A cross-study transcriptional patient map of heart failure defines conserved multicellular coordination in cardiac remodeling

Jan D. Lanzer[1 †], Ricardo O. Ramirez Flores[1,2 †], José Liñares Blanco[2], and Julio Saez-Rodriguez[1,2*]

[1] Heidelberg University, Faculty of Medicine, and Heidelberg University Hospital, Institute for Computational Biomedicine, Heidelberg, Germany

[2] European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Hinxton, Cambridgeshire, U.K.

* Corresponding author: saezlab@ebi.ac.uk

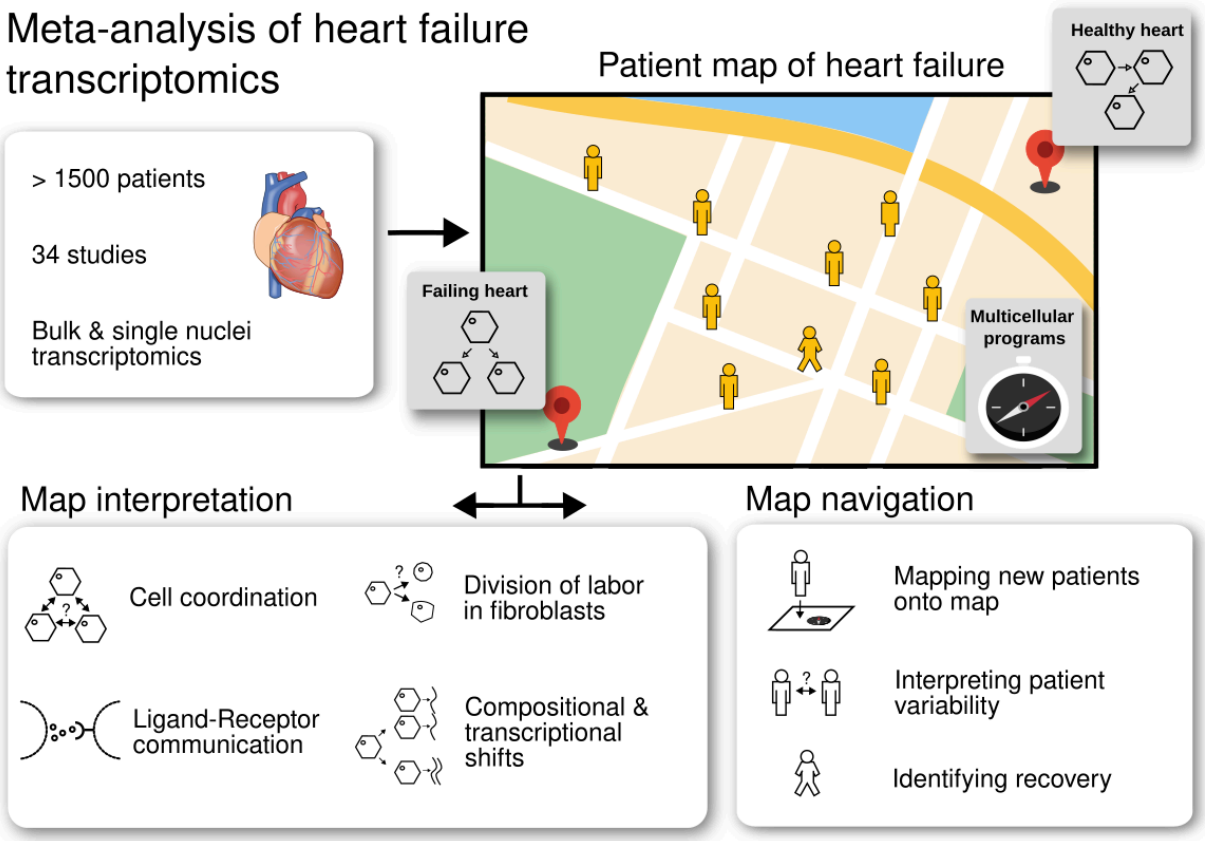
[†] Equal contribution

Abstract

Heart failure (HF) is characterized by severely reduced cardiac function and tissue remodeling, driven by complex multicellular regulatory processes. Extensive studies have generated molecular profiles at both bulk and single-cell levels; however, systematic integration that describes tissue-wide changes as a function of cell type coordination remains challenging. This disconnect hampers our understanding of the complex multicellular interactions driving heart failure, limiting our ability to translate molecular insights into actionable therapeutic strategies. Here, we integrated bulk and single-cell transcriptional profiles from cardiac tissues of HF and control patients across 25 studies, covering 1,524 individuals and seven cell-types, to delineate consensus multicellular transcriptional changes associated with cardiac remodeling. Our analyses revealed conserved cellular coordination events involving fibrotic, metabolic, inflammatory, and hypertrophic mechanisms, with fibroblasts playing a central role in predicting cardiomyocyte stress. Further analysis of fibroblast populations suggested that their activation in HF represents a broad phenotypic shift rather than solely accumulating distinct cell states. The integration of bulk and single-cell data within our data collection indicated that transcriptional responses to HF across cell types occur independently of tissue composition. Mapping independent data into our consensus programs demonstrated that recovery after left ventricular assist device implantation aligns with molecular recovery, highlighting the clinical relevance of the multicellular molecular state. Overall, our work synthesizes independent cardiac transcriptomics studies and makes the conserved HF associated

insights available, establishing a reference for detailed exploration of HF-related multicellular molecular events.

Graphical Abstract



1. Introduction

Molecular descriptions of the inter- and intra-cellular signaling mechanisms of cardiac cells in health and disease are essential to describe the pathophysiology of heart failure (HF). Historically, bulk transcriptomics has been used to compare general functional and compositional characteristics of the healthy and failing hearts. However, technological advances now facilitate the profiling of single cells in tissues, expanding our knowledge on the specific molecular processes carried out by the muscular, vasculature, stromal, neuronal, and immune cells that compose the heart.

Increasing collections of single-cell atlases chart the compositional and functional diversity of cardiac tissue of distinct heart compartments, and during development or disease [1–10]. Each study has established its individual classification of single cells based on lineage (i.e. cell-types) and functional potential (i.e. cell-states), complicating the comparison of the behaviors of cardiac cells across distinct biological or clinical contexts. While efforts to establish unified cell ontologies [11], integrate studies [12], or provide access to the information from distinct single-cell studies [13] exist, they are focused on the generation of cell-taxonomies that lack the contextualization of single-cell behaviors in terms of multicellular processes [14,15].

Understanding multicellular cooperation is critical because it ensures proper functioning of the heart and to what extent distinct pathological processes trigger similar multicellular responses in patients is not known. Furthermore, such a multicellular tissue-centric analysis of single-cell data opens an opportunity to reinterpret bulk transcriptomics data beyond cellular compositions and complement the knowledge of single-cell studies that are limited by the low number of patient samples they profile.

We recently reported a consensus transcriptional disease signature of HF built from the meta-analysis of public bulk transcriptomics datasets that pointed to a convergent transcriptional response of ischemic (ICM) and dilated (DCM) cardiomyopathies [16]. However, given the resolution of bulk transcriptomics, we were unable to quantify to what extent the convergent disease signature of HF was the product of specific cell-types or tissue changes in composition and multicellular coordination.

Here we created a multicellular reference of the HF transcriptome by integrating the single-cell and bulk transcriptional profiles from cardiac tissues of control and HF patients across 25 core studies (1,524 patients) and nine supporting studies (24 mice samples and 667 patients) to provide consensus multicellular processes associated with inflammation, fibrosis, cardiac remodeling, and fetal reprogramming in distinct etiologies. First, we evaluated the consistency of the multicellular molecular

and compositional changes of cardiac tissue during HF from single-cell studies. Then, we built a consensus patient map, a latent space that captures the variability of tissue samples across single-cell studies based on coordinated transcriptional changes occurring in multiple cell-types, here referred to as multicellular programs. We reconstructed networks of cell-dependencies underlying multicellular coordinated responses upon HF and identified a central role of fibroblasts in predicting the behavior of the other cell-types, particularly stressed cardiomyocytes. We further studied the distribution of HF responses across the population of fibroblasts to link the diversity of molecular phenotypes to tissue-level responses. Moreover, we showed that the fibrotic multicellular program of HF can be traced in bulk transcriptomics data. Finally, we highlighted the clinical relevance of our inferred consensus multicellular programs of HF by projecting independent single-cell datasets onto our patient map. We summarized our integration work and made it publicly available at "**R**eference of the **Hea**rt failure **T**ranscriptome" (ReHeaT, https://saezlab.shinyapps.io/reheat2/) to allow researchers to query if individual genes are associated with HF in a cell-type specific matter, facilitating a more nuanced exploration of the disease at the multicellular level.

2. Results

2.1 Study curation for the creation of the reference of the heart failure transcriptome

We queried public repositories of single-cell and bulk transcriptomics data to identify studies that profiled left-ventricle tissue samples of healthy non-failing (NF) controls and HF patients with the same inclusion criteria to ensure consistency with our previous bulk resource (Methods, Figure 1A, Supplemental Table 1). In addition to the previously reported 16 bulk studies, we identified nine new HF studies - four single-nucleus (SN) [4,7–9] and five bulk studies [17] [18] [19] [20,21] - that fulfilled inclusion criteria, resulting in 25 core studies. Nine HF studies that could not be included due to reasons including disease etiology, species, or sample size requirements (Methods) were termed supporting studies and were used to corroborate and reinterpret our findings (Table 1). To facilitate the integration and comparison of all studies we aligned meta data and cell type annotations to a harmonized vocabulary covering seven major cell-types: Cardiomyocytes (CM), fibroblasts (Fib), pericytes (PC), and endothelial (Endo), vascular smooth muscle (vSMCs), myeloid, and lymphoid cells. This data curation resulted in the largest collection of transcriptomic HF tissue samples to date spanning a total of 2,215 samples (1524 tissue samples [SN 132, bulk 1392] from core studies [Figure 1B] , and 691 tissue samples from supporting studies).

StudyID	Data modality	Cohort	Samples	HF Etiology / Model	Species	Citation
Wang22	Bulk	Core Study	15	ICM	Human	[18]
Forte21	Bulk	Core Study	10	ICM	Human	[20,21]
Rao21	Bulk	Core Study	67	HF	Human	[19]
Hua19	Bulk	Core Study	30	ICM	Human	[21]
Flam19	Bulk	Core Study	354	HCM, DCM	Human	[17]
Chaffin2022	Single-nucleus	Core Study	42	HCM, DCM	Human	[8]
Koenig2022	Single-nucleus	Core Study	38	DCM	Human	[7]
Reichart2022	Single-nucleus	Core Study	37	DCM, ARVC	Human	[4]
Simonson2023	Single-nucleus	Core Study	15	ICM	Human	[9]
Hill2022	Single-nucleus	Supporting Study	12	CHD	Human	[10]
Nicin2022	Single-nucleus	Supporting Study	5	HCM	Human	[22]
Liu2022	Single-nucleus	Supporting Study	15	Sarcoidosis, ICM	Human	[23]
Kuppe2022	Single-nucleus	Supporting Study	20	MI	Human	[5]
Amrute2023	Single-nucleus	Supporting Study	13	LVAD	Human	[6]
Mehdiabadi2022	Single-nucleus	Supporting Study	10	pediatric DCM, fetal	Human	[24]
Litviňuková2020	Single-nucleus	Supporting Study	14	Healthy	Human	[1]
McLellan2020	Single-cell	Supporting Study	8	AngII	Mouse	[25]
Ren2020	Single-cell	Supporting Study	16	TAC	Mouse	[26]

Table 1. Overview of data sets used in this study.

ICM, ischemic cardiomyopathy; HF, heart failure; DCM, dilated cardiomyopathy; HCM, hypertrophic cardiomyopathy; ARVC, arrhythmogenic right ventricular cardiomyopathy; CHD, congenital heart disease; MI, myocardial infarction; LVAD, left ventricle assist device; AngII, angiotensin II; TAC, transverse aortic constriction.

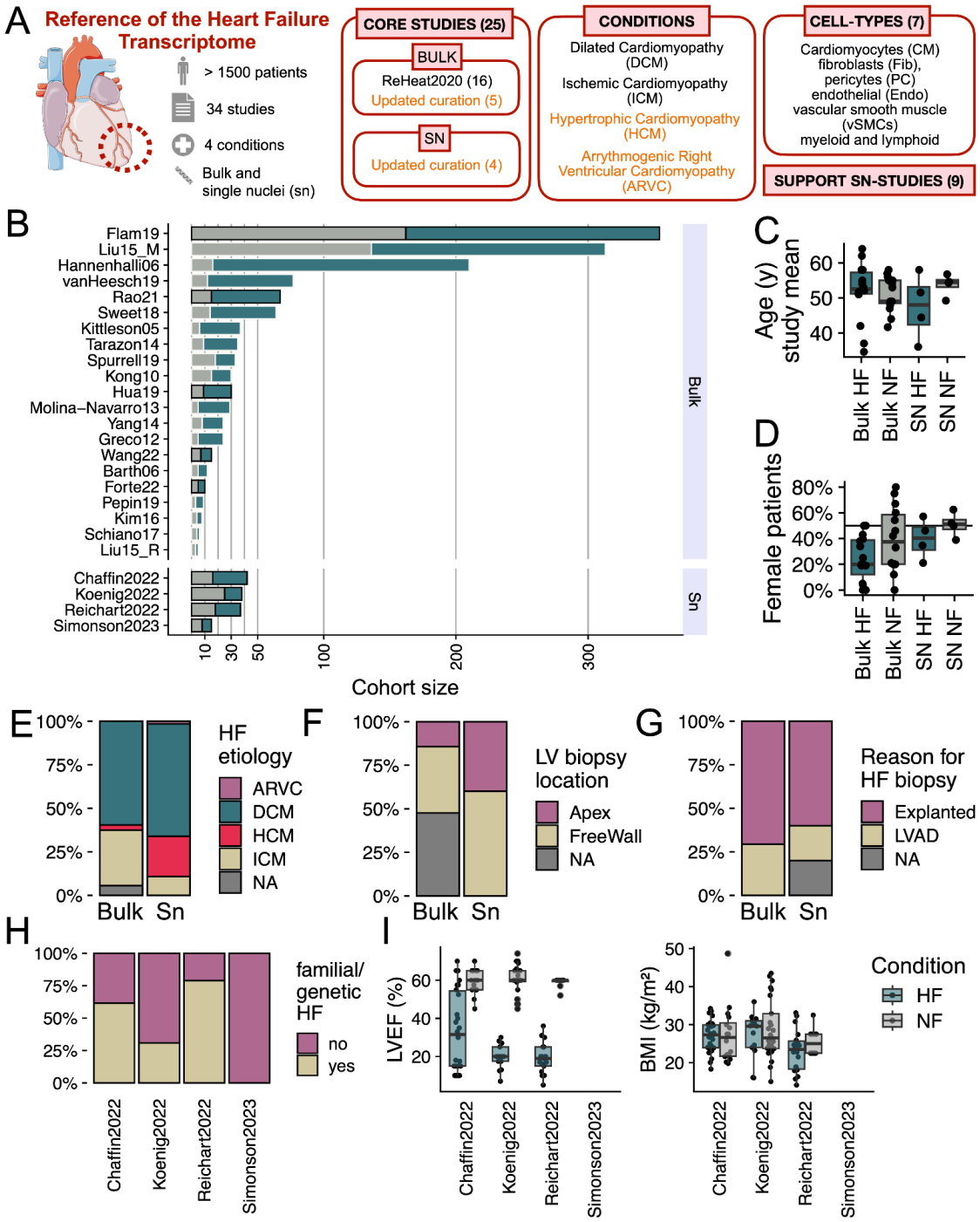


Figure 1. Study and data overview and meta data presentation.

A. Our updated reference integrates the molecular profiling of heart tissues of patients with HF and control donors from bulk and single-nucleus (SN) technologies, across distinct end-stage conditions, and cell-types. In addition, we collected supporting data from studies profiling related heart conditions, smaller sample sizes, and mice models.

B. Sample size (x-axis) of core HF studies separated for bulk and SN technologies. Control patients (NF) in gray and HF (HF) patients in green, black outline indicate new data added to the reference.

C. Distribution of mean patient age per study, categorized by sequencing modality and HF status.
D. Percentage of female patients per study, categorized by sequencing modality and HF status.
E-G. HF patient samples across bulk and SN studies categorized by (E) HF etiology, (F) sampled location, (G) and reason for biopsy.
H. HF patient samples in SN studies categorized by genetic or familial cardiomyopathy.
I. Distribution of left ventricular ejection fraction (LVEF, left) and body mass index (BMI, right) across heart tissue samples in SN studies.

2.1.1 Clinical presentation of 24 core heart failure studies

To assess the clinical diversity of the distinct patient cohorts included in our curation, we compared core studies based on reported patient-level meta-data. HF and non-failing (NF) patients were similar in age, with a mean of 51.7 years for HF and 52.1 years for NF (Figure 1C, Supplementary Figure 1A). Studies undersampled female patients, with a total mean of 27% for HF and 44% for NF, though recent single-nucleus studies showed a higher proportion of female patients (bulk studies: HF 24%, NF 42%; SN studies: HF 40%, NF 51%) (Figure 1D, Supplementary Figure 1B). Dilated cardiomyopathy (DCM) was the most common etiology across studies, accounting for 60% of HF patients (59% in bulk, 65% in SN studies), followed by ischemic cardiomyopathy (ICM) at 31% (32% bulk, 11% SN) and hypertrophic cardiomyopathy (HCM) at 4% (3% bulk, 23% SN) (Figure 1E, Supplementary Figure 1C). The left ventricular (LV) free wall was the most sampled location (73% of total HF patients, 74% bulk, 58% SN), with the LV apex less frequently sampled (10% of total HF patients, 7% bulk, 42% SN) (Figure 1F). No septal samples were included in the collection. HF samples were acquired after cardiectomy (83% of total HF patients, 85% in bulk, 59% in SN) or during left ventricular assist device (LVAD) implantation (16% total, 15% bulk, 22% SN) (Figure 1G, Supplementary Figure 1D). Cardiomyopathies can arise from primary (familial) or secondary (acquired) factors. SN studies reported 60% of DCM and HCM patients (and one arrhythmogenic right ventricular cardiomyopathy [ARVC] patient) were suspected to have a primary etiology (Figure 1H, Supplementary Figure 1E). Left ventricular ejection fraction (LVEF) and body mass index (BMI) were documented in three of the four SN studies (Figure 1I), with a median BMI of 27 kg/m² for NF patients and 26 kg/m² for HF patients, and a median LVEF of 60% in NF and 20% in HF patients. Notably, race was reported for only 52% of SN patient samples, with 42% of those identified as Caucasian, 8% as African American, and less than 1% as Asian or South Asian (Supplementary Figure 1F). In summary, our curation reveals notable sampling biases in gender, age, race, and disease stage, likely due to challenges in cardiac biopsy acquisition.

2.2 Comparison of the core heart failure studies in bulk and single cell for molecular convergence

First we compared the molecular similarity of HF in the curated core bulk and single-nucleus transcriptomics studies. These comparisons focused on three tissue characteristics: bulk expression profiles from bulk transcriptomics data, which represent the overall molecular state of the tissue; cell-type compositions from single-nucleus transcriptomics, which describe the structure of the profiled samples; and multicellular programs, which capture coordinated molecular activities across different cell types within the tissue. These multicellular programs represent interactions and shared molecular activities between various cell types, helping to define the functional states of the tissue. We hypothesized that if two independent studies using the same technology captured similar molecular processes in HF, a disease classifier trained on one study's data could predict the disease status of heart tissue samples in the other (Figure 2A). This approach allowed us to assess how well independent studies with different technical and clinical characteristics aligned at various scales.

2.2.1 Comparison and meta-analysis of the core bulk transcriptomics studies

We compared the bulk studies in a pairwise manner to assess their agreement on gene expression changes reported in HF (methods) and we found that the overlap of differentially expressed genes of the new studies was low (mean Jaccard index 0.055, Supplementary Figure 2A). However, transcriptional directionality was conserved, as shown by cross-study enrichment analysis (upregulated mean enrichment score 0.58, downregulated mean enrichment score -0.54, Supplementary Figure 2B) and pairwise study classification (mean AUROC 0.89, Figure 2B; see Methods), consistent with our earlier findings [16]. This indicated that the new studies captured the directionality of transcriptional changes in HF, and were thus included to update the consensus signature through a gene-level meta-analysis (Figure 2C). Additional analysis of the updated consensus signature, including comparisons with the previous version in terms of gene set overlap and classifier performance, showed that its predictive value for HF remained stable despite the reordering of the top genes (Supplementary Note 1, Supplementary Figure 2C-F). Overall, these results confirmed the existence of a stable transcriptional response in HF from bulk transcriptomics.

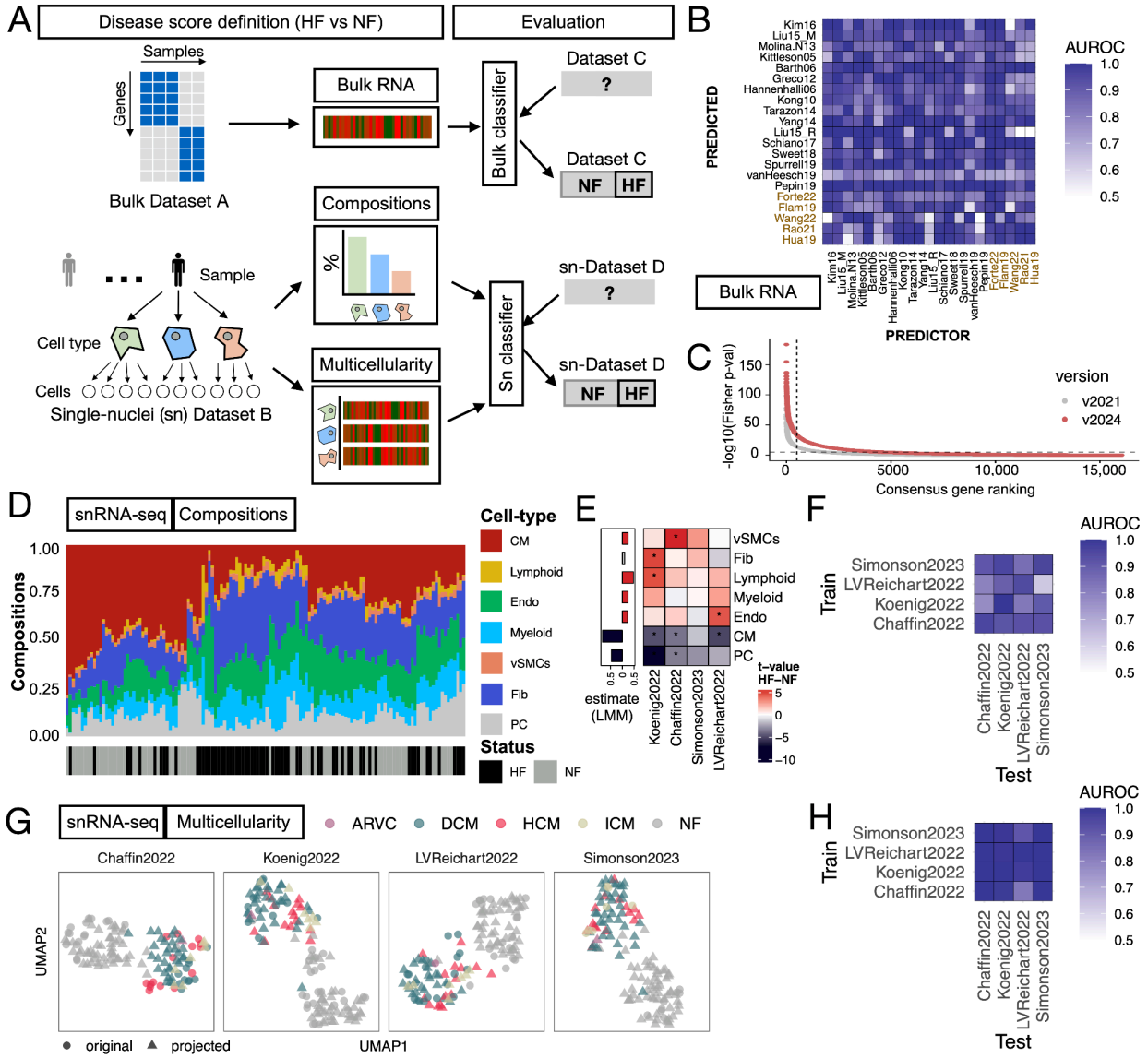


Figure 2. Comparison of HF molecular descriptions across studies

A. Similarities of the molecular descriptions of HF between pairs of bulk or single-nucleus (SN) transcriptomics studies were approximated with the performance of disease classifiers built with each study of our core curation. The disease classifier from bulk studies used the changes in expression of each gene, while the classifiers of SN-studies were built with cell-type compositions or multicellular programs.

B. Area under the receiver operating characteristic curve (AUROC) of pairwise predictions of disease classifiers built from all individual bulk studies included in our core collection using the top 500 differentially expressed genes. Red labels mark the study expansion.

C. Gene ranking of conservation of gene deregulation events in HF across bulk transcriptomics studies based on the adjusted p-value of a Fisher combined analysis. Dotted line shows the top-500 genes.

D. Hierarchical clustering of all SN-transcriptomics tissue samples based on the composition of the seven major cell-types used in our ontology.

E. T-statistics (heatmap) and estimate of difference between failing and non-failing hearts from the differential compositional analysis using t-tests and linear mixed-models, respectively. Stars denote an adj. p-value < 0.05

F. AUROC of pairwise predictions from disease classifiers built from all core SN studies using cell-type compositions.
G. Uniform Manifold Approximation and Projections (UMAP) of the multicellular programs describing the variability of the respective tissue samples of each core SN study. Tissue samples of the rest of the studies are projected into each latent space and distinguished with the shape of the dot. Colors highlight different HF etiologies.
H. AUROC of pairwise predictions from disease classifiers built from all core SN-studies using multicellular programs.

In all panels Cardiomyocytes (CM), fibroblasts (Fib), pericytes, and endothelial (Endo), vascular smooth muscle (vSMCs) cells. Heart failure (HF), and non-failing (NF) hearts.

2.2.2 Comparison of the single-nucleus transcriptomics core studies

We then compared single-nucleus studies by summarizing each individual atlas into a patient-level description of the molecular profiles for each cell type in our ontology, as well as their relative compositions (Figure 2A). To enable cross-study comparisons, we generated pseudobulk expression profiles for each sample and cell type. As expected, we observed variation in the number of genes measured and the number of cells captured across studies, reflecting technical differences (Supplementary Figure 3A-C, Supplementary Note 2). Additionally, we quantified the levels of gene expression contamination within each pseudobulk profile to account for the effects of ambient molecules, which were comparable across studies and could influence the estimation of cell-type specific disease signatures (Supplementary Figure 4A-D, Supplementary Note 2). Next, we evaluated the consistency of cell-type annotations across studies, after harmonizing them to a joint ontology, by comparing gene expression markers for each cell type calculated independently in each study. We observed agreement of markers between cells aligned to the same type (median Jaccard Index 0.48), indicating that the studies classified cells into comparable types. In contrast, comparisons between cells of different types showed very low agreement (median Jaccard Index 0.002), further supporting the accuracy of the original cell annotations (Supplementary Figure 3D). Finally, we generated a consensus set of cell-type markers across studies by performing a meta-analysis of the differentially expressed genes for each cell type (Supplementary Table 2).

Pathologic tissue remodeling often manifests as shifts in the abundance of resident cell lineages. To identify such a compositional disease signature of HF, we quantified and compared cell-type compositions across all tissue samples in the core SN studies. Based on the proportions of the seven cell types defined in our ontology, we observed modest grouping of patients by HF status (one-sided t-test, adj. p-value < 0.05) and no clear grouping by study (median silhouette width for HF, NF, and study = 0.14, 0.19, and -0.02, respectively, where higher values indicate well-defined clusters). We next focused

on how compositional differences between failing and non-failing hearts varied across the core studies. First we inferred compositional disease signatures (i.e. differences in cell-type proportions between failing and non-failing hearts) for each study independently using differential compositional analysis across cell-types (Figure 2E). We then tested how well these signatures classified failing and non-failing hearts in the other core studies, and we summarized these results using receiver operating characteristic (ROC) curves (Figure 2F). Pairwise classification tests showed an average area under the ROC curve (AUROC) of 0.86, suggesting that the direction of cell-type compositional changes was comparable across studies. Finally, to identify the most consistent compositional changes in HF across studies, we used linear mixed models of cell-type proportions using as predictors the HF status of a patient while controlling for individual studies treated as random effects. We observed consistent patterns of cell-composition changes across studies, despite low effect sizes and high variability between studies (Figure 2E). All cell-types changed between failing and non-failing hearts except for fibroblasts (adj. p-value < 0.05, mean proportion of explained variance associated with study across cells = 0.22), with a decrease in cardiomyocytes (CMs) and pericytes (PCs) and an increase in lymphoid cells being the most characteristic changes in failing hearts. Although cell-type proportions varied between studies, the results consistently revealed specific compositional changes in HF. However, the imperfect separation of conditions suggests that, while these proportions offer some insight into disease status, they could be influenced by study-specific factors, such as tissue sampling design, and/or that composition varies biologically between patients independently of HF status.

Next, we explored whether the observed differences in tissue composition among patient groups influenced the comparability of the gene expression profiles of the cell-types in tissue samples from failing and non-failing hearts across single-nucleus studies. We reasoned that, since cells in tissues function as collectives, the molecular state of a tissue sample could be represented by coordinated transcriptional events across multiple cell types where gene expression changes of one cell-type relates to the changes of other cell-types, referred to here as multicellular programs (MCPs). We applied multicellular factor analysis [27] independently to each study in our core SN collection to infer and compare their multicellular programs that describe patient variability within each study (Methods). The multicellular programs of individual studies captured variance related to HF and other clinical covariates, clearly separating failing from non-failing hearts. These results suggested the presence of multicellular programs that reflect HF-associated tissue remodeling within each study (Supplementary Note, Figure 2G, Supplementary Figure 5B-D).

When examining the individual cell types participating in these programs, fibroblasts showed the highest explained variance associated with HF (mean explained variance = 25%, one-sided t-test p-value < 0.05), while lymphoid cells exhibited the lowest (mean explained variance = 6%, one-sided t-test p-value < 0.05). These results contrast with our previous compositional analyses, where lymphoid cells showed significant compositional changes in HF, while fibroblast compositions remained stable across studies. This contrast highlights the complex interplay between compositional and molecular changes in failing heart tissues: although cell-type compositional shifts are associated with myocardial remodeling in HF, they do not necessarily align with the independent, coordinated molecular changes observed.

To compare the similarity of multicellular programs in predicting HF across studies, we projected the samples from the other three core studies into the multicellular space of each individual study and evaluated their ability to differentiate failing from non-failing hearts using pre-trained classifiers (Figure 2G-H, Methods). The classification tests yielded a mean AUROC of 0.98, indicating high concordance in the molecular state of the tissues across studies. When assessing the contribution of individual cell types to the classification task (Methods, Supplementary Figure 5D), we found that all cell types, except lymphoid cells, had a mean AUROC of 0.85 or higher. This suggests that most cell types contribute equally to the multicellular program underlying HF. These findings suggest that, despite technical and clinical differences among the core SN studies, these studies share common multicellular processes that distinguish failing and non-failing hearts.

2.3 Multicellular patient map of end-stage heart failure

Given the observed multicellular molecular similarity across the core SN studies, we built a joint multicellular patient map using multicellular factor analysis to identify both shared and specific axes of variation in the tissue samples (Figure 3A). We decomposed the variability in gene expression across cell-types and patients into 10 multicellular programs, with a median R^2 across cell types of 31% (Supplementary Figure 6A-C). The multicellular space showed little to no variability associated with body mass index, age, or sex (mean R^2 of 0.3%, 3%, and 0%, respectively). However, 17% of the gene expression variability across cell types was associated with HF (Supplementary Figure 6D). Among all cell types, fibroblasts exhibited the highest percentage of variance explained by HF, with an average of 24.5% across studies (t-test p-value = 0.01), despite not having the largest number of genes in the model (Supplementary Figure 6A). This suggests that fibroblasts undergo the most pronounced molecular changes in response to HF and that multicellular coordination of other cell-types.

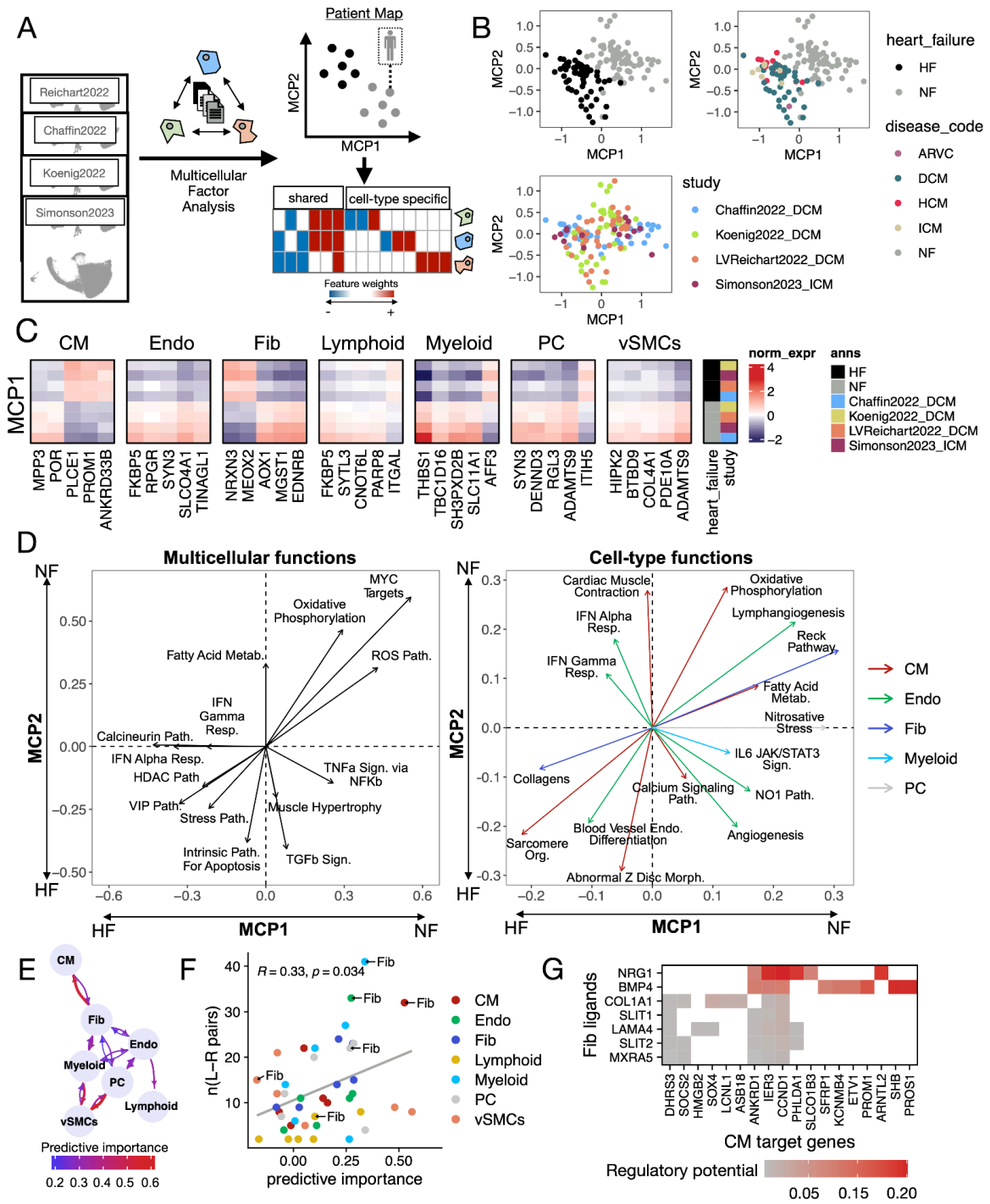


Figure 3. Consensus multicellular landscape of human HF

A. Multicellular factor analysis was used to integrate the patients profiled across the single-nucleus core studies. The integrative model represents each sample in terms of latent variables, referred here as multicellular programs (MCP), that capture gene expression variability across cell-types, patients, and studies. Each MCP can be understood as a collection of genes whose expression is coordinated across cell-types.

B. Patient map built from MCP1 and MCP2 with samples colored by their disease status, etiology, and study

of origin.

C. Mean standardized gene expression of top 5 genes in MCP1 whose expression was captured in all studies. Data was grouped by disease status and study of origin.

D. Functional dissection of the patient map built from MCP1 and MCP2 from a multicellular (left) or cell-type (right) perspective. Each "functional vector" represents the level of enrichment of a function in the location of the map where the arrow points to. The larger the arrow, the more enriched the function. Gene sets were manually selected for representation, from a set of gene sets enriched in either MCP1 and/or MCP2 (adj. p-value < 0.1).

E. Multicellular information network of HF processes captured by MCP1, where each arrow describes how important the expression of a given cell-type is to predict the expression profile of another one (Methods). Predictive importances come from linear mixed models of cell-type signatures of MCP1. Importances below 0.2 were not included.

F. Correlation between the predictive importance of a pair of cell-types (sender and target) and the number of potential ligand-receptor coexpression events. Pairs of cells are coloured by their target cell-type and highlighted when fibroblasts (Fibs) are the sender cell-type.

G. Regulatory potential score, as estimated by NicheNet, representing the potential of fibroblasts' ligands in contributing to the regulation of cardiomyocyte genes (MCP1 gene loading > 0.2).

In all panels Cardiomyocytes (CM), fibroblasts (Fib), pericytes (PC), and endothelial (Endo), vascular smooth muscle (vSMCs) cells. Heart failure (HF), and non-failing (NF) hearts.

Within our model, we identified two major multicellular programs (MCP1 and MCP2) that describe the coordinated multicellular differences between non-failing and failing hearts (Figure 3B). MCP1 explained differences consistently across all studies (ANOVA adj. p-value < 0.05), while MCP2 separated conditions only in Reichart2022 and Koenig2022, indicating that our model is able to capture via the different programs both convergent and study-specific multicellular responses in HF (Supplementary Figure 6E). However, we also observed differences in the distribution of HF samples across MCP1 and MCP2 that were associated with the study of origin (ANOVA adj. p-value ≤ 0.001). This suggests that, despite the generalizability of the multicellular HF programs, their variability is still influenced by technical factors. By fitting linear mixed models to both MCPs using patient information as predictors and accounting for studies as random effects, we found no association between the distribution of HF patients across MCP1 and MCP2 and their organ sampling sites, acquisition mode, primary or secondary causes of HF, or etiology. Additionally, there was no correlation between the MCP1 and MCP2 scores of HF samples and the compositions of the seven major cell-types analyzed, indicating that the activation of these multicellular programs is independent of tissue composition.

The MCPs capture coordinated dysregulation of gene expression across cell types and can be interpreted at the gene level to identify conserved markers of cell-type-specific deregulation in HF (Figure 3C). As such, the MCPs provide a rich resource for generating hypotheses about stable disease markers, such as PLCE1 in CMs [28] or FKBP5 in Endothelial cells [29]. We observed that 46% and 49% of the genes associated with MCP1 and MCP2 (absolute weight > 0.1), respectively, were relevant to more than a

single cell type, indicating that HF transcriptional changes reflect a combination of both cell-type-specific and multicellular processes.

To better interpret the two-dimensional HF patient map spanned by MCP1 and MCP2, we enriched prior knowledge gene sets that reflect cellular and molecular functions, considering both cell-type-specific and multicellular perspectives (Figure 3D). The multicellular response captured by MCP1 was associated with pathways related to hypertrophy and fibrosis, including calcineurin [30,31], histone deacetylases [32,33], vasoactive intestinal polypeptide [34,35] and interferon-α and -γ signaling [36]. In contrast, MCP2 was linked to inflammation, with pathways involving TNF-α, TGF-β, TNF/stress responses, and apoptosis. Many cell-type-specific pathways were shared between MCP1 and MCP2. In cardiomyocytes, upregulated genes were related to sarcomere organization and z-disc morphology, while genes associated with oxidative phosphorylation and fatty acid metabolism were downregulated. In fibroblasts, collagens were upregulated in both MCPs, while genes in the RECK pathway (including TIMPs and MMPs) were downregulated, indicating increased extracellular matrix (ECM) production and decreased degradation. Endothelial cells showed upregulation of genes related to cell differentiation, with MCP1 particularly enriched for interferon-α and -γ genes, while lymphangiogenesis genes were downregulated. Pericytes downregulated their responses to nitrogen stress in MCP1, and myeloid cells showed a weak association with JAK-STAT signaling in MCP2. These results highlight key processes in HF, such as fibrosis, vascularization, and cardiac remodeling, reflecting expected biological pathways in cardiac tissue. While we observed general convergence and conservation of molecular profiles across HF samples from different studies, the molecular patient map built from multicellular programs revealed variable expression levels and combinations of cellular processes across major cell types.

2.4 Blueprint of multicellular processes and cell-cell communication in heart failure

Multicellular programs represent global patterns of cellular coordination, allowing us to extract insights into cell dependencies by examining co-expression and ligand-receptor interactions across cell types. MCP1 was associated with heart failure in a larger proportion of patients while capturing more of the transcriptional variance than MCP2 and was selected to investigate these cell dependencies in greater detail. For this, we inferred from MCP1 two cell-cell dependency networks with directed edges that describe to what extent the molecular profile of one cell-type could predict the profile of each other cell-type in both failing and non-failing hearts (Methods). We found a Spearman correlation of 0.68 (p-value = 1e-06) between the predictor importances of the two networks, suggesting that their overall organization is similar (Supplementary Figure 7A). However, the strength of the predictor importances

changed between failing and non-failing hearts (one sample t-test, adj. p-value < 0.05), with cardiomyocytes' predictors showing the largest shift in importance, and fibroblasts taking on a greater predictive role. The failing heart co-expression network showed strong connectivity within the vascular niche (involving endothelial cells, pericytes, vascular smooth muscle, and myeloid cells), while cardiomyocytes were primarily predicted by fibroblasts (Figure 3E).

To investigate potential mechanisms of cell-cell communication underlying the failing-heart co-expression network captured by MCP1, we inferred ligand-receptor interactions across cell types. We observed a modest correlation between the number of potential ligand-receptor pairs in a cell type pair and their predictive importance in the co-expression network (Spearman correlation = 0.33, p-value = 0.034), suggesting that a minimal set of communication interactions could drive multicellular coordination (Figure 3F, Supplementary Figure 7B). The highest number of ligand-receptor interactions in HF was observed between fibroblasts and cardiomyocytes, as well as between myeloid and endothelial cells, consistent with the results seen in the co-expression network. Given the strong coordination between fibroblasts and cardiomyocytes, the higher number of inferred ligand-receptor interactions, and the role of fibrosis in myocardial remodeling, we used NicheNet [37] to estimate the potential regulatory effect of ligands secreted by Fibs on the expression of HF-associated genes in CMs (Supplementary Figure 7C, D). Many ligand-receptor pairs were mediated through extracellular matrix (ECM) components, such as COL1A1, COL3A1, LAMA4, and MXRA5, supporting the idea that ECM remodeling influences CM transcription and phenotype, especially in response to stress, stiffness, or hypertrophy [38]. Additionally, ligands like BMP4 and NRG1 were identified as potential regulators of the CM stress response within MCP1 (e.g., ANKRD1, PROS1, ACKR3, CCND1) (Figure 3G). BMP4 is known to play a role in HF, particularly in cardiomyocyte transdifferentiation [39], while NRG1 has been implicated in cardiomyocyte division and migration [40]. These ligands could be key regulators of coordinated multicellular fibrotic processes in HF and hence potential modulators of myocardial remodeling. In summary, we reconstructed a multicellular blueprint of cell-cell interactions in HF, revealing the coordinated processes captured by MCP1 and their potential ligand-receptor mechanisms (Supplemental Table 3). Our analysis highlights the central coordinating role of fibroblasts in HF, particularly the increased communication between fibroblasts and cardiomyocytes, supported by both ligand-receptor interactions and predictive models.

2.5 Division of labor in fibroblast during heart failure

Fibroblasts are highly versatile cells that perform a wide range of biological functions, including immune modulation and extracellular matrix (ECM) remodeling. Due to their central role in regulating tissue remodeling, fibroblasts have become a major focus of therapeutic interest in disease contexts [41–44]. Since our analyses indicated that fibroblasts coordinate multicellular responses associated with myocardial remodeling and the vasculature niche, we then investigated how the distinct functions of fibroblasts are distributed across the population of single-cells to gain deeper insight into the multicellular coordination of HF processes of MCP1.

First, we defined common fibroblast cell-states by integrating 242,045 fibroblasts from core single-nucleus studies (Supplementary Figure 8A-G). We identified six conserved fibroblasts states via clustering (Figure 4A), with all patient samples contributing cells to each state (Supplementary Figure 8E). We characterized states by their expression of ECM components as well as pathway and cytokine activity signatures (Figure 4B, Supplementary Figure 8F-G, Supplementary Figure 9A-D). Briefly, Fib0 ($COL4A1^+$) expressed basement membrane components, suggesting a role in tissue homeostasis. Fib1 ($POSTN^+$ and $THBS4^+$), characterized by the expression of matrifibrocyte markers [45], was enriched for TGFβ signaling and core matrisome-related genes. Fib2 ($KAZN^+$) exhibited an immune-related expression profile, including TNFα and interleukin signaling, and expressed secreted ECM factors. Fib3 ($PCOLCE2^+$ and $SCARA5^+$) was associated with secretory factors and angiogenesis, along with signatures of activity of BMP4. Interestingly, Fib4 and Fib5 were the only states that did not strongly express ECM-related genes. Fib4 was characterized by Hedgehog signaling, IL4, and BMP6 signatures of activity, while Fib5 displayed signatures of cytokines such as TWEAK and IL2. By analyzing the compositional changes of fibroblast states between failing and non-failing hearts, we observed a consistent pattern across studies. Fibroblast states Fib1 and Fib4 expanded, while Fib0, Fib2, and Fib3 decreased in abundance in HF (linear mixed model p-value < 0.05, Figure 4C, Supplementary Figure 9E). Our single-cell integration of fibroblasts highlights the diverse functional roles of this cell-type in the heart and suggests that upon HF, distinct populations associated with an increased ECM production and inflammatory responses are favored.

We next investigated how the transcriptional response of fibroblasts in HF, as captured by their component in MCP1, related to the distinct fibroblast cell-states. We hypothesized that the multicellular division of fibroblast functions during HF could be regulated in two complementary ways: first, through

compositional regulation, where the population size of specific cell states adjusts to fulfill necessary functions in the tissue; and second, through molecular regulation, where global gene deregulation occurs across different cell states as a result of myocardial remodeling. To explore compositional regulation, we analyzed the genes captured in the fibroblast component of MCP1. We found that the enrichment of state-specific markers in MCP1 strongly correlated with the mean compositional changes of those states (Pearson's correlation = -0.98, $p = 10^{-16}$) (Figure 4D). This suggests that MCP1 reflects shifts in fibroblast state composition by representing markers in proportion to their abundance. Next, we examined the molecular regulation of genes in MCP1 by comparing the enrichment of MCP1 genes in pseudobulk expression profiles across patient samples and cell states. While we observed variability in enrichment scores between cell states, a consistent difference emerged between failing and non-failing tissues (Figure 4E). Linear mixed models of the enrichment scores revealed that MCP1 expression was more strongly associated with HF status (semi-partial R^2 0.90) than with cell state identity (semi-partial R^2 0.57). These findings characterize a phenotypic shift in fibroblasts toward an HF-specific program, as captured by MCP1. Although this shift involves an increase in the composition of Fib1 and Fib4, it is more accurately characterized by the broader acquisition of MCP1 across all fibroblast states, in line with the observations that fibroblast activation is a presumably continuous process and thus cannot be fully explained by the accretion of a state [43].

Finally, given the observed compositional and molecular regulation of fibroblast functions in HF, as captured by MCP1, we investigated the expression variability of individual genes across different cell-states and disease statuses. We applied linear mixed models to the pseudobulk expression of fibroblast states in each patient sample, using HF status and cell-state label as fixed effects, with studies treated as random effects. Each gene was then categorized into three expression groups—specialist, generalist, or acquired generalist—based on the partial explained variance attributed to cell-state identity and HF status (Methods, Figure 4F). This classification highlighted whether a gene's variability was driven primarily by cell-state differences (e.g., CDK8), HF (e.g., COL24A1), or both (e.g., FGF14) (Figure 4G, Supplemental Table 5). Notably, established matrifibrocyte (Fib1) markers, such as POSTN, THBS4, and CILP, were regulated in both cell-state and disease contexts, suggesting that matrifibrocyte characteristics are broadly acquired across fibroblast states (Supplemental Figure 9F). Enrichment of the distinct expression groups in MCP1 and the HF bulk reference indicated a higher enrichment of genes with acquired generalist regulation (Supplementary Note 4, Supplementary Figure 9G-H). Altogether, our results suggest that the expression of MCP1 fibroblast component genes is distributed differently across fibroblast states, with some genes acting as specialists in non-failing fibroblasts but becoming broadly

expressed across states in HF. These acquired generalist genes represent the prioritized program in both multicellular responses and bulk profiling.

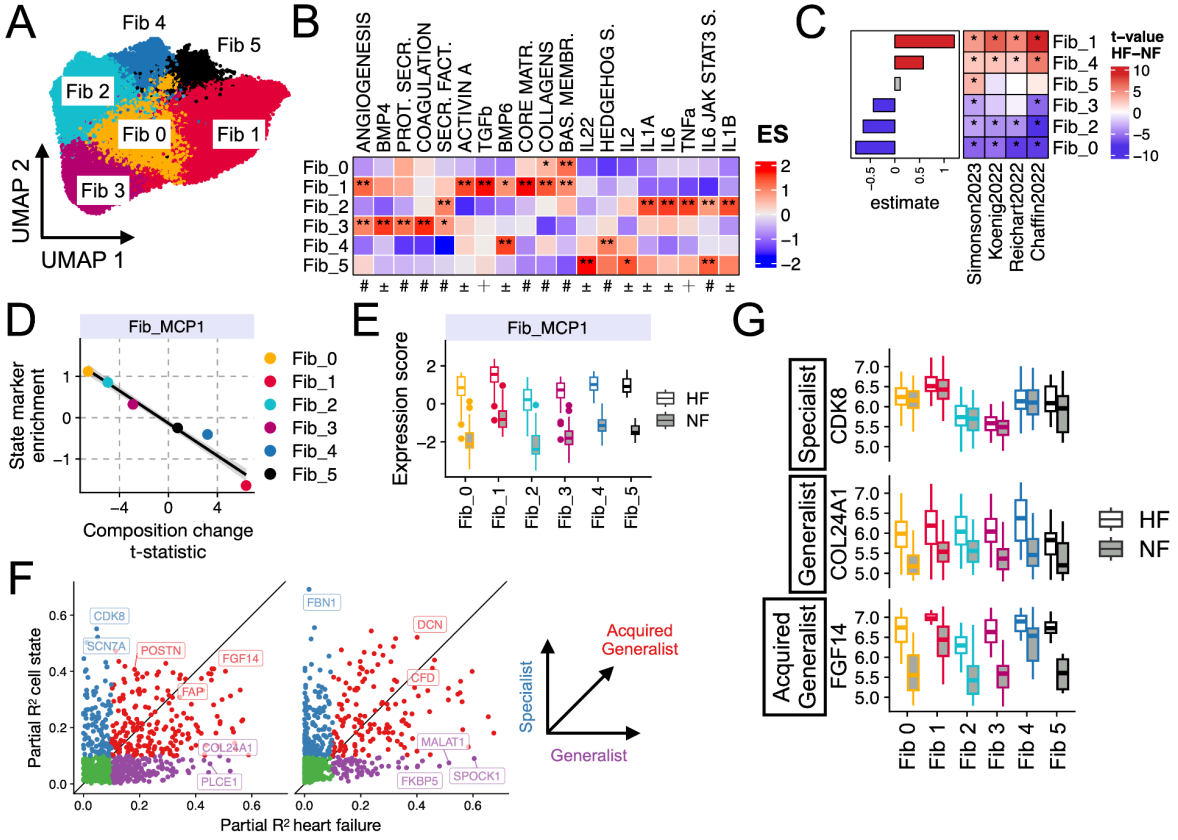


Figure 4. Exploring the division of labor in fibroblast activation in HF

A. Uniform Manifold Approximation and Projection (UMAP) of 242,045 fibroblasts integrated from the four core HF studies. Cells are colored by clusters.

B. Mean enrichment scores of prior knowledge gene sets in pseudobulked states. Star indicate adj. p-value <0.01. The symbols represent the database of origin: # = MSigDB, ± = Cytosig, + = PROGENy.

C. Compositional changes of fibroblast states (y-axis) between control and HF patients from different studies (x-axis) assessed by t-statistics. Side bar plots display the fixed effect estimates from a linear mixed model. Stars indicate adj. p-value <0.05.

D. Comparison of the mean compositional change (t-statistic, x-axis) with the enrichment of the top 200 state markers in the fibroblasts component of MCP1.

E. Enrichment scores of the fibroblast component of MCP1 in pseudobulked patient profiles per fibroblast states (color) separated by failing (HF) and non-failing (NF) hearts.

F. Modelling individual gene's expression with HF and cell state covariates with linear mixed models and comparing the partial R^2 values for each covariate to characterize gene's expression pattern.

G. Normalized expression of pseudobulked patient profiles per state for selected representative markers of different division of labor programs: specialist (CDK8), acquired generalist (FGF14), and generalist (COL24A1).

2.6 Conservation of multicellular responses in heart failure in bulk transcriptomics

MCP1 was derived from 132 patients across four core single-nucleus studies and captured a shared multicellular transcriptional variation that could be functionally interpreted through a blueprint of cell dependencies, providing insights into fibroblast state coordination. However, the limited size of the patient cohorts could limit its generalizability to a broader patient population. To address this, we investigated whether the MCPs could be detected in our compiled larger bulk HF cohort approximately 10 times the size of the SN collection. We enriched bulk transcriptomics samples from each study in our core collection with MCP1 and MCP2 cell-type signatures and compared their ability to distinguish failing from non-failing patients (Figure 5A). MCP1 showed stronger performance, with a median AUROC of 0.84 and a median silhouette score of 0.4 across patient groups, while MCP2 had a median AUROC of 0.64 and a median silhouette score of 0.04, indicating broader applicability of MCP1 across studies and technologies.

Based on these results, we hypothesized that the molecular information captured by the core HF bulk data reflects both cell-type composition changes and multicellular gene expression responses of MCP1 within the tissue. Specifically, changes in gene expression from bulk data in HF could be driven by shifts in the abundance of cell types that express the gene (compositional regulation), transcriptional regulation that occurs independently of cell-type composition (molecular regulation, either within a single cell type or across multiple cell types), or a combination of both mechanisms. To identify which process drives the consensus transcriptional bulk signature of HF, we correlated the consensus effect size of differentially expressed genes in bulk data with compositional and molecular deregulation scores derived from core single-cell studies (Methods). For the top 8,942 genes in the bulk consensus signature (adj. Fisher p-value < 0.05), we observed a Spearman correlation of 0.14 with compositional scores and 0.6 with molecular scores, indicating that multicellular transcriptional coordination, captured by MCP1, is the dominant factor influencing gene expression changes in HF bulk transcriptomics (Figure 5C).

We annotated each of the top 8,942 genes in the bulk consensus signature (adj. Fisher p-value $<$ 0.05) as potentially deregulated by compositional, molecular, or a combination of both mechanisms. This classification was based on the agreement between their differential expression size-effect in bulk and the molecular and compositional scores derived from single-nucleus data (Figure 5C). For 29% of the genes, we were unable to assign a mechanism due to incongruences in the direction of regulation or limitations in single-nucleus gene coverage. However, 5% of the genes were linked to compositional regulation, 56%

to molecular regulation, and 10% to both, underscoring that bulk transcriptional changes in HF cannot be reduced to shifts in tissue composition alone (Figure 5D).

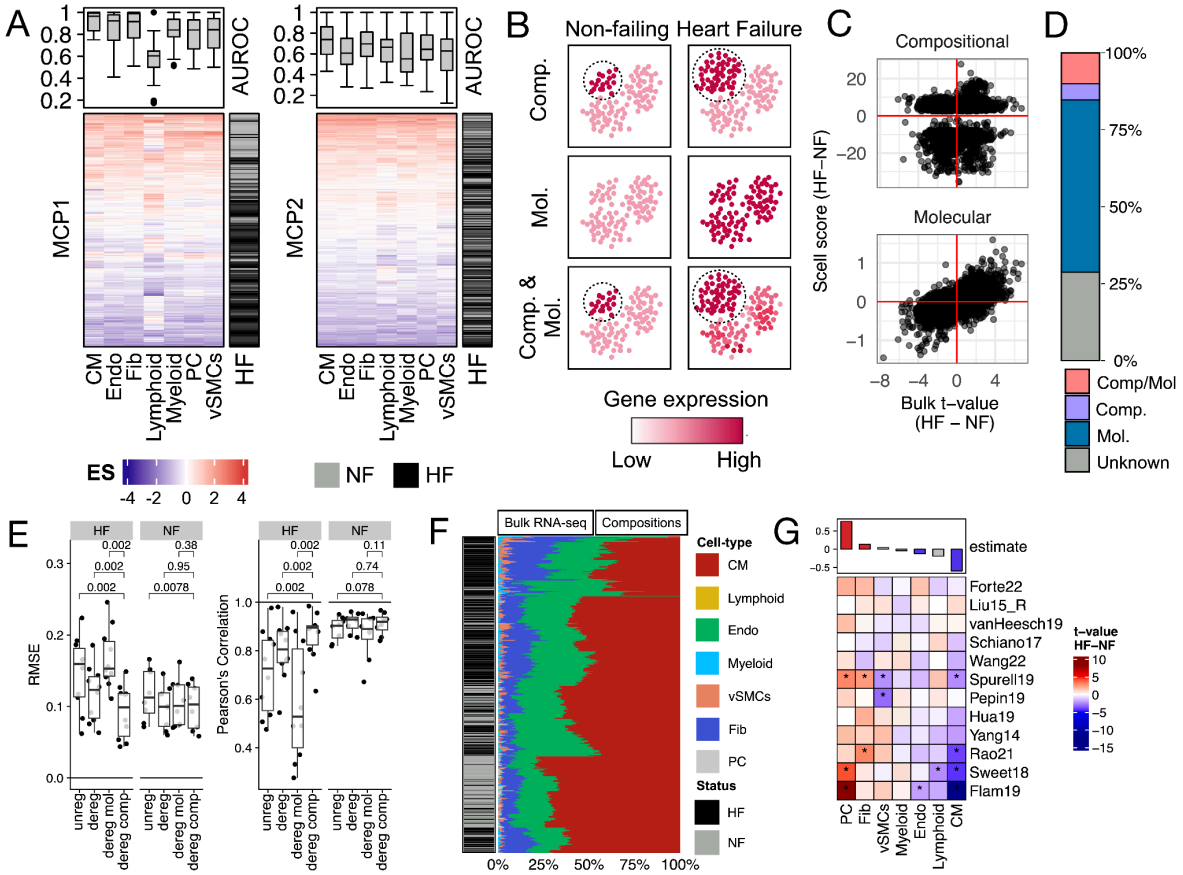


Figure 5. Integrating HF multicellular programs (MCPs) with bulk transcriptomics

A. Cell-type specific transcriptional processes associated with MCP1 and MCP2 enriched in every patient tissue sample from the bulk core study collection. Upper panels represent the distribution of areas under the receiver operating characteristic curve (AUROC) used to evaluate the classification performance of cell-type specific responses in classifying non-failing hearts in each study independently.

B. Schematic representing our intuition behind potential tissue-level gene regulatory mechanisms that explain observed changes in gene expression in bulk transcriptomics. We assumed that an increase in expression of a given gene in HF captured in bulk transcriptomics, could be associated with the increase in composition of a cell-type that specifically express this gene (Compositional regulation (Comp.)), with a generalized increased expression of this genes across cells (Molecular (Mol.)), or a combination of both (Comp. & Mol.). By extension, gene expression downregulation follows similar principles.

C. Consensus gene-level statistics of changes in gene expression upon HF from bulk transcriptomics (x-axis) and consensus gene-level scores (y-axis) of compositional (upper) and molecular regulation (lower) from single-nucleus transcriptomics. We show the collection of genes with available information from the top 8,942 genes of the consensus bulk ranking (adj. Fisher p-value < 0.05) (Figure 2C).

D. Annotation of events of gene expression deregulation from the combination of bulk and single-nucleus transcriptomics studies. The annotation comes from the top 8,942 genes of the consensus ranking (adj. Fisher p-value < 0.05).

E. Root mean square error and Pearson correlation of the deconvolution of pseudobulk expression profiles built from the collection of core and supporting single-nucleus studies in contrast to the known proportions. Tests were separated by type of tissue that was deconvoluted, failing (HF) and non-failing (NF) hearts, and

genes used in the reference for deconvolution. Unreg, unregulated markers; dereg, deregulated markers; dereg mol, molecularly deregulated markers; dereg comp, compositionally deregulated markers.

F. Hierarchical clustering of bulk-transcriptomics tissue samples profiled with RNAseq, based on the deconvoluted composition using a healthy reference SN data set subset to compositionally regulated genes and the seven major cell-types used in our ontology.

G. T-statistics (heatmap) and estimate of difference between failing and non-failing hearts from the differential compositional analysis using t-tests and linear mixed-models, respectively. Stars denote an adj. p-value < 0.05

In all panels cardiomyocytes (CM), fibroblasts (Fib), pericytes (PC), and endothelial (Endo), vascular smooth muscle (vSMCs) cells. Heart failure (HF), and non-failing (NF) hearts.

2.7 Cell type composition estimation in bulk transcriptomics

The multicellular transcriptional responses during HF captured by MCP1 were consistently observed in patient cohorts profiled using both single-nucleus and bulk RNA-seq studies. However, it remained unclear whether the changes in cell type composition observed in single-nucleus studies during HF could be generalized to larger bulk cohorts. To address this, we estimated cell-type compositions from bulk expression data in the core bulk collection using deconvolution methods. These methods infer cell-type proportions by learning the relationship between cell-type marker expression and cell-type abundance from single-cell data. To achieve reliable performance in HF data, it is crucial to properly define cell-type markers and account for their expression changes in HF.

We first noted that 33-47% of cell-type markers identified from our core single-nucleus studies (i.e., genes whose expression is characteristic of a cell type) were deregulated in HF (Supplementary Figure 10A-B). In the previous section, we found that many deregulated genes could not be explained solely by changes in cell type composition. For example, NPPA, a cardiomyocyte marker, is upregulated in failing heart tissue despite a decrease in cardiomyocyte abundance, indicating that NPPA is regulated more by molecular mechanisms than compositional ones. We hypothesized that molecularly regulated cell-type markers like NPPA could reduce deconvolution accuracy, as these markers would not reliably reflect true cell type composition. To test this, we evaluated different sets of cell-type markers in estimating cell-type compositions from ten single-cell studies aggregated into pseudobulks with known compositions (Supplementary Figure 10A,C; Supplementary Note 5). Our benchmarks confirmed that molecularly regulated markers performed poorly as indicators of cell-type composition in disease. Moreover, using only compositionally regulated markers significantly improved deconvolution results in HF samples (Figure 5E).

Using compositionally regulated cell-type markers, we estimated cell-type compositions across all 697 patient tissues from the subset of 12 RNA-seq studies in the core bulk collection. We found that

cardiomyocytes followed by endothelial cells and fibroblasts were the most abundant cell types (Figure 5F). However, cell-type compositions only modestly grouped failing and non-failing patients (median silhouette width HF = 0.19, NF = 0.4, one-sample t-test, p-values: HF < $10^{-}8$, NF < $10^{-}39$) (Supplementary Figure 10F). Differential compositional analysis between failing and non-failing patients showed little agreement on cell-type changes between studies (Figure 5G). There was a modest trend showing a decrease in cardiomyocytes and endothelial cells, and an increase in pericytes and fibroblasts (linear mixed model adjusted p-value < 0.05). However, variability between studies was high, particularly in cardiomyocytes, fibroblasts, and endothelial cells (Supplementary Figure 10G).

These findings did not align well with the composition changes observed in single-cell data, except for a shared decrease in cardiomyocytes. When comparing single-cell and bulk compositions, we found that 80% of the variance was associated with data modality and study labels, while only 53% was associated with HF status based on principal component analysis (Supplementary Figure 10H). In summary, the agreement between bulk studies on cell type composition changes was modest, and the compositional changes did not match those from single-nucleus studies. These discrepancies might be due to technical factors, such as cell isolation protocols or errors in bulk deconvolution. On the other hand, we observed a strong agreement between MCP1 and the bulk cohort. This difference between molecular and compositional changes suggests that the observed molecular shift is primarily driven by shared multicellular coordination within the tissue, independent of cell type composition.

2.8 Reinterpreting independent data with the multicellular program of heart failure

Finally, to illustrate the potential of the robust multicellular programs in contextualizing and integrating independent datasets, we projected supporting single-nucleus and single-cell studies onto our patient map (Figure 6A-B). We confirmed the generalizability of the HF processes described by MCP1 and MCP2, with perfect classification of 12 tissue samples from HF patients across hypertrophic cardiomyopathy (HCM), ischemic cardiomyopathy (ICM), and cardiac sarcoidosis etiologies, drawn from two studies without healthy reference patients (Figure 6B). Projection of myocardial infarction and congenital heart disease patients from two additional studies highlighted the activation of apoptotic processes related to MCP2 (Supplementary Figure 11A-B, Supplementary Note 6). Additionally, we observed little agreement between the multicellular responses in mouse models of heart disease and those captured by the core SN HF studies (Supplementary Figure 11C-D, Supplementary Note 6). By projecting these supporting datasets onto the multicellular patient map of HF, we aimed to describe an analysis

strategy that allows for the comparison of HF studies across diverse etiologies, biological contexts, and species.

We used two supporting studies to examine myocardial remodeling and its reversal, specifically the reactivation of fetal programs and cardiac recovery following left ventricular assist device (LVAD) implantation. Tissue samples from fetal hearts, pediatric dilated cardiomyopathy patients, and non-failing hearts showed differences in MCP1 activation consistent with HF progression (ANOVA, p-value = 0.04), with the strongest difference observed between fetal and non-failing hearts (t-test, adj. p-value = 0.047; Figure 6C). These findings link the generalist multicellular program of HF to the reactivation of fetal myocardial programs, a hallmark of the disease. Direct comparison of gene activation in fetal versus failing hearts revealed shared patterns in cardiomyocytes, including reduced activity of PPAR-α and PPAR-δ footprints and decreased fatty acid oxidation, alongside increased BMP6 footprints. In fibroblasts, we observed increased Hedgehog signaling, TGFB1, IL4, and GLI3 footprints, and enhanced collagen fibril organization (Figure 6D), providing detailed insights into fetal reprogramming captured by MCP1. Heart tissue samples from recovered and non-recovered dilated cardiomyopathy (DCM) patients, before and after LVAD implantation, also aligned with MCP1 activation (Figure 6E). Among recovered patients, we observed a significant change in MCP1 activation after LVAD implantation, where post-implantation samples resembled those from non-failing hearts (paired t-test, adj. p-value = 0.003). Pre-implantation samples also showed differences in MCP1 activation between recovered and non-recovered patients (t-test, adj. p-value = 0.012), indicating that this multicellular program reflects disease severity and can offer insights into recovery potential. Comparison of gene expression between recovered and non-recovered patients before implantation revealed alignment with MCP1 primarily in endothelial cells, fibroblasts, and myeloid cells. In all three cell types, we observed coordinated downregulation of inflammatory responses, including cytokines (IL17a, IL2, IL15, TNFα) and growth factors (EGF) (Figure 6F). These results suggest that HF patients with lower levels of inflammation exhibit greater recovery benefits from LVAD implantation.

In conclusion, the robust multicellular programs captured in this study, particularly MCP1, provide a powerful framework for understanding the molecular mechanisms underlying HF, with broad applicability across various HF etiologies, biological contexts, and species. By enabling the integration and comparison of independent studies, this multicellular reference map offers a valuable tool for clinical applications, such as identifying patients who may benefit most from targeted interventions, predicting disease trajectories, or evaluating animal models of disease.

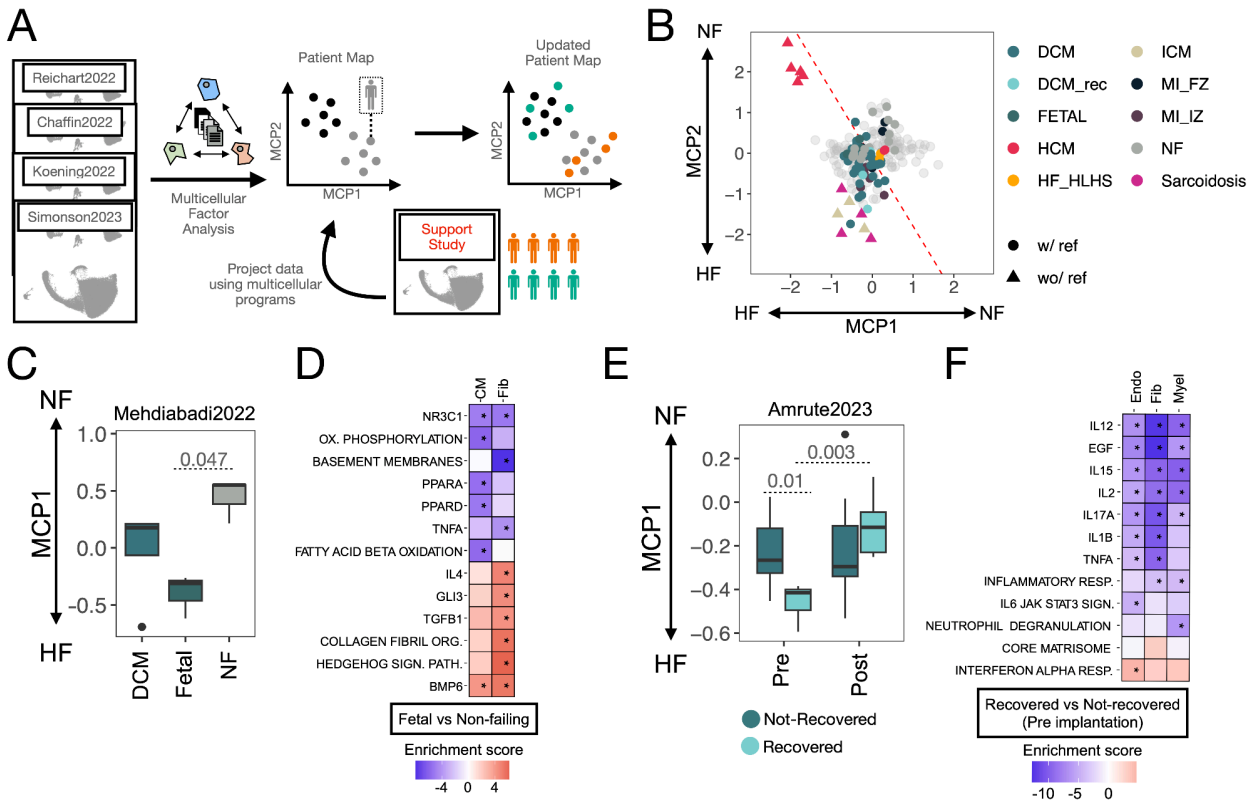


Figure 6. Reinterpretation of patient map using supporting external studies

A. The patient map built with the single-nucleus (SN) core studies can be used for the reanalysis of datasets of similar contexts. One possibility is to project the tissue samples of an independent dataset into the latent space defined by multicellular programs (MCPs) describing the variability of tissue samples from the SN core studies. The values of activation of the different MCPs then can be evaluated for tissue samples of the projected studies.

B. Tissue samples of supporting studies projected into the patient map of HF built from the SN-core-studies (light grey). Color of the points denote distinct etiologies, and shape of the points if the study did or did not contain or not reference control samples. Dotted line represents a decision boundary for the classification of HF samples. coming from a logistic regression model fitted in the SN-core-study collection (see Methods).

C. MCP1 values of projected heart tissue samples of dilated cardiomyopathy pediatric patients, fetal hearts, and non-failing controls. Adjusted p-value of t-test between fetal and non-failing samples.

D. Enrichment scores of differentially active or inactive processes in fetal hearts over non-failing donor hearts captured by MCP1.

E. MCP1 values of projected paired HF samples obtained from dilated cardiomyopathy patients before (pre) and after (post) implantation of a left ventricular assist device. Colors denote the response to treatment. Adjusted p-values come from a t-test of pre samples, and a paired t-test of recovered samples pre and post implantation.

F. Enrichment scores of differentially active or inactive processes in recovered patients versus not recovered captured by MCP1.

3. Discussion

The molecular processes driving HF have been extensively studied through large-scale transcriptomics, both at bulk and single-cell resolution. However, these studies often present a fragmented and sometimes inconsistent view of transcriptional regulation, with variability arising from differences in patient cohorts and technologies. To fully benefit from these molecular profiling efforts and create a meaningful clinical impact, an integrative approach is needed to identify robust, consistent disease processes in HF. This integration must overcome two main challenges: first, the harmonization of diverse data acquisition protocols, storage formats, and accessibility to ensure effective use by both computational and clinical scientists; and second, the use of statistical methods that account for the multicellular nature of tissues, enabling joint analysis of bulk and single-cell data while supporting the integration of new datasets and technologies. In this study, we addressed these challenges by combining single-nucleus transcriptomics with bulk transcriptomics from larger patient cohorts to build a consensus of multicellular gene expression changes during cardiac left ventricular remodeling. This expanded upon our previous reference of the HF transcriptome [16] and uncovered reproducible molecular insights from multiple independent studies. Rather than focusing solely on cell-type taxonomies [46], as is common in single-cell analyses, our work captures multicellular programs (MCPs) that describe coordinated gene expression across multiple cell types. This approach shifts attention to patient-level variability and tissue-wide molecular coordination, offering new insights into how HF affects the heart as an integrated system.

Our comprehensive comparison of single-nucleus and bulk transcriptomics datasets revealed a strong consistency in gene expression responses during HF, despite variations in technical and clinical factors across studies. The addition of new bulk transcriptomics data allowed us to refine the ranking of genes most closely associated with HF, reinforcing our previous findings that coordinated molecular responses during end-stage HF are conserved across patient cohorts [16]. By analyzing MCPs derived from single-nucleus data, we were able to break down these molecular responses into cell-type-specific programs, which were also detectable in bulk transcriptomics, further confirming the robustness of these convergent molecular responses.

Inference of cellular dependencies from MCPs provided new insights into previously unexplored aspects of gene regulation in HF. Analysis of MCP1, the largest axis of conserved variation, showed that cells in both failing and non-failing hearts follow similar coordination patterns across cell types. Fibroblasts emerged as central to MCP1, having the strongest influence on the gene expression of other

cell types, particularly cardiomyocytes. These dependencies suggest that when fibroblasts express their specific MCP1 program, other cell types, such as cardiomyocytes, respond by expressing their corresponding set of genes. However, we observed a weak correlation between ligand-receptor coexpression and gene expression coordination, indicating that cellular coordination in HF may extend beyond direct cell-to-cell communication through ligand-receptor signaling [47–49] and reflecting the limitations of studying ligand-receptor interactions with transcriptomics [50]. Further investigation into cardiomyocyte-fibroblast dependencies identified ligand-receptor pairs associated with ECM components, reflecting how ECM changes can drive transcriptional, morphological, and functional shifts in cardiomyocytes during HF [38,51]. In addition, we found that these ligands could represent a fraction of potential inducers of the stress response of cardiomyocytes, traced in the expression of genes such as ANKRD1 or CCND1. Our findings suggest that this novel multicellular representation of molecular processes in HF provides a valuable framework for studying disease at the tissue level and could guide the identification of therapeutic targets that stabilize multicellular coordination and maintain tissue homeostasis.

Multicellularity allows tissues to perform functions that individual cells cannot achieve alone, with a key strategy being the division of labor, where tasks are distributed among cells. This can range from highly specialized cells performing specific tasks to generalist cells covering a broad range of functions. Previous studies have shown that tissue function typically arises from a balance of specialist and generalist cells, which can dynamically shift roles depending on the context, such as in disease [52 53–55]. By comparing gene expression across cell types within MCP1, we found that most dysregulated genes were cell-type specific, indicating that cell lineages act as specialists. However, when analyzing fibroblast function in greater detail, this pattern shifted. MCP1 showed a broad activation across all fibroblast states. Upon further decomposition of this Fib component of MCP1 into generalist or specialist expression patterns, we found that matrifibrocyte markers formed an acquired generalist program, as all fibroblast states upregulated these markers in HF. This suggests that generalist gene programs, rather than highly specialized ones, drive fibroblast adaptation to HF, which was confirmed by their high ranking in the consensus bulk signature. Additionally, we identified a purely generalist program that included less studied genes, such as COL24A1, previously linked to osteoblast differentiation [56], but with an unknown role in cardiac function. While fibroblasts are expected to reallocate resources to meet tissue demands, different regulatory strategies likely drive the expression of these gene groups through distinct mechanisms. These findings improve our understanding of tissue dynamics in cardiac fibrosis and may help identify biomarkers for diagnostic or therapeutic use by prioritizing robust markers that generalize to

the broader HF population. However, this analysis is limited by the resolution of the defined cell states and the integration strategy used, and further work at the single-cell level, along with experimental validation, is needed to fully explore these regulatory patterns.

The integration of bulk and single-nucleus data, in addition, opened the opportunity to study potential tissue-level mechanisms of deregulation of genes consistently associated with HF. Particularly, we were interested in understanding whether the source of gene deregulation in HF was linked to changes in tissue composition (i.e., an increase or decrease in specific cell types) or molecular mechanisms independent of cell-type composition. We found that conserved disease programs primarily operated independently of tissue composition, suggesting that the multicellular coordination events during cardiac remodeling represent a broader disease mechanism affecting the heart at a systemic level. For example, the cardiomyocyte marker NPPA, a recognized HF biomarker, showed increased expression despite a reduction in the number of cardiomyocytes. This highlights how gene expression changes can occur through mechanisms beyond shifts in cell type abundance. Moreover, identifying the source of expression changes not only provides biological insights but also improves the accuracy of computational methods for estimating cell-type proportions in HF from bulk transcriptomics. Our ability to trace multicellular disease processes inferred from single-nucleus data in bulk transcriptomics supports the notion that these molecular processes can occur independently of local tissue composition. This is further evidenced by the convergence of molecular signals across tissue samples collected from distinct locations in the left ventricle of different patients. However, it is important to acknowledge that cell-type compositions in these analyses were approximated from bulk and single-nucleus data, and may be subject to biases.

The integration of transcriptomics data across scales enabled us to build a patient map that captures heart tissue variability across cell types, based on the activation of two distinct multicellular programs (MCP1 and MCP2). Interpreting this variability with clinical covariates could help identify molecular markers of heart disease. We showed that MCP1 and MCP2 successfully distinguish failing from non-failing hearts across core studies. Although the limited clinical data in these studies did not reveal strong associations with covariates such as LVEF, age, sex, or BMI, we propose that the patient map has translational potential as a reference for analyzing single-cell data from related HF contexts. By projecting independent datasets onto MCP1 and MCP2, we facilitated the reinterpretation of studies on fetal samples, myocardial infarction, animal models, and LVAD-treated patients. For example, MCP1 separated fetal from control samples by identifying shared deregulation of genes, such as BMP6 footprints in cardiomyocytes and hedgehog signaling in fibroblasts. Early cardiac remodeling after infarction, however, aligned more with MCP2. Our findings also indicated a lack of concordance between the

multicellular disease processes of HF in human tissue with the ones observed in commonly used mice models (TAC, Angiotensin II). This could be linked to the difficulties in determining disease time-points in mice that are aligned with end-stage HF patients, the lack of large animal cohorts, the larger amount of tissue profiled in mice compared to human tissue or species differences in pathway activation. Importantly, the decision to implant a left ventricular assist device (LVAD) is complex and could be facilitated from more accurate predictions of recovery success. While studies have used clinical data to predict risk and benefit, molecular data provides another layer of insight. While the patient cohort in the study by Amrute et al. [6] was clinically indistinguishable, their recovery trajectory was captured by a regression on MCP1, suggesting a molecular profile that could predict recovery pre-implant. These findings indicate that the molecular state of a patient, as described by her position within the HF patient map, may provide clinically meaningful information beyond conventional clinical metrics like LVEF or demographic factors. In this regard, molecular phenogrouping studies have indicated that molecular data can correlate with distinct clinical subgroups [57,58], an approach of great interest for personalizing treatment for HF patients. Our study aimed at identifying and quantifying a shared signal, thus this molecular heterogeneity between individuals could be disregarded as statistical noise. We propose that these approaches, capturing both consensus and heterogeneity, are not mutually exclusive. Rather, the consensus derived from our study is more likely to generalize to the HF patient population and can enhance precision medicine approaches, by capturing HF patient heterogeneity along the identified axes of multicellular programs, as demonstrated by the molecular reinterpretation of myocardial infarction, fetal reprogramming, and LVAD response.

Through our curation efforts, we also gathered insights into the clinical spectrum of heart failure patients studied over recent decades, revealing a predominance of Caucasian males with end-stage disease. This demographic bias may limit the generalizability of findings to a broader HF population—particularly patients of diverse geographic origins, female sex, those with HF with preserved ejection fraction, early-stage HF, or HF arising from etiologies such as inflammatory or infiltrative cardiomyopathies. Additionally, our work highlighted the critical need for sharing comprehensive metadata to support clinically relevant reinterpretation of the data. Addressing these two challenges is essential for advancing inclusivity and clinical applicability in the field.

Overall, our study provides insights into the conserved mechanisms of HF remodeling and multicellular organization supported by the largest collection of heart-specific multi-scale transcriptomics to date. We anticipate that this resource will serve as a valuable reference for understanding the coordinated transcriptomic changes in end-stage HF. To facilitate its use by both computational and

clinical scientists, we have made this resource available in various formats through an accessible platform (https://saezlab.shinyapps.io/reheat2/), enabling data-driven translational research.

4. Methods

4.1 Study Inclusion Criteria of single-nucleus studies

We collected human HF single-nucleus transcriptomic (snRNA-seq) studies reported in literature with available gene expression count matrices in public repositories. We included studies whose experimental design consisted in profiling biopsies of human heart samples of the left ventricle with free wall or apex sublocations from patients with end-stage HF and patients with non failing hearts. In addition, we prioritized studies that profiled at least 10 patients per group. Protocol type and links to source data are presented in Table 1 and Supplementary Table 1.

Single-nucleus studies that did not fit the inclusion criteria because of the lack of independent profiling of patients with non failing hearts, acute HF, and pediatric or fetal samples were analyzed independently. In addition, two studies profiling established models of HF in mice (Angiotensin II-Induced (AngII) and transverse aortic constriction (TAC)) were curated to evaluate its consistency with human data. Information of these supporting studies are provided in Table 1 and Supplementary Table 1.

4.2 Processing of single-nucleus studies

Expression count matrices of the selected studies were downloaded directly from their specific repository links. Studies originally provided as R objects were transformed into anndata objects using Zellkonverter. Studies providing count matrices with Ensembl IDs (Reichart2022, Kuppe2022) were transformed into gene symbols using biomaRt v2.58.0 [59] and summing all reads of Ensembl IDs assigned to the same symbol. Across studies we filtered out samples belonging to right ventricles if available, and the unit of each study was considered to be the patient in case multiple samples were collected, as in Reichart2022. No further processing was performed to studies, since the original authors provided processed data. To ensure the comparability of the analysis across atlases, we defined a heart cell ontology that included the following cell types: Cardiomyocytes (CMs), fibroblasts (Fibs), endothelial cells (Endos), pericytes (PCs), vascular smooth muscle cells (vSMCs), myeloid and lymphoid cells. Regular expressions over original cell annotations provided by the selected studies were performed to align each dataset to our proposed ontology. Unannotated cells were discarded. Single-nucleus studies were transformed into collections of pseudobulk expression profiles by summing up the counts of all cells belonging to each of

the cell types defined in our ontology for each sample. Quality metrics including the number of cells used, the total number of reads, and number of genes with available counts were calculated for each pseudobulk sample.

To estimate the agreement between the new assigned cell annotations, we compared the overlap of marker genes of the seven cell-types estimated from each study independently using Jaccard Indexes. Marker genes of each cell type per study were calculated using differential expression analysis between the pseudobulk expression profiles across samples of one cell-type versus the rest using edgeR v4.0.2 with default parameters of gene filtering. Genes with a log-fold-change over 2 and a false discovery rate (FDR) lower than 0.01 were considered cell-type markers.

To estimate the levels of noise expression values within the pseudobulk expression profiles across studies, we defined a score that quantified the amount of contamination either by the misannotation of single cells or by background expression. Our contamination score was defined for each pseudobulk expression profile as the ratio of the number of reads belonging to contaminating genes to those belonging to marker genes. Given a pseudobulk expression profile belonging to a cell type, we defined as contaminating genes all the marker genes of the rest of the cell types. In case that a marker gene was expressed in more than a single cell type, the reads of the gene were always assigned to the marker gene set of the cell type tested. The gene set of marker and contaminating genes was defined for each study independently.

4.3 Processing of supporting studies

The seven supporting studies (Table1) were processed depending on available data formats. For studies where only cell ranger output files were available (Liu2022, Mehdiabadi2022, McLellan2020), we used a uniform processing pipeline with the Seurat v5.0.3 R-package [60]. This included a filtering procedure per sample, doublet detection with DoubletFinder [61], cell filtering (mitochondrial percentage <0.1 in single-nucleus and <0.2 in single cell, genes per cell >300 , RNA counts per cell >500, ribosomal genes per cell <0.1) in samples with at least 10 cells and 250 features. Samples were integrated on joint highly variable genes with Harmony v1.2.0 R-package [62]. Leiden clustering was applied to identify cell lineages and overrepresentation analysis of consensus cell type markers from core HF studies was used to annotate clusters; cell clusters that did not match the seven cell lineages were discarded. For studies with available count matrices and cell lineage annotations (Hill2022, Nicin2022, Ren2020, Litviňuková2020, Kuppe2022), we mapped annotations to our harmonized vocabulary. All studies were then summed to pseudobulk profiles and normalized via edgeR v3.36.0.

4.4. Processing and analysis of core HF bulk studies

For the addition of the five bulk studies we applied a similar strategy as reported previously [16]. In brief, we identified five studies that fulfilled inclusion criteria for bulk studies. We downloaded sequencing files in FASTQ format from Wang22 and Forte22 and realigned data to the human genome as implemented in the ArchS4 pipeline in biojupies[63]. Raw counts were downloaded for Flam19, Rao21, Hua19 from repositories (Supplemental Table 1). For all studies meta data was aligned to a shared vocabulary. Low expressed genes were filtered and counts were normalized with TMM (edgeR v3.36.0), log scaled, and voom-transformed (limma v3.50.3). Differential gene expression analysis, study comparison, classification and gene expression meta analysis was performed as described [16]. The updated consensus was compared to the previous ranking by jaccard indices. For a given ranking segment, genes were selected from both rankings and tested for their classification performance and their enrichment scores across studies[16].

4.5 Definition of consensus cell-type gene markers from single-cell data

For each cell-type, we combined the FDR of the differential expression analysis for all genes that were measured in at least 3 studies using a Fisher's combined probability test with survcomp v1.5 [64]. The degrees of freedom for the significance test of each gene were defined by the number of data sets that included it. A ranking of marker genes was generated by the corrected p-value of the combined test and the mean log fold change across studies. Correction was performed with the Benjamini-Hochberg (BH) procedure.

4.6 Compositional data analysis of core single-cell studies

Unsupervised analysis of cell type compositions was done using hierarchical clustering over the Euclidean distance matrix of all samples across studies. General and study-specific HF compositional signatures were generated using differential compositional analysis. First, cell-type compositions of each sample were transformed into centered-log-ratios (clr) using *compositions v2.0-8* [65]. Then, for each study and cell-type, a t-test comparing the means of clr values of failing and non-failing patient samples was performed. To test the ability of study-specific compositional disease signatures to classify samples of other studies, we defined a compositional disease score. Following our previously defined disease score, we linearly combined the scaled clr values of the samples of one study with the compositional disease signature of the rest, captured by the t-values of their compositional differential analysis. Areas under the receiver operating characteristic curve (AUROC) were used to test the classification accuracy, where the

non-failing class was defined as the response variable, and to measure the conservation of HF compositional changes. General HF compositional signatures across studies were estimated by modeling the difference between the means of clr values between failing and non-failing hearts using linear mixed models. For each cell-type a model was fitted assuming random effects from studies. Across tests, p-values were corrected using the BH procedure.

4.7 Multicellular factor analysis of individual studies

We summarized the molecular variability across cell-types and patient samples for each study in terms of multicellular programs (MCPs). MCPs are latent variables that capture a certain fraction of gene expression variability across samples and can be calculated from the collection of pseudobulk expression matrices of each cell type for each study. For a given MCP, each gene of each cell-type gets assigned a weight that represents its importance in defining the MCP. At the same time, the activation values of a MCP across samples define a range of molecular phenotypic variability that can be associated with patient annotations to facilitate interpretability. Moreover, MCPs capture fractions of the total variance of the dataset and these are specific for each cell type. Finally, the latent space formed by MCPs define a patient map where new data can be projected.

We estimated 10 MCPs for each individual study with Multicellular Factor Analysis [27] using MOFA2 v1.12 [66,67]. In each factor analysis model we included only pseudobulk expression profiles calculated from at least 20 cells and cell types profiled in at least 40% of the patient samples with more than 50 genes. Genes with less than a minimum of 20 counts in a single sample or detected in less than 40% of the remaining samples were discarded. Each filtered pseudobulk matrix was normalized using the trimmed-mean of M values method in edgeR v4.0.2 with a scale factor of 1 million and log-transformed. In each pseudobulk expression matrix of each cell type, we filtered out marker genes of each other cell type to reduce the levels of contamination. Finally, we filtered out samples within each pseudobulk expression matrix of each cell type with less than 97% of the genes measured to avoid MCPs related purely to coverage. Feature-wise sparsity was not included in the model to obtain a greater number of genes associated with MCPs. We associated the MCPs values with clinical covariates using Analysis of Variance (ANOVAs) or linear models for categorical and continuous covariates, respectively (corrected p-value lower or equal to 0.05). Across tests, p-values were corrected using the BH procedure. To quantify the total amount of explained variance (R2) associated with a clinical covariate, we summed the R2 of each MCP associated with that covariate.

4.8 Comparison of multicellular programs across studies

To compare the conservation of multicellular responses associated with HF across the different core single-cell studies, we extended our disease score strategy to multicellular programs. Similarly, as our bulk and compositional disease score, the idea behind this strategy is to show that disease signatures of one study are sufficient to classify HF samples from any other study. We interpreted the overall classification performance of the disease signatures as their level of generalizability. First, for each study we trained a classifier of failing and non-failing hearts using the MCPs' scores of each patient and linear discriminant analysis (LDA). Then, to test the performance of each study to classify patient samples from the rest of the studies, we projected the samples of the rest of the studies into the MCPs' latent space and then predicted their disease status with the LDA classifier. In detail, to project data into an MCP latent space, we multiplied the Moore-Penrose generalized inverse matrix [68] of the concatenated gene weights across MCPs of the reference study with the scaled and normalized gene expression data of the target study. Normalization was performed as described previously. MASS v7.3-57 function ginv() was used to calculate the inverse matrix of the feature weights. AUROCs were used to test the classification accuracy as mentioned previously. To evaluate the agreement between multicellular programs at the cell-type level, we constrained pairwise study projections to use only the gene weight across MCPs to contain only the information of one cell-type at a time. MCPs latent space was visualized in two dimensions using Uniform Manifold Approximation and Projection (UMAP).

4.9 Consensus multicellular program estimation across studies

To estimate the consensus multicellular programs describing the variability of patient samples across studies, we fitted a multicellular factor analysis model to the joint collection of pseudobulk normalized expression profiles of cell types across studies. Gene expression processing was identical to the one performed for the models of individual studies, with the difference being the definition of background marker genes. Marker genes of each cell type were obtained from the consensus marker genes obtained from the combined test as previously described, with an adjusted p-value less or equal than 0.0001 and a mean log fold change greater than 2. A MOFA joint model using an extended group-wise prior hierarchy was used to integrate all studies in the inference of the MCPs. Associations of the MCPs' scores with patient covariates was done with ANOVAs and linear models as described previously. To evaluate the association of MCPs' scores to clinical features in HF patients, we used linear mixed models to model MCP scores with each clinical covariate independently and using the studies as random effects. P-values were adjusted using the BH procedure.

4.10 Functional interpretation of MCP1 and MCP2

To facilitate the interpretation of the MCPs we curated a collection of gene sets from MSigDB [69] representing molecular and cellular functions across three major processes: 1) muscle-related (to enrich functions in CMs), 2) extracellular matrix and fibrosis (to enrich functions in Fib, PC, Endo, and vSMCs), and 3) vasculature (to enrich functions in PC, Endo, and vSMCs). For each collection, we defined a collection of search words that were used in MSigDB's search engine with boolean operators and wildcards. The complete collection of search terms are available in Supplementary Table 5. On top of this curation, we included MSigDB's curated hallmarks for additional gene sets to be enriched in all cell-types.

We performed functional enrichment analyses of the MCPs from a cell-type-specific or a multicellular perspective. As described previously, each gene of each cell-type included in the model gets a weight assigned for each MCP, representing its relevance in the definition of the latent variable. For each MCP and cell-type, we kept genes with an absolute weight greater than 0 for the rest of the analysis. Given that gene weights are centered in 0, for each MCP and cell type it is possible to identify two directed gene sets (positive and negative), which represent their levels of expression in non-failing and failing heart tissues, respectively. For each directed gene set, in addition, we identified genes being relevant in more than a single cell-type and we annotated them as the multicellular component of the program, while the rest of the genes were assigned to their specific cell-type. To test if molecular or cellular functions were overrepresented in each MCP-associated gene set we performed hypergeometric tests. P-value corrections using the BH were performed for each functional gene set collection, independently. Overrepresentation of all functional gene set collections were performed for the directed multicellular MCP-associated gene sets. For cell-type-specific gene sets, overrepresentation of functional sets were specific to their expected functions as described before.

Finally, to identify differential multicellular and cell-type specific functions across MCP1 and MCP2, we estimated the position of "functional vectors" in the 2D space defined by the MCPs. The functional vector is the result of the addition of enrichment vectors across MCPs and directions (four vectors in total: two MCPs and two directions). To determine the vector of each MCP and direction for each gene set, we multiplicated the fraction of genes of the MCP-associated program that overlapped with the functional gene set with the complement of the p-value of the hypergeometric test of enrichment. These scores prioritize well represented and enriched sets in each direction. Addition of the enrichment vectors were

done for all functional gene sets with an adjusted p-value lower or equal to 0.1 in at least one MCP-associated gene set. For visualization purposes, representative pathways were manually selected.

4.11 Multicellular information networks

MCPs describe global coordination across cell-types, however they do not explicitly describe the degree of dependency of each pair of cells. Thus, we inferred from MCP1 a cell-cell dependency network with directed edges that describe to what extent the molecular profile of one cell-type could predict the profile of each other cell-type. First, we generated cell-type signature matrices per patient sample and study, by enriching the cell-type specific gene weights of MCP1 into the pseudobulk profiles of their respective cell-type. Enrichment was performed with linear models as implemented in decoupleR v2.6.0 [70]. Gene weights with an absolute value lower than 0.1 were excluded from the signature to be enriched. To estimate the weight of the incoming edges of the multicellular information networks, we built linear mixed models to predict the signatures of a specific cell-type with the signatures of the rest. In these models, the study of origin was treated as a random effect. For a given predicted cell-type, the coefficient estimate of each predictor cell-type multiplied by the overall fit of the model was used as the final edge weight.

4.12 Ligand-receptor coexpression networks

We extracted curated ligand-receptor pairs with liana v0.1.13 R-package [50]. To calculate ligand-receptor interactions for each cell type pair we extracted the expressed ligand and receptor genes with an absolute gene loading >0.1 in MCP1 and summed both loadings to a ligand-receptor interaction score. To connect these ligands with possible downstream targets, we used the updated Nichenet resource [37]. In detail, for a given cell type A and cell type B pair, we selected all ligands expressed in cell type A and predicted target genes as defined by the top 10% highest absolute gene loadings of the MCP1 from cell type B. Background genes were defined as the 30% genes with MCP1 loadings closest to 0. Then, corrected AUPR was calculated as implemented in nichenetr v.2.0.4. R-package and together with L-R interaction score used for ligand prioritization.

4.13 Study projection of supporting studies

Supporting single-cell studies were projected into the latent space formed by the MCP latent space estimated from the core studies using the gene weights of the MCPs as described previously. Associations of clinical covariates or patient conditions with MCP1 and MCP2 were tested as in individual and combined models. Pseudobulk expression matrices of projected mice studies (AngII and TAC) were

translated into human gene symbols using biomaRt v2.58.0. We filtered out genes with ambiguous mapping including duplications. To project HF patient samples of studies without a healthy reference into the core MCP latent space, we first normalized their pseudobulk expression data using an independent healthy heart single-cell human cell atlas [71]. Processing of the healthy single-cell atlas followed identical processing and normalization as the core and supporting studies. The gene expression data of the studies to be projected was then standardized relative to the healthy atlas. For each gene, we subtracted the mean expression of its corresponding gene in the healthy atlas and divided it by the healthy standard deviation. To evaluate that the projected HF samples were located appropriately in the MCP latent space we created a decision boundary to classify failing from non-failing tissue samples using the information of the core studies. We trained a logistic linear regression to predict HF using the patient values of MCP1 and MCP2 across the four core studies. We then predicted for the projected samples their disease status and calculated the classification error.

4.14 Mapping MCPs to collection of bulk transcriptomics data

To test if MCPs were traceable in bulk transcriptomics data, we calculated the enrichment of cell-type-specific signatures of each MCP into each study of our bulk transcriptomics data curation. For a given bulk sample within a study, we used decoupleR v2.6.0 to estimate enrichment scores of each cell-type signature using linear models. The gene expression of each study was centered and scaled across samples before the enrichment estimation. AUROCs for each bulk study and cell-type were calculated to test if the estimated cell-type-specific signatures were discriminative of failing and non-failing tissue patient samples. Enriched signatures were sorted by their value and non-failing samples were used as response variables to calculate AUROCs.

4.15 Annotation of deregulation processes of bulk data

To integrate the observations of gene deregulation in HF from bulk and single-cell transcriptomics data, we explored to what extent a gene reported to be increasing or decreasing its levels of expression in bulk could be explained by changes in tissue composition or by cell-type specific/multicellular regulatory processes (here referred to in general as molecular process). We assumed that a change in expression associated with changes in tissue compositions, could be traced in genes that have specific expression in cell-types that showed a change in compositions in HF. Marker genes of one or more cell-types could potentially be annotated in this way. In contrast, we assumed that a change in expression associated with a purely molecular process, would be reflected by an agreement between the direction of deregulation in bulk data with the overall change across cell types as reflected by our MCP1. Following these

assumptions, we devised two scores from the core single-cell studies that reflect both compositional and molecular processes. The compositional score was calculated for all marker genes of all cell-types as a product between the mean log fold change of marker expression across studies (Section 4.3) with the mean t-value across studies of the compositional differential analysis of their specific cell-type (Section 4.4). For genes that were markers of more than a single cell-type, we summed the scores. The molecular score of all genes modeled with Multicellular Factor Analysis, was the mean gene weight across cell-types for MCP1, given its proven traceability in bulk. Finally, every gene in the bulk consensus signature was annotated to be potentially deregulated from a compositional, molecular, compositional and molecular or unknown processes, depending on the agreement between the mean t-value of the change in expression between failing and non-failing hearts in bulk transcriptomics datasets and the compositional and molecular scores estimated from the core single-cell studies. If genes had identical signs in the bulk log fold change and the single-cell score, they were annotated with the compositional, molecular, or compositional and molecular label. Genes not profiled with single-cell data or not agreeing with any score were considered unknown.

4.16 Bulk cell type deconvolution

We performed bulk deconvolution to i) test the impact of cell type marker regulation in HF and ii) to infer compositions in bulk RNAseq datasets to identify conserved compositional changes in HF. We used CIBERSORT [72] to estimate cell type compositions in mixed cardiac samples. This method requires a signature matrix that contains marker genes for each cell type. To design this matrix we selected consensus cell type markers derived from our core HF-studies by selecting markers with an adjusted fisher p-value $<10e^{-}50$. We then transferred the annotations of bulk deregulation to these markers and identified four subset of cell type markers: i) unregulated markers, ii) generally deregulated markers, iii) molecularly deregulated markers and iv) compositionally deregulated markers. We used a healthy cardiac single-nucleus reference data set [1], that was filtered for cells from LV location, and manually aligned to our vocabulary of the seven major cell lineages, then transformed to pseudobulks and finally normalized to transcripts per million (TPM) and kept in linear scale, to build the signature matrix according to the authors recommendations [72]. We sampled without replacement 30% of all cells to generate the TPM expression values, generating three different signature matrices, which were merged after deconvolution by taking the average predicted compositions. For the benchmarking of the signature matrices we selected the 4 core and 6 supplemental HF sc-studies and summed gene counts for each patient across all cell lineages to get mixtures with known cell type proportions and applied the same normalization procedure as for the signature matrices. The estimated proportions were then evaluated by calculating

root-mean-squared-error and pearson's correlation per study. To evaluate the estimated proportions we performed differential compositional analysis as described in 4.4. We assigned classes of compositional increase, decrease (t-test p-value <0.05), or no change (t-test p-value >= 0.05) and calculated F1-scores for each class. For this evaluation of compositional change we only tested seven of the ten studies as Amrute2023, Nicin2022, and Liu2022 shared the control samples with the healthy reference.

The HF core bulk studies were then deconvoluted with the compositional signature matrix, and compositional changes were again evaluated as described in 4.4. Cell-type composition matrix from the single-nucleus and bulk cohort was then analyzed together via principal component analysis, while each principal component (PC) was tested for association with HF, technology or study variable and explained variance of significantly associated PCs (linear model p< 0.001) were summed for total estimate of associated variance.

4.17 Single cell integration of fibroblasts

We integrated fibroblast single cell data from the four core HF studies. Each data set was processed individually by applying common thresholds, filtering genes with at least 10 cells and cells with at least 200 genes and counts were log-normalized. Prior to integration, top 200 cell lineage markers from non-fibroblast cell lineages were removed to avoid background contamination, then highly variable genes were calculated. The top 2000 highly variable genes, common to multiple data sets were selected for integration. In this feature space, we regressed out total counts per cell, scaled data and performed PCA. This PCA was then adjusted for the sample label by Harmony [62]. In the resulting embedding, we calculated a nearest neighbor graph and performed leiden clustering to identify cell states. We detected a cluster with low gene counts and expression of cardiomyocyte marker TTN or RYR2, which did not display enrichment of the fibroblast program captured by the MCP1. We removed these cells and repeated the integration steps, resulting in a final integrated fibroblast atlas. We calculated state markers with Wilcoxon tests implemented in SCANPY v1.9.5. [73]. We performed differential abundance testing of cell states between conditions as described in 4.4.

4.18 State characterization and division of labor in fibroblasts

We summed counts per patient and per cell state to pseudobulks which were filtered and normalized with TMM implemented in edgeR [74]. We enriched prior knowledge genesets from CytoSig [75], PROGENy [76], and MSIG DB [69]. We scaled enrichment scores and selected variable genesets (ANOVA p-value <0.05) to apply t-tests in “one vs. the rest” manner to determine significantly upregulated gene sets after BH

correction. We selected top 200 cell state markers based on Wilcoxon p-value and enriched them in the gene loading vectors of MCP1 and MCP2. We selected the mean t-statistics of studies calculated from the differential abundance testing of cell states and fit a linear model to characterize association between state marker enrichment in the MCPs and compositional changes. To characterize the expression pattern of top genes from MCP1 and 2 in different cell states, we enriched the normalized cell state pseudobulks for the top genes from MCP1 and MCP2 (absolute loading >0.1) using the loadings as weights. For this analysis we multiplied the weights with -1, such that positive values would associate with HF. To characterize the variability of enrichment scores between cell states and patient cohorts, we fit a linear mixed model for each MCP with the formula:

$$ES \sim HF + CS + (1 \mid Study + Patient_ID)$$

where ES represents the enrichment score, HF represents HF status, and CS represents cell state. Random intercepts were included for both Study and Patient_ID. We then extracted semi-partial R^2 values for the fixed effects HF and CS with the r2glmm v0.1.2 R-package [77] to compare the expression pattern across cell states and HF status. To determine an individual gene's expression pattern, the normalized gene expression value was modeled instead of the enrichment score. We assigned division of labor groups based on combination of semi-partial R^2 values: specialist (R^2 cell state >0.1 AND R^2 HF <0.1) generalist (R^2 cell state <0.1 AND R^2 HF >0.1), and acquired generalist (R^2 cell state >0.1 AND R^2 HF >0.1). These gensets were further separated in HF or nonfailing associated by using the sign of the gene's loading in MCP1. The MCP gene loadings and the consensus bulk signature were then used as directed rankings and enriched for the division of labor gene sets. All enrichment analyses were performed via *run_ulm()* from the decoupleR v2.7.1 R-package [70].

5. Acknowledgements

RORF acknowledges the support of the German Science foundation (DFG) through the CRC 1550 “Molecular Circuits of Heart Disease”. JDL acknowledges the support of the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung) through CureFib. JLB is supported by the Galician Government through the fellowship ED481B_072. We thank Leonie Küchenhoff for critical feedback on the manuscript. Importantly, we acknowledge all data authors that are cited in this study.

6. Data and code availability

All code associated with this publication is available at github.com/saezlab/reheat2_pub. The original data can be accessed from individual studies in various repositories, as summarized in Supplemental Table 1. Processed data is available on Zenodo (DOI:10.5281/zenodo.13946108). An interactive query of MCP1 and the consensus heart failure signature is accessible via a web application at https://saezlab.shinyapps.io/reheat2/ with the source code available at github.com/saezlab/reheat2_shiny.

7. Conflict of interests

JSR reports funding from GSK, Pfizer and Sanofi and fees/honoraria from Travere Therapeutics, Stadapharm, Astex, Pfizer, Grunenthal and Owkin.

8. Authors contributions

JDL: Data Curation, Formal Analysis, Investigation, Methodology, Resources, Software, Writing – Original Draft Preparation.

RORF: Data Curation, Formal Analysis, Investigation, Methodology, Resources, Software, Writing – Original Draft Preparation.

JLB: Formal Analysis, Validation, Writing - Revision.

JSR: Supervision, Project Administration, Funding Acquisition, Writing - Revision.

9. Supplementary materials

9.1 Supplementary table captions

Supplementary table 1. Source links of used data
Source links for data repositories for core and supporting studies.
Supplementary table 2. Consensus cell-type gene markers
Results of the Fisher combined test aggregating cell-type markers across studies. Mean log-fold changes (LFC) and adjusted p-values are provided.
Supplementary table 3. Search terms for functional analysis
Search terms Cardiomyocytes, Vasculature and Fibrosis gene sets to select genes sets from MSigDB for functional interpretation
Supplementary table 4. Ligand-Receptor interaction in non failing and failing tissue
Ligand-Receptor interactions between cell type pairs based on the assigned gene loadings within MCP1
Supplementary table 5. Division of labor gene programs
Results from the linear mixed model to assign division of labor groups to fibroblast genes within MCP1
Supplementary table 6. Consensus signature of heart failure
Genes ranked by consensus deregulation in HF, including MCP1 gene weights, log-fold changes for cell-type marker genes, and annotations of potential deregulation mechanisms.
Supplementary table 7. Cardiac cell type signature matrix for deconvolution tasks
Cell type signature matrix of compositionally regulated genes with TPM normalized expression values from a healthy reference atlas

9.2 Supplementary notes

Supplementary Note 1: Update of the consensus bulk HF signature

We compared the updated consensus signature with the previous version by assessing overlap of top gene sets and their classifier performance. Although the updated ranking consensus signature reordered the top 500 genes (Jaccard index 0.49, Supplementary Figure 2D), the classification performance these genes did not change significantly and remained high (mean AUC 0.962 vs 0.968 [old version], t-test p=0.235, Supplementary Figure 2E). The enrichment scores of the updated consensus signature also did not change significantly compared to the previous version (Supplementary Figure 2F). However, the number of genes reported under the adjusted Fisher combined test p-value $<10^{-5}$ increased from 1,809 to 4,518 substantially, including a higher gene coverage (increased from 14,041 to 16,036) (Figure 2C). Thus, the inclusion of new studies increased the robustness of the conserved events of deregulation captured by the gene ranking, while suggesting that classification performance cannot be substantially improved by adding additional data sets.

Supplementary Note 2: Comparison of pseudobulk expression matrices of single-nucleus studies.

Core studies had a median of 17,523 non-zero genes measured in each cell-type pseudobulked data and were calculated from a median of 744 cells with significant but not substantial differences in both, number of cells and genes, across studies (ANOVA adj. p-value < 0.001, Supplementary Figure 3A-C), reflecting expected differences at the technical level.

Ambient molecules from droplet-based single-nucleus experiments can influence gene expression profiles and affect the accuracy of cell-type specific disease signatures [78]. For the same reason, we quantified the levels of gene expression contamination in each pseudobulk profile (Supplementary Figure 3A-D). We assumed that for a pseudobulk profile of a given cell-type, a simple measure of the levels of contamination in its gene expression could be approximated from the expression of "unexpected" genes relative to "expected" genes. Here we used marker genes of the cell-type from which the pseudobulk profile was created to define the set of "expected" genes, and the marker genes of the other six cell-types as "unexpected" ones. Then, we defined a contamination score for each pseudobulk profile as a ratio between the read counts of background and marker genes. CM profiles had the lowest contamination scores (mean = 0.124) of all cell-types (t-test adj. p-value < 0.05), and in all studies the gene expression profiles of vSMCs contained more contamination reads than marker genes (one-sample t-test adj. p-value < 0.05, null hypothesis: contamination score mean = 1, Supplementary Figure 3A,C). In addition, we observed that CMs were the cells that contributed the highest fraction (mean fraction of 0.4, Supplementary Figure 3B) of contaminating reads to other cell-types (t-test adj. p-value < 0.05). Two studies provided background corrected gene expression data and we observed differences in the contamination scores in all cell-types between background corrected and not-corrected studies (t-test adj. p-value < 0.05). However, depending on the cell-type analyzed, background correction would increase (Myeloid, Fib, vSMCs) or decrease (CM, Endo, PC, Lymphoid) contamination (Supplementary Figure 3C-D). Our results point to a shared technical limitation in these studies that should be taken into account when comparing gene expression profiles between non-failing and failing hearts.

Supplementary Note 3: Multicellular factor analysis of individual studies

We observed on average that the multicellular programs of individual studies captured 17% of explained variance associated with HF (ANOVA adj. p-value <= 0.05, Supplementary Figure 4B), 20% with left ventricle ejection fraction and 5% to age (linear model adj. p-value <= 0.05, Supplementary Figure 5C). Other patient covariates such as sex or body mass index had no associations with the multicellular space. Visualization of samples in a 2-dimensional Uniform Manifold Approximation and

Projection (UMAP) space built from the multicellular programs showed clear separations of failing and non-failing hearts in all studies (Figure 2G).

Supplementary Note 4: Enrichment of specialist, generalist, and acquired generalist functions of fibroblasts to MCP1 and bulk data

To characterize which expression group (specialist, generalist or acquired generalist) predominates in the Fib component of MCP1, we compared their respective gene loadings. The acquired generalist program was assigned the highest median loading (Wilcoxon's tests, vs. generalist genes $p=2.9e^{-}11$, Supplementary Figure 9G), suggesting that the key genes in the MCP1 were characterized by variability between states but accommodating a generalist expression pattern by upregulation across multiple states. Next, we enriched these programs in the HF consensus bulk signature and found that the acquired generalist program displayed the highest enrichment scores (Supplementary Figure 9H), suggesting that their importance generalizes to a larger HF cohort and can be detected in bulk signatures.

Supplementary Note 5: Evaluation of sets of cell-type marker genes for cell deconvolution from bulk transcriptomics

When deconvoluting diseased samples, we found that markers regulated at the compositional level performed better than every other marker set (root mean squared error [RMSE]: 0.117, paired wilcoxon-test p-value = 0.002; Correlation: 0.895, paired wilcoxon-test p-value = 0.002; Figure 5E). For the deconvolution of healthy samples, this effect was mitigated. When comparing deconvolution errors between cell types, we found that the compositional genes also lowered the RMSE compared to molecular genes in all cell types except for lymphoid cells and pericytes (RMSE, paired wilcoxon-test p-value <0.05) (Supplementary Figure 10D). A downstream goal of bulk deconvolution is the assessment of compositional changes between conditions. We found that molecularly regulated markers failed to reliably predict compositional changes (global average F1-score of 0.44) while compositional genes performed best (global average F1-scores 0.64) (Supplementary Figure 10E). Taken together, we found that molecularly regulated cell type markers are poor indicators of cell type composition in disease, and thus identifying compositionally regulated cell type markers can improve deconvolution results.

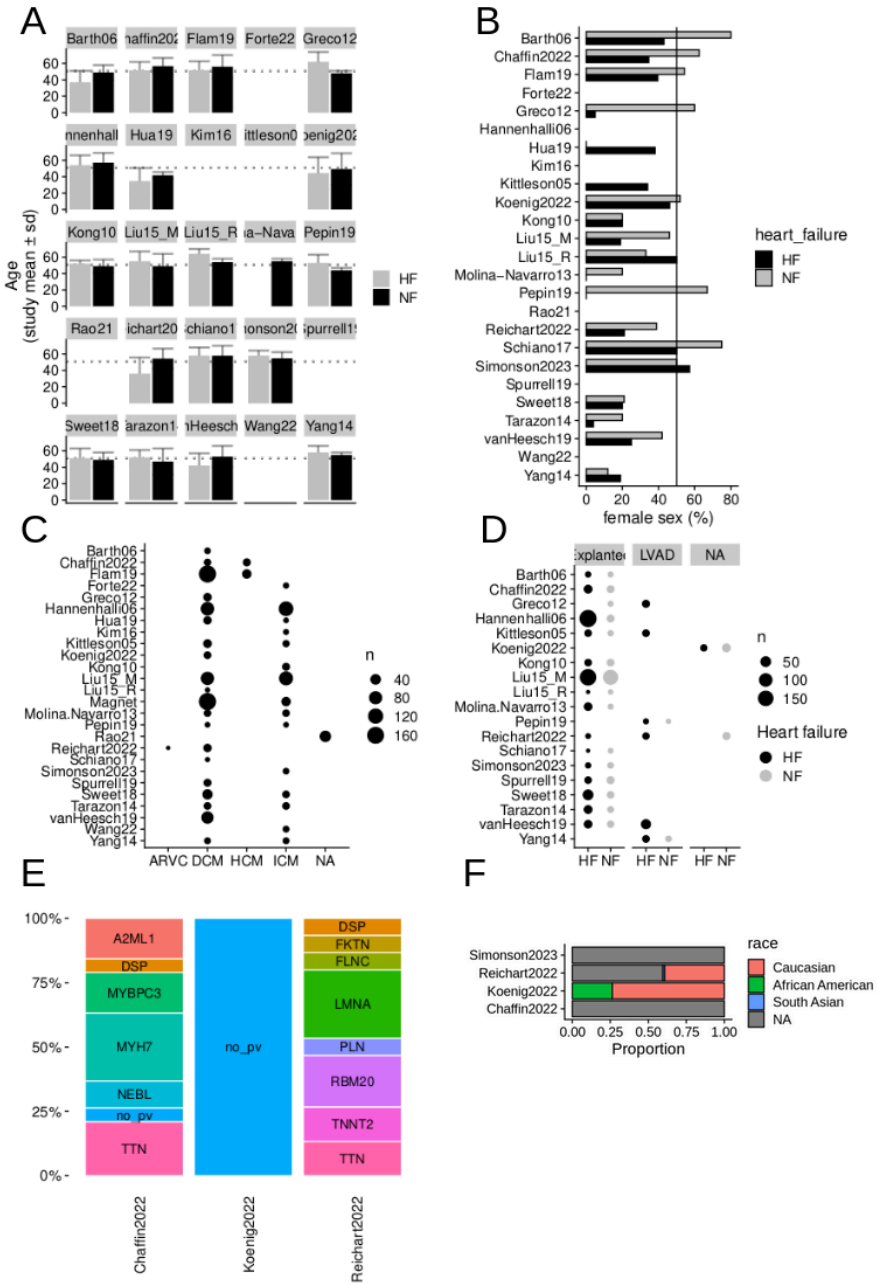
Supplementary Note 6: Projection of supporting studies into the HF multicellular patient map

Other supporting studies let to insights regarding the MCP2, where we observed differences in activation of MCP2 across distinct physiological time-points upon myocardial infarction aligned to the trajectory of HF samples where fibrotic heart tissues located in between ischemic and control tissues (ANOVA, p-value = 0.00005; t-test, adj. p-value < 0.05; Supplementary Figure 11A), potentially related

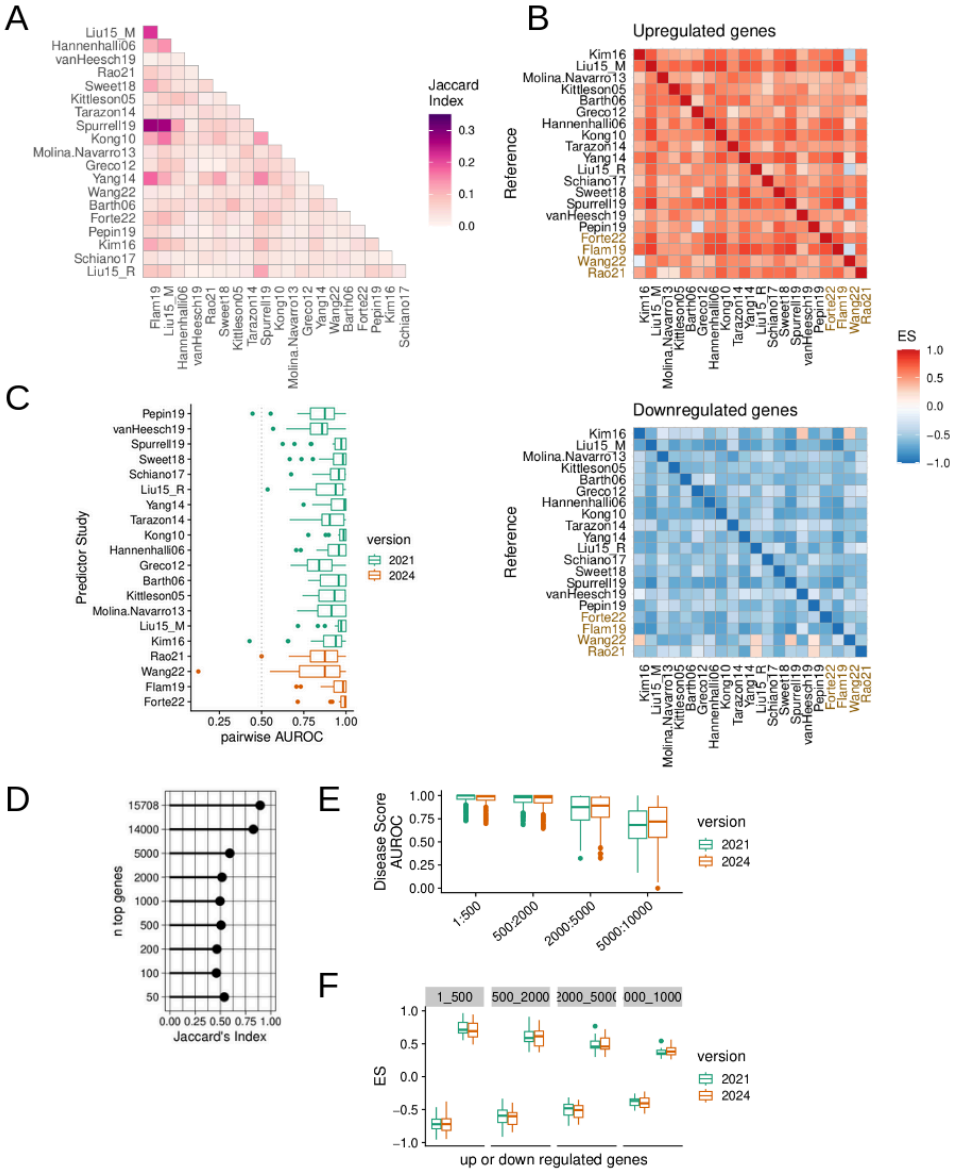
to the activation of apoptotic processes of this multicellular program. A less clear separation between congenital heart disease samples and donor hearts across MCP2 was observed (t-test, p-value = 0.07; Supplementary Figure 11B).

Finally, we projected data of two mice models of heart disease, Angiotensin II-Induced (AngII) and transverse aortic constriction (TAC), to evaluate to what extent the mice phenotypic characteristics of cardiac hypertrophy and failure were aligned to HF patients (Supplementary Figure 11C-D). We did not capture associations between the activation of MCP1 and MCP2, and the disease progression of the TAC model or AngII mice representing HF. These results showed a disagreement between the major axes of disease variability between mice and human samples.

9.3 Supplementary figures



Supplementary Figure 1. Extended metadata presentation.
A-D) Presenting metadata per HF core study, including A) age distribution, B) sex distribution, C) HF etiology, D) Reason for biopsy.
E) Comparing the pathogenic variants of HF patients with familial or genetic cardiomyopathy per single-nucleus study.
F) Comparing reported race of patients per single-nucleus study.



Supplementary Figure 2. Expansion of the consensus bulk transcriptional signature of HF

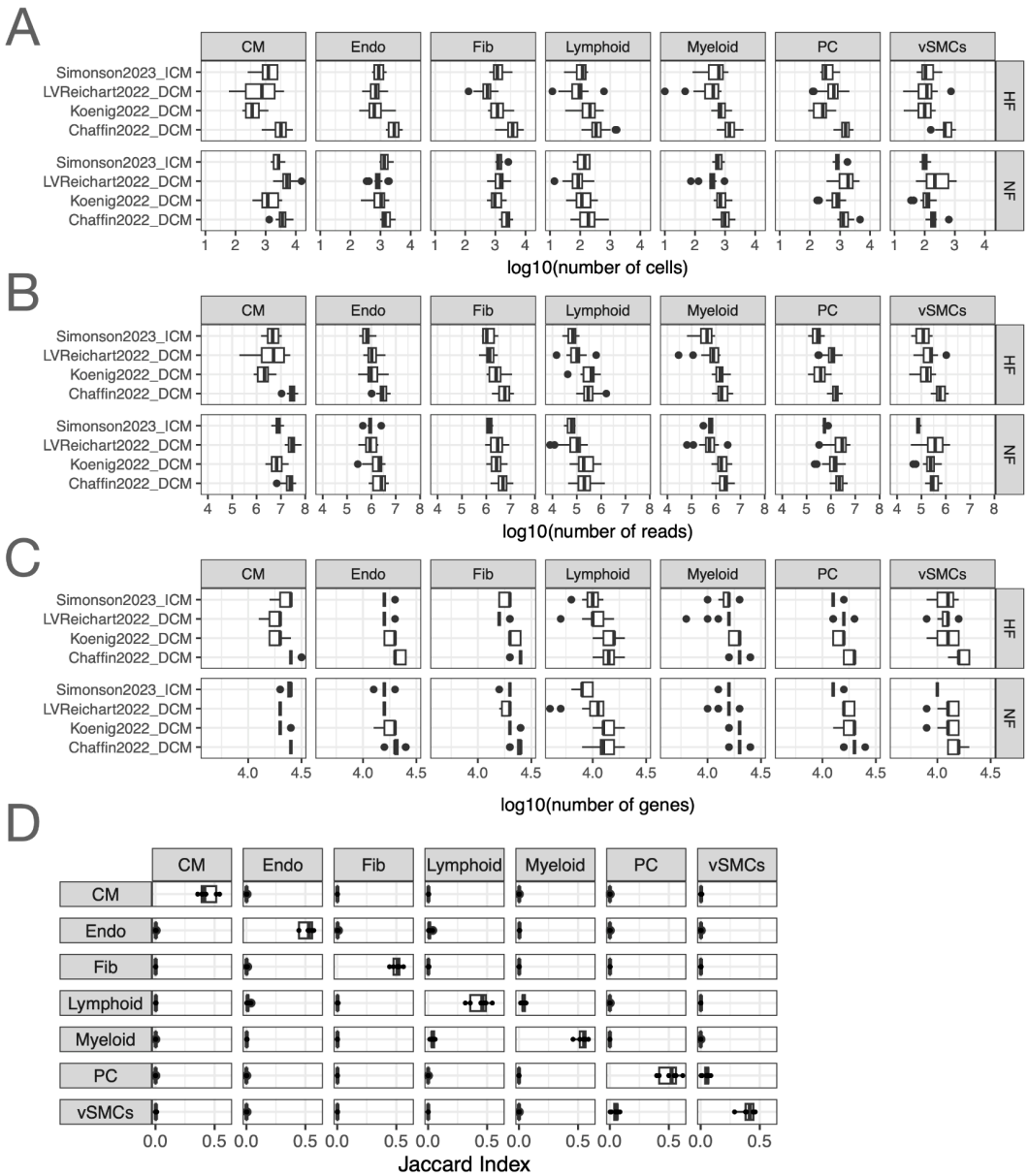
A. Pairwise quantification of the overlap of the top 500 differentially expressed genes between failing and non-failing hearts between core HF bulk studies using the Jaccard index.

B. Enrichment score (ES) of the top 500 differentially expressed genes between failing and non-failing hearts of each study in the sorted gene-level statistics list of each other study. Colored study names are new studies added.

C. Comparison of the updated consensus ranking with the previous version. Jaccard indices were calculated between different number of top genes.

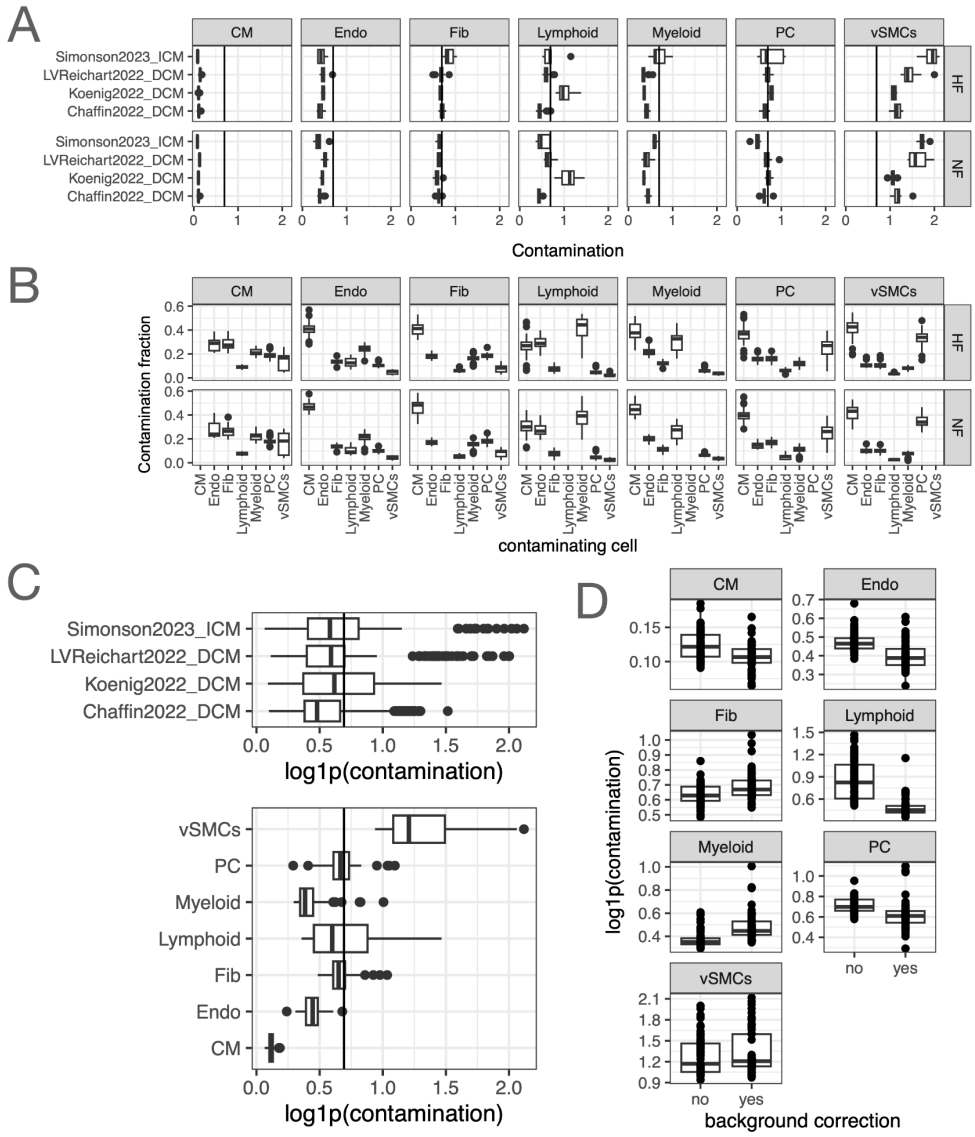
D. Comparison of classification via disease scores calculated with gene setss of different size from the updated and previous ranking.

E. Comparison of enrichment scores of top genes extracted from the consensus rankings in individual studies.



Supplementary Figure 3. Quality of pseudobulk expression profiles of the collection of core single-nucleus (SN) transcriptomics studies.

A-C. Distribution of A) the log10(number of cells) used to calculate the pseudobulk expression profile, B) the number of reads, and C) the number of genes
D. Distribution of Jaccard Indices representing the pairwise similarities between sets of gene expression markers of cell-types (Methods). Each dot represents a comparison between the markers of cell-type X with markers of cell-type Y for every combination of SN-core studies.
In all panels Cardiomyocytes (CM), fibroblasts (Fib), pericytes, and endothelial (Endo), vascular smooth muscle (vSMCs) cells.



Supplementary Figure 4. Contamination of pseudobulk expression profiles of the collection of core single-nucleus (SN) transcriptomics studies.

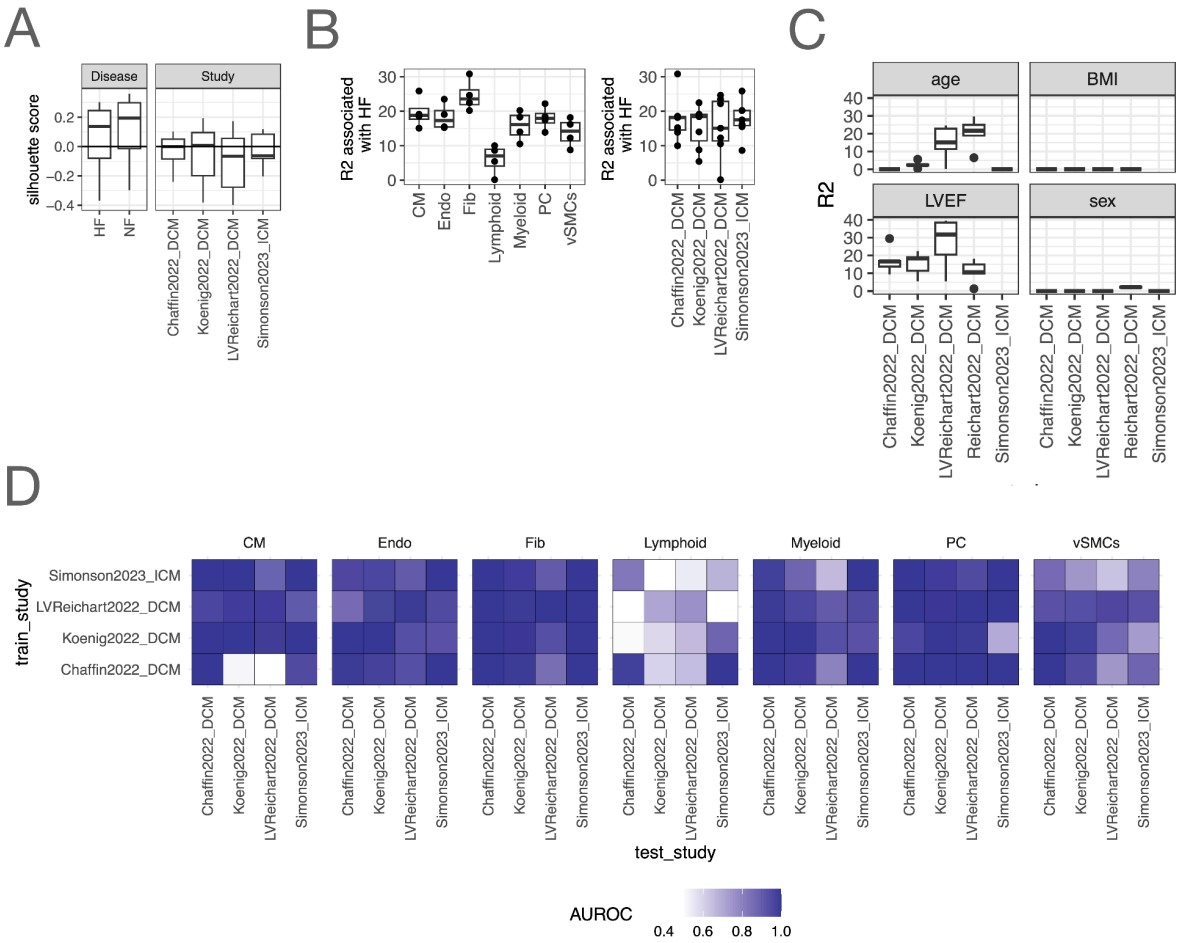
A. Distribution of the contamination score of the tissue samples of each core SN study, divided by cell-type as defined by our ontology and disease status. The line denotes the score where the amount of cell-type reads and contaminating reads is identical.

B. Read contamination fraction of each cell-type in the x-axis to the expression profile of the cell-type in the header, across studies.

C. Distribution of the contamination scores of each tissue sample across SN-core studies (upper) or cell-types (lower). The line denotes the score where the amount of cell-type reads and contaminating reads is identical.

D. Distribution of the contamination scores of each tissue sample across cell-types separated in two groups based on previous background correction of the study.

In all panels Cardiomyocytes (CM), fibroblasts (Fib), pericytes, and endothelial (Endo), vascular smooth muscle (vSMCs) cells.



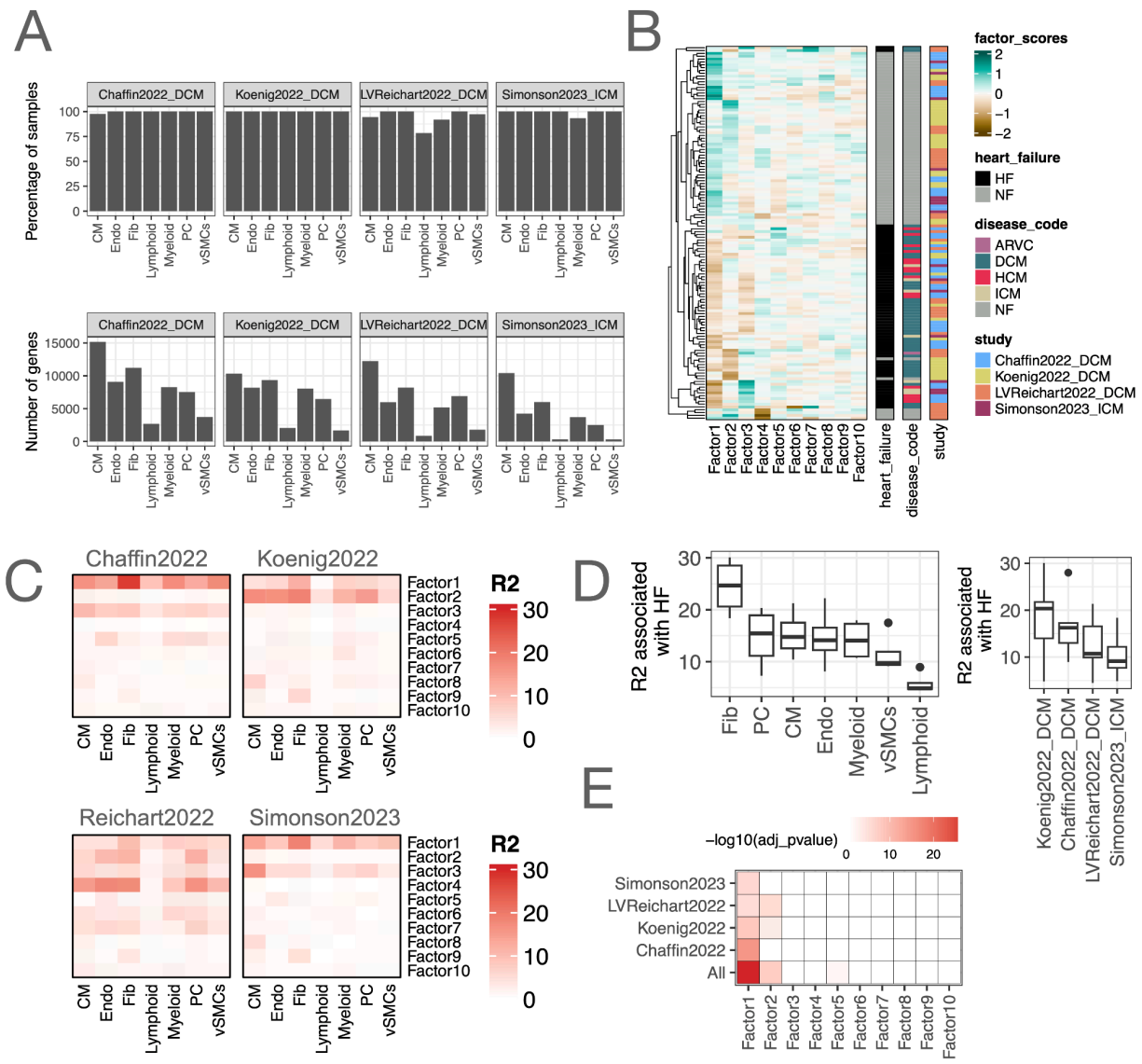
Supplementary Figure 5. Comparison of the single-nucleus (SN) core studies from compositions and multicellular programs

A. Distribution of silhouette scores of patient heart samples of SN core studies grouped by disease status (left) and study (right).

B. Percentage of explained variance (R2) associated with HF (HF) captured by the multicellular factor analysis models fitted to each study independently. Each dot represents one study (left) or one cell-type (right).

C. Percentage of explained variance (R2) associated with age, body mass index (BMI), left ventricular ejection fraction (LVEF) and sex captured by the multicellular factor analysis models fitted to each study independently.

D. Area under the receiver operating characteristic curve (AUROC) of pairwise predictions of disease classifiers built from all core SN-studies using cell-type specific information within multicellular programs.



Supplementary Figure 6. Quality metrics of joint multicellular factor analysis of single-nucleus (SN) core studies

A. Percentage of patient samples and number of genes included in the multicellular factor analysis model separated by study

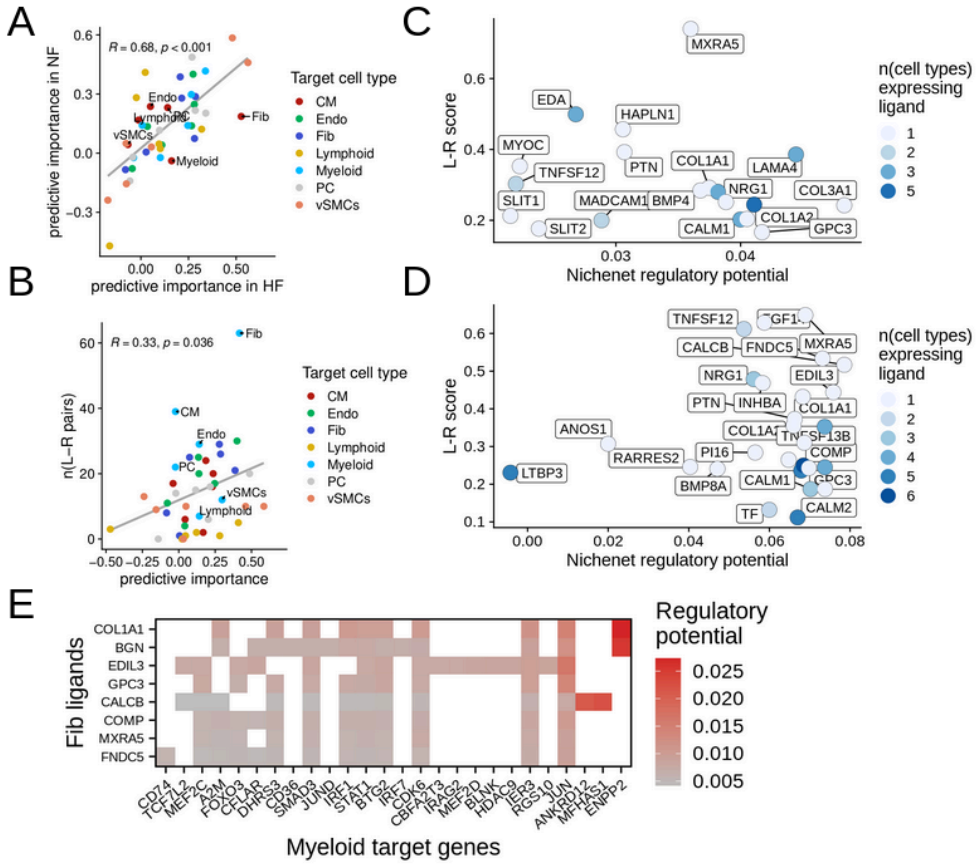
B. Hierarchical clustering of the patient-level values of the 10 multicellular programs inferred in the joint multicellular factor analysis model. Annotations of disease status, etiology, and study of origin are provided.

C. Percentage of explained variance (R2) that each multicellular program captured for each cell-type across studies.

D. Percentage of explained variance (R2) associated with HF (HF) captured by the joint multicellular factor analysis model. Each dot represents one study (left) or one cell-type (right).

E. -log10 of the adjusted p-values of an analysis of variance used to test for association between the disease status of patients and their multicellular program activation across all studies and with their union.

In all panels Cardiomyocytes (CM), fibroblasts (Fib), pericytes, and endothelial (Endo), vascular smooth muscle (vSMCs) cells. Heart failure (HF), and non-failing (NF) hearts.



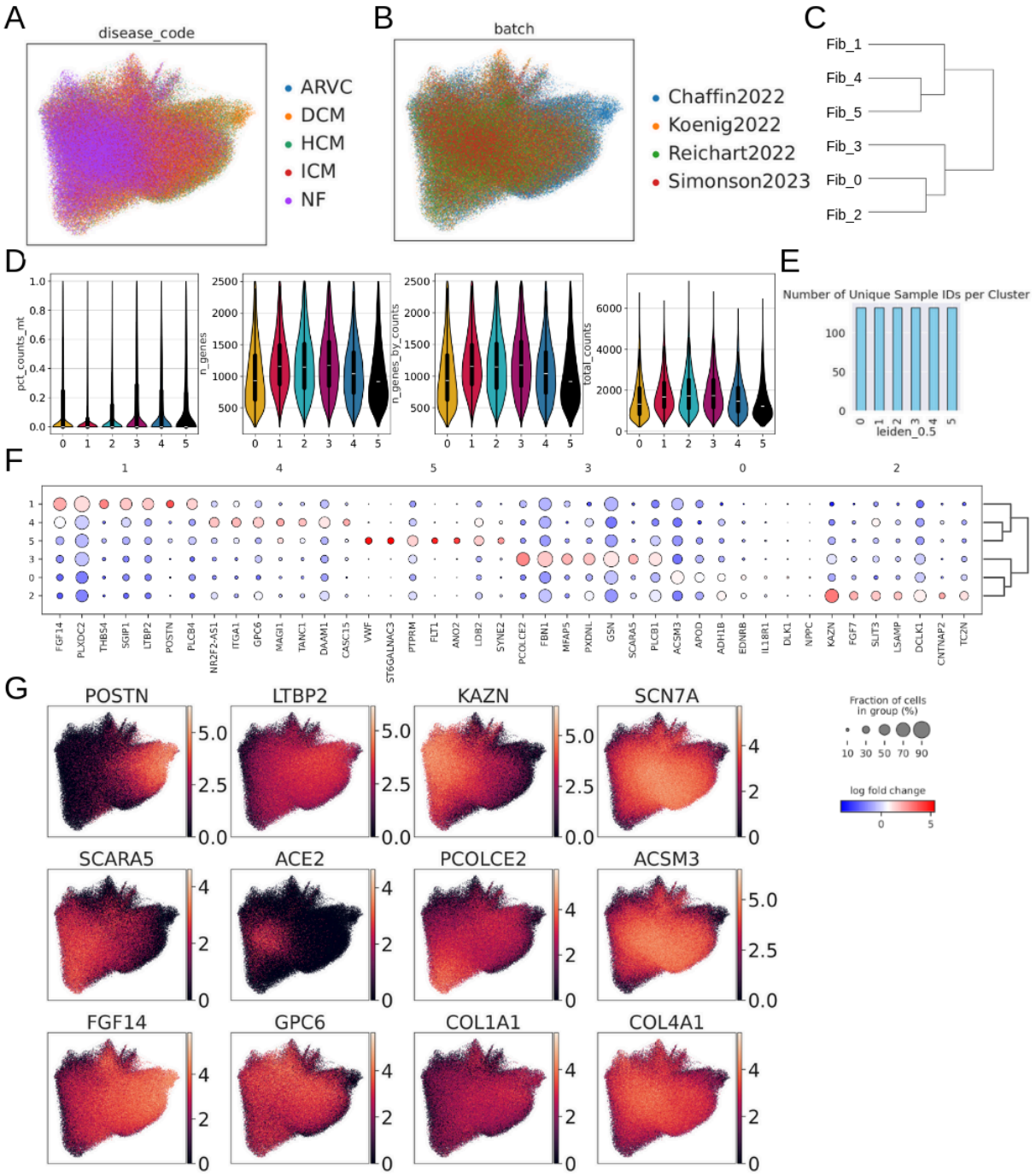
Supplementary Figure 7. Extended cell communication results

A. Comparing the predictive importances of the MCP1 expression in patients with HF (x-axis) and controls (y-axis). Each dot represents a directed cell type pair with target cell types colored. R, Pearson's correlation coefficient.

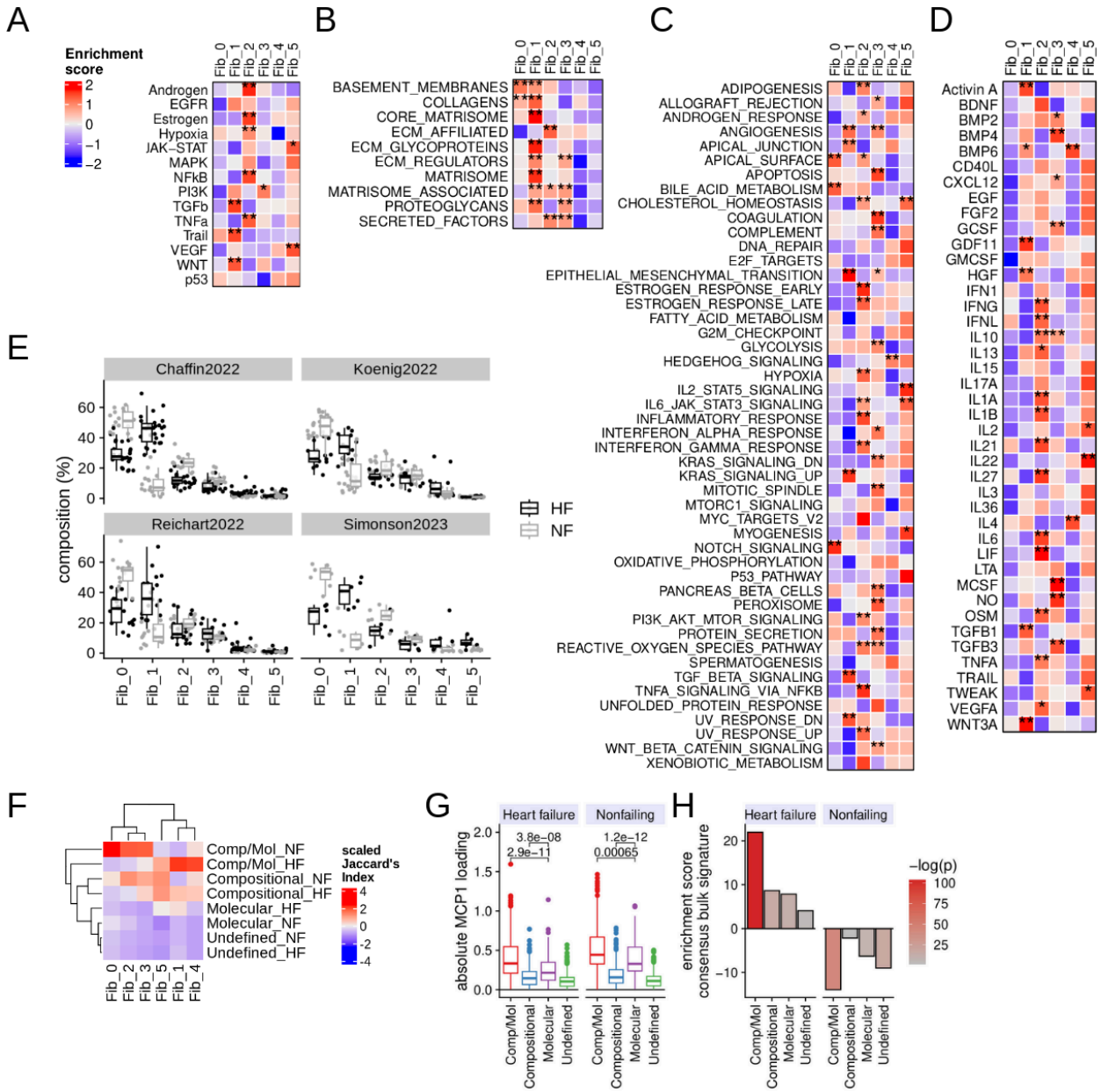
B. Comparing the predictive importance of MCP1 expression in nonfailing patients (x-axis) with the number of ligand receptor pairs inferred. Each dot represents a directed cell type pair with target cell types colored. R, Pearson's correlation coefficient.

C. C+D) Nichenet results for top L-R pairs between Fibroblasts and cardiomyocytes (C) and fibroblasts and myeloids (D). With Nichenet we calculated a regulatory potential (corrected AUPR) of a given ligand to deregulate a gene signature which here represents gene loadings of the target cell type taken from the MCP1. L-R score (y-axis) represents the mean gene loading of the ligand and the receptor coding gene of a given cell type pair. If a ligand connected with multiple receptors, the median L-R score was calculated. Color represents the number of cell types that also express this ligand in HF (one represents a cell type specificity).

D. E) Regulatory potential score, as estimated by NicheNet, representing the potential of fibroblasts' ligands in contributing to the regulation of Myleoid genes. In all panels Cardiomyocytes (CM), fibroblasts (Fib), pericytes, and endothelial (Endo), vascular smooth muscle (vSMCs) cells. Heart failure (HF), and non-failing (NF) hearts.



Supplementary Figure 8. Integrating fibroblasts single-nucleus RNAseq data.
A+B) UMAP embedding colored by disease code (A) and batch (B).
C) Dendogram of hierarchical clustering of fibroblast states.
D) Panels showing mitochondrial genes as percentage for each cell, number of unique genes per cell, number of unique genes with at least 10 counts per cell, and total counts per cell.
E) Number of unique patients per cluster.
F) Dot plot representing state marker expression per cell state.
G) UMAP visualization of expression of selected marker genes.



Supplementary Figure 9. Integrating fibroblasts single-nucleus RNAseq data.

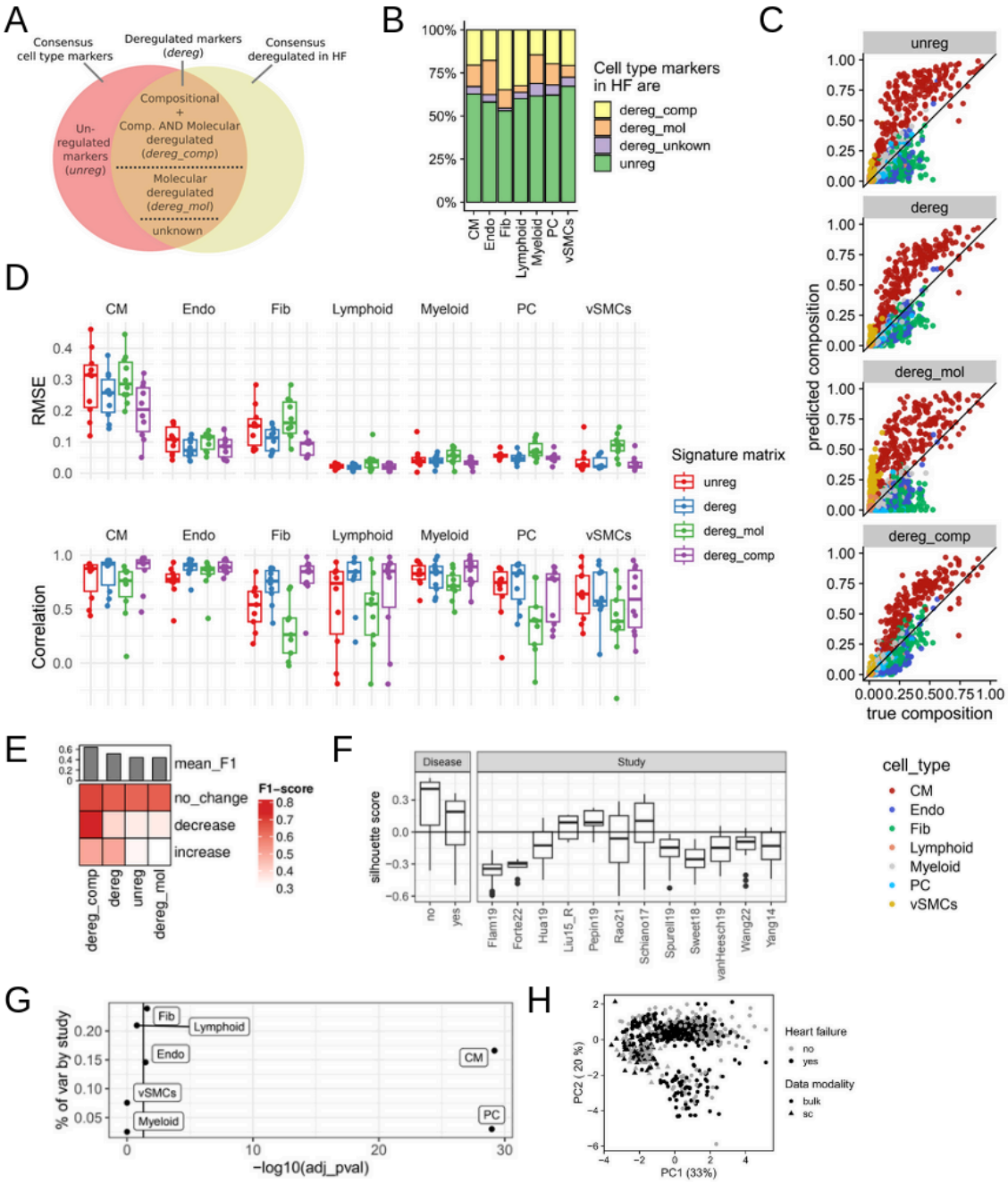
A-D) Different prior knowledge gene sets were enriched in pseudobulked cell state and patient profiles. Mean scaled enrichment scores across patients shown, t-test were performed on enrichment scores via one vs. all with adjusted p-value *<0.01, **0.001, ***0.0001. The different prior knowledge gene sets included (A) PROGENy pathways, (B) ECM genes (NABA geneset from MSIGDB), (C) MSIGDB Hallmarks, (D) Cytosig.

E) Sample composition of fibroblast cell states separated by HF (color) and studies (panels).

F) Jaccard Indices scaled per cell state to visualize intersections of top 150 state markers with division of labor groups.

G) Absolute loadings of multicellular program 1, comparing division of labor programs for heart failure and non-failing associated genes. Wilcoxon's test p-values are provided for selected comparisons.

H) Enrichment scores (y-axis) for different division of labor programs associated with heart failure and non-failing genes in the bulk consensus signature of heart failure. Enrichment was performed using univariate linear models, with negative log10-transformed p-values displayed as coloring.



Supplementary Figure 10. Cell type deconvolution of pseudobulk and bulk.

A. Schematic of annotation of cell type markers deregulated in HF. Different subsets of markers and their deregulation characteristics were used to build signature matrices for cell type deconvolution. The abbreviations for the four resulting signature matrices are written in italic.

B. Quantification of annotations described in (A) per cell type.

C. Comparing true (x-axis) and estimated compositions (y-axis) from pseudobulked profiles of patients, colored by cell type. Each panel represents estimated compositions by using a different signature matrix.

D. Quantification of deconvolution performance per cell type and signature matrix. Root mean square error (RMSE, top) and Pearson's correlation (bottom) were calculated (each data point represents one study).

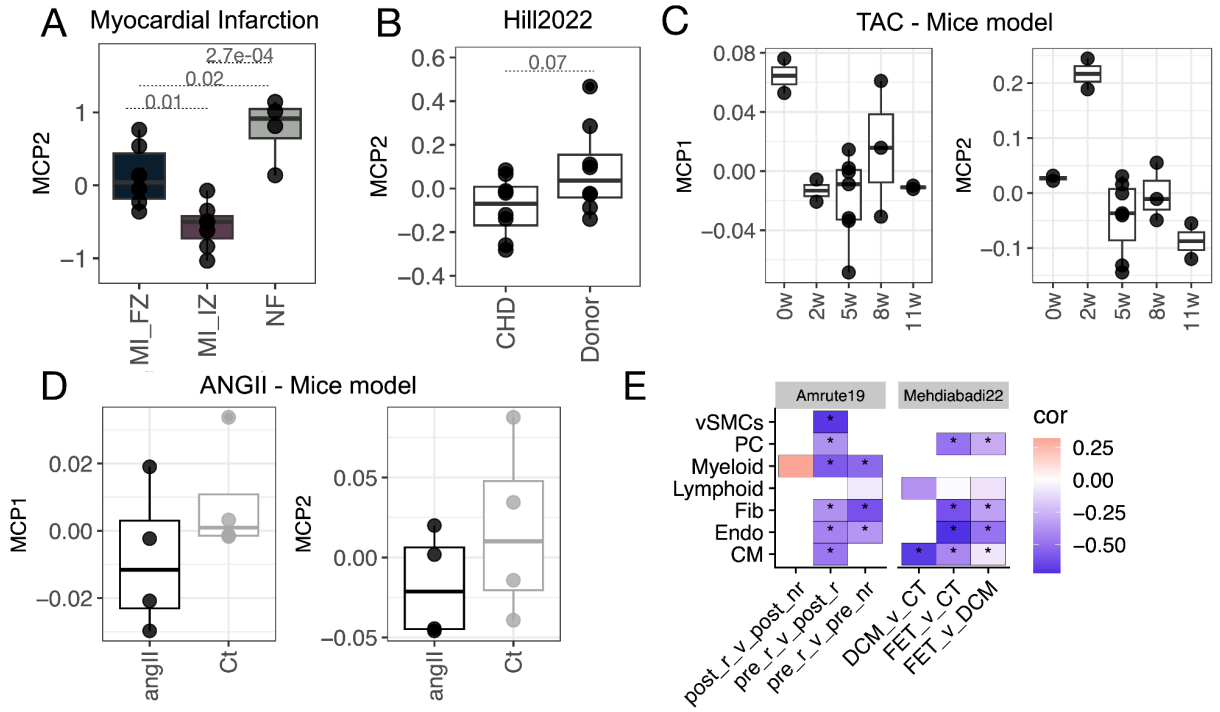
E. F1-scores for the prediction of cell type composition changes based on estimated compositions. Significant

composition changes of the seven cell lineages were classified into three categories, *increase*, *decrease* or *no change* based on a t-test p-value <0.05. Top barplot displays global average of F1 for each signature matrix.

F. Estimated cell type compositions from core HF bulk studies are compared for their clustering of HF status (left panel) and study label (right panel) via silhouette scores.

G. A linear mixed model was applied to meta-analyze composition changes of cell types in deconvoluted bulk studies. The variance explained by the study random effect (y-axis) is compared with the log-transformed p-value of the cell type fixed effect (x-axis).

H. Principal component (PC) analysis of cell type composition profile of single-nucleus (sc) and bulk cohort. PC one and two (with explained variance) on x and y-axi, respectively.



Supplementary Figure 11. Projection of tissue samples into the patient map of multicellular programs (MCPs) of HF.

A. MCP2 values of HF samples obtained from myocardial infarction (MI) patients at different time-points of the disease (FZ = fibrotic, IZ = ischemic) and control non-failing (NF) donor samples.

B. MCP2 values of HF samples obtained from congenital heart disease (CHD) and control non-failing donor patients.

C. MCP1 and 2 values of heart tissue samples obtained from mice with induced HF via transverse aortic constriction (TAC) at different time points.

D. MCP1 and 2 values of heart tissue samples obtained from mice with (Angiotensin II)-induced HF and control mice.

E. Correlating t-statistics of significantly deregulated genes reported by two individual studies per major cell lineage with gene loadings from the MCP1. Pearson correlation coefficients. * p-value <0.05. Since negative loadings are associated with HF, negative correlations indicate shared regulatory direction in HF.

10. Glossary

BH	Benjamini Hochberg
DCM	Dilated cardiomyopathy
HF	Heart failure
HCM	Hypertrophic cardiomyopathy
ICM	Ischemic cardiomyopathy
MCP	Multicellular program
NF	Nonfailing
Sn	Single-nucleus RNAseq

9. References

1. Litviňuková, M. *et al.* Cells of the adult human heart. *Nature* **588**, 466–472 (2020).
2. Kanemaru, K. *et al.* Spatially resolved multiomics of human cardiac niches. *Nature* **619**, 801–810 (2023).
3. Cranley, J. *et al.* Multiomic analysis reveals developmental dynamics of the human heart in health and disease. *BioRxiv* (2024) doi:10.1101/2024.04.29.591736.
4. Reichart, D. *et al.* Pathogenic variants damage cell composition and single cell transcription in cardiomyopathies. *Science* **377**, eabo1984 (2022).
5. Kuppe, C. *et al.* Spatial multi-omic map of human myocardial infarction. *Nature* **608**, 766–777 (2022).
6. Amrute, J. M. *et al.* Defining cardiac functional recovery in end-stage heart failure at single-cell resolution. *Nat. Cardiovasc. Res.* **2**, 399–416 (2023).
7. Koenig, A. L. *et al.* Single-cell transcriptomics reveals cell-type-specific diversification in human heart failure. *Nat. Cardiovasc. Res.* **1**, 263–280 (2022).
8. Chaffin, M. *et al.* Single-nucleus profiling of human dilated and hypertrophic cardiomyopathy. *Nature* **608**, 174–180 (2022).
9. Simonson, B. *et al.* Single-nucleus RNA sequencing in ischemic cardiomyopathy reveals common transcriptional profile underlying end-stage heart failure. *Cell Rep.* **42**, 112086 (2023).
10. Hill, M. C. *et al.* Integrated multi-omic characterization of congenital heart disease. *Nature* **608**, 181–191 (2022).
11. Xu, C. *et al.* Automatic cell-type harmonization and integration across Human Cell Atlas datasets. *Cell* **186**, 5876-5891.e20 (2023).
12. Li, M. *et al.* DISCO: a database of Deeply Integrated human Single-Cell Omics data. *Nucleic Acids Res.* **50**, D596–D602 (2022).

13. Dona, M. S. I. *et al.* CLARA: A web portal for interactive exploration of the cardiovascular cellular landscape in health and disease. *BioRxiv* (2021) doi:10.1101/2021.07.18.452862.

14. Ramirez Flores, R. O., Schäfer, P. S. L., Küchenhoff, L. & Saez-Rodriguez, J. Complementing Cell Taxonomies with a Multicellular Analysis of Tissues. *Physiology (Bethesda)* **39**, 0 (2024).

15. Adler, M., Chavan, A. R. & Medzhitov, R. Tissue biology: in search of a new paradigm. *Annu. Rev. Cell Dev. Biol.* **39**, 67–89 (2023).

16. Ramirez Flores, R. O. *et al.* Consensus Transcriptional Landscape of Human End-Stage Heart Failure. *J. Am. Heart Assoc.* **10**, e019667 (2021).

17. Flam, E. *et al.* Integrated landscape of cardiac metabolism in end-stage human nonischemic dilated cardiomyopathy. *Nat. Cardiovasc. Res.* **1**, 817–829 (2022).

18. Wang, C. *et al.* Alterations of Cardiac Protein Kinases in Cyclic Nucleotide-Dependent Signaling Pathways in Human Ischemic Heart Failure. *Front. Cardiovasc. Med.* **9**, 919355 (2022).

19. Rao, M. *et al.* Resolving the intertwining of inflammation and fibrosis in human heart failure at single-cell level. *Basic Res. Cardiol.* **116**, 55 (2021).

20. Forte, E. *et al.* Adult mouse fibroblasts retain organ-specific transcriptomic identity. *eLife* **11**, (2022).

21. Hua, X. *et al.* Multi-level transcriptome sequencing identifies COL1A1 as a candidate marker in human heart failure progression. *BMC Med.* **18**, 2 (2020).

22. Nicin, L. *et al.* A human cell atlas of the pressure-induced hypertrophic heart. *Nat. Cardiovasc. Res.* (2022) doi:10.1038/s44161-022-00019-7.

23. Liu, J. *et al.* Transcriptional and immune landscape of cardiac sarcoidosis. *Circ. Res.* **131**, 654–669 (2022).

24. Mehdiabadi, N. R. *et al.* Defining the Fetal Gene Program at Single-Cell Resolution in Pediatric Dilated Cardiomyopathy. *Circulation* **146**, 1105–1108 (2022).

25. McLellan, M. A. *et al.* High-Resolution Transcriptomic Profiling of the Heart During Chronic Stress

Reveals Cellular Drivers of Cardiac Fibrosis and Hypertrophy. *Circulation* **142**, 1448–1463 (2020).

26. Ren, Z. *et al.* Single-Cell Reconstruction of Progression Trajectory Reveals Intervention Principles in Pathological Cardiac Hypertrophy. *Circulation* **141**, 1704–1719 (2020).
27. Ramirez Flores, R. O., Lanzer, J. D., Dimitrov, D., Velten, B. & Saez-Rodriguez, J. Multicellular factor analysis of single-cell data for a tissue-centric understanding of disease. *eLife* **12**, (2023).
28. Li, W. *et al.* PLCE1 promotes myocardial ischemia-reperfusion injury in H/R H9c2 cells and I/R rats by promoting inflammation. *Biosci. Rep.* **39**, (2019).
29. Wang, X. *et al.* Downregulation of FKBP5 promotes atrial arrhythmogenesis. *Circ. Res.* **133**, e1–e16 (2023).
30. Molkentin, J. D. *et al.* A calcineurin-dependent transcriptional pathway for cardiac hypertrophy. *Cell* **93**, 215–228 (1998).
31. Shimoyama, M. *et al.* Calcineurin plays a critical role in pressure overload-induced cardiac hypertrophy. *Circulation* **100**, 2449–2454 (1999).
32. Cao, D. J. *et al.* Histone deacetylase (HDAC) inhibitors attenuate cardiac hypertrophy by suppressing autophagy. *Proc Natl Acad Sci USA* **108**, 4123–4128 (2011).
33. Gillette, T. G. HDAC inhibition in the heart: erasing hidden fibrosis. *Circulation* **143**, 1891–1893 (2021).
34. Xi, Y. *et al.* Ionic mechanisms underlying the effects of vasoactive intestinal polypeptide on canine atrial myocardium. *Circ. Arrhythm. Electrophysiol.* **6**, 976–983 (2013).
35. Szema, A. M., Hamidi, S. A., Smith, S. D. & Benveniste, H. VIP gene deletion in mice causes cardiomyopathy associated with upregulation of heart failure genes. *PLoS ONE* **8**, e61449 (2013).
36. Sun, H. & Wang, Y. Interferon regulatory factors in heart. *Hypertension* **63**, 663–664 (2014).
37. Browaeys, R. *et al.* MultiNicheNet: a flexible framework for differential cell-cell communication analysis from multi-sample multi-condition single-cell transcriptomics data. *BioRxiv* (2023)

doi:10.1101/2023.06.13.544751.

38. Münch, J. & Abdelilah-Seyfried, S. Sensing and responding of cardiomyocytes to changes of tissue stiffness in the diseased heart. *Front. Cell Dev. Biol.* **9**, 642840 (2021).
39. Taha, M. F., Valojerdi, M. R. & Mowla, S. J. Effect of bone morphogenetic protein-4 (BMP-4) on cardiomyocyte differentiation from mouse embryonic stem cell. *Int. J. Cardiol.* **120**, 92–101 (2007).
40. Grego-Bessa, J. *et al.* Nrg1 regulates cardiomyocyte migration and cell cycle in ventricular development. *Circ. Res.* **133**, 927–943 (2023).
41. Venugopal, H., Hanna, A., Humeres, C. & Frangogiannis, N. G. Properties and functions of fibroblasts and myofibroblasts in myocardial infarction. *Cells* **11**, (2022).
42. Oatmen, K. E., Cull, E. & Spinale, F. G. Heart failure as interstitial cancer: emergence of a malignant fibroblast phenotype. *Nat. Rev. Cardiol.* **17**, 523–531 (2020).
43. Bretherton, R., Bugg, D., Olszewski, E. & Davis, J. Regulators of cardiac fibroblast cell state. *Matrix Biol.* **91–92**, 117–135 (2020).
44. Rurik, J. G. *et al.* CAR T cells produced in vivo to treat cardiac injury. *Science* **375**, 91–96 (2022).
45. Fu, X. *et al.* Specialized fibroblast differentiated states underlie scar formation in the infarcted mouse heart. *The Journal of Clinical Investigation* (2018).
46. Unable to find information for 14940087.
47. Unable to find information for 15733960.
48. Luthria, G., Lauffenburger, D. & Miller, M. A. Cell-Cell Communication Networks in Tissue: Toward Quantitatively Linking Structure with Function. *Current Opinion in Systems Biology* (2021) doi:10.1016/j.coisb.2021.05.002.
49. Unable to find information for 7484681.
50. Dimitrov, D. *et al.* LIANA+ provides an all-in-one framework for cell-cell communication inference. *Nat. Cell Biol.* **26**, 1613–1622 (2024).

51. Silva, A. C., Pereira, C., Fonseca, A. C. R. G., Pinto-do-Ó, P. & Nascimento, D. S. Bearing my heart: the role of extracellular matrix on cardiac development, homeostasis, and injury response. *Front. Cell Dev. Biol.* **8**, 621644 (2020).

52. Miyara, S. *et al.* Circuit to target approach defines an autocrine myofibroblast loop that drives cardiac fibrosis. *BioRxiv* (2023) doi:10.1101/2023.01.01.522422.

53. Adler, M., Korem Kohanim, Y., Tendler, A., Mayo, A. & Alon, U. Continuum of Gene-Expression Profiles Provides Spatial Division of Labor within a Differentiated Cell Type. *Cell Syst.* **8**, 43-52.e5 (2019).

54. Baghdassarian, H. M. & Lewis, N. E. Resource allocation in mammalian systems. *Biotechnol. Adv.* **71**, 108305 (2024).

55. Lanzer, J. D. *et al.* Single-cell transcriptomics reveal distinctive patterns of fibroblast activation in heart failure with preserved ejection fraction. *Basic Res. Cardiol.* (2024) doi:10.1007/s00395-024-01074-w.

56. Matsuo, N. *et al.* Collagen XXIV (Col24a1) gene expression is a specific marker of osteoblast differentiation and bone formation. *Connect. Tissue Res.* **49**, 68–75 (2008).

57. Hahn, V. S. *et al.* Myocardial gene expression signatures in human heart failure with preserved ejection fraction. *Circulation* **143**, 120–134 (2021).

58. Verdonschot, J. A. J. *et al.* Clustering of cardiac transcriptome profiles reveals unique: subgroups of dilated cardiomyopathy patients. *JACC Basic Transl. Sci.* **8**, 406–418 (2023).

59. Durinck, S., Spellman, P. T., Birney, E. & Huber, W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc.* **4**, 1184–1191 (2009).

60. Hao, Y. *et al.* Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat. Biotechnol.* **42**, 293–304 (2024).

61. Germain, P.-L., Lun, A., Garcia Meixide, C., Macnair, W. & Robinson, M. D. Doublet identification

in single-cell sequencing data using scDblFinder. *F1000Res.* **10**, 979 (2021).

62. Korsunsky, I. *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
63. Torre, D., Lachmann, A. & Ma'ayan, A. BioJupies: Automated Generation of Interactive Notebooks for RNA-Seq Data Analysis in the Cloud. *Cell Syst.* **7**, 556-561.e3 (2018).
64. Schröder, M. S., Culhane, A. C., Quackenbush, J. & Haibe-Kains, B. survcomp: an R/Bioconductor package for performance assessment and comparison of survival models. *Bioinformatics* **27**, 3206–3208 (2011).
65. van den Boogaart, K. G. & Tolosana-Delgado, R. "compositions": A unified R package to analyze compositional data. *Comput. Geosci.* **34**, 320–338 (2008).
66. Argelaguet, R. *et al.* MOFA+: a statistical framework for comprehensive integration of multi-modal single-cell data. *Genome Biol.* **21**, 111 (2020).
67. Argelaguet, R. *et al.* Multi-Omics Factor Analysis-a framework for unsupervised integration of multi-omics data sets. *Mol. Syst. Biol.* **14**, e8124 (2018).
68. Penrose, R. & Todd, J. A. A generalized inverse for matrices. *Math. Proc. Camb. Phil. Soc.* **51**, 406 (1955).
69. Liberzon, A. *et al.* The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015).
70. Badia-I-Mompel, P. *et al.* decoupleR: ensemble of computational methods to infer biological activities from omics data. *Bioinformatics Advances* **2**, vbac016 (2022).
71. Litvinukova, M. *et al.* Cells and gene expression programs in the adult human heart. *BioRxiv* (2020) doi:10.1101/2020.04.03.024075.
72. Newman, A. M. *et al.* Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* **12**, 453–457 (2015).

73. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).

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

75. Jiang, P. *et al.* Systematic investigation of cytokine signaling activity at the tissue and single-cell levels. *Nat. Methods* **18**, 1181–1191 (2021).

76. Schubert, M. *et al.* Perturbation-response genes reveal signaling footprints in cancer gene expression. *Nat. Commun.* **9**, 20 (2018).

77. Jaeger, B. C., Edwards, L. J., Das, K. & Sen, P. K. AnR^2 statistic for fixed effects in the generalized linear mixed model. *J. Appl. Stat.* **44**, 1086–1105 (2017).

78. Fleming, S. J. *et al.* Unsupervised removal of systematic background noise from droplet-based single-cell experiments using CellBender. *Nat. Methods* **20**, 1323–1335 (2023).