

Article | Published: 22 June 2022

Single-nucleus profiling of human dilated and hypertrophic cardiomyopathy

Mark Chaffin, Irinna Papangeli, Bridget Simonson, Amer-Denis Akkad, Matthew C. Hill, Alessandro Arduini, Stephen J. Fleming, Michelle Melanson, Sikander Hayat, Maria Kost-Alimova, Ondine Atwa, Jiangchuan Ye, Kenneth C. Bedi Jr, Matthias Nahrendorf, Virendar K. Kaushik, Christian M. Stegmann, Kenneth B. Margulies, Nathan R. Tucker & Patrick T. Ellinor 

Nature (2022)

4960 Accesses | 142 Altmetric | [Metrics](#)

Abstract

Heart failure encompasses a heterogeneous set of clinical features that converge on impaired cardiac contractile function^{1,2} and presents a growing public health concern. Previous work has highlighted changes in both transcription and protein expression in failing hearts^{3,4}, but may overlook molecular changes in less prevalent cell types. Here we identify extensive molecular alterations in failing hearts at single-cell resolution by performing single-nucleus RNA sequencing of nearly 600,000 nuclei in left ventricle samples from 11 hearts with dilated cardiomyopathy and 15 hearts with hypertrophic cardiomyopathy as well as 16 non-failing hearts. The transcriptional profiles of dilated or hypertrophic cardiomyopathy hearts broadly converged at the tissue and cell-type level. Further, a subset of hearts from patients with cardiomyopathy harbour a unique population of activated fibroblasts that is almost entirely absent from non-failing samples. We performed a CRISPR-knockout screen in primary human cardiac fibroblasts to evaluate this fibrotic cell state transition; knockout of genes associated with

fibroblast transition resulted in a reduction of myofibroblast cell-state transition upon TGFβ1 stimulation for a subset of genes. Our results provide insights into the transcriptional diversity of the human heart in health and disease as well as new potential therapeutic targets and biomarkers for heart failure.

This is a preview of subscription content

Access options

Subscribe to Nature+

Get immediate online access to the entire Nature family of 50+ journals

\$29.99

monthly

Subscribe

Buy article

Get time limited or full article access on ReadCube.

\$32.00

Buy

All prices are NET prices.

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

Subscribe

All prices are NET prices.
VAT will be added later in the checkout.

Additional access options:

- [Log in](#)
- [Learn about institutional subscriptions](#)

Data availability

Processed single-nucleus transcriptomic data are available through the Broad Institute's Single Cell Portal (https://singlecell.broadinstitute.org/single_cell) under project ID SCP1303 (https://singlecell.broadinstitute.org/single_cell/study/SCP1303/). Raw sequence data are available to authorized users through dbGaP (the database of Genotypes and Phenotypes) accession number phs001539. Use of the raw sequencing data is limited to health, medical and/or biomedical purposes, including methods development research. [Source data](#) are provided with this paper.

References

1. Savarese, G. & Lund, L. H. Global Public Health Burden of Heart Failure. *Card. Fail. Rev.* **03**, 7 (2017).
 2. Oktay, A. A. et al. *Diabetes, Cardiomyopathy, and Heart Failure. Endotext* (MDText.com, Inc., 2000).
 3. Liu, Y. et al. RNA-Seq identifies novel myocardial gene expression signatures of heart failure. *Genomics* **105**, 83–89 (2015).
 4. Chen, C. Y. et al. Suppression of detyrosinated microtubules improves cardiomyocyte function in human heart failure. *Nat. Med.* **24**, 1225–1233 (2018).
-

5. Maron, B.J. et al. Contemporary definitions and classification of the cardiomyopathies: An American Heart Association Scientific Statement from the Council on Clinical Cardiology, Heart Failure and Transplantation Committee; Quality of Care and Outcomes Research and Functional Genomics and Translational Biology Interdisciplinary Working Groups; and Council on Epidemiology and Prevention. *Circulation* **113**, 1807–1816 (2006).

6. Sweet, M. E. et al. Transcriptome analysis of human heart failure reveals dysregulated cell adhesion in dilated cardiomyopathy and activated immune pathways in ischemic heart failure. *BMC Genomics* **19**, 812 (2018).

7. Tucker, N. R. et al. Transcriptional and Cellular Diversity of the Human Heart. *Circulation* **142**, 466–482 (2020).

8. Litviňuková, M. et al. Cells of the adult human heart. *Nature* **588**, 466–472 (2020).

9. Wang, L. et al. Single-cell reconstruction of the adult human heart during heart failure and recovery reveals the cellular landscape underlying cardiac function. *Nat. Cell Biol.* **22**, 108–119 (2020).

10. Ashburner, M. et al. Gene ontology: Tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000).

11. Carbon, S. et al. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res.* **47**, D330–D338 (2019).

12. Kalucka, J. et al. Single-Cell Transcriptome Atlas of Murine Endothelial Cells. *Cell* **180**, 764–779.e20 (2020).

13. Crinier, A. et al. High-Dimensional Single-Cell Analysis Identifies Organ-Specific Signatures and Conserved NK Cell Subsets in Humans and Mice. *Immunity* **49**, 971–986.e5 (2018).

14. Tirosh, I. et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* **352**, 189–196 (2016).

15. Bajpai, G. et al. The human heart contains distinct macrophage subsets with divergent origins and functions. *Nat. Med.* **24**, 1234–1245 (2018).

16. Tallquist, M. D. & Molkentin, J. D. Redefining the identity of cardiac fibroblasts. *Nat. Rev. Cardiol.* **14**, 484–491 (2017).

17. Cucoranu, I. et al. NAD(P)H oxidase 4 mediates transforming growth factor- β 1-induced differentiation of cardiac fibroblasts into myofibroblasts. *Circ. Res.* **97**, 900–907 (2005).

18. Tillmanns, J. et al. Fibroblast activation protein alpha expression identifies activated fibroblasts after myocardial infarction. *J. Mol. Cell. Cardiol.* **87**, 194–203 (2015).

19. Shinde, A. V. & Frangogiannis, N. G. Mechanisms of Fibroblast Activation in the Remodeling Myocardium. *Curr. Pathobiol. Rep.* **5**, 145–152 (2017).

20. 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).

21. Forte, E. et al. Dynamic Interstitial Cell Response during Myocardial Infarction Predicts Resilience to Rupture in Genetically Diverse Mice. *Cell Rep.* **30**, 3149–3163 (2020).

22. Frolova, E. G. et al. Thrombospondin-4 regulates fibrosis and remodeling of the myocardium in response to pressure overload. *FASEBJ.* **26**, 2363–2373 (2012).

23. Liu, X. et al. Long non-coding and coding RNA profiling using strand-specific RNA-seq in human hypertrophic cardiomyopathy. *Sci. Data* **6**, 90 (2019).

24. Bengtsson, E. et al. The Leucine-rich Repeat Protein PRELP Binds Perlecan and Collagens and May Function as a Basement Membrane Anchor. *J. Biol. Chem.* **277**, 15061–15068 (2002).

25. Li, L. et al. The role of JAZF1 on lipid metabolism and related genes in vitro. *Metabolism.* **60**, 523–530 (2011).

26. Guang-feng, M. et al. JAZF1 can regulate the expression of lipid metabolic genes and inhibit lipid accumulation in adipocytes. *Biochem. Biophys. Res. Commun.* **445**, 673–680 (2014).

27. Yuan, L. et al. Transcription factor TIP27 regulates glucose homeostasis and insulin sensitivity in a PI3-kinase/Akt-dependent manner in mice. *Int. J. Obes.* **39**, 949–958 (2015).

28. Koch, M. et al. A Novel Marker of Tissue Junctions, Collagen XXII. *J. Biol. Chem.* **279**, 22514 (2004).

29. Watanabe, T. et al. A Human Skin Model Recapitulates Systemic Sclerosis Dermal Fibrosis and Identifies COL22A1 as a TGF β Early Response Gene that Mediates Fibroblast to Myofibroblast Transition. *Genes (Basel)*. **10**, 75 (2019).
-
30. Ma, Y. et al. Cardiomyocyte d-dopachrome tautomerase protects against heart failure. *JCI Insight* **4**, e128900 (2019).
-
31. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10 (2011).
-
32. Fleming, S. J., Marioni, J. C. & Babadi, M. CellBender remove-background: A deep generative model for unsupervised removal of background noise from scRNA-seq datasets. Preprint at *bioRxiv* <https://doi.org/10.1101/791699> (2019).
-
33. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: Large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).
-
34. Korsunsky, I. et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
-
35. Wolock, S. L., Lopez, R. & Klein, A. M. Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. *Cell Syst.* **8**, 281–291.e9 (2019).
-
36. Lun, A. T. L. & Marioni, J. C. Overcoming confounding plate effects in differential expression analyses of single-cell RNA-seq data. *Biostatistics* **18**, 451–464 (2017).
-
37. Chen, Y., Lun, A. T. L. & Smyth, G. K. From reads to genes to pathways: Differential expression analysis of RNA-Seq experiments using Rsubread and the edgeR quasi-

likelihood pipeline. *F1000Research* **5**, 1438 (2016).

38. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

39. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. Voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15**, R29 (2014).

40. Falcon, S. & Gentleman, R. Using GOSTATS to test gene lists for GO term association. *Bioinformatics* **23**, 257–258 (2007).

41. Büttner, M., Ostner, J., Müller, C. L., Theis, F. J. & Schubert, B. scCODA is a Bayesian model for compositional single-cell data analysis. *Nat. Commun.* **12**, 1–10 (2021).

42. Brill, B., Amir, A. & Heller, R. Testing for differential abundance in compositional counts data, with application to microbiome studies. Preprint at *arXiv* <https://arxiv.org/abs/1904.08937> (2019).

43. Jassal, B. et al. The reactome pathway knowledgebase. *Nucleic Acids Res.* **48**, D498–D503 (2020).

44. Korotkevich, G., Sukhov, V. & Sergushichev, A. Fast gene set enrichment analysis. Preprint at *bioRxiv* <https://doi.org/10.1101/060012> (2016).

45. Yu, G. & He, Q. Y. ReactomePA: An R/Bioconductor package for reactome pathway analysis and visualization. *Mol. Biosyst.* **12**, 477–479 (2016).

46. Street, K. et al. Slingshot: cell lineage and pseudotime inference for single-cell

transcriptomics. *BMC Genomics* **19**, 477 (2018).

47. Hafemeister, C. & Satija, R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol.* **20**, 296 (2019).

48. La Manno, G. et al. RNA velocity of single cells. *Nature* **560**, 494–498 (2018).

49. Bergen, V., Lange, M., Peidli, S., Wolf, F. A. & Theis, F. J. Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat. Biotechnol.* **38**, 1408–1414 (2020).

50. Van den Berge, K. et al. Trajectory-based differential expression analysis for single-cell sequencing data. *Nat. Commun.* **11**, 1201 (2020).

51. Newman, A. M. et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat. Biotechnol.* **37**, 773–782 (2019).

52. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).

53. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. Preprint at *arXiv* <https://arxiv.org/abs/1303.3997> (2013).

54. Poplin, R. et al. Scaling accurate genetic variant discovery to tens of thousands of samples. Preprint at *bioRxiv* <https://doi.org/10.1101/201178> (2017).

55. McLaren, W. et al. The Ensembl Variant Effect Predictor. *Genome Biol.* **17**, 122 (2016).
-
56. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
-
57. Sanjana, N. E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat. Methods* **11**, 783–784 (2014).
-
58. Sanson, K. R. et al. Optimized libraries for CRISPR-Cas9 genetic screens with multiple modalities. *Nat. Commun.* **9**, 1–15 (2018).
-
59. Doench, J. G. et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.* **34**, 184–191 (2016).
-

Acknowledgements

The Precision Cardiology Laboratory is a joint effort between the Broad Institute and Bayer AG. This work was supported by the Fondation Leducq (14CVD01), and by grants from the National Institutes of Health to P.T.E. (1R01HL092577, R01HL128914 and K24HL105780), N.R.T. (5K01HL140187), K.B.M. (1R01HL105993) and M.C.H. (5T32HL007208-42). This work was also supported by a grant from the American Heart Association Strategically Focused Research Networks to P.T.E. (18SFRN34110082).

Author information

Authors and Affiliations

Precision Cardiology Laboratory and the Cardiovascular Disease Initiative, The Broad Institute, Cambridge, MA, USA
Mark Chaffin, Bridget Simonson, Matthew C. Hill, Alessandro Arduini, Stephen J. Fleming, Ondine Atwa, Jiangchuan Ye & Patrick T. Ellinor

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

Irinna Papangeli, Amer-Denis Akkad, Sikander Hayat & Christian M. Stegmann

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

Matthew C. Hill, Matthias Nahrendorf & Patrick T. Ellinor

Data Sciences Platform, The Broad Institute, Cambridge, MA, USA

Stephen J. Fleming

Center for the Development of Therapeutics, The Broad Institute, Cambridge, MA, USA

Michelle Melanson, Maria Kost-Alimova & Virendar K. Kaushik

Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

Kenneth C. Bedi Jr & Kenneth B. Margulies

Center for Systems Biology, Department of Radiology, Massachusetts General Hospital, Boston, MA, USA

Matthias Nahrendorf

Masonic Medical Research Institute, Utica, NY, USA

Nathan R. Tucker

Demoulas Center for Cardiac Arrhythmias, Massachusetts General Hospital, Boston, MA, USA

Patrick T. Ellinor

Contributions

M.C., N.R.T., C.M.S. and P.T.E. conceived and designed the study. I.P., B.S., A.-D.A. and A.A. generated snRNA-seq data. M.C. and S.J.F. analysed the snRNA-seq data. I.P., M.C.H., S.H., J.Y., M.N. and P.T.E. interpreted snRNA-seq analyses. B.S. performed RNA in situ hybridization and immunofluorescence staining. M.M., M.K.-A., M.C.H., O.A. and V.K.K. developed, executed, validated and interpreted cardiac fibrosis cellular assay. K.C.B. and K.B.M. acquired LV samples for this study. M.C., I.P., B.S. and P.T.E. drafted the manuscript, which was revised by all authors.

Corresponding author

Correspondence to [Patrick T. Ellinor](#).

Ethics declarations

Competing interests

A.-D.A. is an employee of Bayer US LLC (a subsidiary of Bayer AG) and may own stock in Bayer AG. I.P., S.H. and C.M.S. were full-time employees of Bayer when this work was

performed. C.M.S. is now a full-time employee of Vifor Pharma Management. V.K.K. was a co-founder of Dogma Therapeutics and Apolo1 Bio, and is now an employee of Anji Pharmaceuticals. P.T.E. has received sponsored research support from Bayer AG and IBM Health, and he has served on advisory boards or consulted for Bayer AG, MyoKardia and Novartis. K.B.M. has research grant funding from Sanofi-Aventis, USA and has also served on advisory boards for MyoKardia and Pfizer. M.N. has received funds or material research support from Lilly, Alnylam, Biotronik, CSL Behring, GlycoMimetics, GSK, Medtronic, Novartis and Pfizer, as well as consulting fees from Biogen, Gimv, IFM Therapeutics, Molecular Imaging, Sigilon and Verseau Therapeutics. The Broad Institute, Massachusetts General Hospital and Bayer AG have a pending patent application (US application no. 17/669,904) on the identification of activated fibroblasts for the diagnosis and treatment of cardiomyopathy based on the results from this work with M.C., I.P., A.-D.A., C.M.S. and P.T.E. listed as co-inventors. All remaining authors declare no competing interests.

Peer review

Peer review information

Nature thanks Eva van Rooij and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer review reports are available.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Sample quality control assessment.

a, Distribution of key quality control metrics from CellRanger count as boxplots stratified by non-failing (NF, $n = 32$), dilated cardiomyopathy (DCM, $n = 24$), and hypertrophic cardiomyopathy (HCM, $n = 32$). Cutoffs for low quality are shown by the dashed line, and samples that fail on a given metric are labeled. Center line, median; box limits, upper and lower quartiles; whiskers, 1.5x interquartile range; points, outliers. **b**, Exemplar high quality UMI (unique molecular identifier) decay curves (green) and low quality UMI decay curves (red) are shown. Cell barcodes are ranked by the total UMI. **c**, Total counts from Y chromosome transcripts in phenotypically classified females (red) and males (blue) across all samples ($n = 88$).

[Source data](#)

Extended Data Fig. 2 Nuclei quality control assessment.

a, Uniform manifold and projection (UMAP) representation of all CellBender non-empty droplets ($n = 885,944$), colored by Leiden clustering. **b**, Distribution of the median of four quality control metrics across clusters ($n = 47$) in **a** as boxplots, including percent of unique molecular identifiers mapping to mitochondrial genes (%MT), fraction of reads mapping exclusively to exons (Exonic fraction), entropy, and Scrublet estimated doublet score. Outlier clusters removed based on the criteria are highlighted in red. Center line, median; box limits, upper and lower quartiles; whiskers, 1.5x interquartile range; points, outliers. **c**, Distribution of number of unique molecular identifiers (nUMI), number of unique genes (nGene), %MT, and entropy across nuclei of each unique cell type ($n = 13$). Center line, median; box limits, upper and lower quartiles; whiskers, 1.5x interquartile range. **d**, UMAP representation of non-empty droplets after removal of low-quality clusters identified in **a** and **b** ($n = 745,342$). Additional low-quality nuclei as detected per cluster and per-sample are colored in red. **e**, Proportion of each sample (top, $n = 80$) and cluster (bottom, $n = 47$) removed during the quality control procedure. **f**, UMAP representation of non-empty droplets after removal of low-quality clusters and per-cluster quality control ($n = 605,314$). Red nuclei were deemed as misclassified or low-quality nuclei based on sub-cluster analysis within each cluster. **g**, The average score for each sub-cluster (x-axis) based on marker genes for each major cell type (y-axis), see [Methods](#). When a sub-cluster scores highly for an unrelated cell

type (black border), it was removed. CM, Cardiomyocyte; FB, Fibroblast; EC, Endothelial; PC, Pericyte; MP, Macrophage; VSMC, Vascular smooth muscle cell; LC, Lymphocyte; EndoC, Endocardial; AD, Adipocyte; NRN, Neuronal; LEC, Lymphatic endothelial; ActFB, Activated fibroblast; MC, Mast Cell; ProfMP, Proliferating macrophage; EpiC, Epicardial; MT, Mitochondrial.

[Source data](#)

Extended Data Fig. 3 Marker genes and cell type clustering.

a, Dot plots showing the expression profile of the top 4 marker genes for each cell type in single nuclei RNA-sequencing data ($n = 592,689$). The size of the dot reflects the percent of nuclei expressing the gene at non-zero levels and the shade reflects the mean log-normalized expression. **b**, Hierarchical clustering of cell types and expression profiles of the top 2,000 most highly variable genes. Avg Expr, Average log-normalized expression; Pct Nuclei Expr > 0, Percent of nuclei expressing the gene at non-zero levels; VSMC, Vascular smooth muscle cell.

[Source data](#)

Extended Data Fig. 4 Compositional changes of global cell types.

Proportion of each cell type across patients stratified by disease status. Statistically credible changes in cell types comparing dilated cardiomyopathy ($n = 11$) or hypertrophic cardiomyopathy ($n = 15$) to non-failing patients ($n = 16$), as tested with *scCODA*, are denoted with a *, see [Methods](#). Less prevalent cell types are shown in inset to improve readability. Center line, median; box limits, upper and lower quartiles; whiskers, 1.5x interquartile range; VSMC, vascular smooth muscle cell.

[Source data](#)

Extended Data Fig. 5 Sample-level principal component analysis by cell type.

Principal component analysis (PCA) using the top 500 most highly variables genes after summing expression counts for each sample ($n_{NF} = 16$, $n_{DCM} = 11$, $n_{HCMrEF} = 7$, $n_{HCMpEF} = 8$), for pseudo-bulk (**a**), cardiomyocyte (**b**), fibroblast (**c**), endothelial cell I (**d**), pericyte (**e**), macrophage (**f**), vascular smooth muscle cell (VSMC) (**g**), and lymphocyte (**h**). The

analysis was performed across all samples (left) and restricted to cardiomyopathy samples (right). The percent of total variation for each principal component (PC) is shown in parentheses. DCM, Dilated cardiomyopathy; HCMrEF, Hypertrophic cardiomyopathy with reduced ejection fraction; HCMpEF, Hypertrophic cardiomyopathy with preserved ejection fraction; LVEF, left ventricular ejection fraction.

[Source data](#)

Extended Data Fig. 6 Whole genome sequencing analysis of cardiomyopathy patients.

a, Number of deleterious variant carriers of known clinical cardiomyopathy testing panel genes by disease state ($n_{\text{DCM}} = 10$, $n_{\text{HCM}} = 15$, $n_{\text{NF}} = 15$). Grey indicates patients with no loss-of-function (LOF) or pathogenic variant in known cardiomyopathy genes. **b, c, d** Volcano plots displaying the log fold-change ($\log(\text{FC})$) and two-sided p -value from a limma-voom differential expression analysis between dilated cardiomyopathy (DCM) *TTN* LOF carriers and DCM non-carriers ($n = 4$ vs 6) (**b**), hypertrophic cardiomyopathy (HCM) *MYBPC3* LOF carriers and HCM non-carriers ($n = 3$ vs 12) (**c**), and HCM *MYH7* pathogenic variant carriers and HCM non-carriers ($n = 5$ vs 10) (**d**) based on CellBender remove-background counts. Dots are colored by cell type with outlined dots representing genes with $\text{FDR} < 0.01$ as calculated with the Benjamini-Hochberg procedure. Only genes deemed to have a low probability of background contamination are displayed. VSMC, vascular smooth muscle cell.

[Source data](#)

Extended Data Fig. 7 Sub-clustering of abundant cell types.

a-f, Sub-clustering results for cardiomyocytes ($n = 158,469$) (**a**), fibroblasts ($n = 147,219$) (**b**), endothelial cells ($n = 112,977$) (**c**), pericytes ($n = 69,304$) (**d**), vascular smooth muscle cells (VSMC, $n = 18,137$) (**e**), and lymphocytes ($n = 16,246$) (**f**). Uniform manifold approximation and projection (UMAP) visualization colored by Leiden clusters is shown on the left. The distribution of sub-populations across patients by disease status represented as box plots with statistically credible changes indicated with a * (middle,

see [Methods](#)). Sub-population labels are colored as in the UMAP visualization. Center line, median; box limits, upper and lower quartiles; whiskers, 1.5x interquartile range; points, outliers. Dot plots of the most selective markers for each sub-population compared to all other sub-populations (right). The size of the dot reflects the percent of nuclei expressing the gene at non-zero levels and the shade reflects the mean log-normalized expression. Avg norm expr, Average log-normalized expression scaled to the max value for each gene; Pct Nuclei Expr > 0, Percent of nuclei expressing the gene at non-zero levels; Act. FB, Activated fibroblast; L-EC, Lymphatic endothelial.

[Source data](#)

Extended Data Fig. 8 Macrophage sub-populations.

a, Uniform manifold approximation and projection (UMAP) representation of 53,730 nuclei classified as macrophage or proliferating macrophage in the global analysis colored by sub-population. **b**, UMAP plot with the S phase cell cycle score overlaid (see [Methods](#)). **c**, UMAP plot with the G2M phase cell cycle score overlaid (see [Methods](#)). **d**, Expression of markers for each sub-population. The size of each dot reflects the percent of nuclei expressing the gene at non-zero levels and the shade reflects the mean log-normalized expression. **e**, Distribution of macrophage sub-populations across patients ($n_{\text{DCM}} = 11$, $n_{\text{HCM}} = 15$, $n_{\text{NF}} = 16$) by disease status, with statistically credible differences denoted with a *, see [Methods](#). Center line, median; box limits, upper and lower quartiles; whiskers, 1.5x interquartile range; points, outliers. **f**, Representative immunofluorescence staining for macrophage marker CD163, cycling marker MKI67, and nuclei with DAPI from single tissue sections of four patients. The percent of macrophages with MKI67 expression across full tissue sections are shown on each image. **g**, Dendrogram demonstrating similarity of sub-population centroids based on the top 2000 most highly variable genes using Euclidean distance and the Ward method. **h-i**, Expression of marker genes for CCR2 positive (CCR2+) (**h**) and CCR2 negative (CCR2-) (**i**) cardiac macrophages obtained from Bajpai et al., 2018¹⁵. The size of each square represents the percent of nuclei expressing the gene at non-zero levels while the shade represents a log fold-change ($\log(\text{FC})$) estimate comparing the expression in the given sub-population to all other sub-populations. Genes significantly

up- or down-regulated in a sub-population (FDR < 0.01, as calculated with the Benjamini-Hochberg procedure) are shown with a black border. M ϕ , Macrophage; NF, Non-failing; DCM, Dilated cardiomyopathy; HCM, Hypertrophic cardiomyopathy; Pct nuclei expr > 0; Avg norm expr, Average normalized expression scaled to the max expression for each gene across all sub-populations.

[Source data](#)

Extended Data Fig. 9 Imaging validation.

a, Additional immunofluorescence staining of four patients for macrophage marker CD163 and cycling marker MKI67 from single tissue sections of each patient shown in Extended Data Fig. 8f. Images are displayed both with, and without, DAPI to allow better visualizing of cells co-expressing CD163 and MKI67. **b**, in situ hybridization with RNAscope showing localization of canonical fibroblast marker, *DCN* (green/blue), and activated fibroblast marker, *COL22A1* (red), across several images of a single section from each of four patients: dilated cardiomyopathy patient P1304, hypertrophic cardiomyopathy patient P1425, and non-failing patients P1515 and P1516. Nuclear localization is shown with hematoxylin (blue). Negative control sections are shown in the second row for each sample. NF, non-failing; DCM, dilated cardiomyopathy; HCM, hypertrophic cardiomyopathy.

Extended Data Fig. 10 Validation of computational deconvolution analysis for activated fibroblasts.

a, *in situ* hybridization with RNAscope showing localization of canonical fibroblast marker *DCN* (green/blue), and activated fibroblast marker *COL22A1* (red), across single sections derived from 6 patients. Nuclear localization is shown with hematoxylin (blue). The computationally predicted percent of activated fibroblasts from CIBERSORTx (% Act. FB) for each patient is shown above their respective image. **b**, **c**, Uniform manifold and projection representation of all nuclei from both the primary analysis ($n_{\text{patient}} = 42$) and validation single-nuclei RNA-sequencing ($n_{\text{patient}} = 3$) (**b**), and separately for validation samples alone (**c**). The total number of patients (n_{patients}) and total nuclei (n_{nuclei}) are shown above each respective figure. **d**, Relative contribution of fibroblasts

and activated fibroblasts to each primary analysis patient (left) and validation patients only (right). DCM, Dilated cardiomyopathy; HCM, Hypertrophic cardiomyopathy; NF, Non-failing; VSMC, Vascular smooth muscle cell.

[Source data](#)

Extended Data Fig. 11 Replication of cardiac fibroblast activation assay across up to five screens.

The percent change in the fraction of myofibroblasts across up to five screens for each sgRNA (independent sgRNAs denoted as “G”). Within each screen, well values were normalized to the median effect of *TGFBR1* sgRNAs and non-targeting control sgRNAs. The specific screen replicate (Rep.) is shown above each respective bar where “P” indicates the primary screen, and r1-r4 indicate replicate screens 1–4. The number of wells included in each screen (Nwell) are displayed below each bar with individual dots displayed for each well value. Error bars depict standard error to the mean.

[Source data](#)

Extended Data Fig. 12 Global expression profiles of genes included in the cardiac fibroblast activation assay.

Dot plot representation of expression profiles ($n_{\text{patient}} = 42$; $n_{\text{nuclei}} = 592,689$) of all genes included in cardiac fibroblast activation assay. The size of the dot reflects the percent of nuclei expressing the gene at non-zero levels and the shade reflects the mean log-normalized expression. Avg expr, Average log-normalized expression; Pct nuclei expr > 0, Percent of nuclei expressing the gene at non-zero levels; VSMC, Vascular smooth muscle cell.

[Source data](#)

Supplementary information

[Reporting Summary](#)

[Peer Review File](#)

Supplementary Tables

This folder contains Supplementary Tables 1–16 and a DOC file containing their accompanying legends.

Source data

Source Data Fig. 1

Source Data Fig. 2

Source Data Fig. 3

Source Data Fig. 4

Source Data Extended Data Fig. 1

Source Data Extended Data Fig. 2

Source Data Extended Data Fig. 3

Source Data Extended Data Fig. 4

Source Data Extended Data Fig. 5

Source Data Extended Data Fig. 6

Source Data Extended Data Fig. 7

Source Data Extended Data Fig. 8

Source Data Extended Data Fig. 10

Source Data Extended Data Fig. 11

Source Data Extended Data Fig. 12

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Chaffin, M., Papangeli, I., Simonson, B. *et al.* Single-nucleus profiling of human dilated and hypertrophic cardiomyopathy. *Nature* (2022). <https://doi.org/10.1038/s41586-022-04817-8>

Received

24 February 2021

Accepted

27 April 2022

Published

22 June 2022

DOI

<https://doi.org/10.1038/s41586-022-04817-8>

Subjects

[Cardiomyopathies](#)

• [Gene expression](#)

• [Transcriptomics](#)

Nature (*Nature*) | ISSN 1476-4687 (online) | ISSN 0028-0836 (print)