EG, Lage K, Launer LJ, Levy D, Lundby A, Macfarlane PW, May D, Meitinger T, Metspalu A, Nappo S, Naitza S, Neph S, Nord AS, Nutile T, Okin PM, Olsen JV, Oostra BA, Penninger JM, Pennacchio LA, Pers TH, Perz S, Peters A, Pinto YM, Pfeufer A, Pilia MG, Pramstaller PP, Prins BP, Raitakari OT, Raychaudhuri S, Rice KM, Rossin EJ, Rotter JI, Schafer S, Schlessinger D, Schmidt CO, Sehmi J, Silljé HHW, Sinagra G, Sinner MF, Slowikowski K, Soliman

EZ, Spector TD, Spiering W, Stamatoyannopoulos JA, Stolk RP, Strauch K, Tan S-T, Tarasov KV, Trinh B, Uitterlinden AG, Boogaard M van den, Duijn CM van, Gilst WH van, Viikari JS, Visscher PM, Vitart V, Völker U, Waldenberger M, Weichenberger CX, Westra H-J, Wijmenga C, Wolffenbuttel BH, Yang J, Bezzina CR, Munroe PB, Snieder H, Wright AF, Rudan I, Boyer LA, Asselbergs FW, Veldhuisen DJ van, Stricker BH, Psaty BM, Ciullo M, Sanna S, Lehtimäki T, Wilson JF, Bandinelli S, Alonso A, Gasparini P, Jukema JW, Kääb S, Gudnason V, Felix SB, Heckbert SR, Boer RA de, Newton-Cheh C, Hicks AA, Chambers JC, Jamshidi Y, Visel A, Christoffels VM, Isaacs A, Samani NJ, Bakker PIW de. 2016. 52 Genetic Loci Influencing Myocardial Mass. *Journal of the American College of Cardiology* **68**:1435–1448. doi:10.1016/j.jacc.2016.07.729

Jagla K, Frasch M, Jagla T, Dretzen G, Bellard F, Bellard M. 1997a. ladybird, a new component of the cardiogenic pathway in Drosophila required for diversification of heart precursors. *Dev Camb Engl* **124**:3471–9.

Jagla K, Jagla T, Heitzler P, Dretzen G, Bellard F, Bellard M. 1997b. ladybird, a tandem of homeobox genes that maintain late wingless expression in terminal and dorsal epidermis of the Drosophila embryo. *Development (Cambridge, England)* **124**:91–100.

Jin H, Stojnic R, Adryan B, Ozdemir A, Stathopoulos A, Frasch M. 2013. Genome-wide screens for in vivo Tinman binding sites identify cardiac enhancers with diverse functional architectures. *PLoS Genetics* **9**:e1003195. doi:10.1371/journal.pgen.1003195

Jin SC, Homsy J, Zaidi S, Lu Q, Morton S, DePalma SR, Zeng X, Qi H, Chang W, Sierant MC, Hung W-C, Haider S, Zhang J, Knight J, Bjornson RD, Castaldi C, Tikhonoa IR, Bilguvar K, Mane SM, Sanders SJ, Mital S, Russell MW, Gaynor JW, Deanfield J, Giardini A, Porter GA, Srivastava D, Lo CW, Shen Y, Watkins WS, Yandell M, Yost HJ, Tristani-Firouzi M, Newburger JW, Roberts AE, Kim R, Zhao H, Kaltman JR, Goldmuntz E, Chung WK, Seidman JG, Gelb BD, Seidman CE, Lifton RP, Brueckner M. 2017. Contribution of rare inherited and de novo variants in 2,871 congenital heart disease probands. *Nature Genetics* **49**:1593–1601. doi:10.1038/ng.3970

Junion G, Spivakov M, Girardot C, Braun M, Gustafson EH, Birney E, Furlong EEM. 2012. A Transcription Factor Collective Defines Cardiac Cell Fate and Reflects Lineage History. *Cell* **148**:473–486. doi:10.1016/j.cell.2012.01.030

Kathiriya IS, Rao KS, Iacono G, Devine WP, Blair AP, Hota SK, Lai MH, Garay BI, Thomas R, Gong HZ, Wasson LK, Goyal P, Sukonnik T, Hu KM, Akgun GA, Bernard LD, Akerberg BN, Gu F, Li K, Speir ML, Haeussler M, Pu WT, Stuart JM, Seidman CE, Seidman JG, Heyn H, Bruneau BG. 2020. Modeling Human TBX5 Haploinsufficiency Predicts Regulatory Networks for Congenital Heart Disease. *Dev Cell*. doi:10.1016/j.devcel.2020.11.020

Lai EC, Bodner R, Posakony JW. 2000. The enhancer of split complex of Drosophila includes four Notch-regulated members of the bearded gene family. *Dev Camb Engl* **127**:3441–55.

Lai Z, Fortini ME, Rubin GM. 1991. The embryonic expression patterns of zfh-1 and zfh-2, two Drosophila genes encoding novel zinc-finger homeodomain proteins. *Mech Develop* **34**:123–134. doi:10.1016/0925-4773(91)90049-c

Larkin A, Marygold SJ, Antonazzo G, Attrill H, dos Santos G, Garapati PV, Goodman JL, Gramates LS, Millburn G, Strelets VB, Tabone CJ, Thurmond J, Consortium F, Perrimon N, Gelbart SR, Agapite J, Broll K, Crosby M, Santos G dos, Falls K, Gramates LS, Jenkins V, Longden I, Matthews B, Sutherland C, Tabone CJ, Zhou P, Zytkovicz M, Brown N, Antonazzo G, Attrill H, Garapati P, Larkin A, Marygold S, McLachlan A, Millburn G, Pilgrim C, Ozturk-Colak A, Trovisco V, Kaufman T, Calvi B, Goodman J, Strelets V, Thurmond J, Cripps R, Lovato T. 2020. FlyBase: updates to the Drosophila melanogaster knowledge base. *Nucleic Acids Res* **49**:D899–D907. doi:10.1093/nar/gkaa1026

Leeuwen J van, Boone C, Andrews BJ. 2017. Mapping a diversity of genetic interactions in yeast. *Curr Opin Syst Biology* **6**:14–21. doi:10.1016/j.coisb.2017.08.002

Lo PCH, Frasch M. 2001. A role for the COUP-TF-related gene seven-up in the diversification of cardioblast identities in the dorsal vessel of Drosophila. *Mechanisms of development* **104**:49–60.

Mann T, Bodmer R, Pandur P. 2009. The Drosophila homolog of vertebrate Islet1 is a key component in early cardiogenesis. *Development (Cambridge, England)* **136**:317–326. doi:10.1242/dev.022533

Panta M, Kump AJ, Dalloul JM, Schwab KR, Ahmad SM. 2020. Three distinct mechanisms, Notch instructive, permissive, and independent, regulate the expression of two different pericardial genes to specify cardiac cell subtypes. *Plos One* **15**:e0241191. doi:10.1371/journal.pone.0241191

Perrin L, Monier B, Ponzielli R, Astier M, Sémériva M. 2004. Drosophila cardiac tube organogenesis requires multiple phases of Hox activity. *Developmental biology* **272**:419–431. doi:10.1016/j.ydbio.2004.04.036

Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, Vilo J. 2019. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res* **47**:W191–W198. doi:10.1093/nar/gkz369

Reim I, Frasch M. 2009. Genetic and Genomic Dissection of Cardiogenesis in the Drosophila Model. *Pediatric cardiology* **31**:325–334. doi:10.1007/s00246-009-9612-1

Reim I, Frasch M. 2005. The Dorsocross T-box genes are key components of the regulatory network controlling early cardiogenesis in Drosophila. *Development (Cambridge, England)* **132**:4911–4925. doi:10.1242/dev.02077

Rotstein B, Paululat A. 2016. On the Morphology of the Drosophila Heart. *Journal of cardiovascular development and disease* **3**:15–14. doi:10.3390/jcdd3020015

Salazar JL, Yamamoto S. 2018. Molecular Mechanisms of Notch Signaling. *Adv Exp Med Biol* **1066**:141–185. doi:10.1007/978-3-319-89512-3_8

Schönbauer C, Distler J, Jährling N, Radolf M, Dodt H-U, Frasch M, Schnorrer F. 2011. Spalt mediates an evolutionarily conserved switch to fibrillar muscle fate in insects. *Nature* **479**:406–409. doi:10.1038/nature10559

Schroeder AM, Allahyari M, Vogler G, Missinato MA, Nielsen T, Yu MS, Theis JL, Larsen LA, Goyal P, Rosenfeld JA, Nelson TJ, Olson TM, Colas AR, Grossfeld P, Bodmer R. 2019. Model system identification of novel congenital heart disease gene candidates: focus on RPL13. *Hum Mol Genet* **28**:3954–3969. doi:10.1093/hmg/ddz213

Seyres D, Ghavi-Helm Y, Junion G, Taghli-Lamallem O, Guichard C, Röder L, Girardot C, Furlong EEM, Perrin L. 2016. Identification and in silico modeling of enhancers reveals new features of the cardiac differentiation network. *Development (Cambridge, England)* **143**:4533–4542. doi:10.1242/dev.140822

Sifrim A, Hitz M-P, Wilsdon A, Breckpot J, Turki SHA, Thienpont B, McRae J, Fitzgerald TW, Singh T, Swaminathan GJ, Prigmore E, Rajan D, Abdul-Khaliq H, Banka S, Bauer UMM, Bentham J, Berger F, Bhattacharya S, Bu'Lock F, Canham N, Colgiu I-G, Cosgrove C, Cox H, Daehnert I, Daly A, Danesh J, Fryer A, Gewillig M, Hobson E, Hoff K, Homfray T, Study I, Kahlert A-K, Ketley A, Kramer H-H, Lachlan K, Lampe AK, Louw JJ, Manickara AK, Manase D, McCarthy KP, Metcalfe K, Moore C, Newbury-Ecob R, Omer SO, Ouwehand WH, Park S-M, Parker MJ, Pickardt T, Pollard MO, Robert L, Roberts DJ, Sambrook J, Setchfield K, Stiller B, Thornborough C, Toka O, Watkins H, Williams D, Wright M, Mital S, Daubeney PEF, Keavney B, Goodship J, Consortium U, Abu-Sulaiman RM, Klaassen S, Wright CF, Firth HV, Barrett JC, Devriendt K, FitzPatrick DR, Brook JD, Study DDD, Hurles ME. 2016. Distinct genetic architectures for syndromic and nonsyndromic congenital heart defects identified by exome sequencing. *Nature Genetics* **48**:1060–1065. doi:10.1038/ng.3627

Theis JL, Vogler G, Missinato MA, Li X, Nielsen T, Zeng X-XI, Martinez-Fernandez A, Walls SM, Kervadec A, Kezos JN, Birker K, Evans JM, O'Byrne MM, Fogarty ZC, Terzic A, Grossfeld P, Ocorr K, Nelson TJ, Olson TM, Colas AR, Bodmer R. 2020. Patient-specific genomics and cross-species functional analysis implicate LRP2 in hypoplastic left heart syndrome. *Elife* **9**:e59554. doi:10.7554/elife.59554

Tong AHY, Lesage G, Bader GD, Ding H, Xu H, Xin X, Young J, Berriz GF, Brost RL, Chang M, Chen Y, Cheng X, Chua G, Friesen H, Goldberg DS, Haynes J, Humphries C, He G, Hussein S, Ke L, Krogan N, Li Z, Levinson JN, Lu H, Ménard P, Munyana C, Parsons AB, Ryan O, Tonikian R, Roberts T, Sdicu A-M, Shapiro J, Sheikh B, Suter B, Wong SL, Zhang LV, Zhu H, Burd CG, Munro S, Sander C, Rine J, Greenblatt J, Peter M, Bretscher A, Bell G, Roth FP, Brown GW, Andrews B, Bussey H, Boone C. 2004. Global Mapping of the Yeast Genetic Interaction Network. *Science* **303**:808–813. doi:10.1126/science.1091317

Ward EJ, Skeath JB. 2000. Characterization of a novel subset of cardiac cells and their progenitors in the Drosophila embryo. *Development (Cambridge, England)* **127**:4959–4969.

Yi P, Johnson AN, Han Z, Wu J, Olson EN. 2008. Heterotrimeric G Proteins Regulate a Noncanonical Function of Septate Junction Proteins to Maintain Cardiac Integrity in Drosophila. *Developmental cell* **15**:704–713. doi:10.1016/j.devcel.2008.10.001

Zinzen RP, Girardot C, Gagneur J, Braun M, Furlong EEM. 2009. Combinatorial binding predicts spatio-temporal cis-regulatory activity. *Nature* **462**:65–70. doi:10.1038/nature08531

Tables

ion channel	CB	CM	*odd*-PC	*tin*-PC
Ca-alpha1D		x		
Ca-beta		x		x
RyR	x	x		
fwe	x	x	x	x
piezo		x		x
GluRIIA		x		
GluRIIB		x		
GluRIIC		x		
GluRIID	x	x		
Clic	x	x	x	x
Clc-b	x	x	x	x
CG3638	x	x	x	x
Sh		x		x
Shal		x		
slo		x		
SK		x		
pain				x
nAChRbeta3	x			
sand	x	x		
CG34396	x	x		x

Table 1. Expression of ion channels in the late embryonic heart. A list of *Drosophila melanogaster* ion channels that are detected in sorted heart cells.

gene	Jin et al. 2013 (early)	Jin et al. 2013 (late)	Junion et al. 2012 (6-8h)
hbs			x
ct	*x*	x	x
CG9650	*x*		x
Fas2			x
Nrt	*x*	x	x
bap	*x*	x	
scrib			x
CG6356		x	x
fng			
emc		x	x
RhoGAP18B			x
Inx2			x
Lac			x
Hesr			
Wnt4			x
Pvf3			x
Ptp10D			x
CG10479	*x*	x	x
CG4393			x
mspo			
CG42340			
DOR			x
sano			
CG2150			
E(spl)m3-HLH			x

Table 2. Tin-PC marker genes and present-calls in Tinman ChIP-on-chip datasets.

Figure Legends

Figure 1. Isolation and sorting of cardiac cells for single-cell sequencing.
(**A**) Examples of embryos homozygous for *mid^{E19}::GFP* at stage 15 and 16 of embryonic development. Note the strong bilateral expression in Tinman-positive cardioblasts, and weaker GFP expression outside the heart. (**A'**) Confocal images of a *mid^{E19}::GFP* heterozygous embryo that also expresses a membrane RFP using *tinD*-Gal4 crossed to UAS-*CD4::tdTomato*. Note the low non-cardiac expression (arrowheads) only visible with high contrast settings. (**B**) Experimental workflow overview. (**B'**) Example scatterplot of cells during cell sorting, obtained from heterozygous *mid^{E19}::GFP* embryos. Two clusters of GFP positive cells can be separated from the bulk of cells, only differing in GFP intensity (low GFP vs. high GFP).

Figure 2. Single-cell sequencing of *mid^{E19}*::GFP cells
(**A**) UMAP plot of the expression profiles from 12,000 sorted GFP+ cells. Highlighted are cardiac cells (cardioblasts and pericardial cells, encircled), as well as engulfed muscle cells (arrowhead). Note the presence of various other cell types including neurons, fat body and muscle cells. (**B**) Examples of expression patterns used to identify clusters.

Figure 3. Analysis of marker gene expression in sorted embryonic heart cells.
(**A**) Expression levels of genes and signaling pathways with known cardiac. (**B**) UMAP plot of subclustered sorted heart cells. (**C**) Examples of markers for each cardiac and pericardial cell type. (**D**) *Mhc* is strongly expressed in later cardiac cells, and the expression of *Mhc* and *Myc* in CBs and CMs is anticorrelated (CB: *Myc* high / *Mhc* low; CMs: *Myc* low / *Mhc* high).

Figure 4. Characterization of Tin-PCs with novel marker genes.
(**A**) Pericardial cells (red) are localized laterally and ventrally to the heart CBs (green). (**A'**) Tin-PCs localize ventrally to the closing heart tube and express high levels of the cell adhesion marker Neurotactin (cyan). (**B**) These PCs can be distinguished from Odd-PCs in their ventral positioning as well as gene expression of other novel markers (following panels). (**C-F, b**) The homeobox transcription factor Cut (red) is a marker for Tin-PCs which are localized ventrally to Mef2-positive CBs (blue). (**G,G', c**) The GPCR *Corazonin receptor* (*CrzR*) is a marker for Tin-PCs. (**H-H''', d**) Expression pattern of *Lachesin* (*Lac*) showing high expression in Tin-PCs and weaker expression in Odd-PCs. *Lac* is already expressed in Tin-PCs at stage 15 (**H**, insets).

Figure 5. Identification of muscle cells engulfed by plasmatocytes.
(**A**) Plasmatocytes expressing *SPARC* and *Pvr* marker genes. Note the subset of cells co-expressing Mhc and wupA representing a cell with 'dual identity' (arrowhead). (**B**) Still-image from time-lapse video showing a phagocytotic event by a plasmatocyte and two other GFP-positive cells.

Figure 6. Dynamic gene expression during somatic muscle differentiation.
(**A**) Subset of marker genes expressed in developing muscle cells (as identified by *lmd* expression). Note contrasting expression of the early marker *twi* and late marker *Mhc*. *Notch*, as well as its direct targets are also expressed in a gradient that is inverse to *Mhc* (E(spl)-C genes *m4*, *malpha* and Brd-C gene *Ocho*). E(spl)m7 is expressed similar to *lmd* throughout these cells. (**B**) UMAP plot of a sample subset of cells with identified clusters including differentiating muscle cells (encircled). (**C**) RNA velocity

vectors show a directionality within the differentiating muscles as well as in other clusters indicating temporal rather than spatial differences between these cells. (**D**) RNA velocity derived example of *Mhc* relative expression of spliced (s) vs. unspliced (u) *Mhc* mRNA. *Mhc* appears upregulated at mid-differentiation (red, right panel) and is downregulated at the end of differentiation (blue, right panel).

Supplemental Figure 1. Dot plot showing marker gene expression for each identified cluster across all cell types identified in the scRNAseq dataset.

Supplemental Figure 2. Gene expression patterns known myoblasts and myocyte differentiation and specification markers as well as *Notch* and Notch signaling targets of the E(spl)-C and Brd-C.

Supplemental Figure 3. eGFP expression visualized by HCR in situ (left) and their representation in the single-cell dataset (right).

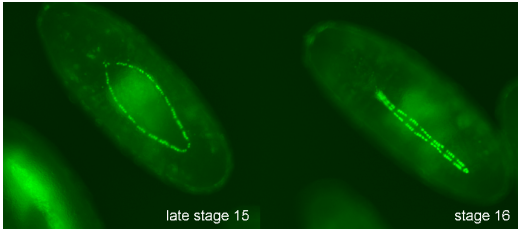
Supplemental Figure 4. GO terms overrepresented in cardiomyocytes (CMs).

Supplemental Figure 5. GO terms overrepresented in cardioblasts (CBs).

Supplemental Movie 1. Time-lapse of GFP and RFP expression in the closing heart tube. Genotype: *tinD-Gal4; mid*^{E19}*::GFP x UAS-CD4::tdTom*

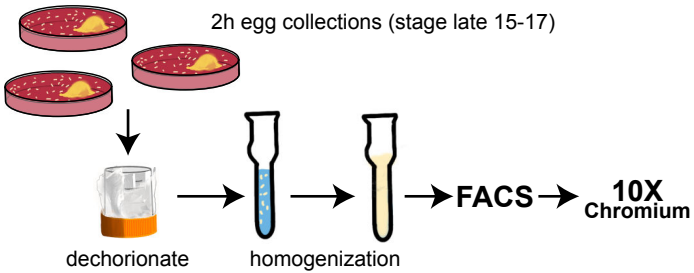
Supplemental Movie 2. Time-lapse of the closing heart tube and a plasmatocyte engulfing other cells. Genotype: *tinC*Δ*4-Gal4 x UAS::act5C::GFP*

A

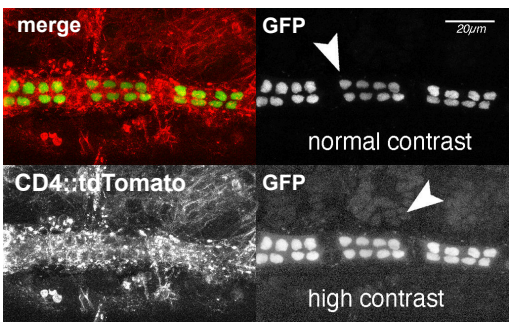


homozygous *mid*^{E19}*::GFP* expression

B

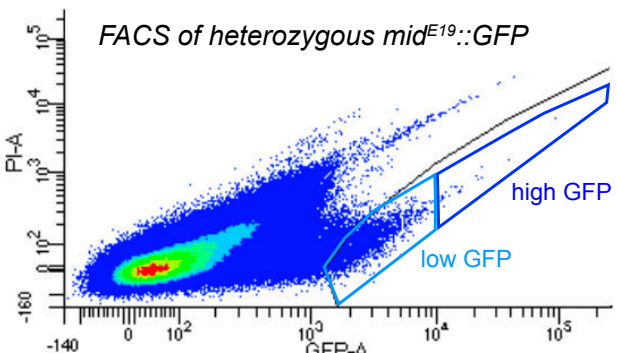


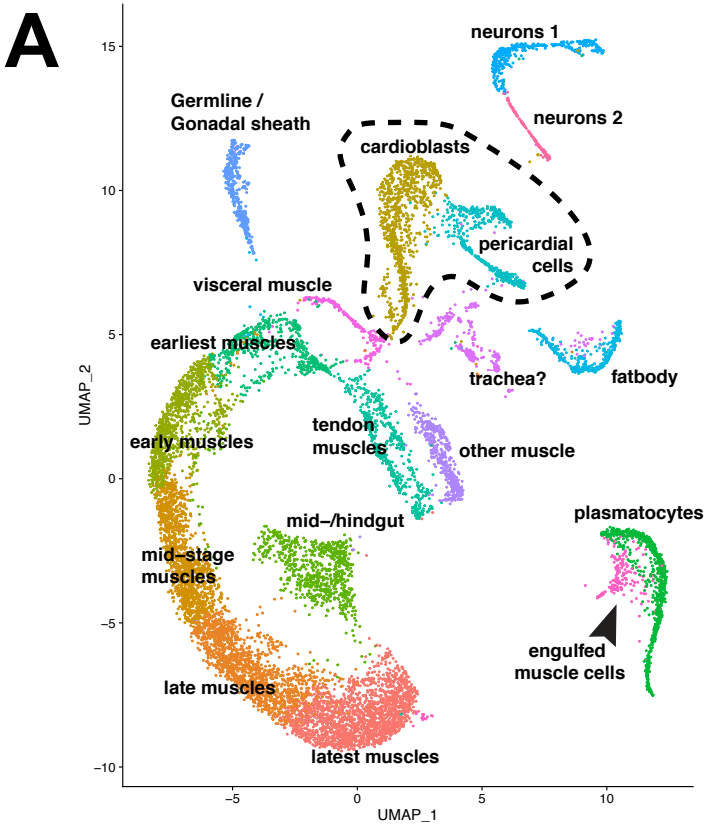
A'

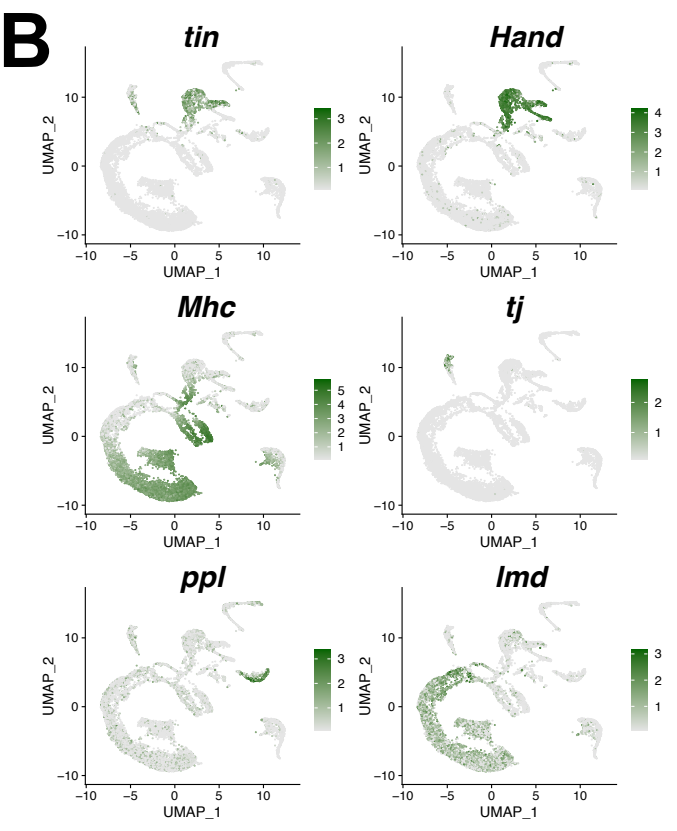


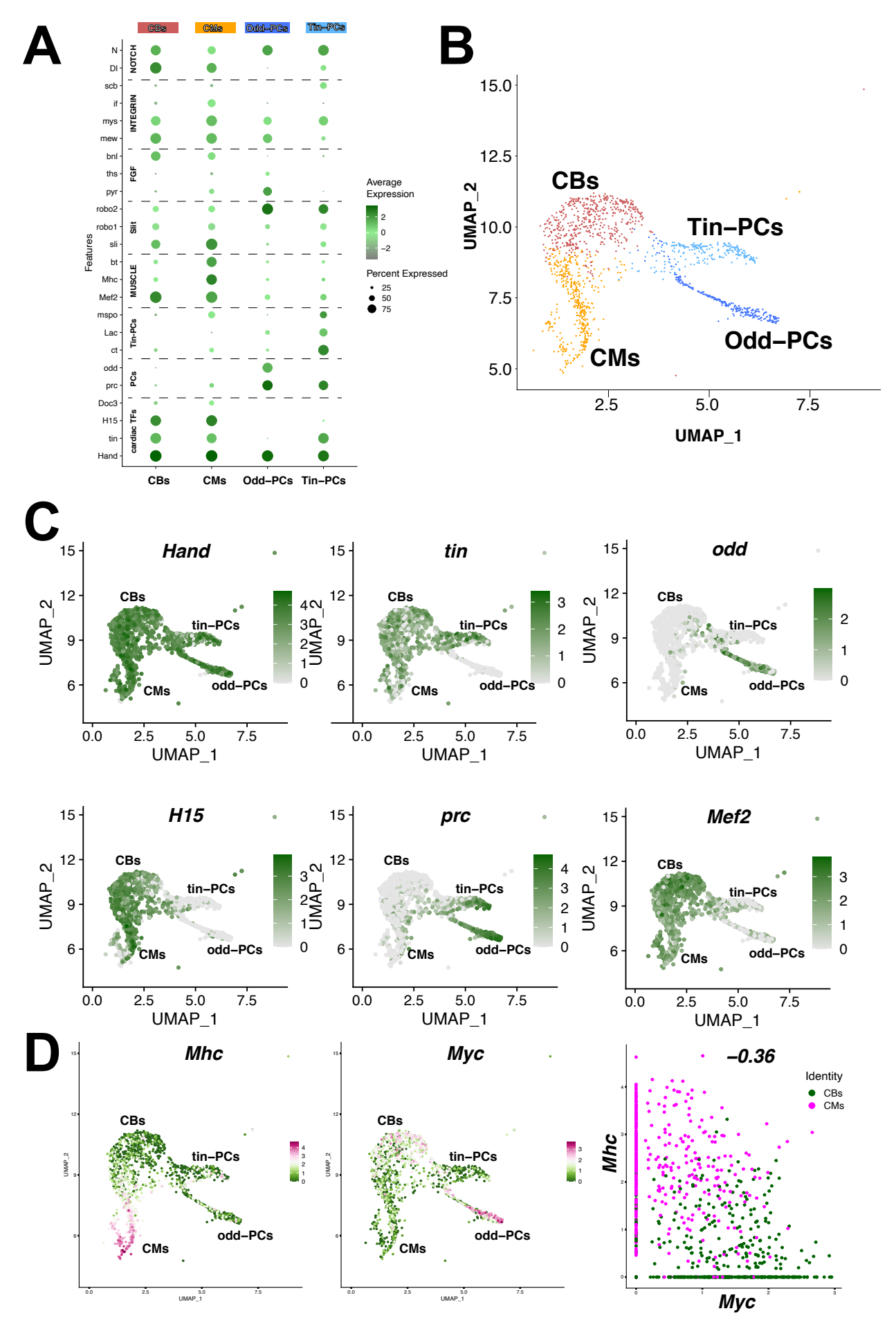
tinD-Gal4 > CD4:tdTomato + heterozygous mid^{E19}*::GFP*

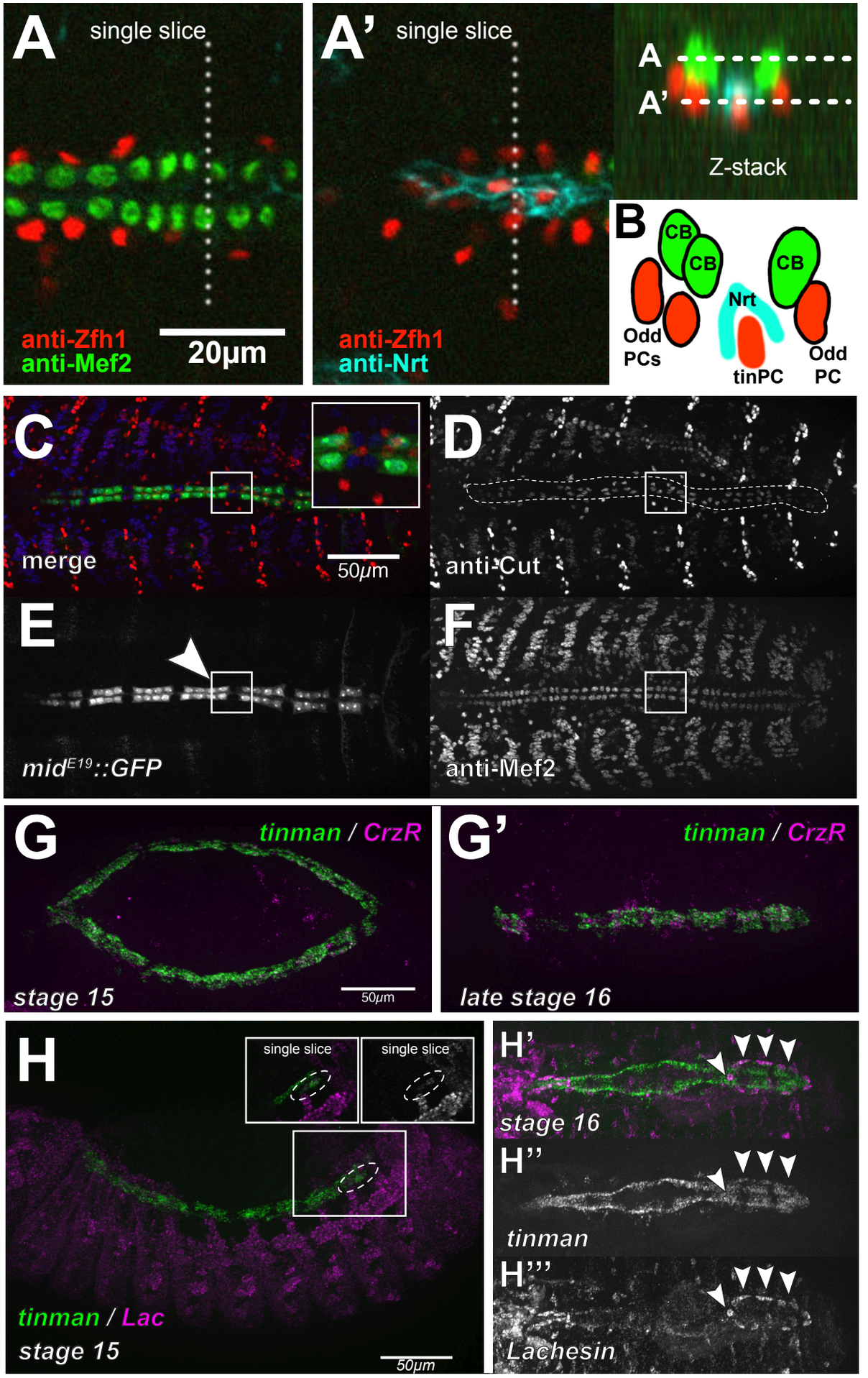
B'

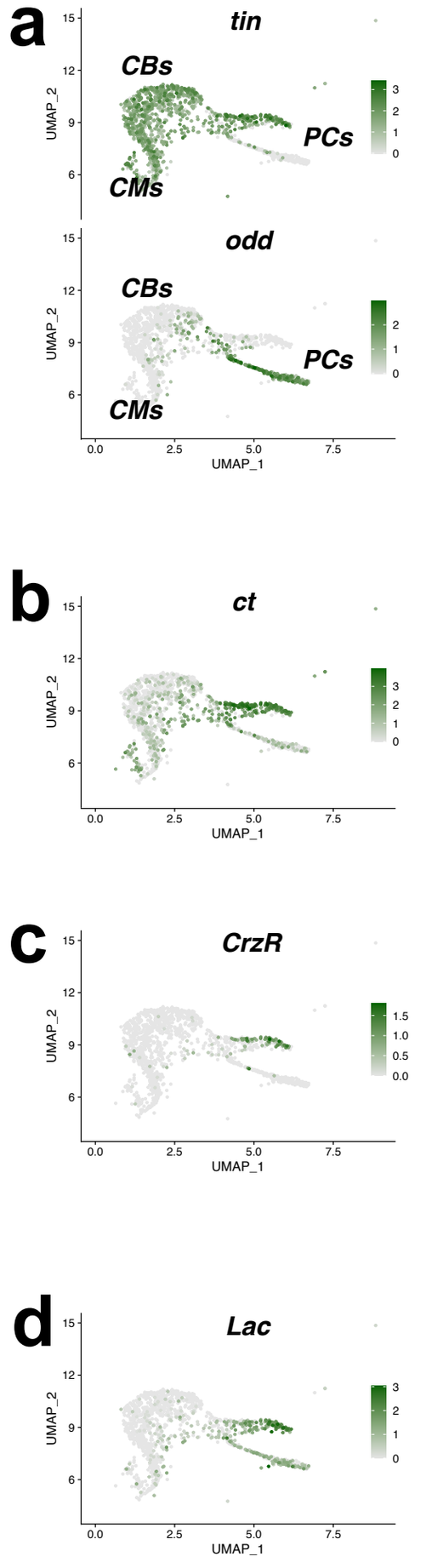


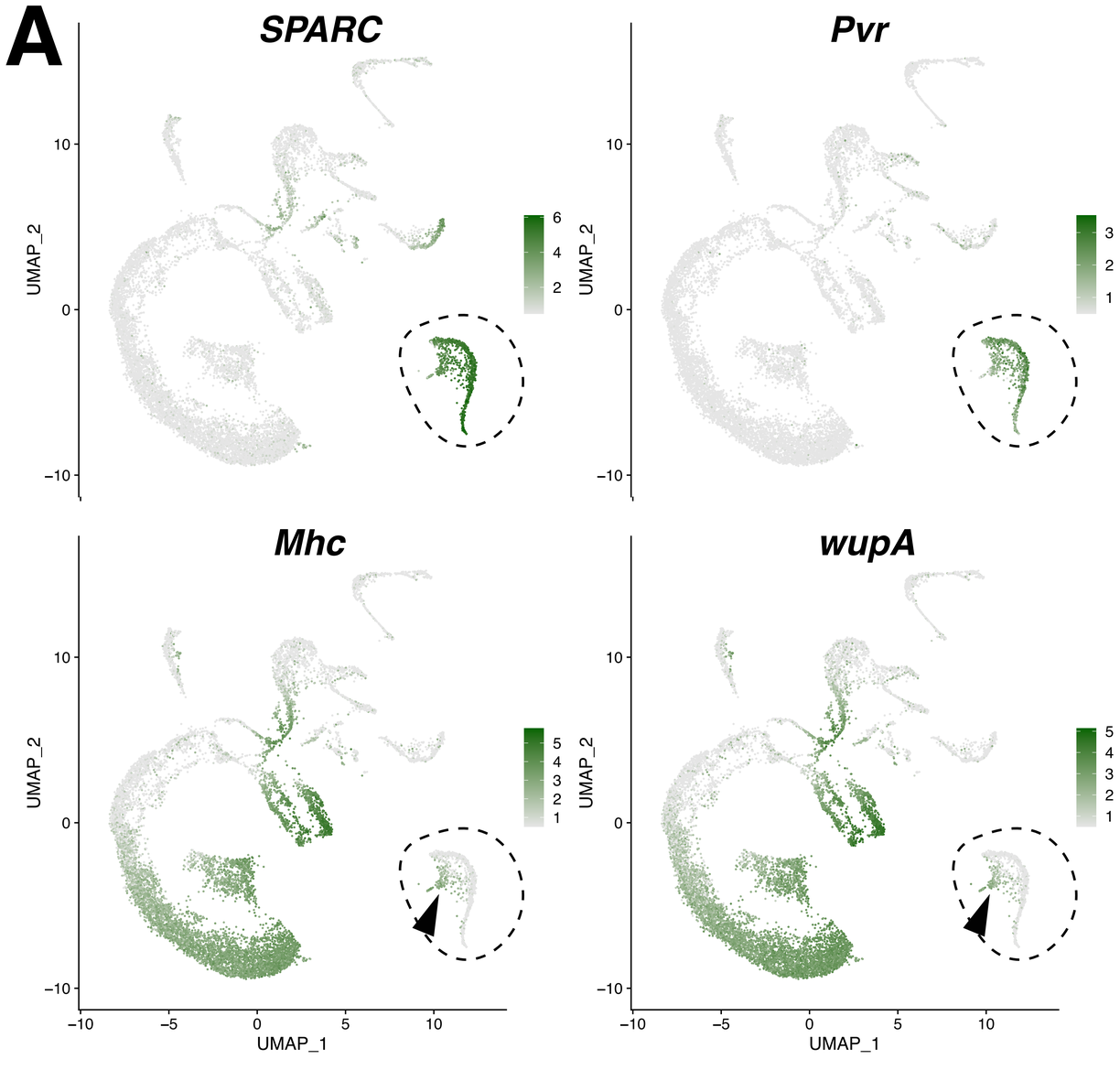

A
neurons 1
neurons 2
Germline /
Gonadal sheath
cardioblasts
pericardial
cells
visceral muscle
earliest muscles
trachea?
fatbody
early muscles
tendon
muscles
other muscle
mid–/hindgut
plasmatocytes
mid–stage
muscles
engulfed
muscle cells
late muscles
latest muscles
UMAP_1
UMAP_2

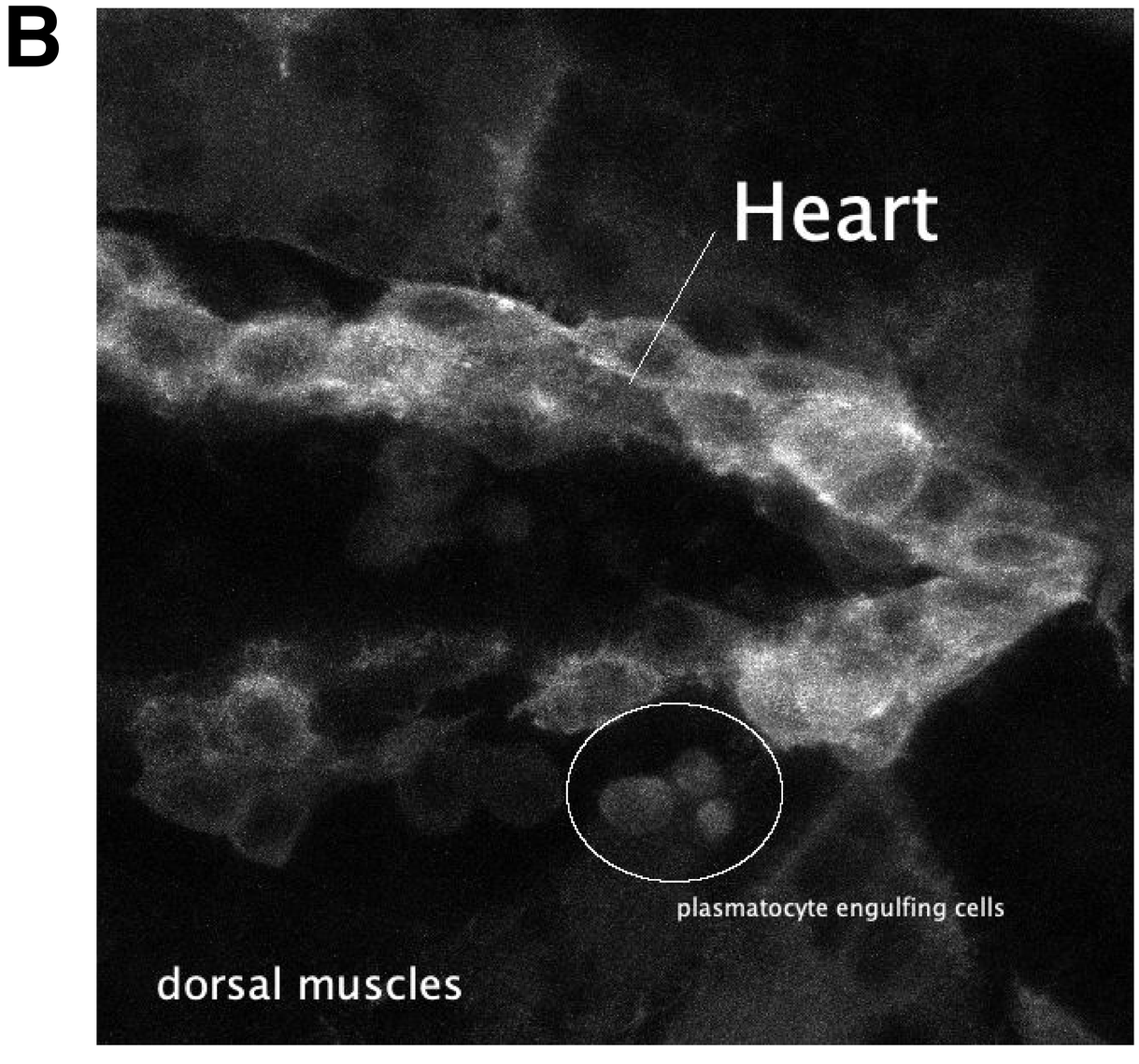

B
tin
Hand
Mhc
tj
ppl
lmd
UMAP_1
UMAP_2

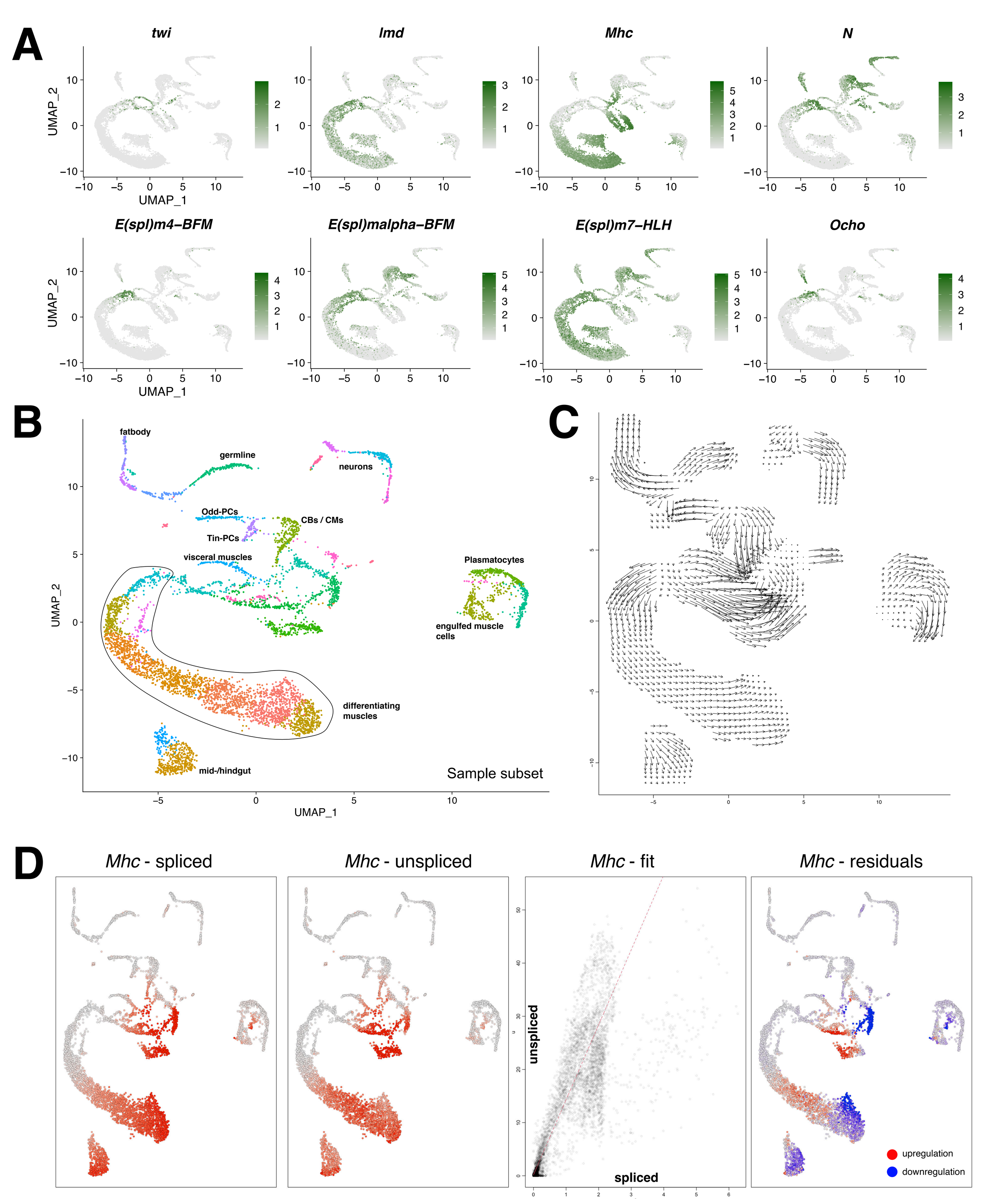

A
CBs
CMs
Odd-PCs
Tin-PCs
NOTCH
INTEGRIN
FGF
Slit
MUSCLE
Tin-PCs
PCs
cardiac TFs
N
Dl
scb
if
mys
mew
bnl
ths
pyr
robo2
robo1
sli
bt
Mhc
Mef2
mspo
Lac
ct
odd
prc
Doc3
H15
tin
Hand
Features
Average Expression
Percent Expressed
B
UMAP_2
UMAP_1
CBs
Tin-PCs
Odd-PCs
CMs
C
Hand
tin
odd
H15
prc
Mef2
tin-PCs
odd-PCs
D
Mhc
Myc
−0.36
Identity
CBs
CMs


A
single slice
anti-Zfh1
anti-Mef2
20µm
A'
single slice
anti-Zfh1
anti-Nrt
A
A'
Z-stack
B
CB
CB
CB
Nrt
Odd PCs
tinPC
Odd PC
C
merge
50µm
D
anti-Cut
E
mid^E19::GFP
F
anti-Mef2
G
tinman / CrzR
stage 15
50µm
G'
tinman / CrzR
late stage 16
H
single slice
single slice
tinman / Lac
stage 15
50µm
H'
stage 16
H''
tinman
H'''
Lachesin


a
tin
CBs
PCs
CMs
odd
CBs
PCs
CMs
b
ct
c
CrzR
d
Lac
UMAP_1
UMAP_2


A
SPARC
Pvr
Mhc
wupA
UMAP_1
UMAP_2
10
0
−10
−5
5
6
4
2
3
1


B
Heart
plasmatocyte engulfing cells
dorsal muscles


A
twi
lmd
Mhc
N
E(spl)m4–BFM
E(spl)malpha–BFM
E(spl)m7–HLH
Ocho
UMAP_1
UMAP_2
B
fatbody
germline
neurons
Odd-PCs
CBs / CMs
Tin-PCs
visceral muscles
Plasmatocytes
engulfed muscle cells
differentiating muscles
mid-/hindgut
Sample subset
UMAP_1
UMAP_2
C
D
Mhc - spliced
Mhc - unspliced
Mhc - fit
Mhc - residuals
unspliced
spliced
upregulation
downregulation