



Published in final edited form as:

Cell Rep. 2023 February 28; 42(2): 112086. doi:10.1016/j.celrep.2023.112086.

Single-nucleus RNA sequencing in ischemic cardiomyopathy reveals common transcriptional profile underlying end-stage heart failure

Bridget Simonson^{1,2,7}, Mark Chaffin^{1,2,7}, Matthew C. Hill^{1,2,3}, Ondine Atwa^{1,2}, Yasmine Guedira^{1,2}, Harshit Bhasin^{1,2}, Amelia W. Hall^{1,4}, Sikander Hayat⁵, Simon Baumgart⁵, Kenneth C. Bedi Jr.⁶, Kenneth B. Margulies⁶, Carla A. Klattenhoff⁵, Patrick T. Ellinor^{1,2,3,8,*}

¹Precision Cardiology Laboratory, The Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA

²Cardiovascular Disease Initiative, The Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA

³Cardiovascular Research Center, Massachusetts General Hospital, Boston, MA 02114, USA

⁴Gene Regulation Observatory, The Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA

⁵Precision Cardiology Laboratory, Bayer US, LLC, Cambridge, MA 02142, USA

⁶Penn Cardiovascular Institute, University of Pennsylvania, Philadelphia, PA 19104, USA

⁷These authors contributed equally

⁸Lead contact

In brief

Simonson et al. compare non-failing and ischemic cardiomyopathy human left ventricles using snRNA-seq and identify an increase in lymphatic and angiogenic endothelial cells in disease. Transcriptionally, ischemic cardiomyopathy is similar to other cardiomyopathies. Gene targets

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

*Correspondence: ellinor@mgh.harvard.edu.

AUTHOR CONTRIBUTIONS

M.C., B.S., C.A.K., and P.T.E. conceived and designed the study. B.S. generated snRNA-seq data. M.C. analyzed the snRNA-seq data. B.S., M.C., M.C.H., A.W.H., S.H., S.B., and P.T.E. interpreted snRNA-seq analyses. O.A. performed RNA *in situ* hybridization. K.C.B. and K.B.M. acquired LV samples for this study. M.C., B.S., M.C.H., and P.T.E. drafted the manuscript, which was revised by all authors.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.celrep.2023.112086>.

DECLARATION OF INTERESTS

C.A.K. is an employee of Bayer US LLC (a subsidiary of Bayer AG) and may own stock in Bayer. S.H., S.B., and C.A.K. were full-time employees of Bayer when this work was performed. P.T.E. has received sponsored research support from Bayer AG, IBM Health, Bristol Myers Squibb, and Pfizer; he has also served on advisory boards or consulted for Bayer AG, MyoKardia, and Novartis. K.B.M. has research grant funding from Amgen, USA and has also served on advisory boards for MyoKardia, Bristol-Myers Squibb, and Pfizer.

INCLUSION AND DIVERSITY

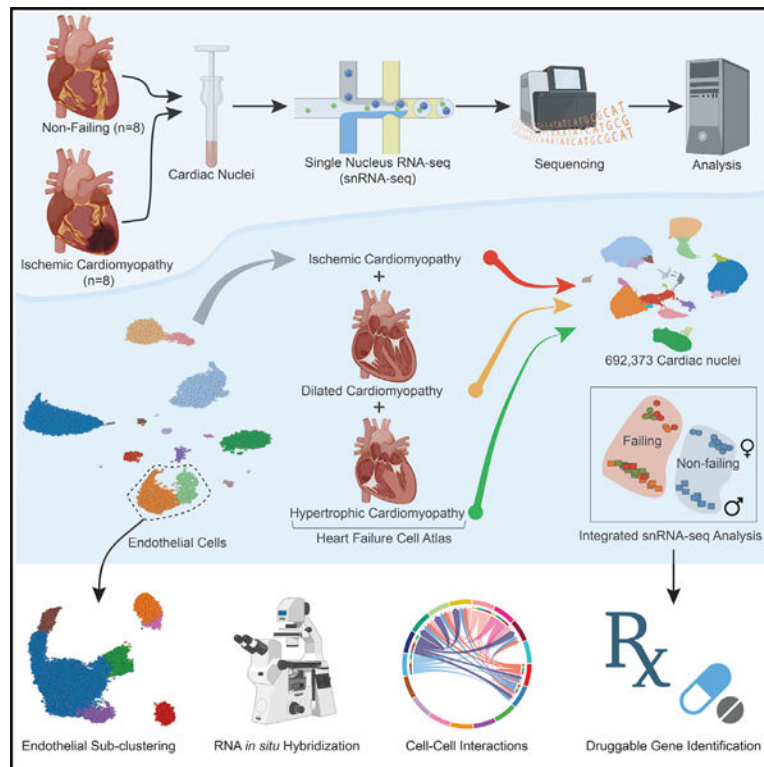
We support inclusive, diverse, and equitable conduct of research.

from this shared expression profile could be explored for therapeutic benefit in end-stage heart failure.

SUMMARY

Ischemic cardiomyopathy (ICM) is the leading cause of heart failure worldwide, yet the cellular and molecular signature of this disease is largely unclear. Using single-nucleus RNA sequencing (snRNA-seq) and integrated computational analyses, we profile the transcriptomes of over 99,000 human cardiac nuclei from the non-infarct region of the left ventricle of 7 ICM transplant recipients and 8 non-failing (NF) controls. We find the cellular composition of the ischemic heart is significantly altered, with decreased cardiomyocytes and increased proportions of lymphatic, angiogenic, and arterial endothelial cells in patients with ICM. We show that there is increased LAMININ signaling from endothelial cells to other cell types in ICM compared with NF. Finally, we find that the transcriptional changes that occur in ICM are similar to those in hypertrophic and dilated cardiomyopathies and that the mining of these combined datasets can identify druggable genes that could be used to target end-stage heart failure.

Graphical abstract



INTRODUCTION

Ischemic cardiomyopathy (ICM) is the most prevalent type of heart disease, affecting 126 million people worldwide and causing 9 million deaths in 2017.¹ ICM is most commonly caused by atherosclerotic coronary artery disease inhibiting blood flow to the heart. This

leads to decreased oxygen in the myocardium causing cardiomyocyte death and cardiac remodeling via formation of a fibrotic scar. Over time the lack of oxygen and formation of the scar leads to a loss of contractile function, cardiac hypertrophy, dilation, and heart failure.² Current therapeutics for ICM include medical interventions such as ACE inhibitors and beta blockers, and surgical revascularization therapy; however, for patients progressing to heart failure the long-term prognosis remains poor.^{3,4}

ICM has been studied in detail in animal models where disease progression can be monitored closely over time.^{4,5} These efforts have identified specific cell types, genes, and pathways that play crucial roles in ICM. One important cardiac cell type responsible for responding to ischemic injury is vascular endothelial cells. Endothelial cells play an important role throughout the remodeling of the heart after ischemia.⁶ During the initial immune response, they contribute to white blood cell infiltration to the infarcted region, and they release cytokines and chemokines, which contribute to fibroblast activation and cardiomyocyte health.⁷ The formation of functional new vessels from the pre-existing vascular network of endothelial cells, known as angiogenesis, is essential to supply damaged tissue with oxygen and nutrients.⁸ During post-ischemic cardiac remodeling lymphangiogenesis also occurs, and lymphatic endothelial cells play a role in clearing fluid buildup, debris, and dead cells from the infarct.⁹ Both angiogenesis and lymphangiogenesis are processes that have been targeted as having therapeutic potential for ICM, with mixed results shown in animal studies and in clinical trials.^{10,11}

As developing therapeutic targets for such a dynamic cardiovascular disease is challenging, there is a need for greater cellular and transcriptional information from human ICM patients. Advances in single-cell and single-nucleus RNA sequencing (snRNA-seq) have allowed a greater understanding of the underlying transcriptional profile of the healthy^{12,13} and diseased^{14,15} human heart. Here, we utilize snRNA-seq to profile the left ventricle (LV) of transplant recipients with ICM and non-failing (NF) controls.

There were three aims for this study. First, to broadly characterize the cellular composition and transcriptional changes seen in the non-infarct region of the LV of ICM patients. Second, to assess the status of the healthy and ischemic vasculature. And finally, to compare the transcriptional landscape of ICM with other forms of end-stage heart failure, including hypertrophic cardiomyopathy (HCM) and dilated cardiomyopathy (DCM). Together, our results show that there are large changes in both cellular composition and transcription in ICM compared with NF patients. These changes are similar to those seen in HCM and DCM, suggesting that these three types of heart failure share a similar end-stage transcriptional profile. Further investigation into this shared transcriptional profile may lead to identification of new potential therapeutic targets for end-stage heart failure.

RESULTS

In this study, we performed snRNA-seq on samples from the non-infarcted region of the LV of human patients with long-term ICM (n = 8) and NF controls (n = 8). After strict quality control, 1 ICM sample was removed leaving a final dataset of 8 control and 7 ICM human samples (Table 1). After removal of low-quality nuclei, 99,684 nuclei (39,339 ICM; 60,345

NF) across 16 clusters remained (Figures 1A and S1; Table S1). Cell-type labels were assigned to clusters based on the expression of canonical marker genes and the enrichment of gene ontology-based biological processes (Tables S2 and S3). As shown previously, the most abundant cell types were cardiomyocytes (31.6%), fibroblasts (22.6%), and endothelial cells (20.5%)¹³ (Figure 1B). Compared with NF controls, we observed significant decreases in cardiomyocytes and endothelial II populations in ICM, with a non-significant increase in fibroblasts (Figure 1C). Masson trichrome staining of tissue sections from these samples showed a large range of fibrosis in the ICM samples, which, when quantified, corresponded with the proportion of fibroblasts from the sample-matched snRNA-seq data (Figures S2A–S2C). We next performed an exploratory differential expression test to identify transcriptional changes in fibroblasts correlating with the degree of fibrosis in ICM samples. The strongest changes were observed for *BMP6*, *KCNMA1*, *CLSTN2*, *TEX41*, *MALRD1*, and *ACSM3* (Figure S2D). Interestingly, *BMP6* is a member of the TGF- β superfamily,¹⁶ and *KCNMA1* is a subunit of the BK channel which regulates the degree of α -SMA expression after TGF- β stimulation in pulmonary fibroblasts.¹⁷

Changes in gene expression and pathways in ICM compared with NF

To identify transcriptional differences between ICM and NF, we carried out differential gene expression testing on the main cell clusters by aggregating nuclei across samples into a pseudo-bulk representation (see STAR Methods). We observed the highest number of differentially expressed genes (DEGs) in fibroblasts ($n = 1,525$) followed by cardiomyocytes ($n = 513$), endothelial I cells ($n = 457$), and macrophages ($n = 415$) (Figures 1D and S3A; Table S4).

The top up- and downregulated genes for each cell type are shown in Figure S3A. Interesting DEGs in fibroblasts include: *NTM*, which has been identified as a novel biomarker for heart failure,¹⁸ and *ANO1*, which plays an important role in fibrosis in a rat model of myocardial infarction.¹⁹ In cardiomyocytes, *PLCE1* has been shown to play a role in cardiac hypertrophy via its interaction with the scaffolding protein mAKAP6.²⁰ *AZGP1* is a secreted protein that has been shown to have antifibrotic effects in the kidney and the heart²¹ and *CALCOCO2*²² is a key protein in autophagy and its expression is associated with cardiac dysfunction in DCM. Finally, in the macrophages, *ABCA9* is an important gene in monocyte to macrophage differentiation²³ and *CD163L1* expression in macrophages is associated with tissue resident anti-inflammatory macrophages in cancer.²⁴

To determine if these transcriptional changes in ICM were enriched in specific pathways, we performed pathway analysis using the gene ontology biological processes²⁵ and Reactome pathway databases.²⁶ In ICM, differentially regulated pathways were identified in cardiomyocytes, fibroblasts, endothelial I and II cells, macrophages, and pericytes (Figures S3B–S3E; Table S5). In fibroblasts, there was increased expression of genes involved in the extracellular matrix organization and collagen formation pathways. In cardiomyocytes, there was increased expression of genes involved in immune effector processes, and decreased expression of genes involved in metabolism. Notably, in endothelial cells we observed increased expression of genes involved in cell proliferation and angiogenesis, including

activation of vascular endothelial growth factor (VEGF) signaling pathways, and regulation of cellular response to growth factor stimulus.

Endothelial cell diversity in ICM

Endothelial cells play a crucial role in the remodeling of the heart post myocardial infarction.²⁷ Due to the enrichment of pathways involved in endothelial cell proliferation and angiogenesis in late-stage ICM, we further investigated the role of vascular cells in ICM. Specifically, we sought to characterize the cellular diversity of endothelial cells and look for subpopulations enriched in ICM. To do this, we performed subclustering on the combination of the endothelial I, endothelial II, endothelial III, and lymphatic endothelial populations, and from this we identified seven distinct subclusters (Figures 2A and 2B; Tables S6 and S7). We assigned endothelial subtype classifications for six subclusters (EC-angiogenic, EC-arterial, EC-endocardial I, EC-endocardial II, EC-general capillary, EC-lymphatic) by comparing their expression profiles with published data^{28–30} and assessing enriched ontologies (Figures S4A and S4B). One additional subcluster, EC-*KCNIP4*, could not be assigned to known endothelial subtype classifications.

We observed distinct shifts in endothelial cell composition in ICM compared with NF samples, including an increase in EC-lymphatic, EC-angiogenic, and EC-arterial cells (Figures 2C and 2D). Increases in EC-lymphatic cells were consistent across most ICM samples, whereas increases in EC-arterial cells were mostly found in one ICM sample (P1579). Increases in EC-angiogenic cells were mostly found in ICM samples P1579 and P1364. We also saw a non-significant increase in EC-endocardial I and II subclusters in NF compared with ICM, which was mostly derived from sample P1785 (Figure 2C). To validate the shifts observed in EC-lymphatic and EC-angiogenic cells we carried out *in situ* hybridization. Using the EC-lymphatic marker gene *CCL21* ($\log_{2}FC = 5.2$, $p = 4.0 \times 10^{-22}$) we observed increased numbers of EC-lymphatic cells in the ICM samples that were clustered in fibrotic regions in the tissue (Figures 2E, S5A, and S5B). To validate the EC-angiogenic cells, we co-stained with the endothelial cell marker *PECAM1* and EC-angiogenic cell-specific marker *MYO1B* ($\log_{2}FC = 2.15$, $p = 1.7 \times 10^{-14}$). The double-positive EC-angiogenic cells were largely found in the capillaries throughout the tissue (Figures 2F and S5C).

To further understand the differences in endothelial cells in ICM compared with NF we examined the top differentially regulated genes from each subcluster (Figure 3; Table S8). Some genes were differentially expressed between ICM and NF patients in most subclusters of endothelial cells. For example, *HIF3A*, a hypoxia-induced transcription factor, was significantly decreased in ICM patients compared with NF patients in EC-angiogenic, EC-*KCNIP4*, EC-arterial, EC-general capillary, and EC-endocardial I cells. *HIF3A* is a negative regulator of HIF1a/2a response to hypoxia.³¹ Knockout of a variant of *HIF3A* that is expressed in neonatal and embryonic cells in mice leads to the enlargement of the right ventricle of mice, and increased endothelin-1 expression in lung endothelial cells.³² We also saw decreased expression of *SYNE1* in ICM patients compared with NF patients in EC-general capillary, EC-*KCNIP4*, and EC-angiogenic cells. *SYNE1* encodes the nesprin 1

protein, which plays an important role in regulating cell shape, angiogenic loop formation, and migration of endothelial cells.³³

We next examined DEGs in the three subclusters with increased prevalence in ICM compared with NF: EC-arterial, EC-angiogenic, and EC-lymphatic cells (Figure 3). Comparing ICM with NF patients in the EC-arterial population, there was an increase in *INSR* (insulin receptor), an important gene in angiogenesis.³⁴ There was also an increase in *XAF1* and decrease in *GSN*, two genes involved in regulation of apoptosis.^{35,36} Finally, *SMAD6*, an anti-inflammatory and a negative regulator of BMP and TGF- β signaling, was decreased in ICM patients compared with NF patients.³⁷

Comparing ICM with NF patients in the EC-angiogenic population, we saw an increase in *INSR* and a decrease in *PIK3C2A*, a protein previously shown to be decreased in the plasma of patients after an acute myocardial infarction.³⁸ *AFAP1L1*, which is involved in actin stress fiber formation,³⁹ was increased in both the EC-angiogenic and EC-general capillary populations (Figure 3).

Within the EC-lymphatic subcluster, *MMRNI*, was decreased in ICM compared with NF. *MMRNI* is an ECM protein that is a marker gene for lymphatic endothelial cells and binds to collagens I, II, and III, which leads to enhanced activated platelet adhesion and thrombus formation after injury.⁴⁰ EC-lymphatic and EC-endocardial I nuclei also had high expression of the tumor suppressor gene *PKHD1L1* in NF samples, which was strongly decreased in ICM. *PKHD1L1* plays a role in inhibiting proliferation and cell invasion of cancer cells⁴¹ (Figure 3).

We were also interested in understanding if there were any changes in cell-cell communication between endothelial cells and the other cell types in the heart. To do this, we performed an analysis using CellChat.⁴² We find an increase in overall communication involving endothelial cells in ICM compared with NF (Figure 4A), which is particularly pronounced in endocardial and angiogenic endothelial cells (Figure 4B). When considering the flow of information from endothelial sources to other cell types, we see a significant increase in CDH5, JAM, LAMININ, and COLLAGEN signaling (Figure 4C). Of note, there is little communication based on LAMININ signaling from endothelial cells in NF samples, which expands drastically when looking at the ICM condition (Figure 4D). This seems largely driven by an increase in the ligands *LAMA4*, *LAMB1*, and *LAMC1* in several endothelial subclusters, leading to broad communication with many other cell types (Figure 4E).

Proteomic comparison identifies genes and proteins associated with fibrosis

Next, to understand how transcriptional changes at single-nuclei resolution translate to bulk level protein expression, we compared our snRNA-seq dataset with a previously published bulk proteomic data from ICM patients.⁴³ When comparing the two datasets we found a number of genes that were significantly dysregulated between ICM and NF in either the snRNA-seq data or the proteomic data, but not changed in both (Figure S6A). However, there was a subset of genes that were significantly changed in both datasets ($n = 41$) (Figures S6A–S6C; Table S9). Interestingly, many genes that were increased in both datasets

were differentially expressed in fibroblasts and associated with ECM formation and fibrosis (*THBS4*, *AEBPI*, *COL14A1*, *VCAN*, *ACTNL1*). We also saw a concordant increase in *ARHGAP1* in cardiomyocytes from ICM. *ARHGAP1* is a promoter of apoptosis and has previously been shown to be increased in cardiomyocytes in response to hypoxia.⁴⁴

We identified 17 genes that were discordant. A high proportion of the genes that were increased in the proteomic data and decreased in the snRNA-seq data were secreted proteins (12/15: *VCAN*, *PRELP*, *C1QB*, *MFGE8*, *A2M*, *FBLN5*, *ERP44*, *AZGP1*, *C3*, *SERPING1*, *SERPINF1*, *SERPINF2*). The two genes that were increased in the snRNA-seq and decreased in the proteomic data were both mitochondrial genes: *HIBADH*, which was differentially expressed in macrophages, and *ALDH5A1*, which was differentially expressed in fibroblasts. This analysis suggests that there is some correlation between snRNA-seq and proteomics data that may be useful for furthering determining important genes and pathways changed in disease.

Late-stage ICM is transcriptionally similar to late-stage HCM and DCM

Finally, we compared the single-nuclei transcriptional profiles in ICM with a large, recently reported study of snRNA-seq on HCM and DCM samples.¹⁴ In brief, 692,373 nuclei from the ICM dataset and the HCM/DCM dataset were combined and jointly visualized on a shared UMAP representation (Figure 5A). When considering the aggregation of all cell types per sample in a principal-component analysis, we found that the ICM, DCM, and HCM samples clustered together, separate from the NF controls (Figure 5B). We then carried out a cellular composition test of our joint dataset (Figure S7). Overall, cellular composition was similar between DCM, HCM, and ICM, and we observed similar shifts compared with NF patients.

We next wanted to identify genes that were uniquely changed in ICM compared with DCM and HCM. To do this, we identified genes that were significantly changed in ICM vs. NF, as well as differentially expressed between ICM vs. DCM or ICM vs. HCM (Figures 5C and 5D). To control for technical differences between the two studies, we included a covariate representing the study in our differential expression test (see STAR Methods). This identified 17 unique genes, 4 that were significantly differentially expressed both between ICM vs. DCM and ICM vs. HCM, 11 that were significantly differentially expressed only in ICM vs. DCM, and 2 that were significantly differentially expressed only in ICM vs. HCM (Figure 5D). In addition, a small subset of genes were differentially expressed between DCM or HCM vs. NF as well as between DCM or HCM vs. ICM, but not between ICM and NF (Figures 5C and 5D). This included *MDM2*, *KIF9*, and *RSPH6A* in cardiomyocytes, *SELP* in endothelial II, and *TRIM73* in lymphocytes.

There were several interesting genes that were identified in this analysis that were unique to ICM and have been previously associated with ischemia. *CELSR1* was increased in ICM compared with NF and ICM compared with DCM in macrophages. *CELSR1* is an atypical cadherin that is associated with susceptibility to ischemic stroke in Japanese populations⁴⁵ and congenital heart disease associated pulmonary arterial hypertension.⁴⁶ *IREB2* was increased in pericytes in ICM compared with NF and in ICM compared with DCM and is involved in ferroptosis, a pathway that has been implicated in myocardial

infarction and heart disease.⁴⁷ *TMLHE* is the first enzyme in the carnitine/acylcarnitine biosynthesis pathway and was decreased in endothelial III cells in ICM compared with NF and in ICM compared with DCM. Knockdown of *TMLHE* is protective in mouse models of ischemia/reperfusion due to decreased long-chain acylcarnitine and ROS production, inhibiting mitochondria and cardiac damage.⁴⁸

There were also changes in genes that have not been previously linked to ICM or heart disease, but still had interesting biological functions. For example, *NPNT* was increased in ICM compared with NF and also in ICM compared with DCM in endothelial II nuclei. *NPNT* is an ECM protein that interestingly has been shown to promote cytokinesis of postnatal CMs.⁴⁹ Finally, *DTX3* was decreased in VSMC in ICM compared with NF but increased in both HCM and DCM compared with NF. *DTX3* is an E3 ligase that targets *NOTCH2* and activates proliferation in cancer.^{50,51}

We next wanted to explore the similarities observed in the three datasets and looked at the common pathways that were enriched for shared genes between ICM, HCM, and DCM (Figure 5E). We found that many mitochondrial-related pathways were decreased in CMs in disease. We saw an increase in kinase signaling pathways in endothelial I nuclei, with a decrease in signaling by BMP. Interestingly, there were no significant changes observed in fibroblast populations.

Identifying druggable genes for end-stage heart failure

Because of the limited pharmacologic treatments available to treat end-stage heart failure, we sought to identify druggable targets that were present in all three cardiomyopathies. To do so, we compared the list of genes that were strongly, and consistently, up- or downregulated in all cardiomyopathies with the druggable genome.⁵² The druggable genome was derived from Finan et al. and includes three tiers of genes based on how similar their features are to known drug targets.⁵² We identified 122 druggable genes that were downregulated in all 3 cardiomyopathies, and 52 that were upregulated in all 3 cardiomyopathies (Table S10). We further filtered the list for cell-type specificity (Figure 6A). This led to 10 genes that were druggable and upregulated (Figure 6B) in all 3 cardiomyopathies and 29 genes that were druggable and downregulated in all 3 cardiomyopathies (Figure 6C). Of these genes, *CYD2J2* plays an important compensatory role in heart disease⁵³ and several were collagens (*COL21A1*, *COL24A1*, *COL14A1*, and *COL23A1*). *MYH6*, a gene well known to be associated with heart failure, was also identified.⁵⁴ Several genes were identified that have been previously shown to play important roles in disease that may be druggable and have positive outcomes for heart disease, such as *SMOC2*,⁵⁵ *SULF1*,⁵⁶ *SPARCL1*,⁵⁷ *NRK*,⁵⁸ and *PRELP*⁵⁹ (Table S11).

Overall, these data suggest that, while there are some differences in late-stage ICM compared with late-stage HCM and DCM, there is a convergence of pathways in late-stage heart disease. By mining the shared enriched pathways and differently expressed genes we are able to identify potential druggable targets for end-stage heart disease.

DISCUSSION

Single cell RNA sequencing and snRNA-seq have allowed the indepth study of the transcriptome of the healthy and diseased human heart. Here, we carry out snRNA-seq on non-ischemic regions of LV samples from patients with ICM and NF controls. We created a high-quality map of nearly 100,000 nuclei and describe the cellular, transcriptional, and pathway changes observed in ICM compared with NF.

We observed that there was a population of angiogenic endothelial cells (EC-angiogenic) present in capillaries in end-stage disease. Angiogenesis occurs during the initial wound-healing phase of ICM and has been shown to be generally decreased after scar formation.²⁷ Interestingly, the two samples with the highest proportion of EC-angiogenic cells were the two samples that had the most recent myocardial infarctions (MIs) before transplantation (P1579, 9 months; P1364, 5 months). Increasing blood flow to ischemic tissue by increasing angiogenesis has been a focus of therapeutic development of ICM. Methods such as cell therapy and treatment with VEGF have shown promise in clinical trials, with some improvement in cardiac function observed.⁶⁰ The presence of these cell types in late-stage disease suggests that there is still active repair or compensation occurring in the LV of ICM patients.

We were also interested to see the increased number of lymphatic endothelial cells in fibrotic regions of the tissue in ICM compared with NF, especially as these samples are from non-infarcted regions of the LV. Lymphangiogenesis occurs during scar formation after the injury event to help clear fluid, immune cells, and debris from the heart.²⁷ Like angiogenesis, there has been much interest in increasing lymphangiogenesis as a therapy for ICM. Many studies in animal models have stimulated lymphangiogenesis using VEGF-C after MI, leading to increased clearing of immune cells from damaged tissue, smaller scars, and better heart function^{9,61}; however, this response has not always been observed^{62,63} and more work is needed in order to understand whether this would be a useful therapeutic pathway to pursue.

Using CellChat, we found a striking increase in LAMININ signaling in ICM from endothelial cells compared with NF, driven largely by increases in the ligands *LAMA4*, *LAMB1*, and *LAMC1* in disease. LAMININ signaling plays a crucial role in attachment of endothelial cells to the basement membrane, and influences the function of cardiomyocytes and other cell types in the heart.⁶⁴ Interestingly, mutations and decreased expression of *LAMA4* have been associated with heart failure.⁶⁵ *LAMB1* expression has been shown to increase in aging and in a model of acute MI in mice,⁶⁶ and has been hypothesized to play a role in pathophysiology in the heart by increasing endothelial cell to mesenchymal cell transition, leading to increased fibrosis.⁶⁶ Investigation into the role of endothelial cell signaling to other cell types through LAMININ signaling could help to further characterize the role that endothelial cells play in ICM through their interactions with other cell types.

A limitation of snRNA-seq data is that they only capture the transcriptional activity in the nucleus rather than the whole cell, which more closely represents the translated proteins. We therefore compared our snRNA-seq dataset and a previously published proteomic study

that also used samples taken from non-infarct regions of the LV of transplant recipients with ICM.⁴³ We found that there was some correlation of gene and protein expression in these samples. As expected, a large number of the proteins that were identified were changed at the mRNA level in cardiomyocytes, fibroblasts, and endothelial cells, the cells that make up the largest portion of the heart. We also saw proteins that showed correlated changes in smaller cell types such as pericytes and neuronal cells. Many of the concordant genes and proteins identified were in fibroblasts. As we observed increased fibrosis in ICM with trichrome staining, and fibrosis plays a large role in ICM this is not particularly surprising. However, more notable was the observation that many of the discordant genes that were decreased at the gene level but increased at the protein level were proteins that are secreted and play a role in the ECM. This could suggest that there is temporal regulation of the mRNA in the nucleus and an accumulation of the secreted proteins in the ECM. One example of a discordant ECM protein was *PRELP*, which was also identified as a druggable gene that was decreased in fibroblasts in all three cardiomyopathies. PRELP protein is increased in the LV of late-stage but not early-stage ICM in pigs⁶⁷ and in late-stage ICM in humans⁶⁸ and has been suggested to be a biomarker for end-stage heart disease.⁵⁹ *PRELP* is thought to be increased in an intermediary state along the fibroblast myofibroblast transition, and knockout of *PRELP* in fibroblasts leads to decreased activation of fibroblasts after TGB- β stimulation.¹⁴

The paradigm of the “final common pathway” of heart failure theorizes that all heart failures with reduced ejection fraction converge on the same biological pathway.⁶⁹ However, many studies have refuted this idea; for example, bulk RNA-seq⁷⁰ and proteomic⁴³ studies of ICM, DCM, and HCM transplant recipients have shown that there are significant differences between ICM and DCM. Our analysis of ICM compared with HCM and DCM found that, at the nuclear transcriptional level, there are minimal differences in ICM compared with HCM and DCM, supporting the idea of a final common pathway. Current therapies for end-stage heart failure such as implantation of ventricular assist devices and heart transplantation are limited, costly, and highly invasive; therefore, further investigation into the role of these shared pathways and genes in late-stage heart disease may identify areas that could be targeted for therapeutic development. For example, the role of mitochondria dysfunction in the failing heart has been well established,⁷¹ and our data show several mitochondrial-related pathways downregulated in cardiomyocytes across all three cardiomyopathies. Improving mitochondria health and function is a challenging, but active area of interest for pharmacological intervention in heart disease.⁷² Integration of snRNA-seq datasets into existing knowledge around mitochondrial dysfunction in heart disease could help to identify key genes and proteins that could be targeted to restore cardiomyocyte bioenergetics.

In summary, we provide an in-depth analysis of the viable left ventricular myocardium of patients with ICM. We found an increased ratio of fibroblasts to cardiomyocytes and many cell-type-specific transcriptional changes, including activation of angiogenesis in endothelial cells and increased lymphatic endothelial cells. We also show changes in signaling from endothelial cells to other cell types in ICM, with an increase in LAMININ signaling. Finally, within each cell type, including cardiomyocytes and fibroblasts, our data support a

convergence of the transcriptional profile in advanced heart failure, and we have utilized the druggable genome to identify potential therapeutic targets for end-stage heart failure.

Limitations of the study

We recognize that this study has several limitations, many of which have been discussed previously.¹³ The samples we used in this study were from the non-infarct region of the free wall of the LV from ICM patients undergoing heart transplantation or from regions of the free wall of the LV of NF controls. The samples were taken from similar locations from each patient; however, as it is impossible to be exact, there may be variability if there are location-based transcriptional differences, especially in regard to where the sample was taken in relation to the infarct region. Our comparison of snRNA-seq data with proteomic data could be confounded due to the limited number of samples included in the analysis and the samples being taken from different patients. As the samples for this study and the HCM/DCM study were from patients undergoing transplantation, our comparison of transcriptional differences between diseases may be limited as all patients were at end-stage heart failure. ICM may diverge from other forms of heart disease within the infarct and peri-infarct regions. Finally, the composition of the cell types in the ICM samples may be influenced by the fibrotic nature of some of the samples impacting the experimental isolation of nuclei.

STAR★METHODS

RESOURCE AVAILABILITY

Lead contact—Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Patrick Ellinor (ellinor@mgh.harvard.edu).

Materials availability—This study did not generate new unique reagents.

Data and code availability

- Processed single-nuclei transcriptomic data have been deposited at the Broad Institute's Single Cell Portal: https://singlecell.broadinstitute.org/single_cell/study/SCP1849/ and are publicly available as of the date of publication. Accession numbers are listed in the key resources table. Raw sequence data have been deposited at dbGaP (the database of Genotypes and Phenotypes, accession number phs001539.v4.p1) and Sequence Read Archive (SRA) and are available to authorized users. These accession numbers are also listed in the key resources table. Microscopy data reported in this paper will be shared by the lead contact upon request.
- Original code for snRNA-seq analysis, including the global map construction, differential expression analysis, and endothelial subclustering has been deposited at Zenodo: <https://doi.org/10.5281/zenodo.7469682> and GitHub: https://github.com/mark-chaffin/snRNAseq_ischemic_cardiomyopathy. DOIs are listed in the key resources table. Any additional analysis was performed using

publicly available programs and analytical packages as outlined in the STAR Methods and key resources table.

- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human samples—Adult human left ventricular tissue was collected from heart transplant recipients (for the ICM cases) and from deceased organ donors (for the NF cases), as previously described.¹³ All hearts were arrested *in situ* with at least 1 L of ice-cold crystalloid cardioplegia solution, and transported in ice-cold cardioplegia solution to the lab where they were cryopreserved (within 4 h). Specimens were collected in a transmural fashion (epicardium - > endocardium) from the LV free wall below the level of the papillary muscle and adjacent to the septum and above the LV apex. The donor and transplant recipients were collected identically, though it is impossible to be from the exact same spot across samples at different times. We did not purposefully collect the border zone and infarct region in this study. The age and sex of the samples were as follows: NF: M 52, F 64, M 58, M 58, F 47, F 63, F 53, M 42, ICM: F 64, F 55, M 55, F 62, F 47, M 64, F 62. All other relevant patient information can be found in Table 1. Samples collected in the HCM and DCM study were collected in a similar manner.

Written informed consent for research use of donated tissue was obtained from the transplant recipients or next of kin (for the NF cases). Research use of tissues were approved by the relevant institutional review boards at the University of Pennsylvania, Gift-of-Life Donor Program, and Massachusetts General Hospital and the Broad Institute.

METHOD DETAILS

Single-nuclei RNA sequencing—Single nucleus suspensions were generated as previously described.¹³ In brief, approximately 100 mg of left ventricular tissue was cryosectioned (100 μ m) and placed into 2 mL of ice-cold NIM2 buffer (250 mM sucrose, 10 mM Tris pH 8.0, 25 mM KCl, 5 mM MGCL2, 1 μ M DTT, 0.01% IGEPAL). Samples were triturated with a cut tip, transferred to a 7 mL dounce and homogenized on ice. Samples were made up to 4 mL with NIM2 buffer and large debris pelleted by centrifugation at 40 x g at 4° for 2 min. The supernatant was filtered through stacked 40 μ m and 10 μ m filters. Filters were washed with 6 mL of PBS wash buffer (0.01% BSA, 5 mM MgCl2, PBS) and the nuclei pelleted by centrifugation at 550 x g at 4° for 5 min. The nuclei pellet was washed by removal of the supernatant, and addition of 6 mL of PBS wash buffer, followed by centrifugation at 550 x g at 4° for 5 min. The supernatant was removed and nuclei resuspended in 150 μ L of nuclei resuspension buffer (0.01% BSA, 5 mM MgCl2, 0.4 U/ μ L RNase inhibitor), and the nuclei were counted using a hemocytometer. Approximately 6000 nuclei per sample were loaded onto a 10x genomics microfluidic chip and samples were processed according to the 10x genomics protocol (Chromium Single cell 3', v3) with the following changes. First, nuclei were incubated at 4C for 15 min after emulsion generation to promote nuclear lysis. Second, the reverse transcription protocol was modified for one of the two replicates to be 42C for 20 min then 53C for 120 min. Libraries were multiplexed at

an average of 16 libraries per flow cell on an Illumina Novaseq S1 at the Broad Institute's Genomics Platform.

Single-nuclei RNA sequencing data processing—Initial data processing was performed using the 10x Genomics toolkit CellRanger v4.0.0. First, FASTQ files were generated from de-multiplexed raw base call files with *cellranger mkfastq*. Homopolymer repeats greater than 30 base pairs as well as the template switch oligo and its reverse complement (AAGCAGTGGTATCAACGCAGAGTACATGGG, CCCATGTACTCTGCGTTGATACCACTGC TT) were trimmed using *cutadapt*⁷³ with parameters `max_error_rate = 0.1`, `min_overlap = 20` and `max_error_rate = 0.07`, `min_overlap = 10`, respectively. Reads were then aligned to the GRCh38 pre-mRNA human reference (v2020-A) with *cellranger count* with default parameters setting `–expect-cells 6000`.

Samples were assessed for quality using three criteria. First, *cellranger count* metrics were compared across samples to identify extreme outliers (Table S1). One sample, P1795, had an abnormally low percent of valid barcodes (<90%) and was removed. Second, the relationship between total unique molecular identifiers (UMIs) per nucleus and the rank of nuclei by total UMI was visually inspected to ensure a clear distinction between non-empty and empty droplets and acceptable levels of ambient RNA. Third, individual Uniform Manifold Approximation and Projection (UMAP) plots were inspected for each sample to identify any lacking clear structure. Finally, the total count of Y chromosome genes was calculated for each sample and correlated with phenotypic sex. In total, 1 of 16 samples was removed.

For the remaining samples, we applied the tool *remove-background* from CellBender v0.2, <https://github.com/broadinstitute/CellBender>⁷⁴ to call non-empty droplets and subtract ambient background RNA contamination. Parameters were set to: `–expected-cells 6000`, `–z-dim 100`, `–total-droplets-included 25,000`, `–epochs 150`, `–learning_rate 1e-4`, `–fpr 0.01`. For one sample, P1693, the `–total-droplets-included` parameter was set to 30,000 to ensure model convergence.

Finally, we calculated the fraction of reads from each cell barcode that mapped purely to exonic regions using the tool *scR-Invex* (<https://github.com/broadinstitute/scrinvex>). This metric can serve as a surrogate for cytoplasmic contamination as reads derived from the cytoplasmic should map predominantly to exonic regions, whereas reads from the nucleus will span both introns and exons.

Single-nuclei RNA sequencing map aggregation—Single-nuclei RNA sequencing map aggregation was performed analogously to Chaffin et al.¹⁴ All analysis was performed in *scanpy* v1.7.2⁷⁵ unless otherwise specified. First, non-empty droplets called with CellBender *remove-background* were aggregated across samples. The top 2000 highly variable genes were identified using the methodology from Seurat V3.⁸⁸ Counts were log-normalized per nucleus to 10,000 counts and scaled per gene to unit variance and zero mean. The first 50 principal components were estimated using 2000 highly variable genes and aligned across samples to account for biological and technical heterogeneity using *Harmony* as implemented in *harmony-pytorch* v0.1.4.⁷⁶ A neighborhood graph was computed with

sc.pp.neighbors() using cosine distance. Data was projected using the UMAP method with the effective minimum distance between embedded points set to 0.2. High resolution Leiden clustering (resolution = 2.0) was performed to cluster nuclei into groups.

To remove low quality nuclei, four quality control QC metrics were employed: 1. The fraction of UMI mapping to mitochondrial genes based on CellRanger counts (%MT), 2. The fraction of reads mapping exclusively to exonic regions (exon_prop), 3. A measure of nuclei-based entropy as calculated using the *ndd* python library (<https://pypi.org/project/ndd/1.6.3/>)⁷⁷ and 4. A predicted doublet score as estimated using *Scrublet*⁷⁸ with default settings based on CellRanger counts (doublet_score). Clusters from this preliminary map with elevated levels of %MT, exon_prop, or doublet_score or abnormally low levels of entropy were removed which accounted for 47,115/176,518 (26.7%) nuclei. Remaining clusters were further cleaned by performing a per sample, per cluster procedure based on the distribution of each QC metric defined by the first quartile (Q1), the third quartile (Q3), and the interquartile range (IQR). For each sample/cluster combination, outlier nuclei were identified as those with 1. Total UMI (n_umi) greater than $Q3 + 1.5 \times IQR$ or less than $Q1 - 1.5 \times IQR$, 2. Number of unique genes (n_genes) greater than $Q3 + 1.5 \times IQR$ or less than $Q1 - 1.5 \times IQR$, 3. %MT greater than $Q3 + 1.5 \times IQR$, 4. Entropy greater than $Q3 + 1.5 \times IQR$ or less than $Q1 - 1.5 \times IQR$, 5. exon_prop greater than $Q3 + 1.5 \times IQR$, 6. doublet_score greater than $Q3 + 1.5 \times IQR$, or 7. $\log(n_genes) \times \text{entropy}$ greater than $Q3 + 1.5 \times IQR$. When a given sample/cluster combination had less than 30 nuclei, hard thresholds were set at 150 n_umi, 15,000 n_genes, 6000 %MT, 5% entropy, 8 exon_prop, 0.18 doublet_score, 0.30 $\log(n_gene) \times \text{entropy}$. This removed an additional 28,879 nuclei.

The remaining 102,524 nuclei were re-aggregated using the same approach as above after reducing the resolution of Leiden clustering to 0.5. As a final pruning step, a sub-clustering procedure was performed. First, similar clusters were joined, and marker genes were calculated for each major cell type based on an area under the receiver operating characteristic curve (AUC) comparing the expression in each major cell type to all other cell types. Each major cell type was then sub-clustered by recalculating the neighborhood graph on the harmony adjusted PCs. Then, Leiden clustering was performed at increasing resolutions of 0.05–1.0, in increments of 0.05. At each step, marker genes for each sub-cluster compared to all other sub-clusters were calculated using the AUC metric. When a sub-cluster emerged and had no genes with AUC greater than 0.6, sub-clustering was halted. All remaining sub-clusters were then scored based on the top 50 marker genes of the global cell types, as determined by AUC, as well as all mitochondrial genes using the *sc.tl.score_genes()* function in *scanpy*^{89,90}. Any sub-clusters of a cell type enriched in scores from other major cell types or mitochondrial genes were deemed low quality or misclassified and subsequently removed. This resulted in the removal of an additional 2,840 nuclei and a final snRNA-seq map containing 99,684 nuclei.

Single-nuclei RNA marker gene calculation—To assign cell type labels to each cluster, marker genes for each Leiden cluster were calculated compared to all other clusters using two approaches. First, an AUC was calculated for each gene comparing the log-normalized expression at the nuclei level in each cluster against all other clusters. Second, a formal differential expression model controlling for the correlation amongst nuclei from

the same individual was performed by summing gene counts across all nuclei in a cluster within an individual patient and treating the data as a bulk RNA sequencing experiment. Nuclei were only aggregated in an individual if there were more than 20 nuclei of the given cluster. Next, a differential expression model was performed using the limma-voom framework^{79,80} with a model of $\sim 0 + \text{cluster} + \text{individual}$ with DESeq2 normalization⁸¹ to control for the fact that each individual is represented across clusters. An individual contrast was extracted comparing each cluster to all other clusters. Marker genes for each cluster were then identified as protein coding genes with an AUC >0.60, a Benjamini-Hochberg corrected p value from the limma-voom model <0.01, a log fold-change estimate from the limma-voom model greater than 2 and expressed in at least 25% of nuclei in the given cluster.

Composition analysis—To identify differences in cell type composition between ischemic and non-failing samples, we applied the Bayesian method from *scCODA* v0.1.2.post1.⁸² For this analysis, a reference cell type must be selected that is deemed to be relatively invariant across samples. In the global compositional test, pericytes were selected as they had the smallest variation across samples and were reasonably prevalent in the dataset. When testing for changes in endothelial cell composition, the reference group was selected as *EC-KCNIP4*.

Differential expression testing—We performed differential expression between ICM and NF samples treating each patient as the unit of biological replication. To control for the fact that nuclei derived from a given patient will be correlated, we took the approach of Lun and Marioni.⁹¹ For each patient, we calculated the sum across nuclei for each gene within a given cell type of interest and treated the data as if it were derived from a bulk RNA-sequencing experiment. We required a patient to have at least 25 nuclei of the given cell type to be included. We removed genes in <1% of nuclei from both ICM and NF samples, separately, or identified as lowly expressed by `filterByExpr()` in edgeR based on the summed counts.⁸³ We further removed mitochondrial and ribosomal genes from differential expression testing. Expression was normalized using DESeq2⁸¹ and differential expression testing was performed using the limma-voom testing framework.^{79,80} All tests were adjusted for age and sex. We applied false discovery rate (FDR) correction using the Benjamini-Hochberg procedure. Our primary analysis was performed on CellBender adjusted counts, but as a sensitivity analysis, we repeated our test using CellBender unadjusted counts. Only genes with FDR-adjusted p-values <0.05 in both tests were considered significant.

Additionally, we tested for an association between the percent of fibrosis, as estimated using Trichrome staining, and gene expression among the ICM patients. Only the 6 ICM patients with successful Trichrome staining were included. Differential expression was tested as described above, adjusting for age and sex. As this was considered an exploratory analysis due to the limited power of the statistical test, we focus on genes that have an unadjusted p-value <0.005 from both CellBender and CellRanger.

To minimize the role of background contamination in dictating our differential expression results, we calculated a per gene, per cell type background contamination heuristic. For example, if a gene is highly expressed in a prevalent cell type (e.g. cardiomyocytes)

and differentially expressed between ICM and NF patients, we may observe differential expression from the background in other cell types that express low levels of the gene. This heuristic consisted of two components: 1. A measure of overall expression for each gene in the dataset, and 2. A measure of specificity of each gene to the given cell type being tested. For the first component, we divided the total UMI of each gene in the entire dataset by the total UMI in the dataset. We then estimated a value (“bkg_prob”) for each gene along the cumulative distribution of these values. For the second component, we calculated two positive predictive values for each gene, standardized for equal prevalence of cell type, by dichotomizing expression as >0 (PPV0) or >1 (PPV1) and predicting classification into the cell type where differential expression test was performed. We then calculated “nontarget_prob” as $1 - \text{mean}(\text{PPV0}, \text{PPV1})$. Our final heuristic for a given gene in a given cell type was calculated as $\text{bkg_prob} * \text{nontarget_prob}$. Genes with values >0.4 were considered as having a high probability of being derived from background and flagged or removed from downstream analyses.

Pathway/ontology analysis—Pathway analyses were performed using two databases, gene ontology (GO) biological processes (BP)^{25,92} and Reactome pathways.²⁶ Marker gene pathway enrichment was performed on GO BP with topGO⁸⁴ using both a Kolmogorov-Smirnov (KS) test and Fisher’s exact test with the weight01 algorithm to account for the GO graph. For the KS test, genes for each cluster were ranked on the t-statistic from the differential expression test. For Fisher’s exact test, selected marker genes for each cluster were compared against all genes tested. Only GO terms with greater than 15 and less than 500 genes were considered. Ontologies with a p-value <0.001 from the Fisher’s exact test and odds ratio >2 were considered significant. A multiple testing correction was not applied because GO term enrichment is calculated conditionally on neighboring terms and therefore not all tests are independent.

Pathway enrichment for genes differentially expressed in the ICM versus NF comparison was tested using both GO BP databases with topGO⁸⁴ and Reactome pathways using ReactomePA.⁸⁵ Genes were separated into significantly upregulated in ICM and significantly downregulated in ICM prior to enrichment testing. All genes with a background contamination heuristic >0.4 in a given cell type were removed. For GO BP terms, we used a Fisher’s exact test with the weight01 algorithm to account for the GO graph in topGO and only considered ontologies with greater than 15 and less than 500 genes. Ontologies with a p-value <0.001 and an odds ratio for enrichment >2 were considered significant. For Reactome pathways, we employed a hypergeometric test using ReactomePA with a minimum gene set size of 10 and a maximum gene set size of 500. A Benjamini-Hochberg correction was applied to ReactomePA results across all cell type/direction combinations and pathways with FDR <0.05 were considered significant.

Endothelial subclustering analysis—All 20,462 nuclei from global clusters Endothelial I, Endothelial II, Endothelial III, and Lymphatic endothelial were sub-clustered by the following procedure. First, we recalculated Harmony-adjusted principal components as described in the global map construction but restricted to just these 20,462 nuclei. Next, we calculated a neighborhood graph using cosine distance and 10 nearest neighbors.

Clusters were determined with the Leiden algorithm at various resolutions, starting at 0.10 and increasing by increments of 0.10. When a cluster emerged with no genes that had an AUC >0.60 separating it from all other clusters, sub-clustering was halted. Sub-clusters were inspected for enrichment in non-endothelial cell markers and mitochondrial gene expression. Any misclassified or low-quality sub-clusters were removed. This process was repeated 3 times and removed a total of 1,826 (9%) of nuclei. To determine the final cluster identities, we performed hierarchical clustering with Euclidean distance and the Ward method based on the mean expression of the 2000 highly variable genes within each sub-cluster. Sub-clusters with distance less than $(0.25 * \text{max inter-cluster distance})$ were merged.

Marker genes were calculated as in the global map. As sub-clusters tend to be more related than global cell populations, the marker gene criteria were loosened to protein coding genes with an AUC >0.50, a Benjamini-Hochberg corrected p value from the limma-voom model <0.01, a log fold-change estimate from the limma-voom model greater than 0, and expressed in at least 15% of nuclei in the given cluster. Gene ontology enrichment for marker genes was calculated as described in the Pathway/Ontology analysis section. Additionally, to compare sub-clusters of endothelial cells to previously published data, we scored all endothelial cells based on profiles from up to 50 of the top marker genes (as determined by log fold-change) of endothelial sub-clusters from Feng et al.,²⁸ Kalucka et al.,²⁹ and Schupp et al.³⁰ using the *sc.tl.score_genes()* function. Mouse genes were mapped to human orthologs biomaRt v2.36.1.⁸⁶

To identify up- and down-regulated genes in ICM patients compared to NF patients with endothelial sub-clusters, we performed differential expression testing as described above. To calculate the background contamination heuristic, we separated nuclei into endothelial cells (Endothelial I, Endothelial II, Endothelial III, and Lymphatic endothelial) and non-endothelial cells and performed the same calculation as described above.

Cell-cell communication analysis—Communication between endothelial cell subtypes and other cell types was quantified using CellChat.⁴² Expression was log-normalized, and nuclei were separated into ICM and NF. Cell types or subtypes with less than 10 nuclei in either disease group were excluded. We used CellChat to then infer the cell-cell communication network within each disease state based on upregulated ligands/receptors in cell types. We then compared the signaling networks in ICM to the signaling networks in NF, focused on endothelial subtypes. Specifically, we compared the number of interactions involving each endothelial subtype broadly, as well as changes in communication between specific pairs of cell types. We then looked for differences in signaling pathways by comparing the information flow from endothelial cells which was defined as the sum of communication probabilities with any endothelial source to any cell type.

Proteomics comparison—We compared changes in at the RNA level in each cell type between ICM and NF based on our snRNA-seq data to changes in protein expression in left ventricles between ICM and NF reported in Chen et al.⁴³ All log fold-change and p value estimates for genes in the snRNA-seq in each cell type were compared directly to the published estimates from the bulk proteomics data.

Joint ICM, DCM, and HCM snRNA-seq map—We constructed a joint map of snRNA-seq data coming from ICM and NF patients in this study with additional snRNA-seq derived from dilated cardiomyopathy (DCM), hypertrophic cardiomyopathy (HCM), and additional NF patients from Chaffin et al.¹⁴ In brief, CellBender-adjusted counts from the 99,684 ICM and NF nuclei were aggregated with an additional 592,689 nuclei from DCM, HCM, and NF patients. The map was constructed analogously as described in “Single-nuclei RNA sequencing map aggregation”, aligning across the patient identifier of nuclei using *Harmony*. Leiden clustering was performed at resolution 0.6.

Principal component analysis was then performed on the patient-level by summing gene counts across all nuclei per patient and treating the data as a bulk experiment. Mitochondrial genes, ribosomal genes, and genes expressed in <1% of nuclei were removed. Remaining genes were normalized with DESeq2 and genes with less than 10 total counts were removed. A variance stabilizing transformation was applied with *vst()*.⁸¹ Principal components were calculated with *prcomp()* in *R*.

Differential expression analysis was performed on a cell-type basis for 6 pairwise comparisons: ICM vs NF, DCM vs NF, HCM vs NF, ICM vs DCM, ICM vs HCM, and DCM vs HCM. Testing was performed as described in “Differential expression testing between ICM and NF” by first summing counts across all nuclei per patient and the fitting the model $\text{Expr} \sim 1 + \text{disease} + \text{sex} + \text{age} + \text{study}$ using the limma-voom framework^{79,80} where study indicates whether the sample is derived from the ICM vs NF study or the DCM/HCM vs NF study. This last covariate allows us to control for technical differences between studies by assuming that NF patients from each study should be comparable. For each cell type, a contrast was specified for each of the 6 pairwise comparisons described above. Nuclei were only summed in a patient if there were at least 25 nuclei of the given cell type and cell types were only tested if there were at least 3 samples in the two groups compared. Multiple testing correction was performed using the Benjamini-Hochberg procedure within each pairwise comparison at an FDR <0.05.

To identify shared pathways between ICM, DCM, and HCM, we performed a Reactome pathway enrichment analysis focusing on genes that were differentially expressed in all cardiomyopathies, per cell type. Genes significantly dysregulated in all three cardiomyopathies (based on both CellBender and CellRanger counts) were split into those consistently upregulated across cardiomyopathy subtypes and those consistently downregulated across cardiomyopathy subtypes. Pathway analysis was performed on these sets using ReactomePA as described above. A Benjamini-Hochberg correction was applied to ReactomePA results across all cell type/direction combinations and pathways with FDR <0.05 were considered significant.

To identify shared druggable targets across all cardiomyopathies, we performed a series of filtering steps on genes tested for differential expression. For each cell type, we identified genes that were significantly differentially expressed (FDR <0.05 based on Benjamini-Hochberg procedure) in ICM versus NF, DCM versus NF, and HCM versus NF. We split these into up- and down-regulated genes based on the log fold-change estimate. We further removed genes that were significantly differentially expressed between any pair

of cardiomyopathies, reduced this to genes with a strong effect (absolute value of log fold-change >1), and reasonably expressed in either cardiomyopathy or NF (>5% of nuclei). We then annotated these genes by whether they were druggable based on Finan et al.⁵² and looked for cell-type specificity based on AUC.

Trichrome staining—Trichrome staining was carried out on 10 µm sections following product protocol (Abcam: ab150686) and whole sections were imaged and tiled at 10x (Zeiss Axio Observer Z1). Area of fibrosis was quantified using thresholding on (Fuji is just) ImageJ v2.1.0.

RNAscope validation—To image lymphatic endothelial cells and angiogenic endothelial cells, colorimetric RNAscope was carried out on 10 µm fresh frozen tissue sections using the RNAscope 2.5 HD Duplex kit (ACDBio: 322430) following manufacturers protocols. As nearly all angiogenic endothelial cell marker genes were expressed appreciably in other cell types, we used a combination of two probes to identify this population. First, we identified 6 genes that were highly selective to angiogenic endothelial cells compared to other endothelial cell populations (AUC >0.6, log fold-change >1, percent of nuclei expression >20%) and that were not expressed in more than 25% of any other endothelial cell population. As all 6 of these genes showed some expression in other non-endothelial cell types, we calculated the selectivity of these 6 genes combined with expression of the global endothelial marker *PECAM1*. We determined that 23% of angiogenic endothelial cells in our snRNA-seq data expressed *MYO1B* and *PECAM1* whereas double-positives in other endothelial sub-populations and non-endothelial global populations occurred in no more than 6% of nuclei. Therefore, the probes used for this study were hs-CCL21 (474371), hs-PECAM1 (487381-C2), and hs-MYO1B (444301). To quantify the number of lymphatic endothelial cells, (Fuji is just) ImageJ v2.1.0⁸⁷ was used to manually count the number of CCL21 positive cells, and then thresholded to automatically count the number of nuclei per section.

QUANTIFICATION AND STATISTICAL ANALYSIS

Graphpad Prism 9.3.1 was used to carry out an unpaired t test on the number of lymphatic endothelial cells identified in the RNA-scope experiments (Figure S5). Statistical significance defined as a p value less than 0.5 All additional statistical analyses were performed in Python or R, with details of these analyses described in the main text, STAR Methods, figure legends, and supplemental tables.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

The Precision Cardiology Laboratory is a joint effort between the Broad Institute and Bayer AG. P.T.E. is supported by grants from the National Institutes of Health (1R01HL092577, 1R01HL157635, 5R01HL139731), from the American Heart Association Strategically Focused Research Networks (18SFRN34110082), and from the European Union (MAESTRIA 965286). K.B.M. was supported by grants from the National Institutes of Health (1R01HL105993 and 1R01HL149891). Graphical abstract created with art from BioRender.com.

REFERENCES

1. Khan MA, Hashim MJ, Mustafa H, Baniyas MY, Al Suwaidi SKBM, AlKatheeri R, Alblooshi FMK, Almatrooshi MEAH, Alzaabi MEH, Al Darmaki RS, and Lootah SNAH (2020). Global epidemiology of ischemic heart disease: results from the global burden of disease study. *Cureus* 12, e9349. 10.7759/cureus.9349. [PubMed: 32742886]
2. Sutton MG, and Sharpe N (2000). Left ventricular remodeling after myocardial infarction: pathophysiology and therapy. *Circulation* 101, 2981–2988. 10.1161/01.cir.101.25.2981. [PubMed: 10869273]
3. Levy D, Kenchaiah S, Larson MG, Benjamin EJ, Kupka MJ, Ho KKL, Murabito JM, and Vasan RS (2002). Long-term trends in the incidence of and survival with heart failure. *N. Engl. J. Med* 347, 1397–1402. 10.1056/NEJMoa020265. [PubMed: 12409541]
4. Fu X, Khalil H, Kanisicak O, Boyer JG, Vagnozzi RJ, Maliken BD, Sargent MA, Prasad V, Valiente-Alandi I, Blaxall BC, and Molkentin JD (2018). Specialized fibroblast differentiated states underlie scar formation in the infarcted mouse heart. *J. Clin. Invest* 128, 2127–2143. 10.1172/JCI98215. [PubMed: 29664017]
5. Molenaar B, Timmer LT, Droog M, Perini I, Versteeg D, Kooijman L, Monshouwer-Kloots J, de Ruiter H, Gladka MM, and van Rooij E (2021). Single-cell transcriptomics following ischemic injury identifies a role for B2M in cardiac repair. *Commun. Biol* 4, 146. 10.1038/s42003-020-01636-3. [PubMed: 33514846]
6. Singhal AK, Symons JD, Boudina S, Jaishy B, and Shiu Y-T (2010). Role of endothelial cells in myocardial ischemia-reperfusion injury. *Vasc. Dis. Prev* 7, 1–14. 10.2174/1874120701007010001. [PubMed: 25558187]
7. Segers VFM, Brutsaert DL, and De Keulenaer GW (2018). Cardiac remodeling: endothelial cells have more to say than just NO. *Front. Physiol* 9, 382. 10.3389/fphys.2018.00382. [PubMed: 29695980]
8. Ren G, Michael LH, Entman ML, and Frangogiannis NG (2002). Morphological characteristics of the microvasculature in healing myocardial infarcts. *J. Histochem. Cytochem* 50, 71–79. 10.1177/002215540205000108. [PubMed: 11748296]
9. Shimizu Y, Polavarapu R, Eskla K-L, Pantner Y, Nicholson CK, Ishii M, Brunnhoelzl D, Mauria R, Husain A, Naqvi N, et al. (2018). Impact of lymphangiogenesis on cardiac remodeling after ischemia and reperfusion injury. *J. Am. Heart Assoc* 7, e009565. 10.1161/JAHA.118.009565. [PubMed: 30371303]
10. Hou L, Kim JJ, Woo YJ, and Huang NF (2016). Stem cell-based therapies to promote angiogenesis in ischemic cardiovascular disease. *Am. J. Physiol. Heart Circ. Physiol* 310, H455–H465. 10.1152/ajpheart.00726.2015. [PubMed: 26683902]
11. Song L, Chen X, Swanson TA, LaViolette B, Pang J, Cunio T, Nagle MW, Asano S, Hales K, Shipstone A, et al. (2020). Lymphangiogenic therapy prevents cardiac dysfunction by ameliorating inflammation and hypertension. *Elife* 9, e58376. 10.7554/eLife.58376. [PubMed: 33200983]
12. Litviuková M, Talavera-López C, Maatz H, Reichart D, Worth CL, Lindberg EL, Kanda M, Polanski K, Heinig M, Lee M, et al. (2020). Cells of the adult human heart. *Nature* 588, 466–472. 10.1038/s41586-020-2797-4. [PubMed: 32971526]
13. Tucker NR, Chaffin M, Fleming SJ, Hall AW, Parsons VA, Bedi KC, Akkad A-D, Herndon CN, Arduini A, Papangelis I, et al. (2020). Transcriptional and cellular diversity of the human heart. *Circulation* 142, 466–482. 10.1161/CIRCULATIONAHA.119.045401. [PubMed: 32403949]
14. Chaffin M, Papangelis I, Simonson B, Akkad A-D, Hill MC, Arduini A, Fleming SJ, Melanson M, Hayat S, Kost-Alimova M, et al. (2022). Single-nuclei profiling of human dilated and hypertrophic cardiomyopathy. *Nature* 608, 174–180. 10.1038/s41586-022-04817-8. [PubMed: 35732739]
15. Koenig AL, Shchukina I, Amrute J, Andhey PS, Zaitsev K, Lai L, Bajpai G, Bredemeyer A, Smith G, Jones C, et al. (2022). Single-cell transcriptomics reveals cell-type-specific diversification in human heart failure. *Nat. Cardiovasc. Res* 1, 263–280. 10.1038/s44161-022-00028-6. [PubMed: 35959412]
16. Hanna A, and Frangogiannis NG (2019). The role of the TGF- β super-family in myocardial infarction. *Front. Cardiovasc. Med* 6, 140. 10.3389/fcvm.2019.00140. [PubMed: 31620450]

17. Scruggs AM, Grabauskas G, and Huang SK (2020). The role of KCNMB1 and BK channels in myofibroblast differentiation and pulmonary fibrosis. *Am. J. Respir. Cell Mol. Biol* 62, 191–203. 10.1165/rcmb.2019-0163OC. [PubMed: 31486669]
18. Cao TH, Quinn PA, Sandhu JK, Voors AA, Lang CC, Parry HM, Mohan M, Jones DJL, and Ng LL (2015). Identification of novel biomarkers in plasma for prediction of treatment response in patients with heart failure. *Lancet* 385, S26. 10.1016/S0140-6736(15)60341-5.
19. Tian X, Sun C, Wang X, Ma K, Chang Y, Guo Z, and Si J (2020). ANO1 regulates cardiac fibrosis via ATI-mediated MAPK pathway. *Cell Calcium* 92, 102306. 10.1016/j.ceca.2020.102306. [PubMed: 33075549]
20. Zhang L, Malik S, Kelley GG, Kapiloff MS, and Smrcka AV (2011). Phospholipase C scaffolds to muscle-specific A kinase anchoring protein (makap β) and integrates multiple hypertrophic stimuli in cardiac myocytes. *J. Biol. Chem* 286, 23012–23021. 10.1074/jbc.M111.231993. [PubMed: 21550986]
21. Sörensen-Zender I, Bhayana, Susnik N, Rolli V, Batkai S, Baisantray A, Bahram S, Sen P, Teng B, Lindner R, et al. (2015). Zinc- α 2-Glycoprotein exerts antifibrotic effects in kidney and heart. *J. Am. Soc. Nephrol* 26, 2659–2668. 10.1681/ASN.2014050485. [PubMed: 25788525]
22. Gil-Cayuela C, López A, Martínez-Dolz L, González-Juanatey JR, Lago F, Roselló-Lletí E, Rivera M, and Portols M (2019). The altered expression of autophagy-related genes participates in heart failure: NRBP2 and CALCOCO2 are associated with left ventricular dysfunction parameters in human dilated cardiomyopathy. *PLoS One* 14, e0215818. 10.1371/journal.pone.0215818. [PubMed: 31009519]
23. Piehler A, Kaminski WE, Wenzel JJ, Langmann T, and Schmitz G (2002). Molecular structure of a novel cholesterol-responsive A subclass ABC transporter, ABCA9. *Biochem. Biophys. Res. Commun* 295, 408–416. 10.1016/S0006-291X(02)00659-9. [PubMed: 12150964]
24. González-Domínguez É, Samaniego R, Flores-Sevilla JL, Campos-Campos SF, Gómez-Campos G, Salas A, Campos-Peña V, Corbí ÁL, Sánchez-Mateos P, and Sánchez-Torres C (2015). CD163L1 and CLEC5A discriminate subsets of human resident and inflammatory macrophages in vivo. *J. Leukoc. Biol* 98, 453–466. 10.1189/jlb.3H1114-531R. [PubMed: 25877931]
25. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet* 25, 25–29. 10.1038/75556. [PubMed: 10802651]
26. Jassal B, Matthews L, Viteri G, Gong C, Lorente P, Fabregat A, Sidiropoulos K, Cook J, Gillespie M, Haw R, et al. (2020). The reactome pathway knowledgebase. *Nucleic Acids Res* 48, D498–D503. 10.1093/nar/gkz1031. [PubMed: 31691815]
27. Ishikawa Y, Akishima-Fukasawa Y, Ito K, Akasaka Y, Tanaka M, Shimokawa R, Kimura-Matsumoto M, Morita H, Sato S, Kamata I, and Ishii T (2007). Lymphangiogenesis in myocardial remodelling after infarction. *Histopathology* 51, 345–353. 10.1111/j.1365-2559.2007.02785.x. [PubMed: 17727476]
28. Feng W, Chen L, Nguyen PK, Wu SM, and Li G (2019). Single cell analysis of endothelial cells identified organ-specific molecular signatures and heart-specific cell populations and molecular features. *Front. Cardiovasc. Med* 6, 165. 10.3389/fcvm.2019.00165. [PubMed: 31850371]
29. Kalucka J, de Rooij LPMH, Goveia J, Rohlenova K, Dumas SJ, Meta E, Conchinha NV, Taverna F, Teuwen L-A, Veys K, et al. (2020). Single-cell transcriptome atlas of murine endothelial cells. *Cell* 180, 764–779.e20. 10.1016/j.cell.2020.01.015. [PubMed: 32059779]
30. Schupp JC, Adams TS, Cosme C, Raredon MSB, Yuan Y, Omote N, Poli S, Chioccioli M, Rose K-A, Manning EP, et al. (2021). Integrated single-cell atlas of endothelial cells of the human lung. *Circulation* 144, 286–302. 10.1161/CIRCULATIONAHA.120.052318. [PubMed: 34030460]
31. Heikkilä M, Pasanen A, Kivirikko KI, and Myllyharju J. (2011). Roles of the human hypoxia-inducible factor (HIF)-3 α variants in the hypoxia response. *Cell. Mol. Life Sci* 68, 3885–3901. 10.1007/s00018-011-0679-5. [PubMed: 21479871]
32. Yamashita T, Ohneda O, Nagano M, Iemitsu M, Makino Y, Tanaka H, Miyauchi T, Goto K, Ohneda K, Fujii-Kuriyama Y, et al. (2008). Abnormal heart development and lung remodeling in mice lacking the hypoxia-inducible factor-related basic helix-loop-helix PAS protein NEPAS. *Mol. Cell Biol* 28, 1285–1297. 10.1128/MCB.01332-07. [PubMed: 18070924]

33. King SJ, Nowak K, Suryavanshi N, Holt I, Shanahan CM, and Ridley AJ (2014). Nesprin-1 and nesprin-2 regulate endothelial cell shape and migration. *Cytoskeleton* 71, 423–434. 10.1002/cm.21182. [PubMed: 24931616]
34. Escudero CA, Herlitz K, Troncoso F, Guevara K, Acurio J, Aguayo C, Godoy AS, and González M (2017). Pro-angiogenic role of insulin: from physiology to pathology. *Front. Physiol* 8, 204. 10.3389/fphys.2017.00204. [PubMed: 28424632]
35. Li GH, Shi Y, Chen Y, Sun M, Sader S, Maekawa Y, Arab S, Dawood F, Chen M, De Couto G, et al. (2009). Gelsolin regulates cardiac remodeling after myocardial infarction through DNase I-mediated apoptosis. *Circ. Res* 104, 896–904. 10.1161/CIRCRESAHA.108.172882. [PubMed: 19246681]
36. Liston P, Fong WG, Kelly NL, Toji S, Miyazaki T, Conte D, Tamai K, Craig CG, McBurney MW, and Korneluk RG (2001). Identification of XAF1 as an antagonist of XIAP anti-Caspase activity. *Nat. Cell Biol* 3, 128–133. 10.1038/35055027. [PubMed: 11175744]
37. Galvin KM, Donovan MJ, Lynch CA, Meyer RI, Paul RJ, Lorenz JN, Fairchild-Huntress V, Dixon KL, Dunmore JH, Gimbrone MA, et al. (2000). A role for smad6 in development and homeostasis of the cardiovascular system. *Nat. Genet* 24, 171–174. 10.1038/72835. [PubMed: 10655064]
38. Tan B, Liu M, Yang Y, Liu L, and Meng F (2019). Low expression of PIK3C2A gene: a potential biomarker to predict the risk of acute myocardial infarction. *Medicine (Baltim.)* 98, e15061. 10.1097/MD.00000000000015061.
39. Teixeira SC, Lopes DS, Gimenes SNC, Teixeira TL, da Silva MS, Brígido RTES, da Luz FAC, da Silva AA, Silva MA, Florentino PV, et al. (2017). Mechanistic insights into the anti-angiogenic activity of trypanosoma cruzi protein 21 and its potential impact on the onset of chagasic cardiomyopathy. *Sci. Rep* 7, 44978. 10.1038/srep44978. [PubMed: 28322302]
40. Leatherdale A, Parker D, Tasneem S, Wang Y, Bihan D, Bonna A, Hamaia SW, Gross PL, Ni H, Doble BW, et al. (2021). Multimerin 1 supports platelet function in vivo and binds to specific GPAGPOGPX motifs in fibrillar collagens that enhance platelet adhesion. *J. Thromb. Haemost* 19, 547–561. 10.1111/jth.15171. [PubMed: 33179420]
41. Zheng C, Quan R, Xia E-J, Bhandari A, and Zhang X (2019). Original tumour suppressor gene polycystic kidney and hepatic disease 1-like 1 is associated with thyroid cancer cell progression. *Oncol. Lett* 18, 3227–3235. 10.3892/ol.2019.10632. [PubMed: 31452800]
42. Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan C-H, Myung P, Plikus MV, and Nie Q (2021). Inference and analysis of cell-cell communication using CellChat. *Nat. Commun* 12, 1088. 10.1038/s41467-021-21246-9. [PubMed: 33597522]
43. Chen CY, Caporizzo MA, Bedi K, Vite A, Bogush AI, Robison P, Heffler JG, Salomon AK, Kelly NA, Babu A, et al. (2018). Suppression of detyrosinated microtubules improves cardiomyocyte function in human heart failure. *Nat. Med* 24, 1225–1233. 10.1038/s41591-018-0046-2. [PubMed: 29892068]
44. Zhang X, Guan J, Guo M, Dai H, Cai S, Zhou C, Wang Y, and Qin Q (2019). Rho GTPase-activating protein 1 promotes apoptosis of myocardial cells in an ischemic cardiomyopathy model. *Kardiol. Pol* 77, 1163–1169. 10.33963/KP.15040. [PubMed: 31664016]
45. Yamada Y, Fuku N, Tanaka M, Aoyagi Y, Sawabe M, Metoki N, Yoshida H, Satoh K, Kato K, Watanabe S, et al. (2009). Identification of CELSR1 as a susceptibility gene for ischemic stroke in Japanese individuals by a genome-wide association study. *Atherosclerosis* 207, 144–149. 10.1016/j.atherosclerosis.2009.03.038. [PubMed: 19403135]
46. Eyries M, Girerd B, Savale L, Soubrier F, Humbert M, and Montani D (2021). A CELSR1 variant in a patient with pulmonary arterial hypertension. *Clin. Genet* 100, 771–772. 10.1111/cge.14046. [PubMed: 34435352]
47. Wu X, Li Y, Zhang S, and Zhou X (2021). Ferroptosis as a novel therapeutic target for cardiovascular disease. *Theranostics* 11, 3052–3059. 10.7150/thno.54113. [PubMed: 33537073]
48. Liepinsh E, Kuka J, Vilks K, Svalbe B, Stelfa G, Vilskersts R, Sevostjanovs E, Goldins NR, Groma V, Grinberga S, et al. (2021). Low cardiac content of long-chain acylcarnitines in TMLHE knockout mice prevents ischaemia-reperfusion-induced mitochondrial and cardiac damage. *Free Radic. Biol. Med* 177, 370–380. 10.1016/j.freerad-biomed.2021.10.035. [PubMed: 34728372]

49. Wu C-C, Jeratsch S, Graumann J, and Stainier DYR(2020). Modulation of mammalian cardiomyocyte cytokinesis by the extracellular matrix. *Circ. Res* 127, 896–907. 10.1161/CIRCRESAHA.119.316303. [PubMed: 32564729]
50. Ding X-Y, Hu H-Y, Huang K-N, Wei R-Q, Min J, Qi C, Tang H, and Qin X (2020). Ubiquitination of NOTCH2 by DTX3 suppresses the proliferation and migration of human esophageal carcinoma. *Cancer Sci* 111, 489–501. 10.1111/cas.14288. [PubMed: 31854042]
51. Valla M, Opdahl S, Ytterhus B, and Bofin AM (2021). DTX3 copy number increase in breast cancer: a study of associations to molecular subtype, proliferation and prognosis. *Breast Cancer Res. Treat* 187, 57–67. 10.1007/s10549-021-06138-2. [PubMed: 33616774]
52. Finan C, Gaulton A, Kruger FA, Lumbers RT, Shah T, Engmann J, Galver L, Kelley R, Karlsson A, Santos R, et al. (2017). The druggable genome and support for target identification and validation in drug development. *Sci. Transl. Med* 9, eaag1166. 10.1126/scitranslmed.aag1166. [PubMed: 28356508]
53. He Z, Zhang X, Chen C, Wen Z, Hoopes SL, Zeldin DC, and Wang DW (2015). Cardiomyocyte-specific expression of CYP2J2 prevents development of cardiac remodelling induced by angiotensin II. *Cardiovasc. Res* 105, 304–317. 10.1093/cvr/cvv018. [PubMed: 25618409]
54. Carniel E, Taylor MRG, Sinagra G, Di Lenarda A, Ku L, Fain PR, Boucek MM, Cavanaugh J, Miodic S, Slavov D, et al. (2005). Alpha-myosin heavy chain: a sarcomeric gene associated with dilated and hypertrophic phenotypes of cardiomyopathy. *Circulation* 112, 54–59. 10.1161/CIRCULATIONAHA.104.507699. [PubMed: 15998695]
55. Xin C, Lei J, Wang Q, Yin Y, Yang X, Moran Guerrero JA, Sabbisetti V, Sun X, Vaidya VS, and Bonventre JV (2021). Therapeutic silencing of SMOC2 prevents kidney function loss in mouse model of chronic kidney disease. *iScience* 24, 103193. 10.1016/j.isci.2021.103193. [PubMed: 34703992]
56. Wu J, Subbaiah KCV, Xie LH, Jiang F, Khor E-S, Mickelsen D, Myers JR, Tang WHW, and Yao P (2020). Glutamyl-Prolyl-tRNA synthetase regulates proline-rich pro-fibrotic protein synthesis during cardiac fibrosis. *Circ. Res* 127, 827–846. 10.1161/CIRCRESAHA.119.315999. [PubMed: 32611237]
57. Schellings MWM, Vanhoutte D, Swinnen M, Cleutjens JP, Debets J, van Leeuwen REW, d'Hooge J, Van de Werf F, Carmeliet P, Pinto YM, et al. (2009). Absence of SPARC results in increased cardiac rupture and dysfunction after acute myocardial infarction. *J. Exp. Med* 206, 113–123. 10.1084/jem.20081244. [PubMed: 19103879]
58. Wei T, Song J, Liang K, Li L, Mo X, Huang Z, Chen G, Mao N, and Yang J (2021). Identification of a novel therapeutic candidate, NRK, in primary cancer-associated fibroblasts of lung adenocarcinoma microenvironment. *J. Cancer Res. Clin. Oncol* 147, 1049–1064. 10.1007/s00432-020-03489-z. [PubMed: 33387038]
59. Ahmed A, Ahmed S, Arvidsson M, Bouzina H, Lundgren J, and Rådegran G (2020). Prolargin and matrix metalloproteinase-2 in heart failure after heart transplantation and their association with haemodynamics. *ESC Heart Fail* 7, 223–234. 10.1002/ehf2.12560. [PubMed: 31858729]
60. Johnson T, Zhao L, Manuel G, Taylor H, and Liu D (2019). Approaches to therapeutic angiogenesis for ischemic heart disease. *J. Mol. Med* 97, 141–151. 10.1007/s00109-018-1729-3. [PubMed: 30554258]
61. Henri O, Poueche C, Houssari M, Galas L, Nicol L, Edwards-Lévy F, Henry J-P, Dumesnil A, Boukhalfa I, Banquet S, et al. (2016). Selective stimulation of cardiac lymphangiogenesis reduces myocardial edema and fibrosis leading to improved cardiac function following myocardial infarction. *Circulation* 133, 1484–1497. 10.1161/CIRCULATIONAHA.115.020143. [PubMed: 26933083]
62. Houssari M, Dumesnil A, Tardif V, Kivelä R, Pizzinat N, Boukhalfa I, Godefroy D, Schapman D, Hemanthakumar KA, Bizou, et al. (2020). Lymphatic and immune cell cross-talk regulates cardiac recovery after experimental myocardial infarction. *Arterioscler. Thromb. Vasc. Biol* 40, 1722–1737. 10.1161/ATVBAHA.120.314370. [PubMed: 32404007]
63. Keller TCS, Lim L, Shewale SV, McDaid K, Martí-Pàmies Í, Tang AT, Wittig C, Guerrero AA, Sterling S, Leu NA, et al. (2021). Genetic blockade of lymphangiogenesis does not impair cardiac function after myocardial infarction. *J. Clin. Invest* 131, e147070. 10.1172/JCI147070. [PubMed: 34403369]

64. Boland E, Quondamatteo F, and Van Agtmael T (2021). The role of basement membranes in cardiac biology and disease. *Biosci. Rep* 41, BSR20204185. 10.1042/BSR20204185. [PubMed: 34382650]
65. Chen J, Ma Y, Li H, Lin Z, Yang Z, Zhang Q, Wang F, Lin Y, Ye Z, and Lin Y (2022). Rare and potential pathogenic mutations of LMNA and LAMA4 associated with familial arrhythmogenic right ventricular cardiomyopathy/dysplasia with right ventricular heart failure, cerebral thromboembolism and hereditary electrocardiogram abnormality. *Orphanet J. Rare Dis* 17, 183. 10.1186/s13023-022-02348-z. [PubMed: 35526016]
66. Wagner JUG, Chavakis E, Rogg E-M, Muhly-Reinholz M, Glaser SF, Günther S, John D, Bonini F, Zeiher AM, Schaefer L, et al. (2018). Switch in laminin b2 to laminin b1 isoforms during aging controls endothelial cell functions—brief report. *Arterioscler. Thromb. Vasc. Biol* 38, 1170–1177. 10.1161/ATVBAHA.117.310685. [PubMed: 29599141]
67. Barallobre-Barreiro J, Didangelos A, Schoendube FA, Drozdov I, Yin X, Fernández-Caggiano M, Willeit P, Puntmann VO, Aldama-López G, Shah AM, et al. (2012). Proteomics analysis of cardiac extracellular matrix remodeling in a porcine model of ischemia/reperfusion injury. *Circulation* 125, 789–802. 10.1161/CIRCULATIONAHA.111.056952. [PubMed: 22261194]
68. Yi X, Jiang D-S, Feng G, Jiang X-J, and Zeng H-L (2019). An altered left ventricle protein profile in human ischemic cardiomyopathy revealed in comparative quantitative proteomics. *Kardiol. Pol* 77, 951–959. 10.33963/KP.14936. [PubMed: 31434864]
69. Bowles NE, Bowles KR, and Towbin JA (2000). The “final common pathway” hypothesis and inherited cardiovascular disease. The role of cytoskeletal proteins in dilated cardiomyopathy. *Herz* 25, 168–175. 10.1007/s000590050003. [PubMed: 10904835]
70. Sweet ME, Cocciolo A, Slavov D, Jones KL, Sweet JR, Graw SL, Reece TB, Ambardekar AV, Bristow MR, Mestroni L, and Taylor MRG (2018). Transcriptome analysis of human heart failure reveals dysregulated cell adhesion in dilated cardiomyopathy and activated immune pathways in ischemic heart failure. *BMC Genom* 19, 812. 10.1186/s12864-018-5213-9.
71. Knowlton AA, Chen L, and Malik ZA (2014). Heart failure and mitochondrial dysfunction: the role of mitochondrial fission/fusion abnormalities and new therapeutic strategies. *J. Cardiovasc. Pharmacol* 63, 196–206. 10.1097/01.fjc.0000432861.55968.a6. [PubMed: 23884159]
72. Brown DA, Perry JB, Allen ME, Sabbah HN, Stauffer BL, Shaikh SR, Cleland JGF, Colucci WS, Butler J, Voors AA, et al. (2017). Mitochondrial function as a therapeutic target in heart failure. *Nat. Rev. Cardiol* 14, 238–250. 10.1038/nrcardio.2016.203. [PubMed: 28004807]
73. Martin M (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. j* 17, 10.
74. Fleming SJ, Marioni JC, and Babadi M (2019). CellBender remove-background: a deep generative model for unsupervised removal of background noise from scRNA-seq datasets. Preprint at bioRxiv 10.1101/791699.
75. Wolf FA, Angerer P, and Theis FJ (2018). SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol* 19, 15. 10.1186/s13059-017-1382-0. [PubMed: 29409532]
76. Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, Baglaenko Y, Brenner M, Loh P-R, and Raychaudhuri S (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* 16, 1289–1296. 10.1038/s41592-019-0619-0. [PubMed: 31740819]
77. Nemenman I, Shafee F, and Bialek W (2002). Entropy and inference, revisited. Preprint at arXiv 10.48550/arXiv.physics/0108025.
78. Wolock SL, Lopez R, and Klein AM (2019). Scrublet: computational identification of cell doublets in single-cell transcriptomic data. *Cell Syst* 8, 281–291.e9. 10.1016/j.cels.2018.11.005. [PubMed: 30954476]
79. Law CW, Chen Y, Shi W, and Smyth GK (2014). voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol* 15, R29. 10.1186/gb-2014-15-2-r29. [PubMed: 24485249]
80. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, and Smyth GK (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 43, e47. 10.1093/nar/gkv007. [PubMed: 25605792]

81. Love MI, Huber W, and Anders S (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15, 550. 10.1186/s13059-014-0550-8. [PubMed: 25516281]
82. Büttner M, Ostner J, Müller CL, Theis FJ, and Schubert B (2021). scCODA is a Bayesian model for compositional single-cell data analysis. *Nat. Commun* 12, 6876. 10.1038/s41467-021-27150-6. [PubMed: 34824236]
83. Chen Y, Lun ATL, and Smyth GK (2016). From reads to genes to pathways: differential expression analysis of RNA-Seq experiments using Rsubread and the edgeR quasi-likelihood pipeline. *F1000Res* 5, 1438. 10.12688/f1000research.8987.2. [PubMed: 27508061]
84. Alexa A, Rahnenführer J, and Lengauer T (2006). Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics* 22, 1600–1607. 10.1093/bioinformatics/btl140. [PubMed: 16606683]
85. Yu G, and He Q-Y (2016). ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. *Mol. Biosyst* 12, 477–479. 10.1039/c5mb00663e. [PubMed: 26661513]
86. Durinck S, Spellman PT, Birney E, and Huber W (2009). Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc* 4, 1184–1191. 10.1038/nprot.2009.97. [PubMed: 19617889]
87. Schneider CA, Rasband WS, and Eliceiri KW (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* 9, 671–675. 10.1038/nmeth.2089. [PubMed: 22930834]
88. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, Hao Y, Stoeckius M, Smibert P, and Satija R (2019). Comprehensive integration of single-cell data. *Cell* 177, 1888–1902.e21. 10.1016/j.cell.2019.05.031. [PubMed: 31178118]
89. Satija R, Farrell JA, Gennert D, Schier AF, and Regev A (2015). Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol* 33, 495–502. 10.1038/nbt.3192. [PubMed: 25867923]
90. Tirosh I, Izar B, Prakadan SM, Wadsworth MH, Treacy D, Trombetta JJ, Rotem A, Rodman C, Lian C, Murphy G, et al. (2016). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* 352, 189–196. 10.1126/science.aad0501. [PubMed: 27124452]
91. Lun ATL, and Marioni JC (2017). Overcoming confounding plate effects in differential expression analyses of single-cell RNA-seq data. *Biostatistics* 18, 451–464. 10.1093/biostatistics/kxw055. [PubMed: 28334062]
92. The Gene Ontology Consortium (2019). The gene ontology resource: 20 years and still GOing strong. *Nucleic Acids Res* 47, D330–D338. 10.1093/nar/gky1055. [PubMed: 30395331]

Highlights

- Cellular expression and composition are altered in ischemic cardiomyopathy
- LAMININ signaling from endothelial cells to other cell types increases in disease
- Cell-type transcriptional profiles converge across multiple cardiomyopathy subtypes
- Potential druggable targets identified for end-stage heart failure

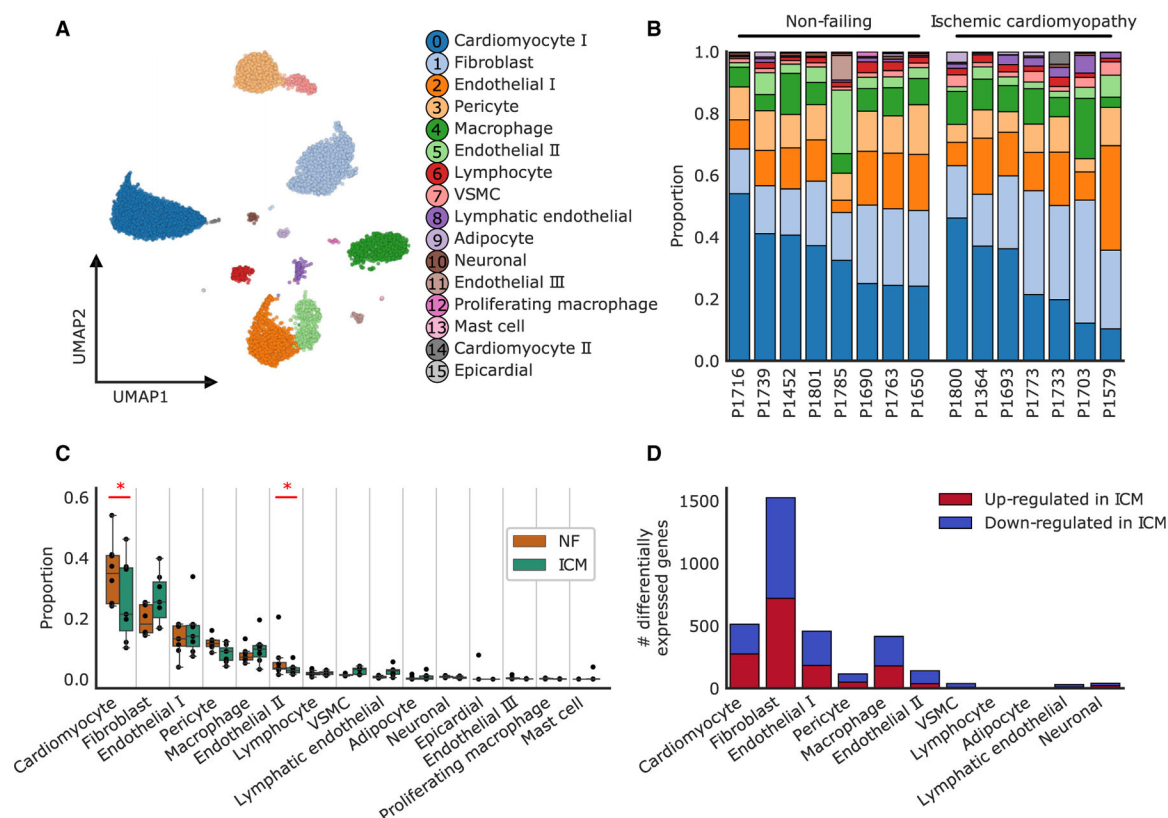


Figure 1. Cellular composition of non-failing and failing ischemic cardiomyopathy

(A) UMAP representation of 99,684 nuclei from 8 NF samples and 7 ICM samples showing 16 cell clusters identified with unsupervised Leiden clustering. VSMC, vascular smooth muscle cell.

(B) Cell-type composition of each individual sample, with colors representing cell types labeled in (A).

(C) Changes in composition observed between NF and ICM samples. Statistically credible changes are noted with an * (see STAR Methods). Boxplots are represented as: center line, median; box limits, upper and lower quartiles; whiskers, 1.5× interquartile range; points, outliers.

(D) Number of significantly differentially expressed genes in each cell type in ICM compared with NF. Upregulated genes are shown in red and downregulated genes are shown in blue.

See also Figures S1–S3 and Tables S1, S2, S3, S4, and S5.

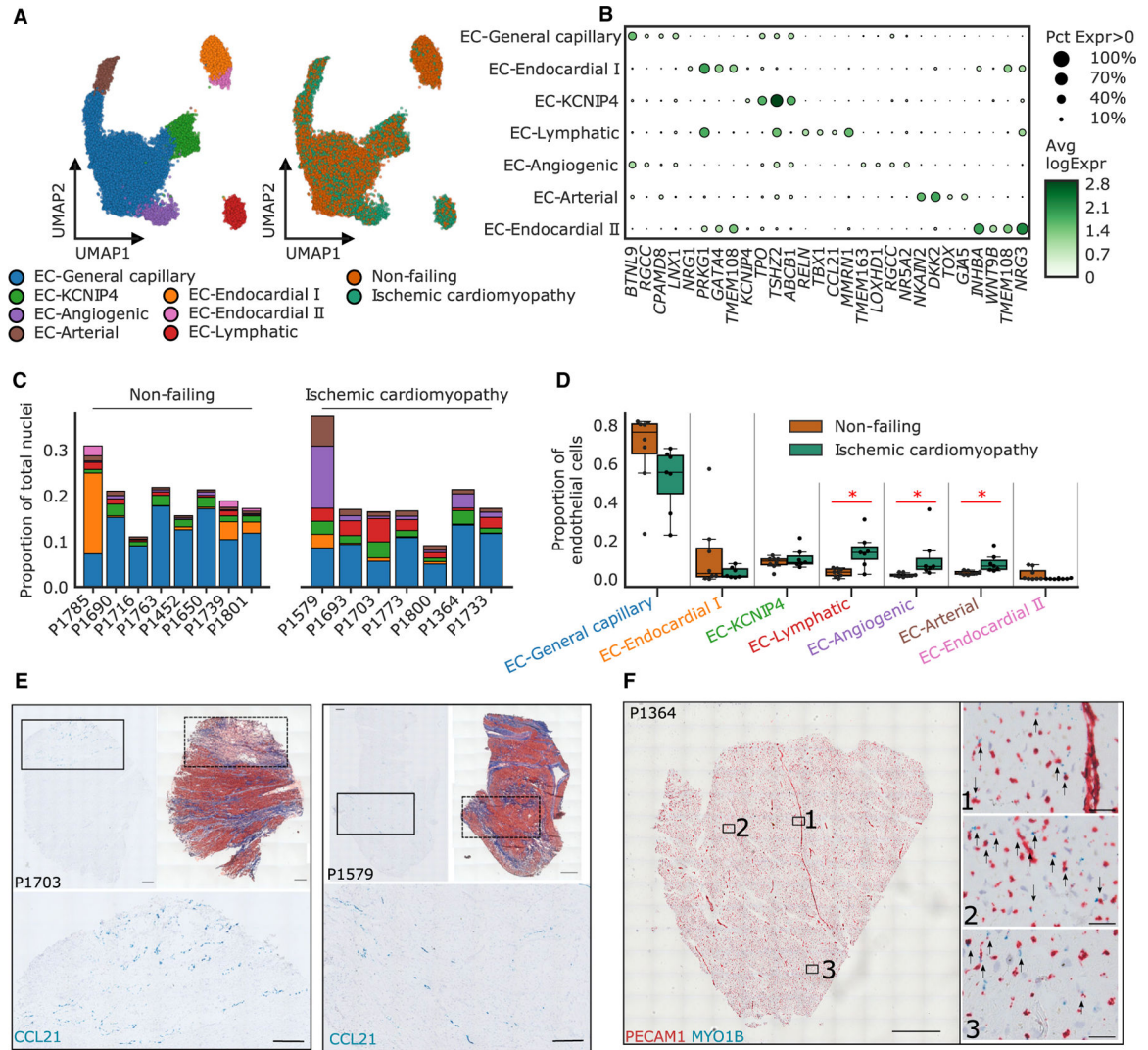


Figure 2. Subclustering of endothelial cells identifies seven distinct subclusters

(A) UMAP of 18,636 endothelial cell identifies 7 subclusters (left) represented across both disease states (right). EC, endothelial cell.

(B) Dot plot of the top marker genes of each endothelial cell subcluster. The size of the dot represents the percent of nuclei expressing the gene at non-zero levels (Pct Expr > 0) and the shading represents the average log-normalized expression (Avg logExpr) of the gene.

(C) Breakdown of endothelial cell composition for each sample relative to the entire composition of all cell types. Subclusters are colored as in (A).

(D) Compositional changes in subclusters between ischemic cardiomyopathy and non-failing samples. Statistically credible increases in EC-lymphatic, EC-angiogenic, and EC-arterial cells were observed in ischemic cardiomyopathy and denoted with an *. Boxplots are represented as: center line, median; box limits, upper and lower quartiles; whiskers, 1.5× interquartile range; points, outliers.

(E) Representative in situ hybridization images of EC-lymphatic cells probed with *CCL21* (blue) alongside trichrome staining of a consecutive section and a cropped close up of an

area with high cell numbers. Nuclei are stained with hematoxylin. Scale bars, 1 mm and 250 μm (on close up images). n = 7 (NF) and 6 (ICM).

(F) Representative *in situ* hybridization images of EC-angiogenic cells probed with *PECAM1* (red) and *MYO1B* (blue). Boxes show where close up images were taken and arrows identify cells stained with both *PECAM1* and *MYO1B*. Scale bars, 1 mm and 50 μm (on close up images). Nuclei are stained with hematoxylin (blue). n = 7 (NF) and 6 (ICM). See also Figures S4 and S5 and Tables S6 and S7.

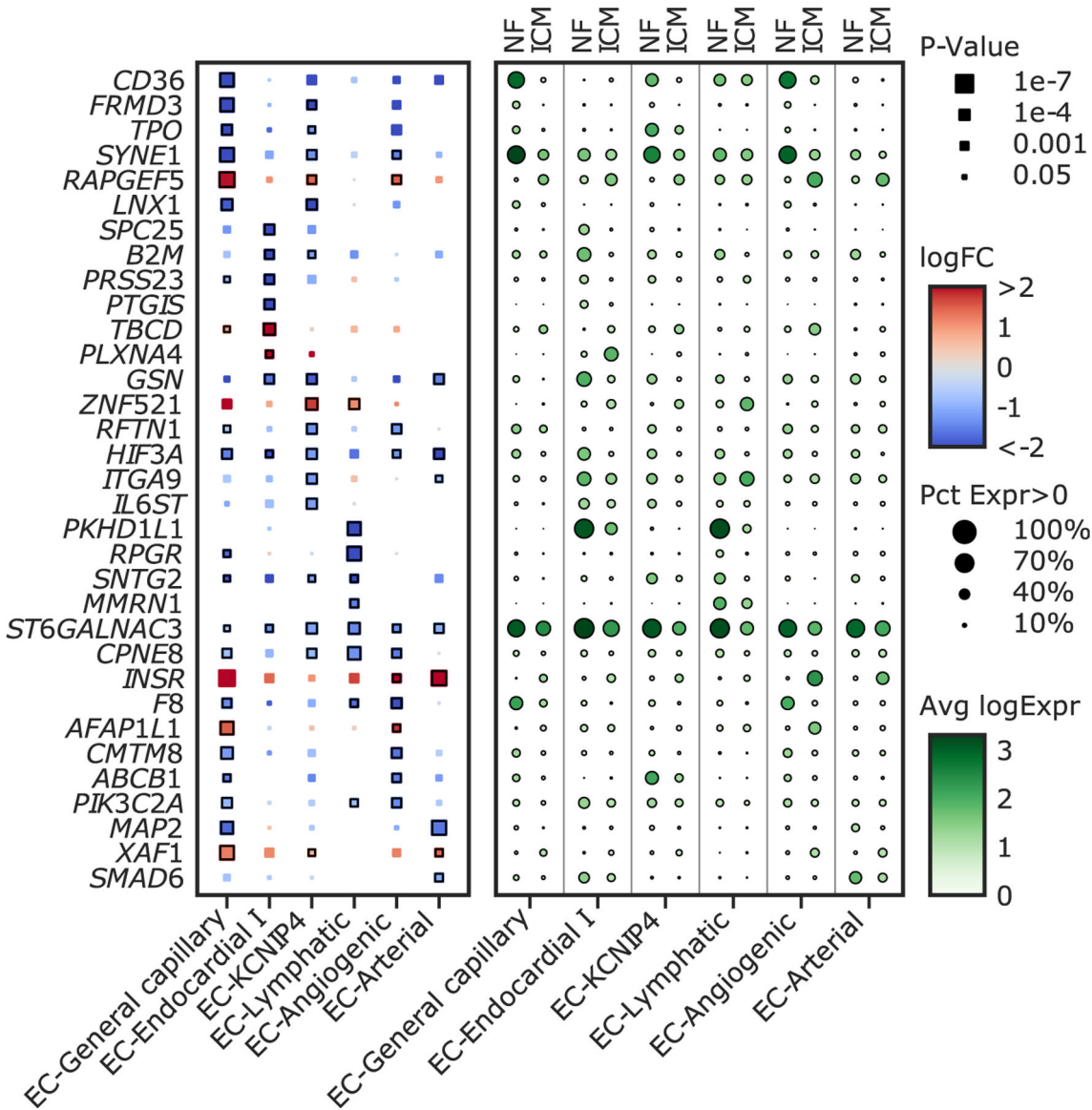


Figure 3. Top differentially expressed protein coding genes in endothelial cell subclusters
Top differentially expressed protein coding genes for each endothelial subcluster. The log fold change (shading) and p value (size) for each gene comparing ICM with NF within each endothelial subcluster is shown on the left. The relative expression in each endothelial subcluster separated by ICM and NF is shown on the right with the size of the dot representing the percent of nuclei expressing the gene and the shade representing the average log-normalized expression. Genes were selected as those expressed in at least 30% of nuclei from either the non-failing or ischemic cardiomyopathy groups with logFC > 1, FDR-adjusted p value < 0.05, and a low probability of being derived from background contamination. NF, non-failing; ICM, ischemic cardiomyopathy; EC, endothelial cell; logFC, log fold change comparing ICM with NF; Pct Expr, percent of nuclei expressing at non-zero levels; Avg logExpr, average log-normalized expression. See also Table S8.



(D) Chord diagram demonstrating the LAMININ signaling pathway within endothelial cell types and all other cell types, highlighting the increase in signaling in ICM compared with NF. The thickness of each connection represents the strength of communication from the origin cell type to the receiver cell type.

(E) Expression of ligand and receptors involved in LAMININ signaling in endothelial cell subclusters and all other cell clusters. The size of each dot represents the percent of nuclei expression the gene and the shading represents the average normalized expression. Red boxes represent changes between NF and ICM that are significant ($p < 0.01$).

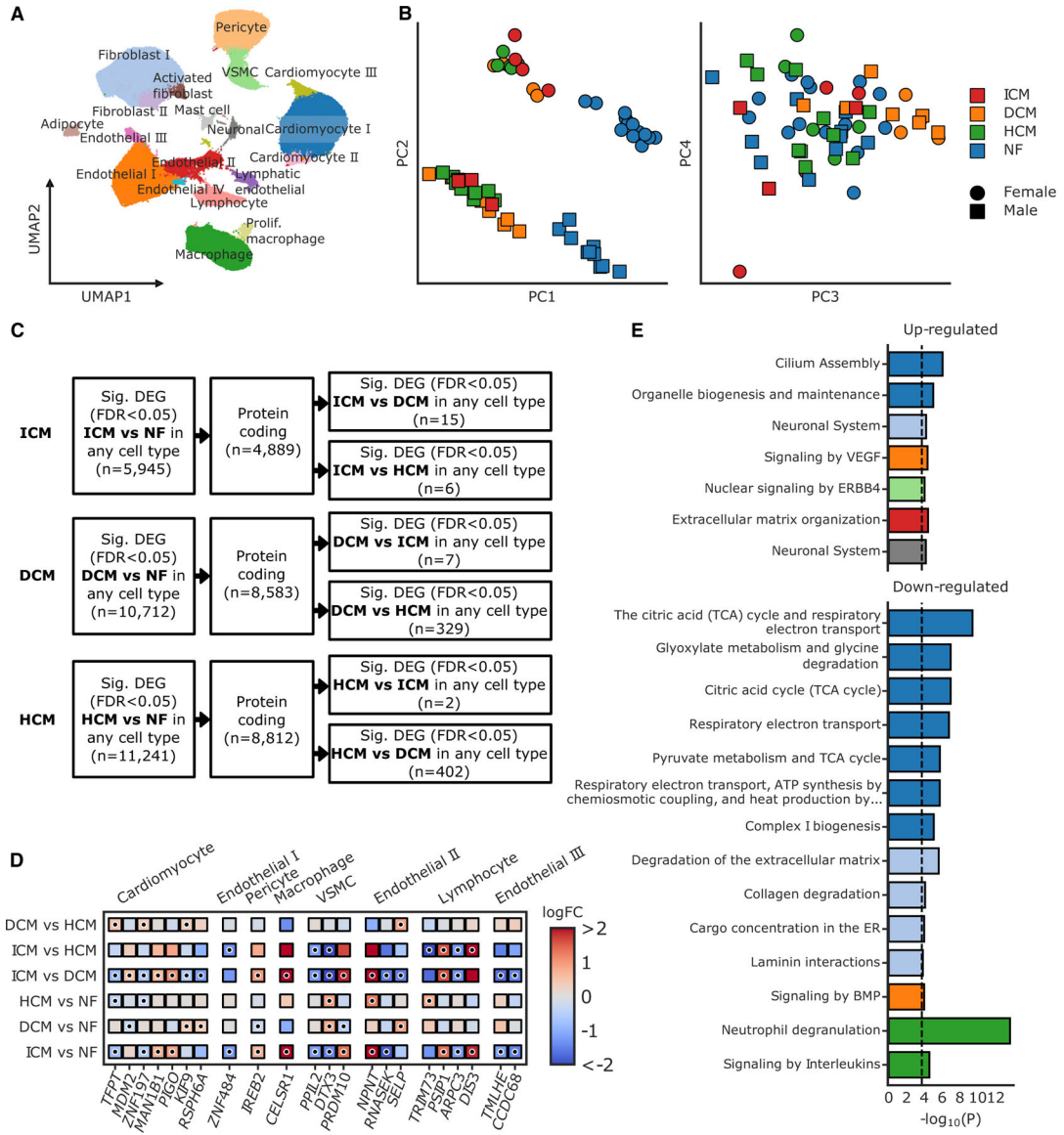


Figure 5. snRNA-seq data on nearly 700,000 nuclei from end-stage cardiomyopathies
 (A) Joint UMAP representation of nuclei from hypertrophic cardiomyopathy (HCM), (n_{patients} = 15), dilated cardiomyopathy (DCM) (n_{patients} = 11), ischemic cardiomyopathy (ICM) (n_{patients} = 7), and non-failing (NF) (n_{patients} = 24) samples (n = 692,373) colored by cell-type clusters.
 (B) Principal-component (PC) analysis of *pseudo-bulk* expression of all samples (n = 57) as calculated by summing expression across all nuclei in each patient. Colors represent the disease state and shape represents the patient sex.
 (C) Flow chart outlining filtering steps carried out to identify protein coding genes that were significantly differentially expressed between any cardiomyopathy and NF, as well as between pairs of cardiomyopathies. DEG, differentially expressed gene; FDR, false discovery rate.

(D) Heatmap of effect size estimates for protein coding genes that were significantly differentially expressed between ICM and NF, as well as ICM and DCM or ICM and HCM (FDR-adjusted p value < 0.05, expressed in >1% of nuclei from either group, low background probability). Also included are genes that were significantly differentially expressed between DCM or HCM and NF, as well as between DCM or HCM and ICM. Shading represents the log fold change (logFC) estimates for each patient population comparison and significant changes are denoted with a dot. VSMC, vascular smooth muscle cell.

(E) Reactome pathways with significant enrichment for genes that were up- or downregulated in all three cardiomyopathies. Multiple testing correction was applied across all cell types/direction combinations using a Benjamini-Hochberg correction (FDR < 0.05, dotted line). Bar colors indicate the cell type in which the significant enrichment was identified, colored as in (A).

See also Figure S7.

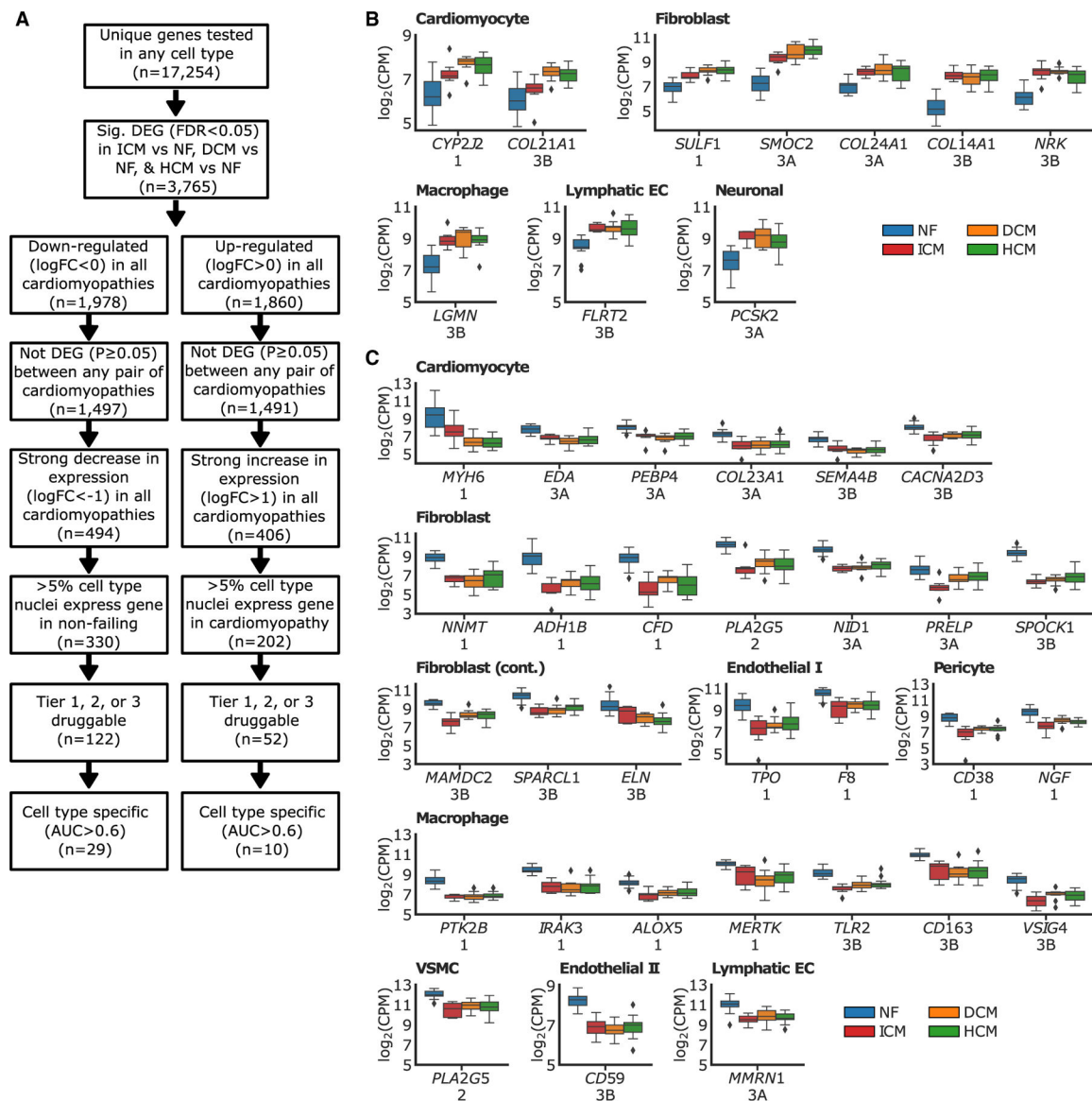


Figure 6. Druggable genes shared by ischemic, hypertrophic, and dilated cardiomyopathies

(A) Flow chart of filtering steps carried out to identify cell-specific druggable genes that are up- or downregulated in ischemic, hypertrophic, and dilated cardiomyopathies. DEG, differentially expressed gene; FDR, false discovery rate; ICM, ischemic cardiomyopathy; NF, non-failing; DCM, dilated cardiomyopathy; HCM, hypertrophic cardiomyopathy; logFC, log fold change; AUC, area under the receiver operating characteristic curve.

(B) Upregulated cell-type-specific druggable genes, with druggable tier noted below.

(C) Downregulated cell-type-specific druggable genes, with druggable tier noted below.

Boxplots are represented as: center line, median; box limits, upper and lower quartiles; whiskers, 1.5× interquartile range; points, outliers. CPM, count per million; EC, endothelial cell; VSMC, vascular smooth muscle cell.

See also Tables S10 and S11.

Table 1.

Clinical characteristics of study participants

	Cases	Controls
Total (n)	7	8
Males (%)	2 (33)	4 (50)
Age (mean \pm SD)	58 (6.2)	54.63 (7.6)
BMI (mean \pm SD)	29.9 (4.5)	30.2 (8.4)
History of hypertension n (%)	5 (71)	2 (25)
History of ICM (%)	7 (100)	0
History of diabetes (%)	3 (43)	0
History of pacemaker device (%)	7 (100)	0
History of CABG (%)	3 (43)	0
Creatinine (mean \pm SD)	1.03 (0.25)	1.9 (1.58)
Initial infarct location n (%)		
Anterior wall	2 (29)	0
Anteroseptal	1 (14)	0
Anterior apical wall	1 (14)	0
Diffuse	2 (29)	0
Time between first infarct and transplant (years, range)	8.32 (0.4–25)	0
Medications n (%)		
Anticoagulants	4 (57)	1 (12.5)
Digoxin	2 (29)	0
Amiodarone	1 (14)	0
ACE inhibitor	6 (87)	1 (12.5)
Beta blocker	7 (100)	0
Angiotensin receptor blockers	1 (14)	0
Calcium channel blockers	2 (29)	0
Nitrate	2 (29)	0
Diuretics	7 (100)	0
Lipid lowering	7 (100)	0
Antiplatelet	4 (57)	0
Acetylsalicylic acid (aspirin)	5 (71)	1 (12.5)
Echocardiography		
LVEF % (mean \pm SD)	23 (17.4)	62.44 (4.8) ****
LVESD cm (mean \pm SD)	4.85 (1.7)	2.56 (0.15) *
LVEDD cm (mean \pm SD)	5.704 (1.5)	3.960 (0.31) *
LVM g (mean \pm SD)	267 (65.4)	197 (40.3) *
LVMI (Male)	119 (25.41)	112.8 (12.16)
LVMI (Female)	149.8 (44.52)	82.77 (4.58) *

ACE, angiotensin-converting enzyme; BMI, body mass index; CABG, coronary artery bypass graft; LVEF, left ventricular ejection fraction; LVESD, left ventricular end systolic dimension; LVEDD, left ventricular end diastolic dimension; LVM, left ventricular mass; LVMI, left ventricular mass index.

*
 $p < 0.05$

**
 $p < 0.01$

 $p < 0.0001$.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
Sucrose	Sigma Aldrich	Cat# 57903–250g
Tris HCl	Sigma Aldrich	Cat# 10812846001
KCl	Sigma Aldrich	Cat# P9333–1kg
MgCl2.6H2O	Sigma Aldrich	Cat# M2670–500g
DTT	Thermo Scientific	Cat# 20291
IGEPAL	Sigma Aldrich	Cat# 18896–50mL
10% BSA	Miltenyi Biotec	Cat# 130–091–376
RNAse Inhibitor Murine	New England Biolabs	Cat# MO314L
16% PFA	Electron Microscopy Sciences	Cat# 15710
Permanent mounting media	Vectamout	Cat# H-5000
Critical commercial assays		
Chromium Single Cell 3' reagent kit, v 3	10x Genomics	Cat# 1000075
Chromium Single Cell B Chip kit, 48 rxn	10x Genomics	Cat# 1000073
Chromium i7 Multiplex Kit, 96 rxn	10s Genomics	Cat# 120262
Trichrome stain kit	Abcam	Cat# 150686
RNAscope 2.5 HD Duplex reagent Kit	ACD bio	Cat# 322430
RNAscope probe hs-CCl2l	ACD bio	Cat# 474371
RNAscope probe hs-PECAM1	ACD bio	Cat# 487381-C2
RNAscope probe hs-MYO1B	ACD bio	Cat# 444301
Deposited data		
Processed snRNA-seq data h5ad	This paper	Broad Institute's Single Cell Portal: SCP1849 (https://singlecell.broadinstitute.org/single_cell/study/SCP1849)
Raw snRNA-seq FASTQ	This paper	dbGaP: phs001539.v4.p1; SRA: PRJNA433594
Software and algorithms		
Custom snRNA-seq analysis code	This paper	Zenodo: https://doi.org/10.5281/zenodo.7469682 ; GitHub: https://github.com/mark-chaffin/snrnaseq_ischemic_cardiomyopathy

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

REAGENT or RESOURCE	SOURCE	IDENTIFIER
R 3.6.0; 4.1	R Foundation	https://www.r-project.org/
Python 3.7.3	https://www.python.org/	https://www.python.org/
Cell Ranger 4.0.0	10X Genomics	https://support.10xgenomics.com/single-cell-gene-expression/software/overview/welcome
cutadapt 1.18	Martin, 2011 ⁷³	https://github.com/marcelm/cutadapt
CellBender 0.2	Fleming et al., 2019 ⁷⁴	https://github.com/broadinstitute/CellBender
scR-Innex	https://github.com/broadinstitute/scrinvex	https://github.com/broadinstitute/scrinvex
scanpy 1.7.2	Wolf et al., 2018 ⁷⁵	https://github.com/scverse/scanpy
harmony-pytorch 0.1.4	Korsunsky et al., 2019 ⁷⁶	https://github.com/lilab-bcb/harmony-pytorch
ndd 1.6.3	Nemenman et al., 2002 ⁷⁷	https://github.com/simomarsili/ndd
Scrublet 0.2.1	Wolock et al., 2019 ⁷⁸	https://github.com/svoolock/scrublet
limma 3.44.3	Law et al., 2014 ⁷⁹	https://doi.org/10.18129/B9.bioc.limma
DESeq2 1.20.0	Love et al., 2014 ⁸¹	https://doi.org/10.18129/B9.bioc.DESeq2
scCODA 0.1.2.post	Büttner et al., 2021 ⁸²	https://github.com/theislab/scCODA
edgeR 3.30.3	Chen et al., 2016 ⁸³	https://doi.org/10.18129/B9.bioc.edgeR
topGO 2.32.0	Alexa et al., 2006 ⁸⁴	https://doi.org/10.18129/B9.bioc.topGO
ReactomePA 1.24.0	Yu and He, 2016 ⁸⁵	https://doi.org/10.18129/B9.bioc.ReactomePA
biomaRt 2.36.1	Durinck et al., 2009 ⁸⁶	https://doi.org/10.18129/B9.bioc.biomaRt
CellChat 1.5.0	Jin et al., 2021 ⁴¹	https://github.com/sqjin/CellChat
GraphPad Prism	Dotmatics	https://imagej.nih.gov
(Fuji is just) ImageJ	Schneider et al., 2012 ⁸⁷	Graphpad.com