GENETICS

Cardiac cell type–specific gene regulatory programs and disease risk association

James D. Hocker[1,2,3], Olivier B. Poirion[4†], Fugui Zhu[5†], Justin Buchanan[4], Kai Zhang[3], Joshua Chiou[2,6], Tsui-Min Wang[7], Qingquan Zhang[5], Xiaomeng Hou[4], Yang E. Li[3], Yanxiao Zhang[3], Elie N. Farah[2,5], Allen Wang[4], Andrew D. McCulloch[7,8], Kyle J. Gaulton[6,9], Bing Ren[3,4,9,10]*, Neil C. Chi[5,10]*, Sebastian Preissl[4]*

Misregulated gene expression in human hearts can result in cardiovascular diseases that are leading causes of mortality worldwide. However, the limited information on the genomic location of candidate cis-regulatory elements (cCREs) such as enhancers and promoters in distinct cardiac cell types has restricted the understanding of these diseases. Here, we defined >287,000 cCREs in the four chambers of the human heart at single-cell resolution, which revealed cCREs and candidate transcription factors associated with cardiac cell types in a region-dependent manner and during heart failure. We further found cardiovascular disease–associated genetic variants enriched within these cCREs including 38 candidate causal atrial fibrillation variants localized to cardiomyocyte cCREs. Additional functional studies revealed that two of these variants affect a cCRE controlling *KCNH2/HERG* expression and action potential repolarization. Overall, this atlas of human cardiac cCREs provides the foundation for illuminating cell type–specific gene regulation in human hearts during health and disease.

INTRODUCTION

Disruption of gene regulation is an important contributor to cardiovascular disease, the leading cause of morbidity and mortality worldwide (*1*). Cis-regulatory elements such as enhancers and promoters are crucial for regulating gene expression (*2–5*). Mutations in transcription factors and chromatin regulators can result in heart disease (*6*, *7*), and genetic variants associated with risk of cardiovascular disease are enriched within annotated candidate cis-regulatory elements (cCREs) in the human genome (*8*). However, a major barrier to understanding the genetic and molecular basis of cardiovascular diseases is the paucity of maps and tools to interrogate gene regulatory programs in the distinct cell types of the human heart. Recent single-cell/nucleus RNA sequencing (RNA-seq) (*9–11*) and spatial transcriptomic (*12*) studies have revealed gene expression patterns in distinct cardiac cell types across developmental and adulthood stages in the human heart, including some that display gene expression patterns that are cardiac chamber/region specific (*10*, *11*). However, the transcriptional regulatory programs responsible for cell type–specific and chamber-specific gene expression, and their potential links to noncoding risk variants for cardiovascular diseases and traits, remain to be fully defined.

cCREs have been annotated in the human genome with the use of chromatin immunoprecipitation sequencing (ChIP-seq), deoxyribonuclease sequencing (DNase-seq), assay for transposase-accessible chromatin using sequencing (ATAC-seq), global run-on sequencing, etc. in a broad spectrum of human tissues including in bulk heart tissues and in purified cardiomyocytes (*2–5*, *13–18*). These maps have provided important insights into dynamic gene regulation during heart failure (*15*, *16*, *18*) and begun to shed light on the function of noncoding cardiovascular disease variants (*8*, *13*, *16*, *18*, *19*). However, major limitations of these studies including their focus on particular chambers/regions of the heart and failure to interrogate cis-regulatory elements across all distinct cardiac cell types have restricted their utility in understanding how specific gene regulatory mechanisms may affect distinct cell types and regions of human hearts in health and disease. Although recent single-cell genomic tools provide the opportunity to interrogate cis-regulatory elements at single-cell resolution (*20–23*), their application to mammalian hearts has been limited to one single-cell ATAC-seq dataset from adult mouse heart (*24*), fewer than 200 total cells from mouse fetal hearts (*25*), and fetal human heart (*26*). Thus, to comprehensively investigate cis-regulatory elements in the specific cell types of the adult human heart, we profiled chromatin accessibility in ~80,000 heart cells using single-nucleus ATAC-seq (snATAC-seq) (*21*, *27*) and created a comprehensive cardiac cell atlas of cCREs annotated by cell type and putative target genes. Integration of these data with single-nucleus RNA-seq (snRNA-seq) datasets from matched specimens revealed gene regulatory programs in nine major cardiac cell types. Using this human cardiac cCRE atlas, we further observed the remodeling of cell type–specific candidate enhancers during heart failure and the enrichment of cardiovascular disease–associated genetic variants in cCREs of specific cell types. Last, we showed that a cardiomyocyte-specific enhancer harboring risk variants for atrial fibrillation (AF) is necessary for cardiomyocyte *KCNH2* expression and regulation of cardiac action potential repolarization.

[1]Medical Scientist Training Program, University of California, San Diego, La Jolla, CA, USA. [2]Biomedical Sciences Graduate Program, University of California, San Diego, La Jolla, CA, USA. [3]Ludwig Institute for Cancer Research, La Jolla, CA, USA. [4]Center for Epigenomics, University of California, San Diego, La Jolla, CA, USA. [5]Division of Cardiology, Department of Medicine, University of California, San Diego, La Jolla, CA, USA. [6]Department of Pediatrics, Pediatric Diabetes Research Center, University of California, San Diego, La Jolla, CA, USA. [7]Departments of Bioengineering and Medicine, University of California, San Diego, La Jolla, CA, USA. [8]Institute for Engineering in Medicine, University of California, San Diego, La Jolla, CA, USA. [9]Department of Cellular and Molecular Medicine, University of California, San Diego School of Medicine, La Jolla, CA, USA. [10]Institute for Genomic Medicine, University of California, San Diego, La Jolla, CA, USA.
*Corresponding author. Email: biren@health.ucsd.edu (B.R.); nchi@health.ucsd.edu (N.C.C.); spreissl@health.ucsd.edu (S.P.)
†These authors contributed equally to this work.

RESULTS

Single nucleus analysis of chromatin accessibility and transcriptome in adult human hearts

To assess the accessible chromatin landscape of distinct cardiovascular cell types, we performed snATAC-seq (27), also known as sci-ATAC-seq (21), on all cardiac chambers from four adult human hearts without known cardiovascular disease (table S1). We obtained accessible chromatin profiles for 79,515 nuclei, with a median of 2682 fragments mapped per nucleus (Fig. 1, A and B; fig. S1; and table S2). We also performed snRNA-seq for a subset of the above heart samples to complement the accessible chromatin data and obtained 35,936 nuclear transcriptomes, with a median of 2184 unique molecular identifiers and 1286 genes detected per nucleus (Fig. 1, A and C; fig. S2, A to F; and table S3). Using SnapATAC (28) and Seurat (29), we identified nine clusters from snATAC-seq (Fig. 1B) and 12 major clusters from snRNA-seq (Fig. 1C and fig. S2, G and H), which were annotated on the basis of chromatin accessibility at promoter regions or expression of known lineage-specific marker genes, respectively (Fig. 1, D and E, and table S4)

(10, 11). For example, chromatin accessibility and gene expression of atrial and ventricular cardiomyocyte markers such as *NPPA* and *MYH7* (30) were used to classify these two cardiomyocyte subtypes (Fig. 1, D and E). Although gene expression patterns of lineage markers strongly correlated with accessibility at promoter regions across annotated cell types (Fig. 1F) and single-cell integration analysis (29) revealed 93% concordance in annotation between snATAC-seq and snRNA-seq datasets (fig. S3 and table S3), some cellular subtypes identified from snRNA-seq including endocardial cells and myofibroblasts were not detected by snATAC-seq (Fig. 1F). Notably, cluster correlation and integration analysis showed that these cell types are present within the snATAC-seq data as part of the endothelial and smooth muscle clusters, respectively (Fig. 1F, fig. S3, and table S3). The discrepancy in clustering may be attributable to the conservative snATAC-seq clustering parameters or the sparse nature of snATAC-seq data (31). In addition, atrial and ventricular cardiomyocyte nuclei from the left and right regions of the heart could be further clustered by transcriptome but not chromatin accessibility (fig. S2, I and J). We noted that cell type composition
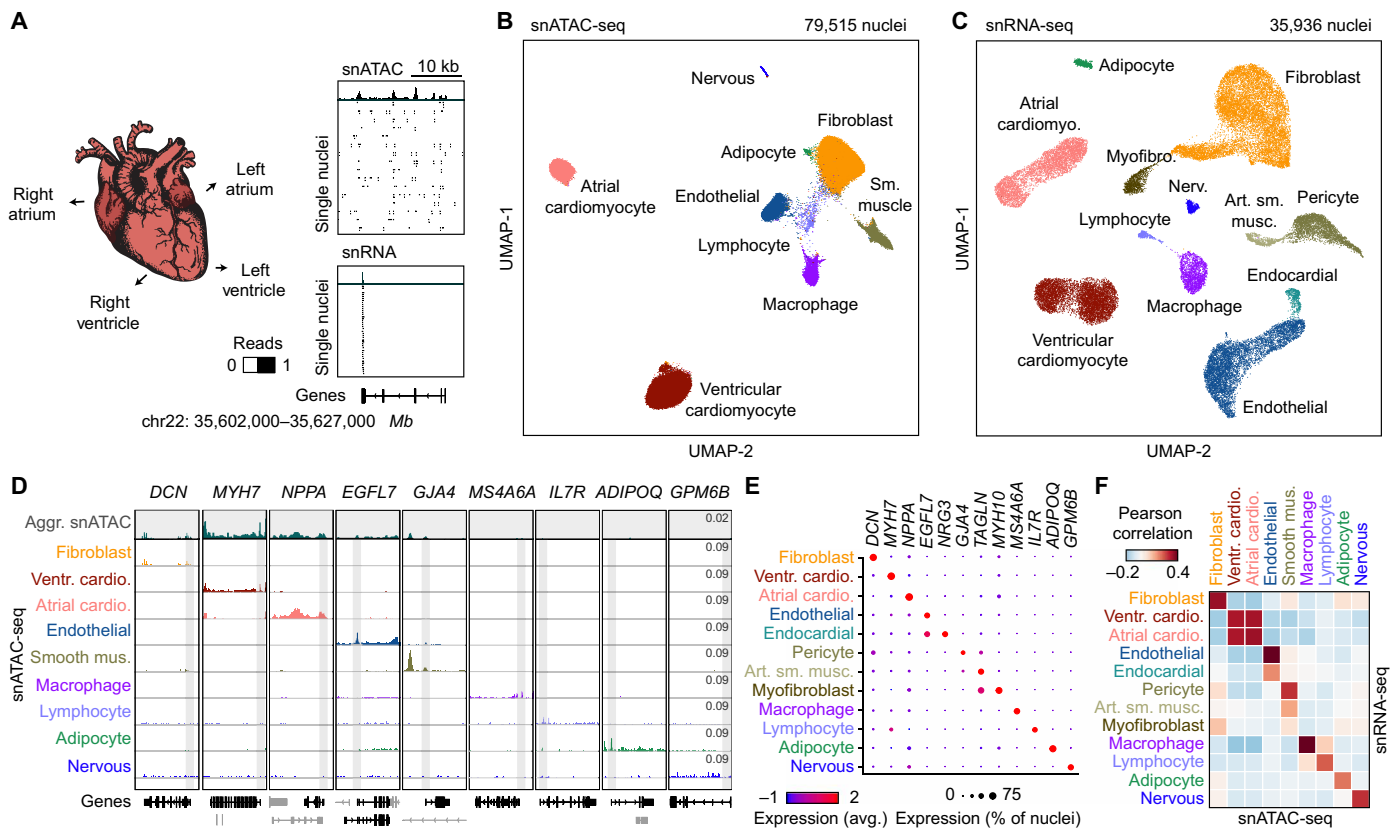

Fig. 1. Single-nucleus chromatin accessibility and transcriptome profiling of human hearts. (**A**) snATAC-seq and snRNA-seq were performed on nuclei isolated from cardiac chambers from four human donors without cardiovascular pathology. snATAC-seq: n = 4 (left ventricle), n = 4 (right ventricle), n = 3 (left atrium), and n = 2 (right atrium); snRNA-seq: n = 2 (left ventricle), n = 2 (right ventricle), n = 2 (left atrium), and n = 1 (right atrium). (**B**) Uniform manifold approximation and projection (UMAP) (108) and clustering analysis of snATAC-seq data reveal nine clusters. Each dot represents a nucleus colored by cluster identity. (**C**) UMAP (108) and clustering analysis of snRNA-seq data reveal 12 major clusters. Each dot represents a nucleus colored by cluster identity. Nerv., nervous. Art. sm. musc., arterial smooth muscle. (**D**) Genome browser tracks (141) of aggregate chromatin accessibility profiles [scale: reads per million (RPM)] at selected representative marker gene examples for individual clusters and for all nuclei pooled together into an aggregated heart dataset (top track, gray). Black genes below tracks represent the indicated marker genes, nonmarker genes are grayed. (**E**) Dot plot illustrating expression of representative marker gene examples in individual snRNA-seq clusters. (**F**) Heatmap illustrating the correlation between clusters defined by chromatin accessibility and transcriptomes. Pearson correlation coefficients were calculated between chromatin accessibility at cCREs within 2 kbp of annotated promoter regions (76) and expression of the corresponding genes for each cluster.

varied substantially between biospecimens and donors, highlighting the importance of single-cell approaches to limit biases due to cell proportion differences in bulk assays (fig. S4 and tables S2 and S3). In summary, we identified and annotated cardiac cell types using both chromatin accessibility and nuclear transcriptome profiles.

Identification of cCREs in distinct cell types of the human heart

To discover the cCREs in each cell type of the human heart, we aggregated snATAC-seq data from nuclei comprising each cell cluster individually and determined accessible chromatin regions with MACS2 (32). We then merged the peaks from all nine cell clusters into a union of 287,415 cCREs, which covered 4.7% of the human

genome (Fig. 2A and table S5). Sixty-seven percent of the cCREs identified in the current study overlapped previously annotated cCREs from a broad spectrum of human tissues and cell lines (fig. S5A) (5), and the union of heart cCREs captured 98.6 and 95.4% of candidate human heart enhancers reported in two previous bulk studies (fig. S5, B and C) (13, 15). Furthermore, 75% of cCREs in the union were at least 2 kilo–base pairs (kbp) away from annotated promoter regions, and 19,447 displayed high levels of cell type specificity [false discovery rate (FDR) < 0.01; Fig. 2B and table S6]. Gene ontology analysis (33) revealed that these cell type–specific cCREs were proximal to genes involved in relevant biological processes, including collagen fibril organization for cardiac fibroblast–specific cCREs (K1) and myofibril organization for ventricular cardiomyocyte–specific cCREs
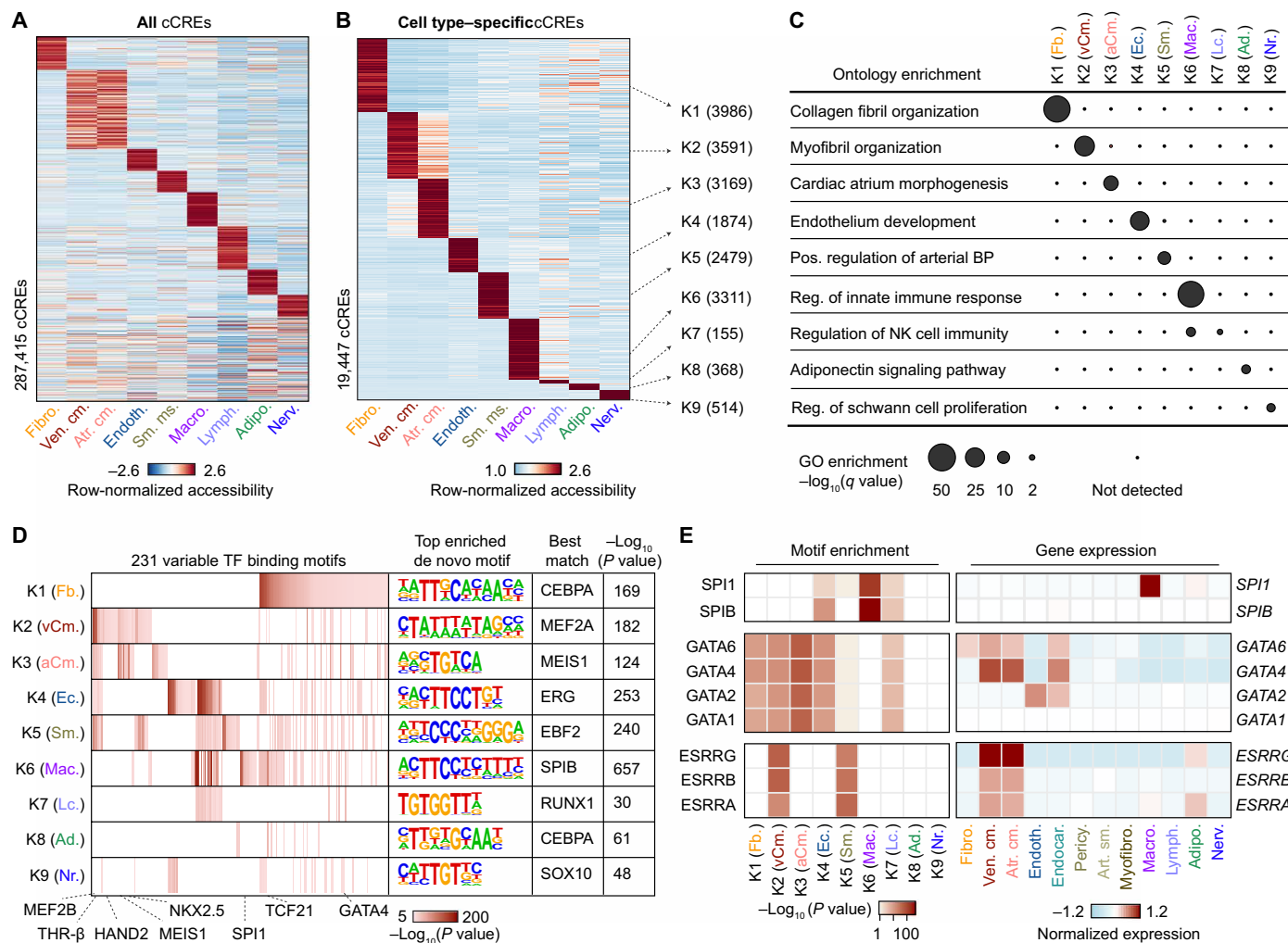

Fig. 2. Characterization of gene regulatory programs in cardiac cell types. (A) Heatmap illustrating row-normalized chromatin accessibility values for the union of 287,415 cCREs. K-means clustering was performed to group cCREs on the basis of relative accessibility patterns. (B) Heatmap showing row-normalized chromatin accessibility of 19,447 cell type–specific cCREs (FDR < 0.01 after Benjamini-Hochberg correction; fold change > 1.2). K-means clustering was performed to group cCREs on the basis of relative accessibility patterns. Number of cCREs per K can be found in brackets. (C) Genomic Regions Enrichment of Annotations Tool (GREAT) ontology analysis (33) of cell type–specific cCREs. Q value for enrichment indicates Bonferroni-adjusted P value. NK, natural killer; GO, gene ontology; BP, blood pressure. (D) Transcription factor motif enrichment (35) for known and de novo motifs within cell type–specific cCREs. The heatmap in shows motifs with an enrichment P value of $<10^{-5}$ in at least one cluster. For de novo transcription factor motifs, the best matches for the top motifs are displayed. Statistical test for motif enrichment: hypergeometric test. P values were not corrected for multiple testing. (E) Combination of transcription factor motif enrichment and gene expression shows cell type–specific roles for members of transcription factor families. Displayed are heatmaps for known motif enrichment in cell type–specific cCREs (left) and gene expression across clusters (right). (Fb., fibroblast; vCm., ventricular cardiomyocyte; aCm., atrial cardiomyocyte; Ec., endothelial; Sm., smooth muscle; Mac., macrophage; Lc., lymphocyte; Ad., adipocyte; Nr., nervous.)

(K2; Fig. 2C and table S7). Using chromVAR (table S8) (*34*) and HOMER (table S9) (*35*), we detected cell type–dependent enrichment for 231 transcription factor–binding signatures, such as MEF2A/B, NKX2.5, and THR-β sequence motifs in cardiomyocyte-specific cCREs and TCF21 motifs in cardiac fibroblast–specific cCREs (Fig. 2D). To discover the transcription factors that may bind to these sites, we combined corresponding snRNA-seq data with sequence motif enrichments to correlate the expression of these transcription factors with motif enrichment patterns across cell types (Fig. 2E). As an example, we found strong enrichment of the binding motif for the macrophage transcription factor SPI1/PU.1 (*36*) in macrophage-specific cCREs, and *SPI1* was exclusively expressed in macrophages (Fig. 2E and table S4). In addition, we observed that transcription factor family members were expressed in cell type–specific combinations. For instance, while GATA Family of transcription factors (GATA-binding factor) displayed similar motif enrichment patterns across sets of cell type–specific cCREs, we found that endothelial

cells and cardiac fibroblasts expressed *GATA2* and *GATA6*, respectively, whereas cardiomyocytes expressed both *GATA4* and *GATA6* and endocardial cells expressed *GATA2*, *GATA4*, and *GATA6* (Fig. 2E and table S4). In summary, these results establish a resource of cCREs for interrogation of cardiac cell type–specific gene regulatory programs.

Cardiac cell type–specific gene regulatory programs implicated in chamber-specific structure and function

Each cardiac chamber performs a unique role that is crucial to system-level heart function (*37*). To investigate the gene regulatory programs underlying chamber-specific gene expression and cellular functions in distinct cardiac cell types, we tested cCREs for differential accessibility across five of the most abundant cell types of the heart: cardiomyocytes, cardiac fibroblasts, endothelial cells, smooth muscle cells, and macrophages. We found 16,451 differentially accessible (DA) cCREs between pooled atria and ventricles, most of which were detected in cardiomyocytes (Fig. 3, A to C, and table S10).
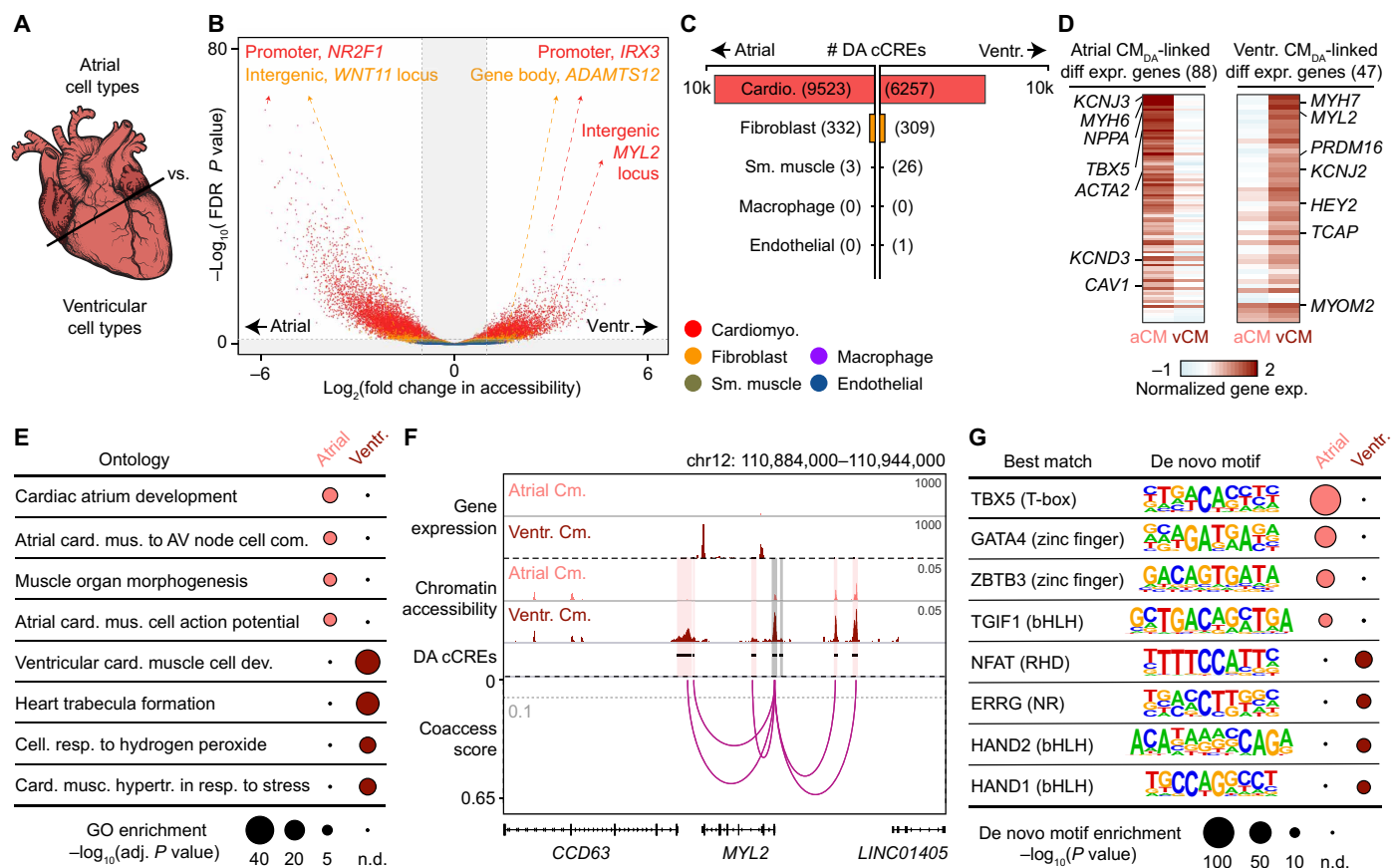

Fig. 3. Cardiomyocyte cCREs display chamber-dependent differences in chromatin accessibility. (**A**) Scheme for comparison of major cell types across heart chambers. (**B**) Volcano plot showing DA cCREs in each cell type between atria and ventricles. Each dot represents a cCRE, and the color indicates the cell type. cCREs with log$_2$(fold change) > 1 and FDR < 0.05 after Benjamini-Hochberg correction (outside the shaded area) were considered as DA. (**C**) Number of DA cCREs between atria and ventricles by cell type. (**D**) Heatmaps showing normalized gene expression levels of differentially expressed genes between atrial (aCM) and ventricular cardiomyocytes (vCM) that were linked by coaccessibility to distal DA cCREs that were more accessible in atrial cardiomyocytes (Atrial CM$_{DA}$) or ventricular cardiomyocytes (Ventr. CM$_{DA}$), respectively. (**E**) GREAT ontology analysis (*33*) of DA cCREs between atrial and ventricular cardiomyocytes. AV, atrioventricular. *P* values shown are Bonferroni adjusted (n.d., not detected). (**F**) Genome browser tracks (*141*) showing chromatin accessibility (scale, RPM) and gene expression [scale, reads per kilobase per million mapped reads (RPKM)] in atrial and ventricular cardiomyocytes as well as DA cCREs that were coaccessible with the promoter of *MYL2*. Gray dotted line indicates coaccessibility threshold (>0.1). Red shaded areas, distal DA cCREs coaccessible with *MYL2* promoter. Gray shaded areas, DA cCREs overlapping the promoter region of *MYL2*. (**G**) Transcription factor motif enrichment analysis (*35*) of DA cCREs between atrial and ventricular cardiomyocytes. The best matches for the top de novo motifs (score > 0.7) are shown. Statistical test for motif enrichment: hypergeometric test. *P* values were not corrected for multiple testing.

Specifically, 11,159 cCREs displayed differential accessibility between the right atrium and the right ventricle, and 12,962 cCREs exhibited differential accessibility between the left atrium and the left ventricle (fig. S6, A to C, and table S10). Comparing the left and right sides of the heart, we identified 101 DA cCREs between the right and left ventricles (fig. S6D) and 2687 DA cCREs between left and right atria, which, in contrast to comparisons between atria and ventricles, were found primarily in cardiac fibroblasts (fig. S6E and table S10).

Using coaccessibility analysis (38) to link distal DA cCREs (~88% of all DA cCREs) to their putative target genes (table S11; median distance, 88.7 kbp), we observed that distal DA cCREs in cardiomyocytes between atria and ventricles were associated with chamber-specific gene expression of their putative target genes (Fig. 3D; fig. S6, B to E; and table S12), and genes near these DA cCREs were enriched for chamber-specific biological processes (Fig. 3E; fig. S6, B to E; and table S13). Specifically, distal DA cCREs with higher accessibility in atrial cardiomyocytes were associated with genes such as PITX2, a transcriptional regulator of cardiac atrial development (39), as well as the ion channel subunit SCN5A, which regulates cardiomyocyte action potential (Fig. 3E and table S13) (40). Furthermore, we found distal DA cCREs with higher accessibility in atrial cardiomyocytes at the HAMP gene locus, which encodes a key regulator of ion homeostasis and was recently described as a potential previously unknown cardiac gene in the right atrium by single-nucleus transcriptomic analysis (10, 11). Conversely, genes near distal DA cCREs with higher accessibility in ventricular cardiomyocytes were enriched for biological processes such as trabecula formation and ventricular cardiac muscle cell differentiation. For example, several distal DA cCREs with increased accessibility in ventricular cardiomyocytes compared to atrial cardiomyocytes were linked to the promoter region of MYL2, which encodes the ventricular isoform of myosin light chain 2 (Fig. 3F and table S4) (41), a regulator of ventricular cardiomyocyte sarcomere function.

In addition, analysis of distal DA cCREs in cardiac fibroblasts revealed that putative target genes were involved in distinct biological processes between right and left atria. In particular, we found that DA cCREs with higher accessibility in right atrial cardiac fibroblasts were proximal to genes involved in heart development, heart growth, and tube development, whereas DA cCREs with higher accessibility in left atrial cardiac fibroblasts were adjacent to genes involved in biological processes such as wound healing and vasculature development (fig. S6E and table S13). We further found a cardiac fibroblast–specific DA cCRE with higher accessibility in left atria at the fibrinogen FN1 gene locus, potentially indicating a more activated fibroblast state (10, 42). Supporting these findings, we identified several other DA cCREs with higher accessibility in left atrial cardiac fibroblasts adjacent to genes involved in generation of extracellular matrix (ECM) such as MMP2 and FBLN2 (table S13). These observations are consistent with previous findings that a higher fraction of ECM is produced in fibroblasts of the left atrium (10).

Using motif enrichment analysis, we inferred candidate transcriptional regulators involved in chamber-specific cellular specialization, including TBX5 (T-Box Transcription Factor 5), GATA4, and TGIF1 (TGFB Induced Factor Homeobox 1) for atrial cardiomyocytes and nuclear factor of activated T cells (NFAT), ERRG (Estrogen-Related Receptor Gamma), HAND1 (Heart- and Neural Crest Derivatives-Expressed Protein 1), and HAND2 for ventricular cardiomyocytes (Fig. 3G and table S14). While the TBX5 DNA binding motif was strongly enriched in both right and left atrial cardiomyocyte DA cCREs, the NFAT5 motif ranked highest in left ventricular cardiomyocyte DA cCREs and the TBX20 motif was strongly enriched in right ventricular cardiomyocyte DA cCREs (fig. S6, B and C, and table S14). Furthermore, cardiac fibroblast DA cCREs with higher accessibility in the right atrium were enriched for the binding motif of forkhead transcription factors (fig. S6E), whereas cardiac fibroblast DA cCREs with higher accessibility in the left atrium were enriched for the homeobox transcription factor CUX1 motif (fig. S6E and table S14). Together, we identified cCREs and candidate transcription factors associated with specific cardiac chambers, particularly within cardiomyocytes and cardiac fibroblasts.

Cell type specificity of candidate enhancers associated with heart failure

Recent large-scale studies profiling the H3K27ac histone modification in human hearts have uncovered candidate enhancers associated with heart failure (15, 18). However, because these studies either examined heterogeneous bulk heart tissue (15, 18) or focused solely on enriched cardiomyocytes (16), it remains unclear what role, if any, additional cardiac cell types and cCREs may contribute to heart failure pathogenesis. Using our cell atlas of cardiac cCREs, we revealed the cell type specificity of candidate enhancers showing differential H3K27ac signal strength between human hearts from healthy donors and donors with dilated cardiomyopathy (heart failure) (Fig. 4 and fig. S7, A to E) (15). We observed that a large fraction of candidate enhancers that displayed increased activity (45%) during heart failure were accessible primarily in cardiac fibroblasts (Fig. 4A, K2-4$_{up}$, and table S15), whereas most of those exhibiting decreased activity (67%) were accessible primarily in cardiomyocytes (Fig. 4B, K1-3$_{down}$, and table S15). Candidate enhancers with increased activity in cardiac fibroblasts were proximal to genes involved in ECM organization and connective tissue development (Fig. 4A, K2-4$_{up}$, and table S16), whereas those exhibiting decreased activity in cardiomyocytes were proximal to genes involved in the regulation of heart contraction and cation transport (Fig. 4B, K1-3$_{down}$, and table S16). For example, several of these cardiac fibroblast candidate enhancers were present at loci encoding the ECM proteins lumican (LUM) and decorin (DCN) and coaccessible with the promoters of these genes (Fig. 4C). Consistent with these findings, both genes were primarily expressed in cardiac fibroblasts (fig. S7F and table S4), and LUM has been reported to exhibit increased expression in failing hearts compared to control hearts (15). On the other hand, several cardiomyocyte candidate enhancers displaying decreased activity in heart failure were coaccessible with the promoter region of IRX4 (Fig. 4D), which encodes a ventricle-specific transcription factor (43) and is primarily expressed in cardiomyocytes of the left ventricle (fig. S7G and table S4).

To identify potential transcription factors regulating these pathologic responses during heart failure, we performed motif enrichment analysis in cell type–specific subsets of enhancers showing differential H3K27ac signal strength between healthy and failing hearts (table S17). For candidate enhancers exhibiting increased activity in heart failure, we identified enrichment of not only bHLH motifs such as AP4 in cardiac fibroblast candidate enhancers, which matched previous bulk analysis (Fig. 4E, K2-4$_{up}$) (15), but also TEAD3 and MYF6 motifs in cardiomyocyte candidate enhancers (Fig. 4E, K1$_{up}$). Conversely, for candidate enhancers displaying decreased activity in heart failure, we observed enrichment of nuclear receptor motifs
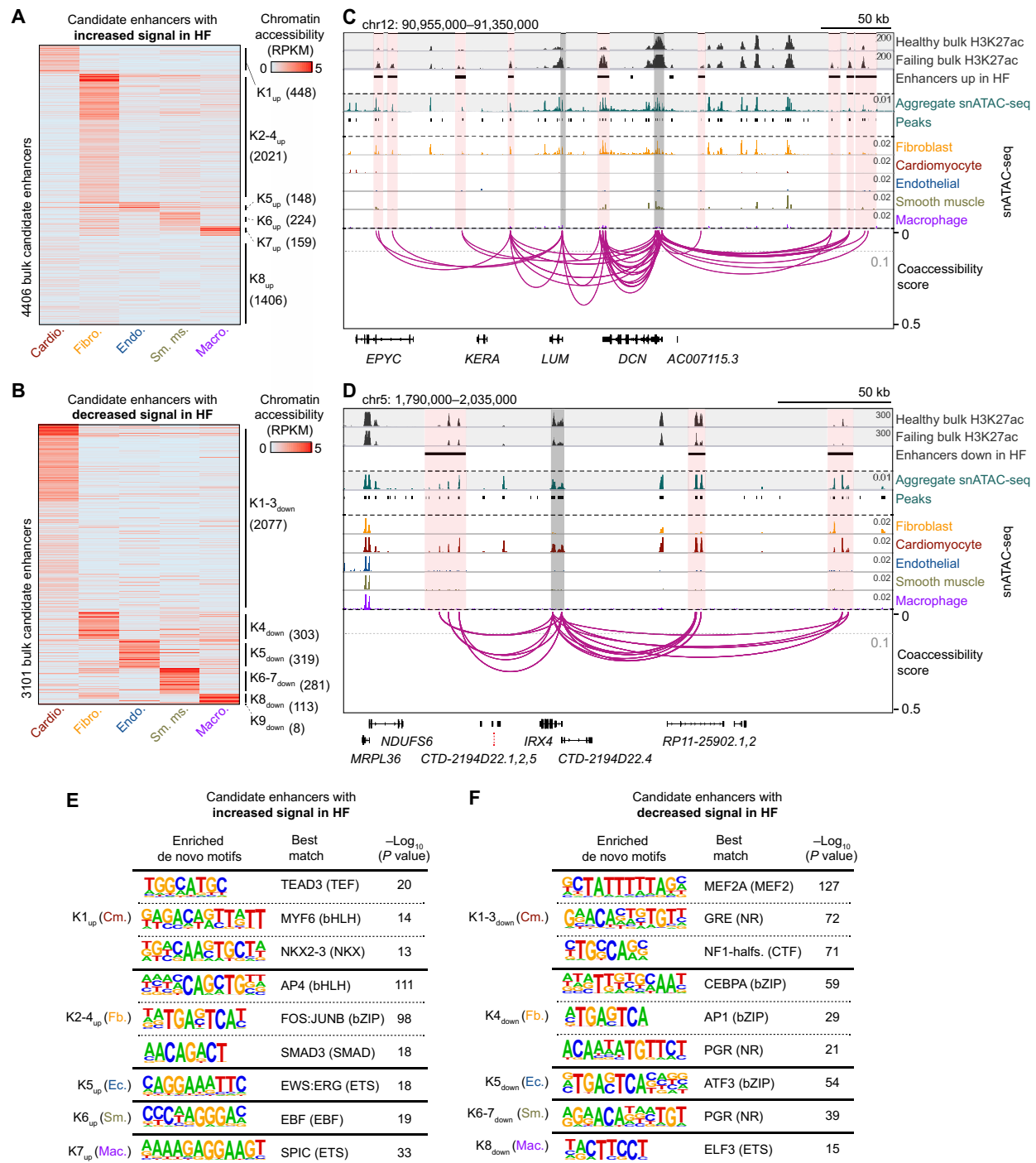
Fig. 4. Cell type specificity of candidate enhancers associated with heart failure. (**A**) Cell type specificity of 4406 candidate enhancers with increased H3K27ac signal in failing left ventricles (*15*). Heatmap displays cell type–resolved chromatin accessibility RPKM values for cell types from left ventricular snATAC-seq datasets. Candidate enhancers were grouped on the basis of chromatin accessibility patterns across cell clusters using *K*-means. HF, heart failure. (**B**) Cell type specificity of 3101 candidate enhancers with decreased H3K27ac signal in failing left ventricles (*15*). (**C**) Genome browser tracks (*141*) showing several candidate enhancers with increased activity during heart failure that were primarily accessible in fibroblasts and coaccessible with the promoters of *LUM* and/or *DCN*. For visualization, linkages between cCREs within candidate enhancers and all gene promoters are shown (coaccessibility > 0.1, gray dotted line). Candidate enhancers coaccessible with gene promoters are indicated by red shaded boxes and promoter regions are indicated by gray shaded boxes (scale, RPM). (**D**) Genome browser tracks (*141*) showing several bulk candidate enhancers with decreased activity in heart failure that were primarily accessible in cardiomyocytes and coaccessible with the promoter of *IRX4* (scale, RPM). (**E** and **F**) Transcription factor motif enrichment (*35*) in the candidate enhancers with (E) increased and (F) decreased activity in failing left ventricles. Analysis was performed on the indicated *K* cluster(s) from (A) and (B), respectively. The best matches for selected de novo motifs (score > 0.7) are shown. Statistical test for motif enrichment: hypergeometric test. *P* values were not corrected for multiple testing.

such as glucocorticoid response element in cardiomyocyte candidate enhancers, which is consistent with previous findings (Fig. 4F, K1-3$_{down}$) (15), as well as other motifs that were not detected in bulk analyses, such as the bZIP transcription factor CEBPA for cardiac fibroblast candidate enhancers (Fig. 4F, K4$_{down}$). Thus, these results show that this cardiac cell atlas of cCREs may be used to assign disease-associated candidate enhancers from bulk assays to their affected cell types and infer transcriptional regulators involved in lineage-specific disease pathogenesis.

Interpreting noncoding risk variants of cardiac diseases and traits

Noncoding genetic variants contributing to risk of complex diseases are enriched within cCREs in a tissue and cell type–dependent manner (24, 44–47). To examine the enrichment of cardiovascular disease variants within cCREs active in cardiac cell types, we performed cell type–stratified LD (linkage disequilibrium) score regression analysis (48) using genome-wide association study (GWAS) summary statistics for cardiovascular diseases (Fig. 5A) (49–53)
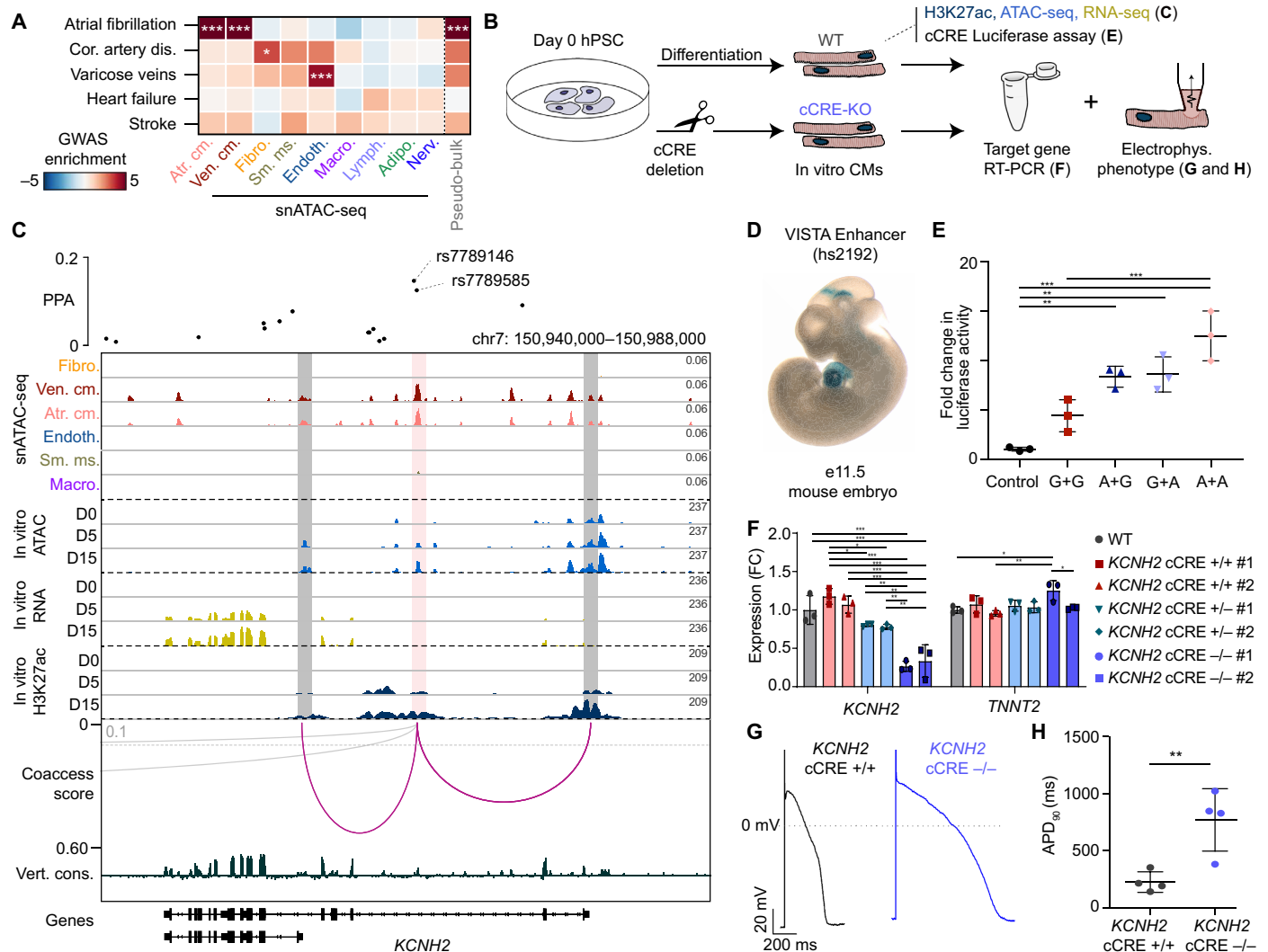

Fig. 5. Identification and characterization of AF-associated variants at the *KCNH2* locus. (**A**) Enrichment of cardiovascular disease variants within cardiac cell type cCREs. *z* scores are shown and were used to compute two-sided *P* values for enrichments. *FDR < 0.05 and ***FDR < 0.001. (**B**) Cardiomyocyte differentiation model schematic. (**C**) Fine mapping (*55*) and molecular characterization of two AF variants. Genome browser tracks (*141*) for snATAC-seq (top; scale, RPM) and indicated molecular features during hPSC-cardiomyocyte differentiation (scale, RPKM). Coaccessibility track shows linkages between the AF variant–containing cCRE and promoters (cutoff > 0.1). PPA, posterior probability of association (*55*). (**D**) Transgenic mouse embryo showing LacZ reporter expression under control of a genomic region [hs2192, VISTA database (*58*)] overlapping the variant-cCRE pair. (**E**) Luciferase reporter assay for the AF variant–harboring cCRE in D15 cardiomyocytes. Genotypes for rs7789146 and rs7789585 were either both G (risk), both A (nonrisk), or a combination. Each dot represents luciferase activity (average of two transfections) in independent replicates of D15 cardiomyocytes. Data are means ± SD. ***P < 0.001 and **P < 0.01 [one-way analysis of variance (ANOVA) and Tukey post hoc test]. Control: minimal promoter. (**F**) Expression of *KCNH2* and *TNNT2* in D25 cardiomyocytes after CRIPSR-Cas9–mediated cCRE deletion. Each dot represents an independent cardiomyocyte differentiation. Data are means ± SD. ***P < 0.001, **P < 0.01, and *P < 0.05 (one-way ANOVA and Tukey post hoc test); WT, unperturbed control. FC, fold change. (**G**) Action potential recordings in hPSC-derived cardiomyocytes with and without cCRE deletion at D25-35. (**H**) APD$_{90}$ (action potential duration at 90% depolarization) at 1-Hz pacing for four independent hPSC-derived cardiomyocytes with and without cCRE deletion at D25-35. Data are means ± SD. **P < 0.01 (unpaired two-sided *t* test).

and control traits (fig. S8A and table S18) by measuring the enrichment of disease-associated variants within all cCREs identified for each cell type. This analysis revealed significant enrichment of AF-associated variants in both atrial ($Z = 5.61$, FDR $= 1.9 \times 10^{-6}$) and ventricular cardiomyocyte cCREs ($Z = 6.80$, FDR $= 2.8 \times 10^{-9}$), varicose vein-associated variants in endothelial cell cCREs ($Z = 4.36$, FDR $= 3.9 \times 10^{-4}$), and coronary artery disease–associated variants in cardiac fibroblast cCREs ($Z = 3.29$, FDR $= 1.7 \times 10^{-2}$; Fig. 5A). Notably, except for AF, these associations were not significant in a pseudo-bulk heart dataset created by combining chromatin accessibility profiles from all cardiac cell types (Fig. 5A). Furthermore, cardiovascular disease variants were not significantly enriched in accessible chromatin from noncardiac tissues (3, 5, 17) or human lung cell types (54), with the exception of a significant enrichment of varicose vein–associated variants in endothelial cells (fig. S8, B and C).

Next, to identify likely causal AF risk variants in cardiomyocyte cCREs, we first determined the probability that variants were causal for AF [posterior probability of association (PPA)] at 111 known loci using Bayesian fine mapping (55). We then intersected fine-mapped AF variants with cCREs detected in atrial and/or ventricular cardiomyocytes and identified 38 variants with PPA > 10% in cardiomyocyte cCREs, including previously reported variants at the HCN4 (13) and SCN10A/SCN5A (56) loci (table S19). We further prioritized AF variants for molecular characterization on the basis of their overlap with cCREs that were primarily accessible in cardiomyocytes, evolutionarily conserved, and coaccessible with promoters of genes expressed in cardiomyocytes. To experimentally validate the molecular functions of cCREs containing AF variants, we used a human pluripotent stem cell (hPSC)–derived cardiomyocyte differentiation model system (Fig. 5B) (57). From the variant prioritization analysis, we found a cCRE in the second intron of the potassium channel gene KCNH2 (HERG), which was coaccessible with the KCNH2 promoter (Fig. 5C) and harbored two variants, rs7789146 and rs7789585, with a combined PPA of 28% (Fig. 5C and fig. S9A). KCNH2 was primarily expressed in ventricular and atrial cardiomyocytes in the adult human heart (fig. S9B). The cCRE appeared to be activated during hPSC-cardiomyocyte differentiation, as evidenced by an increase in H3K27ac signal that correlated with KCNH2 expression (Fig. 5C). Supporting its in vivo role in regulating gene expression in mammalian hearts, a genomic region (hs2192) (58) containing this cCRE was previously shown to drive LacZ reporter expression in mouse embryonic hearts (Fig. 5D) (58).

A cardiomyocyte enhancer of KCNH2 is affected by noncoding risk variants associated with AF

To investigate whether these AF variants may affect enhancer activity and thereby regulate KCNH2 expression and cardiomyocyte electrophysiologic function, we initially carried out reporter assays using a hPSC cardiomyocyte model system. Results from these studies confirmed that in D15 hPSC cardiomyocytes, the KCNH2 enhancer carrying the rs7789146-G/rs7789585-G AF risk allele displayed significantly weaker enhancer activity than when containing the nonrisk variants (Fig. 5E and fig. S9C), thus supporting the functional significance of these AF variants. We next used CRISPR-Cas9 genome editing strategies to remove the enhancer and performed quantitative polymerase chain reaction (qPCR) and electrophysiologic assays to examine its role in KCNH2 expression and function. Supporting the aforementioned findings, CRISPR-Cas9 genome deletion of this cCRE in hPSC cardiomyocytes resulted in decreased

KCNH2 expression in an enhancer dosage–dependent manner (Fig. 5F and fig. S9D). Similar to human cardiomyocytes with loss of KCNH2 function due to mutations in the KCNH2 coding sequence (59) or gene knockdown (60), cellular electrophysiologic studies demonstrated that these cCRE-deleted hPSC cardiomyocytes displayed a significantly prolonged action potential duration (Fig. 5, G and H), thus suggesting that cardiac repolarization abnormalities in atrial cardiomyocytes may lead to AF in an analogous manner to ventricular arrhythmias due to long QT syndrome (59). Together, these results highlight the utility of this single-cell atlas for assigning noncoding cardiovascular disease risk variants to distinct cell types and affected cCREs and functionally interrogating how these variants may contribute to cardiovascular disease risk.

DISCUSSION

The limited ability to interrogate cell type–specific gene regulation in the human heart has been a major barrier for understanding molecular mechanisms of cardiovascular traits and diseases. Here, we report a cell type–resolved atlas of cCREs in the human heart, which was ascertained by profiling accessible chromatin in individual nuclei from all four chambers of multiple human hearts and includes both cell type–specific and heart chamber–specific cCREs. In particular, we observed chamber-specific differences in chromatin accessibility between ventricles and atria as well as left and right atria but notably detected few differences between left and right ventricles. This finding is consistent with recent snRNA-seq analysis in human hearts, which found few differentially expressed genes between left and right ventricles (11). We note that the power to detect chamber-specific differences in chromatin accessibility depends on the number of samples assayed and the total nuclei used as input for differential analysis. Thus, future studies with larger cohorts will likely reveal additional chamber-specific differences between cardiac cell types.

We further highlight the utility of this atlas of heart cCREs to provide previously unidentified insight into aberrant gene regulation during cardiovascular pathology. To this end, we delineated the cell type specificity of enhancers that were differentially active between healthy and failing heart tissue (15) and identified additional transcription factors that may be involved in the pathogenesis of specific cell types during heart failure. Such cell type–specific analysis is particularly important in the context of heart failure because cellular composition can differ between diseased and control hearts (16, 61). This change in cellular composition may, in part, explain the cell type bias that we observed between candidate enhancers exhibiting increased and decreased activity during heart failure (i.e., cardiac fibroblasts and cardiomyocytes, respectively). However, because of the large differences in H3K27ac signal, we suspect that measured changes in candidate enhancer activity could be due to a combination of both enhancer remodeling and shift in cell type composition. Thus, future studies profiling snATAC-seq and H3K27ac in parallel from the same cardiac sample or novel approaches to profile histone modifications in single nuclei (62, 63) will provide greater insight into the extent of changes in chromatin accessibility and enhancer activity in individual cardiac cell types from diseased hearts.

Last, we show how this atlas can be used to not only assign noncoding genetic variants associated with cardiovascular disease risk to cCREs in specific cardiac cell types but also illuminate their cellular and molecular consequences. In particular, we found significant enrichment of AF-associated variants within cardiomyocyte cCREs

and functionally interrogated one of these cCREs by demonstrating its role in regulating *KCNH2* expression and cardiomyocyte repolarization. Similar to electrophysiologic phenotypes of human cardiomyocytes exhibiting *KCNH2* loss of function (*59*, *60*), hPSC cardiomyocytes harboring deletions of this cCRE displayed action potential prolongation, suggesting that cardiac repolarization abnormalities may contribute to AF, possibly through similar mechanisms as to how they may contribute ventricular arrhythmias (*59*). On the other hand, we found no enrichment of variants associated with heart failure in any cardiac cell type. This finding may reflect the heterogeneous etiologies of cardiovascular diseases and the limited number of currently known risk loci for heart failure (*49*). Future GWAS in large cohorts with detailed phenotyping, including biobanks such as the UK Biobank (*64*) and the BioBank Japan Project (*65*) and whole-genome sequencing efforts such as the National Heart, Lung, and Blood Institute Trans-Omics for Precision Medicine program (*66*), will help identify and refine disease association signals. Therefore, this atlas of cardiac cCREs will be a valuable resource for continued discovery of regulatory elements, target genes, and specific cell types that may be affected by noncoding cardiovascular genetic variants.

In summary, we created a human heart cell atlas of >287,000 cCREs, which may serve as a reference to further expand our knowledge of gene regulatory mechanisms underlying cardiovascular disease. To facilitate distribution of these data, we created a web portal at http://catlas.org/humanheart. Integrating this resource with genomic and epigenomic clinical cardiac datasets, we built a systematic framework to interrogate how cis-regulatory elements and genetic variants might contribute to cardiovascular diseases such as heart failure or AF. Overall, such information will have great potential to provide insight into the development of future cardiac therapies that are tailored to affected cell types and thus optimized for treating specific cardiovascular diseases.

MATERIALS AND METHODS
Experimental design
We performed snATAC-seq to define a comprehensive catalog of cCREs for the cell types in four regions of nonfailing human hearts and generated in parallel snRNA-seq datasets for a subset to delineate gene expression patterns. We used the cCRE catalog to computationally assign dynamic enhancers in failing hearts to cell types and to assign cardiovascular disease risk variants to cCREs in individual cardiac cell types. Last, we applied reporter assays, genome editing, and electrophysiological measurements in in vitro differentiated human cardiomyocytes to validate the molecular mechanisms of cardiovascular disease risk variants.

Human tissues
Adult human heart tissues were procured at the time of organ donation using an Institutional Review Board protocol (no. 101021) approved by the University of California, San Diego. Donated hearts were perfused with cold cardioplegia before cardiectomy and then explanted immediately into an ice-cold physiologic solution as we previously described (*67*). Full-thickness samples from each chamber were obtained, and epicardial fat was rapidly removed before immediately flash-freezing samples in liquid nitrogen. Samples were received from the United Network for Organ Sharing. Limited clinical data were obtained for each heart per approved Institutional Review Board protocol (table S1). All samples were stored at −80°C until processing.

Single-nucleus ATAC-seq
Combinatorial barcoding snATAC-seq was performed as described previously (*21*, *27*, *28*) with slight modifications. Nuclei were isolated in gentleMACS M tubes (Miltenyi) on a gentleMACS Octo dissociator (Miltenyi) using the "Protein_01_01" protocol in magnetic-activated cell sorting (MACS) buffer [5 mM $CaCl_2$, 2 mM EDTA, 1× protease inhibitor (Roche, 05-892-970-001), 300 mM MgAc, 10 mM tris-HCl (pH 8), and 0.6 mM dithiothreitol (DTT)]. Nuclei were pelleted with a swinging-bucket centrifuge (500*g*, 5 min, 4°C; 5920R, Eppendorf) and resuspended in 1 ml of nuclear permeabilization buffer [1× phosphate-buffered saline (PBS), 5% bovine serum albumin (BSA), 0.2% IGEPAL CA-630 (Sigma-Aldrich), 1 mM DTT, and 1× protease inhibitor]. Nuclei were rotated at 4°C for 5 min before being pelleted again with a swinging-bucket centrifuge (500*g*, 5 min, 4°C; 5920R, Eppendorf). After centrifugation, permeabilized nuclei were resuspended in 500 µl of high-salt tagmentation buffer [36.3 mM tris acetate (pH 7.8), 72.6 mM potassium acetate, 11 mM Mg acetate, and 17.6% *N*,*N*′-dimethylformamide] and counted using a hemocytometer. Concentration was adjusted to 2000 nuclei/9 µl, and 2000 nuclei were dispensed into each well of a 96-well plate per sample (96 tagmentation wells per sample; samples were processed in batches of two to four samples). For tagmentation, 1 µl of barcoded Tn5 transposomes (table S20) was added using a BenchSmart 96 (Mettler Toledo), mixed five times, and incubated for 60 min at 37°C with shaking (500 rpm). To inhibit the Tn5 reaction, 10 µl of 40 mM EDTA (final 20 mM) was added to each well with a BenchSmart 96 (Mettler Toledo), and the plate was incubated at 37°C for 15 min with shaking (500 rpm). Next, 20 µl of 2× sort buffer (2% BSA and 2 mM EDTA in PBS) were added using a BenchSmart 96 (Mettler Toledo). All wells were combined into a separate fluorescence-activated cell sorter tube for each sample and stained with DRAQ7 at 1:150 dilution (Cell Signaling Technology). Using an SH800 (Sony), 20 nuclei per sample were sorted per well into eight 96-well plates (total of 768 wells) containing 10.5 µl of Elution Buffer (EB; Qiagen) [25 pmol of primer i7, 25 pmol of primer i5, and 200 ng of BSA (Sigma-Aldrich)]. During the sort, nuclei with two to eight copies of DNA (2-8n) were included because cardiomyocyte nuclei in human hearts are often polyploid (*16*). Preparation of sort plates and all downstream pipetting steps were performed on a Biomek i7 automated workstation (Beckman Coulter). After addition of 1 µl of 0.2% SDS, samples were incubated at 55°C for 7 min with shaking (500 rpm). One microliter of 12.5% Triton X was added to each well to quench the SDS. Next, 12.5 µl of NEBNext High-Fidelity 2× PCR Master Mix (New England Biolabs) were added, and samples were PCR-amplified [72°C 5 min, 98°C 30 s, (98°C 10 s, 63°C 30 s, 72°C 60 s) × 12 cycles, held at 12°C]. After PCR, all wells were combined. Libraries were purified according to the MinElute PCR Purification Kit manual (Qiagen) using a vacuum manifold (QIAvac 24 Plus, Qiagen), and size selection was performed with SPRISelect reagent (Beckman Coulter; 0.55× and 1.5×). Libraries were purified one more time with SPRISelect reagent (Beckman Coulter; 1.5×). Libraries were quantified using a Qubit fluorimeter (Life Technologies), and a nucleosomal pattern of fragment size distribution was verified using a TapeStation (High Sensitivity D1000, Agilent). Libraries were sequenced on a NextSeq 500 sequencer (Illumina) using custom sequencing primers with the following read lengths: 50 + 10 + 12 + 50 (read 1 + index 1 + index 2 + read 2). Primer and index sequences are listed in table S20.

Single-nucleus RNA-seq

Nuclei were isolated from heart tissue using a gentleMACS dissociator (Miltenyi). Frozen heart tissue (~40 mg) was suspended in 2 ml of MACS dissociation buffer [5 mM $CaCl_2$ (G-Biosciences, R040), 2 mM EDTA (Invitrogen, 15575-038), 1× protease inhibitor (Roche, 05-892-970-001), 3 mM MgAc (Grow Cells, MRGF-B40), 10 mM tris-HCl (pH 8) (Invitrogen, 15568-075), 0.6 mM DTT (Sigma-Aldrich, D9779), and ribonuclease (RNase) inhibitor (0.2 U/μl; Promega, N251B) in water (Corning, 46-000-CV)] and placed on wet ice. Next, samples were homogenized using a gentleMACS dissociator (Miltenyi) with gentleMACS M tubes (Miltenyi, 130-096-335) and the Protein_01_01 protocol. Suspension was filtered through a 30 μM CellTrics filter (Sysmex, 04-0042-2316). M tube and filter were washed with 3 ml of MACS dissociation buffer and combined with the suspension. Suspension was centrifuged in a swinging-bucket centrifuge (Eppendorf, 5920R) at 500g for 5 min (4°C, ramp speed of 3/3). Supernatant was carefully removed and pellet was resuspended in 500 μl of nuclei permeabilization buffer[(0.1% Triton X-100 (Sigma-Aldrich, T8787), 1× protease inhibitor (Roche, 05-892-970-001), 1 mM DTT (Sigma-Aldrich, D9779), RNase inhibitor (0.2 U/μl; Promega, N251B), and 2% BSA (Sigma-Aldrich, SRE0036) in PBS]. Sample was incubated on a rotator for 5 min at 4°C and then centrifuged at 500g for 5 min (Eppendorf, 5920R; 4°C, ramp speed of 3/3). Supernatant was removed and pellet was resuspended in 600 to 1000 μl of sort buffer [1 mM EDTA and RNase inhibitor (0.2 U/μl) in 2% BSA (Sigma-Aldrich, SRE0036) in PBS] and stained with DRAQ7 (1:100; Cell Signaling Technology, 7406). Seventy-five thousand nuclei were sorted using an SH800 sorter (Sony) into 50 μl of collection buffer [RNase inhibitor (1 U/μl) and 5% BSA (Sigma-Aldrich, SRE0036) in PBS]. Sorted nuclei were then centrifuged at 1000g for 15 min (Eppendorf, 5920R; 4°C, ramp speed of 3/3), and supernatant was removed. Nuclei were resuspended in 18 to 25 μl of reaction buffer [RNase inhibitor (0.2 U/μl) and 1% BSA (Sigma-Aldrich, SRE0036) in PBS] and counted using a hemocytometer. Twelve thousand nuclei were loaded onto a Chromium controller (10x Genomics). Libraries were generated using the Chromium Single Cell 3′ Library Construction Kit v3 (10x Genomics, 1000078) according to the manufacturer specifications. Complementary DNA was amplified for 12 PCR cycles. SPRISelect reagent (Beckman Coulter) was used for size selection and cleanup steps. Final library concentration was assessed by the Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific), and fragment size was checked using TapeStation High Sensitivity D1000 (Agilent) to ensure that fragment sizes were distributed normally around 500 bp. Libraries were sequenced using a NextSeq 500 or HiSeq 4000 (Illumina) using these read lengths: read 1, 28 cycles; read 2, 91 cycles; and index 1, 8 cycles.

hPSC culture

An engineered H9-hTnnTZ-pGZ-D2 hPSC transgenic reporter line was purchased from WiCell and maintained on Geltrex (Gibco) precoated tissue culture plates in E8 medium (68) containing Dulbecco's modified Eagle's medium/F12, L-ascorbic acid-2-phosphate magnesium (64 mg/liter), sodium selenium (14 μg/liter), fibroblast growth factor 2 (100 μg/liter), insulin (19.4 mg/liter), NaHCO₃ (543 mg/liter), transferrin (10.7 mg/liter), and transforming growth factor–β1(2 μg/liter). Cells were passaged every 3 to 5 days upon reaching ~80% confluency. For single-cell passaging experiments, cells were incubated with prewarmed TrypLE Select Enzyme, no phenol red

(1 ml per well of a six-well plate) for 2 to 3 min in a 37°C, 5% CO_2 incubator. Following incubation, cells were triturated to create a single-cell suspension and cultured in E8 medium supplied with Rock inhibitor (69) for 18 to 24 hours after split, followed by daily feeding with E8 medium.

In vitro cardiomyocyte differentiation

The H9-hTnnTZ-pGZ-D2 cell line was differentiated into beating cardiomyocytes using a previously reported Wnt-based monolayer differentiation protocol (70). Briefly, the H9-hTnnTZ-pGZ-D2 cell line was cultured in E8 medium for 3 to 10 passages. Before differentiation, hPSCs were seeded at a density of 350,000 to 400,000 cells per well of a 12-well plate and cultured for 2 days. For direct differentiation, cells were treated with 10 μM CHIR99021 (Fisher Scientific, no. 442350) in RPMI/B-27 without insulin. Fresh RPMI/B-27 without insulin media was replaced at post 24 hours, and cells were then cultured for 2 days. At day 3, cells were treated with 5 μM IWP2 (TOCRIS, no. 353310) in conditional medium and RPMI/B-27 without insulin 1:1 mix medium for another 2 days. At day 5, cells were exposed to fresh RPMI/B-27 without insulin media again for 2 days. Then, fresh RPMI/B-27 with insulin media was used and replenished every 2 days. Contracting cardiomyocytes were usually observed at days 7 and 8. D25 in vitro cardiomyocytes were purified using a PSC-derived cardiomyocyte isolation kit, human (Miltenyi Biotec, 130-110-188) and used for real-time qPCR (RT-qPCR).

Luciferase reporter assay

A genomic region harboring the *KCNH2* intronic enhancer (containing the risk allele rs7789146-G/rs7789585-G) was amplified by nested PCR (*KCNH2*-E-cF, CTGGCTGAAGACACCTTACTTT; *KCNH2*-E-cR, ACGGAGCAGTCAAGGAAAC and *KCNH2*-In-cF, CGGGGTACCCCTCCGTAAATGAGGTGCTATC; *KCNH2*-In-cR, CCCTCGAGACGGAGCAGTCAAGGAAAC) using the genomic DNA of H9-hTnnTZ-pGZ-D2 transgenic cells as a template and cloned into pGL4.23 [luc2/minP] (Promega, catalog no. E8411) luciferase reporter vector. Synthetic DNA containing the *KCNH2* intronic 5′-half enhancer (rs7789045-rs7789690) with the nonrisk/nonrisk allele (rs7789146-A/rs7789585-A), nonrisk/risk allele (rs7789146-A/rs7789585-G), and risk/nonrisk allele (rs7789146-G/rs7789585-A) were purchased from Integrated DNA Technologies. *KCNH2* intronic 3′-half enhancer (rs7789654-rs7790480) was amplified by PCR (*KCNH2*-R-In-cF, GCTGTGCAGTGTCAGGTTAT; *KCNH2*-In-cR, CCCTCGAGACGGAGCAGTCAAGGAAAC). Then, the whole *KCNH2* intronic enhancer with the nonrisk/nonrisk allele (rs7789146-A/rs7789585-A), nonrisk/risk allele (rs7789146-A/rs7789585-G), and risk/nonrisk allele (rs7789146-G/rs7789585-A) were generated by second PCR (*KCNH2*-In-cF, CGGGGTAC-CCCTCCGTAAATGAGGTGCTATC; *KCNH2*-In-cR, CCCTC-GAGACGGAGCAGTCAAGGAAAC). One day before transfection, 3×10^5 of D15 in vitro differentiated cardiomyocytes were plated in a Geltrex-coated 24-well plate. Cardiomyocytes were transfected with 500 ng of pGL4.23 plasmid (either empty, *KCNH2* enhancer with G/G allele, A/A allele, or mix) and 10 ng of TK:Renilla-luc as an internal control using Lipofectamine Stem Transfection Reagent (Invitrogen, no. STEM00003). Media were replaced with fresh media at 24 hours after transfection. At 72 hours after transfection, media were removed and the cells were washed with PBS. Luminescence was measured using a Dual-Luciferase Reporter Assay System (Promega, no. E2920) according to the manufacturer's protocol.

CRISPR-mediated genome editing experiments

To interrogate the functional significance of the AF-associated risk variant–containing cCRE at the *KCNH2* locus, the cCRE sequence was genetically deleted in H9-hTnnTZ-pGZ-D2 transgenic hPSCs using an efficient CRISPR-Cas9–mediated knockout system (*57*, *71*). Two adjacent guide RNAs (gRNAs) (*KCNH2*-enh gRNA-1, CT-CATTTACGGAGGAGCGCA; *KCNH2*-enh gRNA-2, TACAGTG-GCCTTCTAGACGA) targeting the cCRE were designed using a web-based software tool CRISPOR (*72*) based on targeting region of interest and minimizing potential off-target effects. The identified gRNAs were then synthesized in vitro using the GeneArt Precision gRNA Synthesis kit (Invitrogen) according to the manufacturer's protocol. One day before transfection, 1.5×10^5 H9-hTnnTZ-pGZ-D2 hPSCs were seeded in 12-well plates. A pair of ribonucleo-protein complexes containing 1.2 µg of Cas9 protein (New England Biolabs) and 400 ng of in vitro transcribed gRNA was then trans-fected (*73*, *74*) using Lipofectamine stem transfection reagent (Invitrogen). Seventy-two hours after the transfection, cells were diluted and clonally expanded another 7 days. Colonies were picked and lysates were prepared after the first passage for genotyping (*75*) (*KCNH2*-enh extended forward primer, ACACCTTACTTTGG-GTGAGAAG; *KCNH2*-enh extended reverse primer, AGACA-GAGCACAGACCTAGAA; *KCNH2*-enh internal forward primer, GCTGTGCAGTGTCAGGTTAT; *KCNH2*-enh internal reverse primer, TCTCCCTCCTTCTCTCTCATTC). After confirmation of genome-edited clones by Sanger sequencing, two transfected wild-type (WT) clones, two heterozygote clones, and two homozygote clones were selected for further functional analysis.

Real-time quantitative polymerase chain reaction

Total RNA was isolated from the cells using TRIzol reagent (Invitrogen). One microgram of total RNA was reverse-transcribed using the iScript Reverse Transcription Supermix kit (Bio-Rad) for RT-qPCR. RT-qPCR was performed using PowerUP SYBR Green Master Mix (Applied Biosystems) in the CFX Connect Real-Time System (Bio-Rad). The results were normalized to the *TBP* gene. The primers used for RT-qPCR are listed in table S21.

Electrophysiology of cardiomyocytes

Both WT and *KCNH2* enhancer knockout D15 in vitro cardiomyocytes were purified using the PSC-derived cardiomyocyte isolation kit, human (Miltenyi Biotec, 130-110-188) and cultured for another 10 to 20 days in a low density before electrophysiological measurements. The single-pipette, whole-cell patch current-clamp technique was used for recordings. Action potentials were recorded with a patch-clamp amplifier (Axopatch 200B, Axon), and experiments were performed at a temperature of 35° ± 0.5°C. Current-clamp command pulses were generated by a digital-to-analog converter (DigiData 1440, Axon), which was controlled by the pCLAMP software (10.3; Axon). Pipettes (resistance, 3 to 5 megohm) were pulled using a micropipette puller (Model P-87, Sutter Instrument Co.). Several minutes after seal formation, the membrane was ruptured by gentle suction to establish the whole-cell configuration for voltage clamping. Subsequently, the amplifier was switched to the current-clamp mode. Cells were paced with 1 Hz and injected current stimuli from 3 to 15 nA for 5-ms duration. Cells were superfused with extracellular solution containing the following: 140 mM NaCl, 5.4 mM KCl, 1.8 mM $CaCl_2$, 1.0 mM $MgCl_2$, 5.5 mM glucose, and 5.0 mM Hepes (pH 7.4 adjusted with NaOH). Pipette solution contained the following: 120 mM K-gluconate, 10 mM KCl, 5 mM NaCl, 10 mM Hepes, 5 mM phosphocreatine, 5 mM ATP-Mg_2, and amphotericin (0.44 µM; pH 7.2 adjusted with KOH).

Data analysis
Demultiplexing of snATAC-seq reads

For each sequenced snATAC-Seq library, we obtained four FASTQ files, two for paired-end DNA reads as well as the combinatorial indexes for i5 (768 different PCR indices) and T7 (96 different tag-mentation indices; table S20). We selected all reads with ≤2 mistakes per individual index (Hamming distance between each pair of indices is 4) and subsequently integrated the full barcode at the beginning of the read name in the demultiplexed FASTQ files (https://gitlab.com/Grouumf/ATACdemultiplex/).

Filtering of snATAC-seq profiles by transcriptional start site enrichment and unique fragments

TSS (transcriptional start site) positions were obtained from the GENCODE database v31 (*76*). Tn5-corrected insertions were aggregated ±2000 bp around each TSS genome wide. Then, this profile was normalized to the mean accessibility ± (1900 to 2000) bp from the TSS and smoothed every 11 bp. The maximum value of the smoothed profile was taken as the TSS enrichment. We selected all nuclei that had at least 1000 unique fragments and a TSS enrichment of at least 7 for all datasets.

Clustering strategy for snATAC-seq datasets

We used two rounds of clustering analysis to identify clusters. The first round of clustering analysis was performed on individual samples. We divided the genome into 5-kbp consecutive bins and then scored each nucleus for any insertions in these bins, generating a bin-by-cell binary matrix for each sample. We filtered out those bins that were generally accessible in all nuclei for each sample using a z score threshold of 1.65 (equivalent to a one-tailed P value < 0.05). On the basis of the filtered matrix, we then carried out dimensionality reduction followed by graph-based clustering to identify cell clusters. We called peaks using MACS2 (*32*) for each cluster using the aggregated profile of accessibility and then merged the peaks from all clusters to generate a union peak list. On the basis of the peak list, we generated a cell-by-peak count matrix and used Scrublet (*77*) to remove potential doublets with default parameters. Doublet scores returned by Scrublet (*77*) were then used to fit a two-component Gaussian mixture model using the BayesianGaussianMixture function from the python package scikit-learn (*78*). Nuclei in the component with the larger mean doublet score were removed from downstream analysis because they likely reflected doublets.

Next, to carry out the second round of clustering analysis, we merged peaks called from all samples to form a reference peak list. We generated a binary cell-by-peak matrix using nuclei from all samples and again performed dimensionality reduction followed by graph-based clustering to obtain the final cell clusters across the entire dataset.

Dimensionality reduction and batch correction of snATAC-seq data

For processing of snATAC-seq data, we adapted our previously published method, SnapATAC (*28*). To reduce the dimensionality of the peak by cell count matrix, SnapATAC uses spectral embedding for dimensionality reduction. To further increase the performance and scalability of spectral embedding, we applied the Nyström method (*79*) to enable handling of large datasets. Specifically, we first randomly sampled 35,000 nuclei as training data. We computed the

matrix $P = D^{-1}S\,S$, where D is the diagonal matrix such that $D_{ii} = \sum_j S_{ij}$. The eigendecomposition was performed on P, and the eigenvector with eigenvalue 1 was discarded. From the rest of the eigenvectors, we took k of them corresponding to the largest eigenvalues as the spectral embedding of the training data. We used the Nyström method (*79*) to extend the embedding to the data outside the training set. Given a set of unseen samples, we computed the similarity matrix S' between the new samples and the training set. The embedding of the new samples is given by $' = S'\,U\Lambda^{-1}$, where U and Λ are the eigenvectors and eigenvalues of P obtained in the previous step. To correct for donor/batch-specific effects, after dimensionality reduction, we performed cell grouping on individual samples using k-mean clustering with k equal to 20. We then constructed k-NN graphs for each sample and used the MNN (Mutual Nearest Neighbor) correction method to identify mutual nearest neighbors (*80*). These mutual nearest neighbors were used as anchors to match the cells between different samples and correct for donor/batch effects as described previously (*80*).

Clustering of snATAC-seq data

We constructed the k-nearest neighbor graph (k-NNG) using low-dimensional embedding of the nuclei with k equal to 50. We then applied the Leiden algorithm (*81*) with constant Potts model to find communities in the k-NNG corresponding to the cell clusters. The Leiden algorithm can be configured to use different quality functions. The modularity model is a popular choice, but it is hampered by the resolution limit, particularly when the network is large (*82*). Therefore, we used the modularity model only in the first round of clustering analysis to identify initial clusters. In the final round of clustering, we chose the constant Potts model as the quality function because it is resolution limit free and is better suited for identifying rare populations in a large dataset (*82*). Nuclei from two small clusters (280 and 254 nuclei) with low reproducibility and stability were discarded from downstream analysis. Thirty-four nuclei that formed clusters of 1 and 2 nuclei were discarded as well.

Processing and clustering analysis of snRNA-seq datasets

Raw sequencing data were demultiplexed and preprocessed using the Cell Ranger software package v3.0.2 (10x Genomics). Raw sequencing files were first converted from Illumina BCL files to FASTQ files using cellranger mkfastq. Demultiplexed FASTQs were aligned to the GRCh38 reference genome (10x Genomics), and reads for exonic and intronic reads mapping to protein coding genes, long noncoding RNA, antisense RNA, and pseudo-genes were used to generate a counts matrix using cellranger count; expect-cells parameter was set to 5000. A separate counts matrix for each sample was also generated using only reads mapped to intronic regions.

Next, exon and intron count matrices for individual datasets were processed using the Seurat v3.1.4 R package (*29*) (https://satijalab.org/seurat/) to assess dataset quality. Features represented in at least three cells and barcodes with between 500 and 4000 genes were used for downstream processing. In addition, barcodes with mitochondrial read percentages greater than 5% were removed. Counts were log-normalized and scaled by a factor of 10,000 using NormalizeData. To identify variable genes, FindVariableFeatures was run with default parameters except for nfeatures = 3000 to return the top 3000 variable genes. All genes were then scaled using ScaleData, which transforms the expression values for downstream analysis. Next, principal components analysis was performed using RunPCA with default parameters and the top 3000 variable features as input. The first 20 principal components were used to run clustering using FindNeighbors and FindClusters (parameter res = 0.4). To generate uniform manifold approximation and projection (UMAP) coordinates, RunUMAP was run using the first 20 principal components and with parameters umap.method = "umap-learn" and metric = "correlation". Doublet scores [pANN (proportion of Artificial Nearest Neighbors)] were generated for cell barcodes using DoubletFinder (https://github.com/chris-mcginnis-ucsf/DoubletFinder) (*83*) using the parameters pN = 0.15 and pK = 0.005; the anticipated collision rate was set by specifying 2% collisions per thousand nuclei for individual datasets.

Individual datasets were merged together using the merge function in Seurat to combine the count matrices and designate unique barcodes. Cell barcodes with pANN scores greater than 0 were removed from downstream analysis. Metadata were also encoded for each barcode, and the merged dataset was processed in a similar manner as described above; clusters were identified using FindNeighbors and FindClusters (res = 0.8). To generate the UMAP coordinates, the first 14 principal components were used in RunUMAP; the UMAP algorithm for Seurat v3.1.4 uses the uwot R package and that setting was used to generate the coordinates here. To regress out donor specific effects, the Harmony R package (https://github.com/immunogenomics/harmony) (*84*) was used, and the recomputed principal components were used to recluster the cells and rerun UMAP using the above parameters. For downstream analysis and comparison to snATAC-seq data, we manually combined ventricular cardiomyocyte clusters, atrial cardiomyocyte clusters, fibroblast clusters, and endothelial cell clusters on the basis of shared gene expression patterns (fig. S2, G and H). Cluster-specific genes in the all-transcripts dataset were identified in a global differential gene expression test using FindAllMarkers with parameters logFC = 0.25, min.pct = 0.25, and only.pos = FALSE.

Integration of snRNA-seq and snATAC-seq data

The snRNA-seq and snATAC-seq datasets were used to perform label transfer from the RNA cells onto the snATAC-seq dataset using the Seurat v3.1.4 R package (https://satijalab.org/seurat/) (*29*). Gene activity scores were calculated using chromatin accessibility in regions from the promoter up to 2 kb upstream for each ATAC nucleus. Activity scores were log-normalized and scaled using NormalizeData and ScaleData. To compare the snRNA and snATAC datasets and identify anchors, FindTransferAnchors was run considering the top 3000 variable features from the snRNA-seq dataset. Anchor pairs were used to assign RNA-seq labels to the snATAC-seq cells using TransferData, with the weight.reduction parameter set to the principal components used in snATAC-seq clustering. The efficacy of integration was assessed by examining the distribution of the maximum prediction scores output by TransferData and the distribution of annotated snATAC-seq identities to the corresponding predicted label.

Creation of a consensus list of heart cCREs

MACS2 (v2.1.2) (*32*) was used to identify accessible chromatin sites for each cluster with the following parameters: *-q 0.01 --nomodel --shift -100 --extsize 200 -g 2789775646 --call-summits --keepdup-all*. Estimated genome size was determined to be 2,789,775,646 bp and was indicated by the *-g* parameter. We next filtered out peaks overlapping with the ENCODE blacklist (hg38; https://github.com/Boyle-Lab/Blacklist/) (*85*). To generate the union of heart cCREs, we merged the blacklist-filtered peaks obtained for each cluster using the BEDtools merge command with default settings (v2.25.0) (*86*).

Computing relative accessibility scores for cCREs

To correct biases arising from differential read depth among cells and cell types, we derived a procedure that normalizes chromatin

accessibility at cCREs identified by MACS2 peak calling (v2.1.2) (*32*). We define the set of accessible loci by *L*, and we define a peak *p* as a subset of related loci *l* from *L*. Let a_l be the accessibility of accessible locus *l* and *P* the set of nonoverlapping peaks used to define the loci. For a given cell type $S_i \in S$, we computed the median med_j number of reads sequenced per cell. For each feature $p_j \in P$, we computed m_{ij}, the average number of reads sequenced from S_i, and overlapping p_j. We then defined the activity a_{ij} of loci p_j in S_i as $a_{ij} = 10^6 \times \frac{\sqrt[1/\text{med}_j]{(1 - m_{ij})}}{\sum_{j \in P} \sqrt[1/\text{med}_j]{(1 - m_{ij})}}$. We then define the relative accessibility score (RAS) $A_{ij} = \frac{a_{ij}}{\sum_{i \in S} a_{ij}}$.

K-means clustering of cCREs

We clustered the union of 287,415 cCREs using a *K*-means clustering procedure. We first created a sparse cell × peak matrix that was transformed into a RAS-normalized cell type × peak matrix. We then performed *K*-means on the normalized matrix with *K* from 2 to 12 and computed the Davies-Bouldin (DB) index for each *K* (*87*). Let $R_{xy} = \frac{(s_x + s_y)}{d_{xy}}$ with s_x the average distance of each cell of cluster *x* and d_{xy} the distance between the centroids of clusters *x* and *y*. The DB index is defined as $\text{DB} = \frac{1}{K} \sum_{x,y \in K} \max_{x \neq y}(R_{xy})$. We selected *K* = 9 because it resulted in the lowest DB index that indicates the best partition. We used the python library scikit-learn (*78*) to compute the *K*-means algorithm and the DB index (*87*).

Cell type annotation

We annotated snATAC-seq and snRNA-seq clusters on the basis of chromatin accessibility at promoter regions or expression of known lineage marker genes, respectively. We annotated atrial and ventricular cardiomyocytes on the basis of differential chromatin accessibility and gene expression at *NPPA*, *MYH6*, *KCNJ3*, *MYL7*, *MYH7*, *HEY2*, *MYL2*, and other reported markers of atrial and ventricular cardiomyocytes (*30*, *88*, *89*). We used, for example, the gene *DCN* to annotate cardiac fibroblasts (*90*), *VWF* and *EGFL7* for endothelial cells (*91*, *92*), *GJA4* and *TAGLN* for smooth muscle cells (*93*, *94*), *CD163* and *MS4A6A* for macrophages (*95*, *96*), *IL7R* and *THEMIS* for lymphocytes (*97*, *98*), *ADIPOQ* and *CIDEA* for adipocytes (*99*, *100*), and *NRXN3* and *GPM6B* for a cluster of nervous cells with neuronal and Schwann-like gene expression and chromatin accessibility signatures (*10*, *11*, *101*). From snRNA-seq, we identified a population of endothelial-like cells with specific expression of endocardial cell markers *NRG3* and *NPR3* (*102*, *103*). We also identified subtypes of mesenchymal cells that included myofibroblasts with characteristic expression of embryonic smooth muscle actin *MYH10* (*104*, *105*) as well as arterial smooth muscle cells with preferential expression of *ACTA2* and *TAGLN* relative to a larger cluster of pericytes (table S4) (*106*). snRNA-seq annotations were consistent with recent single-cell transcriptomic analyses of adult human heart tissue (*10*, *11*).

Identification of cell type–specific cCREs

We used edgeR (version 3.24) in R (*107*) to identify cell type–specific cCREs. For each cCRE, accessibility within a cell type was compared to average accessibility in all other clusters. For each cell type, we created a count table for each cCRE using the following strategy: Each sample was described with a donor and a chamber ID. For each sample ID, we reported read count within (i) the cell type and (ii) the rest of the cell types in aggregate. We used this count matrix as input for edgeR analysis (*107*) and performed a likelihood ratio test. We considered cCREs with fold change > 1.2 and FDR < 0.01 after Benjamini-Hochberg correction as cell type specific.

Coaccessibility analysis using Cicero

We used the R package Cicero (*38*) to infer coaccessible chromatin loci. For each chromosome, we used as input the corresponding peaks from our 287,415 cCRE union set and the coordinates of the snATAC-seq UMAP (*108*). We randomly subsampled 15,000 cells from our aggregate snATAC-seq dataset to construct input matrices for Cicero analysis. We used ±250 kbp as cutoff for coaccessibility interactions. All other settings were default.

Correlation of gene expression and promoter accessibility

We defined promoter regions as TSS ±2 kbp. TSS were extracted from annotation files from GENCODE release 33 (*76*). We identified promoter-overlapping peaks using BEDtools (*86*) and a custom script (see the "Code availability" section). For each overlapping pair (peak and promoter) identified, we kept only the open chromatin site closest to the TSS to obtain a 1:1 correspondence between genes and open chromatin peaks. We then used the RAS and the cluster-scaled gene expression (*29*) to create feature × cell type matrices for RNA-seq and ATAC-seq datasets. We then used these matrices to create heatmaps and to perform ATAC-seq/RNA-seq cluster correlation analysis using the Pearson similarity metric. For each cell type, we computed the Pearson correlation score between the RAS vector of the 7081 promoters and the scaled expression vector of the corresponding 7081 genes identified via the 1:1 correspondence method described above.

DA cCREs between heart chambers

Between-heart chamber differential accessibility analysis was performed for five cell types from our aggregated snATAC-seq dataset. We considered only cell types that had a representation of at least 50 nuclei per dataset and at least 300 nuclei across each tested condition. The cell types that met these inclusion criteria included cardiomyocytes, fibroblasts, endothelial cells, smooth muscle cells, and macrophages. Within each cell type, a generalized linear model framework was used using the R package edgeR (*107*). All fragments for a given cell type were aggregated in the .bed format. MACS2 (*32*) was used to call peaks on the aggregate .bed file for each cell type with the parameters specified above. NarrowPeak output bed files were used for differential accessibility testing. The aggregate .bed file for each cell type was then partitioned on the basis of dataset of origin using nuclear barcodes. The "coverage" option of the BEDtools package (*86*) was applied with default settings to count the total number of chromatin fragments from each dataset overlapping narrowPeaks called on the aggregate .bed file for the corresponding cell type. This yielded a raw count matrix in the format of snATAC-seq datasets (columns) by narrowPeaks (rows) for each cell type. The raw count matrix was used as input for edgeR analysis. To filter low-coverage peaks from our analysis, we used the "filterByExpr" command within edgeR with default settings. We applied an average prior count of one during fitting of the generalized linear model to avoid inflated fold changes in instances for which peaks lacked coverage for one but not both tested conditions. We modeled chromatin accessibility at each peak as a function of heart chamber (group) with sex as a covariate. The generalized linear model was expressed as follows in edgeR notation:

design <− model.matrix(~sex + group)
y <− estimateDisp(y, design, prior.count = 1)
glmFit(y, design)

Significance was tested using a likelihood ratio test. To account for testing multiple hypotheses, a Benjamini-Hochberg significance correction was applied for all cCREs tested within each considered cell type. Any cCRE with an absolute $\log_2$(fold change) > 1 and an FDR-corrected P value < 0.05 was considered significant.

Gene expression analysis of genes coaccessible with DA cCREs
To compare the expression of genes coaccessible with heart chamber–dependent distal DA cCREs (outside ±2 kb of TSS) in cardiomyocytes and fibroblasts, we performed differential expression testing for all genes between indicated heart chambers using Wilcoxon rank sum test in Seurat (*29*). Genes with an absolute fold change > 1.5 and an FDR-adjusted P value < 0.05 were considered differentially expressed. We then tested the resulting genes for coaccessibility (*38*) with distal DA cCREs at a coaccessibility score threshold of 0.1 and displayed scaled gene expression values from Seurat for the indicated differentially expressed genes linked to chamber-dependent distal DA cCREs.

Genomic Regions Enrichment of Annotations Tool ontology analysis
The Genomic Regions Enrichment of Annotations Tool (GREAT; http://great.stanford.edu/public/html/index.php) (*33*) was used with default settings for indicated cCREs or candidate enhancers in the .bed format. Biological process enrichments are reported. P values shown for enrichment are Bonferroni-corrected binomial P values.

Motif enrichment analysis
For de novo and known motif enrichment analysis of cluster-specific cCREs, the findMotifsGenome.pl utility of the HOMER package was used with default settings (*35*). For display of enrichment patterns for motifs from the JASPAR (*109*) database with evidence of enrichment in at least one set of cell type–specific cCREs, motifs with an enrichment P value < 10^{-5} in at least one set of cluster-specific cCREs were selected. For motif enrichment within DA cCREs, narrowPeak calls from MACS2 were used as input, with peaks called on the corresponding cell type (as described above) used as background. For enrichment of motifs within cell type–attributed bulk enhancers, snATAC-seq peaks from the union of snATAC-seq peaks were used. Summits were extracted from peaks that overlapped bulk enhancer annotations and extended by 250 bp on either side to obtain fixed-width peaks. We also computed motif enrichment scores at single-cell resolution using chromVAR (*34*).

For input to chromVAR, we used the summits of the 287,415 peaks in our consensus list extended by 250 bp in either direction and a set of 870 nonredundant motifs as input (https://github.com/GreenleafLab/chromVARmotifs). To identify differentially enriched motifs in each cell type, we used the following strategy: For each cell type and each motif, we computed a rank sum test between the chromVAR z score distributions from cells within the cell type and outside of the cell type. Tests were run using a random sampling of 40,000 cells. Then, for each cell type, we used 1×10^{-8} as P value cutoff. In addition, we applied a Bonferroni correction to account for multiple testing, which resulted in selection of significant motifs with P value < 1×10^{-11}.

Measuring single-cell chromatin accessibility signal within bulk candidate heart enhancers
We obtained published candidate heart enhancers annotated by H3K27ac ChIP-seq from a recently reported bulk survey of healthy left ventricular tissue from 18 human donors (*15*). Candidate enhancers were defined per the study as H3K27ac ChIP-seq peaks that were at least 1 kb away from a TSS and present in two or more

donors. Because these reference annotations were derived from bulk profiling of healthy left ventricles, we selected only left ventricular nuclei from our aggregate dataset for comparison. We limited our analysis to cell types that comprised at least 5% of nuclei by proportion in our aggregate dataset. These included cardiomyocytes, fibroblasts, endothelial cells, smooth muscle cells, and macrophages. We first combined all fragments for each cell type from left ventricular datasets. The coverage option of BEDtools (*86*) was applied with default settings to count the total number of chromatin fragments from each ventricular cell type overlapping the candidate enhancer annotations. This yielded a raw count matrix in the format of snATAC-seq cell types (columns) by candidate enhancers (rows). The raw count matrix was normalized to RPKM (reads per kilobase per million mapped reads) for each candidate enhancer. We next used Cluster3.0 (*110*) to k-means cluster the 31,033 healthy heart candidate enhancers into K groups between 2 and 12 with the following settings (Method = k-Means, Similarity Metric = *Euclidian distance*, number of runs = 100). We calculated the DB index (*87*) as described above for each clustering using the index.DB function of the R package clusterSim (http://keii.ue.wroc.pl/clusterSim/). We selected a k-means of 8, which yielded the lowest DB index, indicating the best partitioning.

We repeated the above analysis for 4406 candidate enhancers reported to have increased bulk H3K27ac ChIP signal and 3101 candidate enhancers reported to have decreased signal in 18 late-stage idiopathic dilated cardiomyopathy (heart failure) left ventricles versus 18 healthy control left ventricles reported in the same study. We again clustered the candidate enhancers for both groups into k groups between 2 and 12 as above and selected the clustering that yielded the lowest DB index (*87*).

GWAS variant enrichment analysis
We used LD score regression (*48*, *111*) to estimate genome-wide enrichment for GWAS traits using annotation sets from single-cell chromatin accessibility from the heart or lung (*54*) or bulk DNase hypersensitivity sites for tissues from ENCODE (*3*, *5*, *17*). For bulk DNase-seq datasets, peak annotations were merged across biological replicates from the same tissue type. We compiled published GWAS summary statistics for cardiovascular diseases (*49–53*), other diseases (*112–122*), and nondisease traits (*123–132*) using the European subset from transethnic studies where applicable. We created custom LD score files by using peaks from each cell type or tissue as a binary annotation. As background, we used baseline annotations included in the baseline-LD model v2.2. For each trait, we used LD score regression to estimate enrichment coefficient z scores for each annotation relative to the background. Using these z scores, we computed two-sided P values for enrichment and used the Benjamini-Hochberg procedure to correct for multiple tests within each set of annotations.

Fine mapping for AF
We obtained published AF GWAS summary statistics and index variants for 111 disease-associated loci (*50*). To construct credible sets of variants for each locus, we first extracted all variants in LD ($r^2 > 0.1$ using the EUR (European) subset of 1000 Genomes phase 3) (*133*) in a large window (±2.5 Mb) around each index variant. We next calculated approximate Bayes factors (ABF) (*55*) for each variant using effect size and SE estimates. We then calculated PPAs for each variant by dividing its ABF by the sum of ABF for all variants within the locus. For each locus, we then defined 99% credible sets by sorting variants by descending PPA and retaining variants that

added up to a cumulative PPA of >0.99. This resulted in an output of 6014 candidate causal variants.

Variant prioritization for functional validation

To prioritize variants for functional validation, we refined our list of candidate causal variants from fine mapping analysis to only those with a PPA > 0.1 (216 remaining of 6014). We used BEDtools (86) to intersect these variants with ATAC-seq peaks called on an aggregate .bed file for atrial and ventricular cardiomyocyte snATAC-seq clusters (cardiomyocyte cCREs). This resulted in 40 fine-mapped variants that resided within 38 candidate cardiomyocyte cCREs (38 cCRE-variant pairs).

We assessed each remaining cCRE-variant pair via the following criteria:

1) cCREs primarily accessible in cardiomyocytes

2) presence of a corresponding ATAC-seq peak at a testable time point in the in vitro hPSC-cardiomyocyte differentiation model system

3) sequence conservation in 100 vertebrates [genome browser track generated using phyloP of the PHAST5 package downloaded from UCSC genome browser (134), http://hgdownload.soe.ucsc.edu/goldenPath/hg38/phyloP100way/]

4) predicted coaccessibility of candidate enhancer with a gene promoter

5) expression of putative target gene associated with cCRE appearance (chromatin accessibility and H3K27ac) during hPSC-cardiomyocyte differentiation (57)

A candidate cCRE-variant pair at the *KCNH2* locus was prioritized for functional experimentation.

ChIP-seq data processing

Reads were mapped to the human genome reference GRCh38 using Bowtie2 (version 2.2.6) (135) and reads with MAPQ (mapping quality) > 30 selected using SAMtools (version 1.3.1) (136). PCR duplicates were removed using the MarkDuplicates function of Picard tools (version 1.119) (137). RPKM-normalized signal tracks were generated using the BamCoverage function in deepTools (version 2.4.1) (138).

RNA-seq data processing

Reads were mapped to the human genome reference GRCh38 using STAR (version 020201) (139) and reads with MAPQ > 30 selected using SAMtools (version 1.3.1) (136). PCR duplicates were removed using the MarkDuplicates function of Picard tools (version 1.1.19) (137). RPKM-normalized signal tracks were generated using the BamCoverage function in deepTools (version 2.4.1) (138).

ATAC-seq data processing

Reads were mapped to the human genome reference GRCh38 using Bowtie2 (version 2.2.6) (135) and reads with MAPQ > 30 selected using SAMtools (version 1.3.1) (136). PCR duplicates were removed using SAMtools (version 1.3.1) (136). RPKM-normalized signal tracks were generated using the BamCoverage function in deepTools (version 2.4.1) (138).

Statistical analysis

No statistical methods were used to predetermine sample sizes. There was no randomization of the samples, and investigators were not blinded to the specimens being investigated. However, clustering of single nuclei based on chromatin accessibility was performed in an unbiased manner, and cell types were assigned after clustering. Low-quality nuclei and potential barcode collisions were excluded from downstream analysis as outlined above. Cluster specificity at each cCRE was tested using edgeR (107) as described above, with *P* values corrected via the Benjamini-Hochberg procedure. To identify DA sites between the heart chambers and for each cell type, a likelihood ratio test was used and the resulting *P* value was corrected using the Benjamini-Hochberg procedure. For significance of ontology enrichments using GREAT, Bonferroni-corrected binomial *P* values were used (33). For significance testing of enrichment of de novo and known motifs, a hypergeometric test was used without correction for multiple testing (35). For luciferase and qPCR data, we performed one-way ANOVA (analysis of variance) analysis with post hoc Tukey test using GraphPad Prism version 8.0.0 for Windows (GraphPad Software, San Diego, CA, USA; www.graphpad.com).

External datasets

Cardiomyocyte differentiation: RNA-seq, H3K27ac day 0 (hPSC); day 5 (cardiac mesoderm); and day 15 (primitive cardiomyocytes) were downloaded from GSE116862 (57). Signal tracks for heart H3K27ac ChIP-seq data were downloaded from https://portal.nersc.gov/dna/RD/heart/. List of candidate enhancers was downloaded from supplementary tables (15). H3K27ac ChIP-seq data for cardiomyocyte nuclei from nonfailing donors (NF1) were downloaded from National Center for Biotechnology Information Sequence Read Archive BioProject ID PRJNA353755 (140). We acquired snATAC-seq data for human lung from GSE161383 (54) and bulk DNase-seq datasets for human tissues from ENCODE (3, 5, 17) with the following identifiers: ENCSR053ZKP, ENCSR259GYP, ENCSR277KRY, ENCSR458AOS, ENCSR597NVK, ENCSR422IIZ, ENCSR968TPO, ENCSR788IZL, ENCSR859LTL, ENCSR060HPL, ENCSR783OCW, ENCSR171ADO, ENCSR171ETY, ENCSR520BAD, ENCSR686WJL, ENCSR791BHE, ENCSR856XLJ, ENCSR361DND, ENCSR579KDC, ENCSR693UHT, ENCSR365NDK, ENCSR712PYJ, ENCSR930PDT, ENCSR178JBL, ENCSR595HZQ, ENCSR261RWJ, ENCSR455GUW, ENCSR689DSM, ENCSR954AJK, ENCSR564FZH, ENCSR000ELO, ENCSR909HFI, ENCSR931UQB, ENCSR128GBN, ENCSR006IMH, ENCSR163PKT, ENCSR641ZPF, ENCSR782SSS, ENCSR101QXF, ENCSR709IYR, ENCSR866ODX, ENCSR195ONB, ENCSR450PWF, ENCSR549NRK, ENCSR749MUH, ENCSR080ISA, ENCSR090IDV, ENCSR102RSU, ENCSR401ESD, ENCSR484UAU, ENCSR508FVM, ENCSR340MRJ, ENCSR760QZM, ENCSR763AKE, ENCSR923JYH, ENCSR164WOF, ENCSR323UTX, ENCSR650FLQ, ENCSR702DPD, ENCSR129BZE, and ENCSR437AYW.

Code availability

The pipeline for processing snATAC-seq data is available as a part of the Taiji software: https://taiji-pipeline.github.io/. Custom code used for demultiplexing and downstream analysis for snATAC data is available here: https://gitlab.com/Grouumf/ATACdemultiplex/-/tree/master/ATACdemultiplex and https://gitlab.com/Grouumf/ATACdemultiplex/-/blob/master/scripts/.

REFERENCES AND NOTES

1. WHO, Cardiovascular Diseases (CVDs) (2017). The World Health Organization https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds).
2. Roadmap Epigenomics Consortium, A. Kundaje, W. Meuleman, J. Ernst, M. Bilenky, A. Yen, A. Heravi-Moussavi, P. Kheradpour, Z. Zhang, J. Wang, M. J. Ziller, V. Amin, J. W. Whitaker, M. D. Schultz, L. D. Ward, A. Sarkar, G. Quon, R. S. Sandstrom, M. L. Eaton, Y.-C. Wu, A. R. Pfenning, X. Wang, M. Claussnitzer, Y. Liu, C. Coarfa, R. A. Harris, N. Shoresh, C. B. Epstein, E. Gjoneska, D. Leung, W. Xie, R. D. Hawkins, R. Lister, C. Hong, P. Gascard,

A. J. Mungall, R. Moore, E. Chuah, A. Tam, T. K. Canfield, R. S. Hansen, R. Kaul, P. J. Sabo, M. S. Bansal, A. Carles, J. R. Dixon, K.-H. Farh, S. Feizi, R. Karlic, A.-R. Kim, A. Kulkarni, D. Li, R. Lowdon, G. N. Elliott, T. R. Mercer, S. J. Neph, V. Onuchic, P. Polak, N. Rajagopal, P. Ray, R. C. Sallari, K. T. Siebenthall, N. A. Sinnott-Armstrong, M. Stevens, R. E. Thurman, J. Wu, B. Zhang, X. Zhou, A. E. Beaudet, L. A. Boyer, P. L. De Jager, P. J. Farnham, S. J. Fisher, D. Haussler, S. J. M. Jones, W. Li, M. A. Marra, M. T. McManus, S. Sunyaev, J. A. Thomson, T. D. Tlsty, L.-H. Tsai, W. Wang, R. A. Waterland, M. Q. Zhang, L. H. Chadwick, B. E. Bernstein, J. F. Costello, J. R. Ecker, M. Hirst, A. Meissner, A. Milosavljevic, B. Ren, J. A. Stamatoyannopoulos, T. Wang, M. Kellis, Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317–330 (2015).

3. ENCODE Project Consortium, An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).

4. R. E. Thurman, E. Rynes, R. Humbert, J. Vierstra, M. T. Maurano, E. Haugen, N. C. Sheffield, A. B. Stergachis, H. Wang, B. Vernot, K. Garg, S. John, R. Sandstrom, D. Bates, L. Boatman, T. K. Canfield, M. Diegel, D. Dunn, A. K. Ebersol, T. Frum, E. Giste, A. K. Johnson, E. M. Johnson, T. Kutyavin, B. Lajoie, B.-K. Lee, K. Lee, D. London, D. Lotakis, S. Neph, F. Neri, E. D. Nguyen, H. Qu, A. P. Reynolds, V. Roach, A. Safi, M. E. Sanchez, A. Sanyal, A. Shafer, J. M. Simon, L. Song, S. Vong, M. Weaver, Y. Yan, Z. Zhang, Z. Zhang, B. Lenhard, M. Tewari, M. O. Dorschner, R. S. Hansen, P. A. Navas, G. Stamatoyannopoulos, V. R. Iyer, J. D. Lieb, S. R. Sunyaev, J. M. Akey, P. J. Sabo, R. Kaul, T. S. Furey, J. Dekker, G. E. Crawford, J. A. Stamatoyannopoulos, The accessible chromatin landscape of the human genome. *Nature* **489**, 75–82 (2012).

5. ENCODE Project Consortium, J. E. Moore, M. J. Purcaro, H. E. Pratt, C. B. Epstein, N. Shoresh, J. Adrian, T. Kawli, C. A. Davis, A. Dobin, R. Kaul, J. Halow, E. L. Van Nostrand, P. Freese, D. U. Gorkin, Y. Shen, Y. He, M. Mackiewicz, F. Pauli-Behn, B. A. Williams, A. Mortazavi, C. A. Keller, X.-O. Zhang, S. I. Elhajjajy, J. Huey, D. E. Dickel, V. Snetkova, X. Wei, X. Wang, J. C. Rivera-Mulia, J. Rozowsky, J. Zhang, S. B. Chhetri, J. Zhang, A. Victorsen, K. P. White, A. Visel, G. W. Yeo, C. B. Burge, E. Lécuyer, D. M. Gilbert, J. Dekker, J. Rinn, E. M. Mendenhall, J. R. Ecker, M. Kellis, R. J. Klein, W. S. Noble, A. Kundaje, R. Guigó, P. J. Farnham, J. M. Cherry, R. M. Myers, B. Ren, B. R. Graveley, M. B. Gerstein, L. A. Pennacchio, M. P. Snyder, B. E. Bernstein, B. Wold, R. C. Hardison, T. R. Gingeras, J. A. Stamatoyannopoulos, Z. Weng, Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature* **583**, 699–710 (2020).

6. S. Smemo, L. C. Campos, I. P. Moskowitz, J. E. Krieger, A. C. Pereira, M. A. Nobrega, Regulatory variation in a TBX5 enhancer leads to isolated congenital heart disease. *Hum. Mol. Genet.* **21**, 3255–3263 (2012).

7. S. Zaidi, M. Brueckner, Genetics and genomics of congenital heart disease. *Circ. Res.* **120**, 923–940 (2017).

8. M. T. Maurano, R. Humbert, E. Rynes, R. E. Thurman, E. Haugen, H. Wang, A. P. Reynolds, R. Sandstrom, H. Qu, J. Brody, A. Shafer, F. Neri, K. Lee, T. Kutyavin, S. Stehling-Sun, A. K. Johnson, T. K. Canfield, E. Giste, M. Diegel, D. Bates, R. S. Hansen, S. Neph, P. J. Sabo, S. Heimfeld, A. Raubitschek, S. Ziegler, C. Cotsapas, N. Sotoodehnia, I. Glass, S. R. Sunyaev, R. Kaul, J. A. Stamatoyannopoulos, Systematic localization of common disease-associated variation in regulatory DNA. *Science* **337**, 1190–1195 (2012).

9. Y. Cui, Y. Zheng, X. Liu, L. Yan, X. Fan, J. Yong, Y. Hu, J. Dong, Q. Li, X. Wu, S. Gao, J. Li, L. Wen, J. Qiao, Single-cell transcriptome analysis maps the developmental track of the human heart. *Cell Rep.* **26**, 1934–1950.e5 (2019).

10. M. Litvinukova, C. Talavera-López, H. Maatz, D. Reichart, C. L. Worth, E. L. Lindberg, M. Kanda, K. Polanski, M. Heinig, M. Lee, E. R. Nadelmann, K. Roberts, L. Tuck, E. S. Fasouli, D. M. De Laughter, B. M. Donough, H. Wakimoto, J. M. Gorham, S. Samari, K. T. Mahbubani, K. Saeb-Parsy, G. Patone, J. J. Boyle, H. Zhang, H. Zhang, A. Viveiros, G. Y. Oudit, O. A. Bayraktar, J. G. Seidman, C. E. Seidman, M. Noseda, N. Hubner, S. A. Teichmann, Cells of the adult human heart. *Nature* **588**, 466–472 (2020).

11. N. R. Tucker, M. Chaffin, S. J. Fleming, A. W. Hall, V. A. Parsons, K. C. Bedi Jr., A. D. Akkad, C. N. Herndon, A. Arduini, I. Papangeli, C. Roselli, F. Aguet, S. H. Choi, K. G. Ardlie, M. Babadi, K. B. Margulies, C. M. Stegmann, P. T. Ellinor, Transcriptional and cellular diversity of the human heart. *Circulation* **142**, 466–482 (2020).

12. M. Asp, S. Giacomello, L. Larsson, C. Wu, D. Fürth, X. Qian, E. Wärdell, J. Custodio, J. Reimegård, F. Salmén, C. Österholm, P. L. Ståhl, E. Sundström, E. Åkesson, O. Bergmann, M. Bienko, A. Månsson-Broberg, M. Nilsson, C. Sylvén, J. Lundeberg, A spatiotemporal organ-wide gene expression and cell atlas of the developing human heart. *Cell* **179**, 1647–1660.e19 (2019).

13. D. E. Dickel, I. Barozzi, Y. Zhu, Y. Fukuda-Yuzawa, M. Osterwalder, B. J. Mannion, D. May, C. H. Spurrell, I. Plajzer-Frick, C. S. Pickle, E. Lee, T. H. Garvin, M. Kato, J. A. Akiyama, V. Afzal, A. Y. Lee, D. U. Gorkin, B. Ren, E. M. Rubin, A. Visel, L. A. Pennacchio, Genome-wide compendium and functional assessment of in vivo heart enhancers. *Nat. Commun.* **7**, 12923–12923 (2016).

14. D. May, M. J. Blow, T. Kaplan, D. McCulley, B. C. Jensen, J. A. Akiyama, A. Holt, I. Plajzer-Frick, M. Shoukry, C. Wright, V. Afzal, P. C. Simpson, E. M. Rubin, B. L. Black, J. Bristow, L. A. Pennacchio, A. Visel, Large-scale discovery of enhancers from human heart tissue. *Nat. Genet.* **44**, 89–93 (2011).

15. C. H. Spurrell, I. Barozzi, B. J. Mannion, M. J. Blow, Y. Fukuda-Yuzawa, S. Y. Afzal, J. A. Akiyama, V. Afzal, S. Tran, I. Plajzer-Frick, C. S. Novak, M. Kato, E. Lee, T. H. Garvin, Q. T. Pham, A. N. Harrington, M. Lisgo, J. Bristow, T. P. Cappola, M. P. Morley, K. B. Margulies, L. A. Pennacchio, D. E. Dickel, A. Visel, Genome-wide fetalization of enhancer architecture in heart disease. *bioRxiv*, 591362 (2019).

16. R. Gilsbach, M. Schwaderer, S. Preissl, B. A. Grüning, P. Kranzhöfer, P. Schneider, T. G. Nührenberg, S. Mulero-Navarro, D. Weichenhan, C. Braun, M. Dreßen, A. R. Jacobs, H. Lahm, T. Doenst, M. Krane, B. D. Gelb, L. Hein, Distinct epigenetic programs regulate cardiac myocyte development and disease in the human heart in vivo. *Nat. Commun.* **9**, 391 (2018).

17. J. Vierstra, J. Lazar, R. Sandstrom, J. Halow, K. Lee, D. Bates, M. Diegel, D. Dunn, F. Neri, E. Haugen, E. Rynes, A. Reynolds, J. Nelson, A. Johnson, M. Frerker, M. Buckley, R. Kaul, W. Meuleman, J. A. Stamatoyannopoulos, Global reference mapping of human transcription factor footprints. *Nature* **583**, 729–736 (2020).

18. W. L. W. Tan, C. G. Anene-Nzelu, E. Wong, C. J. M. Lee, H. S. Tan, S. J. Tang, A. Perrin, K. X. Wu, W. Zheng, R. J. Ashburn, B. Pan, M. Y. Lee, M. I. Autio, M. P. Morley, W. L. Tam, C. Cheung, K. B. Margulies, L. Chen, T. P. Cappola, M. Loh, J. Chambers, S. Prabhakar, R. S.-Y. Foo, J. Assa, P. Ellinor, J. Felix, P. van der Harst, B. Meder, G. J. Smith, R. Tappu, Epigenomes of human hearts reveal new genetic variants relevant for cardiac disease and phenotype. *Circ. Res.* **127**, 761–777 (2020).

19. C. G. Anene-Nzelu, W. L. W. Tan, C. J. M. Lee, Z. Wenhao, A. Perrin, A. Dashi, Z. Tiang, M. I. Autio, B. Lim, E. Wong, H. S. Tan, B. Pan, M. P. Morley, K. B. Margulies, T. P. Cappola, R. S.-Y. Foo, Assigning distal genomic enhancers to cardiac disease-causing genes. *Circulation* **142**, 910–912 (2020).

20. J. D. Buenrostro, B. Wu, U. M. Litzenburger, D. Ruff, M. L. Gonzales, M. P. Snyder, H. Y. Chang, W. J. Greenleaf, Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature* **523**, 486–490 (2015).

21. D. A. Cusanovich, R. Daza, A. Adey, H. A. Pliner, L. Christiansen, K. L. Gunderson, F. J. Steemers, C. Trapnell, J. Shendure, Multiplex single cell profiling of chromatin accessibility by combinatorial cellular indexing. *Science* **348**, 910–914 (2015).

22. A. T. Satpathy, J. M. Granja, K. E. Yost, Y. Qi, F. Meschi, G. P. McDermott, B. N. Olsen, M. R. Mumbach, S. E. Pierce, M. R. Corces, P. Shah, J. C. Bell, D. Jhutty, C. M. Nemec, J. Wang, L. Wang, Y. Yin, P. G. Giresi, A. L. S. Chang, G. X. Y. Zheng, W. J. Greenleaf, H. Y. Chang, Massively parallel single-cell chromatin landscapes of human immune cell development and intratumoral T cell exhaustion. *Nat. Biotechnol.* **37**, 925–936 (2019).

23. C. A. Lareau, F. M. Duarte, J. G. Chew, V. K. Kartha, Z. D. Burkett, A. S. Kohlway, D. Pokholok, M. J. Aryee, F. J. Steemers, R. Lebofsky, J. D. Buenrostro, Droplet-based combinatorial indexing for massive-scale single-cell chromatin accessibility. *Nat. Biotechnol.* **37**, 916–924 (2019).

24. D. A. Cusanovich, A. J. Hill, D. Aghamirzaie, R. M. Daza, H. A. Pliner, J. B. Berletch, G. N. Filippova, X. Huang, L. Christiansen, W. S. De Witt, C. Lee, S. G. Regalado, D. F. Read, F. J. Steemers, C. M. Disteche, C. Trapnell, J. Shendure, A single-cell atlas of in vivo mammalian chromatin accessibility. *Cell* **174**, 1309–1324.e18 (2018).

25. G. Jia, J. Preussner, X. Chen, S. Guenther, X. Yuan, M. Yekelchyk, C. Kuenne, M. Looso, Y. Zhou, S. Teichmann, T. Braun, Single cell RNA-seq and ATAC-seq analysis of cardiac progenitor cell transition states and lineage settlement. *Nat. Commun.* **9**, 4877 (2018).

26. S. Domcke, A. J. Hill, R. M. Daza, J. Cao, D. R. O'Day, H. A. Pliner, K. A. Aldinger, D. Pokholok, F. Zhang, J. H. Milbank, M. A. Zager, I. A. Glass, F. J. Steemers, D. Doherty, C. Trapnell, D. A. Cusanovich, J. Shendure, A human cell atlas of fetal chromatin accessibility. *Science* **370**, eaba7612 (2020).

27. S. Preissl, R. Fang, H. Huang, Y. Zhao, R. Raviram, D. U. Gorkin, Y. Zhang, B. C. Sos, V. Afzal, D. E. Dickel, S. Kuan, A. Visel, L. A. Pennacchio, K. Zhang, B. Ren, Single-nucleus analysis of accessible chromatin in developing mouse forebrain reveals cell-type-specific transcriptional regulation. *Nat. Neurosci.* **21**, 432–439 (2018).

28. R. Fang, S. Preissl, Y. Li, X. Hou, J. Lucero, X. Wang, A. Motamedi, A. K. Shiau, X. Zhou, F. Xie, E. A. Mukamel, K. Zhang, Y. Zhang, M. Margarita Behrens, J. R. Ecker, B. Ren, Comprehensive analysis of single cell ATAC-seq data with SnapATAC. *Nat. Commun.* **21**, 1337 (2021).

29. T. Stuart, A. Butler, P. Hoffman, C. Hafemeister, E. Papalexi, W. M. Mauck III, Y. Hao, M. Stoeckius, P. Smibert, R. Satija, Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902.e21 (2019).

30. S. Y. Ng, C. K. Wong, S. Y. Tsang, Differential gene expressions in atrial and ventricular myocytes: Insights into the road of applying embryonic stem cell-derived cardiomyocytes for future therapies. *Am. J. Physiol. Cell Physiol.* **299**, C1234–C1249 (2010).

31. H. Chen, C. Lareau, T. Andreani, M. E. Vinyard, S. P. Garcia, K. Clement, M. A. Andrade-Navarro, J. D. Buenrostro, L. Pinello, Assessment of computational methods for the analysis of single-cell ATAC-seq data. *Genome Biol.* **20**, 241 (2019).

32. Y. Zhang, T. Liu, C. A. Meyer, J. Eeckhoute, D. S. Johnson, B. E. Bernstein, C. Nussbaum, R. M. Myers, M. Brown, W. Li, X. S. Liu, Model-based Analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).

33. C. Y. McLean, D. Bristor, M. Hiller, S. L. Clarke, B. T. Schaar, C. B. Lowe, A. M. Wenger, G. Bejerano, GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.* **28**, 495–501 (2010).

34. A. N. Schep, B. Wu, J. D. Buenrostro, W. J. Greenleaf, chromVAR: Inferring transcription-factor-associated accessibility from single-cell epigenomic data. *Nat. Methods* **14**, 975–978 (2017).

35. S. Heinz, C. Benner, N. Spann, E. Bertolino, Y. C. Lin, P. Laslo, J. X. Cheng, C. Murre, H. Singh, C. K. Glass, Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).

36. D. E. Zhang, C. J. Hetherington, H. M. Chen, D. G. Tenen, The macrophage transcription factor PU.1 directs tissue-specific expression of the macrophage colony-stimulating factor receptor. *Mol. Cell. Biol.* **14**, 373–381 (1994).

37. A. F. M. Moorman, V. M. Christoffels, Cardiac chamber formation: Development, genes, and evolution. *Physiol. Rev.* **83**, 1223–1267 (2003).

38. H. A. Pliner, J. S. Packer, J. L. McFaline-Figueroa, D. A. Cusanovich, R. M. Daza, D. Aghamirzaie, S. Srivatsan, X. Qiu, D. Jackson, A. Minkina, A. C. Adey, F. J. Steemers, J. Shendure, C. Trapnell, Cicero predicts cis-regulatory DNA interactions from single-cell chromatin accessibility data. *Mol. Cell* **71**, 858–871.e8 (2018).

39. C. Liu, W. Liu, M. F. Lu, N. A. Brown, J. F. Martin, Regulation of left-right asymmetry by thresholds of Pitx2c activity. *Development* **128**, 2039–2048 (2001).

40. M. R. Rivaud, M. Delmar, C. A. Remme, Heritable arrhythmia syndromes associated with abnormal cardiac sodium channel function: Ionic and non-ionic mechanisms. *Cardiovasc. Res.* **116**, 1557–1570 (2020).

41. J. Veevers, E. N. Farah, M. Corselli, A. D. Witty, K. Palomares, J. G. Vidal, N. Emre, C. T. Carson, K. Ouyang, C. Liu, P. van Vliet, M. Zhu, J. M. Hegarty, D. C. Deacon, J. D. Grinstein, R. J. Dirschinger, K. A. Frazer, E. D. Adler, K. U. Knowlton, N. C. Chi, J. C. Martin, J. Chen, S. M. Evans, Cell-surface marker signature for enrichment of ventricular cardiomyocytes derived from human embryonic stem cells. *Stem Cell Rep.* **11**, 828–841 (2018).

42. L. Hortells, A. K. Z. Johansen, K. E. Yutzey, Cardiac fibroblasts and the extracellular matrix in regenerative and nonregenerative hearts. *J. Cardiovasc. Dev. Dis.* **6**, 29 (2019).

43. B. G. Bruneau, Z.-Z. Bao, M. Tanaka, J. J. Schott, S. Izumo, C. L. Cepko, J. G. Seidman, C. E. Seidman, Cardiac expression of the ventricle-specific homeobox gene *Irx4* is modulated by Nkx2-5 and dHand. *Dev. Biol.* **217**, 266–277 (2000).

44. M. R. Corces, A. Shcherbina, S. Kundu, M. J. Gloudemans, L. Frésard, J. M. Granja, B. H. Louie, T. Eulalio, S. Shams, S. T. Bagdatli, M. R. Mumbach, B. Liu, K. S. Montine, W. J. Greenleaf, A. Kundaje, S. B. Montgomery, H. Y. Chang, T. J. Montine, Single-cell epigenomic analyses implicate candidate causal variants at inherited risk loci for Alzheimer's and Parkinson's diseases. *Nat. Genet.* **52**, 1158–1168 (2020).

45. P. W. Hook, A. S. McCallion, Leveraging mouse chromatin data for heritability enrichment informs common disease architecture and reveals cortical layer contributions to schizophrenia. *Genome Res.* **30**, 528–539 (2020).

46. J. Chiou, C. Zeng, Z. Cheng, J. Y. Han, M. Schlichting, M. Miller, R. Mendez, S. Huang, J. Wang, Y. Sui, A. Deogaygay, M.-L. Okino, Y. Qui, Y. Sun, P. Kudtarkar, R. Fang, S. Preissl, M. Sander, D. U. Gorkin, K. J. Gaulton, Single-cell chromatin accessibility identifies pancreatic islet cell type– and state-specific regulatory programs of diabetes risk. *Nat. Genet.* **53**, 455–466 (2021).

47. A. Nott, I. R. Holtman, N. G. Coufal, J. C. M. Schlachetzki, M. Yu, R. Hu, C. Z. Han, M. Pena, J. Xiao, Y. Wu, Z. Keulen, M. P. Pasillas, C. O'Connor, C. K. Nickl, S. T. Schafer, Z. Shen, R. A. Rissman, J. B. Brewer, D. Gosselin, D. D. Gonda, M. L. Levy, M. G. Rosenfeld, G. McVicker, F. H. Gage, B. Ren, C. K. Glass, Brain cell type–specific enhancer–promoter interactome maps and disease-risk association. *Science* **366**, 1134–1139 (2019).

48. B. K. Bulik-Sullivan, P.-R. Loh, H. K. Finucane, S. Ripke, J. Yang; Schizophrenia Working Group of the Psychiatric Genomics Consortium, N. Patterson, M. J. Daly, A. L. Price, B. M. Neale, LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).

49. M. Arvanitis, E. Tampakakis, Y. Zhang, W. Wang, A. Auton; 23andMe Research Team, D. Dutta, S. Glavaris, A. Keramati, N. Chatterjee, N. C. Chi, B. Ren, W. S. Post, A. Battle, Genome-wide association and multi-omic analyses reveal ACTN2 as a gene linked to heart failure. *Nat. Commun.* **11**, 1122 (2020).

50. J. B. Nielsen, R. B. Thorolfsdottir, L. G. Fritsche, W. Zhou, M. W. Skov, S. E. Graham, T. J. Herron, S. McCarthy, E. M. Schmidt, G. Sveinbjornsson, I. Surakka, M. R. Mathis, M. Yamazaki, R. D. Crawford, M. E. Gabrielsen, A. H. Skogholt, O. L. Holmen, M. Lin, B. N. Wolford, R. Dey, H. Dalen, P. Sulem, J. H. Chung, J. D. Backman, D. O. Arnar, U. Thorsteinsdottir, A. Baras, C. O'Dushlaine, A. G. Holst, X. Wen, W. Hornsby, F. E. Dewey, M. Boehnke, S. Kheterpal, B. Mukherjee, S. Lee, H. M. Kang, H. Holm, J. Kitzman, J. A. Shavit, J. Jalife, C. M. Brummett, T. M. Teslovich, D. J. Carey, D. F. Gudbjartsson, K. Stefansson, G. R. Abecasis, K. Hveem, C. J. Willer, Biobank-driven genomic discovery yields new insight into atrial fibrillation biology. *Nat. Genet.* **50**, 1234–1239 (2018).

51. C. P. Nelson, A. Goel, A. S. Butterworth, S. Kanoni, T. R. Webb, E. Marouli, L. Zeng, I. Ntalla, F. Y. Lai, J. C. Hopewell, O. Giannakopoulou, T. Jiang, S. E. Hamby, E. D. Angelantonio, T. L. Assimes, E. P. Bottinger, J. C. Chambers, R. Clarke, C. N. A. Palmer, R. M. Cubbon, P. Ellinor, R. Ermel, E. Evangelou, P. W. Franks, C. Grace, D. Gu, A. D. Hingorani, J. M. M. Howson, E. Ingelsson, A. Kastrati, T. Kessler, T. Kyriakou, T. Lehtimäki, X. Lu, Y. Lu, W. März, R. M. Mcpherson, A. Metspalu, M. Pujades-Rodriguez, A. Ruusalepp, E. E. Schadt, A. F. Schmidt, M. J. Sweeting, P. A. Zalloua, K. A. Ghalayini, B. D. Keavney, J. S. Kooner, R. J. F. Loos, R. S. Patel, M. K. Rutter, M. Tomaszewski, I. Tzoulaki, E. Zeggini, J. Erdmann, G. Dedoussis, J. L. M. Björkegren; EPIC-CVD Consortium; CARDIoGRAMplusC4D; UK Biobank Cardio Metabolic Consortium CHD working group, H. Schunkert, M. Farrall, J. Danesh, N. J. Samani, H. Watkins, P. Deloukas, Association analyses based on false discovery rate implicate new loci for coronary artery disease. *Nat. Genet.* **49**, 1385–1391 (2017).

52. A. S. Shadrina, S. Z. Sharapov, T. I. Shashkova, Y. A. Tsepilov, Varicose veins of lower extremities: Insights from the first large-scale genetic study. *PLOS Genet.* **15**, e1008110 (2019).

53. R. Malik, G. Chauhan, M. Traylor, M. Sargurupremraj, Y. Okada, A. Mishra, L. Rutten-Jacobs, A.-K. Giese, S. W. van der Laan, S. Gretarsdottir, C. D. Anderson, M. Chong, H. H. H. Adams, T. Ago, P. Almgren, P. Amouyel, H. Ay, T. M. Bartz, O. R. Benavente, S. Bevan, G. B. Boncoraglio, R. D. Brown Jr., A. S. Butterworth, C. Carrera, C. L. Carty, D. I. Chasman, W.-M. Chen, J. W. Cole, A. Correa, I. Cotlarciuc, C. Cruchaga, J. Danesh, P. I. W. de Bakker, A. L. De Stefano, M. den Hoed, Q. Duan, S. T. Engelter, G. J. Falcone, R. F. Gottesman, R. P. Grewal, V. Gudnason, S. Gustafsson, J. Haessler, T. B. Harris, A. Hassan, A. S. Havulinna, S. R. Heckbert, E. G. Holliday, G. Howard, F.-C. Hsu, H. I. Hyacinth, M. A. Ikram, E. Ingelsson, M. R. Irvin, X. Jian, J. Jiménez-Conde, J. A. Johnson, J. W. Jukema, M. Kanai, K. L. Keene, B. M. Kissela, D. O. Kleindorfer, C. Kooperberg, M. Kubo, L. A. Lange, C. D. Langefeld, C. Langenberg, L. J. Launer, J.-M. Lee, R. Lemmens, D. Leys, C. M. Lewis, W.-Y. Lin, A. G. Lindgren, E. Lorentzen, P. K. Magnusson, J. Maguire, A. Manichaikul, P. F. McArdle, J. F. Meschia, B. D. Mitchell, T. H. Mosley, M. A. Nalls, T. Ninomiya, M. J. O'Donnell, B. M. Psaty, S. L. Pulit, K. Rannikmäe, A. P. Reiner, K. M. Rexrode, K. Rice, S. S. Rich, P. M. Ridker, N. S. Rost, P. M. Rothwell, J. I. Rotter, T. Rundek, R. L. Sacco, S. Sakaue, M. M. Sale, V. Salomaa, B. R. Sapkota, R. Schmidt, C. O. Schmidt, U. Schminke, P. Sharma, A. Slowik, C. L. M. Sudlow, C. Tanislav, T. Tatlisumak, K. D. Taylor, V. N. S. Thijs, G. Thorleifsson, U. Thorsteinsdottir, S. Tiedt, S. Trompet, C. Tzourio, C. M. van Duijn, M. Walters, N. J. Wareham, S. Wassertheil-Smoller, J. G. Wilson, K. L. Wiggins, Q. Yang, S. Yusuf; AFGen Consortium; Cohorts for Heart; Aging Research in Genomic Epidemiology (CHARGE) Consortium; International Genomics of Blood Pressure (iGEN-BP) Consortium; INVENT Consortium; STARNET, J. C. Bis, T. Pastinen, A. Ruusalepp, E. E. Schadt, S. Koplev, J. L. M. Björkegren, V. Codoni, M. Civelek, N. L. Smith, D. A. Trégouët, I. E. Christophersen, C. Roselli, S. A. Lubitz, P. T. Ellinor, E. S. Tai, J. S. Kooner, N. Kato, J. He, P. van der Harst, P. Elliott, J. C. Chambers, F. Takeuchi, A. D. Johnson; BioBank Japan Cooperative Hospital Group; COMPASS Consortium; EPIC-CVD Consortium; EPIC-Inter ActConsortium; International Stroke Genetics Consortium (ISGC); METASTROKE Consortium; Neurology Working Group of the CHARGE Consortium; NINDS Stroke Genetics Network (SiGN); UK Young Lacunar DNA Study; MEGASTROKE Consortium, D. K. Sanghera, O. Melander, C. Jern, D. Strbian, I. Fernandez-Cadenas, W. T. Longstreth Jr., A. Rolfs, J. Hata, D. Woo, J. Rosand, G. Pare, J. C. Hopewell, D. Saleheen, K. Stefansson, B. B. Worrall, S. J. Kittner, S. Seshadri, M. Fornage, H. S. Markus, J. M. M. Howson, Y. Kamatani, S. Debette, M. Dichgans, Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nat. Genet.* **50**, 524–537 (2018).

54. A. Wang, J. Chiou, O. B. Poirion, J. Buchanan, M. J. Valdez, J. M. Verheyden, X. Hou, P. Kudtarkar, S. Narendra, J. M. Newsome, M. Guo, D. A. Faddah, K. Zhang, R. E. Young, J. Barr, E. Sajti, R. Misra, H. Huyck, L. Rogers, C. Poole, J. A. Whitsett, G. Pryhuber, X. Xu, K. J. Gaulton, S. Preissl, X. Sun; NHLBI LungMap Consortium, Single-cell multiomic profiling of human lungs reveals cell-type-specific and age-dynamic control of SARS-CoV2 host genes. *eLife* **9**, e62522 (2020).

55. J. Wakefield, Bayes factors for genome-wide association studies: Comparison with P-values. *Genet. Epidemiol.* **33**, 79–86 (2009).

56. M. van den Boogaard, S. Smemo, O. Burnicka-Turek, D. E. Arnolds, H. J. G. van der Werken, P. Klous, D. McKean, J. D. Muehlschlegel, J. Moosmann, O. Toka, X. H. Yang, T. T. Koopmann, M. E. Adriaens, C. R. Bezzina, W. de Laat, C. Seidman, J. G. Seidman, V. M. Christoffels, M. A. Nobrega, P. Barnett, I. P. Moskowitz, A common genetic variant within SCN10A modulates cardiac SCN5A expression. *J. Clin. Invest.* **124**, 1844–1852 (2014).

57. Y. Zhang, T. Li, S. Preissl, J. Grinstein, E. N. Farah, E. Destici, A. Y. Lee, S. Chee, Y. Qiu, K. Ma, Z. Ye, R. Hu, Q. Zhu, H. Huang, R. Fang, L. Yu, J. C. I. Belmonte, J. Wu, S. M. Evans, N. C. Chi, B. Ren, 3D chromatin architecture remodeling during human cardiomyocyte differentiation reveals a role of HERV-H in demarcating chromatin domains. *bioRxiv*, 485961 (2019).

58. A. Visel, S. Minovitsky, I. Dubchak, L. A. Pennacchio, VISTA Enhancer Browser--a database of tissue-specific human enhancers. *Nucleic Acids Res.* **35**, D88–D92 (2007).

59. M. E. Curran, I. Splawski, K. W. Timothy, G. M. Vincen, E. D. Green, M. T. Keating, A molecular basis for cardiac arrhythmia: HERG mutations cause long QT syndrome. *Cell* **80**, 795–803 (1995).

60. D. K. Jones, F. Liu, R. Vaidyanathan, L. L. Eckhardt, M. C. Trudeau, G. A. Robertson, hERG 1b is critical for human cardiac repolarization. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 18073–18077 (2014).

61. R. Gilsbach, S. Preissl, B. A. Grüning, T. Schnick, L. Burger, V. Benes, A. Würch, U. Bönisch, S. Günther, R. Backofen, B. K. Fleischmann, D. Schübeler, L. Hein, Dynamic DNA methylation orchestrates cardiomyocyte development, maturation and disease. *Nat. Commun.* **5**, 5288 (2014).

62. Q. Wang, H. Xiong, S. Ai, X. Yu, Y. Liu, J. Zhang, A. He, CoBATCH for high-throughput single-cell epigenomic profiling. *Mol. Cell* **76**, 206–216.e7 (2019).

63. H. S. Kaya-Okur, S. J. Wu, C. A. Codomo, E. S. Pledger, T. D. Bryson, J. G. Henikoff, K. Ahmad, S. Henikoff, CUT&Tag for efficient epigenomic profiling of small samples and single cells. *Nat. Commun.* **10**, 1930 (2019).

64. C. Sudlow, J. Gallacher, N. Allen, V. Beral, P. Burton, J. Danesh, P. Downey, P. Elliott, J. Green, M. Landray, B. Liu, P. Matthews, G. Ong, J. Pell, A. Silman, A. Young, T. Sprosen, T. Peakman, R. Collins, UK biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **12**, e1001779 (2015).

65. K. Ishigaki, M. Akiyama, M. Kanai, A. Takahashi, E. Kawakami, H. Sugishita, S. Sakaue, N. Matoba, S. K. Low, Y. Okada, C. Terao, T. Amariuta, S. Gazal, Y. Kochi, M. Horikoshi, K. Suzuki, K. Ito, S. Koyama, K. Ozaki, S. Niida, Y. Sakata, Y. Sakata, T. Kohno, K. Shiraishi, Y. Momozawa, M. Hirata, K. Matsuda, M. Ikeda, N. Iwata, S. Ikegawa, I. Kou, T. Tanaka, H. Nakagawa, A. Suzuki, T. Hirota, M. Tamari, K. Chayama, D. Miki, M. Mori, S. Nagayama, Y. Daigo, Y. Miki, T. Katagiri, O. Ogawa, W. Obara, H. Ito, T. Yoshida, I. Imoto, T. Takahashi, C. Tanikawa, T. Suzuki, N. Sinozaki, S. Minami, H. Yamaguchi, S. Asai, Y. Takahashi, K. Yamaji, K. Takahashi, T. Fujioka, R. Takata, H. Yanai, A. Masumoto, Y. Koretsune, H. Kutsumi, M. Higashiyama, S. Murayama, N. Minegishi, K. Suzuki, K. Tanno, A. Shimizu, T. Yamaji, M. Iwasaki, N. Sawada, H. Uemura, K. Tanaka, M. Naito, M. Sasaki, K. Wakai, S. Tsugane, M. Yamamoto, K. Yamamoto, Y. Murakami, Y. Nakamura, S. Raychaudhuri, J. Inazawa, T. Yamauchi, T. Kadowaki, M. Kubo, Y. Kamatani, Large-scale genome-wide association study in a Japanese population identifies novel susceptibility loci across different diseases. *Nat. Genet.* **52**, 669–679 (2020).

66. NHLBI, Trans-Omics for Precision Medicine (TOPMed) Program (2014).

67. J. W. Smyth, T.-T. Hong, D. Gao, J. M. Vogan, B. C. Jensen, T. S. Fong, P. C. Simpson, D. Y.-R. Stainier, N. C. Chi, R. M. Shaw, Limited forward trafficking of connexin 43 reduces cell-cell coupling in stressed human and mouse myocardium. *J. Clin. Invest.* **120**, 266–279 (2010).

68. G. Chen, D. R. Gulbranson, Z. Hou, J. M. Bolin, V. Ruotti, M. D. Probasco, K. Smuga-Otto, S. E. Howden, N. R. Diol, N. E. Propson, R. Wagner, G. O. Lee, J. Antosiewicz-Bourget, J. M. C. Teng, J. A. Thomson, Chemically defined conditions for human iPSC derivation and culture. *Nat. Methods* **8**, 424–429 (2011).

69. K. Watanabe, M. Ueno, D. Kamiya, A. Nishiyama, M. Matsumura, T. Wataya, J. B. Takahashi, S. Nishikawa, S. I. Nishikawa, K. Muguruma, Y. Sasai, A ROCK inhibitor permits survival of dissociated human embryonic stem cells. *Nat. Biotechnol.* **25**, 681–686 (2007).

70. X. Lian, J. Zhang, S. M. Azarin, K. Zhu, L. B. Hazeltine, X. Bao, C. Hsiao, T. J. Kamp, S. P. Palecek, Directed cardiomyocyte differentiation from human pluripotent stem cells by modulating Wnt/β-catenin signaling under fully defined conditions. *Nat. Protoc.* **8**, 162–175 (2013).

71. A. Teumer, A. Tin, R. Sorice, M. Gorski, N. C. Yeo, A. Y. Chu, M. Li, Y. Li, V. Mijatovic, Y.-A. Ko, D. Taliun, A. Luciani, M.-H. Chen, Q. Yang, M. C. Foster, M. Olden, L. T. Hiraki, B. O. Tayo, C. Fuchsberger, A. K. Dieffenbach, A. R. Shuldiner, A. V. Smith, A. M. Zappa, A. Lupo, B. Kollerits, B. Ponte, B. Stengel, B. K. Krämer, B. Paulweber, B. D. Mitchell, C. Hayward, C. Helmer, C. Meisinger, C. Gieger, C. M. Shaffer, C. Müller, C. Langenberg, D. Ackermann, D. Siscovick; DCCT/EDIC, E. Boerwinkle, F. Kronenberg, G. B. Ehret, G. Homuth, G. Waeber, G. Navis, G. Gambaro, G. Malerba, G. Eiriksdottir, G. Li, H. E. Wichmann, H. Grallert, H. Wallaschofski, H. Völzke, H. Brenner, H. Kramer, I. M. Leach, I. Rudan, H. L. Hillege, J. S. Beckmann, J. C. Lambert, J. Luan, J. H. Zhao, J. Chalmers, J. Coresh, J. C. Denny, K. Butterbach, L. J. Launer, L. Ferrucci, L. Kedenko, M. Haun, M. Metzger, M. Woodward, M. J. Hoffman, M. Nauck, M. Waldenberger, M. Pruijm, M. Bochud, M. Rheinberger, N. Verweij, N. J. Wareham, N. Endlich, N. Soranzo, O. Polasek, P. van der Harst, P. P. Pramstaller, P. Vollenweider, P. S. Wild, R. T. Gansevoort, R. Rettig, R. Biffar, R. J. Carroll, R. Katz, R. J. F. Loos, S.-J. Hwang, S. Coassin, S. Bergmann, S. E. Rosas, S. Stracke, T. B. Harris, T. Corre, T. Zeller, T. Illig, T. Aspelund, T. Tanaka, U. Lendeckel, U. Völker, V. Gudnason, V. Chouraki, W. Koenig, Z. Kutalik, J. R. O'Connell, A. Parsa, I. M. Heid, A. D. Paterson, I. H. de Boer, O. Devuyst, J. Lazar, K. Endlich, K. Susztak,

J. Tremblay, P. Hamet, H. J. Jacob, C. A. Böger, C. S. Fox, C. Pattaro, A. Köttgen, Genome-wide association studies identify genetic loci associated with albuminuria in diabetes. *Diabetes* **65**, 803–817 (2016).

72. M. Haeussler, K. Schönig, H. Eckert, A. Eschstruth, J. Mianné, J. B. Renaud, S. Schneider-Maunoury, A. Shkumatava, L. Teboul, J. Kent, J. S. Joly, J. P. Concordet, Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biol.* **17**, 148 (2016).

73. S. Kim, D. Kim, S. W. Cho, J. Kim, J. S. Kim, Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res.* **24**, 1012–1019 (2014).

74. J. A. Zuris, D. B. Thompson, Y. Shu, J. P. Guilinger, J. L. Bessen, J. H. Hu, M. L. Maeder, J. K. Joung, Z. Y. Chen, D. R. Liu, Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. *Nat. Biotechnol.* **33**, 73–80 (2015).

75. D. P. Santos, E. Kiskinis, K. Eggan, F. T. Merkle, Comprehensive protocols for CRISPR/Cas9-based gene editing in human pluripotent stem cells. *Curr. Protoc. Stem Cell Biol.* **38**, 5b.6.1–5b.6.60 (2016).

76. J. Harrow, F. Denoeud, A. Frankish, A. Reymond, C.-K. Chen, J. Chrast, J. Lagarde, J. G. R. Gilbert, R. Storey, D. Swarbreck, C. Rossier, C. Ucla, T. Hubbard, S. E. Antonarakis, R. Guigo, GENCODE: Producing a reference annotation for ENCODE. *Genome Biol.* **7** Suppl 1, S4–S54 (2006).

77. S. L. Wolock, R. Lopez, A. M. Klein, Scrublet: Computational identification of cell doublets in single-cell transcriptomic data. *Cell Syst.* **8**, 281–291.e9 (2019).

78. G. V. Fabian Pedregosa, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, É. Duchesnay, Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **12** (85), 2825–2830 (2011).

79. D. B. Bouneffouf, I, Theoretical analysis of the Minimum Sum of Squared Similarities sampling for Nyström-based spectral clustering, in *2016 International Joint Conference on Neural Networks (IJCNN)* (2016), pp. 3856–3862.

80. L. Haghverdi, A. T. L. Lun, M. D. Morgan, J. C. Marioni, Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nat. Biotechnol.* **36**, 421–427 (2018).

81. V. A. Traag, L. Waltman, N. J. van Eck, From Louvain to Leiden: Guaranteeing well-connected communities. *Sci. Rep.* **9**, 5233 (2019).

82. V. A. Traag, P. Van Dooren, Y. Nesterov, Narrow scope for resolution-limit-free community detection. *Phys. Rev. E Stat. Nonlinear Soft Matter Phys.* **84**, 016114 (2011).

83. C. S. McGinnis, L. M. Murrow, Z. J. Gartner, DoubletFinder: Doublet detection in single-cell RNA sequencing data using artificial nearest neighbors. *Cell Syst.* **8**, 329–337.e4 (2019).

84. I. Korsunsky, N. Millard, J. Fan, K. Slowikowski, F. Zhang, K. Wei, Y. Baglaenko, M. Brenner, P. R. Loh, S. Raychaudhuri, Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).

85. H. M. Amemiya, A. Kundaje, A. P. Boyle, The ENCODE Blacklist: Identification of problematic regions of the genome. *Sci. Rep.* **9**, 9354 (2019).

86. A. R. Quinlan, I. M. Hall, BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).

87. D. L. Davies, D. W. Bouldin, A cluster separation measure. *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-1**, 224–227 (1979).

88. F. Sheikh, R. C. Lyon, J. Chen, Functions of myosin light chain-2 (MYL2) in cardiac muscle and disease. *Gene* **569**, 14–20 (2015).

89. S. W. Kubalak, W. C. Miller-Hance, T. X. O'Brien, E. Dyson, K. R. Chien, Chamber specification of atrial myosin light chain-2 expression precedes septation during murine cardiogenesis. *J. Biol. Chem.* **269**, 16961–16970 (1994).

90. M. B. Furtado, M. W. Costa, E. A. Pranoto, E. Salimova, A. R. Pinto, N. T. Lam, A. Park, P. Snider, A. Chandran, R. P. Harvey, R. Boyd, S. J. Conway, J. Pearson, D. M. Kaye, N. A. Rosenthal, Cardiogenic genes expressed in cardiac fibroblasts contribute to heart development and repair. *Circ. Res.* **114**, 1422–1434 (2014).

91. D. Nichol, H. Stuhlmann, EGFL7: A unique angiogenic signaling factor in vascular development and disease. *Blood* **119**, 1345–1352 (2012).

92. J. Kalucka, L. P. M. H. de Rooij, J. Goveia, K. Rohlenova, S. J. Dumas, E. Meta, N. V. Conchinha, F. Taverna, L.-A. Teuwen, V. Veys, M. García-Caballero, S. Khan, V. Geldhof, L. Sokol, R. Chen, L. Treps, M. Borri, P. Zeeuw, C. Dubois, T. K. Karakach, K. D. Falkenberg, M. Parys, X. Yin, S. Vinckier, Y. Du, R. A. Fenton, L. Schoonjans, M. Dewerchin, G. Eelen, B. Thienpont, L. Lin, L. Bolund, X. Li, Y. Luo, P. Carmeliet, Single-cell transcriptome atlas of murine endothelial cells. *Cell* **180**, 764–779.e20 (2020).

93. C. M. Shanahan, P. L. Weissberg, J. C. Metcalfe, Isolation of gene markers of differentiated and proliferating vascular smooth muscle cells. *Circ. Res.* **73**, 193–204 (1993).

94. The Tabula Muris Consortium; Overall coordination, N. Schaum, J. Karkanias, N. F. Neff, A. P. May, S. R. Quake, T. Wyss-Coray, S. Darmanis, J. Batson, O. Botvinnik, M. B. Chen, S. Chen, F. Green, R. C. Jones, A. Maynard, L. Penland, A. O. Pisco, R. V. Sit, G. M. Stanley, J. T. Webber, F. Zanini, A. S. Baghel, I. Bakerman, I. Bansal, D. Berdnik, B. Bilen,

D. Brownfield, C. Cain, M. B. Chen, S. Chen, M. Cho, G. Cirolia, S. D. Conley, S. Darmanis, A. Demers, K. Demir, A. de Morree, T. Divita, H. du Bois, L. B. T. Dulgeroff, H. Ebadi, F. H. Espinoza, M. Fish, Q. Gan, B. M. George, A. Gillich, F. Green, G. Genetiano, X. Gu, G. S. Gulati, Y. Hang, S. Hosseinzadeh, A. Huang, T. Iram, T. Isobe, F. Ives, R. C. Jones, K. S. Kao, G. Karnam, A. M. Kershner, B. M. Kiss, W. Kong, M. E. Kumar, J. Y. Lam, D. P. Lee, S. E. Lee, G. Li, Q. Li, L. Liu, A. Lo, W.-J. Lu, A. Manjunath, A. P. May, K. L. May, O. L. May, A. Maynard, M. McKay, R. J. Metzger, M. Mignardi, D. Min, A. N. Nabhan, N. F. Neff, K. M. Ng, J. Noh, R. Patkar, W. C. Peng, L. Penland, R. Puccinelli, E. J. Rulifson, N. Schaum, S. S. Sikandar, R. Sinha, R. V. Sit, K. Szade, W. Tan, C. Tato, K. Tellez, K. J. Travaglini, C. Tropini, L. Waldburger, L. J. van Weele, M. N. Wosczyna, J. Xiang, S. Xue, J. Youngyunpipatkul, F. Zanini, M. E. Zardeneta, F. Zhang, L. Zhou, I. Bansal, S. Chen, M. Cho, G. Cirolia, S. Darmanis, A. Demers, T. Divita, H. Ebadi, G. Genetiano, F. Green, S. Hosseinzadeh, F. Ives, A. Lo, A. P. May, A. Maynard, M. McKay, N. F. Neff, L. Penland, R. V. Sit, W. Tan, L. Waldburger, J. Youngyunpipatkul, J. Batson, O. Botvinnik, P. Castro, D. Croote, S. Darmanis, J. L. De Risi, J. Karkanias, A. O. Pisco, G. M. Stanley, J. T. Webber, F. Zanini, A. S. Baghel, I. Bakerman, J. Batson, B. Bilen, O. Botvinnik, D. Brownfield, M. B. Chen, S. Darmanis, K. Demir, A. de Morree, H. Ebadi, F. H. Espinoza, M. Fish, Q. Gan, B. M. George, A. Gillich, X. Gu, G. S. Gulati, Y. Hang, A. Huang, T. Iram, T. Isobe, G. Karnam, A. M. Kershner, B. M. Kiss, W. Kong, C. S. Kuo, J. Y. Lam, B. Lehallier, G. Li, Q. Li, L. Liu, W.-J. Lu, D. Min, A. N. Nabhan, K. M. Ng, P. K. Nguyen, R. Patkar, W. C. Peng, L. Penland, E. J. Rulifson, N. Schaum, S. S. Sikandar, R. Sinha, K. Szade, S. Y. Tan, K. Tellez, K. J. Travaglini, C. Tropini, L. J. van Weele, B. M. Wang, M. N. Wosczyna, J. Xiang, H. Yousef, L. Zhou, J. Batson, O. Botvinnik, S. Chen, S. Darmanis, F. Green, A. P. May, A. Maynard, A. O. Pisco, S. R. Quake, N. Schaum, G. M. Stanley, J. T. Webber, T. Wyss-Coray, F. Zanini, P. A. Beachy, C. K. F. Chan, A. de Morree, B. M. George, G. S. Gulati, Y. Hang, K. C. Huang, T. Iram, T. Isobe, A. M. Kershner, B. M. Kiss, W. Kong, G. Li, Q. Li, L. Liu, W.-J. Lu, A. N. Nabhan, K. M. Ng, P. K. Nguyen, W. C. Peng, E. J. Rulifson, N. Schaum, S. S. Sikandar, R. Sinha, K. Szade, K. J. Travaglini, C. Tropini, B. M. Wang, K. Weinberg, M. N. Wosczyna, S. M. Wu, H. Yousef, B. A. Barres, P. A. Beachy, C. K. F. Chan, M. F. Clarke, S. Darmanis, K. C. Huang, J. Karkanias, S. K. Kim, M. A. Krasnow, M. E. Kumar, C. S. Kuo, A. P. May, R. J. Metzger, N. F. Neff, R. Nusse, P. K. Nguyen, T. A. Rando, J. Sonnenburg, B. M. Wang, K. Weinberg, I. L. Weissman, S. M. Wu, S. R. Quake, T. Wyss-Coray, Single-cell transcriptomics of 20 mouse organs creates a *Tabula Muris*. *Nature* **562**, 367–372 (2018).

95. B. O. Fabriek, C. D. Dijkstra, T. K. van den Berg, The macrophage scavenger receptor CD163. *Immunobiology* **210**, 153–160 (2005).

96. F. O. Martinez, S. Gordon, M. Locati, A. Mantovani, Transcriptional profiling of the human monocyte-to-macrophage differentiation and polarization: New molecules and patterns of gene expression. *J. Immunol.* **177**, 7303–7311 (2006).

97. J. J. Peschon, P. J. Morrissey, K. H. Grabstein, F. J. Ramsdell, E. Maraskovsky, B. C. Gliniak, L. S. Park, S. F. Ziegler, D. E. Williams, C. B. Ware, J. D. Meyer, B. L. Davison, Early lymphocyte expansion is severely impaired in interleukin 7 receptor-deficient mice. *J. Exp. Med.* **180**, 1955–1960 (1994).

98. R. Lesourne, S. Uehara, J. Lee, K.-D. Song, L.-Q. Li, J. Pinkhasov, Y. Zhang, N.-P. Weng, K. F. Wildt, L. Wang, R. Bosselut, P. E. Love, Themis, a T cell–specific protein important for late thymocyte development. *Nat. Immunol.* **10**, 840–847 (2009).

99. V. Puri, S. Ranjit, S. Konda, S. M. C. Nicoloro, J. Straubhaar, A. Chawla, M. Chouinard, C. Lin, A. Burkart, S. Corvera, R. A. Perugini, M. P. Czech, Cidea is associated with lipid droplets and insulin sensitivity in humans. *Proc. Natl. Acad. Sci.* **105**, 7833–7838 (2008).

100. E. Hu, P. Liang, B. M. Spiegelman, AdipoQ is a novel adipose-specific gene dysregulated in obesity. *J. Biol. Chem.* **271**, 10697–10703 (1996).

101. D. A. Skelly, G. T. Squiers, M. A. McLellan, M. T. Bolisetty, P. Robson, N. A. Rosenthal, A. R. Pinto, Single-cell transcriptional profiling reveals cellular diversity and intercommunication in the mouse heart. *Cell Rep.* **22**, 600–610 (2018).

102. Y. Y. Zhao, D. R. Sawyer, R. R. Baliga, D. J. Opel, X. Han, M. A. Marchionni, R. A. Kelly, Neuregulins promote survival and growth of cardiac myocytes: Persistence of ErbB2 and ErbB4 expression in neonatal and adult ventricular myocytes. *J. Biol. Chem.* **273**, 10261–10269 (1998).

103. J. Tang, H. Zhang, L. He, X. Huang, Y. Li, W. Pu, W. Yu, L. Zhang, D. Cai, K. O. Lui, B. Zhou, Genetic fate mapping defines the vascular potential of endocardial cells in the adult heart. *Circ. Res.* **122**, 984–993 (2018).

104. B. D. Southern, L. M. Grove, S. O. Rahaman, S. Abraham, R. G. Scheraga, K. A. Niese, H. Sun, E. L. Herzog, F. Liu, D. J. Tschumperlin, T. T. Egelhoff, S. S. Rosenfeld, M. A. Olman, Matrix-driven Myosin II mediates the pro-fibrotic fibroblast phenotype. *J. Biol. Chem.* **291**, 6083–6095 (2016).

105. J. Baum, H. S. Duffy, Fibroblasts and myofibroblasts: What are we talking about? *J. Cardiovasc. Pharmacol.* **57**, 376–379 (2011).

106. M. Vanlandewijck, L. He, M. A. Mäe, J. Andrae, K. Ando, F. del Gaudio, K. Nahar, T. Lebouvier, B. Laviña, L. Gouveia, Y. Sun, E. Raschperger, M. Räsänen, Y. Zarb, N. Mochizuki, A. Keller, U. Lendahl, C. Betsholtz, A molecular atlas of cell types and zonation in the brain vasculature. *Nature* **554**, 475–480 (2018).

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

108. L. McInnes, J. Healy, N. Saul, L. Großberger, UMAP: Uniform manifold approximation and projection. *J. Open Source Softw.* **3** (29), 861 (2018).

109. A. Mathelier, O. Fornes, D. J. Arenillas, C.-Y. Chen, G. Denay, J. Lee, W. Shi, C. Shyr, G. Tan, R. Worsley-Hunt, A. W. Zhang, F. Parcy, B. Lenhard, A. Sandelin, W. W. Wasserman, JASPAR 2016: A major expansion and update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* **44**, D110–D115 (2016).

110. M. J. L. de Hoon, S. Imoto, J. Nolan, S. Miyano, Open source clustering software. *Bioinformatics* **20**, 1453–1454 (2004).

111. H. K. Finucane, B. Bulik-Sullivan, A. Gusev, G. Trynka, Y. Reshef, P.-R. Loh, V. Anttila, H. Xu, C. Zang, K. Farh, S. Ripke, F. R. Day; ReproGen Consortium; Schizophrenia Working Group of the Psychiatric Genomics Consortium; RACI Consortium, S. Purcell, E. Stahl, S. Lindstrom, J. R. B. Perry, Y. Okada, S. Raychaudhuri, M. J. Daly, N. Patterson, B. M. Neale, A. L. Price, Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat. Genet.* **47**, 1228–1235 (2015).

112. S. Ripke, B. M. Neale, A. Corvin, J. T. R. Walters, K.-H. Farh, P. A. Holmans, P. Lee, B. Bulik-Sullivan, D. A. Collier, H. Huang, T. H. Pers, I. Agartz, E. Agerbo, M. Albus, M. Alexander, F. Amin, S. A. Bacanu, M. Begemann, R. A. Belliveau Jr., J. Bene, S. E. Bergen, E. Bevilacqua, T. B. Bigdeli, D. W. Black, R. Bruggeman, N. G. Buccola, R. L. Buckner, W. Byerley, W. Cahn, G. Cai, D. Campion, R. M. Cantor, V. J. Carr, N. Carrera, S. V. Catts, K. D. Chambert, R. C. K. Chan, R. Y. L. Chen, E. Y. H. Chen, W. Cheng, E. F. C. Cheung, S. A. Chong, C. R. Cloninger, D. Cohen, N. Cohen, P. Cormican, N. Craddock, J. J. Crowley, D. Curtis, M. Davidson, K. L. Davis, F. Degenhardt, J. D. Favero, D. Demontis, D. Dikeos, T. Dinan, S. Djurovic, G. Donohoe, E. Drapeau, J. Duan, F. Dudbridge, N. Durmishi, P. Eichhammer, J. Eriksson, V. Escott-Price, L. Essioux, A. H. Fanous, M. S. Farrell, J. Frank, L. Franke, R. Freedman, N. B. Freimer, M. Friedl, J. I. Friedman, M. Fromer, G. Genovese, L. Georgieva, I. Giegling, P. Giusti-Rodríguez, S. Godard, J. I. Goldstein, V. Golimbet, S. Gopal, J. Gratten, L. de Haan, C. Hammer, M. L. Hamshere, M. Hansen, T. Hansen, V. Haroutunian, A. M. Hartmann, F. A. Henskens, S. Herms, J. N. Hirschhorn, P. Hoffmann, A. Hofman, M. V. Hollegaard, D. M. Hougaard, M. Ikeda, I. Joa, A. Julià, R. S. Kahn, L. Kalaydjieva, S. Karachanak-Yankova, J. Karjalainen, D. Kavanagh, M. C. Keller, J. L. Kennedy, A. Khrunin, Y. Kim, J. Klovins, J. A. Knowles, B. Konte, V. Kucinskas, Z. A. Kucinskiene, H. Kuzelova-Ptackova, A. K. Kähler, C. Laurent, J. L. C. Keong, S. H. Lee, S. E. Legge, B. Lerer, M. Li, T. Li, K.-Y. Liang, J. Lieberman, S. Limborska, C. M. Loughland, J. Lubinski, J. Lönnqvist, M. Macek Jr., P. K. E. Magnusson, B. S. Maher, W. Maier, J. Mallet, S. Marsal, M. Mattheisen, M. Mattingsdal, R. W. McCarley, C. M. Donald, A. M. McIntosh, S. Meier, C. J. Meijer, B. Melegh, I. Melle, R. I. Mesholam-Gately, A. Metspalu, P. T. Michie, L. Milani, V. Milanova, Y. Mokrab, D. W. Morris, O. Mors, K. C. Murphy, R. M. Murray, I. Myin-Germeys, B. Müller-Myhsok, M. Nelis, I. Nenadic, D. A. Nertney, G. Nestadt, K. K. Nicodemus, L. Nikitina-Zake, L. Nisenbaum, A. Nordin, E. O'Callaghan, C. O'Dushlaine, F. A. O'Neill, S.-Y. Oh, A. Olincy, L. Olsen, J. Van Os; Psychosis Endophenotypes International Consortium, C. Pantelis, G. N. Papadimitriou, S. Papiol, E. Parkhomenko, M. T. Pato, T. Paunio, M. Pejovic-Milovancevic, D. O. Perkins, O. Pietiläinen, J. Pimm, A. J. Pocklington, J. Powell, A. Price, A. E. Pulver, S. M. Purcell, D. Quested, H. B. Rasmussen, A. Reichenberg, M. A. Reimers, A. L. Richards, J. L. Roffman, P. Roussos, D. M. Ruderfer, V. Salomaa, A. R. Sanders, U. Schall, C. R. Schubert, T. G. Schulze, S. G. Schwab, E. M. Scolnick, R. J. Scott, L. J. Seidman, J. Shi, E. Sigurdsson, T. Silagadze, J. M. Silverman, K. Sim, P. Slominsky, J. W. Smoller, H.-C. So, C. C. A. Spencer, E. A. Stahl, H. Stefansson, S. Steinberg, E. Stogmann, R. E. Straub, E. Strengman, J. Strohmaier, T. S. Stroup, M. Subramaniam, J. Suvisaari, D. M. Svrakic, J. P. Szatkiewicz, E. Söderman, S. Thirumalai, D. Toncheva, S. Tosato, J. Veijola, J. Waddington, D. Walsh, D. Wang, Q. Wang, B. T. Webb, M. Weiser, D. B. Wildenauer, N. M. Williams, S. Williams, S. H. Witt, A. R. Wolen, E. H. M. Wong, B. K. Wormley, H. S. Xi, C. C. Zai, X. Zheng, P. Zimprich, N. R. Wray, K. Stefansson, P. M. Visscher; Wellcome Trust Case-Control Consortium, R. Adolfsson, O. A. Andreassen, D. H. R. Blackwood, E. Bramon, J. D. Buxbaum, A. D. Børglum, S. Cichon, A. Darvasi, E. Domenici, H. Ehrenreich, T. Esko, P. V. Gejman, M. Gill, H. Gurling, C. M. Hultman, N. Iwata, A. V. Jablensky, E. G. Jönsson, K. S. Kendler, G. Kirov, J. Knight, T. Lencz, D. F. Levinson, Q. S. Li, J. Liu, A. K. Malhotra, S. A. McCarroll, A. M. Quillin, J. L. Moran, P. B. Mortensen, B. J. Mowry, M. M. Nöthen, R. A. Ophoff, M. J. Owen, A. Palotie, C. N. Pato, T. L. Petryshen, P. Posthuma, M. Rietschel, B. P. Riley, D. Rujescu, P. C. Sham, P. Sklar, D. S. Clair, D. R. Weinberger, J. R. Wendland, T. Werge, M. J. Daly, P. F. Sullivan, M. C. O'Donovan, Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421–427 (2014).

113. A. Wiberg, M. Ng, A. B. Schmid, R. W. Smillie, G. Baskozos, M. V. Holmes, K. Künnapuu, R. Mägi, D. L. Bennett, D. Furniss, A genome-wide association analysis identifies 16 novel susceptibility loci for carpal tunnel syndrome. *Nat. Commun.* **10**, 1030 (2019).

114. J. Grove, S. Ripke, T. D. Als, M. Mattheisen, R. K. Walters, H. Won, J. Pallesen, E. Agerbo, O. A. Andreassen, R. Anney, S. Awashti, R. Belliveau, F. Bettella, J. D. Buxbaum, J. Bybjerg-Grauholm, M. Bækvad-Hansen, F. Cerrato, K. Chambert, J. H. Christensen,

C. Churchhouse, K. Dellenvall, D. Demontis, S. De Rubeis, B. Devlin, S. Djurovic, A. L. Dumont, J. I. Goldstein, C. S. Hansen, M. E. Hauberg, M. V. Hollegaard, S. Hope, D. P. Howrigan, H. Huang, C. M. Hultman, L. Klei, J. Maller, J. Martin, A. R. Martin, J. L. Moran, M. Nyegaard, T. Nærland, D. S. Palmer, A. Palotie, C. B. Pedersen, M. G. Pedersen, T. dPoterba, J. B. Poulsen, B. S. Pourcain, P. Qvist, K. Rehnström, A. Reichenberg, J. Reichert, E. B. Robinson, K. Roeder, P. Roussos, E. Saemundsen, S. Sandin, F. K. Satterstrom, G. D. Smith, H. Stefansson, S. Steinberg, C. R. Stevens, P. F. Sullivan, P. Turley, G. B. Walters, X. Xu; Autism Spectrum Disorder Working Group of the Psychiatric Genomics Consortium; BUPGEN, Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium; 23andMe Research Team, K. Stefansson, D. H. Geschwind, M. Nordentoft, D. M. Hougaard, T. Werge, O. Mors, P. B. Mortensen, B. M. Neale, M. J. Daly, A. D. Børglum, Identification of common genetic risk variants for autism spectrum disorder. *Nat. Genet.* **51**, 431–444 (2019).

115. I. Tachmazidou, K. Hatzikotoulas, L. Southam, J. Esparza-Gordillo, V. Haberland, J. Zheng, T. Johnson, M. Koprulu, E. Zengini, J. Steinberg, J. M. Wilkinson, S. Bhatnagar, J. D. Hoffman, N. Buchan, D. Süveges; arcOGEN Consortium, L. Yerges-Armstrong, G. D. Smith, T. R. Gaunt, R. A. Scott, L. C. McCarthy, E. Zeggini, Identification of new therapeutic targets for osteoarthritis through genome-wide analyses of UK Biobank data. *Nat. Genet.* **51**, 230–236 (2019).

116. K. M. de Lange, L. Moutsianas, J. C. Lee, C. A. Lamb, Y. Luo, N. A. Kennedy, L. Jostins, D. L. Rice, J. Gutierrez-Achury, S.-G. Ji, G. Heap, E. R. Nimmo, C. Edwards, P. Henderson, C. Mowat, J. Sanderson, J. Satsangi, A. Simmons, D. C. Wilson, M. Tremelling, A. Hart, C. G. Mathew, W. G. Newman, M. Parkes, C. W. Lees, H. Uhlig, C. Hawkey, N. J. Prescott, T. Ahmad, J. C. Mansfield, C. A. Anderson, J. C. Barrett, Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nat. Genet.* **49**, 256–261 (2017).

117. N. R. Wray, S. Ripke, M. Mattheisen, M. Trzaskowski, E. M. Byrne, A. Abdellaoui, M. J. Adams, E. Agerbo, T. M. Air, T. M. F. Andlauer, S.-A. Bacanu, M. Bækvad-Hansen, A. F. T. Beekman, T. B. Bigdeli, E. B. Binder, D. R. H. Blackwood, J. Bryois, H. N. Buttenschøn, J. Bybjerg-Grauholm, N. Cai, E. Castelao, J. H. Christensen, T.-K. Clarke, J. I. R. Coleman, L. Colodro-Conde, B. Couvy-Duchesne, N. Craddock, G. E. Crawford, C. A. Crowley, H. S. Dashti, G. Davies, I. J. Deary, F. Degenhardt, E. M. Derks, N. Direk, C. V. Dolan, E. C. Dunn, T. C. Eley, N. Eriksson, V. Escott-Price, F. H. F. Kiadeh, H. K. Finucane, A. J. Forstner, J. Frank, H. A. Gaspar, M. Gill, P. Giusti-Rodríguez, F. S. Goes, S. D. Gordon, J. Grove, L. S. Hall, E. Hannon, C. S. Hansen, T. F. Hansen, S. Herms, I. B. Hickie, P. Hoffmann, G. Homuth, C. Horn, J.-J. Hottenga, D. M. Hougaard, M. Hu, C. L. Hyde, M. Ising, R. Jansen, F. Jin, E. Jorgenson, J. A. Knowles, I. S. Kohane, J. Kraft, W. W. Kretzschmar, J. Krogh, Z. Kutalik, J. M. Lane, Y. Li, Y. Li, P. A. Lind, X. Liu, L. Lu, D. J. Mac Intyre, D. F. Mac Kinnon, R. M. Maier, W. Maier, J. Marchini, H. Mbarek, P. M. Grath, P. M. Guffin, S. E. Medland, D. Mehta, C. M. Middeldorp, E. Mihailov, Y. Milaneschi, L. Milani, J. Mill, F. M. Mondimore, S. Montgomery, S. Mostafavi, N. Mullins, M. Nauck, B. Ng, M. G. Nivard, D. R. Nyholt, P. F. O'Reilly, H. Oskarsson, M. J. Owen, J. N. Painter, C. B. Pedersen, M. G. Pedersen, R. E. Peterson, E. Pettersson, W. J. Peyrot, G. Pistis, D. Posthuma, S. M. Purcell, J. A. Quiroz, P. Qvist, J. P. Rice, B. P. Riley, M. Rivera, S. S. Mirza, R. Saxena, R. Schoevers, E. C. Schulte, L. Shen, J. Shi, S. I. Shyn, E. Sigurdsson, G. B. C. Sinnamon, J. H. Smit, D. J. Smith, H. Stefansson, S. Steinberg, C. A. Stockmeier, F. Streit, J. Strohmaier, K. E. Tansey, H. Teismann, A. Teumer, W. Thompson, P. A. Thomson, T. E. Thorgeirsson, C. Tian, M. Traylor, J. Treutlein, V. Trubetskoy, A. G. Uitterlinden, D. Umbricht, S. Van der Auwera, A. M. van Hemert, A. Viktorin, P. M. Visscher, Y. Wang, B. T. Webb, S. M. Weinsheimer, J. Wellmann, G. Willemsen, S. H. Witt, Y. Wu, H. S. Xi, J. Yang, F. Zhang; eQTLGen; 23andMe, V. Arolt, B. T. Baune, K. Berger, D. I. Boomsma, S. Cichon, U. Dannlowski, E. C. J. de Geus, J. R. De Paulo, E. Domenici, K. Domschke, T. Esko, H. J. Grabe, S. P. Hamilton, C. Hayward, A. C. Heath, D. A. Hinds, K. S. Kendler, S. Kloiber, G. Lewis, Q. S. Li, S. Lucae, P. F. A. Madden, P. K. Magnusson, N. G. Martin, A. M. McIntosh, A. Metspalu, O. Mors, P. B. Mortensen, B. Müller-Myhsok, M. Nordentoft, M. M. Nöthen, M. C. O'Donovan, S. A. Paciga, N. L. Pedersen, B. W. J. H. Penninx, R. H. Perlis, D. J. Porteous, J. B. Potash, M. Preisig, M. Rietschel, C. Schaefer, T. G. Schulze, J. W. Smoller, K. Stefansson, H. Tiemeier, R. Uher, H. Völzke, M. M. Weissman, T. Werge, A. R. Winslow, C. M. Lewis, D. F. Levinson, G. Breen, A. D. Børglum, P. F. Sullivan; Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium, Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat. Genet.* **50**, 668–681 (2018).

118. A. Mahajan, D. Taliun, M. Thurner, N. R. Robertson, J. M. Torres, N. W. Rayner, A. J. Payne, V. Steinthorsdottir, R. A. Scott, N. Grarup, J. P. Cook, E. M. Schmidt, M. Wuttke, C. Sarnowski, R. Mägi, J. Nano, C. Gieger, S. Trompet, C. Lecoeur, M. H. Preuss, B. P. Prins, X. Guo, L. F. Bielak, J. E. Below, D. W. Bowden, J. C. Chambers, Y. J. Kim, M. C. Y. Ng, L. E. Petty, X. Sim, W. Zhang, A. J. Bennett, J. Bork-Jensen, C. M. Brummett, M. Canouil, K.-U. Ec kardt, K. Fischer, S. L. R. Kardia, F. Kronenberg, K. Läll, C.-T. Liu, A. E. Locke, J.'. Luan, I. Ntalla, V. Nylander, S. Schönherr, C. Schurmann, L. Yengo, E. P. Bottinger, I. Brandslund, C. Christensen, G. Dedoussis, J. C. Florez, I. Ford, O. H. Franco, T. M. Frayling,

V. Giedraitis, S. Hackinger, A. T. Hattersley, C. Herder, M. A. Ikram, M. Ingelsson, M. E. Jørgensen, T. Jørgensen, J. Kriebel, J. Kuusisto, S. Ligthart, C. M. Lindgren, A. Linneberg, V. Lyssenko, V. Mamakou, T. Meitinger, K. L. Mohlke, A. D. Morris, G. Nadkarni, J. S. Pankow, A. Peters, N. Sattar, A. Stančáková, K. Strauch, K. D. Taylor, B. Thorand, G. Thorleifsson, U. Thorsteinsdottir, J. Tuomilehto, D. R. Witte, J. Dupuis, P. A. Peyser, E. Zeggini, R. J. F. Loos, P. Froguel, E. Ingelsson, L. Lind, L. Groop, M. Laakso, F. S. Collins, J. W. Jukema, C. N. A. Palmer, H. Grallert, A. Metspalu, A. Dehghan, A. Köttgen, G. R. Abecasis, J. B. Meigs, J. I. Rotter, J. Marchini, O. Pedersen, T. Hansen, C. Langenberg, N. J. Wareham, K. Stefansson, A. L. Gloyn, A. P. Morris, M. Boehnke, M. I. McCarthy, Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nat. Genet.* **50**, 1505–1513 (2018).

119. J. C. Lambert, C. A. Ibrahim-Verbaas, D. Harold, A. C. Naj, R. Sims, C. Bellenguez, A. L. De Stafano, C. Bis, G. W. Beecham, B. Grenier-Boley, G. Russo, T. A. Thorton-Wells, N. Jones, A. V. Smith, V. Chouraki, C. Thomas, M. A. Ikram, D. Zelenika, B. N. Vardarajan, Y. Kamatani, C. F. Lin, A. Gerrish, H. Schmidt, B. Kunkle, M. L. Dunstan, A. Ruiz, M. T. Bihoreau, S. H. Choi, C. Reitz, F. Pasquier, C. Cruchaga, D. Craig, N. Amin, C. Berr, O. L. Lopez, P. L. De Jager, V. Deramecourt, J. A. Johnston, D. Evans, S. Lovestone, L. Letenneur, F. J. Morón, D. C. Rubinsztein, G. Eiriksdottir, K. Sleegers, A. M. Goate, N. Fiévet, M. W. Huentelman, M. Gill, K. Brown, M. I. Kamboh, L. Keller, P. Barberger-Gateau, B. McGuiness, E. B. Larson, R. Green, A. J. Myers, C. Dufouil, S. Todd, D. Wallon, S. Love, E. Rogaeva, J. Gallacher, P. St George-Hyslop, J. Clarimon, A. Lleo, A. Bayer, D. W. Tsuang, L. Yu, M. Tsolaki, P. Bossù, G. Spalletta, P. Proitsi, J. Collinge, S. Sorbi, F. Sanchez-Garcia, N. C. Fox, J. Hardy, M. C. D. Naranjo, P. Bosco, R. Clarke, C. Brayne, D. Galimberti, M. Mancuso, F. Matthews; European Alzheimer's Disease Initiative (EADI); Genetic and Environmental Risk in Alzheimer's Disease; Alzheimer's Disease Genetic Consortium; Cohorts for Heart, Aging Research in Genomic Epidemiology, S. Moebus, P. Mecocci, M. Del Zompo, W. Maier, H. Hampel, A. Pilotto, M. Bullido, F. Panza, P. Caffarra, B. Nacmias, J. R. Gilbert, M. Mayhaus, L. Lannefelt, H. Hakonarson, S. Pichler, M. M. Carrasquillo, M. Ingelsson, D. Beekly, V. Alvarez, F. Zou, O. Valladares, S. G. Younkin, E. Coto, K. L. Hamilton-Nelson, W. Gu, C. Razquin, P. Pastor, I. Mateo, M. J. Owen, K. M. Faber, P. V. Jonsson, O. Combarros, M. C. O'Donovan, L. B. Cantwell, H. Soininen, D. Blacker, S. Mead, T. H. Mosley Jr., D. A. Bennett, T. B. Harris, L. Fratiglioni, C. Holmes, R. F. de Bruijn, P. Passmore, T. J. Montine, K. Bettens, J. I. Rotter, A. Brice, K. Morgan, T. M. Foroud, W. A. Kukull, D. Hannequin, J. F. Powell, M. A. Nalls, K. Ritchie, K. L. Lunetta, J. S. Kauwe, E. Boerwinkle, M. Riemenschneider, M. Boada, M. Hiltuenen, E. R. Martin, R. Schmidt, D. Rujescu, L. S. Wang, J. F. Dartigues, R. Mayeux, C. Tzourio, A. Hofman, M. M. Nöthen, C. Graff, B. M. Psaty, L. Jones, J. L. Haines, P. A. Holmans, M. Lathrop, M. A. Pericak-Vance, L. J. Launer, L. A. Farrer, C. M. van Duijn, C. Van Broeckhoven, V. Moskvina, S. Seshadri, J. Williams, G. D. Schellenberg, P. Amouyel, Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat. Genet.* **45**, 1452–1458 (2013).

120. Y. Okada, D. Wu, G. Trynka, T. Raj, C. Terao, K. Ikari, Y. Kochi, A. Ohmura, A. Suzuki, S. Yoshida, R. R. Graham, A. Manoharan, W. Ortmann, T. Bhangale, J. C. Denny, R. J. Carroll, A. E. Eyler, J. D. Greenberg, J. M. Kremer, D. A. Pappas, L. Jiang, J. Yin, L. Ye, D.-F. Su, J. Yang, G. Xie, E. Keystone, H.-J. Westra, T. Esko, A. Metspalu, X. Zhou, N. Gupta, D. Mirel, E. A. Stahl, D. Diogo, J. Cui, K. Liao, M. H. Guo, K. Myouzen, T. Kawaguchi, M. J. H. Coenen, P. L. C. M. van Riel, M. A. F. J. van de Laar, H.-J. Guchelaar, T. W. J. Huizinga, P. Dieudé, X. Mariette, S. Louis Bridges Jr., A. Zhernakova, R. E. M. Toes, P. P. Tak, C. Miceli-Richard, S.-Y. Bang, H.-S. Lee, J. Martin, M. A. Gonzalez-Gay, L. Rodriguez-Rodriguez, S. Rantapää-Dahlqvist, L. Arlestig, H. K. Choi, Y. Kamatani, P. Galan, M. Lathrop; RACI consortium; GARNET consortium, S. Eyre, J. Bowes, A. Barton, N. de Vries, L. W. Moreland, L. A. Criswell, E. W. Karlson, A. Taniguchi, R. Yamada, M. Kubo, J. S. Liu, S.-C. Bae, J. Worthington, L. Padyukov, L. Klareskog, P. K. Gregersen, S. Raychaudhuri, B. E. Stranger, P. L. De Jager, L. Franke, P. M. Visscher, M. A. Brown, H. Yamanaka, T. Mimori, A. Takahashi, H. Xu, T. W. Behrens, K. A. Siminovitch, S. Momohara, F. Matsuda, K. Yamamoto, R. M. Plenge, Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature* **506**, 376–381 (2014).

121. J. Bentham, D. L. Morris, D. S. Cunninghame Graham, C. L. Pinder, P. Tombleson, T. W. Behrens, J. Martín, B. P. Fairfax, J. C. Knight, L. Chen, J. Replogle, A. C. Syvänen, L. Rönnblom, R. R. Graham, J. E. Wither, J. D. Rioux, M. E. Alarcón-Riquelme, T. J. Vyse, Genetic association analyses implicate aberrant regulation of innate and adaptive immunity genes in the pathogenesis of systemic lupus erythematosus. *Nat. Genet.* **47**, 1457–1464 (2015).

122. A. Aylward, J. Chiou, M. L. Okino, N. Kadakia, K. J. Gaulton, Shared genetic risk contributes to type 1 and type 2 diabetes etiology. *Hum. Mol. Genet.*, (2018).

123. M. den Hoed, M. Eijgelsheim, T. Esko, B. J. J. M. Brundel, D. S. Peal, D. M. Evans, I. M. Nolte, A. V. Segrè, H. Holm, R. E. Handsaker, H.-J. Westra, T. Johnson, A. Isaacs, J. Yang, A. Lundby, J. H. Zhao, Y. J. Kim, M. J. Go, P. Almgren, M. Bochud, G. Boucher, M. C. Cornelis, D. Gudbjartsson, D. Hadley, P. van der Harst, C. Hayward, M. den Heijer, W. Igl, A. U. Jackson, Z. Kutalik, J. Luan, J. P. Kemp, K. Kristiansson, C. Ladenvall, M. Lorentzon, M. E. Montasser, O. T. Njajou, P. F. O'Reilly, S. Padmanabhan, B. S. Pourcain,

T. Rankinen, P. Salo, T. Tanaka, N. J. Timpson, V. Vitart, L. Waite, W. Wheeler, W. Zhang, H. H. M. Draisma, M. F. Feitosa, K. F. Kerr, P. A. Lind, E. Mihailov, N. C. Onland-Moret, C. Song, M. N. Weedon, W. Xie, L. Yengo, D. Absher, C. M. Albert, A. Alonso, D. E. Arking, P. I. W. de Bakker, B. Balkau, C. Barlassina, P. Benaglio, J. C. Bis, N. Bouatia-Naji, S. Brage, S. J. Chanock, P. S. Chines, M. Chung, D. Darbar, C. Dina, M. Dörr, P. Elliott, S. B. Felix, K. Fischer, C. Fuchsberger, E. J. C. de Geus, P. Goyette, V. Gudnason, T. B. Harris, A.-L. Hartikainen, A. S. Havulinna, S. R. Heckbert, A. A. Hicks, A. Hofman, A. Holewijn, F. Hoogstra-Berends, J.-J. Hottenga, M. K. Jensen, A. Johansson, J. Junttila, S. Kääb, B. Kanon, S. Ketkar, K.-T. Khaw, J. W. Knowles, A. S. Kooner, J. A. Kors, M. Kumari, L. Milani, P. Laiho, E. G. Lakatta, C. Langenberg, M. Leusink, Y. Liu, R. N. Luben, K. L. Lunetta, S. N. Lynch, M. R. P. Markus, P. Marques-Vidal, I. M. Leach, W. L. McArdle, S. A. McCarroll, S. E. Medland, K. A. Miller, G. W. Montgomery, A. C. Morrison, M. Müller-Nurasyid, P. Navarro, M. Nelis, J. R. O'Connell, C. J. O'Donnell, K. K. Ong, A. B. Newman, A. Peters, O. Polasek, A. Pouta, P. P. Pramstaller, B. M. Psaty, D. C. Rao, S. M. Ring, E. J. Rossin, D. Rudan, S. Sanna, R. A. Scott, J. S. Sehmi, S. Sharp, J. T. Shin, A. B. Singleton, A. V. Smith, N. Soranzo, T. D. Spector, C. Stewart, H. M. Stringham, K. V. Tarasov, A. G. Uitterlinden, L. Vandenput, S.-J. Hwang, J. B. Whitfield, C. Wijmenga, S. H. Wild, G. Willemsen, J. F. Wilson, J. C. M. Witteman, A. Wong, Q. Wong, Y. Jamshidi, P. Zitting, J. M. A. Boer, D. I. Boomsma, I. B. Borecki, C. M. van Duijn, U. Ekelund, N. G. Forouhi, P. Froguel, A. Hingorani, E. Ingelsson, M. Kivimaki, R. A. Kronmal, D. Kuh, N. Lind, N. G. Martin, B. A. Oostra, N. L. Pedersen, T. Quertermous, J. I. Rotter, Y. T. van der Schouw, W. M. M. Verschuren, M. Walker, D. Albanes, D. O. Arnar, T. L. Assimes, S. Bandinelli, M. Boehnke, R. A. de Boer, C. Bouchard, W. L. M. Caulfield, J. C. Chambers, G. Curhan, D. Cusi, J. Eriksson, L. Ferrucci, W. H. van Gilst, N. Glorioso, J. de Graaf, L. Groop, U. Gyllensten, W.-C. Hsueh, F. B. Hu, H. V. Huikuri, D. J. Hunter, C. Iribarren, B. Isomaa, M.-R. Jarvelin, A. Jula, M. Kähönen, L. A. Kiemeney, M. M. van der Klauw, J. S. Kooner, P. Kraft, L. Iacoviello, T. Lehtimäki, M.-L. L. Lokki, B. D. Mitchell, G. Navis, M. S. Nieminen, C. Ohlsson, N. R. Poulter, L. Qi, O. T. Raitakari, E. B. Rimm, J. D. Rioux, F. Rizzi, I. Rudan, V. Salomaa, P. S. Sever, D. C. Shields, A. R. Shuldiner, J. Sinisalo, A. V. Stanton, R. P. Stolk, D. P. Strachan, J.-C. Tardif, U. Thorsteinsdottir, J. Tuomilehto, D. J. van Veldhuisen, J. Virtamo, J. Viikari, P. Vollenweider, G. Waeber, E. Widen, Y. S. Cho, J. V. Olsen, P. M. Visscher, C. Willer, L. Franke; Global BPgen Consortium; CARDIo GRAM Consortium, J. Erdmann, J. R. Thompson; PR GWAS Consortium, A. Pfeufer; QRS GWAS Consortium, N. Sotoodehnia; QT-IGC Consortium, C. Newton-Cheh; CHARGE-AF Consortium, P. T. Ellinor, B. H. C. Stricker, A. Metspalu, M. Perola, J. S. Beckmann, G. D. Smith, K. Stefansson, N. J. Wareham, P. B. Munroe, O. C. M. Sibon, D. J. Milan, H. Snieder, N. J. Samani, R. J. F. Loos, Identification of heart rate–associated loci and their effects on cardiac conduction and rhythm disorders. *Nat. Genet.* **45**, 621–631 (2013).

124. A. Okbay, J. P. Beauchamp, M. A. Fontana, J. J. Lee, T. H. Pers, C. A. Rietveld, P. Turley, G.-B. Chen, V. Emilsson, S. F. W. Meddens, S. Oskarsson, J. K. Pickrell, K. Thom, P. Timshel, R. de Vlaming, A. Abdellaoui, T. S. Ahluwalia, J. Bacelis, C. Baumbach, G. Bjornsdottir, J. H. Brandsma, M. Pina Concas, J. Derringer, N. A. Furlotte, T. E. Galesloot, G. Girotto, R. Gupta, L. M. Hall, S. E. Harris, E. Hofer, M. Horikoshi, J. E. Huffman, K. Kaasik, I. P. Kalafati, R. Karlsson, A. Kong, J. Lahti, S. J. van der Lee, C. deLeeuw, P. A. Lind, K. O. Lindgren, T. Liu, M. Mangino, J. Marten, E. Mihailov, M. B. Miller, P. J. van der Most, C. Oldmeadow, A. Payton, N. Pervjakova, W. J. Peyrot, Y. Qian, O. Raitakari, R. Rueedi, E. Salvi, B. Schmidt, K. E. Schraut, J. Shi, A. V. Smith, R. A. Poot, B. St Pourcain, A. Teumer, G. Thorleifsson, N. Verweij, D. Vuckovic, J. Wellmann, H. J. Westra, J. Yang, W. Zhao, Z. Zhu, B. Z. Alizadeh, N. Amin, A. Bakshi, S. E. Baumeister, G. Biino, K. Bønnelykke, P. A. Boyle, H. Campbell, F. P. Cappuccio, G. Davies, J. E. de Neve, P. Deloukas, I. Demuth, J. Ding, P. Eibich, L. Eisele, N. Eklund, D. M. Evans, J. D. Faul, M. F. Feitosa, A. J. Forstner, I. Gandin, B. Gunnarsson, B. V. Halldórsson, T. B. Harris, A. C. Heath, L. J. Hocking, E. G. Holliday, G. Homuth, M. A. Horan, J. J. Hottenga, P. L. de Jager, P. K. Joshi, A. Jugessur, M. A. Kaakinen, M. Kähönen, S. Kanoni, J. Keltigangas-Järvinen, L. A. L. M. Kiemeney, I. Kolcic, S. Koskinen, A. T. Kraja, M. Kroh, Z. Kutalik, A. Latvala, L. J. Launer, M. P. Lebreton, D. F. Levinson, P. Lichtenstein, P. Lichtner, D. C. M. Liewald; LifeLines Cohort Study, A. Loukola, P. A. Madden, R. Mägi, T. Mäki-Opas, R. E. Marioni, P. Marques-Vidal, G. A. Meddens, G. McMahon, C. Meisinger, T. Meitinger, Y. Milaneschi, L. Milani, G. W. Montgomery, R. Myhre, C. P. Nelson, D. R. Nyholt, W. E. R. Ollier, A. Palotie, L. Paternoster, N. L. Pedersen, K. E. Petrovic, D. J. Porteous, K. Räikkönen, S. M. Ring, A. Robino, O. Rostapshova, I. Rudan, A. Rustichini, V. Salomaa, A. R. Sanders, A. P. Sarin, H. Schmidt, R. J. Scott, B. H. Smith, J. A. Smith, J. A. Staessen, E. Steinhagen-Thiessen, K. Strauch, A. Terracciano, M. D. Tobin, S. Ulivi, S. Vaccargiu, L. Quaye, F. J. A. van Rooij, C. Venturini, A. A. E. Vinkhuyzen, U. Völker, H. Völzke, J. M. Vonk, D. Vozzi, J. Waage, E. B. Ware, G. Willemsen, J. R. Attia, D. A. Bennett, K. Berger, L. Bertram, H. Bisgaard, D. I. Boomsma, I. B. Borecki, U. Bültmann, C. F. Chabris, F. Cucca, D. Cusi, I. J. Deary, G. V. Dedoussis, C. M. van Duijn, J. G. Eriksson, B. Franke, L. Franke, P. Gasparini, P. V. Gejman, C. Gieger, H. J. Grabe, J. Gratten, P. J. F. Groenen, V. Gudnason, P. van der Harst, C. Hayward, D. A. Hinds, W. Hoffmann, E. Hyppönen, W. G. Iacono, B. Jacobsson, M. R. Järvelin, K. H. Jöckel, J. Kaprio, S. L. R. Kardia, T. Lehtimäki, S. F. Lehrer, P. K. E. Magnusson, N. G. Martin, M. McGue, A. Metspalu, N. Pendleton,

B. W. J. H. Penninx, M. Perola, N. Pirastu, M. Pirastu, O. Polasek, D. Posthuma, C. Power, M. A. Province, N. J. Samani, D. Schlessinger, R. Schmidt, T. I. A. Sørensen, T. D. Spector, K. Stefansson, U. Thorsteinsdottir, A. R. Thurik, N. J. Timpson, H. Tiemeier, J. Y. Tung, A. G. Uitterlinden, V. Vitart, P. Vollenweider, D. R. Weir, J. F. Wilson, A. F. Wright, D. C. Conley, R. F. Krueger, G. Davey Smith, A. Hofman, D. I. Laibson, S. E. Medland, M. N. Meyer, J. Yang, M. Johannesson, P. M. Visscher, T. Esko, P. D. Koellinger, D. Cesarini, D. J. Benjamin, Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* **533**, 539–542 (2016).

125. A. E. Locke, B. Kahali, S. I. Berndt, A. E. Justice, T. H. Pers, F. R. Day, C. Powell, S. Vedantam, M. L. Buchkovich, J. Yang, D. C. Croteau-Chonka, T. Esko, T. Fall, T. Ferreira, S. Gustafsson, Z. Kutalik, J. Luan, R. Mägi, J. C. Randall, T. W. Winkler, A. R. Wood, T. Workalemahu, J. D. Faul, J. A. Smith, J. H. Zhao, W. Zhao, J. Chen, R. Fehrmann, Å. K. Hedman, J. Karjalainen, E. M. Schmidt, D. Absher, N. Amin, D. Anderson, M. Beekman, J. L. Bolton, J. L. Bragg-Gresham, S. Buyske, A. Demirkan, G. Deng, G. B. Ehret, B. Feenstra, M. F. Feitosa, K. Fischer, A. Goel, J. Gong, A. U. Jackson, S. Kanoni, M. E. Kleber, K. Kristiansson, U. Lim, V. Lotay, M. Mangino, I. M. Leach, C. Medina-Gomez, S. E. Medland, M. A. Nalls, C. D. Palmer, D. Pasko, S. Pechlivanis, M. J. Peters, I. Prokopenko, D. Shungin, A. Stančáková, R. J. Strawbridge, Y. J. Sung, T. Tanaka, A. Teumer, S. Trompet, S. W. van der Laan, J. van Setten, J. V. Van Vliet-Ostaptchouk, Z. Wang, L. Yengo, W. Zhang, A. Isaacs, E. Albrecht, J. Ärnlöv, G. M. Arscott, A. P. Attwood, S. Bandinelli, A. Barrett, I. N. Bas, C. Bellis, A. J. Bennett, C. Berne, R. Blagieva, M. Blüher, S. Böhringer, L. L. Bonnycastle, Y. Böttcher, H. A. Boyd, M. Bruinenberg, I. H. Caspersen, Y.-D. I. Chen, R. Clarke, E. W. Daw, A. J. M. de Craen, G. Delgado, M. Dimitriou, A. S. F. Doney, N. Eklund, K. Estrada, E. Eury, L. Folkersen, R. M. Fraser, M. E. Garcia, F. Geller, V. Giedraitis, B. Gigante, A. S. Go, A. Golay, A. H. Goodall, S. D. Gordon, M. Gorski, H.-J. Grabe, H. Grallert, T. B. Grammer, J. Gräßler, H. Grönberg, C. J. Groves, G. Gusto, J. Haessler, P. Hall, T. Haller, G. Hallmans, C. A. Hartman, M. Hassinen, C. Hayward, N. L. Heard-Costa, Q. Helmer, C. Hengstenberg, O. Holmen, J.-J. Hottenga, A. L. James, J. M. Jeff, Å. Johansson, J. Jolley, T. Juliusdottir, L. Kinnunen, W. Koenig, M. Koskenvuo, W. Kratzer, J. Laitinen, C. Lamina, K. Leander, N. R. Lee, P. Lichtner, L. Lind, J. Lindström, K. S. Lo, S. Lobbens, R. Lorbeer, Y. Lu, F. Mach, P. K. E. Magnusson, A. Mahajan, W. McArdle, S. M. Lachlan, C. Menni, S. Merger, E. Mihailov, L. Milani, A. Moayyeri, K. L. Monda, M. A. Morken, A. Mulas, G. Müller, M. Müller-Nurasyid, A. W. Musk, R. Nagaraja, M. M. Nöthen, I. M. Nolte, S. Pilz, N. W. Rayner, F. Renstrom, R. Rettig, J. S. Ried, S. Ripke, N. R. Robertson, L. M. Rose, S. Sanna, H. Scharnagl, S. Scholtens, F. R. Schumacher, W. R. Scott, T. Seufferlein, J. Shi, A. V. Smith, J. Smolonska, A. V. Stanton, V. Steinthorsdottir, K. Stirrups, H. M. Stringham, J. Sundström, M. A. Swertz, A. J. Swift, A.-C. Syvänen, S.-T. Tan, B. O. Tayo, B. Thorand, G. Thorleifsson, J. P. Tyrer, H.-W. Uh, L. Vandenput, F. C. Verhulst, S. H. Vermeulen, N. Verweij, J. M. Vonk, L. L. Waite, H. R. Warren, D. Waterworth, M. N. Weedon, L. R. Wilkens, C. Willenborg, T. Wilsgaard, M. K. Wojczynski, A. Wong, A. F. Wright, Q. Zhang; Life Lines Cohort Study, E. P. Brennan, M. Choi, Z. Dastani, A. W. Drong, P. Eriksson, A. Franco-Cereceda, J. R. Gådin, A. G. Gharavi, M. E. Goddard, R. E. Handsaker, J. Huang, F. Karpe, S. Kathiresan, S. Keildson, K. Kiryluk, M. Kubo, J.-Y. Lee, L. Liang, R. P. Lifton, B. Ma, S. A. McCarroll, A. J. McKnight, J. L. Min, M. F. Moffatt, G. W. Montgomery, J. M. Murabito, G. Nicholson, D. R. Nyholt, Y. Okada, J. R. B. Perry, R. Dorajoo, E. Reinmaa, R. M. Salem, N. Sandholm, R. A. Scott, L. Stolk, A. Takahashi, T. Tanaka, F. M. van't Hooft, A. A. E. Vinkhuyzen, H.-J. Westra, W. Zheng, K. T. Zondervan; ADIPOGen Consortium; AGEN-BMI Working Group; CARDIOGRAMplusCD Consortium; CKDGen Consortium; GLGC; ICBP; MAGIC Investigators; MuTHER Consortium; MIGen Consortium; PAGE Consortium; ReproGen Consortium; GENIE Consortium; International Endogene Consortium, A. C. Heath, D. Arveiler, S. J. L. Bakker, J. Beilby, R. N. Bergman, J. Blangero, P. Bovet, H. Campbell, M. J. Caulfield, G. Cesana, A. Chakravarti, D. I. Chasman, P. S. Chines, F. S. Collins, D. C. Crawford, L. A. Cupples, D. Cusi, J. Danesh, U. de Faire, H. M. den Ruijter, A. F. Dominiczak, R. Erbel, J. Erdmann, J. G. Eriksson, M. Farrall, S. B. Felix, E. Ferrannini, J. Ferrières, I. Ford, N. G. Forouhi, O. H. Franco, R. T. Gansevoort, P. V. Gejman, C. Gieger, C. Gottesman, V. Gudnason, U. Gyllensten, A. S. Hall, T. B. Harris, A. T. Hattersley, A. A. Hicks, L. A. Hindorff, A. Hingorani, A. Hofman, G. Homuth, G. K. Hovingh, S. E. Humphries, S. C. Hunt, E. Hyppönen, T. Illig, K. B. Jacobs, M.-R. Järvelin, K.-H. Jöckel, B. Johansen, P. Jousilahti, J. W. Jukema, A. M. Jula, J. Kaprio, J. J. P. Kastelein, S. M. Keinanen-Kiukaanniemi, L. A. Kiemeney, P. Knekt, J. S. Kooner, C. Kooperberg, P. Kovacs, A. T. Kraja, M. Kumari, J. Kuusisto, T. A. Lakka, C. Langenberg, L. L. Marchand, T. Lehtimäki, V. Lyssenko, S. Männistö, A. Marette, T. C. Matise, C. A. McKenzie, B. M. Knight, F. L. Moll, A. D. Morris, A. P. Morris, J. C. Murray, M. Nelis, C. Ohlsson, A. J. Oldehinkel, K. K. Ong, P. A. F. Madden, G. Pasterkamp, J. F. Peden, A. Peters, D. S. Postma, P. P. Pramstaller, J. F. Price, L. Qi, O. T. Raitakari, T. Rankinen, D. C. Rao, T. K. Rice, P. M. Ridker, J. D. Rioux, M. D. Ritchie, I. Rudan, V. Salomaa, N. J. Samani, J. Saramies, M. A. Sarzynski, H. Schunkert, P. E. H. Schwarz, P. Sever, A. R. Shuldiner, J. Sinisalo, R. P. Stolk, K. Strauch, A. Tönjes, D.-A. Trégouët, A. Tremblay, E. Tremoli, J. Virtamo, M.-C. Vohl, U. Völker, G. Waeber, G. Willemsen, J. C. Witteman, M. C. Zillikens, L. S. Adair, P. Amouyel, F. W. Asselbergs, T. L. Assimes, M. Bochud, B. O. Boehm, E. Boerwinkle, S. R. Bornstein, E. P. Bottinger, C. Bouchard, S. Cauchi, J. C. Chambers,

S. J. Chanock, R. S. Cooper, P. I. W. de Bakker, G. Dedoussis, L. Ferrucci, P. W. Franks, P. Froguel, L. C. Groop, C. A. Haiman, A. Hamsten, J. Hui, D. J. Hunter, K. Hveem, R. C. Kaplan, M. Kivimaki, D. Kuh, M. Laakso, Y. Liu, N. G. Martin, W. März, M. Melbye, A. Metspalu, S. Moebus, P. B. Munroe, I. Njølstad, B. A. Oostra, C. N. A. Palmer, N. L. Pedersen, M. Perola, C. Power, T. Quertermous, R. Rauramaa, F. Rivadeneira, T. E. Saaristo, D. Saleheen, N. Sattar, E. E. Schadt, D. Schlessinger, P. E. Slagboom, H. Snieder, T. D. Spector, U. Thorsteinsdottir, M. Stumvoll, J. Tuomilehto, A. G. Uitterlinden, M. Uusitupa, P. van der Harst, M. Walker, H. Wallaschofski, N. J. Wareham, H. Watkins, D. R. Weir, H.-E. Wichmann, J. F. Wilson, P. Zanen, I. B. Borecki, P. Deloukas, C. S. Fox, I. M. Heid, J. R. O'Connell, D. P. Strachan, K. Stefansson, C. M. van Duijn, G. R. Abecasis, L. Franke, T. M. Frayling, M. I. McCarthy, P. M. Visscher, A. Scherag, C. J. Willer, M. Boehnke, K. L. Mohlke, C. M. Lindgren, J. S. Beckmann, I. Barroso, K. E. North, E. Ingelsson, J. N. Hirschhorn, R. J. F. Loos, E. K. Speliotes, Genetic studies of body mass index yield new insights for obesity biology. *Nature* **518**, 197–I206 (2015).

126. F. R. Day, D. J. Thompson, H. Helgason, D. I. Chasman, H. Finucane, P. Sulem, K. S. Ruth, S. Whalen, A. K. Sarkar, E. Albrecht, E. Altmaier, M. Amini, C. M. Barbieri, T. Boutin, A. Campbell, E. Demerath, A. Giri, C. He, J. J. Hottenga, R. Karlsson, I. Kolcic, P.-R. Loh, K. L. Lunetta, M. Mangino, B. Marco, G. M. Mahon, S. E. Medland, I. M. Nolte, R. Noordam, T. Nutile, L. Paternoster, N. Perjakova, E. Porcu, L. M. Rose, K. E. Schraut, A. V. Segrè, A. V. Smith, L. Stolk, A. Teumer, I. L. Andrulis, S. Bandinelli, M. W. Beckmann, J. Benitez, S. Bergmann, M. Bochud, E. Boerwinkle, S. E. Bojesen, M. K. Bolla, J. S. Brand, H. Brauch, H. Brenner, L. Broer, B. Brüning, J. E. Buring, H. Campbell, E. Catamo, S. Chanock, G. Chenevix-Trench, T. Corre, F. J. Couch, D. L. Cousminer, A. Cox, L. Crisponi, K. Czene, G. D. Smith, E. J. C. N. de Geus, R. de Mutsert, I. De Vivo, J. Dennis, P. Devilee, I. Dos-Santos-Silva, A. M. Dunning, J. G. Eriksson, P. A. Fasching, L. Fernández-Rhodes, L. Ferrucci, D. Flesch-Janys, L. Franke, M. Gabrielson, I. Gandin, G. G. Giles, H. Grallert, D. F. Gudbjartsson, P. Guénel, P. Hall, E. Hallberg, U. Hamann, T. B. Harris, C. A. Hartman, G. Heiss, M. J. Hooning, J. L. Hopper, F. Hu, D. J. Hunter, M. A. Ikram, H. K. Im, M.-R. Järvelin, P. K. Joshi, D. Karasik, M. Kellis, Z. Kutalik, G. L. Chance, C. Lambrechts, C. Langenberg, L. J. Launer, J. S. E. Laven, S. Lenarduzzi, J. Li, P. A. Lind, S. Lindstrom, Y. M. Liu, J. Luan, R. Mägi, A. Mannermaa, H. Mbarek, M. I. McCarthy, C. Meisinger, T. Meitinger, C. Menni, A. Metspalu, K. Michailidou, L. Milani, R. L. Milne, G. W. Montgomery, A. M. Mulligan, M. A. Nalls, P. Navarro, H. Nevanlinna, D. R. Nyholt, A. J. Oldehinkel, T. A. O'Mara, S. Padmanabhan, A. Palotie, N. Pedersen, A. Peters, J. Peto, P. D. P. Pharoah, A. Pouta, P. Radice, I. Rahman, S. M. Ring, A. Robino, F. Rosendaal, I. Rudan, R. Rueedi, D. Ruggiero, C. F. Sala, M. K. Schmidt, R. A. Scott, M. Shah, R. Sorice, M. C. Southey, U. Sovio, M. Stampfer, M. Steri, K. Strauch, T. Tanaka, E. Tikkanen, N. J. Timpson, M. Traglia, T. Truong, J. P. Tyrer, A. G. Uitterlinden, D. R. V. Edwards, V. Vitart, U. Völker, P. Vollenweider, Q. Wang, E. Widen, K. W. van Dijk, G. Willemsen, R. Winqvist, B. H. R. Wolffenbuttel, J. H. Zhao, M. Zoledziewska, M. Zygmunt, B. Z. Alizadeh, D. I. Boomsma, M. Ciullo, F. Cucca, T. Esko, N. Franceschini, C. Gieger, V. Gudnason, C. Hayward, P. Kraft, D. A. Lawlor, P. K. E. Magnusson, N. G. Martin, D. O. Mook-Kanamori, E. A. Nohr, O. Polasek, D. Porteous, A. L. Price, P. M. Ridker, H. Snieder, T. D. Spector, D. Stöckl, D. Toniolo, S. Ulivi, J. A. Visser, H. Völzke, N. J. Wareham, J. F. Wilson; Life Lines Cohort Study; InterAct Consortium; kConFab/AOCS Investigators; Endometrial Cancer Association Consortium; Ovarian Cancer Association Consortium; PRACTICAL consortium, A. B. Spurdle, U. Thorsteindottir, K. S. Pollard, D. F. Easton, J. Y. Tung, J. Chang-Claude, D. Hinds, A. Murray, J. M. Murabito, K. Stefansson, K. K. Ong, J. R. B. Perry, Genomic analyses identify hundreds of variants associated with age at menarche and support a role for puberty timing in cancer risk. *Nat. Genet.* **49**, 834–841 (2017).

127. F. R. Day, K. S. Ruth, D. J. Thompson, K. L. Lunetta, N. Pervjakova, D. I. Chasman, L. Stolk, H. K. Finucane, P. Sulem, B. Bulik-Sullivan, T. Esko, A. D. Johnson, C. E. Elks, N. Franceschini, C. He, E. Altmaier, J. A. Brody, L. L. Franke, J. E. Huffman, M. F. Keller, P. F. McArdle, T. Nutile, E. Porcu, A. Robino, L. M. Rose, U. M. Schick, J. A. Smith, A. Teumer, M. Traglia, D. Vuckovic, J. Yao, W. Zhao, E. Albrecht, N. Amin, T. Corre, J.-J. Hottenga, M. Mangino, A. V. Smith, T. Tanaka, G. Abecasis, I. L. Andrulis, H. Anton-Culver, A. C. Antoniou, V. Arndt, A. M. Arnold, C. Barbieri, M. W. Beckmann, A. Beeghly-Fadiel, J. Benitez, L. Bernstein, S. J. Bielinski, C. Blomqvist, E. Boerwinkle, N. V. Bogdanova, S. E. Bojesen, M. K. Bolla, A.-L. Borresen-Dale, T. S. Boutin, H. Brauch, H. Brenner, T. Brüning, B. Burwinkel, A. Campbell, H. Campbell, S. J. Chanock, J. R. Chapman, Y.-D. I. Chen, G. Chenevix-Trench, F. J. Couch, A. D. Coviello, A. Cox, K. Czene, H. Darabi, I. De Vivo, E. W. Demerath, J. Dennis, P. Devilee, T. Dörk, I. Dos-Santos-Silva, A. M. Dunning, D. Eicher, P. A. Fasching, J. D. Faul, J. Figueroa, D. Flesch-Janys, I. Gandin, M. E. Garcia, M. García-Closas, G. G. Giles, G. G. Girotto, M. S. Goldberg, A. González-Neira, M. O. Goodarzi, M. L. Grove, D. F. Gudbjartsson, P. Guénel, X. Guo, C. A. Haiman, P. Hall, U. Hamann, B. E. Henderson, L. J. Hocking, A. Hofman, G. Homuth, M. J. Hooning, J. L. Hopper, F. B. Hu, J. Huang, K. Humphreys, D. J. Hunter, A. Jakubowska, S. E. Jones, M. Kabisch, D. Karasik, J. A. Knight, I. Kolcic, C. Kooperberg, V.-M. Kosma, J. Kriebel, V. Kristensen, D. Lambrechts, C. Langenberg, J. Li, X. Li, S. Lindström, Y. Liu, J. Luan, J. Lubinski, R. Mägi, A. Mannermaa, J. Manz, S. Margolin, J. Marten, N. G. Martin, C. Masciullo, A. Meindl, K. Michailidou,

E. Mihailov, L. Milani, R. L. Milne, M. Müller-Nurasyid, M. Nalls, B. M. Neale, H. Nevanlinna, P. Neven, A. B. Newman, B. G. Nordestgaard, J. E. Olson, S. Padmanabhan, P. Peterlongo, U. Peters, A. Petersmann, J. Peto, P. D. P. Pharoah, N. N. Pirastu, A. Pirie, G. Pistis, O. Polasek, D. Porteous, B. M. Psaty, K. Pylkäs, P. Radice, L. J. Raffel, F. Rivadeneira, I. Rudan, A. Rudolph, D. Ruggiero, C. F. Sala, S. Sanna, E. J. Sawyer, D. Schlessinger, M. K. Schmidt, F. Schmidt, R. K. Schmutzler, M. J. Schoemaker, R. A. Scott, C. M. Seynaeve, J. Simard, R. Sorice, M. C. Southey, D. Stöckl, K. Strauch, A. Swerdlow, K. D. Taylor, U. Thorsteinsdottir, A. E. Toland, I. Tomlinson, T. Truong, L. Tryggvadottir, S. T. Turner, D. Vozzi, Q. Wang, M. Wellons, G. Willemsen, J. F. Wilson, R. Winqvist, B. B. H. R. Wolffenbuttel, A. F. Wright, D. Yannoukakos, T. Zemunik, W. Zheng, M. Zygmunt, S. Bergmann, D. I. Boomsma, J. E. Buring, L. Ferrucci, G. W. Montgomery, V. Gudnason, T. D. Spector, C. M. van Duijn, B. Z. Alizadeh, M. Ciullo, L. Crisponi, D. F. Easton, P. P. Gasparini, C. Gieger, T. B. Harris, C. Hayward, S. L. R. Kardia, P. Kraft, B. M. Knight, A. Metspalu, A. C. Morrison, A. P. Reiner, P. M. Ridker, J. I. Rotter, D. Toniolo, A. G. Uitterlinden, S. Ulivi, H. Völzke, N. J. Wareham, D. R. Weir, L. M. Yerges-Armstrong; PRACTICAL consortium; kConFab Investigators; AOCS Investigators; Generation Scotland; EPIC-InterAct Consortium; Life Lines Cohort Study, A. L. Price, K. Stefansson, J. A. Visser, K. K. Ong, J. Chang-Claude, J. M. Murabito, J. R. B. Perry, A. Murray, Large-scale genomic analyses link reproductive aging to hypothalamic signaling, breast cancer susceptibility and BRCA1-mediated DNA repair. *Nat. Genet.* **47**, 1294–1303 (2015).

128. A. R. Wood, T. Esko, J. Yang, S. Vedantam, T. H. Pers, S. Gustafsson, A. Y. Chu, K. Estrada, J. Luan, Z. Kutalik, N. Amin, M. L. Buchkovich, D. C. Croteau-Chonka, F. R. Day, Y. Duan, T. Fall, R. Fehrmann, T. Ferreira, A. U. Jackson, J. Karjalainen, K. S. Lo, A. E. Locke, M. Mägi, E. Mihailov, E. Porcu, J. C. Randall, A. Scherag, A. A. E. Vinkhuyzen, H.-J. Westra, T. W. Winkler, T. Workalemahu, J. H. Zhao, D. Absher, E. Albrecht, D. Anderson, J. Baron, M. Beekman, A. Demirkan, G. B. Ehret, B. Feenstra, M. F. Feitosa, K. Fischer, R. M. Fraser, A. Goel, J. Gong, A. E. Justice, S. Kanoni, M. E. Kleber, K. Kristiansson, U. Lim, V. Lotay, J. C. Lui, M. Mangino, I. M. Leach, C. Medina-Gomez, M. A. Nalls, D. R. Nyholt, C. D. Palmer, D. Pasko, S. Pechlivanis, I. Prokopenko, J. S. Ried, S. Ripke, D. Shungin, A. Stancáková, R. J. Strawbridge, Y. J. Sung, T. Tanaka, A. Teumer, S. Trompet, S. W. van der Laan, J. van Setten, J. V. Van Vliet-Ostaptchouk, Z. Wang, L. Yengo, W. Zhang, U. Afzal, J. Arnlöv, G. M. Arscott, S. Bandinelli, A. Barrett, C. Bellis, A. J. Bennett, C. Berne, M. Blüher, J. L. Bolton, Y. Böttcher, H. A. Boyd, M. Bruinenberg, B. M. Buckley, S. Buyske, I. H. Caspersen, P. S. Chines, R. Clarke, S. Claudi-Boehm, M. Cooper, E. W. Daw, P. A. De Jong, J. Deelen, G. Delgado, J. C. Denny, R. Dhonukshe-Rutten, M. Dimitriou, A. S. F. Doney, M. Dörr, N. Eklund, E. Eury, L. Folkersen, M. E. Garcia, F. Geller, V. Giedraitis, A. S. Go, H. Grallert, T. B. Grammer, J. Gräßler, H. Grönberg, L. C. P. G. M. de Groot, C. J. Groves, J. Haessler, P. Hall, T. Haller, G. Hallmans, A. Hannemann, C. A. Hartman, M. Hassinen, C. Hayward, N. L. Heard-Costa, Q. Helmer, G. Hemani, A. K. Henders, H. L. Hillege, M. A. Hlatky, W. Hoffmann, P. Hoffmann, O. Holmen, J. J. Houwing-Duistermaat, T. Illig, A. Isaacs, A. L. James, J. Jeff, B. Johansen, Å. Johansson, J. Jolley, T. Juliusdittir, J. Junttila, A. N. Kho, L. Kinnunen, N. Klopp, T. Kocher, W. Kratzer, P. Lichtner, L. Lind, J. Lindström, S. Lobbens, M. Lorentzon, Y. Lu, V. Lyssenko, P. K. E. Magnusson, A. Mahajan, M. Maillard, W. L. McArdle, C. A. McKenzie, S. M. Lachlan, P. J. McLaren, C. Menni, S. Merger, L. Milani, A. Moayyeri, K. L. Monda, M. A. Morken, G. Müller, M. Müller-Nurasyid, A. W. Musk, N. Narisu, M. Nauck, I. M. Nolte, M. M. Nöthen, L. Oozageer, S. Pilz, N. W. Rayner, F. Renstrom, N. R. Robertson, L. M. Rose, R. Roussel, S. Sanna, H. Scharnagl, S. Scholtens, F. R. Schumacher, H. Schunkert, R. A. Scott, J. Sehmi, T. Seufferlein, J. Shi, K. Silventoinen, J. H. Smit, A. V. Smith, J. Smolonska, A. V. Stanton, K. Stirrups, D. J. Stott, H. M. Stringham, J. Sundström, M. A. Swertz, A.-C. Syvänen, B. O. Tayo, S. Thorleifsson, J. P. Tyrer, S. van Dijk, N. M. van Schoor, N. van der Velde, D. van Heemst, F. V. A. van Oort, S. H. Vermeulen, N. Verweij, J. M. Vonk, L. L. Waite, M. Waldenberger, R. Wennauer, L. R. Wilkens, C. Willenborg, T. Wilsgaard, M. K. Wojczynski, A. Wong, A. F. Wright, Q. Zhang, D. Arveiler, S. J. L. Bakker, J. Beilby, R. N. Bergman, S. Bergmann, R. Biffar, J. Blangero, D. I. Boomsma, S. R. Bornstein, P. Bovet, P. Brambilla, M. J. Brown, H. Campbell, M. J. Caulfield, A. Chakravarti, R. Collins, F. S. Collins, D. C. Crawford, L. A. Cupples, J. Danesh, U. de Faire, H. M. den Ruijter, P. Erbel, J. Erdmann, J. G. Eriksson, M. Farrall, E. Ferrannini, J. Ferrières, I. Ford, N. G. Forouhi, T. Forrester, R. T. Gansevoort, P. V. Gejman, C. Gieger, A. Golay, O. Gottesman, V. Gudnason, U. Gyllensten, D. W. Haas, A. S. Hall, T. B. Harris, A. T. Hattersley, A. C. Heath, C. Hengstenberg, A. A. Hicks, L. A. Hindorff, A. D. Hingorani, A. Hofman, G. K. Hovingh, S. E. Humphries, S. C. Hunt, E. Hyppönen, K. B. Jacobs, M.-R. Jarvelin, P. Jousilahti, A. M. Jula, J. Kaprio, J. J. P. Kastelein, M. Kayser, F. Kee, S. M. Keinanen-Kiukaanniemi, L. A. Kiemeney, J. S. Kooner, C. Kooperberg, S. Koskinen, P. Kovacs, A. T. Kraja, M. Kumari, J. Kuusisto, T. A. Lakka, C. Langenberg, L. L. Marchand, T. Lehtimäki, S. Lupoli, P. A. F. Madden, S. Männistö, P. Manunta, A. Marette, T. C. Matise, B. M. Knight, T. Meitinger, F. L. Moll, G. W. Montgomery, A. D. Morris, A. P. Morris, J. C. Murray, M. Nelis, C. Ohlsson, A. J. Oldehinkel, K. K. Ong, W. H. Ouwehand, G. Pasterkamp, A. Peters, P. P. Pramstaller, J. F. Price, L. Qi, O. T. Raitakari, T. Rankinen, D. C. Rao, T. K. Rice, M. Ritchie, I. Rudan, V. Salomaa, N. J. Samani, J. Saramies, M. A. Sarzynski, P. E. H. Schwarz, S. Sebert, P. Sever, A. R. Shuldiner, A. Sinisalo, V. Steinthorsdittir, R. P. Stolk, J.-C. Tardif, A. Tönjes,

A. Tremblay, E. Tremoli, J. Virtamo, M.-C. Vohl; Electronic Medical Records and Genomics (eMEMERGEGE) Consortium; MIGen Consortium; PAGEGE Consortium; Life Lines Cohort Study, P. Amouyel, F. W. Asselbergs, T. L. Assimes, M. Bochud, B. O. Boehm, E. Boerwinkle, E. P. Bottinger, C. Bouchard, S. Cauchi, J. C. Chambers, S. J. Chanock, R. S. Cooper, P. I. W. de Bakker, G. Dedoussis, L. Ferrucci, P. W. Franks, P. Froguel, L. C. Groop, C. A. Haiman, A. Hamsten, M. G. Hayes, J. Hui, D. J. Hunter, K. Hveem, J. W. Jukema, R. C. Kaplan, M. Kivimaki, D. Kuh, M. Laakso, Y. Liu, N. G. Martin, W. März, M. Melbye, S. Moebus, P. B. Munroe, I. Njølstad, B. A. Oostra, C. N. A. Palmer, N. L. Pedersen, M. Perola, L. Pérusse, U. Peters, J. E. Powell, C. Power, T. Quertermous, R. Rauramaa, E. Reinmaa, P. M. Ridker, F. Rivadeneira, J. I. Rotter, T. E. Saaristo, D. Saleheen, D. Schlessinger, P. E. Slagboom, H. Snieder, T. D. Spector, K. Strauch, M. Stumvoll, J. Tuomilehto, M. Uusitupa, P. van der Harst, H. Völzke, M. Walker, N. J. Wareham, H. Watkins, H.-E. Wichmann, J. F. Wilson, P. Zanen, P. Deloukas, I. M. Heid, C. M. Lindgren, K. L. Mohlke, E. K. Speliotes, U. Thorsteinsdottir, I. Barroso, C. S. Fox, K. E. North, D. P. Strachan, J. S. Beckmann, S. I. Berndt, M. Boehnke, I. B. Borecki, M. I. McCarthy, A. Metspalu, K. Stefansson, A. G. Uitterlinden, C. M. van Duijn, L. Franke, C. J. Willer, A. L. Price, G. Lettre, R. J. F. Loos, M. N. Weedon, E. Ingelsson, J. R. O'Connell, G. R. Abecasis, D. I. Chasman, M. E. Goddard, P. M. Visscher, J. N. Hirschhorn, T. M. Frayling, Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat. Genet.* **46**, 1173–1186 (2014).

129. M. Horikoshi, R. N. Beaumont, F. R. Day, N. M. Warrington, M. N. Kooijman, J. Fernandez-Tajes, B. Feenstra, N. R. van Zuydam, K. J. Gaulton, N. Grarup, J. P. Bradfield, D. P. Strachan, R. Li-Gao, T. S. Ahluwalia, E. Kreiner, R. Rueedi, L.-P. Lyytikäinen, D. L. Cousminer, Y. Wu, E. Thiering, C. A. Wang, C. T. Have, J.-J. Hottenga, N. Vilor-Tejedor, P. K. Joshi, E. T. H. Boh, I. Ntalla, N. Pitkänen, A. Mahajan, E. M. van Leeuwen, R. Joro, V. Lagou, M. Nodzenski, L. A. Diver, K. T. Zondervan, M. Bustamante, P. Marques-Vidal, J. M. Mercader, A. J. Bennett, N. Rahmioglu, D. R. Nyholt, R. C. W. Ma, C. H. T. Tam, W. H. Tam; CHARGE Consortium Hematology Working Group, S. K. Ganesh, F. J. A. van Rooij, S. E. Jones, P.-R. Loh, K. S. Ruth, M. A. Tuke, J. Tyrrell, A. R. Wood, H. Yaghootkar, D. M. Scholtens, L. Paternoster, I. Prokopenko, P. Kovacs, M. Atalay, S. M. Willems, K. Panoutsopoulou, X. Wang, L. Carstensen, F. Geller, K. E. Schraut, M. Murcia, C. E. M. van Beijsterveldt, G. Willemsen, E. V. R. Appel, C. E. Fonvig, C. Trier, C. M. T. Tiesler, M. Standl, Z. Kutalik, S. Bonàs-Guarch, D. M. Hougaard, F. Sánchez, D. Torrents, J. Waage, M. V. Hollegaard, H. G. de Haan, F. R. Rosendaal, C. Medina-Gomez, S. M. Ring, G. Hemani, G. M. Mahon, N. R. Robertson, C. J. Groves, C. Langenberg, J. Luan, R. A. Scott, J. H. Zhao, F. D. Mentch, S. M. MacKenzie, R. M. Reynolds; Early Growth Genetics (EGG) Consortium, W. L. Lowe, A. Tönjes, M. Stumvoll, V. Lindi, T. A. Lakka, C. M. van Duijn, W. Kiess, A. Körner, T. I. A. Sørensen, H. Niinikoski, K. Pahkala, O. T. Raitakari, E. Zeggini, G. V. Dedoussis, Y.-Y. Teo, S.-M. Saw, M. Melbye, H. Campbell, J. F. Wilson, M. Vrijheid, E. J. C. N. de Geus, D. I. Boomsma, H. N. Kadarmideen, J.-C. Holm, T. Hansen, S. Sebert, A. T. Hattersley, L. J. Beilin, J. P. Newnham, C. E. Pennell, J. Heinrich, L. S. Adair, J. B. Borja, K. L. Mohlke, J. G. Eriksson, E. Widén, M. Kähönen, J. S. Viikari, T. Lehtimäki, P. Vollenweider, K. Bønnelykke, H. Bisgaard, O. Mook-Kanamori, A. Hofman, F. Rivadeneira, A. G. Uitterlinden, C. Pisinger, O. Pedersen, C. Power, E. Hyppönen, N. J. Wareham, H. Hakonarson, E. Davies, B. R. Walker, V. W. V. Jaddoe, M.-R. Järvelin, S. F. A. Grant, A. A. Vaag, D. A. Lawlor, T. M. Frayling, G. D. Smith, A. P. Morris, K. K. Ong, J. F. Felix, N. J. Timpson, J. R. B. Perry, D. M. Evans, M. I. McCarthy, R. M. Freathy, Genome-wide associations for birth weight and correlations with adult disease. *Nature* **538**, 248–252 (2016).

130. X. Jiang, P. F. O'Reilly, H. Aschard, Y. H. Hsu, J. B. Richards, J. Dupuis, E. Ingelsson, D. Karasik, S. Pilz, D. Berry, B. Kestenbaum, J. Zheng, J. Luan, E. Sofianopoulou, E. A. Streeten, D. Albanes, P. L. Lutsey, L. Yao, W. Tang, M. J. Econs, H. Wallaschofski, H. Völzke, A. Zhou, C. Power, M. I. McCarthy, E. D. Michos, E. Boerwinkle, S. J. Weinstein, N. D. Freedman, W. Y. Huang, N. M. van Schoor, N. van der Velde, L. C. P. G. M. de Groot, A. Enneman, L. A. Cupples, S. L. Booth, R. S. Vasan, C. T. Liu, Y. Zhou, S. Ripatti, C. Ohlsson, L. Vandenput, M. Lorentzon, J. G. Eriksson, M. K. Shea, D. K. Houston, S. B. Kritchevsky, Y. Liu, K. K. Lohman, L. Ferrucci, M. Peacock, C. Gieger, M. Beekman, E. Slagboom, J. Deelen, D. van Heemst, M. E. Kleber, W. März, I. H. de Boer, A. C. Wood, J. I. Rotter, S. S. Rich, C. Robinson-Cohen, M. den Heijer, M. R. Jarvelin, A. Cavadino, P. K. Joshi, J. F. Wilson, C. Hayward, L. Lind, K. Michaëlsson, S. Trompet, M. C. Zillikens, A. G. Uitterlinden, F. Rivadeneira, L. Broer, L. Zgaga, H. Campbell, E. Theodoratou, S. M. Farrington, M. Timofeeva, M. G. Dunlop, A. M. Valdes, E. Tikkanen, T. Lehtimäki, L. P. Lyytikäinen, M. Kähönen, O. T. Raitakari, V. Mikkilä, M. A. Ikram, N. Sattar, J. W. Jukema, N. J. Wareham, C. Langenberg, N. G. Forouhi, T. E. Gundersen, K. T. Khaw, A. S. Butterworth, J. Danesh, T. Spector, T. J. Wang, E. Hyppönen, P. Kraft, D. P. Kiel, Genome-wide association study in 79,366 European-ancestry individuals informs the genetic architecture of 25-hydroxyvitamin D levels. *Nat. Commun.* **9**, 260 (2018).

131. L. B. L. Wittemans, L. A. Lotta, C. Oliver-Williams, I. D. Stewart, P. Surendran, S. Karthikeyan, F. R. Day, A. Koulman, F. Imamura, L. Zeng, J. Erdmann, H. Schunkert, K. T. Khaw, J. L. Griffin, N. G. Forouhi, R. A. Scott, A. M. Wood, S. Burgess, J. M. M. Howson, J. Danesh, N. J. Wareham, A. S. Butterworth, C. Langenberg, Assessing the causal association of glycine with risk of cardio-metabolic diseases. *Nat. Commun.* **10**, 1060 (2019).

132. P. R. Jansen, K. Watanabe, S. Stringer, N. Skene, J. Bryois, A. R. Hammerschlag, C. A. de Leeuw, J. S. Benjamins, A. B. Muñoz-Manchado, M. Nagel, J. E. Savage, H. Tiemeier, T. White; 23andMe Research Team, J. Y. Tung, D. A. Hinds, V. Vacic, X. Wang, P. F. Sullivan, S. van der Sluis, T. J. C. Polderman, A. B. Smit, J. Hjerling-Leffler, E. J. W. Van Someren, D. Posthuma, Genome-wide analysis of insomnia in 1,331,010 individuals identifies new risk loci and functional pathways. *Nat. Genet.* **51**, 394–403 (2019).

133. 1000 Genomes Project Consortium, A. Auton, L. D. Brooks, R. M. Durbin, E. P. Garrison, H. M. Kang, J. O. Korbel, J. L. Marchini, S. M. Carthy, G. A. McVean, G. R. Abecasis, A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).

134. A. Siepel, G. Bejerano, J. S. Pedersen, A. S. Hinrichs, M. Hou, K. Rosenbloom, H. Clawson, J. Spieth, L. W. Hillier, S. Richards, G. M. Weinstock, R. K. Wilson, R. A. Gibbs, W. J. Kent, W. Miller, D. Haussler, Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res.* **15**, 1034–1050 (2005).

135. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).

136. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin; 1000 Genome Project Data Processing Subgroup, The Sequence Alignment/ Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

137. B. Institute, Picard Toolkit, *Broad Institute, GitHub Repository* (2019).

138. F. Ramirez, F. Dundar, S. Diehl, B. A. Gruning, T. Manke, deepTools: A flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* **42**, W187–W191 (2014).

139. A. Dobin, C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, T. R. Gingeras, STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

140. M. A. Deutsch, S. A. Doppler, X. Li, H. Lahm, G. Santamaria, G. Cuda, S. Eichhorn, T. Ratschiller, E. Dzilic, M. Dreßen, A. Eckart, K. Stark, S. Massberg, A. Bartels, C. Rischpler, R. Gilsbach, L. Hein, B. K. Fleischmann, S. M. Wu, R. Lange, M. Krane, Reactivation of the Nkx2.5 cardiac enhancer after myocardial infarction does not presage myogenesis. *Cardiovasc. Res.* **114**, 1098–1114 (2018).

141. J. T. Robinson, H. Thorvaldsdóttir, W. Winckler, M. Guttman, E. S. Lander, G. Getz, J. P. Mesirov, Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).

Citation: J. D. Hocker, O. B. Poirion, F. Zhu, J. Buchanan, K. Zhang, J. Chiou, T.-M. Wang, Q. Zhang, X. Hou, Y. E. Li, Y. Zhang, E. N. Farah, A. Wang, A. D. McCulloch, K. J. Gaulton, B. Ren, N. C. Chi, S. Preissl, Cardiac cell type–specific gene regulatory programs and disease risk association. *Sci. Adv.* **7**, eabf1444 (2021).