Supplemental Material

Expanded Materials and Methods

Ethical statement and human heart tissue.

The Myocardial Applied Genomics Network (MAGNet; www.med.upenn.edu/magnet), collects and banks human cardiac tissue for genomic research. All subjects or next of kin provided written informed consent for tissue donation, and analyses and all study protocols were approved by relevant institutional review boards. Left ventricular (LV) free-wall tissue was harvested at the time of cardiac surgery from subjects with heart failure undergoing transplantation and from unused donor hearts with apparently normal function. Hearts was perfused immediately with cold in-situ high-potassium cardioplegia prior to cardiectomy to arrest contraction and prevent ischemic damage, and tissue specimens were frozen in liquid nitrogen. LV sample metadata are provided in **Online Table I**.

Human cardiac H3K27ac ChIP-seq

For each ChIP-seq library, 70 mg of snap frozen human heart left ventricle tissue was cut and pounded into fine powder in liquid nitrogen. Tissues were washed twice in cold PBS (Takara, 1x protease inhibitor), cross-linked with formaldehyde (1%) for 5 min at room temperature and subsequently quenched with glycine (125mM) for 5 min at room temperature. Cell pellets were rinsed twice with cold PBS. Cells were lysed in lysis buffer (50mM HEPES-KOH, pH 7.5, 150mM NaCl, 1mM EDTA, 1% Triton X, 0.1% Sodium deoxycholate, 0.1% SDS, Takara 1X protease inhibitors) using a glass douncer tight pestle 'A' with 10 to 15 strokes to release nuclei. Centrifuge was performed at 4000 rpm for 10 min at 4ºC to collect the nuclei pellet. Nuclei were checked under microscopy for each preparation. Nuclei were lysed in nuclei lysis buffer (50mM HEPES-KOH, pH 7.5, 150mM NaCl, 1mM EDTA, 1% Triton X, 0.1% Sodium deoxycholate, 1% SDS, Takara, 1x protease inhibitor) and sonicated with Bioruptor sonicator to obtained chromatin fragments between 200 to 500 bp. Sheared chromatin was immunoprecipitated with 5 µg of H3K27ac antibody (catalogue #: ab4729, Abcam) with 50 µl of protein G beads (Invitrogen) overnight at 4ºC. Beads were washed and eluted in 200 µl of elution buffer (50mM Tris-HCI, pH7.5 and 10mM EDTA), de-crosslink at 65ºC overnight. Pulldown was phenol/chloroform treated and DNA purified by ethanol precipitation. Library preparation was performed using 2 ng ChIP DNA with NEB ultra II library preparation kit, according to manufacturer's protocol. 10-12 PCR cycles were performed using indexed primers and libraries size between 300 to 500 bp were selected. The libraries underwent paired end sequencing at 2 x 100 bp read length on Illumina Hiseq 2500. Samples were processed in batches as enumerated in **Online Table I**. FASTQ files for the ChIP-seq libraries are deposited in EBI (EGA Study ID: EGAS00001003586). ChIP-seq read signal in the bigwig format has also been deposited onto UCSC Genome Browser (link: https://genome.ucsc.edu/s/wlwtan/EGAS00001003586). For comparative analysis, we downloaded H3K27ac ChIP-seq data (Experiment Accession: ENCSR557DFM) and H3K4me1 ChIP-seq data (Experiment Accession: ENCFF072ZDB) from ENCODE. We also downloaded human cardiomyocyte H3K27ac and H3K4me1 ChIP-seq data from NCBI Bioproject PRJNA353755.

Read Alignment and Peak Calling

Paired-end sequencing reads were aligned to human genome (*hg19*) genome using BWA mem version 0.7.5 [51]. Duplicated reads were filtered out using Picard MarkDuplicates function. The mapped reads were fed into peak-caller, Dfilter [17], to identify significant peaks using the following setting: -ks=60, -bs=100, -lpvalue=8. Each of the ChIP-seq biological replicates were assessed for their Duplication Rate, Fraction of Reads in Peaks (FRiP), average correlation with other replicates and SSD (**Online Table II**). Peaks detected in at least 4 individuals were merged across samples (overlap>4) to define a consensus set of 47,321 peaks. All subsequent analyses were performed based on this 47,321 consensus peak set (**Online Table III**).

Peak Height Normalization

Human genome was divided into 100 bp bin. For each library, reads were counted in 100 bp bins, and scaled to normalize for sequencing depth. Binned counts were then adjusted by normalizing their GC-content against the average GC-content of all libraries. In each peak region, the sum of bin-wise

normalized counts was defined as the peak height. In order to reduce technical variation, the peak heights in the consensus peak set were quantile-normalized [52].

Removal of Confounding Factors

Firstly, the normalized peak heights were $\log_2$-transformed. Principal-component analysis (PCA) was performed to identify potential confounding factors by correlating the top 5 principal components (PCs) with the biological covariates (etiology, conditions such as atrial fibrillation, hypertension, diabetes and ventricular tachycardia/ventricular fibrillation, age, gender, height, weight, cardiomyocyte purity) and technical covariates (sequencing batches, median library fragment insert size from paired-end reads, proportion of duplicated reads, sequencing depth and number of peaks per ChIP-seq library) (**Online Figure IIB, C**). Cardiomyocyte purity was estimated using FAC-sorting (see below). For differential acetylation (DA) analysis between end-stage heart failure and non-failing hearts, covariates that were regressed out were age, gender, height, weight, cardiomyocyte purity, and technical covariates (**Online Figure IIB**). For identifying haQTL, we regressed out all technical and biological covariates, crucially including etiology, conditions such as atrial fibrillation, hypertension, diabetes and ventricular tachycardia/ventricular fibrillation (**Online Figure IIC**). PCA was performed again after regression to confirm that no confounding factors correlated strongly with the top 5 PCs. G-SCI test analysis to identify haQTL were based on the peak height matrix after covariate regression.

Differentially Acetylation (DA) analysis between Heart Failure (HF) and control non-failing hearts (NF)

Using the normalized peak height matrix with confounding factors removed, DA peaks (**Online Table IV**) were identified with the following criteria (fold-change > 1.3; FDR < 0.05; Wilcoxon rank sum test; Benjamini-Hochberg correction). Differentially hyper-acetylated and hypo-acetylated, in bed format, were supplied to the GREAT analysis tool for pathway analysis separately using the default parameters (FDR < 0.05). For motif enrichment analysis, we used the HOMER findMotifsGenome.pl script [53]. Motif models were drawn from the TRANSFAC vertebrate database [54] and the analysis was performed separately on increased and decreased DA peaks.

SNP-Calling Pipeline

ChIP-seq reads were passed to the multi-sample SNP-calling pipeline. Reads used for SNP calling were de-duplicated and retained only if they were mapped to the genome in the proper orientation. We performed indel realignment, base-quality-score recalibration and SNP calling using GATK version 2.7-2 [55]. 576,482 SNPs within peaks were called using GATK UnifiedGenotyper at a SNP quality threshold of 50. These SNP calls were filtered out with the following criteria: MQ0Fraction > 0.001, QD < 4.3, within 6 bp of an indel, more than seven SNPs within a 100-bp region, Mapping Quality < 45, Homopolymer Run > 10, MQ0 > 9.5, Dels > 0.255. Moreover, only SNP calls covered by at least 5 non-reference reads across all libraries and 3 or more non-reference reads in at least one library were retained. SNPs that violated Hardy-Weinberg equilibrium with a binomial test P-value ($<1 \times 10^{-3}$) were removed as well. To eliminate mapping artefacts, SNPs in highly paralogous regions of the genome implicated by the ''Self Chain'' track on the UCSC Genome Browser (normalized score ≥ 90) were filtered out. Finally, a high-confidence set of 249,732 SNPs within H3K27ac peaks were obtained. Novel SNPs were identified by comparing the SNP set with gnomAD and dbSNP, we did not perform genotype calling since the G-SCI test does not require prior knowledge of genotype, instead it integrates over the likelihoods of all three genotypes for each individual, given the data. Six samples were discarded when we discovered that they were incorrectly labelled as "European" using LASER **(Online Figure XV1)** [56]. The remainder 64 samples were used for subsequent haQTL analysis.

Identification of haQTL (G-SCI test)

haQTLs were called in the 64 samples using G-SCI test. This was done after all technical and biological covariates, importantly, disease etiology were regressed out. Top PCs which account for more than 5% variance were thus adjusted for. G-SCI test was performed on each of the 249,732 SNPs within peaks. For each SNP, an adjusted *P*-value was computed using a permutation test from 10,000 to 1 million permutations until a nonzero *P*-value was obtained. After 1 million permutations, if the adjusted *P*-value was still 0, it was set to 5×10^{-7}. We then used the Benjamini and Hochberg multiple testing

correction to calculate the FDR. At FDR threshold of 10%, 12,006 candidate haQTLs were identified. To detect possible artificial haQTLs due to different mapping rates to the reference genome between alleles, we simulated all possible 100 bp paired-end reads covering the haQTL and flanking SNPs and indels. The union of our SNP and indel (quality > 50 by GATK) calls and the 1000 Genome EUR SNPs and indels (1000 Genomes Project Consortium, 2012) were used. The fragment length of the simulated paired-end reads was set to be equal to 180 which is the median fragment size of all libraries. The simulated reads were then mapped to the reference genome using BWA. 4,231 haQTLs were discarded because their inferred allelic imbalances from the ChIP-seq data were smaller than five times the mapping bias estimated from the simulation. The remaining haQTLs were further filtered by an effect-size filter which calculated the Pearson correlation between peak height and the fraction of Q30 nonreference bases. haQTLs with $R^2 < 0.1$ were discarded. The final set of 1,680 haQTLs were in the remaining haQTLs after effect-size filter and only the most significant SNP (sentinel SNP) in each ChIP-seq peak was retained.

PCM-1+ cardiomyocyte marker FAC sorting to determine cardiomyocyte cell purity

50 mg of snap frozen human heart left ventricle tissue was cut and pounded into fine powder in liquid nitrogen. Tissues were rinsed in cold PBS (Takara, 1X protease inhibitor) and lyse in 10 ml cell lysis buffer (50mM HEPES-KOH, pH 7.5, 150mM NaCl, 1mM EDTA, 1% Triton X, 0.1% Sodium deoxycholate, 0.1% SDS, Takara 1X protease inhibitors) using a glass douncer with tight pestle 'A' (10-15 strokes). Crude nuclei were filtered through 70 µM and 40 µM cell strainers. Filtered nuclei were subjected to overnight incubation with cardiomyocyte-specific anti-PCM1 antibody (catalogue #: HPA023374, Sigma Aldrich) at 1:1000 dilution. Nuclei were secondary stained with Goat anti-rabbit Alexa Fluor 488 secondary antibody (catalogue #: R37116, Thermofisher scientific) at 1:500 dilution for 30 min. Nuclei were further stained with DAPI (catalogue #: D9542, Thermofisher scientific) at 1:1000 dilution for 5 min before proceeding for FAC sorting on the BD biosciences Aria II sorter.

RNA-seq library construction, sequencing and bioinformatics analysis

Total RNA was extracted using the miRNeasy Kit (Qiagen) including DNase treatment. RNA concentration and quality were determined using the NanoVue Plus™ spectrophotometer (GE Healthcare) and the Agilent 2100 RNA Nano Chip (Agilent). RNA sequencing libraries were prepared using the Illumina TruSeq stranded mRNA kit followed by the Nugen Ovation amplification kit. To avoid confounding by batch effects, libraries were randomly selected into pools of 32, and pools were sequenced on a Hiseq 2500 to a depth of ~30 million 100 bp paired-end reads per biological sample (**Online Table II**). Fastq files were aligned against human reference (hg19/hGRC37) using the STAR aligner [57]. Duplicate reads were removed using MarkDuplicates from Picard tools, and per gene read counts for Ensembl (v75) gene annotations were computed. Expression levels in counts per million (CPM) were normalized and transformed using VOOM in the LIMMA R package [58].

Expression-QTL (eQTL) analysis

eQTL analysis with our LV samples was performed with linear regression analysis. The FPKM expression level was corrected for technical confounding factors including batch, RIN, library size and fragment length, as well as biological factors such as etiology-related traits, sex, height and weight. PCA was performed again after regression to confirm that no confounding factors correlated strongly with the top 5 PCs. We considered cis-eQTL interactions with less than 100 kb separating the SNP and the TSS. We retained 180 eQTL at FDR ≤ 20%. For distal-eQTL, we considered haQTL-gene pairs that were confined within the same TAD, and having minimal distance of 10 kb between haQTL and TSS. In order to refine the search for distal eQTL, only haQTL-gene pairs connected by HiChIP chromatin loops were considered.

Human cardiac Hi-C

In situ Hi-C was performed as previously described [59]. 70 mg of snap frozen human heart left ventricle tissue was cut and pounded into fine powder in liquid nitrogen. Tissues were washed twice in cold PBS, crosslinked in 1.5% formaldehyde for 10 minutes at room temperature after which glycine was added to stop the reaction. Cross-linked cells were lysed in cold lysis buffer (10 mM Tris-HCl pH8.0, 10 mM NaCl, 0.2% (v/v) Igepal CA630, mixed with protease inhibitors) on placed ice for 30 minutes.

Subsequently, we applied 15 strokes using a glass douncer tight pestle 'A' to release nuclei. Pellets were washed twice with cold 1x NEBuffer 2, resuspended in 50 µl of 0.5% SDS and incubated at 62ºC for 10 minutes to solubilize the chromatin. After heating, 120 µl of water and 50 µl of 5% Triton X-100 (catalogue #: 93443, New England Biolabs) was added to quench the SDS. Chromatin was digested overnight by adding 25 µl of 10X NEBuffer 2 and 100U MboI (catalogue #: R0147, New England Biolabs). Digestion efficiency was confirmed by isolating DNA from a 25 µl aliquot of the suspension and performing agarose gel electrophoresis. Digested ends were filled and labelled with biotin by adding 37.5 µl of 0.4 mM Biotin-14-dATP (catalogue #: 19524016, Thermo Fisher Scientific), 1.5 µl of 10mM dCTP, 1.5 µl of 10 mM dGTP, 1.5 µl of 10 mM dTTP, 8 µl of 5U/ µl DNA polymerase I, large Klenow fragment (catalogue #: M0210, New England Biolabs) and incubated at 37ºC for 90 minutes. Blunt end ligation was performed by adding 463 µl of water, 120 µl of 10X NEB T4 DNA ligase buffer, 200 µl of 5% Triton X-100, 12 µl of 10 mg/ml BSA and 5 µl of T4 DNA Ligase (catalogue #:M0202, New England Biolabs,), the solution was incubated at 16ºC for 4 hours. After ligation, crosslink was reversed overnight and DNA was purified using Phenol-chloroform method. Next, Hi-C libraries were treated with T4 DNA Polymerase (New England Biolabs, M0203) to remove biotin from unligated ends. To do this, 5 µg of Hi-C library was combined with 2.5 µl of 1 mM dATP, 2.5 µl of 1 mM dGTP, 10 µl of 10X NEBuffer 2.1, 5 µl of T4 DNA Polymerase (catalogue #: M0203, New England Biolabs) and water was added to a final volume of 100 µl. The reactions were incubated at 20ºC for 4 hours and stopped by incubating at 75ºC for 20 minutes. Samples were sheared to a mean fragment length of 400 bp using Covaris M220. Pull-down of biotinylated DNA was performed using Dynabeads My One T1 Streptavidin beads (catalogue #: 65601, ThermoFisher Scientific), and library prep was performed following NEBNext® Ultra™ II DNA Library Prep Kit for Illumina®. Samples were quantified using KAPA and sequencing performed on a Hiseq 4000.

Hi-C Analysis

All sequencing QC are described in **Online Table II**. FASTQ sequences were mapped and filtered using HiCUP pipeline [60]. Hi-C reads (paired end, 2x151 bp) were aligned against the hg19 genome using Bowtie2. Only unique high-quality alignments were retained. Forward and reverse reads were mapped independently. Sequences which were mapped unambiguously to the genome were retained. Di-tags representative of experimental artefact were filtered out as these invalid di-tags might lead to inaccurate inference concerning genomic structure. These invalid di-tags were detected by positioning the putative di-tags on an in silico digested reference genome. Sequences which were mapped to the same restriction fragment were discarded as they did not describe the multi-dimensional contacts. Di-tags insert sizes which did not fit within the size-selection range were also discard. Lastly, PCR duplicates were removed. Cleaned-up aligned sequenced files in BAM format were processed by HiCPro [61] to generate raw interaction count at various bins of equal sizes (10 kb, 40 kb). Contact matrices with raw interaction counts were normalized using iterative correction of biased based on ICE algorithm [62]. TAD-calling was performed using Domaincaller [30], based on the ICE-normalized count matrix at 40 kb resolution. Insulation-scores were calculated as described in Crane *et al* [63], using a window-size of 10 kb (**Online Figure VIIE**). Fit-hi-C [30] was also used to identify significant chromatin interaction (FDR<0.05) at 10 kb resolution (**Online Figure VIIF**).

Human cardiac HiChIP

HiChIP was adapted from a previously published protocol [64] with some modifications. Briefly, 50 mg of snap frozen human heart left ventricle tissue was cut and pulverized in liquid nitrogen using mortar and pestle. Tissues were washed twice in cold PBS, and Hi-C portion performed following the Arima-HIC protocol described in the Arima Hi-C Kit (Product No A510008). For the ChIP portion, the nuclear pellet was sonicated to obtain chromatin fragments between 200 to 500 bp. Sheared chromatin was immunoprecipitated with 3 µg of H3K27ac antibody (Abcam, ab4729) with 30 µl of protein G beads (Invitrogen) overnight at 4ºC. Beads were washed and eluted in 200 µl of elution buffer (50 mM sodium bicarbonate, 1% SDS), de-crosslinked at 67ºC for 2 hours and DNA extracted using the Qiagen MinElute kit. Library preparation was performed using NEB ultra II library preparation kit, according to manufacturer's protocol. 10-12 PCR cycles were performed using indexed primers and libraries size between 300 to 500 bp were selected. The libraries underwent paired end sequencing at 2 x 151 bp read length on Illumina HiSeq 4000.

HiChIP Analysis

A similar analysis workflow from Hi-C above was applied to HiChIP. FASTQ sequences were mapped and filtered using HiCUP pipeline [60]. Filtered sequenced files in BAM format were analysed by HiCPro to generate raw interaction count. Contact matrices at 10 kb resolution were normalized using iterative correction of biased based on ICE algorithm [62]. Filtered sequenced files in BAM format were also converted into .hic files using pre function in juicer package [65]. High confidence chromatin loops were identified using the Juicer pipeline HiCCUPS tool at 5 kb and 10 kb resolution using the default parameters: hiccups –m 500 –r 5000,10000 –f 0.1,0.1, -p 4,2 –i 7,5, -d 20000,20000 <.hic> <output> [59]. Biological and technical reproducibility assessment between HiChIP samples were performed by comparing the total number of di-tags supporting each chromatin loop called from the merged replicates (**Online Figure IXA**). The Pearson correlation between HiChIP replicates was computed from ICE-normalized count using the R cor () function.

Transcription Factor (TF) Motif Alteration Prediction

R package MotifBreakR and haploREG was used to identify transcription factor motifs that overlapped with the 1680 haQTLs using motif databases from Jaspar 2018 [66] using the default parameters: filterp=TRUE, pwmList=jaspar2018, threshold=1e-4, method="ic". The motifs of transcription factor binding were predicted to be significantly altered are listed in the **Online Table X**. Only TFs that were found to be expressed in the human RNA-seq samples were used for the analysis (FPKM>=1).

Intersection against published GWAS datasets

A total of 10 sets of heart-related GWAS summary statistics were downloaded. LD (r^2>0.8; EUR population from 1000 Genomes project) between QTLs and GWAS SNPs was calculated. LD was calculated using plink software [67]. The analysis for enrichment of GWAS SNPs in haQTL was performed as elaborated below. As negative control, another 4 sets of non-heart-related GWAS [68-71] were used for overlap to show non-significant enrichment (**Online Figure XIII**).

Enrichment of GWAS SNPs in haQTL (Permutation Test)

First, we estimated the number of haQTLs that overlapped/in LD (r^2>0.8) with SNPs in the GWAS datasets. Let this number be *n*. Next, we derived a simulation SNP pool, *S* by subsetting SNPs in the GWAS datasets that are overlapping or LD (r^2>0.8) with our entire SNP call set in enhancer regions. We excluded SNPs in GWAS datasets that were not in LD or not overlapping with our enhancer SNP callset. We also excluded haQTL SNPs from the simulator datasets. *n* number of SNPs were randomly drawn from the simulation SNP pool, *S*, matching to the MAF, number of TSS within 100kb, and distance to nearest TSS of n haQTL SNPs. 1,000 simulator SNP sets were generated. For each simulator SNP sets, total number of random SNPs in LD or overlapping with the GWAS subthreshold SNPs were counted (expected overlapping frequency). Let *x* be the observed number of haQTL that are in LD/overlapping with GWAS subthreshold SNPs. Let y be the expected number of overlaps/LD between each simulator sets and GWAS subthreshold SNPs. Fold-enrichment was derived by taking the ratio between observed number of haQTL that are in LD/overlapping with GWAS subthreshold SNPs over the expected number of overlaps/LD between each simulator set and GWAS subthreshold SNPs. Enrichment p-value was calculated as the simulator sets showing higher or equal number of SNPs in LD/overlapping with GWAS subthreshold SNPs as haQTLs (**Online Table XI**).

Human embryonic stem cell derived cardiomyocyte (hES-CM) differentiation

For feeder-free culture, H1 Human embryonic stem-cell (H1-hESC) line was maintained on Geltrex (catalogue #: A1413202, ThermoFisher Scientific) in mTeSR1 medium (STEMCELL Technologies). Cardiomyocyte differentiation protocol was optimized as previously described [72]. H1-hESCs were dissociated into single cells with Accutase (Invitrogen) in 37 °C for 5 minutes and seeded 600,000 cells onto Geltrex-coated plates in a 12-well format (150,000 cell/cm^2) with 5 μM ROCK inhibitor Y27632 (STEMCELL Technologies). At 80-85% cell confluency, cells were treated with 10uM CHIR99021 (catalogue #: 72054, STEMCELL Technologies) in RPMI/B27 without insulin (Gibco) and fresh media was replaced after 24 hours. At Day 3, 5 μM IWP2 (catalogue #: I0536, Sigma Aldrich) was added into

conditioned media and fresh media (RPMI/B27-insulin) at 1:1 ratio and removed during medium change on day 5. Cells were maintained in RPMI/B27 with insulin (catalogue #: 17504044, Gibco) from day 7 onwards and refreshed every 3 days.

Luciferase Enhancer Reporter Assays

10 haQTL SNP loci were randomly selected for validation based on haQTL candidates that overlapped with heart-related GWAS (**Online Table XI**). Allele-specific enhancer oligos (bearing either reference or alternative SNVs) were synthesised using gBlocks Gene Fragment service from Integrated DNA Technologies IDT. gBlocks consisted of 183 bp of the enhancer sequences centred around the sentinel haQTL, and flanking primer sequences (18 bp and 19 bp respectively), making a total of 220 bp. Primer sequences allowed for PCR amplification and cloning into the Firefly luciferase backbone, and are as listed in **Online Table XIII**. Amplified PCR products were cloned into the *SalI* restriction enzyme site of the pGL4.13 (Promega, catalogue #: E6681), and Sanger sequencing was performed to verify accurate insertion into the plasmid construct.

96-well tissue culture treated plates were coated with 0.1 mL of 0.1% (w/v) gelatin per well and incubated at 37°C for at least 1 hour, after which the gelatin was aspirated. Human iCMs (Cellular Dynamics, catalogue #: R1117) were thawed according to manufacturer's instructions and the cells were plated at a density of 12,000 cells per well of 96-well plates at a volume 100 µL of plating medium. 48 hours post-seeding, iCM plating medium was replaced with 100 µL of Maintenance Medium. 4 days post plating, cells were transfected with respective luciferase reporter constructs. For each well, 95 ng of enhancer firefly Luciferase reporter (pGL4.13, Promega) and 5 ng of Renilla Luciferase transfection control vector (pGL4.73, Promega) were mixed with 10 µL of Opti-MEM (Life Technologies, catalogue #: 51985-034). 0.6 µL of Viafect transfection reagent (Catalogue #: E4981, Promega) was added to the DNA/Opti-MEM mixture. After mixing, the transfection cocktail was incubated at room temperature for 15 min and subsequently 10 µL of the mixture was dispensed into each well with iCMs and plates were transferred to 37°C incubator. For each construct, 3 wells were transfected in each experimental replicate, and independent experiments were performed another day to a total of 6 wells per construct. Luciferase activity was measured 4 days after transfections using the Dual-Luciferase Reporter Assay System (catalogue #: E1980, Promega). Luminescence was measured according to manufacturer's assay protocol for 96-well plates using the Promega GloMax®-Multi detection system. Firefly luciferase activity was normalized to Renilla luciferase activity to control for transfection efficiency in each well. For each enhancer, we collated readings from different experimental replicates by normalizing the activity of reporters to the reference allele. Statistical test was performed using Student's t test.

Electrophoretic Mobility Shift Assay (EMSA)

DNA probes were generated by primer extension using PrimeSTAR MAX polymerase (Clontech) and then purified by gel extraction kit (Qiagen). 10 pmol DNA probe was 5' end labelled with γ-32P-ATP (Perkin Elmer) using T4 PNK (New England Biolabs) followed by G25 column purification (GE healthcare). For each EMSA reaction, 0.1 pmol labelled DNA probe was incubated with 500 ng recombinant MEF2A protein (catalogue #: 204772, Abcam) in 1x EMSA binding buffer (10 mM Tris, 50mM KCl, 5mM MgCl2, 0.05% NP-40, 2.5% glycerol,1mM EDTA and 1mM DTT) supplemented with 1 µg/µl Poly(dI•dC) (Invitrogen) at room temperature for 30 minutes. For competitive binding assay, 10 pmol unlabelled probe was added to binding mixture. Binding mixture was separated by 5% native TBE-PAGE, followed by gel drying and gel exposure to BioMax® MS film (Carestream Kodak). Sequences of probes are listed in **Online Table XIII**.

Lentivirus Generation and Purification Protocol

Lentiviral particles were produced in HEK293T cells cultured in DMEM+5% FBS according to standard protocol. Briefly, 10 µg of transfer plasmid, 22.5 µg of pMDLg/pRRE, 7.5 µg of pRSV-Rev and 7.5 µg pMD2.G (Addgene #12251, #12253 & #12259), were transfected in one 15 cm^2 dish with 150 µl of PEI. Media was refreshed after approximately 12 h with DMEM+5% FBS. Cells were incubated and supernatant was collected and refreshed after a further 24 h, and collected again at 48 h. Viral particles were concentrated using Lenti-Pac™ Lentivirus Concentration Solution (catalogue #:

LPR-LCS-01, GeneCopoeia™) according to manufacturer's instructions. 15 minutes prior to transduction, hES-CM and hES-EC were pre-treated with 8 μg/ml polybrene for increased transduction efficiency.

hES-CM CRISPR-targeted excision and gene knockdown

Targeting sgRNAs were designed by identifying NGG PAM sites flanking the haQTL candidate at the distal enhancer of *NOS1AP* and *C1orf226* (chr1:162033295-162034795). The bioinformatic tool at http://crispor.tefor.net/ was used as previously described [73]. sgRNA constructs were cloned into the *Esp3I*-digested LentiCRISPRv2 plasmid backbone (Addgene #52961) using T4 DNA ligase (catalogue #: M0202, New England Biolabs, NEB), according to manufacturer's instructions. Vectors were transformed into RapidTrans™ TAM1 Competent E. coli (catalogue #: 11096, Active Motif). Plasmid DNA was extracted by Monarch® Plasmid Miniprep Kit (NEB) and successful cloning confirmed by Sanger sequencing. Each vector DNA was packaged into individual lentivirus using the protocol described above. For all control groups, H1-hESC derived cardiomyocytes were transduced with lentivirus carrying the non-targeting control (NTC) sgRNAs: pLentiCRISPRv2-Scramble1 (sgRNA sequence: 5'-AAAACAGGACGATGTGCGGC-3') and/or pLentiCRISPRv2-Scramble2 (sgRNA sequence: 5'-AACGTGCTGACGATGCGGGC-3'). For the excision of *NOS1AP/C1ORF226* distal enhancer (chr1:162033295-162034795), lentiviral pools of independent pairs of guides each were co-transduced at high multiplicity of infection (MOI). Two days post-infection, cells were selected with 3 μg/ml Puromycin (Sigma catalogue #: P9620) and sub-cultured for a further 10 days. Efficient CRISPR-targeted excision was verified by PCR amplification, gel electrophoresis and Sanger sequencing. The sequences of sgRNAs used are listed in **Online Table XIII**.

RNA extraction and RT-qPCR

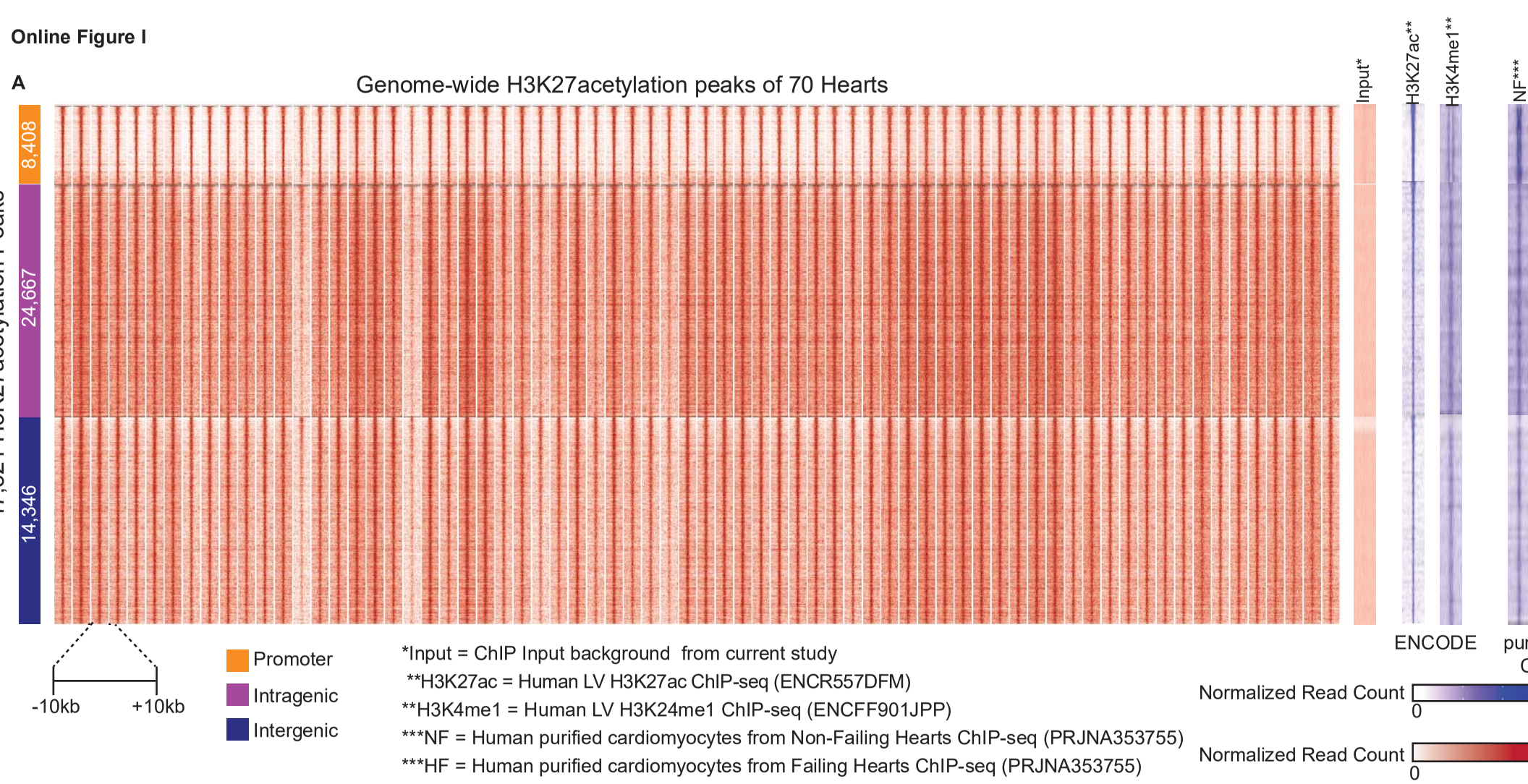
Total RNA was extracted from cardiomyocytes using the RNeasy Mini Kit (Zymo Research), and reverse transcribed using qScript cDNA Supermix (catalogue #: 95049-100, Quantabio), according to manufacturer's instructions. RT-qPCR were performed using PerfeCTa SYBR Green FastMix Low ROX (catalogue #: 101414-288, Quantabio) ABI 7500 Real-Time PCR system (Applied Biosystems), and primers used are listed in **Online Table XIII**. The thermocycling conditions is as follows:

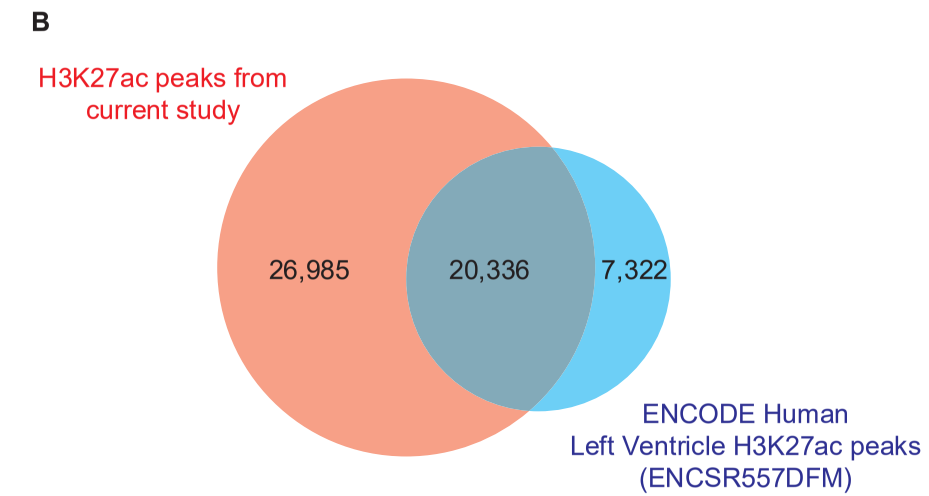
Stage	Temperature	Duration
Initial denaturation	95 °C	30 seconds
PCR cycling (40 cycles)	95 °C	5 seconds
	60 °C	15 seconds
	70 °C	10 seconds
Melt Curve Stage	95 °C	15 seconds
	60 °C	1 minute
	95 °C	15 seconds

Statistics and Software

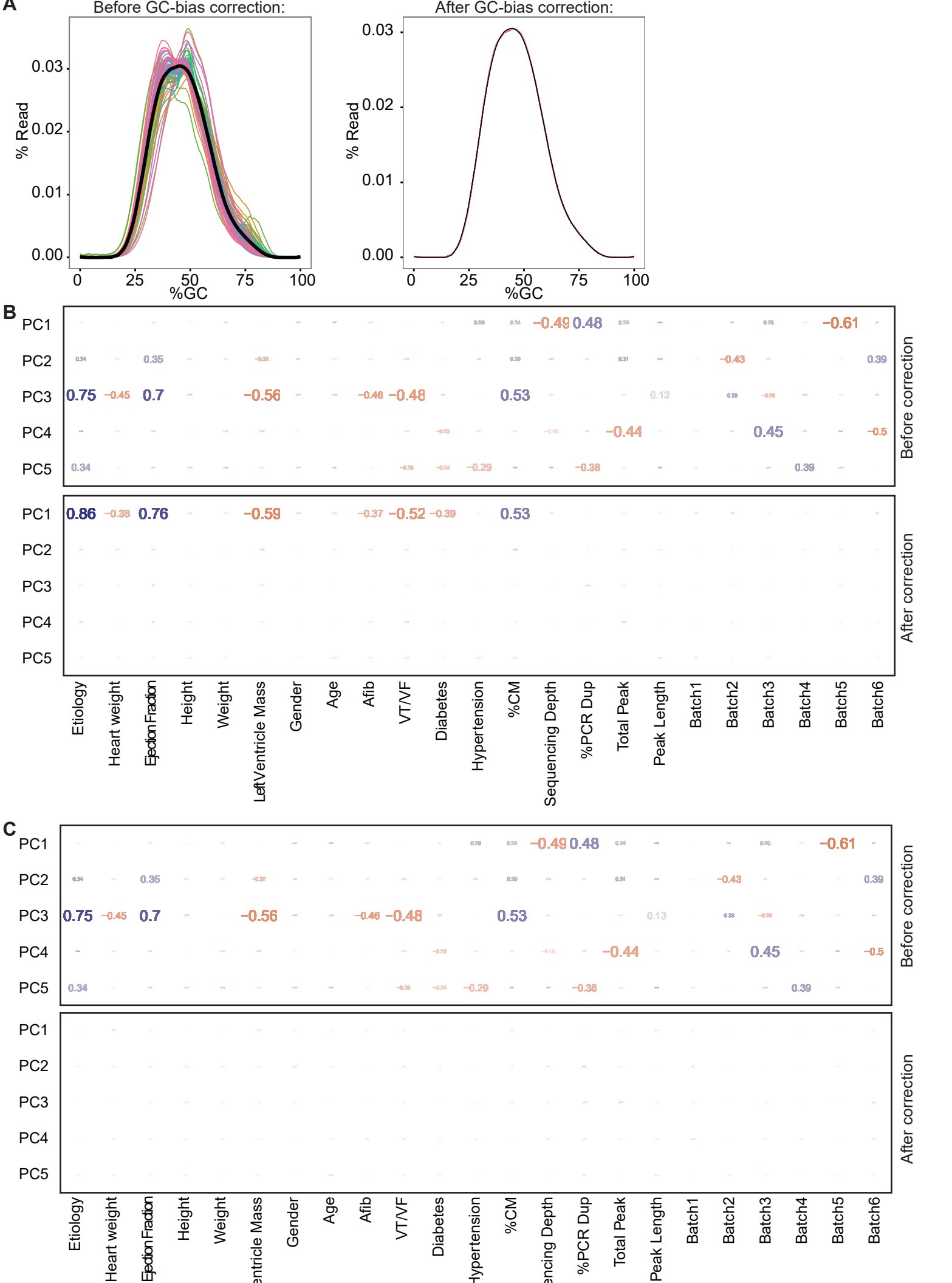
For RT-qPCR and Luciferase Assay analysis, the data are represented as the mean ± SD. Statistical test were performed by two-tailed Student's t-test using Prism v8.0 software. Normal distribution of the data was assessed on GraphPad prism using the Shapiro-Wilk normality test. All bioinformatics statistical analyses were performed in R (version 3.5).

Online Figure I



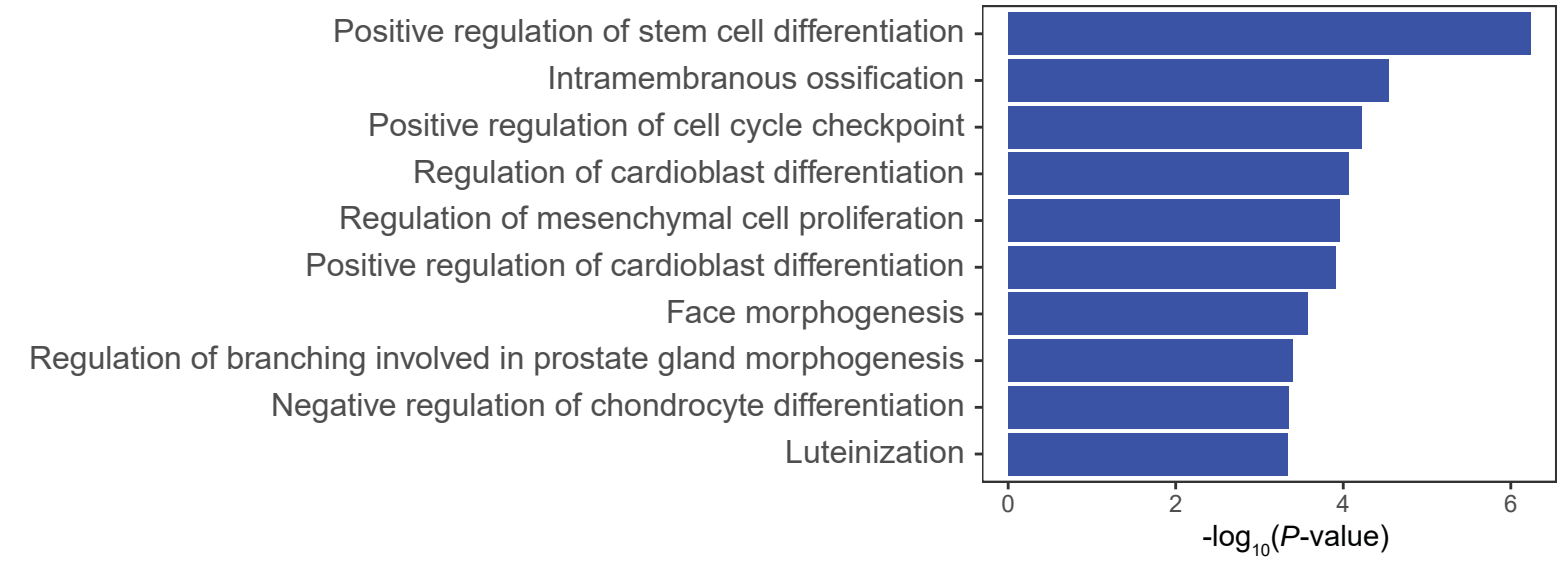


Online Figure I. H3K27ac enhancer profiles of human left ventricles (LV).
(A) Heatmaps (red) of 47,321 consensus H3K27ac ChIP-seq peaks are consistent and reproducible across 70 LV. Heatmap (blue) of H3K27ac and H3K4me1 ChIP-seq in human LV from ENCODE, and H3K27ac ChIP-seq using purified human CM for the same loci as displayed in our samples. Datasets are taken from deposited ID as indicated. Each row represents H3K27ac enrichment at an enhancer locus, centred at the H3K27ac peak and extended ±10 kb.
(B) Venn diagram showing the relationship between H3K27ac enhancer loci found in the current study, overlapped with those in the ENCODE H3K27ac ChIP-seq (Experiment Accession: ENCSR557DFM).

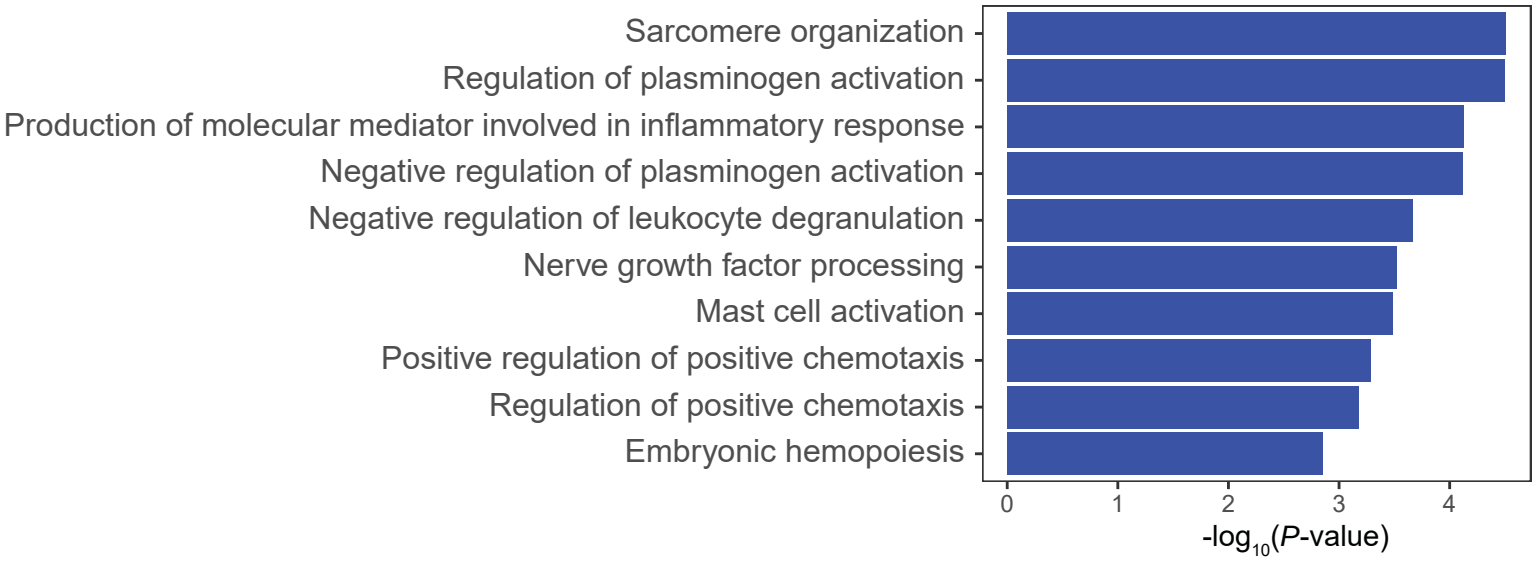

A
Before GC-bias correction:
After GC-bias correction:
% Read
%GC
0.03
0.02
0.01
0.00
0
25
50
75
100
B
C
PC1
PC2
PC3
PC4
PC5
Before correction
After correction
Etiology
Heart weight
Ejection Fraction
Height
Weight
Left Ventricle Mass
Gender
Age
Afib
VT/VF
Diabetes
Hypertension
%CM
Sequencing Depth
%PCR Dup
Total Peak
Peak Length
Batch1
Batch2
Batch3
Batch4
Batch5
Batch6

Online Figure II. Normalization and confounding factor correction of H3K27ac ChIP-seq.
(A) QC metrics showing GC correction (left: before; right: after) applied to H3K27ac peaks.
(B) Plots showing confounding biological and technical variables before (above) and after (below) regression for DA analysis.
(C) Plots showing confounding biological and technical variables before (above) and after (below) regression for haQTL analysis. In the case of the latter, every technical and biological confounding factor was corrected, especially including etiology-related ones (etiology, heart weight, ejection fraction, atrial fibrillation status, VT/VF, diabetes status, and hypertension status) in order to remove the disease effect.

GREAT Analysis of Hyper-acetylated H3K27ac Enriched Regions

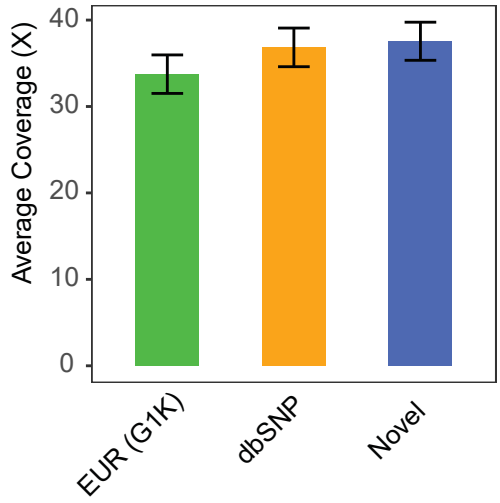


GREAT Analysis of Hypo-acetylated H3K27ac Enriched Regions



Online Figure III. Pathway analysis of Differential Acetylation Peaks. GREAT pathway analysis was performed using the differentially acetylation peaks. Bar plots show the top 10 most enriched pathways in upregulated and downregulated peaks, respectively.

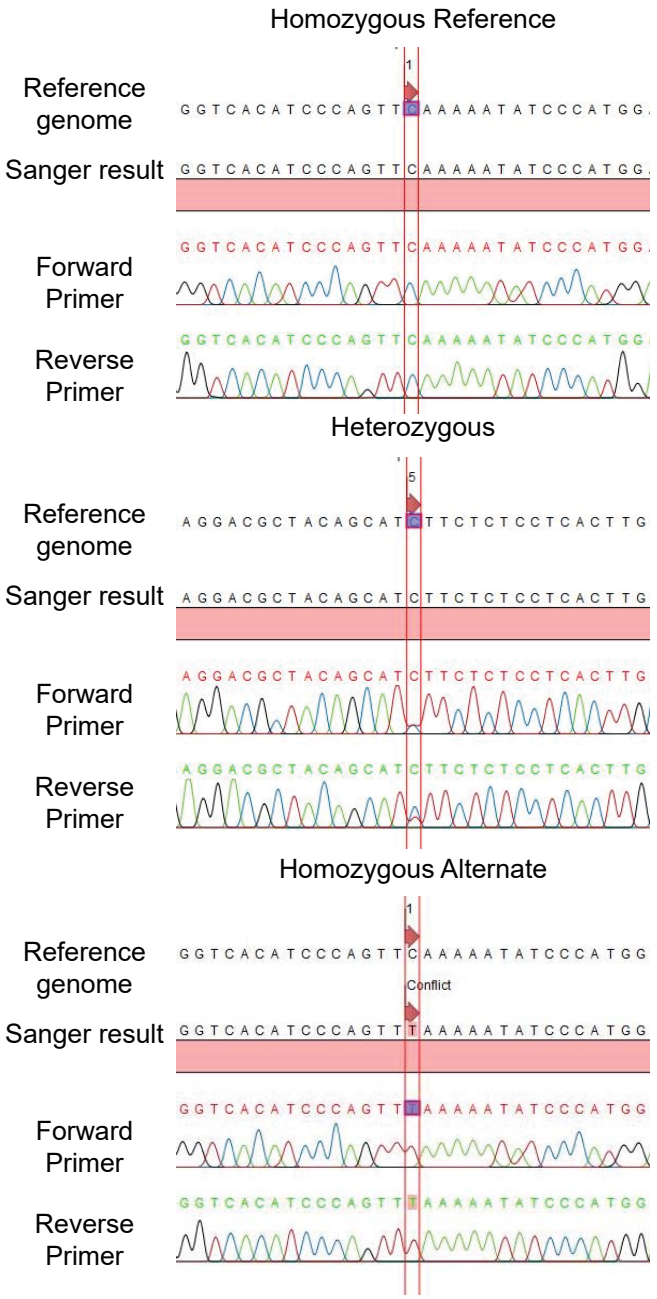
A



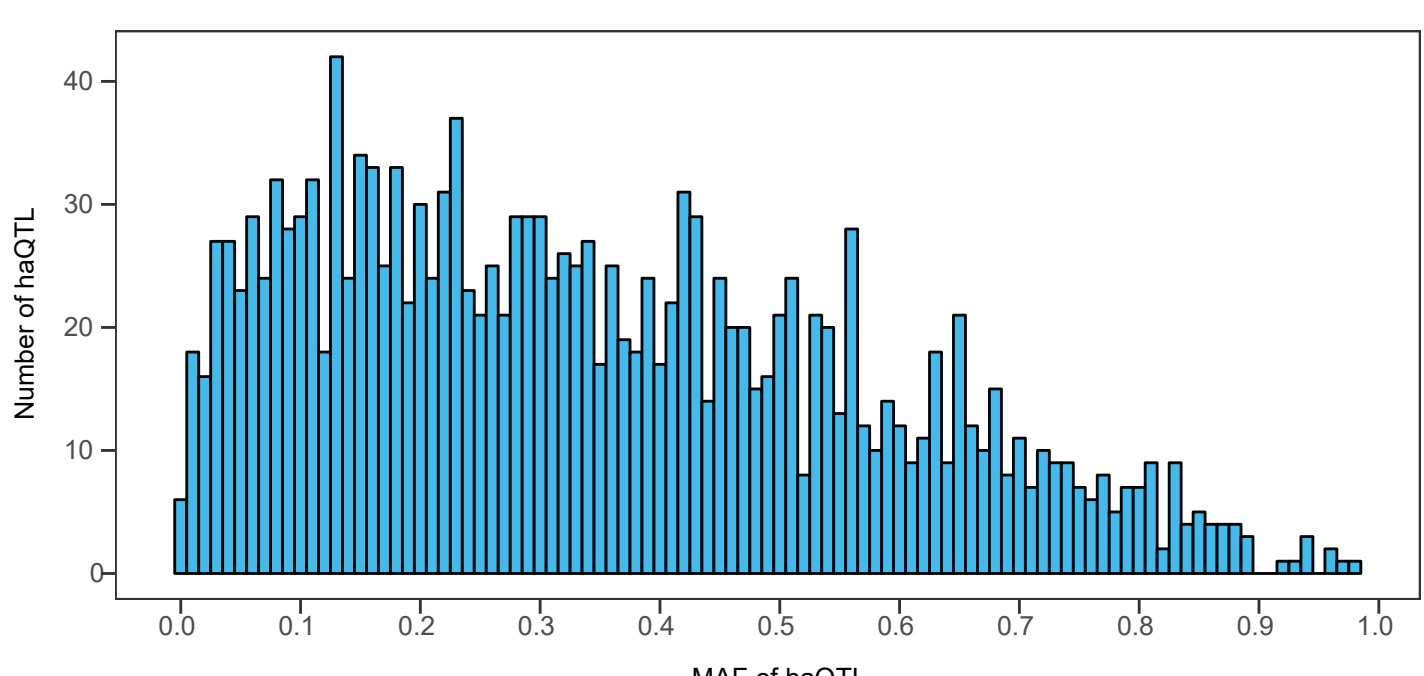
B

SNP locus	LV Sample ID		
	Homozygous Ref	Homozygous Alt	Heterozygous
chr1-162033890 C>T	1185	1252	1122
	1296	1472	1610
chr1-208135982 T>A	1252	1122	1184
	1185	1217	1309
chr2-36683708 G>A	1381	1247	1122
	1406	1330	1217
chr3-53293986 G>A	1122	1279	1217
	1247	1381	1330
chr12-82153090 C>T	1247	1309	1184
	1252	1622	1185
chr12-115358952 T>G	1275	1184	1247
	1622	1266	1472
chr13-74518210 G>A	1381	1275	1252
	1528	1406	1254
chr14-64680848 A>G	1264	1254	1275
	1330	1279	1381
chr16-11692198 G>T	1406	1264	1275
	1296	1330	1279
chr16-15906130 A>G	1184	1254	1264
	1279	1309	1406

C

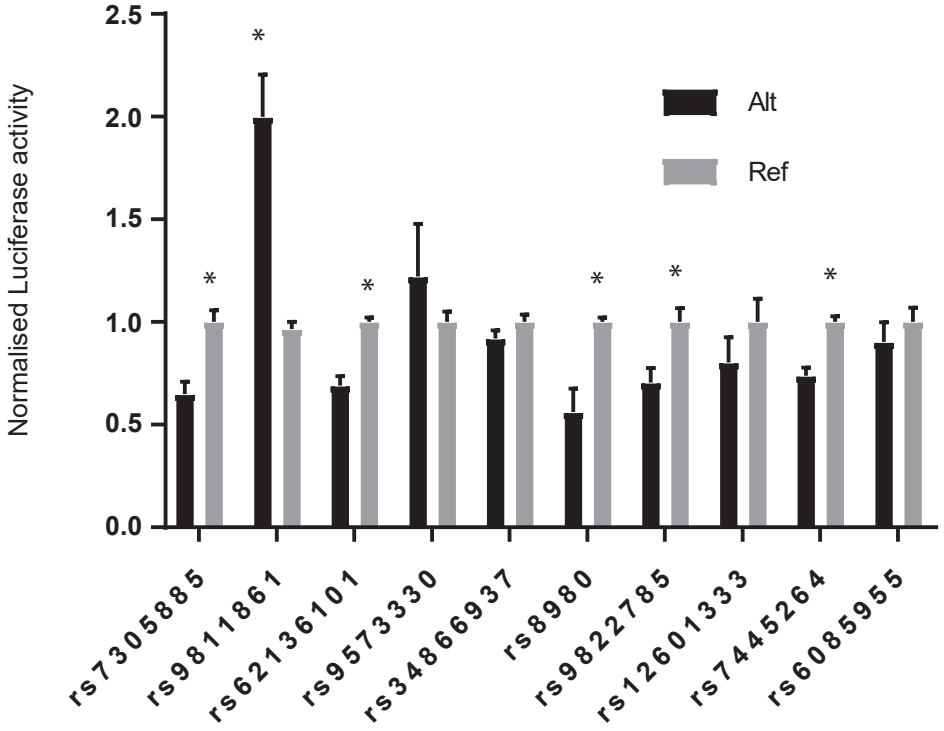


D

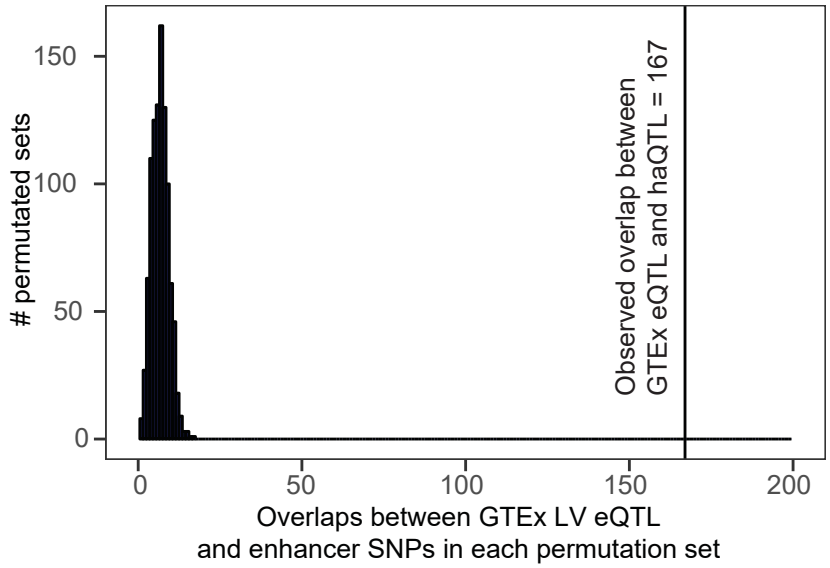


Online Figure IV. Quality control of enhancer SNPs.
(A) Sequencing read counts for SNPs called at H3K27ac enhancer loci, categorised according to whether they were found in 1K Genome (G1K) or dbSNP databases, or novel. Error bars represent s.d.
(B, C) Sanger sequencing validation was performed for randomly selected 60 LV samples; 6 samples for each SNP at 10 loci. Samples were chosen to represent homozygous reference, homozygous alternative and heterozygous status. Examples of Sanger sequencing results are shown in (C).
(D) Allele frequency for SNPs in all 1,680 haQTLs.

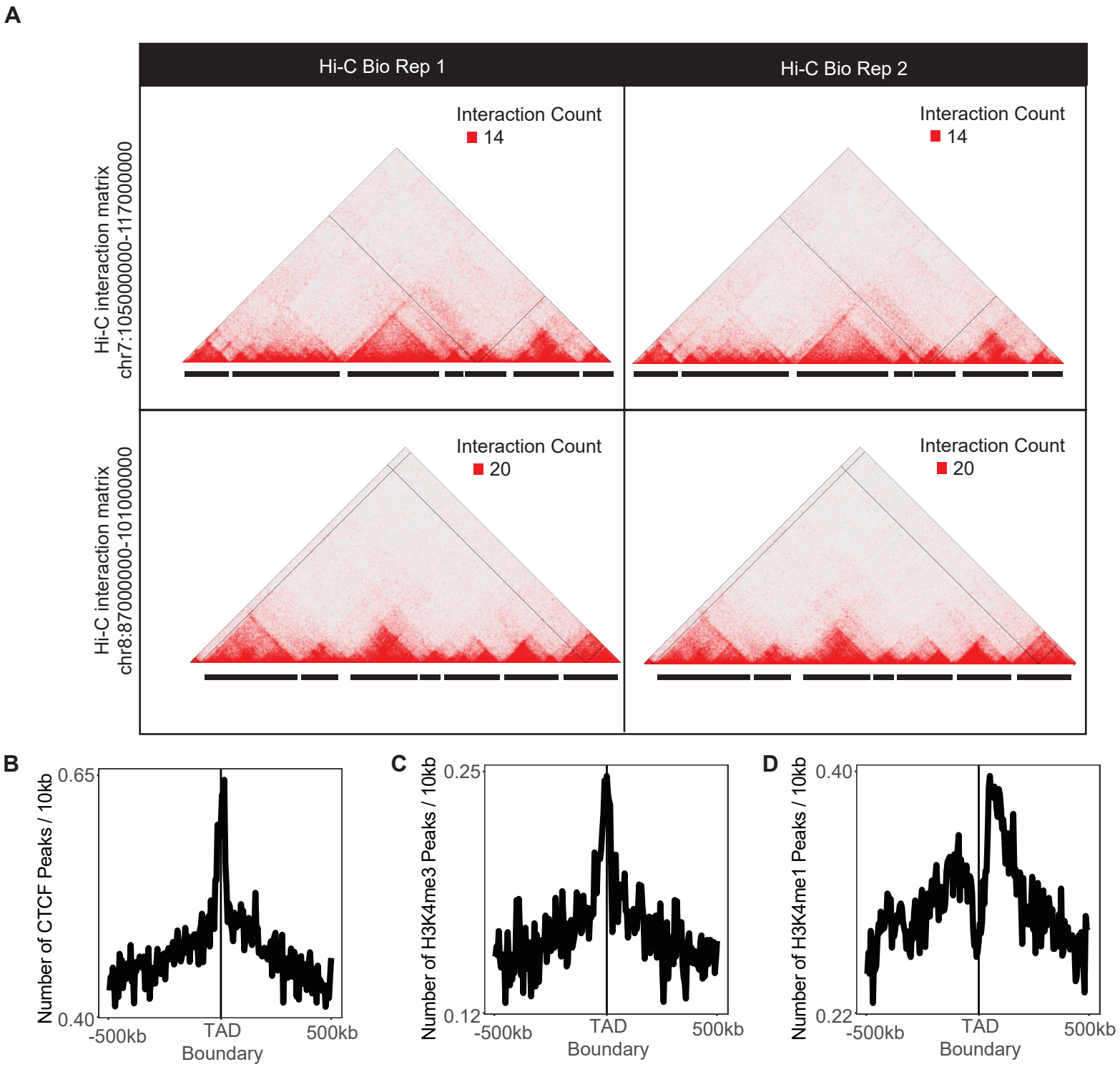
Online Figure V



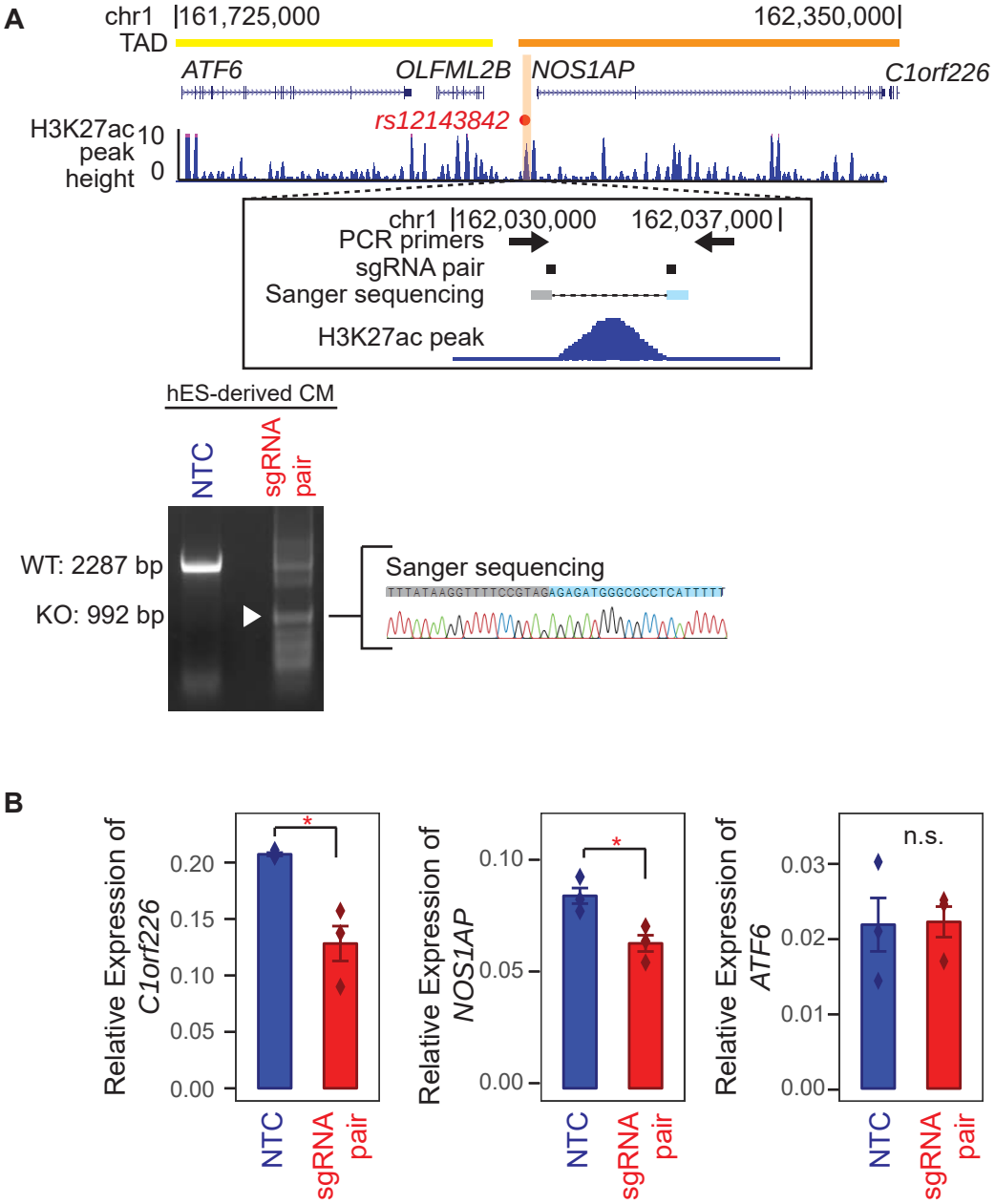
Online Figure V. Luciferase reporter assays. For luciferase reporter validation, 10 haQTLs were randomly selected from the final list of significant overlaps with previous GWAS studies (See also **Figure 6**). For each locus, 183 bp oligo sequence at the enhancer region, centred around either the reference or alternative SNV, was cloned into the Firefly luciferase reporter construct. *Black*: enhancer constructs bearing the alternative SNVs; *Grey*: enhancer constructs bearing the reference SNVs. Error bars represent s.e.m. in 6 biological replicates performed in 2 batches of iPS-derived cardiomyocyte cultures. All 10 loci proved to have enhancer activity, and in the expected concordance for differential enhancer activity for reference and alternative SNVs respectively, with 6 tests meeting statistical significance (*$p < 0.05$ Student's t test). See **Online Table XII** for details on selected haQTL loci.


150
100
50
0
permutated sets
Observed overlap between
GTEx eQTL and haQTL = 167
0
50
100
150
200
Overlaps between GTEx LV eQTL
and enhancer SNPs in each permutation set

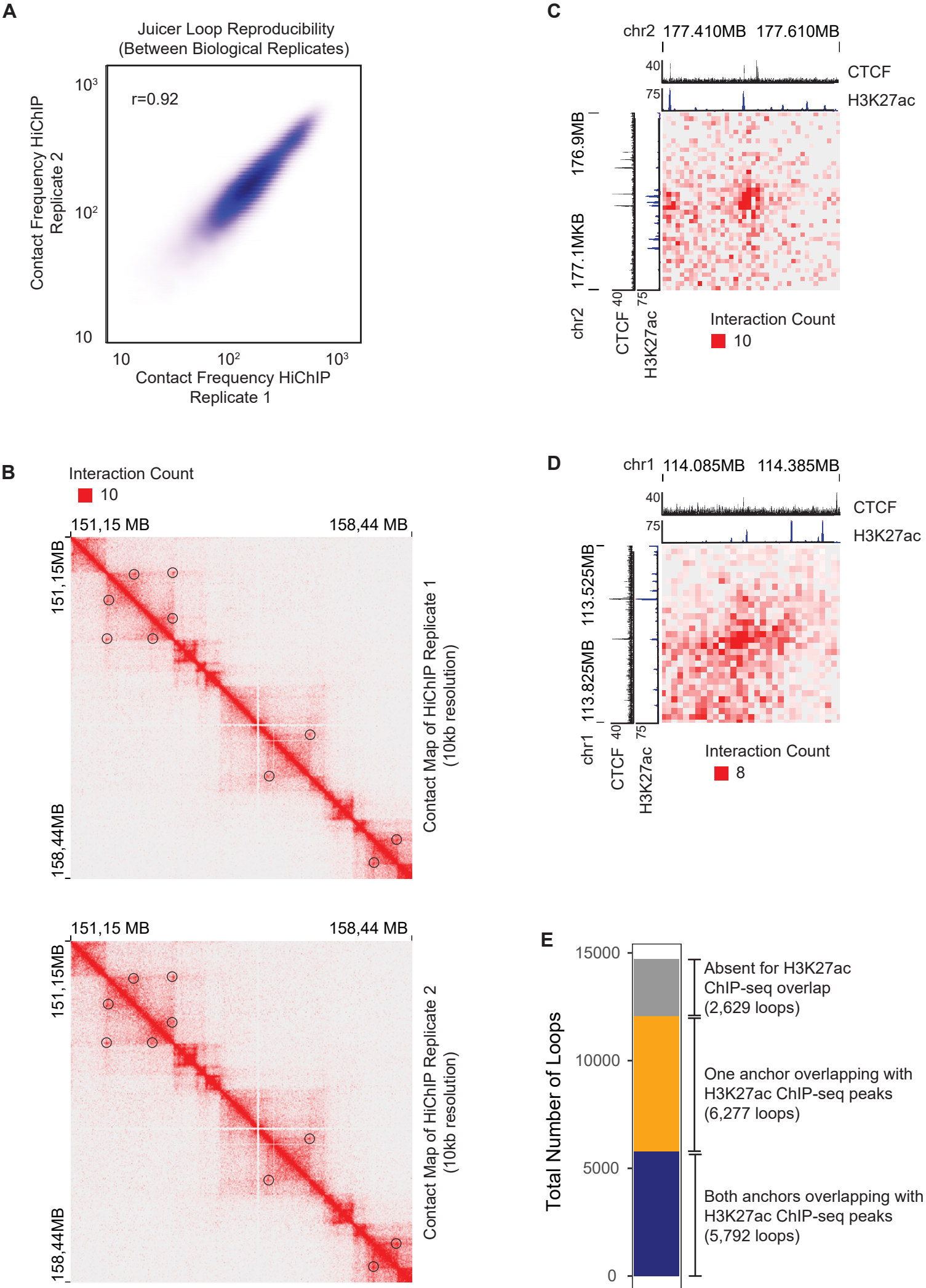
Online Figure VI. Enrichment analysis of haQTL in GTEx eQTL. Random permutation analysis showing that haQTL are significantly more overlapped with eQTL from GTEx project, compared to 1,000 randomly simulated sets of 1,000 SNPs, selected matched by MAF and distance to nearest TSS (see methods for permutation tests).


A
Hi-C Bio Rep 1
Hi-C Bio Rep 2
Hi-C interaction matrix
chr7:105000000-117000000
Interaction Count
14
Hi-C interaction matrix
chr8:87000000-101000000
Interaction Count
20
B
Number of CTCF Peaks / 10kb
0.65
0.40
-500kb
TAD
Boundary
500kb
C
Number of H3K4me3 Peaks / 10kb
0.25
0.12
-500kb
TAD
Boundary
500kb
D
Number of H3K4me1 Peaks / 10kb
0.40
0.22
-500kb
TAD
Boundary
500kb
E
Insulation Score
0.20
0
0.60
-2MB
TAD
Boundary
2MB
F
Outside TAD
Within TAD
Outside TAD
2533 TADs (40kb resolution)
0
>30
Number of Fit-hi-C
Enhancer-Promoter Interactions

Online Figure VII. Quality metrics for Hi-C.
(A) TADs are unchanged and reproducible between Hi-C libraries from 2 different control non-failing LV. The black bars under each Hi-C interaction heatmaps represent TADs.
(B-D) CTCF binding, chromatin modifications (H3K4me3 and H3K4me1) flanking TAD boundaries.
(E) Insulation profile at TAD boundaries. Minima of the insulation profile denote high insulation characteristic of TAD boundaries.
(F) Enhancer-Promoter (E-P) interactions are anchored predominantly inside TADs. Each row represents a TAD. Columns represent scaled distance (0-100%) from the left to the right of TAD boundaries. Blue horizontal lines represent E-P interactions spanning the TAD locus.

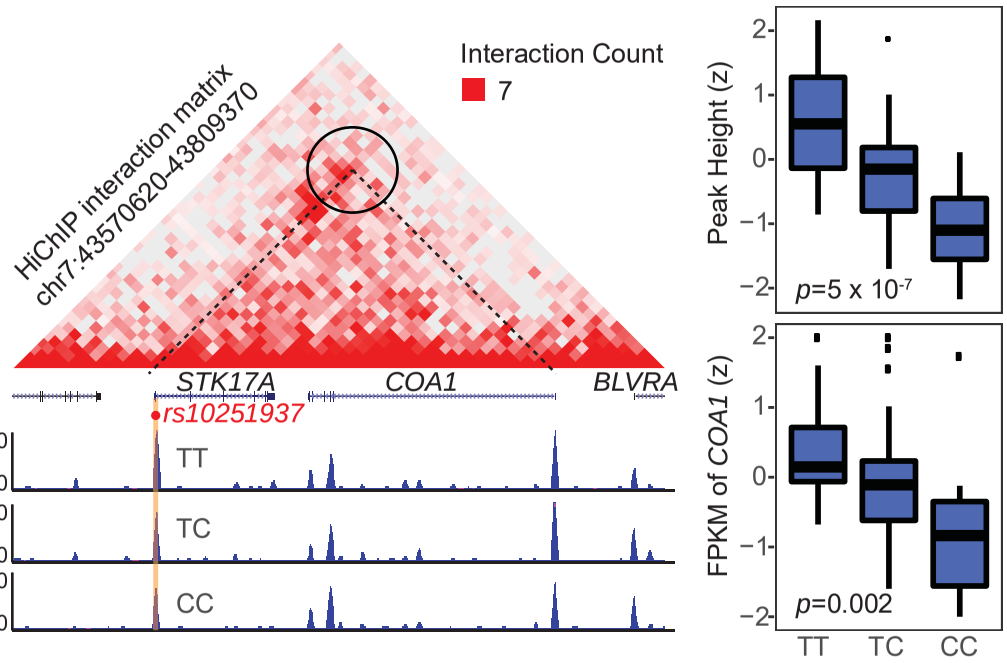

A
chr1 |161,725,000
162,350,000|
TAD
ATF6
OLFML2B
NOS1AP
C1orf226
rs12143842
H3K27ac peak height
10
0
chr1 |162,030,000
162,037,000|
PCR primers
sgRNA pair
Sanger sequencing
H3K27ac peak
hES-derived CM
NTC
sgRNA pair
WT: 2287 bp
KO: 992 bp
Sanger sequencing
TTTATAAGGTTTTCCGTAGAGAGATGGGCGCCTCATTTTT
B
Relative Expression of C1orf226
0.20
0.15
0.10
0.05
0.00
*
NTC
sgRNA pair
Relative Expression of NOS1AP
0.10
0.05
0.00
*
NTC
sgRNA pair
Relative Expression of ATF6
0.03
0.02
0.01
0.00
n.s.
NTC
sgRNA pair

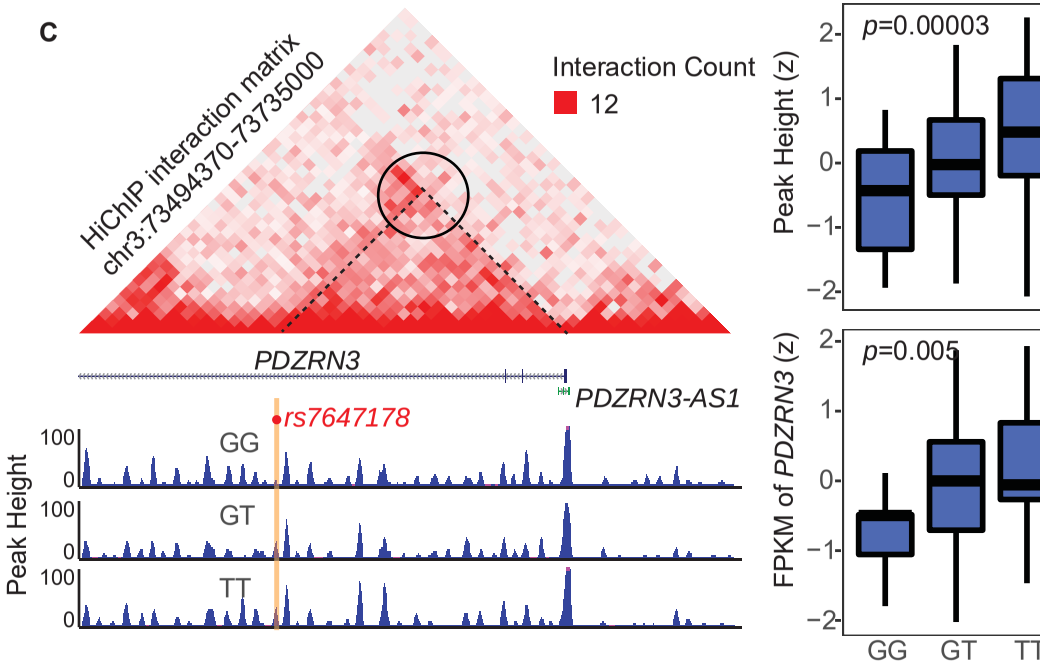
Online Figure VIII. CRISPR/Cas9-targeting of haQTL loci.
(A) Successful distal haQTL excision for rs12143842, in a representative pool of hES-derived CM, using a pair of sgRNA, performed in 3 independent batches of hES-CMs. Representative PCR amplification (gel image) performed using a set of flanking PCR primers, and Sanger sequencing of the amplified KO band with chromatogram displaying the ligated distal sites (grey and light blue). Yellow and orange horizonal bars refer to adjacent TADs containing their respective genes (see also contact matrix in **Figure 3C**).
(B) In hES-CM cells with haQTL (rs12143842) deleted using a pair of sgRNA (red bar), compared to control using non-targeting sgRNA (NTC, blue bar), *NOS1AP* (gene in *cis*, and within the TAD) and distal interacting gene *C1orf226* (also within the TAD), were consistently downregulated. The expression of other adjacent genes outside the TAD: *OLFML2B* (not shown) and *ATF6* were unchanged. *p<0.01. Student's t test. n.s.= not significant. Student's t test.

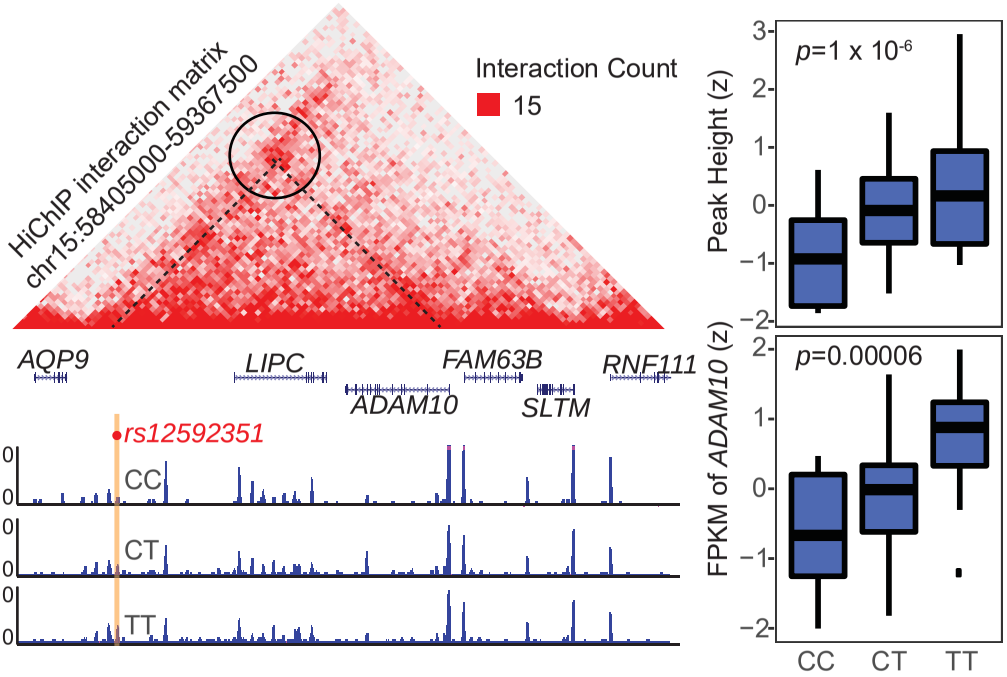

A
Juicer Loop Reproducibility
(Between Biological Replicates)
r=0.92
Contact Frequency HiChIP
Replicate 2
Contact Frequency HiChIP
Replicate 1
10
10²
10³
B
Interaction Count
10
151,15 MB
158,44 MB
151,15MB
158,44MB
Contact Map of HiChIP Replicate 1
(10kb resolution)
Contact Map of HiChIP Replicate 2
(10kb resolution)
C
chr2 177.410MB 177.610MB
CTCF
H3K27ac
176.9MB
177.1MKB
Interaction Count
10
D
chr1 114.085MB 114.385MB
CTCF
H3K27ac
113.525MB
113.825MB
Interaction Count
8
E
Total Number of Loops
15000
10000
5000
0
Absent for H3K27ac
ChIP-seq overlap
(2,629 loops)
One anchor overlapping with
H3K27ac ChIP-seq peaks
(6,277 loops)
Both anchors overlapping with
H3K27ac ChIP-seq peaks
(5,792 loops)

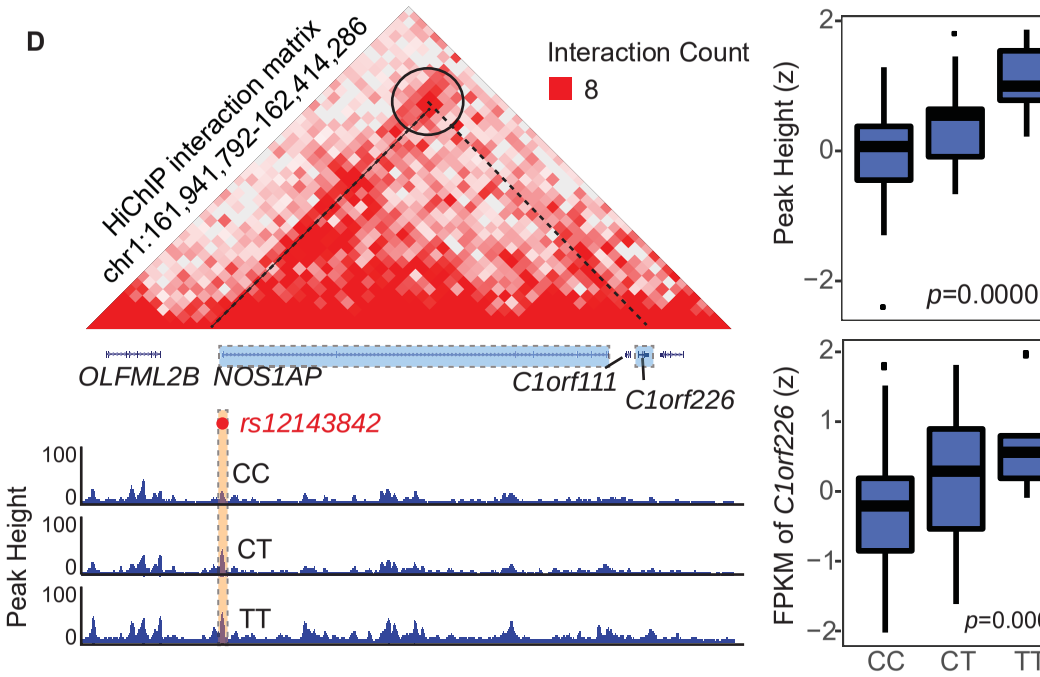
Online Figure IX. Quality metrics for HiChIP.
(A) Reads support reproducibility of loops in two independent H3K27ac HiChIP replicates.
(B) Juicebox visualization of chromatin loops (black circles) in two HiChIP biological replicates.
(C-D) Two examples of HiChIP loops at 10 kb resolution, showing H3K27ac enrichment at their anchors.
(E) Stacked bar plots showing 80% of chromatin loop anchors from H3K27ac HiChIP overlap with H3K27ac ChIP-seq peaks.

nline Figure X

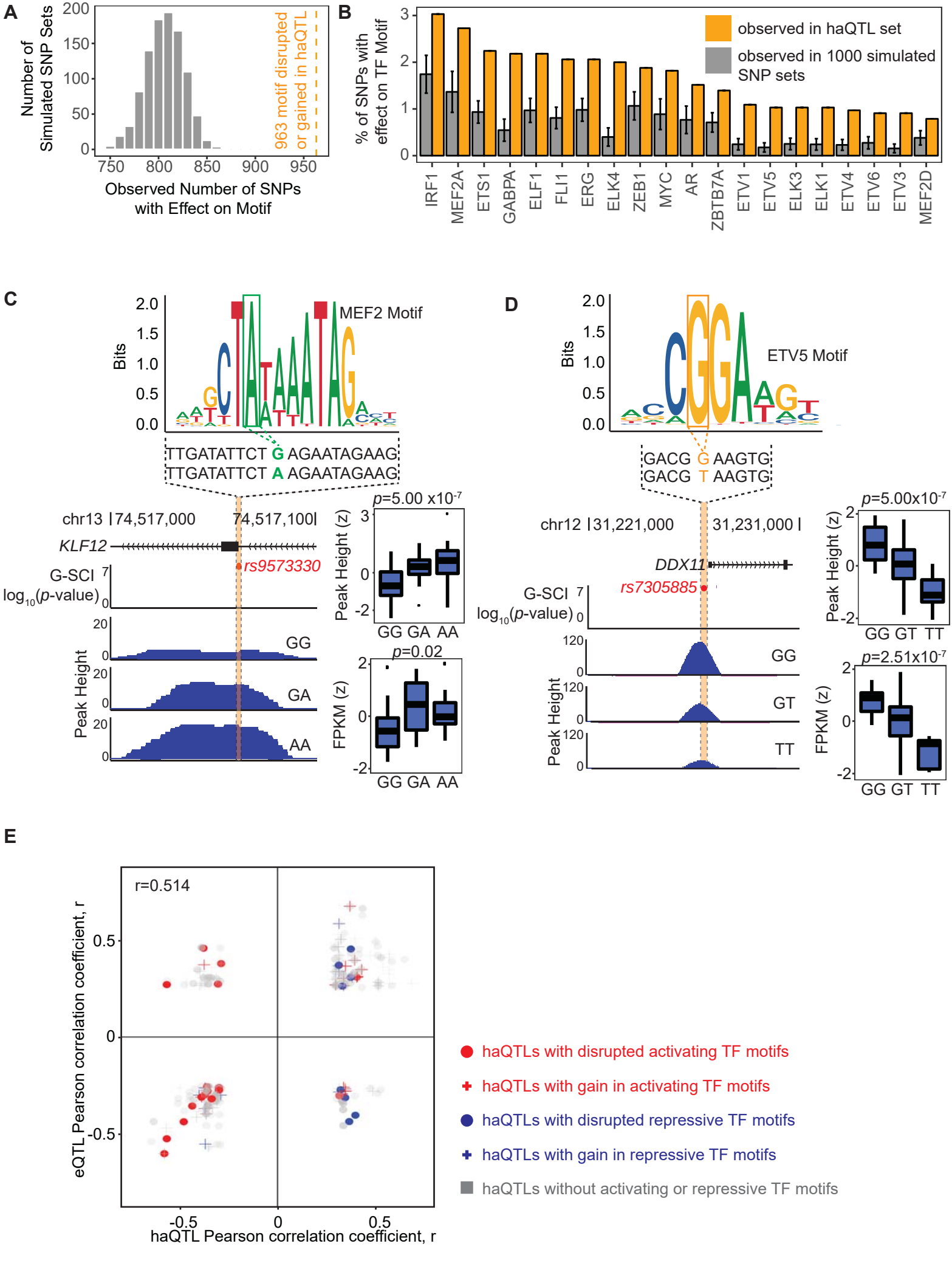




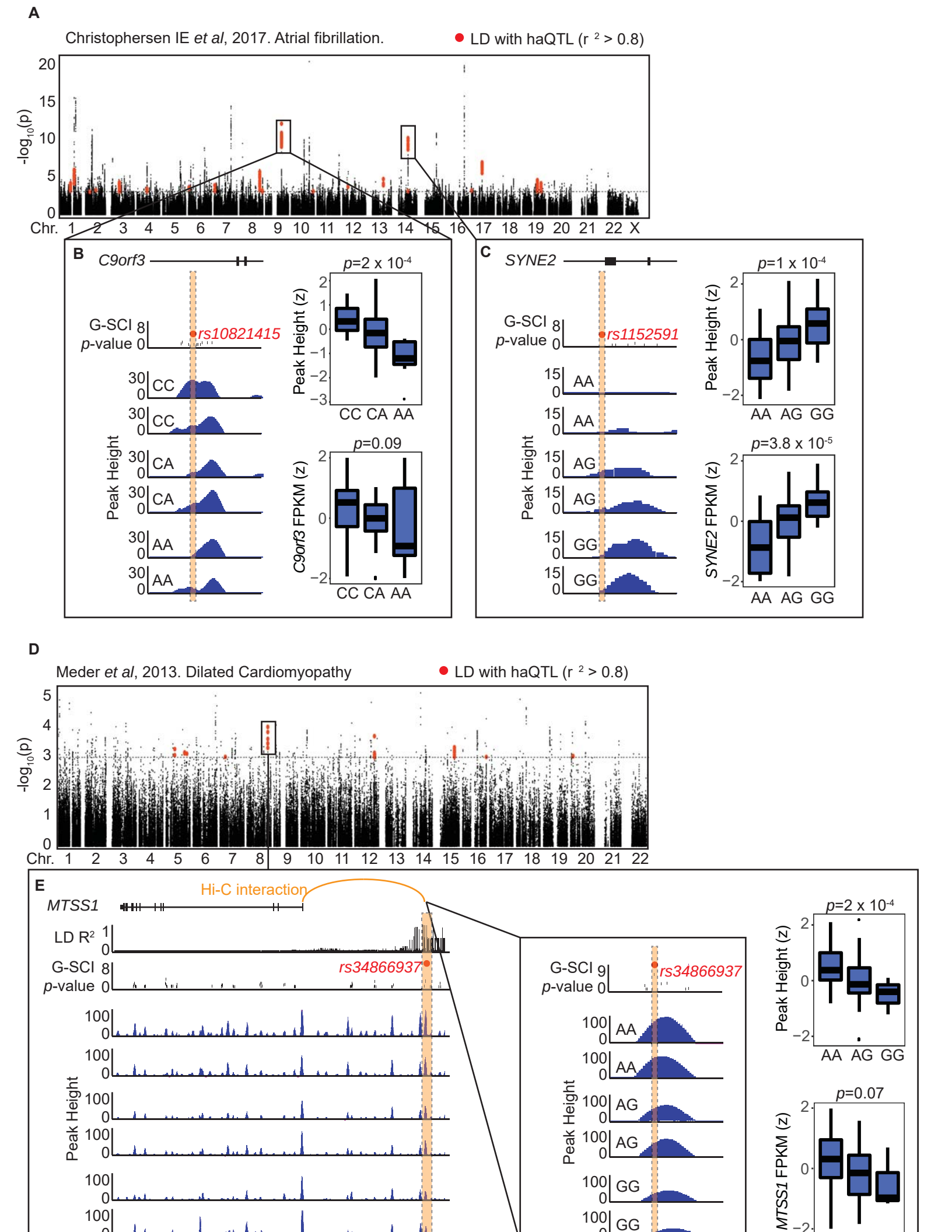




Online Figure X. HiChIP integration with haQTL. Examples of haQTLs with correlated distal gene expression, with distal interactions validated by HiChIP.
(A-D) HiChIP loops (black circles) in the interaction matrices mark out the interactions (dotted black lines) connecting each 2 distal loci. Significant H3K27ac peak height association with the alternative alleles are also associated with concordant distal gene expression.
(D) is a validation using HiChIP analysis here, for the same analysis that used Hi-C in **Figure 3C**.


A
Number of Simulated SNP Sets
963 motif disrupted or gained in haQTL
Observed Number of SNPs with Effect on Motif
B
% of SNPs with effect on TF Motif
observed in haQTL set
observed in 1000 simulated SNP sets
IRF1
MEF2A
ETS1
GABPA
ELF1
FLI1
ERG
ELK4
ZEB1
MYC
AR
ZBTB7A
ETV1
ETV5
ELK3
ELK1
ETV4
ETV6
ETV3
MEF2D
C
MEF2 Motif
Bits
TTGATATTCT G AGAATAGAAG
TTGATATTCT A AGAATAGAAG
chr13 74,517,000
74,517,100
KLF12
rs9573330
G-SCI log10(p-value)
Peak Height
GG
GA
AA
p=5.00 x10-7
Peak Height (z)
p=0.02
FPKM (z)
GG GA AA
D
ETV5 Motif
GACG G AAGTG
GACG T AAGTG
chr12 31,221,000
31,231,000
DDX11
rs7305885
GG
GT
TT
p=5.00x10-7
p=2.51x10-7
E
r=0.514
eQTL Pearson correlation coefficient, r
haQTL Pearson correlation coefficient, r
haQTLs with disrupted activating TF motifs
haQTLs with gain in activating TF motifs
haQTLs with disrupted repressive TF motifs
haQTLs with gain in repressive TF motifs
haQTLs without activating or repressive TF motifs

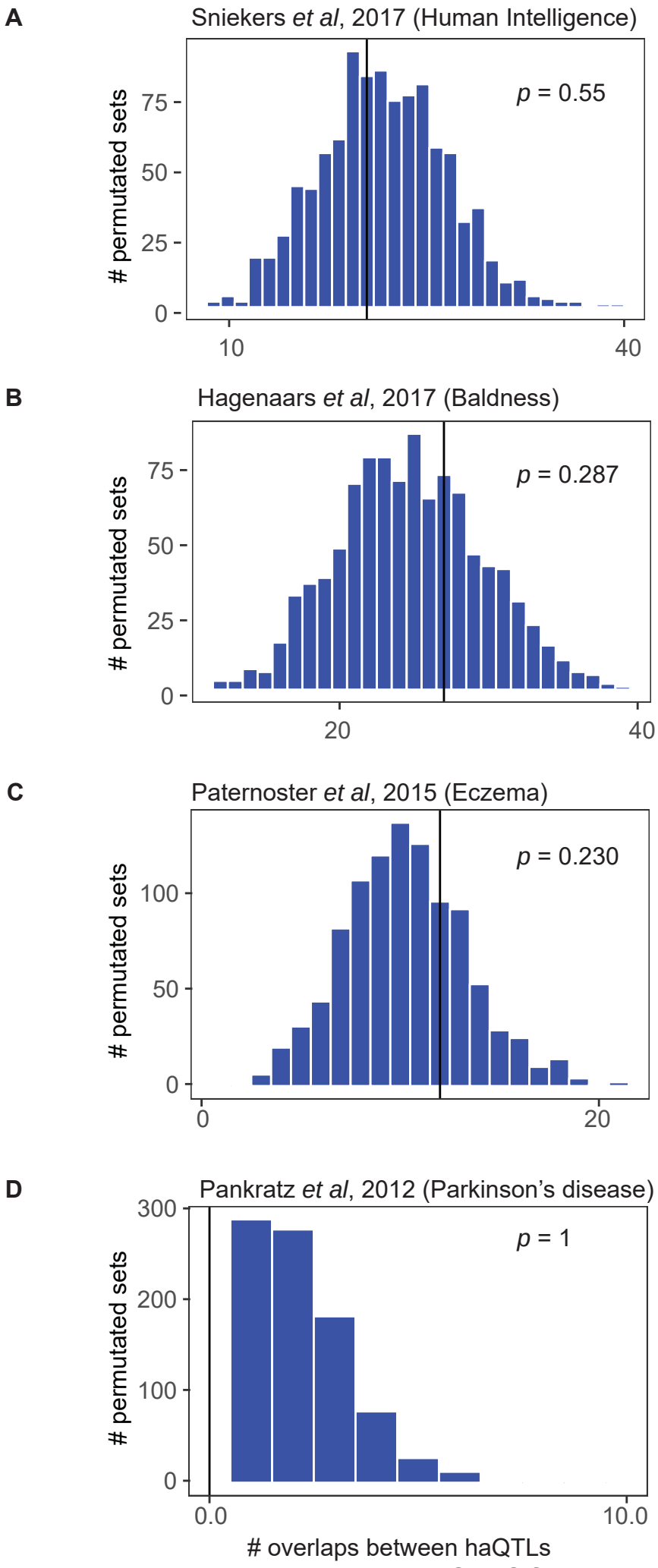
Online Figure XI. haQTLs inferred to alter transcription factor (TF) binding.
(A) Random permutation analysis showing that haQTL are significantly more enriched for SNPs bearing inferred TF motif change.
(B) The effect of alternative sentinel SNPs in haQTLs on randomly selected TF, showing that for each TF, haQTLs bore a greater likelihood of altering the TF motif, compared to simulated SNP sets.
(C) An example of haQTL (sentinel: rs9573330) inferred to produce a gain in MEF2A motif at a H3K27ac locus in the intron of *KLF12* (FDR=1.26 x 10^{-4}). The motif gain is associated with increased H3K27ac peak height, and concordant increase in *KLF12* gene expression (FDR<20%).
(D) An example of haQTL (sentinel: rs7305885) inferred to disrupt ETV5 motif in the promoter of *DDX11*. The alternative allele, T, is associated with reduced acetylation (FDR=1.26 x 10^{-4}), and reduced *DDX11* gene expression (FDR=1.54 x 10^{-4}).
(E) Plot showing the correlation between all 1,680 haQTL and their corresponding gene expression. haQTLs are as labelled, where the alternative alleles contain a SNP that produces an inferred TF motif change (with eQTL FDR<40%) for "activating" and "repressive" TFs as shown in **Figure 4B**. "Activating" TFs are those whose motif gain or disruption was correlated to an increase or decrease in H3K27ac, respectively. "Repressive" TFs are those with the converse relationship.


A
Christophersen IE et al, 2017. Atrial fibrillation.
LD with haQTL (r 2 > 0.8)
-log10(p)
Chr. 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X
B
C9orf3
G-SCI p-value
rs10821415
Peak Height
CC
CC
CA
CA
AA
AA
p=2 x 10-4
Peak Height (z)
CC CA AA
p=0.09
C9orf3 FPKM (z)
CC CA AA
C
SYNE2
G-SCI p-value
rs1152591
Peak Height
AA
AA
AG
AG
GG
GG
p=1 x 10-4
Peak Height (z)
AA AG GG
p=3.8 x 10-5
SYNE2 FPKM (z)
AA AG GG
D
Meder et al, 2013. Dilated Cardiomyopathy
LD with haQTL (r 2 > 0.8)
-log10(p)
Chr. 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
E
Hi-C interaction
MTSS1
LD R2
G-SCI p-value
rs34866937
Peak Height
rs34866937
AA
AA
AG
AG
GG
GG
p=2 x 10-4
Peak Height (z)
AA AG GG
p=0.07
MTSS1 FPKM (z)

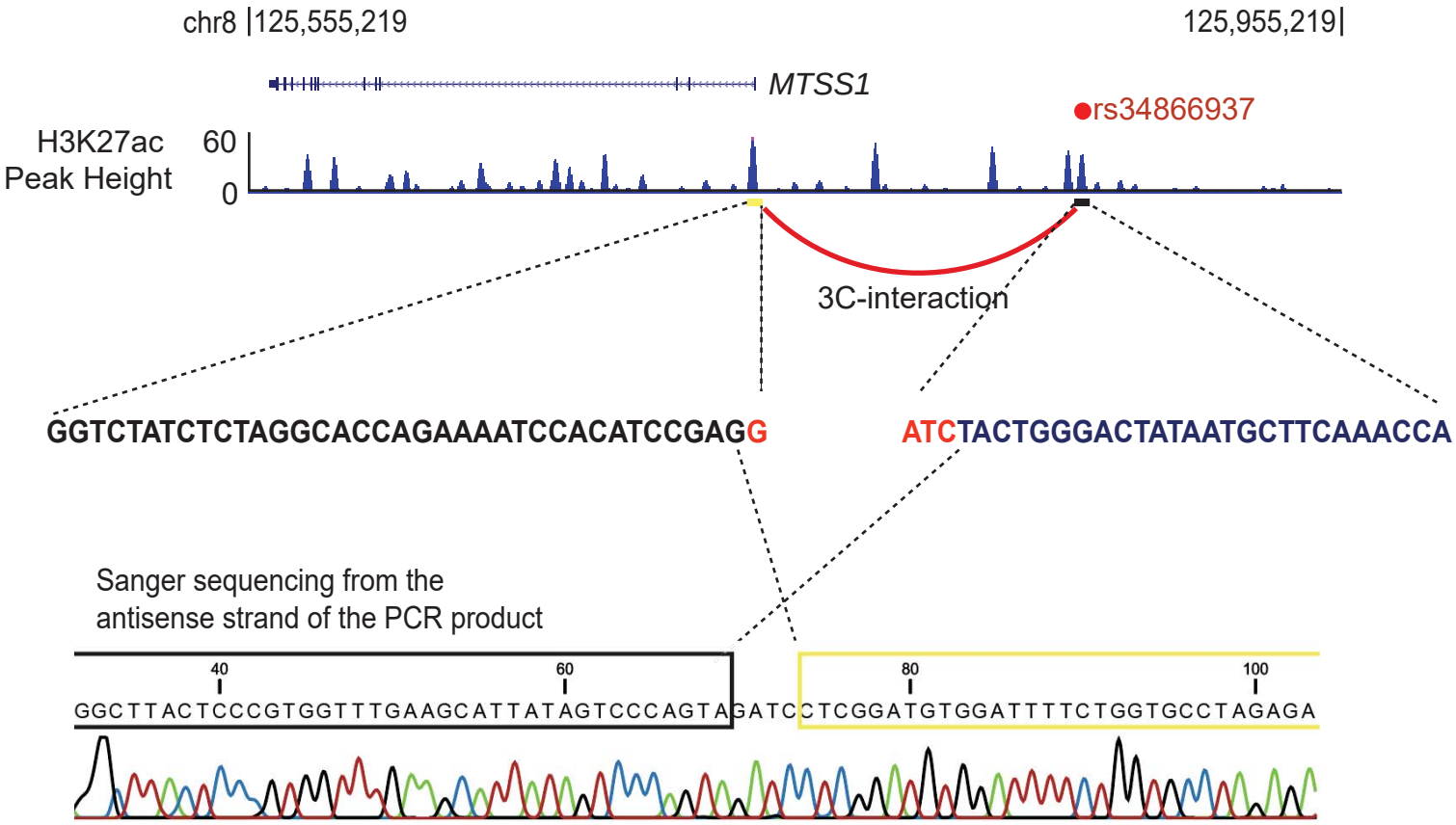
Online Figure XII. Examples of haQTLs identified in subthreshold GWAS SNPs.
(A) As in **Figure 5**, a Manhattan plot of adjusted *P*-values for all SNPs from the GWAS meta-analysis for atrial fibrillation, in LD (r^2>0.8) with SNPs in our 249,732 cardiac enhancers (black). Red represents colocalised SNPs in 22 loci that are either in LD or overlapping with haQTLs.
(B, C) Examples of 2 independent haQTLs (sentinels: rs10821415 and rs1152591) colocalised with atrial fibrillation GWAS sub-threshold SNPs. For **(c)**, the alternative A allele for rs10821415 is associated with reduced H3K27ac peak height (FDR=0.0105) located within a *C9orf3* intron, and a trend of reduced gene expression FDR=0.74) (boxplots). Notably, rs10821415 is a significant eQTL for *C9orf3* in the GTEx project. For **(D)**, the alternative G allele for rs1152591 is associated with increased H3K27ac peak height (FDR=0.0006) located within a *SYNE2* intron, and increased *SYNE2* gene expression (FDR=0.018) (boxplots).
(D) As in **Figure 5**, a Manhattan plot of adjusted *P*-values for all SNPs from the GWAS for dilated cardiomyopathy, in LD (r^2>0.8) with our enhancer SNPs (black). Red represents colocalised SNPs in 9 loci that are either in LD or overlapping with haQTLs.
(E) An example of haQTL (sentinel: rs34866937) colocalised with dilated cardiomyopathy GWAS sub-threshold SNPs. The alternative G allele is associated with reduced H3K27ac peak height (FDR=1.26 x 10^{-4}) located 120 kb from the promoter of *MTSS1*, and a trend of reduced *MTSS1* gene expression (FDR=0.71) (boxplots). Again, rs34866937, is a significant eQTL for *MTSS1* in the GTEx project.

Fox ER et al, 2013. EchoGen Consortium.
LV diastolic internal dimension (LVDD).
LV mass (LVM).
LV systolic dysfunction (LVSD).
LV wall thickness (LVWT).
LV fractional shortening (FS).
LA antero-posterior diameter (LA).
Smith NL et al, 2010. CHARGE: Incident Heart Failure.
Smith JG et al, 2016. CHARGE: Heart Failure Mortality.
-log10(p)
Chr. 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22

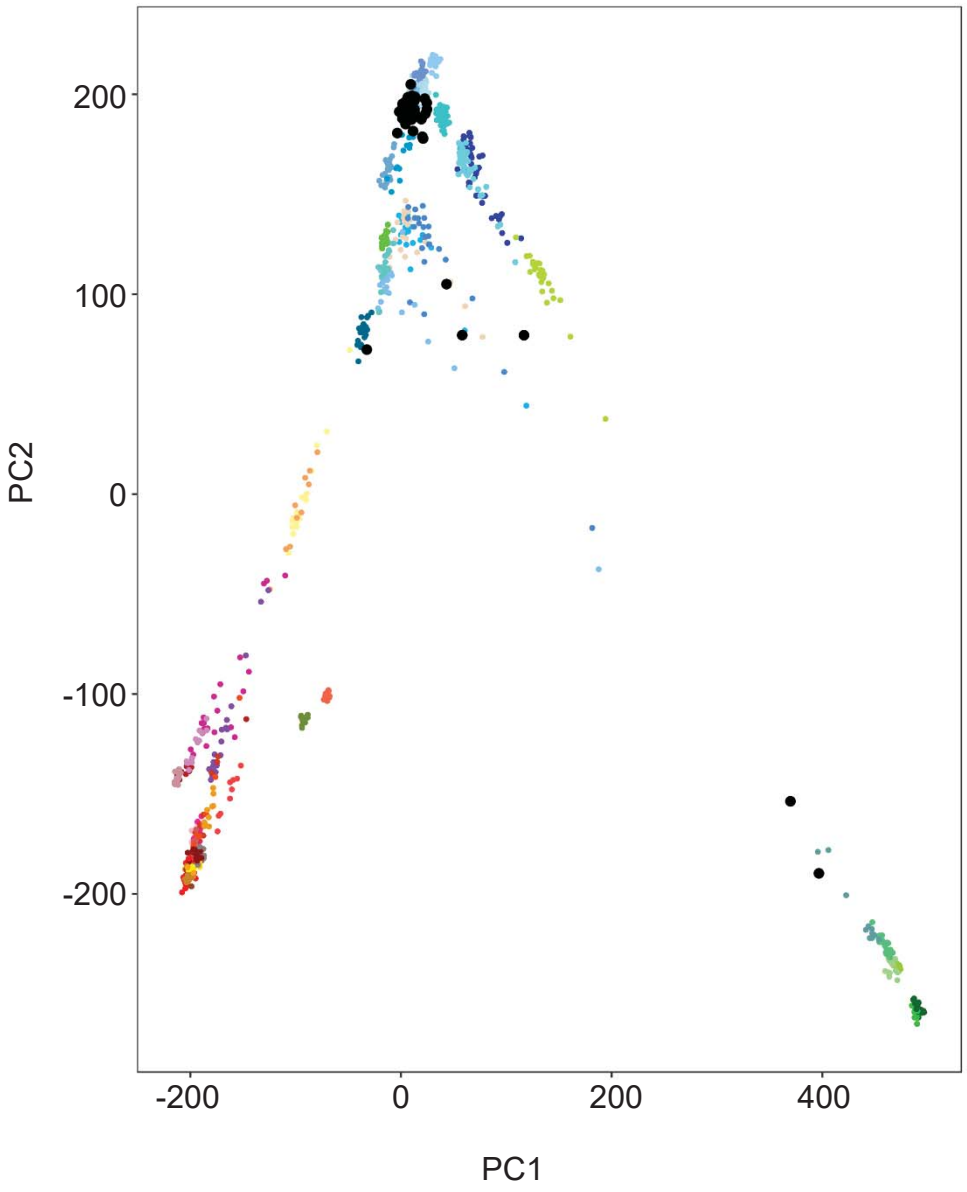
Online Figure XIII. haQTLs enriched for heart GWAS sub-threshold SNPs. In addition to the 2 GWAS shown in **Figure 5** and **Online Figure XII**, intersection between our haQTLs was also carried out with 8 other GWAS datasets, yielding a total of 62 unique colocalised loci. See also **Figure 6** and **Online Table XII**.


A
Sniekers et al, 2017 (Human Intelligence)
p = 0.55
# permutated sets
75
50
25
0
10
40
B
Hagenaars et al, 2017 (Baldness)
p = 0.287
# permutated sets
75
50
25
0
20
40
C
Paternoster et al, 2015 (Eczema)
p = 0.230
# permutated sets
100
50
0
0
20
D
Pankratz et al, 2012 (Parkinson's disease)
p = 1
# permutated sets
300
200
100
0
0.0
10.0
overlaps between haQTLs

Online Figure XIV. haQTLs is not enriched for non-heart GWAS sub-threshold SNPs. Non-significant enrichment of haQTL for SNPs from non-heart-related GWAS as negative control comparison. Datasets from 4 non heart-related GWAS (human intelligence, baldness, eczema and Parkinson's disease) were downloaded and processed through the same permutation tests (as in **Figure 5** and **Figures S12**). These proven to have non-significant enrichment for haQTL for each dataset. (empirical *P*-values 0.287 to 1).


chr8 |125,555,219
125,955,219|
MTSS1
rs34866937
H3K27ac
Peak Height
60
0
3C-interaction
GGTCTATCTCTAGGCACCAGAAAATCCACATCCGAGG
ATCTACTGGGACTATAATGCTTCAAACCA
Sanger sequencing from the
antisense strand of the PCR product
40
60
80
100
GGCTTACTCCCGTGGTTTGAAGCATTATAGTCCCAGTAGATCCTCGGATGTGGATTTTCTGGTGCCTAGAGA

Online Figure XV. 3C-PCR validation of MTSS1 enhancer-promoter interaction. 3C-PCR validation and Sanger sequencing to show the interaction between *MTSS1* promoter and the distal enhancer containing the rs34866937 SNP. Despite being ~120 kb apart, the 3C primers amplified a ~250 bp product with one half aligning to the *MTSS1* promoter (in yellow), and the other half aligning to the distal enhancer (in black). 3C primers used were 5'-GAAGCTCCCTAAATACACAT-3' and 5'-CTCCCTTCTAAAAATATCTCCC-3'. Sanger sequencing was performed from the antisense strand. The chromatogram displays the PCR product with predicted fragments interfaced at the *MboI* restriction enzyme recognition site GATC. *MboI* is the restriction enzyme used for 3C library construction.



Population

- Adygei
- **LV samples from this Study**
- Balochi
- BantuSouthAfrica
- BantuKenya
- Basque
- Bedouin
- BiakaPygmy
- Brahui
- Burusho
- Cambodian
- Dai
- Colombian
- Daur
- Druze
- French
- Han
- Han-NChina
- Hezhen
- Italian
- Japanese
- Kalash
- Karitiana
- Lahu
- Makrani
- Mandenka
- Maya
- MbutiPygmy
- Melanesian
- Miao
- Mixed
- Mongola
- Mozabite
- Naxi
- Orcadian
- Oroqen
- Papuan
- Pathan
- Pima
- Russian
- San
- Sardinian
- She
- Sindhi
- Surui
- Tu
- Tujia
- Tuscan
- Uygur
- Xibo
- Yakut
- Yi
- Yoruba

Online Figure XV1. Ancestry estimation against the HGDP worldwide map for LV samples. Ancestry estimation against the HGDP worldwide map for LV samples used in this study (in black). Six LV samples were found to be inaccurately labelled as “European” and therefore removed from subsequent analysis.

Online Table I Tissue biodata

Online Table II Quality metrics for ChIP-seq, RNA-seq, Hi-C and HiChIP

Online Table III Consensus H3K27ac Peaks

Online Table IV Differential Acetylation Peaks

Online Table V Pathway analysis of Differential Acetylation Peaks

Online Table VI haQTL annotation

Online Table VII eQTL list

Online Table VIII TAD coordinates identified in Hi-C

Online Table IX Chromatin loops identified in HiChIP

Online Table X TF motif alteration predicted

Online Table XI Enrichment level of haQTL in various GWAS

Online Table XII List of overlaps between GWAS and haQTL

Online Table XIII sgRNA designs, primer list