A single cell transcriptomic and clonal analysis depicts valvulogenesis

Batoul Fahrat[1], Ignacio Bordeu[2,3], Bernd Jagla[4], Hugo Blanc[5], Jean Livet[6], Benjamin D Simons[2,3,7], Emmanuel, Beaurepaire[5], and Michel Pucéat[1]

1 INSERM U1251/MMG Aix-Marseille University, Marseille, France

2 *DAMTP, University of Cambridge, Cambridge, UK*
3 *The Wellcome Trust/Cancer Research UK Gurdon Institute, University of Cambridge, Cambridge, UK*
4 Pasteur Institute UtechS CB & Hub de Bioinformatique et Biostatistiques, C3BI, Paris
5 Laboratory for Optics and Biosciences, Ecole polytechnique, CNRS, INSERM, IP Paris, Palaiseau, France
6 Sorbonne Université, INSERM, CNRS, Institut de la Vision, Paris, France
7The Wellcome Trust/Medical Research Council Stem Cell Institute, University of Cambridge, Cambridge, UK

Mitral valve prolapse is often associated with several congenital heart malformations. Endocardial cells of the atrioventricular canal undergo endothelial to mesenchymal transition (EndMT) to give rise to the different mitral valvular cells. Our understanding of the identity and fate decisions of these endocardial cells during development is still lacking to date. Here, using single cell RNA sequencing (scRNA-seq) of genetically labeled AVC endocardial cells at E9.5 in murine heart, we uncovered and genetically characterized a restricted population of pro-EndMT cells and further distinct populations of valve progenitors that contribute to different endothelial and interstitial valvular cells in E16.5 embryonic and P0 postnatal murine mitral valve. Moreover, by genetically labeling endocardial cells using CAG$^{ERT2Cre+/-}$/Brainbow$^{+/-}$ mice, at embryonic stage E8.5 prior to EndMT, we observed specific modes of growth of endocardial derived clones in E16.5 embryonic mitral valve. Collectively, our data reveal the identity of specified endocardial cells and distinct types of clonal contribution of these cells to the formation of the mitral valve. They further open the path toward new cell targets for a therapeutic approach of congenital and acquired valve diseases.

INTRODUCTION

During evolution, as organisms have adapted to new environments, the heart has grown in size and acquired a complex structure, ranging from a simple contractile vessel in Amphoxius to a four chambers organ in mammals, which required the development of valves to ensure unidirectional blood flow.

Malformation of valves accounts for 10% of congenital heart diseases[1]. Mutations of a few genes have been reported as responsible for valves defects. These include genes encoding cytoskeleton proteins fibrillin and filamin A[2,3], TGFβ receptors and downstream signaling smad pathway[4] and more recently Dachsous, potentially LMCD1, tensin and DZIP[5,6]. Altogether, mutations of these genes account for less than 5% of genetic valvulopathies. This suggests that besides genetics, other biological processes such as defects in cell fate determination and lineage specification and/or in cell migration might be responsible for valve diseases.

Valve development is a complex biological process involving several cell lineages first specified and determined by specific gene transcriptional events, influenced by many extrinsic factors and cellular events including growth factors, cell-cell and cell-matrix interactions and mechanical phenomenons[7]. Valves form as early as at embryonic stages E9.0-E10 in mouse embryo and develop in two discrete cardiac regions, namely the atrioventricular canal (AVC) and the outflow tract (OFT). The AVC valves include the mitral and tricuspid valves, while the OFT or semilunar valves include the aortic and pulmonary valves[5]. The mitral valve is the most affected valve in cardiac congenital diseases.

Valve formation commences with the delamination of endocardial cells that get separated from the myocardium by a cardiac jelly. This event is followed by endocardial cell transition toward a mesenchymal phenotype. Cell proliferation within an extracellular matrix forms two cushions. The endothelial-to-mesenchymal transition (endMT) is triggered by a series of morphogenetic factors and signaling pathways including the TGFβ/BMP-smad, Wnt/TCF, Notch4 secreted by the myocardium and NFATc mediated pathways[8] likely governed by a set of specific genes and transcriptional pathways. The most intriguing feature of this biological process is that it is restricted within a set of a few endocardial cells. Besides the signaling pathways, little is known about the biological processes and transcriptional networks required for the cell to initiate local EndMT. The

specific transcriptomic signature or any biological property that confers to the cell a competency for EndMT is still unknown.

Afterwards, a series of morphogenetic events allows the cushion cells to further differentiate into specific cell types (i.e; fibroblast, smooth muscle cells, chondrogenic cells), to proliferate and distribute to build up the valve leaflets. How early does lineage restriction occur and how they proliferate in a clonal dependent manner and distribute within the valve leaflet remain questionable[9].

To address these open questions, we use a single cell-RNA sequencing approach combined with clonal analysis to uncover the origin and transcriptomic signature of valvular endocardial cells, together with the behavior of the valvular endothelial (VEC) and interstitial (VIC) cells during formation of valve leaflets of the mitral valve.

We report that endocardial cells are a heterogeneous cell population within the AVC. A subset of these cells[10] which feature a unique gene profile, are able to acquire EndMT competency and commit toward a specific valvular cell type. Furthermore, based on our clonal and histological analyses, we could estimate the number of self-renewing progenitors required in order to building up a valve leaflet, and we observed multiple behaviors of VIC within the leaflet allowing for the formation of a functional valve.

RESULTS

Single cell transcriptomics reveals cell specification in the atrioventricular canal

To interrogate the endocardial cell heterogeneity within the AVC, we first used a single cell-RNA sequencing approach. The AVC was carefully dissected out from 60 19-22 somites (E9.5) embryos from Rosa26^{tdTomato} females bred with Tie2^{Cre} males.

Cells were enzymatically dissociated, FACS sorted, loaded into the chromium (*10X genomics*) and 1500 cells were sequenced and 1200 used in the analysis. Cell clustering illustrated in the UMAP plot (Fig. 1a) revealed heterogeneity of the whole cell population. The AVC cell population featured an average of 5530 genes/cell and comprised one population of endocardial cells, two populations of hybrid cells undergoing EndMT, a subpopulation of late hybrid /mesenchymal cells as well as chondro-osteogenic progenitors and hemogenic cell. A small cluster of macrophages was also observed (Fig.1a).

The heatmap (Fig. 1b) further revealed a broad spectrum of gene expression reminiscent of future VIC and VEC cell populations.

More specifically, we found one subpopulation of *eng*$^{+}$, *pecam*$^{+}$, *egfl7*$^{+}$, *cdh5*$^{+}$, *nppr3*$^{+}$ endocardial cells (cluster 3), a subpopulation of *zeb2*$^{+}$, *prrx1/2*$^{+}$, *twist1,2*$^{+}$, *msx1*$^{+}$, *cldn7*$^{+}$,*epcam*$^{+}$, *ebfn1*$^{+}$, *efna2*$^{+}$ cells starting the transition from endothelial to mesenchymal fate (early hybrid, cluster 5), two subpopulations^{-} of *Tagln*$^{+}$,*col2a1*$^{+}$ *and twist1,2*$^{-}$ hybrid cells (clusters 0 and 2), and a subpopulation of *Tagln*$^{+}$, *col2a1*$^{+}$, *and Alcam*+ late hybrid cells transitioning toward mesenchymal cells. We also observed subpopulations of *cdh11*$^{+}$, *sox9*$^{+}$, *sparc*$^{+}$, *Tbx18*+, *col3a1*$^{+}$, *Cthrct*$^{+}$, *GATA6*$^{+}$, *Maoa*$^{+}$ chondro-osteoblastic progenitor cells (cluster 4), of *tal1(scl)*$^{+}$,*Gata1*$^{+}$,*Nfe2*$^{+}$, *Klf1*$^{+}$, *Atf1*$^{+}$ hemogenic cells (cluster 1) and a small cluster of *Plek*$^{+}$, *Tyrobp*$^{+}$, *Gata1*$^{+}$, *Nef2*$^{+}$, *Fcer1g*$^{+}$ macrophages .

EndMT is the first step of valvulogenesis. We thus first addressed the question of the EndMT competency of these cells. Epithelial-to-mesenchymal transition (EMT) is no longer understood as a binary process but has been described as a continuum transition[10]. EndMT is likely to follow such a step-by-step transition from endocardial to mesenchymal cell phenotype. We first interrogated a few genes that we surmised to play a role in EMT and expressed in either/or endocardial cells (marked by *pecam1, eng, cdh5* and *npr3*), early hybrid (marked by *twist1* and *prrx2*) and hybrid cells (markets by *acta2* and *tagln*) (top panels). The panel plots in Figure 1c reveals that only part of endocardial cells expressed *twist1* together with *prrx2* early marker of EMT. Interestingly these cells also expressed TGFβ2 while TGFβ1 was more broadly expressed in endocardial cells. *Prrx2* marked also all early hybrid and hybrid cells. The same cells highly expressing *hand2* and *msx1* both known to play a role in EndMT[11,12] lost Zo-1 an adherens junction gene (*tjp1*). Expression of acetyl Co-A transferase (*acly*), was increased in these pro-EMT cells. A differential gene expression analysis between cluster 3 (endocardial cells) and 5 (early hybrid cells) (extended data 1) further revealed an upregulation of *prrx1, cldn6, cldn7,notch2, cxcl12, cxcl13, pdgfrα, wnt6, wnt2, while cldn5, notch1* and *TGFβ1* were downregulated in early hybrid cells.

To further dissect this transition, we used an inference trajectory approach. Interestingly, this bioinformatics approach suggested that two subpopulations of endocardial cells gave rise first to early *twist1*$^{+}$ hybrid cells. The cells then progressed towards hybrid mesenchymal cells still retaining expression of genes reminiscent of endothelial cells. Three branches emerged from these cells engaged into the process of EMT to give rise to either chondro-osteogenic cells, remain as hybrid cells as well as they

moved towards late hybrid and hemogenic cells (Fig.1c). Macrophages remained aside the trajectory indicating that they were likely contaminating cells attached to others.

More specifically, we looked at signaling pathways that could be involved in EndMT.

BMP2 was expressed only in hybrid cells (cluster 0) while TGF*β*2 was indeed found in both early and late hybrid cells (clusters 0 and 5) but not in endocardial cells which expressed TGF*β*1. While *TGFβR2* receptor was expressed at very low level, *TGFβR1* receptor was expressed in both early and late hybrid cells but only at low level in endocardial cells. Late hybrid cells included cells co-expressing TGFβ1, TGFβ2 and the TGFβR1 (Fig. 2b). The VEGFR *flit1* was expressed only in endocardial cells. The *notch2* ligand and *Jag1* receptor genes were expressed mostly in early hybrid and chondrogenic cells (Fig. 2a). *SHH* was expressed in early hybrid, while its receptor *Ptch1* was more broadly expressed in early hybrid and chondrogenic cells but not in endocardial cells (Fig 2a). *IGF1* was also mostly expressed in endocardial cells, while *IGF2* was expressed in all cells. The purinergic receptors *P2rx4* and *P2yr1* the only purinergic receptors expressed were only expressed in rare cells (Fig 2a and b).

Interestingly, gene encoding enzymes of the fatty acid oxidation process, the acetyl-CoA acyltransferase 2 (*caa2*), the acetyl Co-A transferase (*acly*), the acetyl-CoA decarboxylase (*acss2*), the diacylglycerol O-acyltransferase 1 (*dgat1*), and the enoyl-Co A hydratase (*echs1*) were expressed all along the EMT process and in few endocardial cells (*acly, caa2, acss2*). Their expression increased in early hybrid *twist1*$^{+}$ cells, in *vcam1*$^{+}$ cells and to a low extent in late *itgav*+ hybrid cells (Fig.2b).The glycolytic enzymes enolase (*eno3*), phosphoglycerate mutase *pgam2*, acetyl coA dehydrogenase (*acadl*), the phosphofructokinase (*pfkp*) were co-expressed in the hybrid and late hybrid cells but not in endocardial or early hybrid cells. *Ogdh* encoding the Oxoglutarate Dehydrogenase, a citric acid cycle enzyme that plays a key role in replenishing TCA cycle intermediate was expressed all along the EMT process.

We next picked up proteins encoded by genes of interest in the EndMT process and carried out immunofluorescence experiments on sections of E9.5 embryonic heart. Vcam1 and Twist1, a protein involved in early EMT, featured a salt and pepper pattern in the endocardium within the AVC as shown in the panel plot (Fig 1c). TWIST+ cells also co-expressed Pecam1 and TWIST was still located into the cytosol of cells. MSX1, another marker of EndoMT turned on by Bmp2 signaling from the myocardium was observed in delaminated cells migrating from the endocardium to the cushion. The fourth marker, the ATP-citrate lyase (ACLY) was expressed in some endocardial cells, as well as in cells

migrating within the cushion still expressing *pecam1*, namely, early hybrid cells. The fifth marker VCAM1 was expressed in most cells of the cushion as well as the AVC endocardium. Two other markers for late hybrid EndoMT and mesenchymal cells were validated, ITGAV expressed in endocardial and post-EndoMT cells in the cushion and ITGB3 only expressed in rare delaminated endocardial cells, respectively. MSX1 was also expressed in very few cells of delaminated endocardium (Fig.3). An increase in expression of enzymes involved in fatty acid oxidation together with specific expression of TGFβ2 pathway components confer to *hand2*$^{+}$,*Msx1*$^{+}$ endocardial cells a potential to undergo EMT.

Transcriptomic profile of embryonic and postnatal mitral valve depicts early valvulogenesis

We collected embryos generated from Rosa26^{tdTomato} females bred with Tie2^{Cre} males at E16.5. Hearts were dissected out and the annulus of the mitral valve was carefully cut out, enzymatically dissociated and pooled together from 92 embryos at the same stage of development. The isolated cells were sorted by FACS and altogether 954 Tomato+ cells were then subjected to single cell-RNA sequencing and analysis.

The UMAP revealed that E16.5 valvular cells segregated between VEC expressing *pecam1, cdh5, and eng* and VICs expressing markers of fibroblasts, *postn and dcn* (Fig. 4a). VECs distributed into two clusters including one with cells expressing *wt1,* suggesting that these VECs belong to the parietal leaflet in which epicardial cells contributed to[13] (Fig. 4a,4b; extended data Figure 1).

E16.5 valvular cells segregated from the neonatal valvular cells as well as AVC cells segregated from E16.5 and P1 valvular cells as shown in the UMAP plot (Fig. 4c). The heatmap including AVC E16.5 embryonic and P1 neonatal mitral valve cells revealed the presence of numerous valvular genes (Fig. 4d). *Egfl7, pecam1, cdh5, eng, npr3, nos3, tek, ecscr, apold1, tie1, cd93, thbd, mmrn2, ctla2, gata2 ,erg, and emcn* were expressed in endocardial cells *(*AVC cluster 6*),* and VEC of E16.5 mitral valve (cluster 7) and of neonatal mitral valve (clusters 5 and 8). Interestingly the E16.5 mitral valve and cluster 8 of neonatal mitral valve expressed *CD36* and *fabp4*, two markers of fatty acid metabolism

VICs of both E16.5 and P1 embryonic and neonatal mice (clusters 4 and 1) expressed numerous markers of VIC of adult valve layers (spongiosa, fibrosa and atrialis)[7]. Indeed, we interrogated a list of genes specific of spongiosa *(dpt, postn, thbs2, eln, col3a1,*

col1a1), fibrosa (col1a2, thbs1, bgn, Dkk3) or atrialis enriched in elastin *(eln).* All genes were expressed in most cells of both E16.5 and neonatal mitral valves in clusters 4 (E16.5) and 1 (Neonatal) even in cells that did not cluster in a layer-specific manner. Clusters 3 and 16 included cells from the endothelial layer of the mitral valve leaflets at E16.5 or in neonates Immune cells including macrophages which expressed *tyrobp, plek, ctss, laptm5, lcp1, cd83, gp49a, aif1, lyz2 and cx3cr1,* were found in cluster 10 (Fig 4d).

We then applied linear trajectory inference via pseudotime analysis (using Elpigraph) to track the steps of valvulogenesis (from day E9.5 until post-birth). The trajectory described the multistep process by which AVC cells at E9.5 transition towards valvular cells. This analysis revealed a subset of *pecam1+, eng+, cdh5+* endocardial cells that give rise to the endothelial layers (VEC) of both E16.5 and neonatal valves. The endocardial cells also contributed to subsets of VICs. One subset at E16.5 led to the VICs in neonatal valve. The endocardial cells also underwent EMT to give rise to chondrogenic and hemogenic cells (Fig. 4e). We more specifically interrogated ECM-related genes in the three single cell RNA-seq datasets, (AVC, mitral valve at E16.5 and at neonatal stages). The heatmap in extended data Figure 2 showed that clusters 1 and 4 (including only E16.5 and neonatal valvular cells) were enriched in collagen genes (*col1a1, col1a2, col3a1)*, reminiscent of the fibrosa. Clusters 5, 7 and 8 including E16.5 and neonatal mitral valvular cells were enriched in proteoglycans including perlecan, *hspg2*, versican (*vcan),* lumican *(lum),* and elastin, reminiscent of the spongiosa and atrialis. AVC cells from clusters 0, 2, 3 and 6 specifically expressed the galactose kinase (*galk1*), an enzyme required to synthetize proteoglycans, as well as the hyaluronan receptor CD44 and the hyaluronan synthase 2 (*has2*).

We further used Tempora[14] another cell trajectory inference using biological pathways information.

This trajectory first revealed that cells differentiate step by step at each stage of development to finally participate in post-natal valve leaflets.

The endocardial cells (cluster 4) contributed directly to the E16.5 VEC that underwent EMT to give rise to subpopulations of VICs in neonatal valve (clusters 12, 16). They also participated through hybrid cells (cluster 7) to the two clusters of VECs of neonatal valve (clusters 3 and 9). Cluster 12 received cells from the chondrogenic AVC cluster 0. A small cluster of neonatal VICs (cluster 16) resulted from EndMT from E16.5 VEC (cluster 6). Neonatal VICs (cluster 8) received a minor contribution of AVC hemogenic cells (cluster 1). The significant Go terms revealed by the Tempora trajectory inference further pointed the

role of the TGFβ family and smad transducers, Wnt, Notch and Nfκb pathways as well as the PDGRα and cilia formation[15]. Integrin and Rac signaling contributed to the process of valvulogenesis. Dynamic of smooth muscle cells emerged as an important process within the developing valve (extended data Figure 3). The vertex plot of the trajectory (extended data Figure 4) confirms the different contributions of cells within the developing valve. It further highlighted that VEC in E16.5 and neonatal mitral valve feature a slightly different origin within the AVC cells.

Progenitor clones fragment and proliferate along the leaflets axis

The embryonic and cell lineage at the origin of endocardial cells is diverse, which could explain their transcriptomic heterogeneity and the competency to EndMT only restricted to specific cell types. We then questioned how post-EMT cells further acquire their fate and distribute within the valve leaflet. To address these questions, we set up a lineage tracing experiment using CAG^{CreERT2} male bred with CYTBOW females. The Cre recombinase was induced at E8.5 using a low dose of tamoxifen (20 µg/g mouse), which allowed induction at low cell density (extended data Figure 5 a,b). Embryos were first collected at E9.5 to check the number of labelled endocardial cells. Extended data Figure 3a shows the labelling of a single endocardial cell at the proposed tamoxifen dosage. Phosphohistone 3 (PH3) immunostaining was used to score the rate of cell division of AVC endocardial cells at E8.5. A low cell proliferation rate was observed at E9.5 (extended data Fig. 3b), thus limiting cell division right after recombination.

Embryos were collected at E16.5 and labelled cells within the mitral valve leaflets were scored by imaging heart sections. We first monitored the labeled cells in the mitral valve via two-photons microscopy of whole transparized hearts. Such a microscopic approach revealed that labeled cells formed elongated clusters that distributed primarily along the long axis of the leaflet (extended data Fig. 6a). This observation resulted in a high correlation between the 2D-projected area of cell clusters and their 3d volume (Pearson's $r=0.91$, $p=0.01$, extended data Fig. 6b), which allowed us to base our clonal analysis on the 2D-projected area of the mitral valve leaflets (Fig. 6ab).

Within a leaflet we could observe that putative clones of a given color where fragmented into smaller clusters of cells. To infer whether a group of clusters of cells belonged to a single clone, we followed a biostatistical approach[16] that allowed us to estimate the likelihood that a group of clusters was monoclonal (see Methods for details).

Applying this biostatistical approach to the full clonal data, we could estimate both the overall induction frequency, pN (the frequency at which a single leaflet progenitor is induced times the total number of leaflet precursors), and the clone fragmentation rate f (Fig. 5b and extended data Fig. 7). The overall induction frequency was found to be pN=0.5±0.3 (95% C.I.) by considering the combined data from both mural and septal leaflets, while we estimated the fragmentation rate f independently of each leaflet, which was f=1.6±0.1 for the septal, and a slightly lower rate of f=1.1±0.2 for the mural leaflet. We then assessed which groups of cells clusters of a given color in a single leaflet were likely to be monoclonal (see Methods for details). We found that in the septal leaflet, groups of 4 or less cell clusters could be considered belonging to a single clone, while in the mural leaflet this threshold was reduced to 3 fragments (see Fig. 5b). With this consideration, out of the 171 clone candidates, only 55% (58 clones for the septal and 41 clones for the mural leaflet) satisfied the constraint and could be considered clones with statistical confidence.

Fragment and clone size distribution

The distributions of size of the clusters in identified clones for both leaflets were well described by log-normal distributions (Fig.6a), with logarithmic mean μs = 1.3 (2) and variance σs = 0.9 (2) for the septal, and μm = 1.4 (2) and σm = 1.0 (2) for the mural leaflet. The log-normal distribution of cluster sizes is a hallmark of fragmentation/coagulation processes[17-19]. The fact that clusters for both leaflets follow identical distributions (within errors), suggested that both leaflets share a similar growth dynamic.

The distribution of the size of clones (sum of areas of monoclonal fragments) could be fitted almost equally well by both a log-normal and negative binomial distributions (Fig. 6b). A log-normal distribution of clone sizes could result from a growth process in which progenitors undergo a fixed number of rounds of divisions before terminally differentiating. On the other hand a negative-binomial distribution of clone sizes would suggest a growth process were duplication of self-renewing progenitors in the leaflets is complemented by the permanent immigration of new progenitors from the annulus. The latter could be consistent with the development of the mural leaflet, where at E13.5 there is migration of epicardial cells into the leaflet[13], however the out-clonal data did not have the resolution to distinguish between the plausible models of growth.

Estimation of number of progenitors of the mural and septal leaflets

Focusing on the clones identified by our statistical analysis, we proceeded to estimate the number of self-renewing progenitors contributing to the formation of the mural and septal leaflets. For this, we estimated the average number of clones needed to cover the 2D-projected area of each of the leaflets. For the septal leaflet, we estimated that about 60±13 (SEM) self-renewing progenitors were required to cover the cross section of the leaflet, while about 47±10 (SEM) progenitors were required in the case of the mural leaflet. Given the similarities observed in the clonal organization and estimation of number of progenitors in both leaflets, we questioned whether the pool of progenitors could be the same for both leaflets, so that there is a unique common pool of about 60 progenitors that give rise to both leaflets. In this spirit, we analyzed likelihood that both mural and septal leaflets in heart have cells labelled in the same combination of colors, and compared to the fraction expected from random combinations (null model) estimated from the unicolor and bicolor statistics (see extended Fig. 8). When comparing any two regions, and in particular septal and mural leaflets, we observed a significant coincidence in the color expression of both leaflets compared to the expected from random combinations (Fig 6c). Overall these observations suggest that both leaflets may originate from a common pool of about 60 self-renewing progenitors.

Discussion

Single cell RNA sequencing of AVC endocardial cells and early cushion cells revealed their heterogeneity as expected from their multipotency[20]. This further showed that a subset of endocardial cells, competent for EMT were already primed toward a specific valvular cell type (smooth muscle cell, fibroblast, endothelial cells of future endothelial layers of both mural and septal mitral valve leaflet).

First, we found that a small subset of endocardial cells acquired a competency to undergo EndMT. These cells expressed specific markers of EMT recapitulating the different stages of this process as previously found in the oncology field of research[21]. Thus, EndMT which allows the cells to migrating out of the endocardium to form the cushions is a transformation and likely finely tuned cell reprogramming process as later recapitulated by diseased cancer cells[22]. Pathological EMT has been reported to be regulated by metabolism involving acetyl-CoA[23]. We further found that the endocardial cells competent for EndMT turned on a specific transcriptional program encoding proteins involved in

acetyl CoA metabolism such as *acly, acss2* or *dgat.* This transcriptional program which likely underlies the regulation of EndMT by metabolism[24] was turned on at the earliest step of EndMT and lasted all along the process as suggested by the prolonged expression of *acly* (extended data Fig. 1b).

Ptch1 and *TGFβR1* are the two receptors genes together with the genes encoding the ligands SHH, TGFβ2 and Notch2 to be expressed in early *twist1*+, *prrx1,2*+ cells pointing to their role in the initiation of EndMT. Interestingly SHH signaling component early expressed during the process of EMT regulates fatty acid metabolism[25] while *Flt1* mainly expressed in endocardial cells prevents this metabolism rather favoring glycolysis[26]. Our data thus suggest that only cells that specifically harbor *TGFβR1* co-expressed with *twist1* as well as *Notch2*$^+$ and *Ptch1*$^+$ are capable to undergo EMT while switching their metabolism toward fatty acid oxidation as indicated by expression of *acyl, echs1*, *acaa2*. Thus, a response to combinatorial TGFβ2, Notch2 and SHH signaling pathways, a competence to switch metabolism and a specific transcriptional profile including expression of *hand2* and *msx1* are all required for endocardial cells to undergo EndMT and to be committed toward a valvular cell. These cells loss *cldn5*, a claudin that limits endothelial cell motility[27] and turn on expression of *cldn6* and *cldn7* involved in cancer metastasis[28]. This specific and restricted cell signature explains that only a few cells within the collective endocardial AVC cells undergo EndMT. As recently reported, cell state heterogeneity make cells either low or high responder to growth factors and downstream signaling to ensure specific and context-dependent cellular decision in a multicellular setting[29].

The EndMT transcriptional and cell metabolic reprograming events were then followed in mesenchymal cells by a genetic program specific to future valvular cell types. The latter was then designed to allow cells to acquiring a capability to secrete specific extracellular matrix proteins (ECM) (Extended data Fig.1). We found that at E16.5, valvular cells expressed genes encoding ECM proteins found in the layers of the valve leaflets while a cell subset remained endothelial to form the surrounding layer of the leaflets. The process was more advanced at the neonatal stage, the mitral valve featuring two different types of VICs expressing collagen genes as well as proteoglycan genes. Interestingly the VEC could be discriminated between parietal and septal valve leaflets depending upon the presence or absence, respectively of cells expressing *wt1*. Wt1 lineage tracing indeed was reported to contribute to VICs as well as VECs of the parietal mitral leaflet[11] (Extended data Fig 7). This suggests that some cells of (pro)epicardial origin expressing Tie2 at the

time of induction of the recombinase retained expression of *wt1* even after migration into the valve leaflet. The VICs of both parietal and septal leaflets likely originated from the same AVC cells cluster (Fig. 4F) which is in agreement with the clonal analysis that also pointed to the same pool of progenitors shared by both valve leaflets (Fig. 6). Clonal analysis predicted that 60 progenitors were required to build up the septal leaflet and 47 for the parietal leaflet, a number also approximated from the total number of AVC cells which contribute to the E16.5 Mitral leaflets in single cell data calculated from the additive number of cells in each contributing cluster (Fig 4F) divided by the number of embryos and in turn AVC used (140 progenitors). This implied that the total clones size results from a growth process of differentiating progenitors which undergo a fixed number of cell divisions before terminally differentiating in VICs and the division events were not synchronized. Assuming a number of cells in a mitral valve leaflet of 400 in septal and 150 in mural leaflets at birth[15], this suggests that cells should undergo no more than three rounds of division. This is still possible that they even undergo less division as cells such as epicardial cells also migrate into and contribute to at least the mural leaflet of the mitral valve.

When and how do they expand are other questions addressed by the analysis of the summed areas of clones of a given color (Fig.6). Such an analysis revealed a bimodal behavior of the total clone area of given colors, where a fraction of the combined clones have an area of about 5000 μm2. This might indeed also suggest that the progenitors expand prior to commitment to either leaflet during the early steps of the process of the morphogenesis of the leaflet. Upon commitment, an uneven fraction of cells commit to one or the other leaflet depending upon their transcriptomic profile defined as early as in the AVC at E9.5 (Fig. 1), where they continue expanding, conserving the overall clone size. Smaller clones might also correspond to lineages that committed partially, or to endocardial cells already committed within the AVC that migrated late into the leaflet after the completion of formation of the 4 chambers and to non-endocardial cells, such as the ones of epicardial origin.

Altogether, our single cell RNA-sequencing data combined with the clonal analysis revealed that the process of embryonic mitral valve formation starts early during development when the AVC cells either prone to undergo EMT or not already distribute between valve progenitors cells and endothelial/endocardial cells that will participate in lining the chambers during their growth, respectively. The EMT process is well secured by three cell processes (i.e original transcriptomic profile, specific receptor mediated signaling

events, competency to both EMT and to a metabolic cell reprogramming event) to prevent any overgrowth of the cushions and to lead to both an appropriate and a well-tuned size of the leaflets to ensure a perfect closer of the future valve.

Our findings further point to a valve morphogenesis that remains a cell dynamic process. Cells specifically contribute to each anterior or posterior cushion and later to specific mural or septal leaflet. Other cells from likely different embryonic origins[30] migrate later into these respective valve leaflets to complete the whole process of leaflets formation. The 2D distribution of the cell clones already specified toward the spongiosa, the fibrosa or the atrialis likely contributes to the layer formation of the valve.

Dysregulation of any of these processes may lead to a pathological valve.

Acknowledgments:

This study was funded by the Leducq Foundation (MITRAL network of excellence).

We are also grateful to the Leducq Foundation for generously awarding us for cell imaging facility (MP “Equipement de Recherche et Plateformes Technologiques” (ERPT).

Methods

Mouse lines

CAG^{CreERT2} mice and *cytbow* mice were previously described[31]. *Rosa–tdTomato and Tie2cre* mice were obtained from Jackson laboratory (B6;129S6-*Gt (ROSA)26Sortm14(CAG-tdTomato)Hze and B6.Cg-Tg(Tek-cre)12Flv/J, respectively)*. The WT1^{CreERT2} mouse was provided by Sylvia Evans laboratory (UCSD). Mice were kept under standardized conditions (22–24°C temperature; 50%±10% humidity) on a 12 h light/12 h dark cycle in an enriched environment (kraft paper enrichment). Food and tap water were provided ad libitum. The parental strains were bred on a C57BL/6 background.

CAG$^{CreERT2+/-}$ mice were crossed with the transgenic mice *cytbow*. To obtain embryos of the desired stages (E9.5, E16.5), mating females were verified for the presence of vaginal plug every morning. The day of the presence of a vaginal plug was considered as embryonic day E0.5 at noon. Pregnant females were euthanized at the desired stage of gestation.

Tamoxifen (Sigma-Aldrich, France; 20 µg/g) was administrated with corn oil by oral gavage at embryonic day E8.5. Embryos were collected at E9.5 and E16.5, fixed overnight with 4% paraformaldehyde.

Preparation of frozen sections

Embryos at E9.5 and hearts at E16.5 were prepared for immunostainings. Samples were embedded in Optimum Cutting Temperature (O.C.T) compound (Tissue-Tek) for cryoprotection. Embryos at E9.5 and E16.5 were dissected in ice-cold PBS and fixed for 2hrs or overnight at 4°C in 4% PFA in PBS. Embryos were then transferred to 6% sucrose/PBS (w/v), 12.5% sucrose/PBS and 25% sucrose/PBS (1hr each) with gentle shaking at 4°C. Then the embryos were transferred into embedding molds, ad frozen in50:50 (v/v) O.C.T/50% sucrose. Samples were stored at -80°C and sectioned (10µm, 20µm, 30µm) using a Leica Microm HM 500 OM Cryostat (-20°C). Frozen sections were mounted on superfrost plus slides and stored at -80°C before use.

Immunofluorescence analysis

Immunofluorescence was performed on frozen embryonic heart sections (Leica cryostat-10 µm). Sections were fixed with PFA for 10 minutes, washed with PBS 5 minutes, then blocked in PBS with .5% donkey serum, 5% BAS and 0.1% tween 20. Sections were stained with the following antibodies: anti-CD31 (BD550274 Purified Rat anti-mouse CD31; 1/100), anti-alpha actinin sarcomeric (mouse, Sigma-Aldrich A7811; 1/200) anti Twist1 (mouse 1/300 Novusbio NBP2), anti-ACLY (37364SS ACLY rabbit 1/200 Novusbio NBP2 – 67509), anti-Msx1 (rabbit 1/300 Novusbio NBP2 – 57021), anti-VECAM-1 (rabbit 1/250 Abcam ab-134047), anti-Integrin alpha V (rabbit 1/500 Abcam ab-179475) anti-Integrin beta 3 (rabbit 1/500 Abcam ab-119992). Nuclear staining was performed with DAPI (1/1000).

Images were acquired using a confocal LSM 800 Zeiss microscope equipped with an airryscan of 32 detectors. Light was provided by a laser module 405/488,561 and 640 nm wavelengths, respectively.

All images were acquired using the ZEN ZEISS software. Then some images were deconvoluted using Autoquant and reconstructed in 3D using Imaris software (IMARIS). All samples were mounted in Fluoromount™ (Cliniscience, France)

In all figures, tdTomato/mCherry, mEYFP, and mCerulean/mTurquoise2 are rendered in red, yellow, and blue, respectively.

Two-photon imaging of transparised embryonic hearts was performed as previously described[32].

Assessment of the recombination of the *cytbow* reporter by PCR

Genomic DNA was extracted by DNA precipitation in NAOH solution (50nm) from the tail of mice, wild type mouse and adult *cytbow* mouse as a control. Tails of embryos at E16.5 *CAG-Cre/cytbow* of the whole litters were used to genotype embryonic mice with specific primers for both, *CAG*^{Cre} and *cytbow* using the following primers CAG^{cre} forward 5'-GCTAACCATGTTCATGCCTTC and reverse 5'-AGGCAAATTTTGGTGTACGG-3'; TCY forward 5'-CTGTTCCTGTACGGCATGGA-3', and reverse 5'-TTTCAGGTTCAGGGGGAGGT-3'.

Mouse Genotyping Primers

CYT forward 5'-CTGTTCCTGTACGGCATGGA-3' and reverse 5'-TTTCAGGTTCAGGGGGAGGT-3'; CAG^{CreERT2} *5'GCTAACCATGTTCATGCCTTC-3'; 5'-AGGCAAATTTTGGTGTACGG-3'; ROSA*^{tdTomato}*; Wild type forward: 5'-AAGGGAGCTGCAGTGGAGTA-3' and reverse :5-'CTGTTCCTGTACGGCATGG-3', Mutant forward :5-'CCGAAAATCTGTGGGAAGTC-3' and reverse :5'-GGCATTAAAGCAGCGTATCC-3'; Tie-2*^{Cre} forward *5'-CGCATAACCAGTGAAACAGCATTG-3' and reverse 5'-CCCTGTGCTCAGACAGAAATGAGA-3'; WT1*^{CreERT2} *forward 5'-TGAAACAGGGGCAATGGTGCG-3' and reverse 5'-CGGAATAGAGTATGGGGGGCTCAG-3'*

Single cell RNA sequencing

Endocardial cells of the atrio-ventricular canal derived from Tie2^{Cre}.Tdt positive embryos at E9.5 (18-22 somites) and mitral valves were collected at E16.5 and P0, dissociated with TrypLE^{TM} express enzyme (Gibco; Catalog number12604013) into single cells. After sorting the positive tomato cells with FACS, 10x genomics chromium platform was used to capture and barcode cells in order to generate single-cell Gel Beads-in emulsion (GEMs), following the manufacturer's protocol. Thus, using the reverse transcription master mix, cell suspensions of 2 samples of AVC, 1 sample of E16,5,and 2 samples of P0 were loaded onto the 10x Genomics Single Cell 3' Chip. Single cells were then partitioned into GEMs along with gel beads coated with oligonucleotides. Reverse transcription of cDNAs was followed by amplification, and a library was constructed using the Single cell 3' reagent Kit for all the samples, followed by sequencing of all libraries wit Next-seq Illumina sequencer. 10X Genomics pipeline Cell Ranger v1.2.0 was used for sample demultiplexing, barcode processing, and UMI counting. Libraries obtained from Illumina Sequencer were

demultiplexed into reads in FASTQ format using the “cell ranger mkfastq” pipeline. In order to generate a gene-barcode matrix for each library, cellranger count pipeline was used. All the reads were aligned to the *mus musculus* reference genome mm10.

Cell ranger pipeline was used to normalize all the libraries to the same sequencing depth.

A first analysis of the data was performed with cell ranger and loop cell 10xGenomics software’s. C loop clustered the cells using genes that are differentially expressed comparing to all other cells in other clusters with a threshold of Log2 > or equal to 2. A deep and secondary analysis was performed using the SCDE package and SCShinyhub[33]. Low UMI cells were excluded before clustering. We used: for E9.5 samples: sample: 1200 cells/ median gene per cell: 4617. For E16.5 sample: 452 cells/ median gene per cell: 700/; For P0 2 samples were used for a total of 854 cells/ median gene per cell 900.

Clonality assessment and clonal grouping

Cell labeling followed tamoxifen administration at E8.5, and samples were collected at E16.5. At collection, labelled cells had expanded into clones that appeared to be fragmented into separate clusters of cells within a leaflet. To infer whether a group of clusters of cell belonged to a single clone or were the result of merger of two or more clones, we followed a biostatistical approach first introduced by Lescroart *et al*[34], which we describe below.

In this approach we considered the probability p that, following tamoxifen administration at E8.5, a given expressing cell was induced, as shown in extended data Fig.4.The probability $P(m)$, of observing m successful inductions can be approximated, for a large number N of expressing cells, by a Poisson distribution: $P(m) \approx \frac{(pN)^m}{m!} e^{-pN}$ **(1)**, where pN is the overall induction probability. If fragmentation is considered to be Poissonian, the probability *R*(*k*) that a given clone will end up fragmented into *k* fragments at day E16.5 is: $R(k) = \frac{f^{k-1}}{(k-1)!} e^{-f}$ **(2)**, where f corresponds to the expected number of fragments of a clone derived from a single cell, at the time of measurement (E16.5). As argued in[32], the degree of fragmentation *f* could depend on the total size of the clone. To assess this, we measured the relation between total clone size and number of fragments in unicolor hearts (extended data Fig. 6). Although the number of measurements was low, as restricted to leaflets labelled with a single color, there was no clear dependence between these two variables. For this we considered clones in separate leaflets in the same heart to be independent.

Considering that the number of fragmentation events is given by the difference *k-m* (*k≥m*) between the number of fragments *k* and the number of induced cells *m*, then the conditional probability $S(k \mid m)$ of finding a total of *k* labeled fragments of a single color at day E16.5, given that *m* cells of that same color were induced at day E8.5 is: $S(k|m) = \frac{(mf)^{k-m})}{(k-m)!} e^{-mf}$ **(3)**. Combining equations **(1)** and **(3)**, we obtained the joint probability *J* (*k*|*m*) of finding a leaflet with initially *m* induced cells giving rise to *k* fragments of a given color is

$J(k|m) = S(k|m)P\ (m) = \frac{(mf)^{k-m}\ (pN)^m}{(k-m)!\ m!} e^{-mf-pN}$ **(4).**

Taking into account only conditions with induced cells (*m* > 0), we defined the joint distribution of labeled heart σ: $J_{ind}(m,k) = J(m,k)/(1-e^{-pN})$.Then the distribution for the number of fragments in labeled hearts is:

$F(k) = \sum_(m = 1)\hat{}k\ J_ind\ (m,\ k) = e\hat{}(-pN)/(1 - e\hat{}(-pN)\ )\ \sum\ (\ [\![(mf)]\!]\ \hat{}(k - m)\ (pN)\hat{}m)/(k - m)!\ m!\ e\hat{}(-mf)$ **(5).**

To assess the clonality of our lineage tracing data, we used the experimental quantification of the number of fragments *k* and Eq. **(5)**. We first estimated the overall induction frequency *pN* and average fragmentation rate of clones *f* by combining the data from both the mural and septal leaflets. Given the spatial separation between the mural and septal leaflets, we considered the cell clusters from separate leaflets to be independent from each other. Thus, if in a given sample we observed, for example, 4 red fragments in the septal and 3 red fragments in the mural leaflet, we considered these to be two clone candidates, one with 4 and the other with 3 fragments. We analyzed the number of fragments in both leaflets of the mitral valve of 40 hearts, from which we obtained fragments for 171 clone candidates. We analyzed the distribution of the number of fragments (extended data Fig. 7) and fitted the theoretical curve (**5**), from which we obtained an overall induction frequency *pN* = 0.5 (0.3, 95% C.I.) (**2**) and a fragmentation rate *f* = 1.4 (0.3) (**3**). This implied that clones with 4 or less fragments were likely to be monoclonal, when the data for both the mural and septal leaflets was pooled together.

Even though the induction frequency *pN=0.5* can be assumed to be equal for both leaflets, this is not necessarily the case for the fragmentation rate *f*.

To account for possible differences in the fragmentation in the mural and septal leaflets, we analyzed the distribution of number of fragments *F* (*k*) of both leaflets, independently (Fig. 5), for fixed induction frequency *pN* = 0.5, as obtained from the combined data

(extended data Fig. 7). In the case of the septal leaflet, the threshold for monoclonality was found to be k = 4, meaning that groups of four clusters or less were considered monoclonal, while this threshold was lowered to k = 3 for the mural leaflet. The difference originates from the lower fragmentation f in the case of the smaller mural leaflet.

References

1. Hoffman, J.I. & Kaplan, S. The incidence of congenital heart disease. *J Am Coll Cardiol* **39**, 1890-1900 (2002).
2. Kyndt, F. *et al.* Mutations in the gene encoding filamin A as a cause for familial cardiac valvular dystrophy. *Circulation* **115**, 40-49 (2007).
3. Dietz, H.C. *et al.* Marfan syndrome caused by a recurrent de novo missense mutation in the fibrillin gene. *Nature* **352**, 337-339 (1991).
4. Neptune, E.R. *et al.* Dysregulation of TGF-beta activation contributes to pathogenesis in Marfan syndrome. *Nat Genet* **33**, 407-411 (2003).
5. Durst, R. *et al.* Mutations in DCHS1 cause mitral valve prolapse. *Nature* **525**, 109-113 (2015).
6. Toomer, K.A. *et al.* Primary cilia defects causing mitral valve prolapse. *Sci Transl Med* **11** (2019).
7. Luna-Zurita, L. *et al.* Integration of a Notch-dependent mesenchymal gene program and Bmp2-driven cell invasiveness regulates murine cardiac valve formation. *J Clin Invest* **120**, 3493-3507 (2010).
8. Zhou, B. *et al.* Characterization of Nfatc1 regulation identifies an enhancer required for gene expression that is specific to pro-valve endocardial cells in the developing heart. *Development* **132**, 1137-1146 (2005).
9. Combs, M.D. & Yutzey, K.E. Heart valve development: regulatory networks in development and disease. *Circ Res* **105**, 408-421 (2009).
10. Pastushenko, I. & Blanpain, C. EMT Transition States during Tumor Progression and Metastasis. *Trends Cell Biol* **29**, 212-226 (2019).
11. Laurent, F. *et al.* HAND2 Target Gene Regulatory Networks Control Atrioventricular Canal and Cardiac Valve Development. *Cell Rep* **19**, 1602-1613 (2017).
12. Chen, Y.H., Ishii, M., Sucov, H.M. & Maxson, R.E., Jr. Msx1 and Msx2 are required for endothelial-mesenchymal transformation of the atrioventricular cushions and patterning of the atrioventricular myocardium. *BMC Dev Biol* **8**, 75 (2008).
13. Wessels, A. *et al.* Epicardially derived fibroblasts preferentially contribute to the parietal leaflets of the atrioventricular valves in the murine heart. *Dev Biol* **366**, 111-124 (2012).
14. Tran, T.N. & Bader, G.D. Tempora: Cell trajectory inference using time-series single-cell RNA sequencing data. *PLoS computational biology* **16**, e1008205 (2020).
15. Moore, K. *et al.* PDGFRalpha: Expression and Function during Mitral Valve Morphogenesis. *J Cardiovasc Dev Dis* **8** (2021).
16. Wuidart, A. *et al.* Early lineage segregation of multipotent embryonic mammary gland progenitors. *Nat Cell Biol* **20**, 666-676 (2018).
17. Bailey, N. *The elements of stochastic processes*. (john wiley & sons. Inc., New York, 1964).
18. Redner, S. *Statistical theory of fragmentation. In Disorder and fracture*. (1990).
19. Chabab, S. *et al.* Uncovering the Number and Clonal Dynamics of Mesp1 Progenitors during Heart Morphogenesis. *Cell Rep* **14**, 1-10 (2016).
20. Zhang, H., Lui, K.O. & Zhou, B. Endocardial Cell Plasticity in Cardiac Development, Diseases and Regeneration. *Circ Res* **122**, 774-789 (2018).
21. Pastushenko, I. *et al.* Identification of the tumour transition states occurring during EMT. *Nature* **556**, 463-468 (2018).

22. McFaline-Figueroa, J.L. *et al.* A pooled single-cell genetic screen identifies regulatory checkpoints in the continuum of the epithelial-to-mesenchymal transition. *Nat Genet* **51**, 1389-1398 (2019).
23. Sanchez-Martinez, R. *et al.* A link between lipid metabolism and epithelial-mesenchymal transition provides a target for colon cancer therapy. *Oncotarget* **6**, 38719-38736 (2015).
24. Xiong, J. *et al.* A Metabolic Basis for Endothelial-to-Mesenchymal Transition. *Mol Cell* **69**, 689-698 e687 (2018).
25. Bhatia, B., Hsieh, M., Kenney, A.M. & Nahle, Z. Mitogenic Sonic hedgehog signaling drives E2F1-dependent lipogenesis in progenitor cells and medulloblastoma. *Oncogene* **30**, 410-422 (2011).
26. Seki, T. *et al.* Ablation of endothelial VEGFR1 improves metabolic dysfunction by inducing adipose tissue browning. *J Exp Med* **215**, 611-626 (2018). Science 10.1126/science.abf4062 (2022)
27. Yang, Z. *et al.* The tight junction protein Claudin-5 limits endothelial cell motility. *J Cell Sci* **134** (2021).
28. Tabaries, S. & Siegel, P.M. The role of claudins in cancer metastasis. *Oncogene* **36**, 1176-1190 (2017).
29. Kramer, BA, Sarabia del Castillo, J Pelkmans, Multimodal perception links cellular state to decision making in single cellsScience 10.1126/science.abf4062 (2022)
30. Farrar, E.J. *et al.* OCT4-mediated inflammation induces cell reprogramming at the origin of cardiac valve development and calcification. *Sci Adv* **7**, eabf7910 (2021).
31. Abdeladim, L. *et al.* Multicolor multiscale brain imaging with chromatic multiphoton serial microscopy. *Nat Commun* **10**, 1662 (2019).
32. Mahou, P. et al. Multicolor two-photon tissue imaging by wavelength mixing. Nat. Methods 9, 815–818 (2012).
33. Jagla, B. *et al.* SCHNAPPs - Single Cell sHiNy APPlication(s). *J Immunol Methods* **499**, 113176 (2021).
34. Lescroart, F. *et al.* Early lineage restriction in temporally distinct populations of Mesp1 progenitors during mammalian heart development. *Nat Cell Biol* **16**, 829-840 (2014).

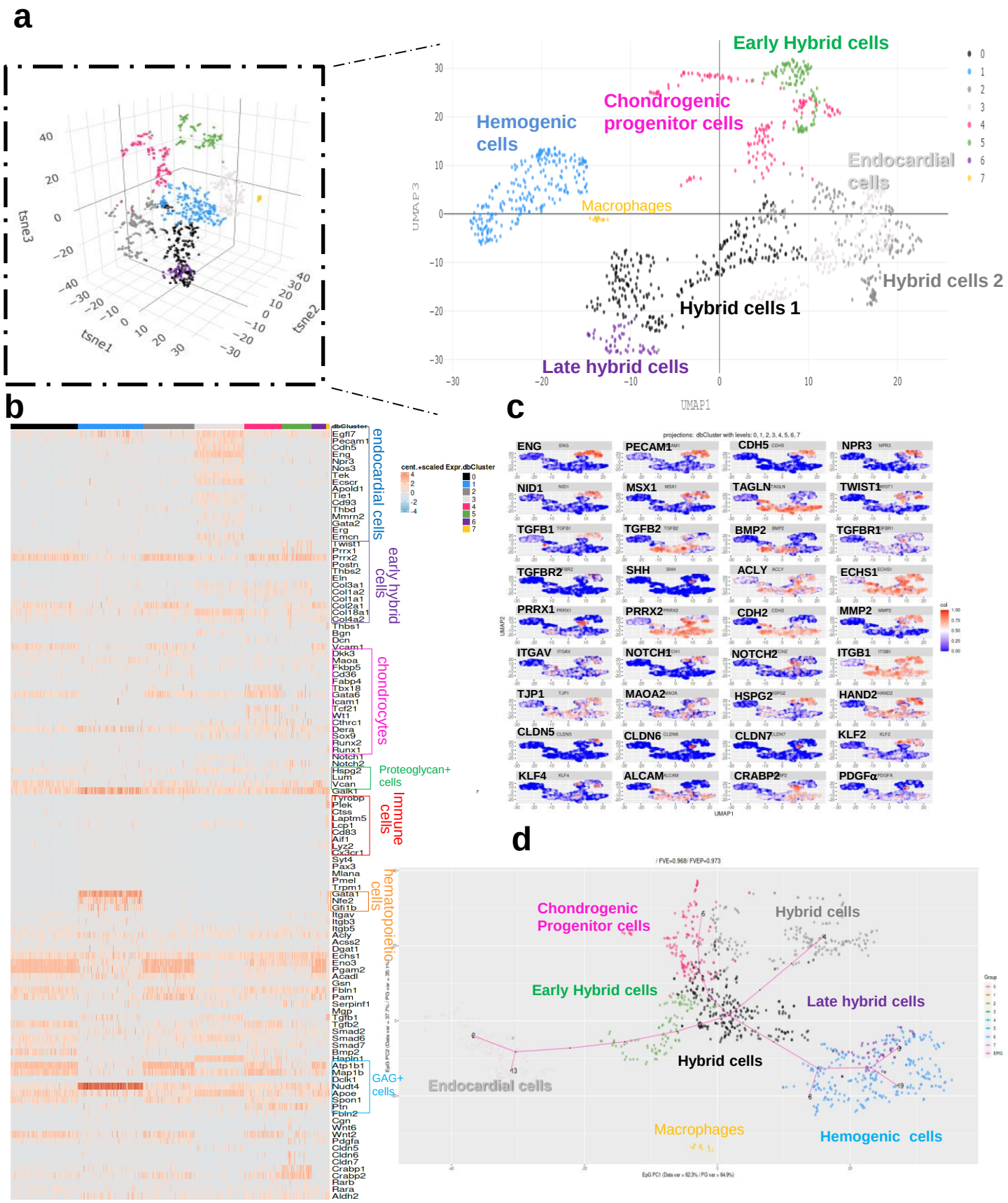


Figure 1 : single cell transcriptomic of E9.5 AVC cells. AVC was dissected from embryos from Rosa26-floxed stop tdTomato females bred with Tie2-Cre males. Enzymatically dissociated cells were FACS-sorted and used for single cell-RNA seq. (a) UMAP; (b) heatmap of gene expression; (c) panel plots of genes involved in EMT in endocardial cells (cluster 3), early hybrid (cluster 5) and hybrid (cluster 0) cells (cluster numbers in first top right panel) (d) elpigraph inference trajectory.

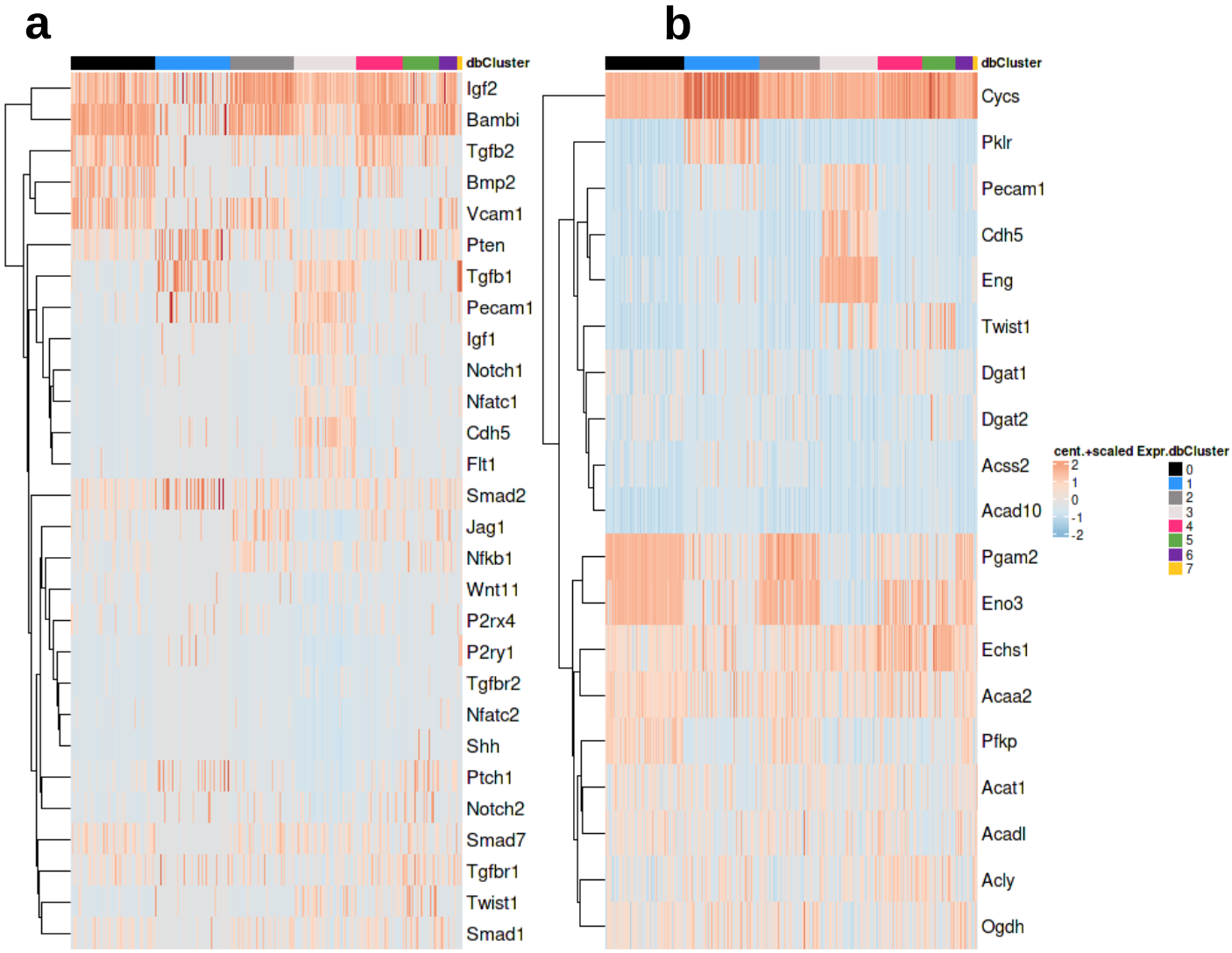


Figure 2: heatmap of AVC cells. Heatmap reavealing expression of genes involved in (a) signaling pathways
(b) heat map of the metabolic pathways proteins involved in fatty acid β-oxydation.

Figure 2

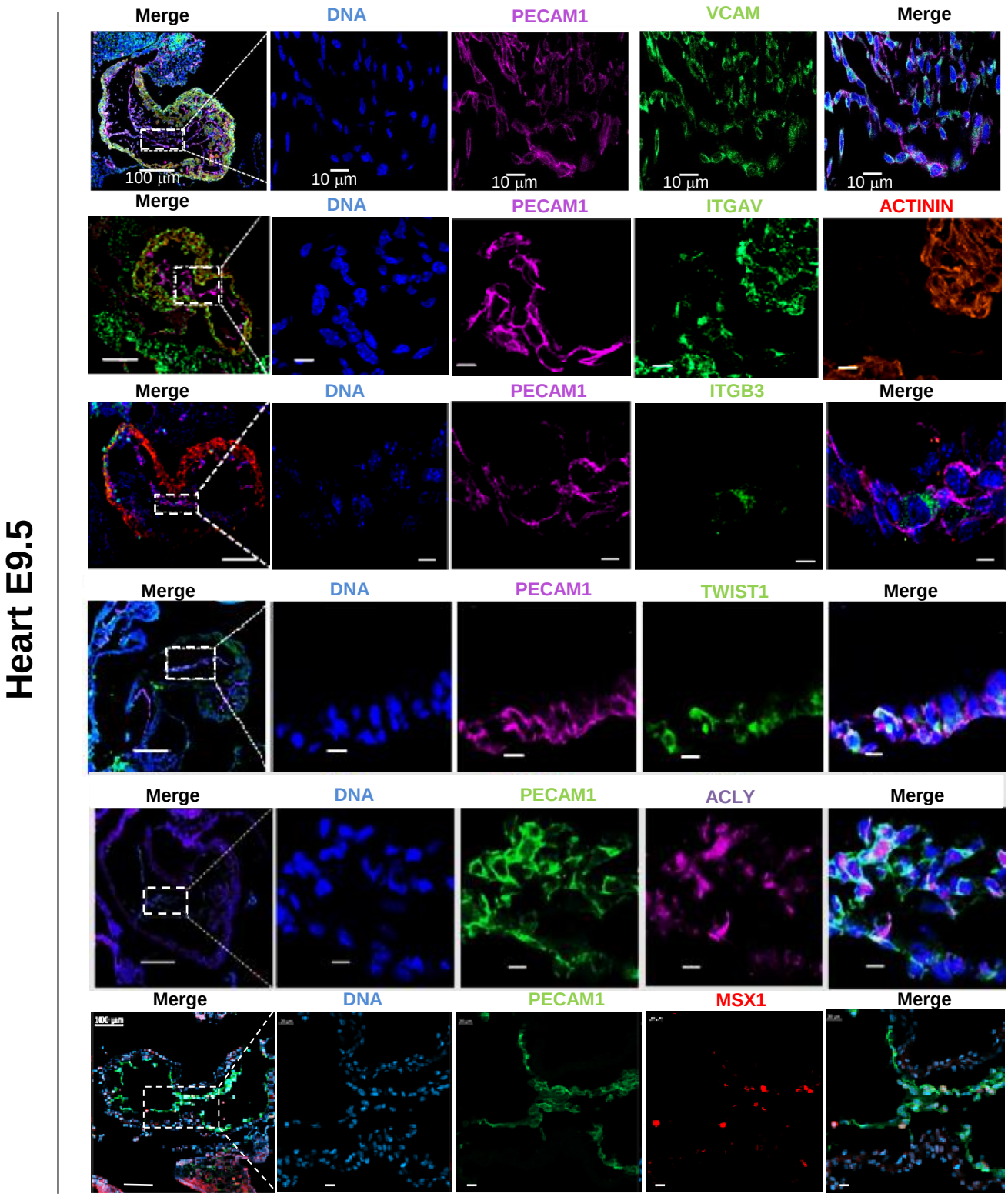


Figure 3: immunofluorescence of proteins expressed in AVC cells in E9.5 mouse heart. The left panels show low magnifications of sections of whole heart. The square marks the AVC only imaged in the other panels; The scale bars in the left column is 100 µm and the others are 10 µm

Figure 3

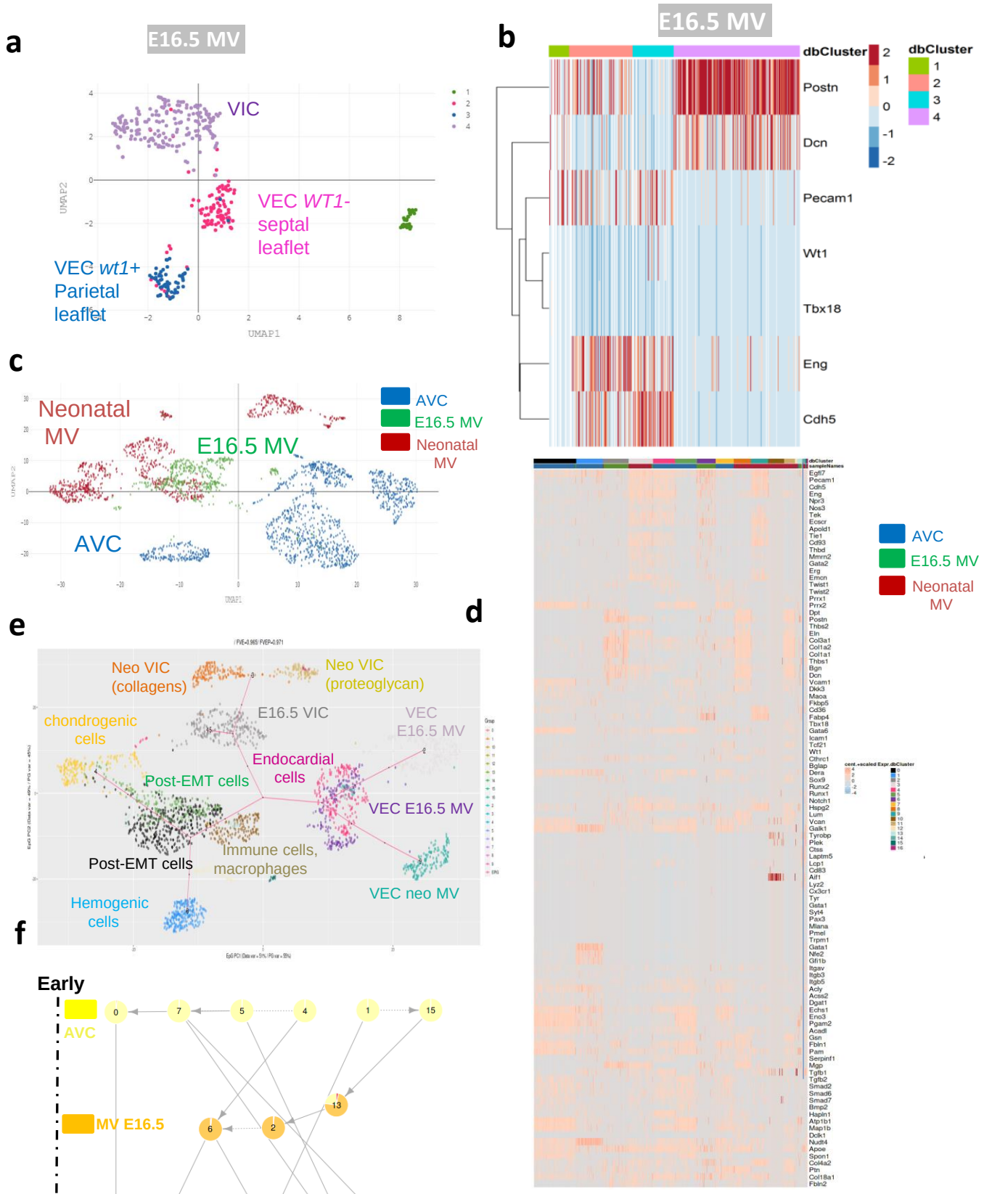
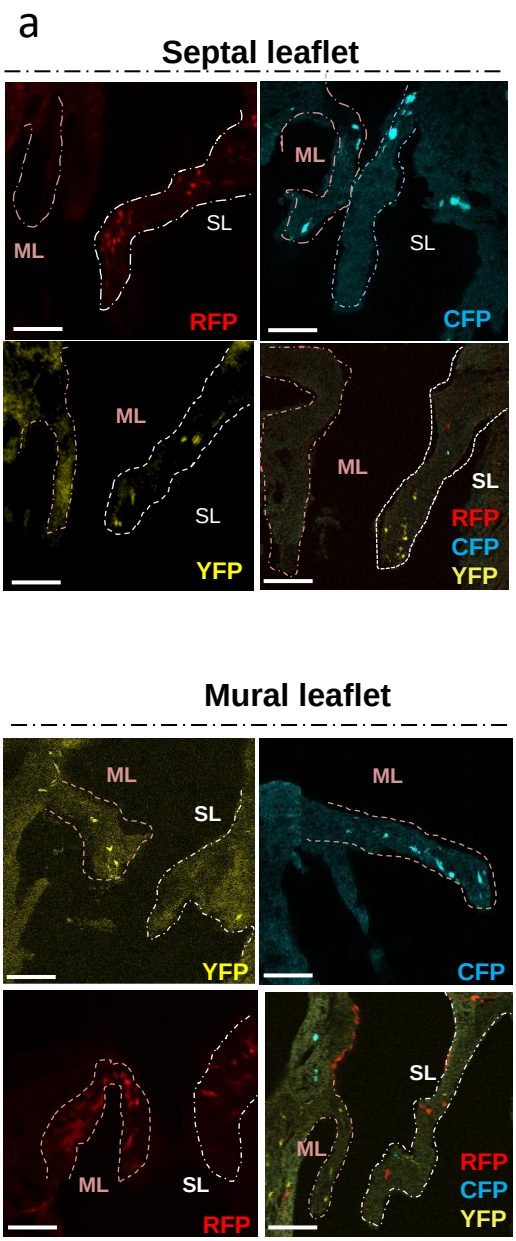


Figure 4 : single cell transcriptomic of E16.5 and neonatal mitral valves. Mitral valve was dissected from embryos from Rosa26-floxed stop tdTomato females bred with Tie2-Cre males. Enzymatically dissociated cells were FACS-sorted and used for single cell-RNA seq. (a) UMAP of E16.5 MV cells , (b) heatmap of gene expression E16.5 MV cells.(c) UMAP of all samples (AVC, E16.5 and neonatal mitral valve cells). (d) heatmap of all samples (e) elpigraph inference trajectory including all datasets. (f) tempora inference trajectory of all samples.



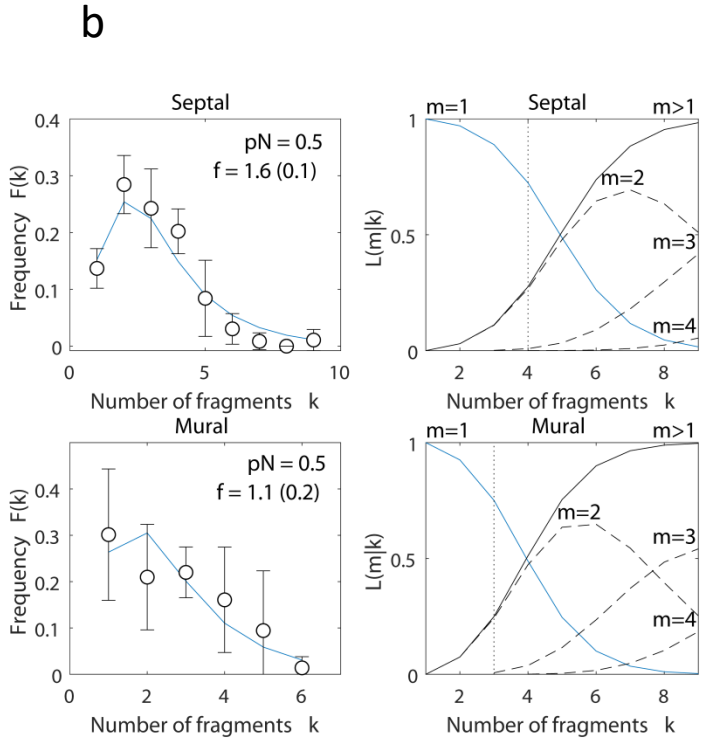


Figure 5. cell clone distribution in embryonic mitral valve a: E16.5 mitral valve leaflets images with cell clones from Cag^{Cre} x cytbow embryos b: Probability F(k) of observing a total of k fragments of any given colour, and (right) conditional probability L(m|k) that m cell where induced given that we observe k fragments in the anchoring point for the (top) septal and (bottom) mural leaflets. Markers correspond to experimental measurements, error bars correspond to the standard deviation from the measurement of F(k) for the three different colors. The solid line was obtained from fitting Eq. (5) in methods for pn = 0.50 and free parameter f**.**

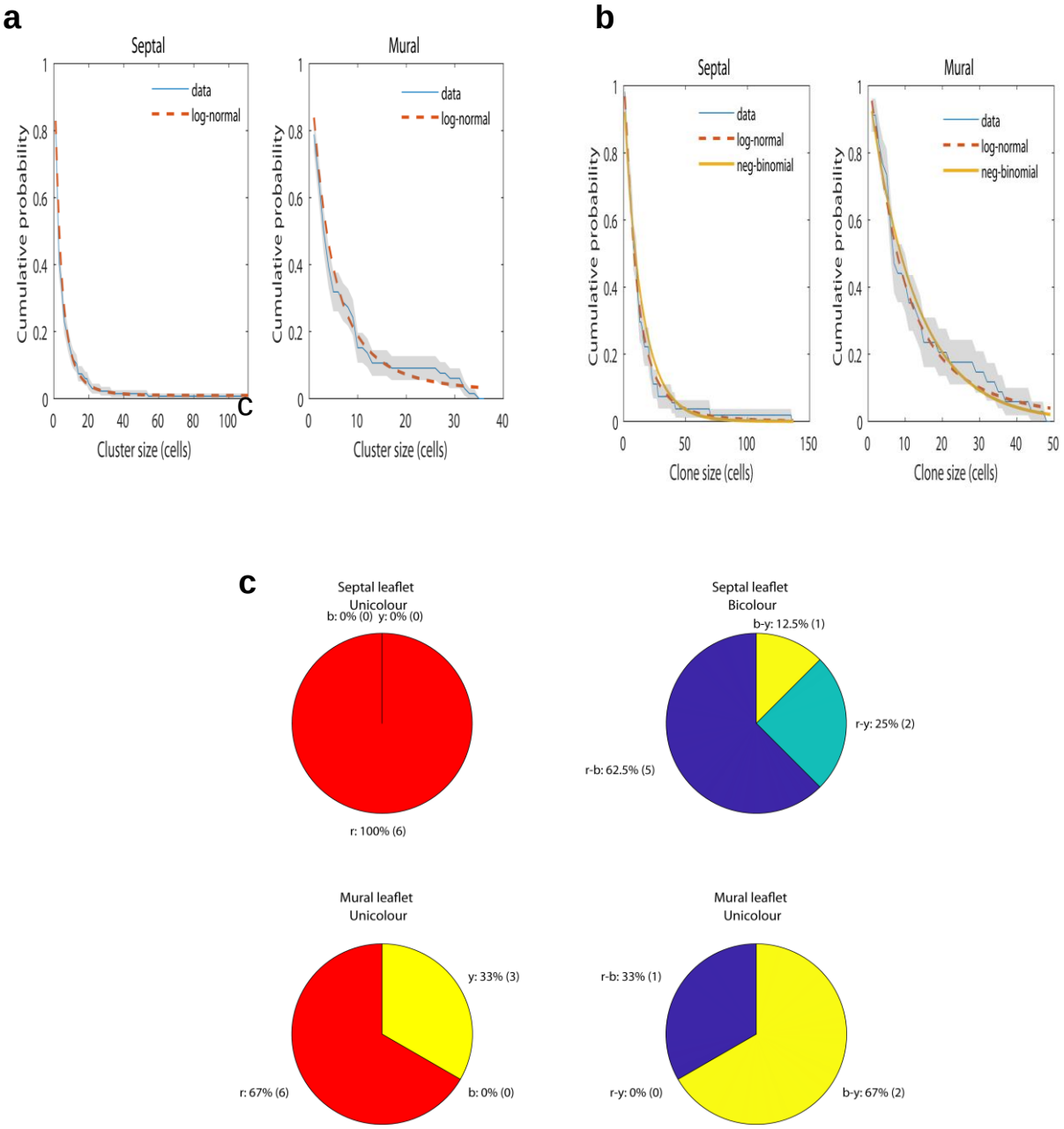


Figure 6 (a): Distribution of cluster (a) and cell clone (b) sizes in both spetal and mural leaflets of E16.5 mitral valve (c) :observed color expression in both septal and mural leaflets compared with ramdom combinations