Published in final edited form as:

Science. 2018 March 09; 359(6380): 1177–1181. doi:10.1126/science.aao4174.

Defining the earliest step of cardiovascular lineage segregation by single cell RNA-seq

Fabienne Lescroart^{1,*}, Xiaonan Wang^{2,3,*}, Xionghui Lin^{1,*}, Benjamin Swedlund¹, Souhir Gargouri¹, Adriana Sánchez-Dànes¹, Victoria Moignard^{2,3}, Christine Dubois¹, Catherine Paulissen¹, Sarah Kinston^{2,3}, Berthold Göttgens^{2,3,#,†}, and Cédric Blanpain^{1,4,#,†}

¹Université Libre de Bruxelles, Laboratory of Stem Cells and Cancer, Brussels B-1070, Belgium.

²Department of Haematology, Cambridge Institute for Medical Research, University of Cambridge, Cambridge CB2 0XY, UK.

³Wellcome and Medical Research Council Cambridge Stem Cell Institute, University of Cambridge, Cambridge, UK

⁴WELBIO, Université Libre de Bruxelles, Brussels B-1070, Belgium.

Abstract

Mouse heart development arises from *Mesp1* expressing cardiovascular progenitors (CPs) that are specified during gastrulation. The molecular processes that control early regional and lineage segregation of CPs remain unclear. Here, we performed single cell RNA-sequencing of WT and *Mesp1* null CPs. We showed that *Mesp1* CPs are molecularly distinct and make the continuum between epiblast and later mesodermal cells including hematopoietic progenitors. Single cell transcriptome analysis of *Mesp1* deficient CPs showed that *Mesp1* is required for the exit from the pluripotent state and the induction of the cardiovascular gene expression program. We identified distinct populations of *Mesp1* CPs that correspond to progenitors committed to different cell lineages and regions of the heart, identifying the molecular features associated with early lineage restriction and regional segregation of the heart at the early stage of gastrulation.

One sentence summary:

Early lineage segregation during gastrulation

The mammalian heart is composed of different regions (ventricles, atria, great vessels) and cell types including cardiomyocytes (CMs), endocardial (ECs), smooth muscle (SMCs) and epicardial cells (EPs) (1). Heart development begins at gastrulation, during which CPs leave the primitive streak (PS) and migrate toward the antero-lateral pole of the embryo (2). From embryonic day 6.25 (E6.25) to E7.25, *Mesp1* marks the early CPs within the PS whereas a day later *Mesp1* is expressed in the somites (2, 3). *Mesp1*⁺ cells give rise to all heart cells, ECs of the aorta and brain, some muscles of the head and neck, as well as to few somitic derivatives and liver cells from its later expression (2, 4–6). Temporally inducible *Mesp1*

[†] Corresponding authors.

^{*} Denotes equal contribution

[#] Co-last authors

lineage tracing shows that at E6.5 *Mesp1*⁺ cells mark left ventricle (LV) progenitors, whereas the right ventricle, atria, outflow and inflow tracts and head muscles arise from *Mesp1*⁺ cells at E7.25, which correspond respectively to the first and second heart fields (FHF and SHF). No somitic or liver derivatives were labeled at these early time points (4, 5, 7). In addition, most of the *Mesp1* CPs differentiate into either CMs or ECs, suggesting that lineage segregation occurs early during gastrulation (4, 5). It remains unknown whether molecular heterogeneity between E6.5 and E7.25 *Mesp1* CPs reflects stochasticity in gene expression, transcriptional priming or early lineage and regional segregation.

To investigate the molecular and cellular basis of the earliest stages of CP specification and diversification, we performed single cell RNA-seq of *Mesp1* CPs at E6.75 and E7.25. To this end, *Mesp1-rtTA/tetO-H2B-GFP* mice were treated with doxycycline at different time points after plug detection to label only early *Mesp1* expressing cells and no somitic derivatives (Fig. 1A), embryos were dissociated into single cells and H2B-GFP-positive CPs were isolated by FACS (fig. S1). A total of 172 and 341 *Mesp1* CPs at E6.75 and E7.25 respectively were sequenced and analyzed further after passing through a stringent quality control pipeline (see Methods). We recently reported single cell transcriptomes for E6.5 epiblast cells, as well as E7.25/7.5 Flk1-expressing progenitors (8). Visualization using dimensionality reduction techniques allowed us to order the cells along developmental progression and assign a timestamp to each cell, demonstrating that the *Mesp1*⁺ CPs at E6.75 and E7.25 likely represent a continuum of differentiation (Fig. 1B–C and fig. S2) (8).

To determine the role of *Mesp1* in regulating the cardiovascular differentiation program and the heterogeneity of early CPs, we performed scRNA-seq of FACS isolated *Mesp1* expressing cells in *Mesp1* knockout (KO) context (2)(fig. S3). We sequenced transcriptomes of 85 single *Mesp1* null cells isolated at E6.75, before phenotypic onset (Fig. 2A). Pseudotime analysis revealed that *Mesp1* KO cells presented a developmental block, being stuck in the gene expression program of epiblast cells (Fig. 2B). PCA analysis showed that principal component 2 captured expression differences between WT and *Mesp1* KO cells (fig. S4) with 206 downregulated and 136 upregulated genes (Table S1). We found a highly significant overlap for genes differentially expressed between WT and *Mesp1* KO cells *in vivo* and genes that are down/up-regulated following *Mesp1*-induced gain of function in embryonic stem cells (ESCs) *in vitro* (9), many of which are direct *Mesp1* target genes (Fig. 2C, fig. S5 and Table S1). Several well-known regulators of pluripotency including *Nanog*, *Eras*, *Pou5f1/Oct4* (10) and markers of the epiblast including *E-Cadherin/Cdh1*, *Epcam*, *Cldn6* and *Cldn7* were upregulated in single *Mesp1* KO cells (Fig. 2D–F and Table S1), consistent with the defect of exiting the pluripotent epiblast stage. In contrast, the genes downregulated in *Mesp1* KO cells were greatly enriched for *Mesp1* target genes controlling EMT (*Snai1*, *Zeb2*), migration (*Rasgrp3*), and cardiovascular commitment (*Etv2*, *Hand1*, *Myf7*, *Gata4*, *Flk1*, *Pdgfra*) (9, 11) (Fig. 2F–G and fig. S5). *Pdgfra*/Flk1 expressing cells that mark *Mesp1* CPs in human and mouse ESC differentiation *in vitro* and during mouse gastrulation *in vivo* (5, 12, 13) were much reduced in *Mesp1* KO cells, supporting the absence of CP specification (fig. S6).

SPRING analysis of WT *Mesp1* expressing cells at E6.75 and E7.25 identified five distinct destination cell types (DCTs) protruding from a core of intermingled cells (Fig. 3A–B). All

cells present within the DCTs came from E7.25 embryos, consistent with cell fate diversification of *Mesp1*-derived lineages during the late stages of gastrulation. To further define the nature of the five DCTs, we identified genes with specifically elevated expression in each of the five groups (Fig. 3B–C, fig. S7 and Table S2). This analysis identified both known and previously unrecognized genes associated with cardiovascular development. DCT1 was enriched among others in *Sox7*, *Etv2*, *Tal1* transcripts, markers of the endothelial or endocardial lineage (14, 15). DCT2 was marked by the expression of *Hand1*, *Bmp4*, *Tnnc1*, *Tbx3*, *Hand2*, *Tbx20*, *Gata4*, *Myl4*, *Mef2c*, which are well-known CM markers (7, 16) (Fig. 3C, fig. S7 and Table S2). *Bmp4* promotes CM differentiation (17). Moreover, *Hand1* lineage tracing showed that *Hand1* expressing cells contribute to the LV and to the myocardial and epicardial lineages with no contribution to the endocardium (16). These data suggest that DCT2 corresponds to CPs committed to the CM lineage. In situ hybridization of *Sox7* (EC marker) and *Bmp4* (CM marker) on *Mesp1-rtTA/tetO-H2B-GFP* embryos showed that these two markers did not co-localize in *Mesp1*-H2B-GFP expressing cells at E7.25, consistent with the notion that DCT1 and DCT2 marks two distinct *Mesp1* populations committed to EC and CM differentiation (Fig. 3 D). Unsupervised hierarchical clustering of DCT1 and DCT2 markers with CM and EC markers identified by scRNA-Seq of more mature mouse hearts at E10.5 (18) further showed that DCT1 and DCT2 cells clustered respectively with EC and CM lineage (Fig. 3E).

DCT3 and DCT4 are enriched in genes expressed and regulating SHF development (*Tbx1*, *Foxc2*, *Hoxb1*, *Hoxa1*) (Fig. 3C, fig. S7 and Table S2) (19–21). Lineage tracing experiments have previously shown that *Tbx1* (DCT3) marks anterior SHF progenitors whereas *Hoxb1* and *Hoxa1* (DCT4) mark posterior SHF (19, 21), suggesting that DCT3 and DCT4 correspond to *Mesp1* CPs committed to the anterior and posterior SHFs. RNA-FISH experiments further showed that *Foxc2* (DCT3) and *Wnt2b* (DCT4) were largely non-overlapping with *Wnt2b* localized closer to the PS whereas *Foxc2* expressing cells were found more antero-laterally at E7.25 (Fig. 3F–G). A subset of SHF progenitors, called cranio-pharyngeal progenitors, express *Tcf21* and contribute to the formation of some head muscles (22). RNA-FISH showed that *Tcf21* is preferentially expressed in a subset of DCT3 *Mesp1* H2B-GFP expressing *Foxc2* whereas little overlap was observed between DCT4 progenitors expressing *Wnt2* and *Tcf21* (fig S8). Consistent with the notion that DCT3 marks cranio-pharyngeal progenitors, the “pharyngeal” cluster found in our previous Flk1+ scRNA-seq (8) are closest to DCT3 (fig. S7). DCT5 expressed endoderm markers such as *Sox17* and *Foxa2*, and may have no relation with cardiac development (fig. S7 and Table S2).

Based on the cumulative evidence suggesting that our scRNA-seq capture the developmental progression from epiblast to early cardiovascular lineage segregation, we next investigated the expression of genes with known biological function, which revealed specific expression of cardiovascular, mesodermal and other genes regulating signaling pathways across the various DCT populations (fig. S9). Analysis of genes dynamically expressed during the trajectory towards DCT1 and DCT2 revealed the existence of gene clusters peaking in expression at sequential points along the pseudotemporal ordering. These gene clusters showed enrichment for gene ontology categories associated with developmental progression where the predicted trajectory provided consistency between the pseudotime and the real

developmental time (fig. S10 and Table S4). Moreover, cells coexpressing genes enriched in both DCT3 and DCT4 presented early pseudotime values than either DCT3 or DCT4 cells, consistent with the presence of immature progenitors that undergo multilineage priming before making cell fate decision into either DCT3 or DCT4 lineages (fig. S11).

To more clearly differentiate between the putative differentiation paths to the DCT1/EC and DCT2/CM progenitors, we determined the genes involved in embryonic development that are specifically up/down regulated in a given DCT (fig. S12). Of particular interest, DCT2 cells showed reduced *Notch1* expression (Fig. 4A). Different studies have shown the importance of Notch1 in the latter stages of cardiovascular development, in regulating endocardium, valve formation, trabeculation and myocardium compaction (23). However, a role for Notch in the early steps of cardiovascular lineage commitment during gastrulation has not been previously described. RNA-FISH and immunostaining showed that *Notch1* was indeed expressed and active in a subset of *Mesp1*-H2B-GFP+ cells (Fig. 4B and fig S13). To determine whether *Notch1* expression in DCT1 cluster marks *Mesp1*+ CPs committed to the EC fate, we induced lineage tracing by administering tamoxifen to *Notch1-CreERT2/Rosa-tdTomato* mice at E6.5 (24) and assessed the fate of marked cells at E12.5 (Fig. 4C–G). Interestingly, tdTomato+ cells in the ventricles were almost exclusively found in the EC (83.9% $\pm$ 3.0) with minor contribution to the CM (6.0% $\pm$ 1.4) and to the EPs (10.2% $\pm$ 2.1) (Fig. 4H), consistent with the notion that DCT1 marks early *Mesp1*/*Notch1* double positive CPs committed to the EC fate. In the atria, *Notch1-CreERT2* marked preferentially the ECs (76.9% $\pm$ 2.9), although a smaller contribution to the CM lineage was also observed (16.6% $\pm$ 2.7) (Fig. 4I), consistent with a lower expression of *Notch1* in DCT2 cluster. In addition to marking ECs of the heart, *Notch1-CreERT2*, similar to *Mesp1-Cre*, also marked, at the early stage of gastrulation, ECs of the aorta, intersomitic and brain vessels (fig. S14).

Altogether, our single cell profiling of early CPs shows that *Mesp1* CPs segregate rapidly from the epiblast into distinct cardiovascular lineages. The analysis of *Mesp1* KO cells showed that *Mesp1* is required to exit the pluripotent state of the epiblast and promotes EMT, migration and cardiovascular specification *in vivo*. Consistent with the early regional and lineage segregation found by clonal analysis (4, 5), our scRNA-seq demonstrates that *Mesp1* CPs are also molecularly heterogeneous as previously suggested by sc-RNAseq during *in vitro* ESC differentiation (25) and identifies temporally and spatially distinct *Mesp1* subpopulations that likely correspond to CPs committed to the different cardiovascular lineages and regions of the heart at the early stages of gastrulation (fig. S15). Future studies will be required to determine whether the early cardiovascular lineage segregation uncovered here is also occurring for the other mesodermal and endodermal cells, and whether defects in the early commitment steps are associated with congenital diseases and heart malformations. Finally, our results will be important to design new strategies to direct the differentiation of ESC into a specific cardiovascular lineage.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the LIMF-imaging facility for help with confocal imaging and Fiona Hamey for help with SPRING analysis. F.L. has been supported by the FNRS, the EMBO long-term fellowship and Leducq Fondation. Work in the Gottgens laboratory is supported by grants from the Wellcome Trust, Bloodwise, Cancer Research UK, NIDDK and core support grants by the Wellcome Trust to the Wellcome Trust -MRC Cambridge Stem Cell Institute. C.B. is an investigator of WELBIO. Work in C.B.'s laboratory was supported by the FNRS, the ULB foundation, the European Research Council (ERC), the foundation Bettencourt Schueller (C.B. and F.L.) and the Leducq Fondation as part of the network « 22q11.2 deletion syndrome: Novel approaches to understand cardiopharyngeal pathogenesis ».

References

1. Martin-Puig S, Wang Z, Chien KR, Lives of a heart cell: tracing the origins of cardiac progenitors. *Cell stem cell* 2, 320–331 (2008). [PubMed: 18397752]
2. Saga Y et al., *MesP1* is expressed in the heart precursor cells and required for the formation of a single heart tube. *Development (Cambridge, England)* 126, 3437–3447 (1999).
3. Bondue A, Blanpain C, *Mesp1*: a key regulator of cardiovascular lineage commitment. *Circulation research* 107, 1414–1427 (2010). [PubMed: 21148448]
4. Devine WP, Wythe JD, George M, Koshiba-Takeuchi K, Bruneau BG, Early patterning and specification of cardiac progenitors in gastrulating mesoderm. *eLife* 3, (2014).
5. Lescroart F et al., Early lineage restriction in temporally distinct populations of *Mesp1* progenitors during mammalian heart development. *Nature cell biology* 16, 829–840 (2014). [PubMed: 25150979]
6. Harel I et al., Distinct origins and genetic programs of head muscle satellite cells. *Developmental cell* 16, 822–832 (2009). [PubMed: 19531353]
7. Buckingham M, Meilhac S, Zaffran S, Building the mammalian heart from two sources of myocardial cells. *Nature reviews. Genetics* 6, 826–835 (2005).
8. Scialdone A et al., Resolving early mesoderm diversification through single-cell expression profiling. *Nature* 535, 289–293 (2016). [PubMed: 27383781]
9. Bondue A et al., *Mesp1* acts as a master regulator of multipotent cardiovascular progenitor specification. *Cell stem cell* 3, 69–84 (2008). [PubMed: 18593560]
10. Martello G, Smith A, The nature of embryonic stem cells. *Annual review of cell and developmental biology* 30, 647–675 (2014).
11. Lindsley RC et al., *Mesp1* coordinately regulates cardiovascular fate restriction and epithelial-mesenchymal transition in differentiating ESCs. *Cell stem cell* 3, 55–68 (2008). [PubMed: 18593559]
12. Bondue A et al., Defining the earliest step of cardiovascular progenitor specification during embryonic stem cell differentiation. *The Journal of cell biology* 192, 751–765 (2011). [PubMed: 21383076]
13. Kattman SJ et al., Stage-specific optimization of activin/nodal and BMP signaling promotes cardiac differentiation of mouse and human pluripotent stem cell lines. *Cell stem cell* 8, 228–240 (2011). [PubMed: 21295278]
14. Behrens T et al., Occupational prestige, social mobility and the association with lung cancer in men. *BMC cancer* 16, 395 (2016). [PubMed: 27388894]
15. Palencia-Desai S et al., Vascular endothelial and endocardial progenitors differentiate as cardiomyocytes in the absence of *Etsrp/Etv2* function. *Development (Cambridge, England)* 138, 4721–4732 (2011).
16. Barnes RM, Firulli BA, Conway SJ, Vincentz JW, Firulli AB, Analysis of the *Hand1* cell lineage reveals novel contributions to cardiovascular, neural crest, extra-embryonic, and lateral mesoderm derivatives. *Developmental dynamics : an official publication of the American Association of Anatomists* 239, 3086–3097 (2010). [PubMed: 20882677]
17. Schultheiss TM, Burch JB, Lassar AB, A role for bone morphogenetic proteins in the induction of cardiac myogenesis. *Genes & development* 11, 451–462 (1997). [PubMed: 9042859]

18. DeLaughter DM et al., Single-Cell Resolution of Temporal Gene Expression during Heart Development. *Developmental cell* 39, 480–490 (2016). [PubMed: 27840107]
19. Huynh T, Chen L, Terrell P, Baldini A, A fate map of Tbx1 expressing cells reveals heterogeneity in the second cardiac field. *Genesis (New York, N.Y. : 2000)* 45, 470–475 (2007).
20. Kume T, Jiang H, Topczewska JM, Hogan BL, The murine winged helix transcription factors, Foxc1 and Foxc2, are both required for cardiovascular development and somitogenesis. *Genes & development* 15, 2470–2482 (2001). [PubMed: 11562355]
21. Bertrand N et al., Hox genes define distinct progenitor sub-domains within the second heart field. *Developmental biology* 353, 266–274 (2011). [PubMed: 21385575]
22. Harel I et al., Pharyngeal mesoderm regulatory network controls cardiac and head muscle morphogenesis. *Proceedings of the National Academy of Sciences of the United States of America* 109, 18839–18844 (2012). [PubMed: 23112163]
23. MacGrogan D, Nus M, de la Pompa JL, Notch signaling in cardiac development and disease. *Current topics in developmental biology* 92, 333–365 (2010). [PubMed: 20816401]
24. Fre S et al., Notch lineages and activity in intestinal stem cells determined by a new set of knock-in mice. *PloS one* 6, e25785 (2011).
25. Chan SS, Chan HHW, Kyba M, Heterogeneity of Mesp1+ mesoderm revealed by single-cell RNA-seq. *Biochemical and biophysical research communications* 474, 469–475 (2016). [PubMed: 27131741]
26. Tumber T et al., Defining the epithelial stem cell niche in skin. *Science (New York, N. Y.)* 303, 359–363 (2004).
27. Madisen L et al., A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nat Neurosci* 13, 133–140 (2010). [PubMed: 20023653]
28. Srinivas S et al., Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. *BMC developmental biology* 1, 4 (2001). [PubMed: 11299042]
29. Picelli S et al., Full-length RNA-seq from single cells using Smart-seq2. *Nature protocols* 9, 171–181 (2014). [PubMed: 24385147]
30. Wu TD, Nacu S, Fast and SNP-tolerant detection of complex variants and splicing in short reads. *Bioinformatics (Oxford, England)* 26, 873–881 (2010).
31. Anders S, Pyl PT, Huber W, HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics (Oxford, England)* 31, 166–169 (2015).
32. Lun AT, McCarthy DJ, Marioni JC, A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. *F1000Research* 5, 2122 (2016). [PubMed: 27909575]
33. Leek JT et al., sva: Surrogate Variable analysis. (2017).
34. Brennecke P et al., Accounting for technical noise in single-cell RNA-seq experiments. *Nature methods* 10, 1093–1095 (2013). [PubMed: 24056876]
35. Becker RA, Chambers JM, Wilks AR, *The New S Language*. Wadsworth & Brooks/Cole., (1998).
36. Mardia KV, Kent JT, Bibby JM, *Multivariate Analysis*. London: Academic Press, (1979).
37. Venables WN, Ripley BD, Eds., *Modern Applied Statistics with S*, (Springer-Verlag, 2002).
38. Angerer P et al., destiny: diffusion maps for large-scale single-cell data in R. *Bioinformatics (Oxford, England)* 32, 1241–1243 (2016).
39. Weinreb C, Wolock S, Klein A, SPRING: a kinetic interface for visualizing high dimensional single-cell expression data. *BioRxiv*, (2017).
40. Qiu X et al., Reverse graph embedding resolves complex single-cell developmental trajectories. *BioRxiv*, (2017).
41. Setty M et al., Wishbone identifies bifurcating developmental trajectories from single-cell data. *Nature biotechnology* 34, 637–645 (2016).
42. Haghverdi L, Buttner M, Wolf FA, Buettner F, Theis FJ, Diffusion pseudotime robustly reconstructs lineage branching. *Nature methods* 13, 845–848 (2016). [PubMed: 27571553]
43. Qiu X et al., Reversed graph embedding resolves complex single-cell trajectories. *Nature methods* 14, 979–982 (2017). [PubMed: 28825705]
44. Love MI, Huber W, Anders S, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology* 15, 550 (2014). [PubMed: 25516281]

45. Chiapparo G et al., Mesp1 controls the speed, polarity, and directionality of cardiovascular progenitor migration. *The Journal of cell biology* 213, 463–477 (2016). [PubMed: 27185833]
46. Langmead B, Trapnell C, Pop M, Salzberg SL, Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10, R25 (2009). [PubMed: 19261174]
47. Zhang Y et al., Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 9, R137 (2008). [PubMed: 18798982]
48. McLean CY et al., GREAT improves functional interpretation of cis-regulatory regions. *Nature biotechnology* 28, 495–501 (2010).
49. Kent WJ et al., The human genome browser at UCSC. *Genome research* 12, 996–1006 (2002). [PubMed: 12045153]
50. DeTomaso D, Yosef N, FastProject: a tool for low-dimensional analysis of single-cell RNA-Seq data. *BMC bioinformatics* 17, 315 (2016). [PubMed: 27553427]
51. Subramanian A et al., Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America* 102, 15545–15550 (2005). [PubMed: 16199517]

Figure 1: scRNA-seq of *Mesp1*+ CPs fills the gap between E6.5 epiblast cells and E7.5 mesodermal cells.

A. Scheme of the experimental strategy used for isolating *Mesp1*-expressing CPs in vivo. Scale: 200µm. **B.** SPRING plot of 892 cells showing *Mesp1* CPs at E6.75 and E7.25 and the published epiblast cells (E6.5_Scia) and E7.5 Flk1+ progenitors (E7.5_Scia) with read count of *Mesp1* > 0. **C.** SPRING plot coloured by the inferred pseudotime time for all 892 cells.

Figure 2: *Mesp1* controls the exit from pluripotency, EMT and cardiovascular specification.

A. SPRING plot of all 892 cells including *Mesp1* KO cells coloured by cell types **B.**

Pseudotime time distribution for WT and *Mesp1* KO cells at E6.75. **C.** Comparison of the genes differentially expressed in scRNA-Seq experiments between control and *Mesp1* KO cells and the genes regulated by *Mesp1* gain of function (GOF) in ESC. The 58 genes in agreement with the scRNA-Seq experiment with FDR < 0.1 were highlighted in red. The significance of the overlap was calculated by hypergeometric test using the phyper function in R. **D.** Gene ontology enrichment for genes downregulated in *Mesp1* KO cells. **E-G.** Violin plots showing the mean and variance difference between WT and *Mesp1* KO cells of genes regulating pluripotency (Nanog, Eras) (**E**), EMT (Cdh1; Snai1) (**F**) and cardiovascular fate (Gata4, Etv2, Myl7) (**G**).

Figure 3: *Mesp1* single cell analysis identifies different progenitors committed to different fates and heart regions.

A. SPRING analysis of the 807 WT *Mesp1*-H2B-GFP+ cells at E6.75 and E7.25 and *Mesp1*+ Scia cells. **B.** The 5 endpoints (EPs) revealed by SPRING analysis were considered as 5 distinct cell types (DCT1–5). **C.** Expression of key genes specific for DCT1–4. **D.** RNA-FISH of *Sox7* and *Bmp4* on sections of *Mesp1*-rtTA/tetO-H2B-GFP embryos at E7.25. **E.** Heatmap of DCT1 and DCT2 endpoint cells based on unsupervised clustering of the expression of CM and EC marker genes identified at E10.5 (24) combined with newly identified genes enriched in DCT1 or DCT2. **F-G.** RNA-FISH of *Foxc2* and *Wnt2b* on sections of *Mesp1*-rtTA/tetO-H2B-GFP embryos at E7.25. Higher magnification is found in G. Scale: 50 μm.

Figure 4: Notch1 marks *Mesp1* progenitors committed to the endocardial fate.

A. Violin plot of *Notch1* expression in DCT1 and DCT2 cells. **B.** *Notch1* RNA-FISH on a section of an E7.25 *Mesp1-rtTA/tetO-H2B-GFP* embryo. scale: 50 μ m. **C.** Experimental strategy used for tracing Notch1 expressing cells at E6.5. **D-G.** Confocal analysis of immunostaining for endoglin (EC marker) (D-F) and cTNT (CM marker) (G) of *Notch1-CreERT2/Rosa-tdTomato* heart sections at E12.5. Lower magnification (D) Scale: 500 μ m. Higher magnifications of the LV (E) and RA (G) showed that most tdTomato + cells are ECs although rare CMs are also marked (asterisks). Scale: 200 μ m. **H, I.** Percentage of tdTomato positive cells in ECs, CMs and EPs in the ventricles (H) (7804 cells counted; n=6 embryos from 3 different litters) and in the atria (I) (4819 cells counted; n= 7 embryos from 3 different litters).