Supplemental Materials and Methods

Mice

Mice were fed with standard mouse diets and kept under a regular 12 hour light/12 hour dark cycle at 18-23°C. Mouse experiments were performed under protocols approved by the Boston Children's Hospital institutional animal care and use committee. GATA4^{flbio} (Jackson Labs #018121)[11], GATA4^{flox} (Jackson Labs #008194)[46], Rosa26^{fsTRAP} (Jackson Labs #022367)[47], Rosa26^{mTmG} (Jackson Labs #007576)[48], Tie2-Cre (Jackson Labs #008863)[16], Cdh5-CreERT2, also known as Cdh5 (Pac)-CreERT2 (Taconic #13073)[17], and Tnnt2-Cre (Jackson Labs #024240)[15] have been reported previously. Rosa26^{fsBirA} was derived from the Rosa26^{fsTRAP} mouse by removal of the frt-TRAP-frt cassette using germline Flp recombination as described previously[13]. CFW mice (Charles River, 024) were used for AAV reporter assays. No animals were excluded from this study. Sex was not considered as a biological variable, and both male and female animals were used in the experiments.

Gata4 knockout in embryonic cardiomyocytes with Tnnt2-Cre

To inactivate *Gata4* in embryonic cardiomyocytes, *Gata4$^{fl/+}$;Tnnt2-Cre*+ male mice were crossed to *Gata4$^{fl/fl}$* female mice. Embryos were collected at E12.5 for immunostaining or H&E staining[49]. Genotype of embryos was determined by PCR of genomic DNA.

Gata4 knockout in Cdh5CreERT2-labeled ECs

Gata4$^{fl/+}$;Rosa26$^{mTmG/+}$;Cdh5-CreERT2+ male mice were crossed to *Gata4$^{fl/fl}$;Rosa26$^{mTmG/+}$* female mice. The activity of Cre was induced with 40 mg tamoxifen/kg body weight of pregnant mice by oral gavage at E9.5 and embryos were collected at E12.5 for immunostaining or E13.5 for H&E staining.

Histology

Tissues were fixed in 4% PFA at 4°C overnight. Paraffin sections (7 μm thick) of H&E stained E13.5 embryos were imaged with a Keyence BZ-X700 microscope. Average left ventricular wall thickness was measured with ImageJ by dividing the area of the free wall by the length of its outer perimeter. For frozen section (10 μm thick) immunostaining, cryosections were stained with GATA4 antibody (Santa Cruz, sc-5031, 1:50 dilution), Cardiac troponin T antibody (Abcam, ab45932; 1:100 dilution), cardiac troponin I antibody (Abcam, ab56357, 1:100 dilution), or phosphohistone H3 antibody (MilliporeSigma, 06-570, 1:50 dilution) for overnight at 4°C and then stained with secondary antibodies (Alexa Fluor Secondary Antibodies, 1:1000 dilution). Images were captured by confocal microscopy (Olympus FV3000). Antibody specificity was confirmed by staining knockout samples (GATA4), by omitting primary antibodies, or by comparison to previously reported staining results. Representative images were selected based on average results of no less than 4 images per group.

Cell culture

293T cells and NIH3T3 cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM, Thermo Fisher, Catalog No.11-965-092) with 10% FBS. HUVEC cells were cultured in EGM™-2 Endothelial Cell Growth Medium-2 (Lonza, Catalog #: CC-3162) for no more than 7 passages.

Lentivirus was prepared with standard protocol by transfecting 293T cells with lentivectors along with packaging plasmid psPAX2 and envelope plasmid pMD2.G. For control NIH/3T3 stable cell line, pCDH-MSCV-MCS-EF1a-GFP-T2A-Puro (System Biosciences, cat#CD713B-1) was used to generate GFP lentivirus. To get the HA-tagged GATA4 overexpression lentivector, mouse Gata4 coding sequence with 2xHA on its C-terminal replaced the coding sequence of GFP in pSico (Addgene, Cat#11578), and followed by a T2A-Hygromycin cassette as drug

selection marker. To get Flag-tagged Nkx2-5 or Ets1 overexpression lentivectors, coding sequence of mouse Nkx2-5 or human Ets1 with N-terminal 3xFlag tag replaced the coding sequence of GFP in pSico, and followed by an EF1a-GFP-T2A-Puro cassette as drug selection marker.

The lentivirus was concentrated with Amicon Ultra Centrifugal Filters (MilliporeSigma, UFC910024) before adding to NIH/3T3 cells. The transduced cells were selected with puromycin (2 μg/mL) and hygromycin (200 μg/mL) for three weeks and then maintained in 0.5 μg/mL puromycin and 50 μg/mL hygromycin for ChIP-seq and ATAC-seq.

Protein precipitation and immunoblots

Immunoblotting was performed using standard protocols. Briefly, precast Bis-Tris protein gels (Invitrogen, Cat# NP0321BOX) were used to resolve proteins, then the proteins were transferred to PVDF membranes. The blots were blocked in TBST buffer supplemented with 5% non-fat milk for one hour at room hour before incubated with primary antibodies at 4°C overnight. The following primary antibodies were diluted 1:1000 in TBST buffer supplemented with 5% non-fat milk: GATA4, Santa Cruz, Cat#sc-9053; BirA, Abcam, Cat#ab14002; Nkx2-5, Santa Cruz, Cat#14033; Ets1, Cell Signaling, Cat#14069; HA-tag, Cell Signaling, Cat#3724; Flag-tag, Sigma, Cat#F1804. Secondary antibodies conjugated with HRP were diluted 1:10,000 in TBST buffer supplemented with 5% non-fat milk. Results are representative of no less than two independent experiments.

For streptavidin pulldown of $GATA4^{flbio}$ from hearts, 12-15 mouse heart ventricles were homogenized in cold PBS and protein lysate was prepared with 0.5% SDS lysis buffer (1xPBS supplemented with 1 mM DTT, protease inhibitor cocktail, and 0.5% SDS). Biotinylated proteins were then pulled down by incubating with Dynabeads M-280 streptavidin (ThermoFisher Scientific, Cat#11206D) at 4°C for one hour and then eluted from the beads with SDS elution buffer (20 mM Tris-HCl pH 8.0, 1% SDS, and 2 mM ETDA) at 98°C for 10 minutes.

For co-immunoprecipitation assays, $2x10^7$ cells were extracted with IP buffer (1xPBS supplemented with 1% Igepal-630, 0.5 mM DTT, 1× Protease Inhibitor Cocktail) and then incubated with HA-tag antibody or Flag-tag antibody (2 μg antibody per 10 mg total protein) overnight at 4°C. 1,000 units of Benzonase nuclease (Sigma, Cat#70664) were added to the protein lysate prior to incubation with HA and Flag antibodies. Then Dynabeads Protein-G (Thermo Fisher, 10004D) were added to the lysate to pull down the antibody and bound proteins. Co-immunoprecipitated proteins were detected by western blot.

Endogenous co-immunoprecipitation of GATA4 and ETS1 was performed with 10 E16.5 $Gata4^{fb/+};Rosa26^{BirA}$ embryonic mouse hearts, $Gata4^{fb/+};Rosa26^{WT}$ embryonic hearts were used as negative control. Ventricular hearts were homogenized in PBS with 1% formaldehyde and incubated at room temperature for 10 minutes. Cross-linking was quenched by adding glycine to 0.125 M. Protein lysate was prepared with 0.5% SDS lysis buffer as described above. Biotinylated proteins and associated proteins were pulled down by Dynabeads M-280 (Thermo Fisher) and then eluted with SDS elution buffer at 98°C for 10 minutes.

AAV reporter assays

Wildtype and mutant 0.4 kb of genomic region of CM-selective GATA4 peak 23 kb downstream of the *Vegfa* transcription start site (TSS) were synthesized (sequences listed in Table S2) and cloned into a self-complementary adeno-associated virus (AAV) reporter vector[12]. Adeno-associated virus serotype 9 (AAV9) was generated by transfecting HEK293T cells and purified using OptiPrep density gradient purification (Sigma). Purified virus (5E11 viral genomes/g body weight) was subcutaneously injected into wild type CFW newborn pups (P0). 8-18 heart ventricles for each virus were harvested at P7. RNA was extracted using TRIzol (Life Technologies) and purified with Zymo RNA Clean kit. Viral DNA was recovered with DirectPCR

lysis reagent (Viagen, Cat#101-T). Reporter activity was determined by the ratio of mCherry RNA/DNA.

H3K27ac HiChIP of P0 cardiomyocytes

About 20 heart ventricles of postnatal day 0 mice with mixed C57BL/6x129 background were dissociated to single cells with Neonatal Cardiomyocyte Isolation kit (Cellutron, cat #nc-6031). CMs were purified with the Neonatal Cardiomyocyte Isolation Kit (Miltenyi, cat #130-105-420). Two million cardiomyocytes were used for HiChIP library preparation using the Arima-HiC+ kit (Arima Genomics, cat #201911-792). The full description of this dataset will be reported in a separate manuscript in preparation.

Luciferase reporter assays

About 200 bp of candidate enhancer regions (Table S2) containing wildtype or mutant GATA4 and ETS1 motifs were cloned into pGL3-Promoter vector (Promega) upstream of SV40 minimal promoter (SV40P). Testing plasmid DNA was mixed with internal control pCMV-Renilla plasmid DNA at 5:1 ratio, along with GFP, or GATA4, or Ets1 overexpression plasmid DNA. For transfection of HUVEC cells, HUVEC cells (Lonza) were cultured to 60–80% confluent in 24-well plates and transfected in duplicate with 250 ng pGL3 candidate enhancer construct, 200 ng GFP or Gata4 plasmid, and 50 ng pCMV-Renilla internal control plasmid, using 0.75 µl jetOPTIMUS (Polyplus). After 24 hours, cells were harvested for analysis using the Dual-Glo luciferase assay system (Promega). NIH3T3 cells were transfected using Lipofectamine 3000 (Thermo Fisher). Luciferase activity was measured using a FlexStation 3 Multi-Mode microplate reader (Molecular Devices). Results are representative of two independent experiments.

bioChIP-seq and ChIP-Seq library construction and data analysis

Library construction: bioChIP-seq was performed as described previously [13]. For each bioChIP-seq of GATA4, 15-20 E12.5 heart ventricles were harvested and crosslinked in 1% formaldehyde in PBS for 10 minutes at room temperature. Chromatin was fragmented using a microtip sonicator (Qsonica, S-3000) at 65% amplitude for 8 minutes. $GATA^{fb}$ and bound chromatin was pulled down by incubation with streptavidin beads (ThermoFisher Scientific, Cat#11206D). Libraries were constructed using the KAPA HyperPrep ChIP-seq library preparation kit (Roche, cat# 07962347001). Sequencing (75 nt single end) was performed on an Illumina NextSeq 500.

For ChIP-seq of NIH3T3 cells, $2x10^7$ cells expressing HA-GATA4, Flag-NKX2-5, and Flag-ETS1 were cross-linked with 1% formaldehyde for 10 min at room temperature. The nuclei were sonicated in Sonication Buffer (20 mM Tris Cl pH 8.0, 2 mM EDTA, 150 mM NaCl, 1% NP40, 0.1% SDS, 1× PIC) using a Q800R3 Sonicator (Qsonica) with 65% amplitude for 40 minutes. Sheared DNA was incubated with 10 µg primary antibody (anti-HA, Cat#3724; anti-Flag, Cat#F1804; anti-H3K27ac, Cat#ab4729) at 4 °C overnight and then incubated with BSA-blocked Dynabeads Protein G for one hour. ChIP-seq libraries were prepared and sequenced as described in bioChIP-seq of GATA4.

Alignment and Peak calling: The raw reads from ChIP-Seq and bioChIP-Seq datasets were mapped to the Mus musculus genome (mm10 build) using Bowtie v1.1.1 with the following parameters: m = 1 (i.e. removes all those alignments with more than one match). MACS2 was used to determine peaks in each replicate with p-value 1e-5 as cutoff and ChIP input as the control sample. The peak list was further filtered to remove all blacklisted regions (as defined by ENCODE for mm10 ChIP-Seq: http://mitra.stanford.edu/kundaje/akundaje/release/blacklists/mm10-mouse/).

High confidence GATA4 bound regions in *Gata4$^{flbio/+}$; Rosa26^{fsBirA}; Tie2-Cre* embryos, *Gata4$^{flbio/+}$; Rosa26^{fsBirA}*; Tnnt2-Cre embryos and pan-lineage Gata4 bioChIP-seq E12.5 *GATA4$^{flbio/+}$* embryos (previously published dataset) were generated based on IDR analysis

(https://sites.google.com/site/anshulkundaje/projects/idr). The IDR analysis consists of creating pseudo-replicates from each ChIP-Seq bam file by randomly splitting it into two equal parts. Then the replicates were merged into pooled files by utilizing the merge function in samtools v1.1. Next, peaks were called using MACS2 under low stringency (cutoff P-value: 5e-2) using input as control. Finally, IDR analysis was run on the top 100,000 peaks (sorted by p value) for each cell type and optimal cutoffs (as defined here: https://sites.google.com/site/anshulkundaje/projects/idr) were used to define Gata4 targets in each of the 3 sample types.

Differential Gata4 peaks: The high confidence peaks obtained from running IDR on each experimental group (Tie2-Cre, Tnnt2-Cre and pan-lineage), were used as the starting point. Union of all 3 peak sets was taken and unique regions were determined (i.e. bound by GATA4 in at least 1 of the 3 experiments). Next, locations were filtered to exclude regions with sub-threshold GATA4 signal (Reads per bp ≤ 0.15 in all groups). For each region, log2FC in RPKM-normalized GATA4 signal (Tie2/TnnT2) was calculated, without input subtraction. All GATA4 regions were divided into sextiles based on the endocardial to cardiomyocyte signal ratio. The two extreme sextiles (1st as EC-selective and 6th as CM-selective) and the 3rd and 4th sextiles (shared) were further studied for GO terms and motif analysis. Alternatively, ChIP read counts overlapping the union of peak regions was extracted from Tie2-Cre (EC) or TnnT2-Cre (CM) samples using bedtools. Differential signal between EC or CM samples were identified using DESeq2[50], with lineage-selective regions defined as Padj<0.01 and abs($\log_2$(Tie2/TnT))>=1.

Motif analysis: Motifs were determined by scanning a 200 bp window centered on the summits of peaks for all possible matches to a reference PWM file using the Homer suite of command line tools[51]. The reference PWM file was generated from the union of the default motifs provided with Homer and experimentally determined TF heterodimer motifs defined in ref. [52]. The motifs were then clustered based on their similarity to each other using STAMP[53] and highly redundant motifs were removed. For analysis of central enrichment of motifs within regions, we used Centrimo[54] with default parameters.

Association of peaks to genes: For gene set ontology analysis, peaks were associated with genes using the default rule for GREAT[45]. For differentially expressed genes, peaks were associated with the nearest gene.

Aggregation plots: RPKM-normalized signal files were used for aggregation plots, without input subtraction.

ATAC-seq library construction and data analysis

Library construction: ATAC-seq of E12.5 CMs was reported previously[12]. For ATAC-seq of Tie2-Cre cell lineages, E12.5 Rosa26^{fsTrap};Tie2Cre mouse heart ventricles were dissected in ice cold PBS and dissociated to single cells with Neonatal Cardiomyocyte Isolation kit (Cellutron, cat#nc-6031). 40K GFP+ cells were sorted (BD FACSMelody) for each ATAC-seq library sample. For ATAC-seq of Cdh5 (Pac)-CreERT2 labeled, *Gata4* mutant and control cell lineages, tamoxifen dissolved in sunflower seed oil was administered to pregnant mice at 40 mg/kg body weight by gavage to induce the activity of Cre at E9.5. E12.5 heart ventricles were dissected in ice cold PBS and dissociated to single cells with the Neonatal Cardiomyocyte Isolation kit (Miltenyi). 20K GFP+ cells were sorted (BD FACSMelody) for each ATAC-seq library sample. For ATAC-seq of 3T3 stable cell lines, 50K cells were used for each replicate. The Optimized ATAC-Seq (OmniATAC-seq[55]) libraries were generated from these cells.

Alignment and Peak calling: The raw reads from ATAC-Seq datasets were mapped to the Mus musculus genome (mm10 build) using Bowtie2 v2.2.9[56] with the following parameters: -X = 2000. Samtools v1.2 [57] and removeChrom.py (https://github.com/jsh58/harvard/blob/master/removeChrom.py) were then used to remove all reads aligned to ChrM. Next, picardtools (http://broadinstitute.github.io/picard/) MarkDuplicates was used to remove all PCR duplicates from the Bam files. MACS2 [58]was then run to determine

peaks in each replicate with the following parameters: --keep-dup all --nomodel --nolambda --shift -100 --extsize 200 -B. The resultant peak list was further filtered to remove all blacklisted regions for mm10 as indicated above and also reads mapping to chrM.

Footprinting: py3DNase[25] as implemented in Anaconda3 v5.0.1 was used for carrying out the footprinting analysis in each of the 3 Gata4 peak-sets: EC-selective, CM-selective and shared (as defined by differential ChIP-Seq analysis). These peak-sets were filtered to obtain only accessible peaks, centered at the summits and then scanned in a 200 bp window using the Wellington method implemented as part of py3DNase. The ATAC-Seq bam files used for this analysis were obtained by merging all replicates so as to maximize coverage. The footprints were determined at p-value: 1E-20. The resultant bed files were then used as input for motif analysis to determine GATA4 cofactors.

Differential Footprinting: We performed differential footprinting using HINT-ATAC v0.13.0.[26]. ATAC-seq peaks that overlapped with EC-selective or CM-selective GATA4 regions were analyzed for footprints, using EC and CM ATAC-seq data merged from all replicates. Motifs analyzed were the non-redundant PWMs described above under Motif Analysis.

RNA-Seq library construction and data analysis

Library construction: E12.5 *Rosa26^{fsTrap};Tie2Cre* mouse heart ventricles were dissected in ice cold PBS and dissociated to single cells with Neonatal Cardiomyocyte Isolation kit (Cellutron, cat#nc-6031). 60K GFP+ cells were sorted (BD FACSMelody) for each RNA sample. For E12.5 cardiomyocyte isolation and purification, E12.5 heart ventricles with mixed C57BL/6x129 background were dissociated to single cells with Neonatal Cardiomyocyte Isolation kit (Cellutron, cat #nc-6031). 100K CMs were purified with the Neonatal Cardiomyocyte Isolation Kit (Miltenyi, cat #130-105-420), which depletes non-CMs with magnetic cell separation and antibodies to yield CMs that are >90% pure. RNA was prepared with RNeasy Micro Kit (Qiagen, Cat#74004). The polyadenylated RNA was purified by binding to oligo (dT) magnetic beads (ThermoFisher Scientific, 61005). RNA-seq libraries were prepared with ScriptSeq v2 kit (Illumina, SSV 21106) according to the manufacturer's instructions.

Alignment: Raw sequencing reads were aligned to mm10 using STAR v2.5 aligner [59] and gencode.vM13 annotation files were used as input. The aligned reads were assigned to the reference transcriptome using featureCounts [60]. Read counts were normalized using DESeq2 and expressed as transcripts per million (TPM).

Differential expression analysis: This was carried out using the DESeq2 package [50] implemented in R v3.5.1. The contrast method was used with the cutoffs: p value <= 0.01 and |log2FC| >=2, to identify differentially expressed genes.

Gene set enrichment analysis: GSEA desktop was used with the pre-ranked option[61]. We used custom gene sets from genes whose TSS were either closest to or within 10 kb of lineage-specific GATA4 peaks.

Relationship between lineage-selective GATA4 bioChIP-seq peaks and gene expression in ECs and CMs

A mosaic plot (mosaicplot function in R) was used to visualize the relationship of lineage-biased gene expression in lineage-selective GATA4 ChIP-seq peak regions. GATA4 peak regions were assigned to nearest genes. Cell widths are proportional to the percentage of GATA4 regions. Cell heights are proportional to the percentage of genes. Color shows the chi-square test residuals, with blue and red indicating higher or lower observed-to-expected ratio, respectively. Gene set enrichment analysis (GSEA[61]) was used to measure the enrichment of genes associated with EC- or CM-selective GATA4 peaks and lineage-biased expression.

Analysis of GATA4 pioneering activity

Overlapping regions bound by GATA4, NKX2-5, or ETS1 in different stable cell lines were merged using Homer's mergePeaks function and clustered by presence or absence of transcription factor binding in the different cell lines. Homer was used to determine enrichment of transcription factor binding motifs in each cluster.

Regions were classified by overlap with ATAC-seq peaks in control versus transcription factor-expressing cells. "Pioneer binding" was defined as transcription factor binding to a region that was inaccessible ("closed") in control cells. "Pioneer opening" was defined as transcription factor binding to a region that was closed in control cells and open in the transcription factor-expressing cell. "Enhancer activation" was defined as a region overlapping an H3K27ac peak in transcription-factor expressing cells but not in control cells.

For quantitative comparison of ATAC-seq signal between cell lines, read counts in each region were determined using deepTools and normalized using DESeq2[50].

E12.5 heart dissociation for scRNA-seq

To prepare single cells for RNA-seq, *Gata4$^{fl/+}$;Rosa26$^{mTmG/+}$;Cdh5(Pac)-CreERT2*+ (Gata4_het) or *Gata4$^{fl/fl}$;Rosa26$^{mTmG/+}$;Cdh5(Pac)-CreERT2*+ (Gata4_KO) mouse embryos were induced with tamoxifen at E9.5 and hearts were collected at E12.5. Heart ventricles were dissociated to single cells with collagenase. scRNA-seq libraries were prepared from 10K of total ventricular cells using 10X Chromium Next GEM Single Cell 3' v3 according to the manufacturer's instructions. ~20K paired-end 150bp reads were obtained for each cell.

scRNA-seq data processing

Data preprocessing and quality control

We aligned reads to the mouse reference genome (mm10) and used the 10x Genomics CellRanger pipeline (v6.1.2)[62] with default parameters to obtain the cell-gene expression matrix. We used the following quality control metrics to ensure the high-quality of single-cell profiles. Genes expressed in less than three cells were filtered out. Cells with gene number more than 200, UMI number more than 400 the proportion of reads from either mitochondrial genome or ribosomal DNA is less than 20% were retained. Doublets were detected and removed using DoubletFinder (v2.0.3)[63]. 41,385 high-quality cells were used for further analysis.

Dimensionality reduction, clustering and visualization

Cell-gene expression matrices were loaded in Seurat (v3.1.2)[64]. Read counts were normalized and log-transformed. We selected 2,000 most variable genes as features to perform dimensionality reduction using principal component analysis (PCA). The cells from different conditions (Gata4 KO and control) were integrated using canonical correlation analysis (CCA) with the top 30 PCs. A shared nearest neighbor graph of individual cells was employed for clustering analysis using the FindClusters function. We identified marker genes for each cell cluster using the FindConservedMarkers function. Uniform Manifold Approximation and Projection (UMAP) analysis was performed for visualization.

Differential expression analysis for endocardial cells

We defined ECCs using gene markers *Pecam1*, *Cdh5*, *Npr3* and *GFP*. To remove potential contamination from cardiomyocytes, cells expressing cardiomyocyte marker genes (*Myl2*, *Tnnt2*, and *Myh7*) were excluded. Pseudo-bulk differential expression analysis was performed in ECCs between *Gata4* KO and control groups by aggregating reads counts of cells for each sample, separately. The R package DESeq2[50] was used to identify differentially expressed genes. To link DEGs with ChIP peaks, peaks were assigned to the nearest gene.

Other bioinformatic analyses

Gene ontology term enrichment analysis of genomic regions was performed using GREAT[45]. RPKM-normalized data were visualized in the Integrated Genome Viewer[65]. Deeptools 3.0[66]

was used to generate signal files, perform correlation analysis of signal files across genomic intervals and visualize signals.

Data Access

Data used in this manuscript and the data sources are summarized in Table S1. Original data in this manuscript have been deposited to gene expression omnibus: GSE156001 (GATA4 bioChIP-seq, ATAC-seq, and bulk RNA-seq in E12.5 mouse hearts), GSE155652 (ChIP-seq in NIH3T3 cells). and GSE208162 (scRNA-seq of E12.5 mouse hearts).

Statistics

Statistical analysis was performed using JMP14 (SAS), Graphpad Prism 9, and R. Bar graphs show mean ± SEM. Samples were tested for normality using the Shapiro-Wilk test. Unless otherwise noted, normal data were analyzed using Welch's t-test (2 groups) or ANOVA (>2 groups).

Table S1. Data sources for this study.

Name	Description	Feature	Sample	Method	Useable, uniquely mapped reads	Accession	PMID	
Tie2_GATA4_input	Heart ventricle apex Gata4fb/+;Rosa26fsBirA/+;Tie2Cre, input	Input	E12.5 heart ventricle	Input	21,774,086		This study	
Tie2-GATA4-R1	Heart ventricle apex GATA4 bioChIP-seq in Tie2Cre lineage	GATA4; Tie2Cre lineage	E12.5 heart ventricle	bioChIP-seq	42,848,092		This study	
Tie2-GATA4-R2	Heart ventricle apex GATA4 bioChIP-seq in Tie2Cre lineage	GATA4; Tie2Cre lineage	E12.5 heart ventricle	bioChIP-seq	30,402,997		This study	
Tie2-GATA4-R3	Heart ventricle apex GATA4 bioChIP-seq in Tie2Cre lineage	GATA4; Tie2Cre lineage	E12.5 heart ventricle	bioChIP-seq	41,073,179		This study	
Tnnt2_GATA4_input	Heart ventricle apex Gata4fb/+;Rosa26fsBirA/+;Tnnt2Cre, input	Input	E12.5 heart ventricle	Input	20,361,887		This study	
Tnnt2-GATA4-R1	Heart ventricle apex GATA4 bioChIP-seq in Tnnt2Cre lineage	GATA4; Tnnt2Cre lineage	E12.5 heart ventricle	bioChIP-seq	43,094,068		This study	
Tnnt2-GATA4-R2	Heart ventricle apex GATA4 bioChIP-seq in Tnnt2Cre lineage	GATA4; Tnnt2Cre lineage	E12.5 heart ventricle	bioChIP-seq	26,742,855		This study	
Tnnt2-GATA4-R3	Heart ventricle apex GATA4 bioChIP-seq in Tnnt2Cre lineage	GATA4; Tnnt2Cre lineage	E12.5 heart ventricle	bioChIP-seq	43,195,564		This study	
EC-ATAC-R1	FACS-sorted Tie2Cre; Rosa26-fsTrap/+ cells	Accessibility	E12.5 heart ventricle	ATAC-seq	138,426,277		This study	
EC-ATAC-R2	FACS-sorted Tie2Cre; Rosa26-fsTrap/+ cells	Accessibility	E12.5 heart ventricle	ATAC-seq	153,282,897		This study	
CM-ATAC-R1	MACS-purified CMs	Accessibility	E12.5 heart ventricle	ATAC-seq	170,856,689	GSE124008	31659164	Akerberg et al., Nat. Commun. 2019
CM-ATAC-R2	MACS-purified CMs	Accessibility	E12.5 heart ventricle	ATAC-seq	166,786,289	GSE124008	31659164	Akerberg et al., Nat. Commun. 2019
G4het-EC-ATAC-R2	FACS-sorted Cdh5CreERT2; Rosa26-fsmTmG; GATA4fl/+ cells	Accessibility	E12.5 heart ventricle	ATAC-seq	14,065,071		This study	
G4het-EC-ATAC-R3	FACS-sorted Cdh5CreERT2; Rosa26-fsmTmG; GATA4fl/+ cells	Accessibility	E12.5 heart ventricle	ATAC-seq	15,487,379		This study	
G4het-EC-ATAC-R4	FACS-sorted Cdh5CreERT2; Rosa26-fsmTmG; GATA4fl/+ cells	Accessibility	E12.5 heart ventricle	ATAC-seq	9,477,936		This study	
G4KO-EC-ATAC-R2	FACS-sorted Cdh5CreERT2; Rosa26-fsmTmG; GATA4fl/fl cells	Accessibility	E12.5 heart ventricle	ATAC-seq	14,200,917		This study	
G4KO-EC-ATAC-R3	FACS-sorted Cdh5CreERT2; Rosa26-fsmTmG; GATA4fl/fl cells	Accessibility	E12.5 heart ventricle	ATAC-seq	12,717,875		This study	
G4KO-EC-ATAC-R4	FACS-sorted Cdh5CreERT2; Rosa26-fsmTmG; GATA4fl/fl cells	Accessibility	E12.5 heart ventricle	ATAC-seq	20,145,752		This study	
EC-RNA-seq-R1	FACS-sorted Tie2Cre; Rosa26-fsTrap/+ cells	Transcriptome	E12.5 heart ventricle	RNA-seq	51,701,002		This study	
EC-RNA-seq-R2	FACS-sorted Tie2Cre; Rosa26-fsTrap/+ cells	Transcriptome	E12.5 heart ventricle	RNA-seq	40,048,473		This study	
CM-RNA-seq-R1	MACS-purified CMs	Transcriptome	E12.5 heart ventricle	RNA-seq	103,830,882	GSE124008	31659164	Akerberg et al., Nat. Commun. 2019
CM-RNA-seq-R2	MACS-purified CMs	Transcriptome	E12.5 heart ventricle	RNA-seq	93,115,606	GSE124008	31659164	Akerberg et al., Nat. Commun. 2019
CM-RNA-seq-R3	MACS-purified CMs	Transcriptome	E12.5 heart ventricle	RNA-seq	98,327,031	GSE124008	31659164	Akerberg et al., Nat. Commun. 2019
GATA4_Ht	Heart ventricle apex GATA4 bioChip-seq all lineages	GATA4	E12.5 heart apex	bioChIP-seq		GSE124008	25249388	He et al., N. Comms 2014
EP300_Ht	Heart ventricle apex EP300 bioChip-seq all lineages	EP300	E12.5 heart apex	bioChIP-seq		GSE88789	28121289	Zhou et al., Elife 2017
H3K27ac_Ht	Heart H3K27ac ChIP-seq all lineages	H3K27ac	E12.5 heart	ChIP-seq		GSE88789	28121289	Zhou et al., Elife 2017
H3K27ac_FB	Forebrain H3K27ac ChIP-seq all lineages	H3K27ac	E12.5 forebrain	ChIP-seq		GSE88789	28121289	Zhou et al., Elife 2017
3T3_G_input	Cells: 3T3-G. Assay: input	Input	3T3-G	Input	9,906,450		This study	
G_G-R1	Cells: 3T3-G. Assay: GATA4. Rep 1.	GATA4 (HA antibody)	3T3-G	ChIP-seq	27,005,293		This study	
G_G-R2	Cells: 3T3-G. Assay: GATA4. Rep 2.	GATA4 (HA antibody)	3T3-G	ChIP-seq	8,183,589		This study	
G_G-R3	Cells: 3T3-G. Assay: GATA4. Rep 3.	GATA4 (HA antibody)	3T3-G	ChIP-seq	12,722,069		This study	
GE_G-R1	Cells: 3T3-GE. Assay: GATA4. Rep 1.	GATA4 (HA antibody)	3T3-GE	ChIP-seq	31,727,355		This study	
GE_G-R2	Cells: 3T3-GE. Assay: GATA4. Rep 2.	GATA4 (HA antibody)	3T3-GE	ChIP-seq	17,569,030		This study	
GE_G-R3	Cells: 3T3-GE. Assay: GATA4. Rep 3.	GATA4 (HA antibody)	3T3-GE	ChIP-seq	19,627,463		This study	
GN_G-R1	Cells: 3T3-GN. Assay: GATA4. Rep 1.	GATA4 (HA antibody)	3T3-GN	ChIP-seq	32,299,408		This study	
GN_G-R2	Cells: 3T3-GN. Assay: GATA4. Rep 2.	GATA4 (HA antibody)	3T3-GN	ChIP-seq	12,513,205		This study	
GN_G-R3	Cells: 3T3-GN. Assay: GATA4. Rep 3.	GATA4 (HA antibody)	3T3-GN	ChIP-seq	12,197,726		This study	
E_E-R1	Cells: 3T3-E. Assay: ETS1. Rep 1.	ETS1 (Flag antibody)	3T3-E	ChIP-seq	25,837,952		This study	
E_E-R2	Cells: 3T3-E. Assay: ETS1. Rep 2.	ETS1 (Flag antibody)	3T3-E	ChIP-seq	12,115,843		This study	
E_E-R3	Cells: 3T3-E. Assay: ETS1. Rep 3.	ETS1 (Flag antibody)	3T3-E	ChIP-seq	9,220,347		This study	
GE_E-R1	Cells: 3T3-GE. Assay: ETS1. Rep 1.	ETS1 (Flag antibody)	3T3-E	ChIP-seq	29,044,500		This study	
GE_E-R2	Cells: 3T3-GE. Assay: ETS1. Rep 2.	ETS1 (Flag antibody)	3T3-E	ChIP-seq	27,352,867		This study	
GE_E-R3	Cells: 3T3-GE. Assay: ETS1. Rep 3.	ETS1 (Flag antibody)	3T3-E	ChIP-seq	31,780,539		This study	
N_N-R1	Cells: 3T3-N. Assay: NKX2-5. Rep1.	NKX2-5 (Flag antibody)	3T3-N	ChIP-seq	16,099,624		This study	
N_N-R2	Cells: 3T3-N. Assay: NKX2-5. Rep2.	NKX2-5 (Flag antibody)	3T3-N	ChIP-seq	7,849,198		This study	
N_N-R3	Cells: 3T3-N. Assay: NKX2-5. Rep3.	NKX2-5 (Flag antibody)	3T3-N	ChIP-seq	8,711,730		This study	
GN_N-R1	Cells: 3T3-GN. Assay: NKX2-5. Rep1.	NKX2-5 (Flag antibody)	3T3-GN	ChIP-seq	25,330,619		This study	
GN_N-R2	Cells: 3T3-GN. Assay: NKX2-5. Rep2.	NKX2-5 (Flag antibody)	3T3-GN	ChIP-seq	12,909,242		This study	
GN_N-R3	Cells: 3T3-GN. Assay: NKX2-5. Rep3.	NKX2-5 (Flag antibody)	3T3-GN	ChIP-seq	12,717,592		This study	
Ctrl_ATAC-R1	Cells: 3T3-GFP. Assay: ATAC. Rep 1.	Accessibility	3T3-control	ATAC-seq	6,581,603		This study	
Ctrl_ATAC-R2	Cells: 3T3-GFP. Assay: ATAC. Rep 1.	Accessibility	3T3-control	ATAC-seq	6,735,160		This study	
Ctrl_ATAC-R3	Cells: 3T3-GFP. Assay: ATAC. Rep 1.	Accessibility	3T3-control	ATAC-seq	8,493,543		This study	
G_ATAC-R1	Cells: 3T3-G. Assay: ATAC. Rep 1.	Accessibility	3T3-G	ATAC-seq	14,007,022		This study	
G_ATAC-R2	Cells: 3T3-G. Assay: ATAC. Rep 2.	Accessibility	3T3-G	ATAC-seq	17,791,218		This study	
G_ATAC-R3	Cells: 3T3-G. Assay: ATAC. Rep 3.	Accessibility	3T3-G	ATAC-seq	19,870,031		This study	
GE_ATAC-R1	Cells: 3T3-GE. Assay: ATAC. Rep 1.	Accessibility	3T3-GE	ATAC-seq	12,851,191		This study	
GE_ATAC-R2	Cells: 3T3-GE. Assay: ATAC. Rep 2.	Accessibility	3T3-GE	ATAC-seq	8,686,404		This study	
GE_ATAC-R3	Cells: 3T3-GE. Assay: ATAC. Rep 3.	Accessibility	3T3-GE	ATAC-seq	9,501,436		This study	
GN_ATAC-R1	Cells: 3T3-GN. Assay: ATAC. Rep 1.	Accessibility	3T3-GN	ATAC-seq	10,519,324		This study	
GN_ATAC-R2	Cells: 3T3-GN. Assay: ATAC. Rep 2.	Accessibility	3T3-GN	ATAC-seq	16,342,561		This study	
GN_ATAC-R3	Cells: 3T3-GN. Assay: ATAC. Rep 3.	Accessibility	3T3-GN	ATAC-seq	8,887,725		This study	
E_ATAC-R1	Cells: 3T3-E. Assay: ATAC. Rep 1.	Accessibility	3T3-E	ATAC-seq	9,160,620		This study	
E_ATAC-R2	Cells: 3T3-E. Assay: ATAC. Rep 2.	Accessibility	3T3-E	ATAC-seq	8,692,877		This study	
E_ATAC-R3	Cells: 3T3-E. Assay: ATAC. Rep 3.	Accessibility	3T3-E	ATAC-seq	6,919,757		This study	
N_ATAC-R1	Cells: 3T3-N. Assay: ATAC. Rep 1.	Accessibility	3T3-N	ATAC-seq	12,949,690		This study	
N_ATAC-R2	Cells: 3T3-N. Assay: ATAC. Rep 2.	Accessibility	3T3-N	ATAC-seq	16,072,649		This study	
N_ATAC-R3	Cells: 3T3-N. Assay: ATAC. Rep 3.	Accessibility	3T3-N	ATAC-seq	14,317,400		This study	
Ctrl_ATAC-R1	Cells: 3T3-GFP. Assay: H3K27ac. Rep 1.	H3K27ac	3T3-control	ChIP-seq	22,565,261		This study	
Ctrl_ATAC-R2	Cells: 3T3-GFP. Assay: H3K27ac. Rep 1.	H3K27ac	3T3-control	ChIP-seq	14,047,211		This study	
Ctrl_ATAC-R3	Cells: 3T3-GFP. Assay: H3K27ac. Rep 1.	H3K27ac	3T3-control	ChIP-seq	23,510,029		This study	
G_H3K27ac-R1	Cells: 3T3-G. Assay: H3K27ac. Rep 1.	H3K27ac	3T3-G	ChIP-seq	41,262,722		This study	
G_H3K27ac-R2	Cells: 3T3-G. Assay: H3K27ac. Rep 2.	H3K27ac	3T3-G	ChIP-seq	17,058,862		This study	
G_H3K27ac-R3	Cells: 3T3-G. Assay: H3K27ac. Rep 3.	H3K27ac	3T3-G	ChIP-seq	16,962,594		This study	
GE_H3K27ac-R1	Cells: 3T3-GE. Assay: H3K27ac. Rep 1.	H3K27ac	3T3-GE	ChIP-seq	34,709,725		This study	
GE_H3K27ac-R2	Cells: 3T3-GE. Assay: H3K27ac. Rep 2.	H3K27ac	3T3-GE	ChIP-seq	18,975,536		This study	
GE_H3K27ac-R3	Cells: 3T3-GE. Assay: H3K27ac. Rep 3.	H3K27ac	3T3-GE	ChIP-seq	25,289,486		This study	
GN_H3K27ac-R1	Cells: 3T3-GN. Assay: H3K27ac. Rep 1.	H3K27ac	3T3-GN	ChIP-seq	34,215,422		This study	
GN_H3K27ac-R2	Cells: 3T3-GN. Assay: H3K27ac. Rep 2.	H3K27ac	3T3-GN	ChIP-seq	17,581,582		This study	
GN_H3K27ac-R3	Cells: 3T3-GN. Assay: H3K27ac. Rep 3.	H3K27ac	3T3-GN	ChIP-seq	20,336,410		This study	

Name	Description	Feature	Sample	Method	Pass QC cells	PMID
Gata4 EC KO vs control scNRA-seq	10x genomics scRNAseq of GATA4 endothelial KO vs control E12.5 heart ventricle	scRNAseq	Het 1	scRNAseq	19,239	This study
Gata4 EC KO vs control scNRA-seq	10x genomics scRNAseq of GATA4 endothelial KO vs control E12.5 heart ventricle	scRNAseq	Het 2	scRNAseq	7,257	This study
Gata4 EC KO vs control scNRA-seq	10x genomics scRNAseq of GATA4 endothelial KO vs control E12.5 heart ventricle	scRNAseq	KO 1	scRNAseq	10,940	This study
Gata4 EC KO vs control scNRA-seq	10x genomics scRNAseq of GATA4 endothelial KO vs control E12.5 heart ventricle	scRNAseq	KO 2	scRNAseq	3,949	This study

Table 2. Oligonucleotides and synthesized sequences used in this study.		
Name	**Sequence (5'->3')**	
Vegfa_+23kb_Wild-type	GACTGTGGTTGGGGATTTGGGACCTCAGAATCCAGAAATAGCAGCCAGGATTTGAACTTGGGCAGTCAGGTGAGCAGGGCCAGAGTGCCCCTCACCTTAT CAGTTCAGTCCCTTCCAGATAGAGGGTTATTAATAGCCACGAGGTGTTTGTTAGCAAGCAGTCAGCAACTCCCAGAGGGCGGCATGGGGTCATGAGACTT CAGCCAGGTCAGCCCCCACCCCCAGCCCCATGGGGGCTCGTCCAAGAAGTTGGCTCCCAGCTGGGCCCCATCCTGCCCCTGGGAGCCCTTACTCAACCAG GCAGCCACCCCGTGACTAGGAAGAATGACCCCTGTCACACTTGTGTGCTTGCGAGCTCGTG	mm10, chr17:46009110-46009470
Vegfa_+23kb_GATA4_mut	GACTGTGGTTGGGGATTTGGGACCTCAGAATCCAGAAATAGCAGCCAGGATTTGAACTTGGGCAGTCAGGTGAGCAGGGCCAGAGTGCCCCTCACCTTAC TAGTTCAGTCCCTTCCAGATAGAGGGTTATTAATAGCCACGAGGTGTTTGTTAGCAAGCAGTCAGCAACTCCCAGAGGGCGGCATGGGGTCATGAGACTT CAGCCAGGTCAGCCCCCACCCCCAGCCCCATGGGGGCTCGTCCAAGAAGTTGGCTCCCAGCTGGGCCCCATCCTGCCCCTGGGAGCCCTTACTCAACCAG GCAGCCACCCCGTGACTAGGAAGAATGACCCCTGTCACACTTGTGTGCTTGCGAGCTCGTG	
Vegfa_+23kb_MEF2_mut	GACTGTGGTTGGGGATTTGGGACCTCAGAATCCAGAGAACTAAGCCAGGATTTGAACTTGGGCAGTCAGGTGAGCAGGGCCAGAGTGCCCCTCACCTTAT CAGTTCAGTCCCTTCCAGATAGAGGGTTATTAATAGCCACGAGGTGTTTGTTAGCAAGCAGTCAGCAACTCCCAGAGGGCGGCATGGGGTCATGAGACTT CAGCCAGGTCAGCCCCCACCCCCAGCCCCATGGGGGCTCGTCCAAGAAGTTGGCTCCCAGCTGGGCCCCATCCTGCCCCTGGGAGCCCTTACTCAACCAG GCAGCCACCCCGTGACTAGGAAGAATGACCCCTGTCACACTTGTGTGCTTGCGAGCTCGTG	
Vegfa_+23kb_NKX_mut	GACTGTGGTTGGGGATTTGGGACCTCAGAATCCAGAAATAGCAGCCAGGATTTGAACTTGGGCAGTCAGGTGAGCAGGGCCGAACTGCCCCTCACCTTAT CAGTTCAGTCCCTTCCAGATAGAGGGTTATTAATAGCCACGAGGTGTTTGTTAGCAAGCAGTCAGCAACTCCCAGAGGGCGGCATGGGGTCATGAGACTT CAGCCAGGTCAGCCCCCACCCCCAGCCCCATGGGGGCTCGTCCAAGAAGTTGGCTCCCAGCTGGGCCCCATCCTGCCCCTGGGAGCCCTTACTCAACCAG GCAGCCACCCCGTGACTAGGAAGAATGACCCCTGTCAACGTTGTGTGCTTGCGAGCTCGTG	
Vegfa_+23kb_TBX_mut	GACTGTGGTTGGGGATTTGGGACCTCAGAATCCAGAAATAGCAGCCAGGATTTGAACTTGGGCAGTCAGTGAGGCAGGGCCAGAGTGCCCCCCTACTTAT CAGTTCAGTCCCTTCCAGATAGAGGGTTATTAATAGCCACGACTGTGTTGTTAGCAAGCAGTCAGCAACTCCCAGAGGGCGGCATGGGGTCATGAGACTT CAGCCAGGTCAGCCCCCACCCCCAGCCCCATGGGGGCTCGTCCAAGAAGTTGGCTCCCAGCTGGGCCCCATCCTGCCCCTGGGAGCCCTTACTCAACCAG GCAGCCACCCCGTGACTAGGAAGAATGACCCCTGTCACACTTGTGTGCTTGCGAGCTCGTG	
Vegfa_+23kb_GMNT_mut	GACTGTGGTTGGGGATTTGGGACCTCAGAATCCAGAGAACTAAGCCAGGATTTGAACTTGGGCAGTCAGTGAGGCAGGGCCGAACTGCCCCCCTACTTAC TAGTTCAGTCCCTTCCAGATAGAGGGTTATTAATAGCCACGACTGTGTTGTTAGCAAGCAGTCAGCAACTCCCAGAGGGCGGCATGGGGTCATGAGACTT CAGCCAGGTCAGCCCCCACCCCCAGCCCCATGGGGGCTCGTCCAAGAAGTTGGCTCCCAGCTGGGCCCCATCCTGCCCCTGGGAGCCCTTACTCAACCAG GCAGCCACCCCGTGACTAGGAAGAATGACCCCTGTCAACGTTGTGTGCTTGCGAGCTCGTG	
mCherry_RT1	TCGTCCATGCCGCCGGTGGAGT	RT primer to generate cDNA of mCherry from viral RNA
mCherry-qf1	ACGGCCACGAGTTTGAGATT	qPCR primers for qRT-PCR of mCherry
mCherry-qr1	CAAGTAGTCGGGGATGTCGG	qPCR primers for qRT-PCR of mCherry
mCherry-qf2	TGAAGGGCGAGATCAAGCAG	qPCR primers for qRT-PCR of mCherry
mCherry-qr2	GGTGTAGTCCTCGTTGTGGG	qPCR primers for qRT-PCR of mCherry
Eln_+8kb_Enh	CTCAGAGACTCAGTGACAATACACAGCTGGGCCTGCCTGGTCACTGGGCCTTGTCAGCCCAGCAGGGTTGGAAATATCTGGCCGACATGGGTAGATTGTG CTAAAGTCCTGGCGCCACCACCTCCTGGGGTCTGTGCCCCCCTCCTCTCCTGACCCTCTAGAGACTACATGCCAGGCCCGGATGGTCTGAGCAATCTGTT CAGAGCTTGGGGGATTCAGGGAGGGGTGACTCTGCAGG	mm10, chr5:134740814-134741051
Eln_+8kb_Enh_mut	CTCAGAGACTCAGTGACAATACACAGCTGGGCCTGCCTGGTCACTGGGCCTTGTCAGCCCAGCAGGGTTGAGATATACTTGGCCGACATGGGTAGATTGT GCTAAAGTCCTGGCGCCACCACCTCCTGGGGTCTGTGCCCCCCTCCTCTCCTGACCCTCTAGAGACTACATGCCAGGCCCGGATGGTCTGAGCAATCTGT TCAGAGCTTGGGGGATTCAGGGAGGGGTGACTCTGCAGG	GATA and ETS motifs mutated
Pdgfra_-14kb_Enh	TGAGATTAGAGATAGGGCCTCAGTTCCCAAACAGACGGTCTGAACGAAGCCCCTGCTTTGATCTGCACACAAAAGAAGTGACCTGAAATAACATGGCGGT GGCTAATCAGGTATTTCCCATTGTTCTCTGCTTGTCCCTATCTTAGCAGGAAAGTGACTTTGAAATGACCGCTCTGTT	mm10, chr5:75137936-75138113
Pdgfra_-14kb_Enh_mut	TGAGATTAGAAGTAGGGCCTCAGCTCTCAAACAGACGGTCTGAACGAAGCCCCTGCTTTGATCTGCACACAAAAGAAGTGACCTGAAATAACATGGCGGT GGCTAATCAGGTATAAGGCATTGTTCTCTGCTTGTCCCTACTTTAGCAAGAGAGTGACTTTGAAATGACCGCTCTGTT	GATA and ETS motifs mutated
Pdgfra_+58kb_Enh	GTGGGAAGTAGACATTTTCCTCCTGTTGTGCTAATGTTTACTTAGTGGGCCTGGCTTATCTCTAGTGCATCATTACACCAGTCCTTTAAACTCCTCTCCA TTGTTTAAGGGAAGTCCCAGTATAGCTCCCAGTCTTCAGCTCCCTGAAGGTCACCAATGCAATGCTCAAGATGGGTGGAGA	mm10, chr5:75210154-75210334
Pdgfra_+58kb_Enh_mut	GTGAGAGGTAGACATTCTCTCTGTTGTGCTAATGTTTACTTAGTGGGCCTGGCTACTTTCTAGTGCATCATTACACCAGTCCTTTAAACTCCTCTCCA TTGTTTAAGAGAGGTCCCAGTATAGCTCCCAGTCTTCAGCTCCCTGAAGGTCACCAATGCAATGCTCAAGATGGGTGGAGA	GATA and ETS motifs mutated

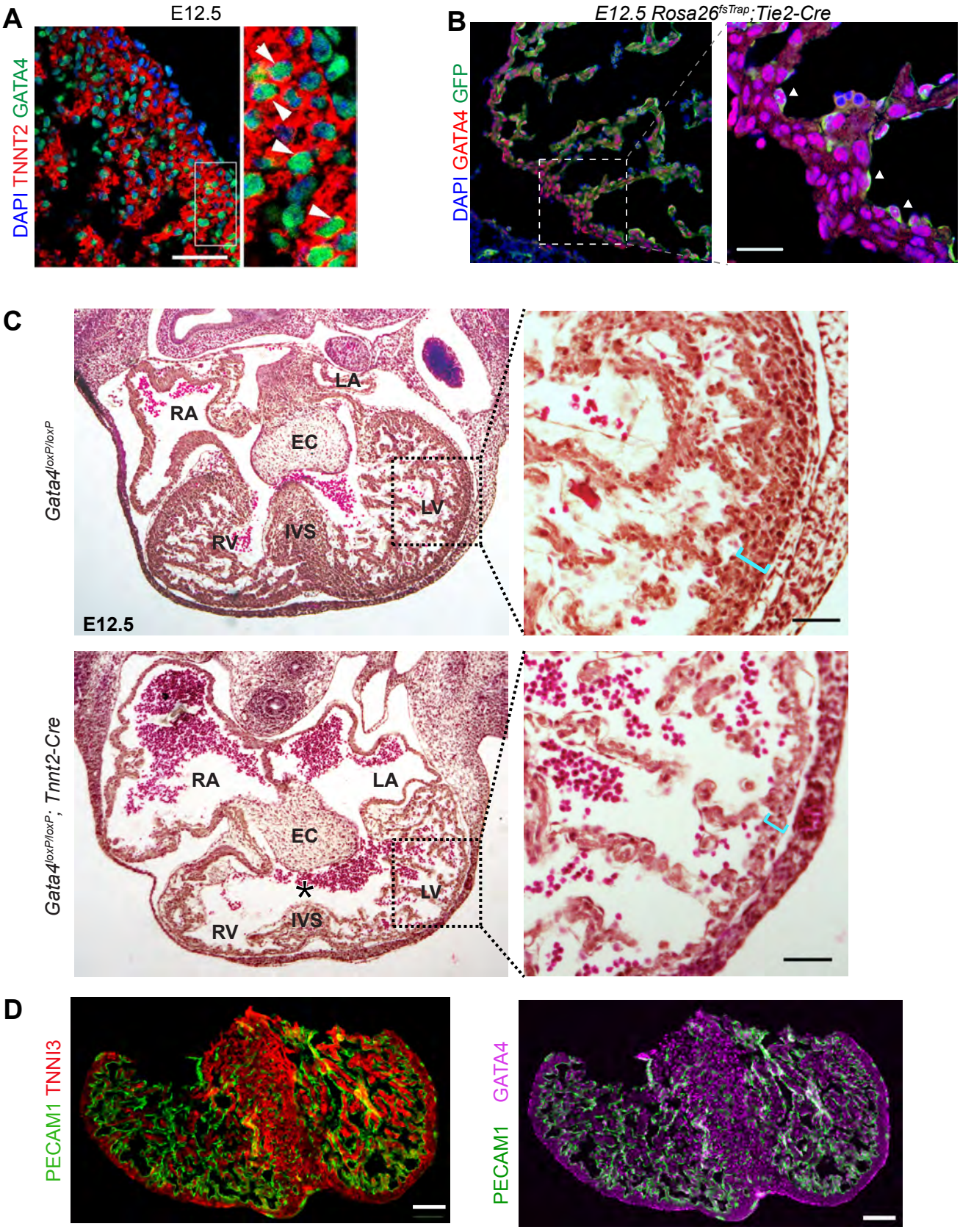


Fig. S1. Gata4 cardiac expression and requirement in cardiomyocytes. A. E12.5 heart ventricle section stained for GATA4 and TNNT2, a cardiomyocyte marker. Bar = 50 μM. **B.** E12.5 heart ventricle section stained for GATA4. GFP was expressed from the Rosa26^{fsTRAP} allele, activated by Tie2Cre in endothelial cells. Bar = 50 μM. **C.** Gata4^{fl} ablation by Tnnt2-Cre impaired heart development. Control (top row) and mutant (bottom row) E12.5 embryo transverse sections through the heart were stained with H&E. Note thinned ventricular wall (blue lines), deficient interventricular septum (asterisk), and reduced trabeculation of mutant hearts. Endocardial cushion development was not affected. RV, right ventricle; LV, left ventricle; RA, right atrium; LA, left atrium; IVS, interventricular septum; EC, endocardial cushion. Bar = 50 μM. **D**. Representative example of E12.5 heart ventricle sample used for GATA4 bioChIP-seq. Heart sections were stained with CM marker TNNI3, endocardial/endothelial marker PECAM1, and GATA4. Bar = 100 μM.

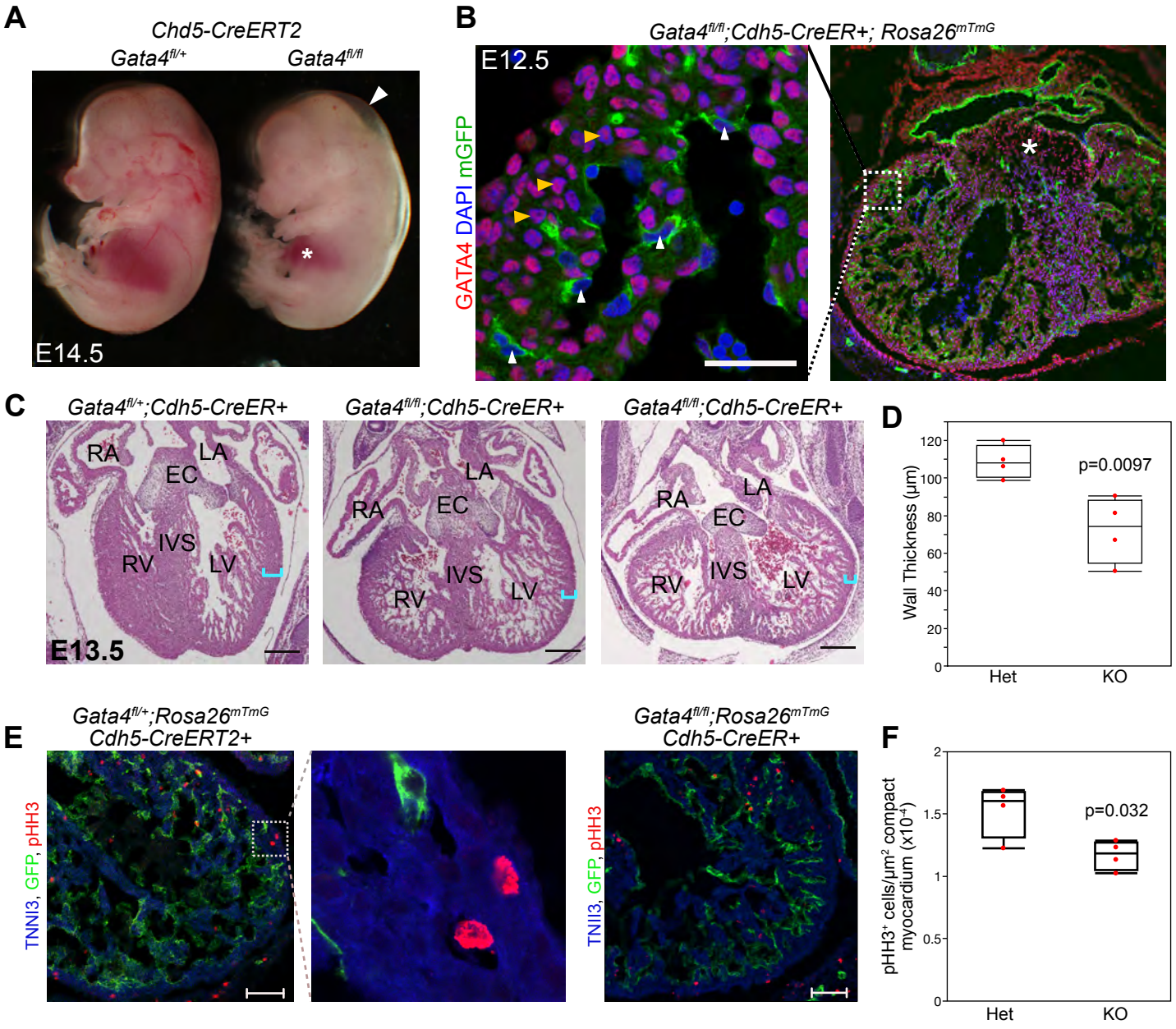


Fig. S2. *Gata4* is required in endocardium for normal heart development. Gata4$^{fl/+}$; Cdh5-CreERT2 and Gata4$^{fl/fl}$; Cdh5-CreERT2 were treated with tamoxifen at E9.5. **A.** Embryos died at around E14.5 with peripheral edema (arrowhead), hypoplastic liver spot (asterisk), and pallor. **B.** Immunostaining at E12.5 demonstrated loss of GATA4 in endocardial cells (white arrowheads). GATA4 immunoreactivity in endocardial cushion mesenchyme was not affected (orange arrowheads). Bar = 50 μM. **C.** Transverse E13.5 embryo sections showing thinning of the compact myocardium (blue lines). RA, right atrium. LA, left atrium. RV, right ventricle. LV, left ventricle. IVS, interventricular septum. EC, endocardial cushion. Blue brackets indicate the width of the LV free wall compact myocardium. Bar = 200 μM. **D.** Average width of the compact myocardium was calculated as the compact myocardial area divided by its outer circumference. 3 to 5 sections were analyzed for each of 4 embryos per group. The average width of mutant compact myocardium was significantly reduced compared to littermate controls. Welch's *t*-test: p=0.01. **E.** Cardiomyocyte proliferation following endocardial *Gata4* ablation. Proliferation was measured by staining for phosphohistone H3 (pHH3), an M-phase marker. Bar = 100 μM. **F.** Quantification of pHH3+ cardiomyocytes. Only pHH3+ (red) nuclei within in TNNI3+ compact myocardium (blue) were counted. The area of compact myocardium was measured with ImageJ. Welch's *t*-test.

EC-selective GATA4 regions

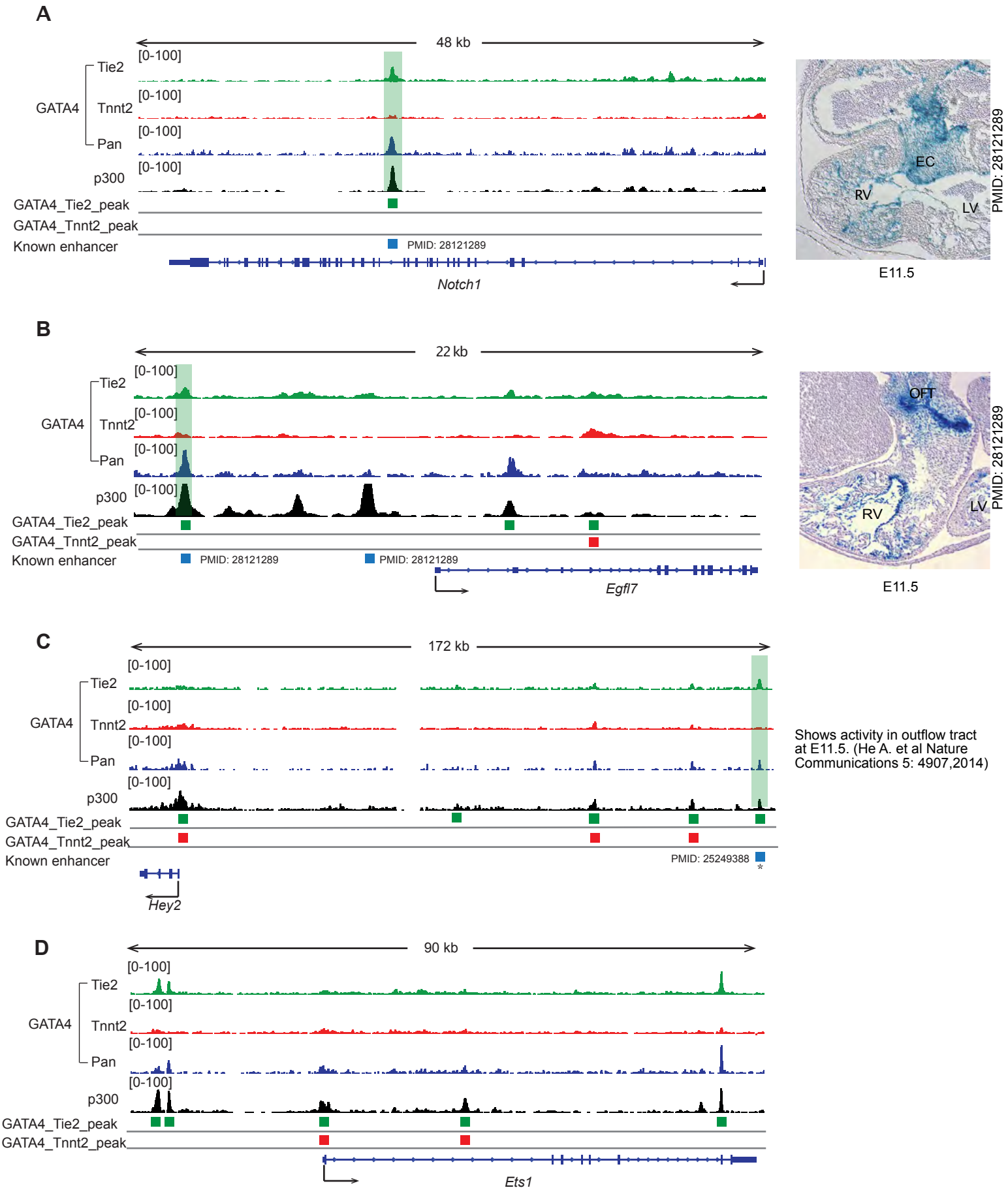


Fig. S3. Examples of EC-selective (A-H) or CM-selective (I-K) GATA4 regions. Genome browser views show GATA4 bioChIP-seq signals and peak calls. Green shading highlights selected regions previously studied in transient transgenic assays. *(continued)*

EC-selective GATA4 regions

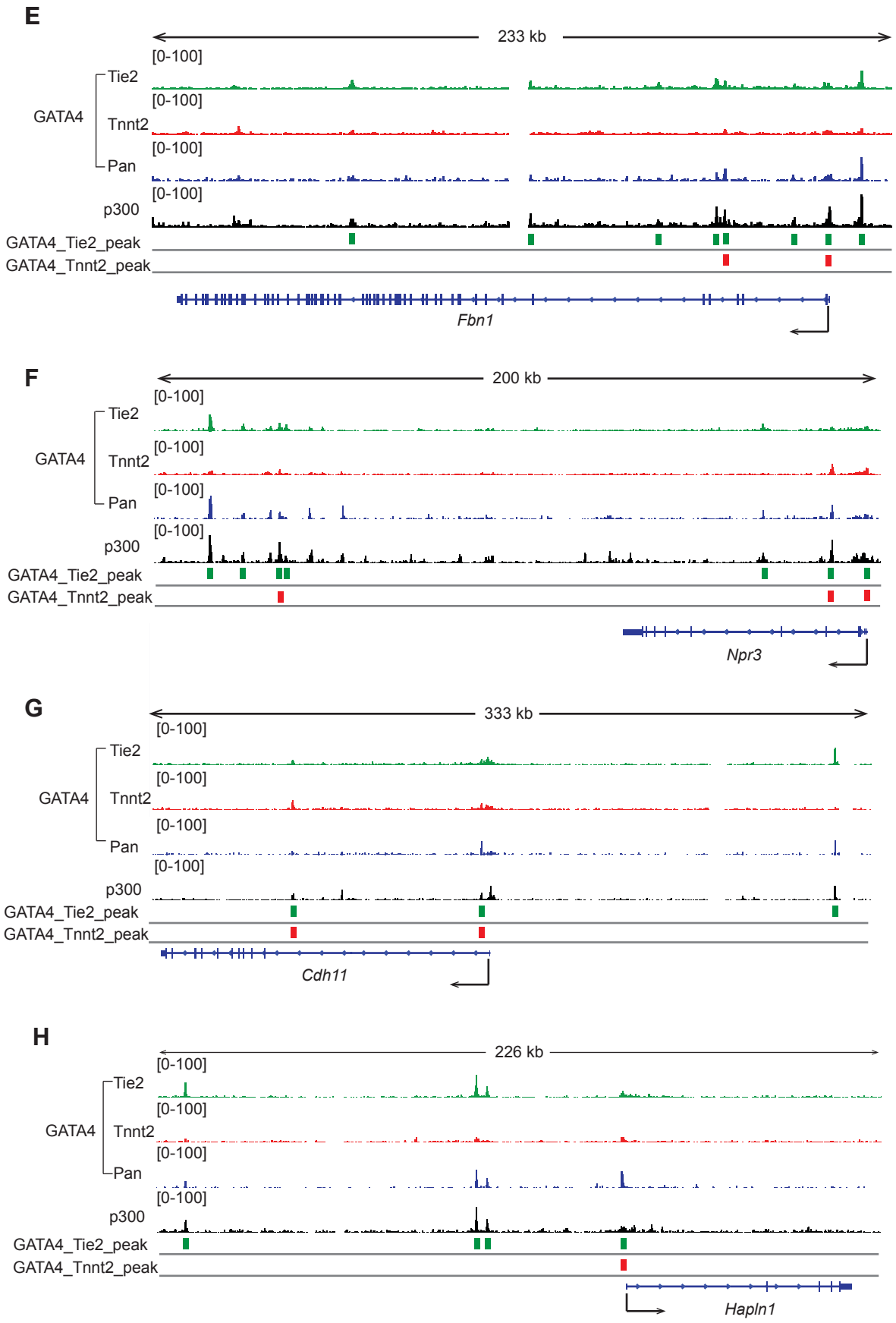


Fig. S3, continued.

CM-selective GATA4 regions

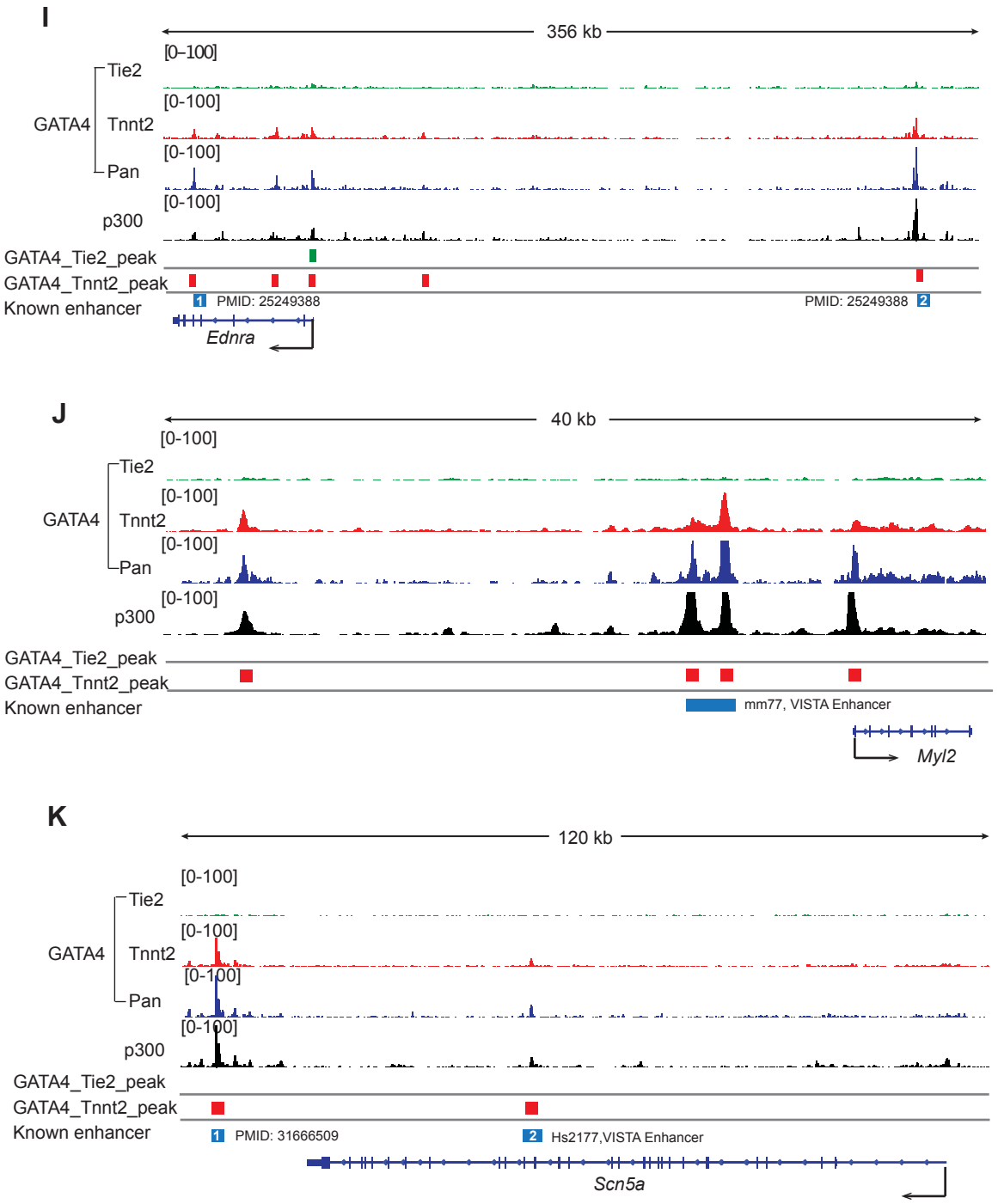


Fig. S3, continued.

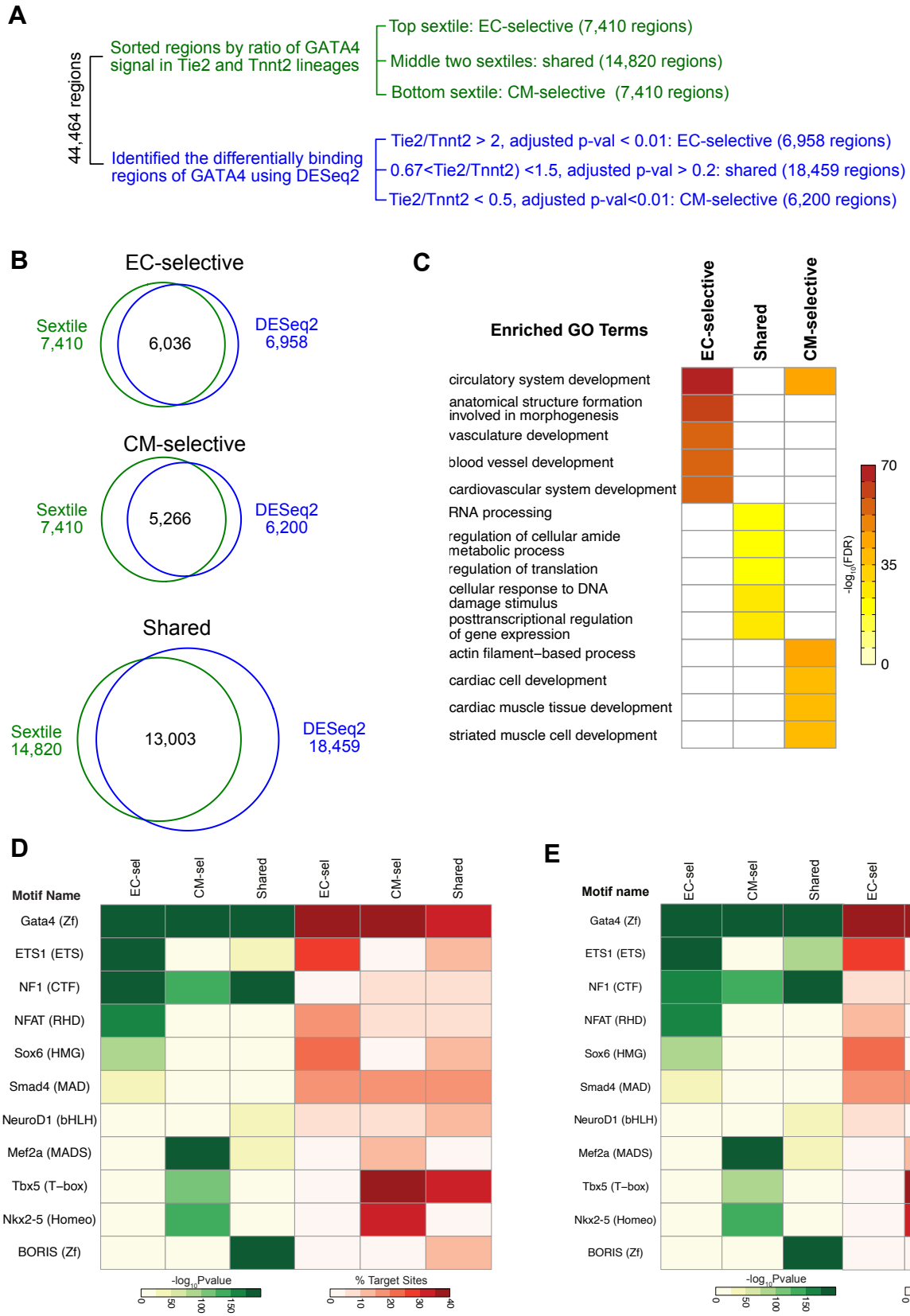


Fig. S4. Lineage-selective GATA4 regions. A. We used two different methods to define lineage-selective GATA4 regions. The first method ranked regions by the average ratio of GATA4 signal between reproducible peaks in Tie2-Cre (EC) and Tnnt2-Cre (CM) lineages ("sextile" method). The second region statistically compared GATA4 signal between reproducible peaks in Tie2-Cre and Tnnt2-Cre, using DESeq2 ("DESeq2" method). **B.** The lineage-selective regions defined by the two methods were highly overlapped. For the sextile method, shared regions were defined as the middle two sextiles, whereas for the DESeq2 method they were defined as regions that were not lineage-selective. **C.** The top 5 Gene Ontology terms among EC-selective, shared, and CM-selective GATA4 regions defined DESeq2. The GO terms were similar to those obtained using the sextile method (Fig. 2C). **D-E.** Enriched TF motifs in lineage-selective or shared GATA4 regions defined by the sextile (D) or DESeq2 (E) methods. The motifs identified by each method were very similar.

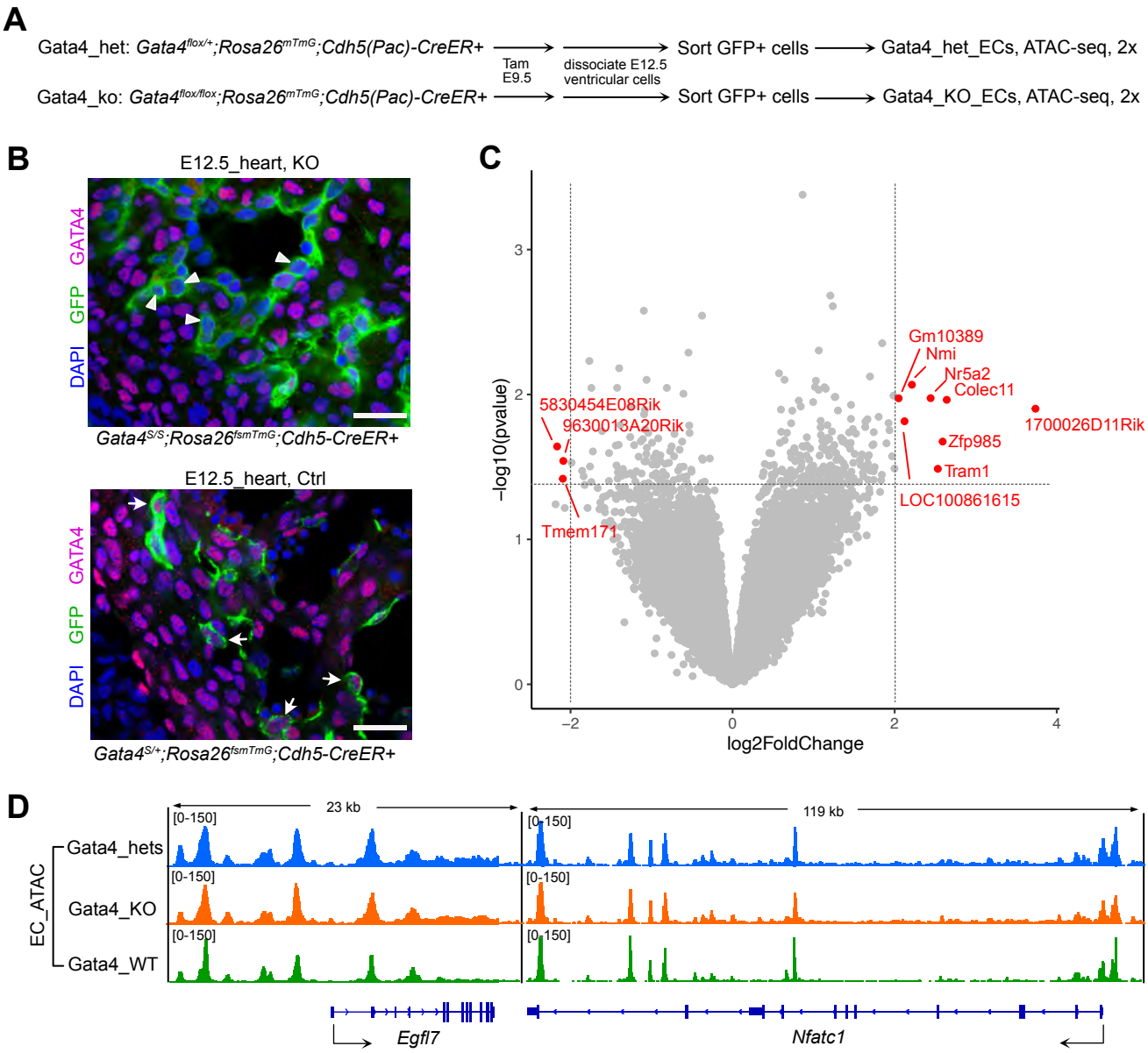
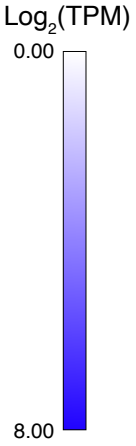


Fig. S5. GATA4 is not required to maintain chromatin accessibility.

A. Schematic for ATAC-seq measurement of chromatin accessibility in control and Gata4_KO EC lineages. Endothelial cell inactivation of *Gata4* by administration of tamoxifen to *Cdh5-CreER;Gata4$^{fl/fl}$;Rosa26^{mTmG}* or *Cdh5-CreERT2;Gata4$^{fl/+}$;Rosa26^{mTmG}* embryos at E9.5. E12.5 EC cells were isolated for ATAC-seq by FACS. **B**. Immunostaining at E12.5 demonstrated loss of GATA4 in endocardial cells (arrowheads) of Gata4_KO (Top), compared to expression of GATA4 in control endocardial cells (arrows) of *Gata4_Ctrl* (Bottom). Bar = 50 μM. **C**. Volcano plot shows inactivation of *Gata4* rarely changed chromatin accessibility. Red points indicate nominal P<0.05 and absolute fold-change >2. Points are labeled with the nearest gene. **D**. Representative genome browser views of ATAC-seq signal in E12.5 ECs of *Gata4_WT*, *Gata4_het*, or *Gata4_KO* hearts.

A



EC	CM	Gene	Family	EC vs CM	GATA4 occupancy
		Ehf	ETS	not DE	EC 100k
		Elf1	ETS	not DE	
		Elf2	ETS	not DE	EC 10k
		Elf3	ETS	not DE	EC 100k
		Elf4	ETS	not DE	CM 10k
		Elk1	ETS	not DE	
		Elk3	ETS	DE	EC 100k
		Elk4	ETS	not DE	EC 100k
		Erf	ETS	DE	
		Erg	ETS	DE	CM 100k
		Ets1	ETS	DE	EC 100k
		Ets2	ETS	DE	
		Etv1	ETS	not DE	
		Etv2	ETS	DE	EC 10k
		Etv3	ETS	not DE	
		Etv4	ETS	not DE	
		Etv5	ETS	not DE	CM 100k
		Etv6	ETS	not DE	CM 100k
		Fev	ETS	DE	EC 10k
		Fli1	ETS	DE	EC 100k
		Gabpa	ETS	not DE	EC 100k
		Spdef	ETS	not DE	
		Spi1	ETS	DE	
		Gata1	GATA	not DE	
		Gata2	GATA	DE	
		Gata3	GATA	DE	
		Gata4	GATA	not DE	
		Gata5	GATA	DE	CM 100k
		Gata6	GATA	not DE	
		Mef2a	MEF2	not DE	CM 10k
		Mef2c	MEF2	not DE	CM 100k
		Mef2d	MEF2	not DE	CM 10k
		Nfat5	NFAT	not DE	EC 10k
		Nfatc1	NFAT	DE	EC 10k
		Nfatc2	NFAT	not DE	EC 10k
		Nfatc3	NFAT	not DE	
		Nfatc4	NFAT	DE	EC 10k
		Nkx2-5	NKX	DE	CM 10k
		Sox2	SOX	DE	EC 100k
		Sox3	SOX	DE	CM 100k
		Sox4	SOX	DE	
		Sox5	SOX	DE	
		Sox6	SOX	not DE	
		Sox7	SOX	DE	EC 100k
		Sox8	SOX	not DE	
		Sox9	SOX	DE	EC 100k
		Sox10	SOX	DE	CM 100k
		Sox11	SOX	not DE	
		Sox12	SOX	DE	
		Sox13	SOX	DE	EC 100k
		Sox15	SOX	not DE	
		Sox17	SOX	not DE	
		Sox18	SOX	DE	EC 10k
		Sox30	SOX	not DE	EC 10k
		Srf	SRF	not DE	CM 10k
		Tbx2	TBX	DE	
		Tbx3	TBX	not DE	CM 100k
		Tbx4	TBX	DE	CM 100k
		Tbx5	TBX	not DE	CM 10k
		Tbx15	TBX	not DE	EC 100k
		Tbx18	TBX	DE	CM 100k
		Tbx20	TBX	not DE	EC 100k
		Tead1	TEAD	not DE	
		Tead2	TEAD	not DE	
		Tead3	TEAD	not DE	
		Tead4	TEAD	not DE	

B

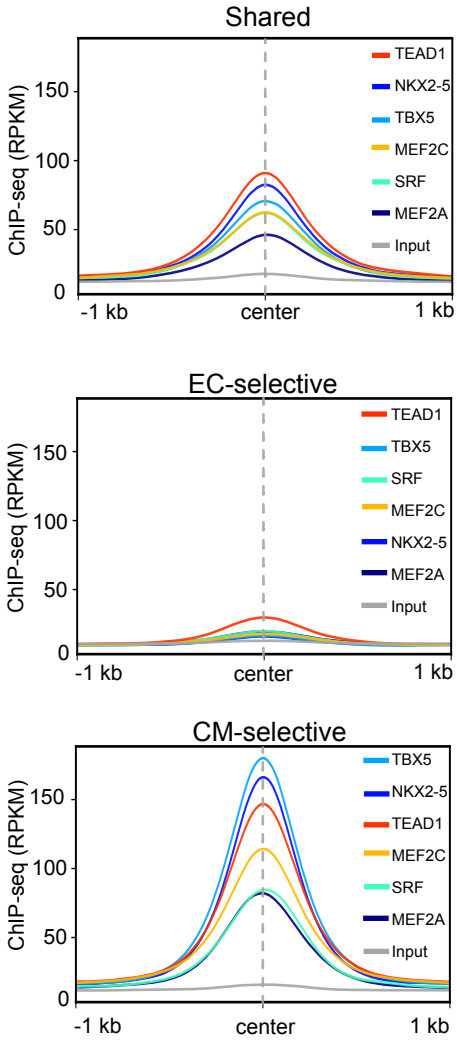


Fig. S6. Expression and occupancy of adjacent sequences by GATA4 for selected TFs in E12.5 heart. **A**. Expression of selected TFs based on RNA-seq of ECs and CMs isolated from E12.5 heart. The table also lists whether the TF was differentially expressed (DE) between these cell types (absolute log2 fold-change > 2 and adjusted P-value < 0.01), and whether the TF is within 10 kb or 100 kb of an EC-selective or CM-selective GATA4 region. Blanks indicate either no adjacent GATA4 region, or adjacency to both EC-selective and CM-selective GATA4 regions. Among ETS family members, ETS1 and ETS2 were the most highly expressed in ECs. **B**. Aggregation plots of E12.5 cardiac TF bioChIP-seq signal at GATA4 regions. CM-selective and Shared but not EC-selective GATA4 regions had significant NKX2-5, TEAD1, TBX5 and MEF2C signal in heart.

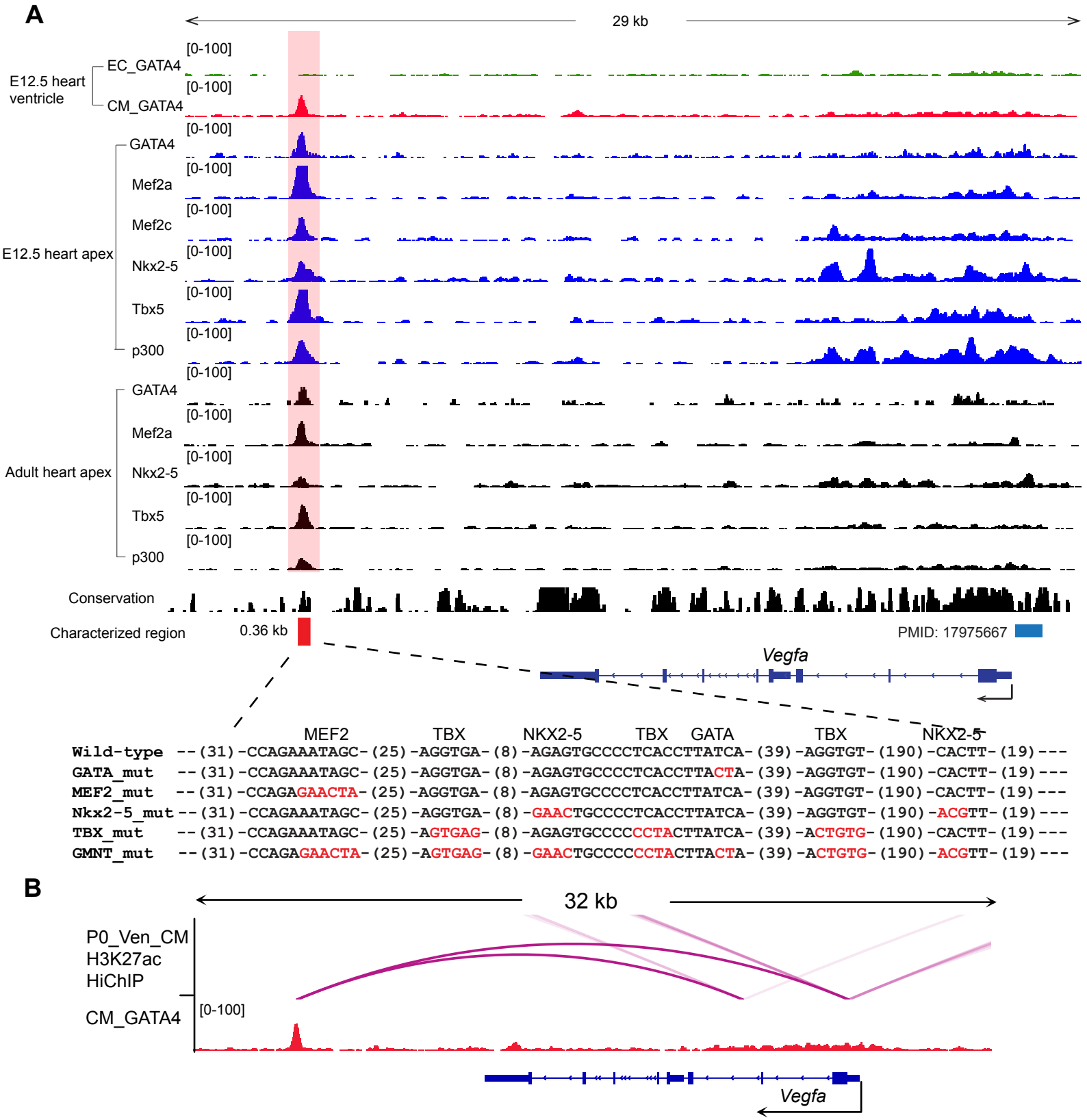
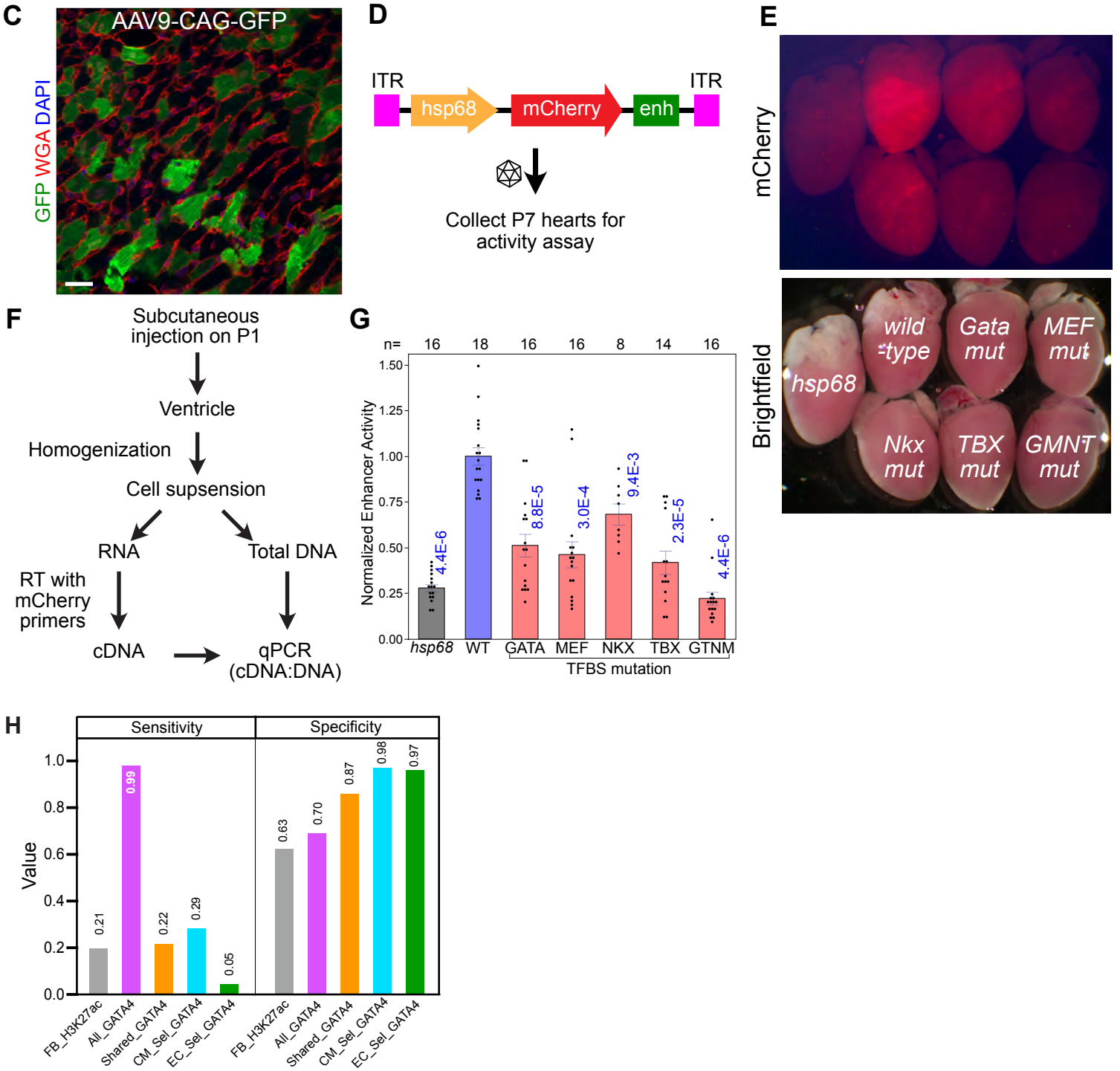


Fig. S7. Characterization of a CM-selective *Vegfa* enhancer.
A. A cardiomyocyte-selective GATA4 peak 23 kb downstream of the *Vegfa* TSS contains motifs for multiple cardiac TFs. Enhancers containing the wild-type candidate enhancer, or mutants with selected TF motifs ablated by nucleotide substitutions (red) were tested for cardiomyocyte enhancer activity. ChIP-seq data from whole E12.5 or adult heart apex is also shown, from He et al., 2014, Zhou et al., 2017, and Akerberg et al., 2019. **B**. H3K27ac HiCHIP identifies chromatin loops between GATA4 -23 kb peak and promoter of *Vegfa* gene. IGV genome browser view of H3K27ac HiCHIP in P0 ventricular CMs and GATA4 bioChIP-seq signal in E12.5 CMs.

(continued)



(Fig. S7, continued.) C. AAV9 seletively transduces CMs in the mouse heart. AAV-CAG-GFP was delivered to neonatal mice. Heart sections show that GFP was confined to cardiomyocytes. WGA, wheat germ agglutinin. Bar, 20 μm. **D.** Schematic diagram of AAV-based reporter assay to measure the activity of putative enhancers. **E.** Vegfa reporter activity. AAV was injected systemically at P1, and mCherry fluorescence was evaluated from whole hearts at P7. Tested enhancers are labeled on the matching whole mount image. **F**. Method to measure normalized enhancer activity as the ratio of mCherry RNA to mCherry DNA. **G.** Relative activity of wild-type and mutant *Vegfa* enhancers. Bars show mean ± SEM. Blue numbers indicate adjusted P-values from Steel multiple non-parametric comparison vs wild type control are shown above bars. Sample size is indicated by numbers above graph. TFBS, transcription factor binding site. **H.** In vivo activity characteristics of ineage-selective GATA4 bioChIP-seq peak regions. Specificity and sensitivity of GATA4 region classes was evaluated using 1371 murine regions tested for heart enhancer activity in the Vista Enhancer Database (Nucleic Acids Res 35:D88-92). E12.5 forebrain (FB) H3K27ac, all GATA4 ChIP-seq peak regions in ECs and CMs, Shared_GATA4 regions, CM_selective regions, and EC_selective regions in Fig. 2A were tested. Out of 175 regions with activity in the developing heart, All_GATA4 identified 173 regions but with modest specificity, CM_selective_GATA4 identified 51 regions but had superb specificity. As expected, EC-selective regions predicted few heart enhancers. Sensitivity=TP/(TP+FN). Specificity,=TN/(TN+FP). TP: true positive; FN: false positive; TN: true negative; FN: false negative.

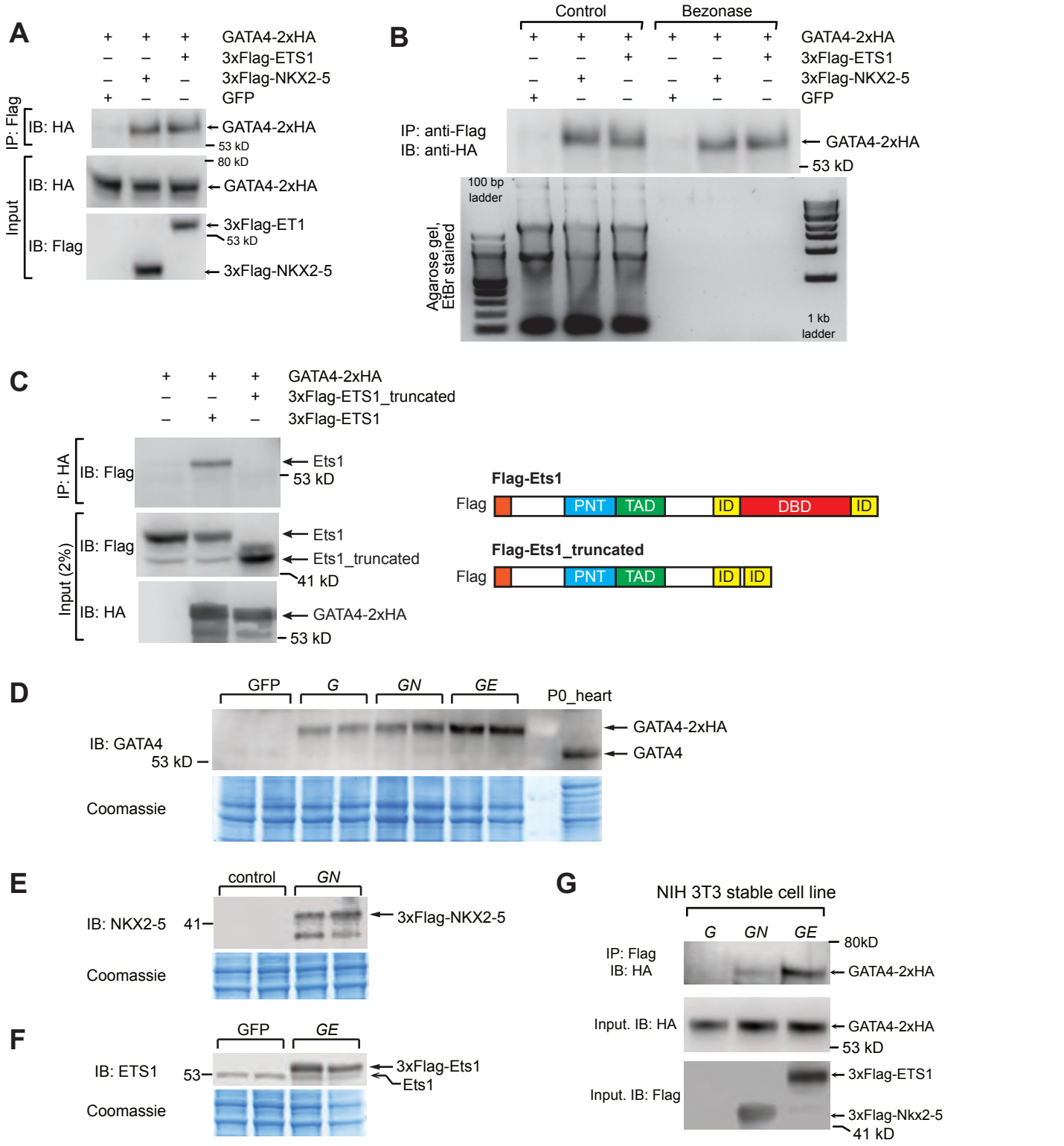


Fig. S8. GATA4-ETS1 and GATA4-NKX2-5 interactions. A. GATA4 interacts with either NKX2-5 or ETS1. 293T cells were transiently transfected with indicated expression plasmids. Pull down of Flag-tagged NKX2-5 or ETS1 co-precipitated approximatey similar amounts of GATA4. **B.** Interaction between GATA4 and ETS1 is not mediated by DNA or RNA. Co-IP in 293T cells was not impaired by treatment with benzonase, a nuclease. DNA degradation was confirmed by ethidium bromide staining of samples run on an agarose gel (bottom). **C.** The ETS DNA-binding domain is essential for GATA4 binding. Deletion of the ETS DNA binding domain (DBD; red, right) abolished its interaction with GATA4. Right panel shows ETS1 domain structure and deletion mutant lacking DNA-binding domain. PNT, pointed domain. TAD, transcriptional activation domain. ID, inhibitor domain. **D-F.** Characterization of protein expression in NIH3T3 and NIH3T3 stable cell lines. Stable NIH3T3 cell lines expressing HA-tagged GATA4, or Flag-tagged ETS1 or NKX2-5. Coomassie stained gels were used for loading control. GATA4 expression was comparable to endogenous levels in neonatal heart. There was no detectable endogenous expression of GATA4 or NKX2-5, but there was low endogenous ETS1 expression. **G.** NKX2-5 or ETS1 co-precipitated GATA4 from NIH3T3 stable cell lines.

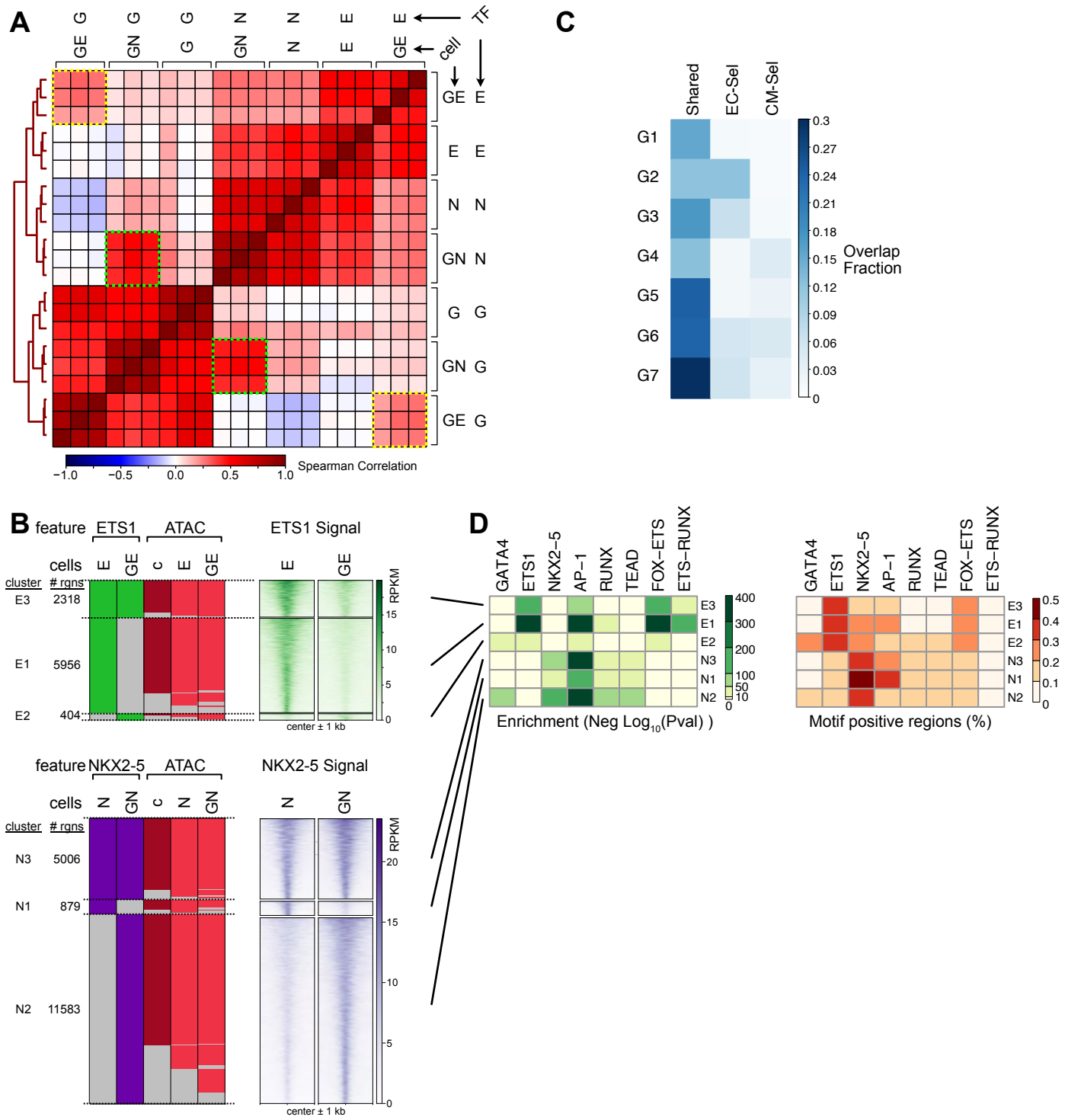


Fig. S9. Chromatin occupancy by GATA4, ETS1, and NKX2-5 stably expressed in NIH3T3 fibroblasts. **A.** Correlation of ChIP-seq signals between groups and replicates. The correlation matrix shows the Spearman correlation of signals at the union of peaks identified in the indicated stable cell lines. The green and yellow boxes highlight correlation between GATA4 and NKX2-5 in 3T3-GN cells and GATA4 and ETS1 in 3T3-GE cells, respectively. **B.** ETS1 and NKX2-5 binding clusters. The union of regions occupied by ETS1 is shown, clustered by ETS1 occupancy in 3T3-E, 3T3-GE, or both. Overlap with ATAC-seq peaks is also shown. Color=positive regions. Grey=negative regions. Heatmaps of ETS1 or NKX2-5 signal in each region are shown to the right. **C.** Pairwise overlap of GATA4 region sets in E12.5 heart ventricles and GATA4 region sets in 3T3 stable cell lines. Note that the regions bound by GATA4 in the presence of ETS1 in NIH3T3 cells (G2, G3, G6) overlapped most with EC-selective regions in the embryo, whereas regions bound by GATA4 plus NKX2-5 in NIH3T3 cells (G4, G5, G6) overlapped most with CM-selective regions in the embryo. **D.** Motif analysis of regions occupied by ETS1 and NKX2-5. The ETS1 motif, and the FOX-ETS and ETS-RUNX heterodimer motifs, were enriched in ETS1-expressing cells, and NKX2-5 was enriched in NKX2-5 expressing cells. In the regions dependent on GATA4 for occupancy (E2; N2), the GATA4 motif was enriched.

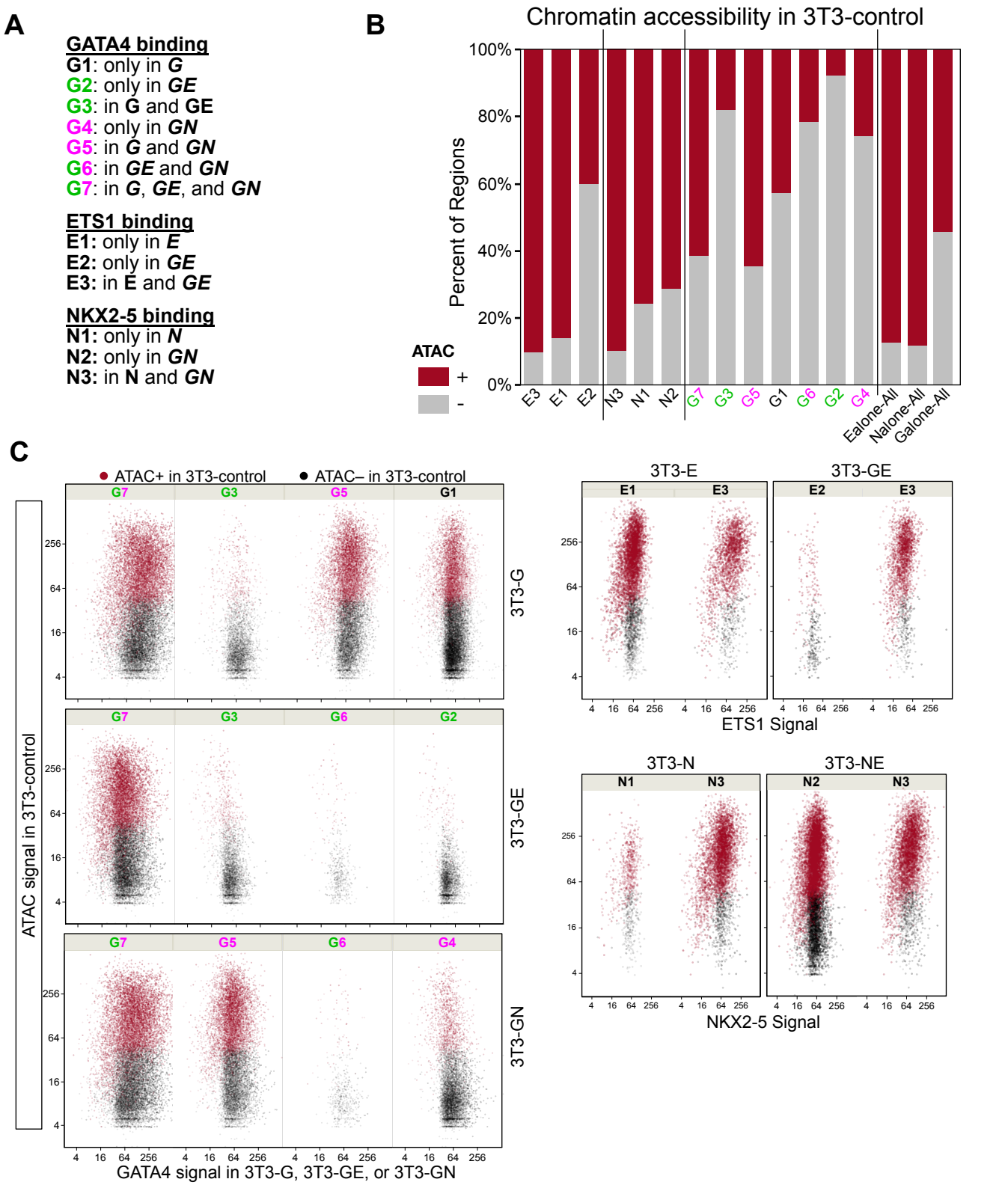


Fig. S10. Influence of ETS1 or NKX2-5 on GATA4 pioneer binding, pioneer opening, and pioneer enhancer activation. A. Categories of GATA4-bound regions. **B.** Pioneer binding activity of GATA4 compared to ETS1 and NKX2-5. At regions bound by ETS1 and NKX2-5 independent of GATA4 (E3, E1; N3, N1), chromatin was predominantly accessible in control cells. In contrast, a large fraction of regions bound by GATA4, or GATA4 in combination with ETS1 or NKX2-5, were inaccessible in control cells. Ealone-All, Galone-All, and Nalone-All refer to all of the regions occupied by ETS1, NKX2-5, or GATA4 alone. **C**. ATAC signal in 3T3-control cells for regions bound by GATA4, ETS1, or NKX2-5 in the indicated cell lines. GATA4-bound regions (G1-G7; E2; N2) contained a higher proportion of regions with low accessibility in 3T3-control cells, consistent with GATA4 having pioneer binding activity. Red and black color indicate classification of regions into accessible and inaccessible in 3T3-control cells by MACS2.

(Continued)

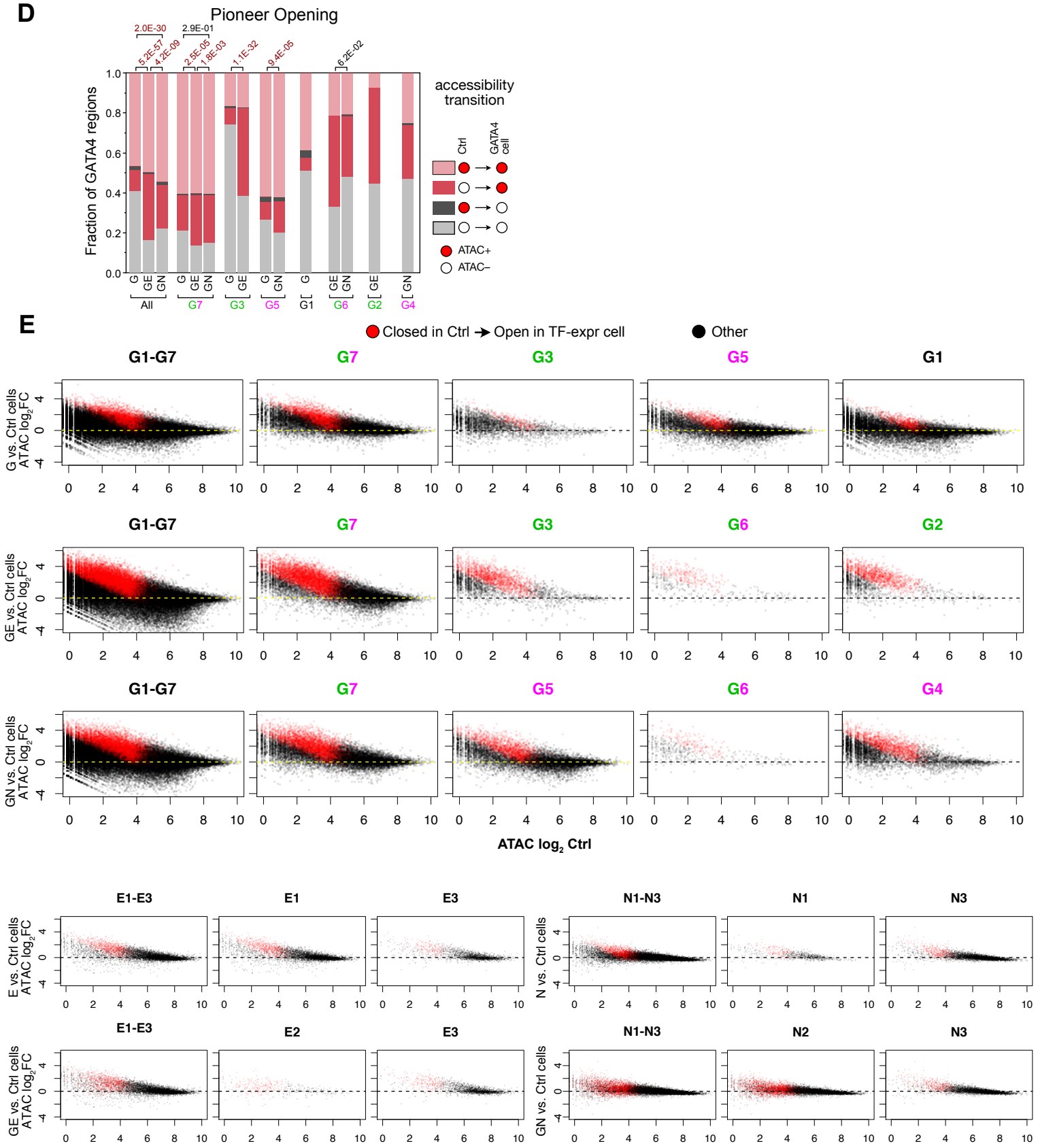


Fig. S10, Continued

D. Change in accessibility with GATA4 occupancy. The distribution of four possible chromatin accessibility transitions is plotted for the GATA4 clusters relevant for each GATA-expressing cell line, compared to the control (GATA4-naive) cell line. The fraction of closed regions in control cells that became open in the GATA4 cells was compared using the Fisher exact test. Nominal Fisher exact p-value between pink and red bars is shown above. **E.** Scatterplot of change in chromatin accessibility between control cells and cells expressing TF combinations as indicated. Regions that were closed in control cells and open in TF expressing cells ("closed to open") are colored red. ATAC-seq signals were normalized between groups using DESeq2. Closed to open regions, which had low signal in control cells and greatest fold-change in TF expressing cells, were most frequent in GE cells and had an intermediate frequency in GN cells. Closed to open regions were not often observed for E or N cells, except in clusters E2 or N2, which dependent upon interaction with GATA4. (Continued)

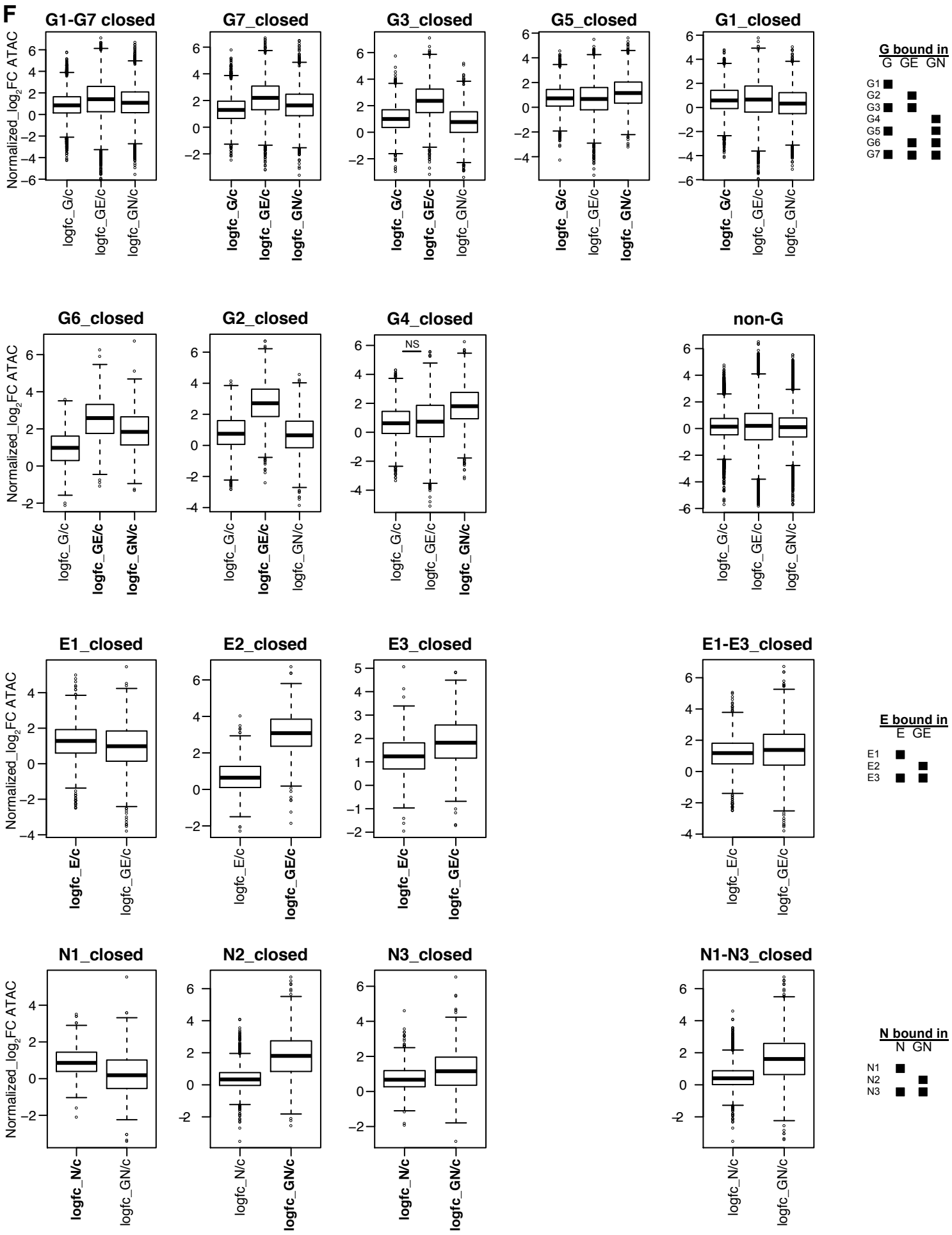


Fig. S10, Continued

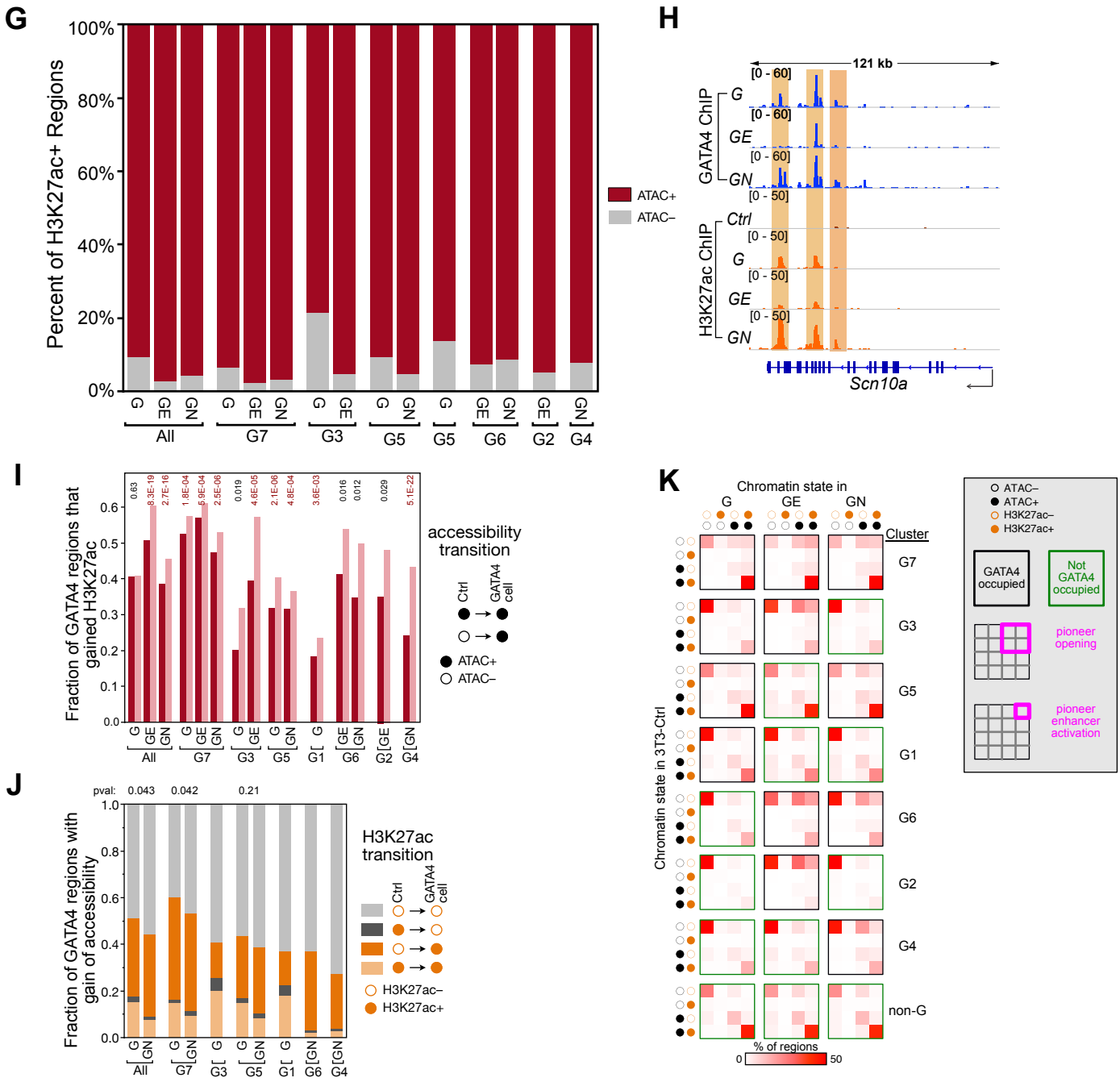


Fig. S10, continued.
F. Change in chromatin accessibility between control cells and cells expressing TF combinations as indicated. Regions were grouped by their pattern of GATA4, ETS1, or NKX2-5 binding, as summarized in the legends to the right. Only regions classified as closed (ATAC negative) in control cells were included. ATAC-seq signals were normalized between groups using DESeq2. Notice that GATA4 plus ETS1 consistently increased ATAC-seq signal more than GATA4 plus NKX2-5 or GATA4 alone, within clusters where these comparisons were relevant (bold labels). All pairwise comparisons were significantly different (Mann-Whitney P<0.001) except as noted. **G**. Fraction of H3K27ac+ regions in the indicated cell line that were ATAC+ or ATAC–. The large majority of H3K27ac+ regions were ATAC+. **H.** Representative genome browser views illustrating GATA4 pioneer enhancer activation (orange shading). **I.** Establishment of enhancers at persistently open versus newly opened in GATA4 expressing cells. Of GATA4 regions that were persistently accessible (pink) or transitioned from closed to open (red), the fraction that gained H3K27ac is plotted. Persistently open regions more frequently gained H3K27ac than regions that were opened by pioneer activity. Nominal Fisher exact p-value between pink and red bars is shown above. **J.** The distribution of H3K27ac transitions for regions that gained accessibility is shown. Fisher p-value indicates comparison between G and GN of the proportion of GATA4 regions with gain of accessibility that undergo enhancer activation. **K.** Summary of chromatin state transitions. Within a cell line, a region can have one of 4 chromatin states depending on whether it is open (ATAC+) or closed (ATAC–), and marked by H3K27ac or not. Between control cells and TF-expressing cells, there are 16 possible chromatin state transitions, summarized for each GATA4 cluster of regions as a 4x4 heatmap representing the frequency of each observed transition. Transitions defined as pioneer opening and pioneer enhancer activation are indicated in purple in the schematics to the right. Green outlines indicate regions not bound by GATA4 in the corresponding cell line and cluster. Non-G are the union of regions with ATAC or H3K27ac signal but no GATA4 occupancy.

A

	Passed QC cells (#)	Mean reads per cell	Median gene number per cell	Median UMI counts per cell	Correlation score between replicates
Gata4_Het_1	19,239	24,066	1,622	4,143	0.98
Gata4_Het_2	7,257	42,344	1,740	10,605	
Gata4_KO_1	10,940	34,778	1,272	5,122	0.98
Gata4_KO_2	3,949	55,480	1,586	6,770	

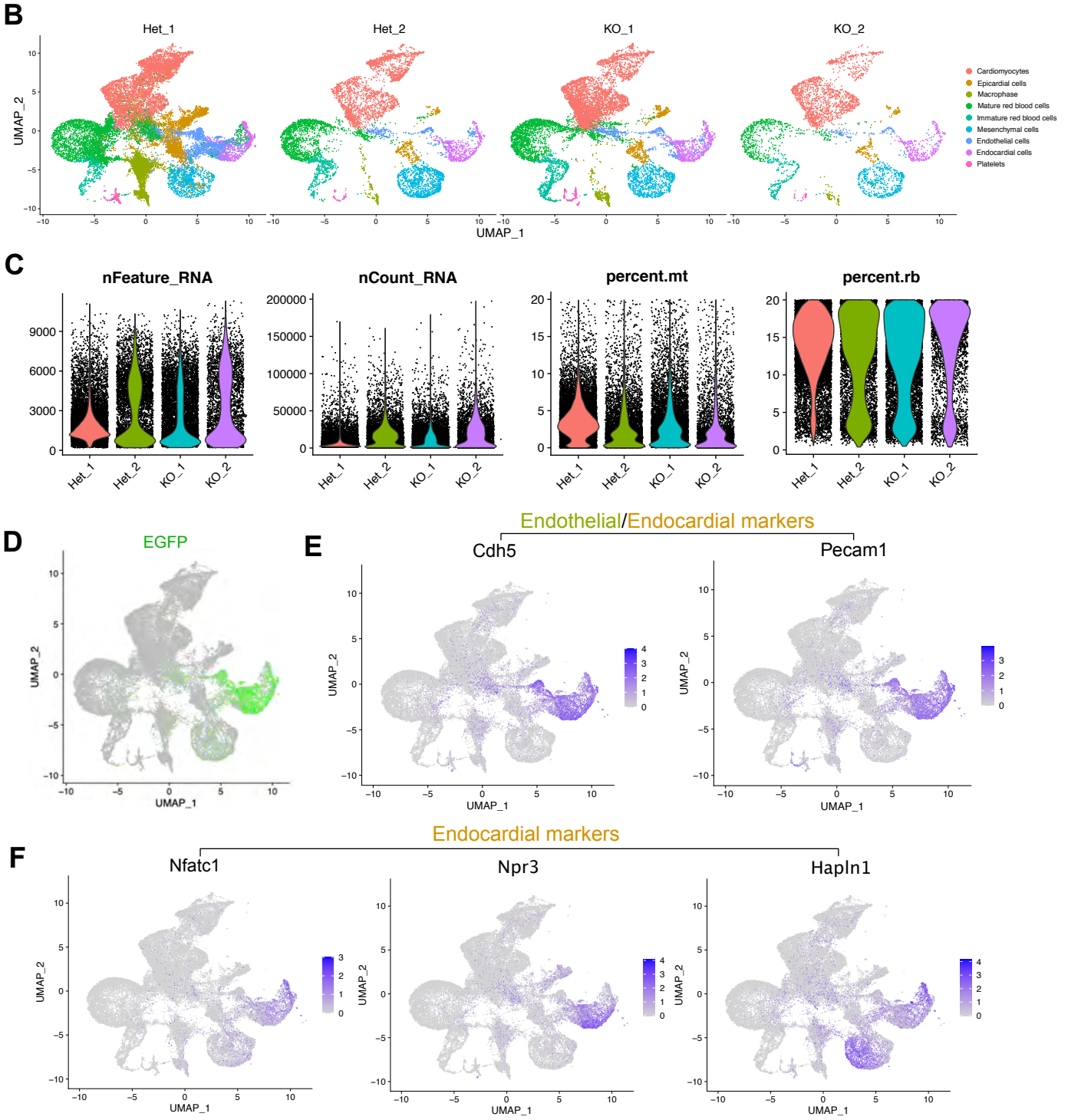
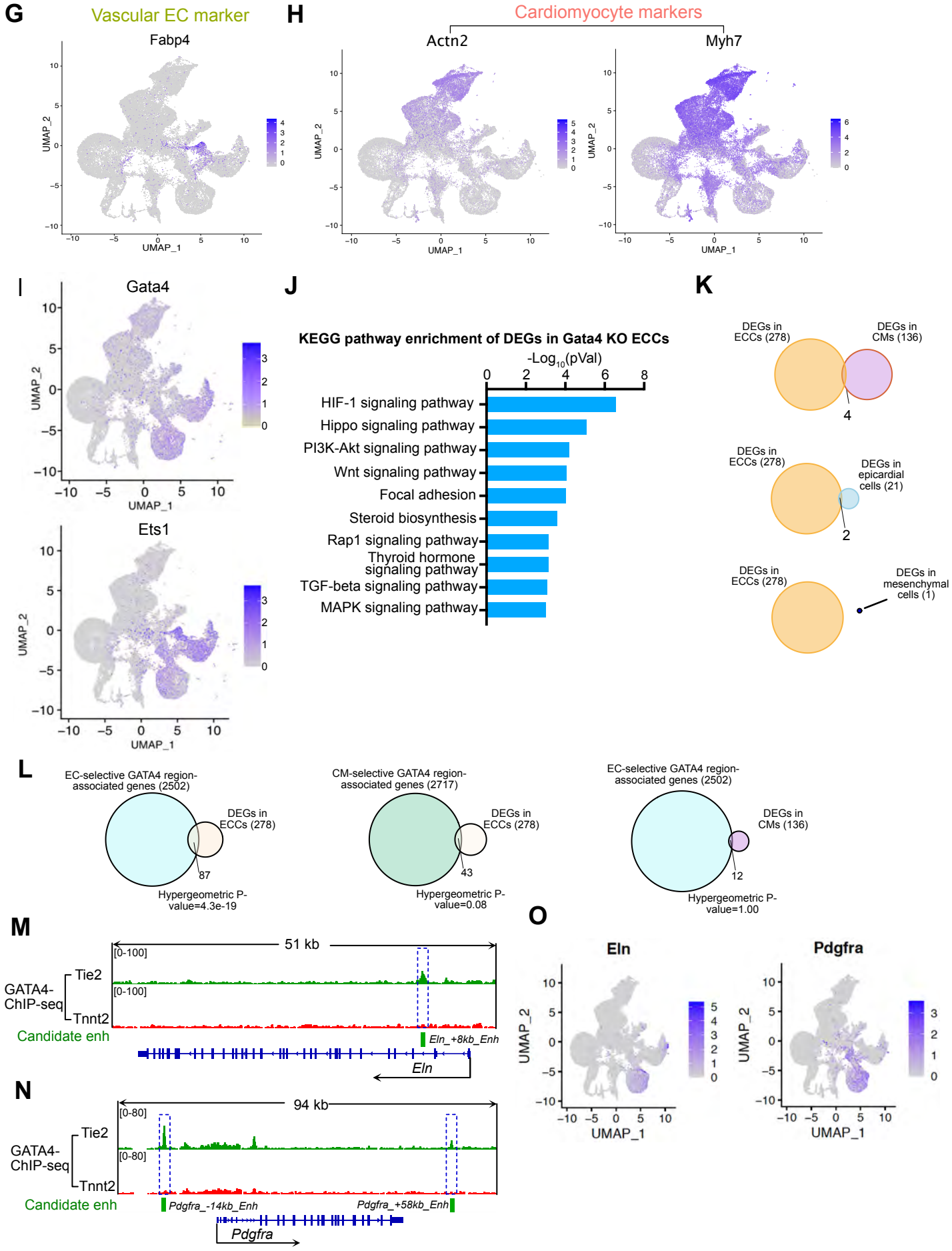


Fig. S11. scRNA-seq and reporter assays. A. Summary of scRNA-seq data of E12.5 heart ventricles. Correlation between biological replicates was determined using Spearman correlation. **B**. UMAP plots of each scRNA-seq replicate. **C**. QC of scRNA-seq data. Number of genes detected in each cell (nFeature_RNA), total number of molecules detected within a cell (nCount_RNA), mitochondrial content (percent.mt), and percentage of ribosomal genes (percent.rb). **D.** GFP reporter in *mTmG* mice is activated by Cdh5-CreERT2 with administration of tamoxifen at E9.5. **E-H.** Marker genes in endotheilal/endocardial cells (E), endocardial cell clusters (F), endothelial cells (G), or cardiomyocytes (H). *(continued)*


G
Vascular EC marker
Fabp4
H
Cardiomyocyte markers
Actn2
Myh7
UMAP_1
UMAP_2
I
Gata4
Ets1
J
KEGG pathway enrichment of DEGs in Gata4 KO ECCs
-Log10(pVal)
HIF-1 signaling pathway
Hippo signaling pathway
PI3K-Akt signaling pathway
Wnt signaling pathway
Focal adhesion
Steroid biosynthesis
Rap1 signaling pathway
Thyroid hormone signaling pathway
TGF-beta signaling pathway
MAPK signaling pathway
K
DEGs in ECCs (278)
DEGs in CMs (136)
4
DEGs in epicardial cells (21)
2
DEGs in mesenchymal cells (1)
L
EC-selective GATA4 region-associated genes (2502)
87
Hypergeometric P-value=4.3e-19
CM-selective GATA4 region-associated genes (2717)
43
Hypergeometric P-value=0.08
12
Hypergeometric P-value=1.00
M
51 kb
[0-100]
GATA4-ChIP-seq
Tie2
Tnnt2
Candidate enh
Eln_+8kb_Enh
Eln
N
94 kb
[0-80]
Pdgfra_-14kb_Enh
Pdgfra_+58kb_Enh
Pdgfra
O
Eln
Pdgfra

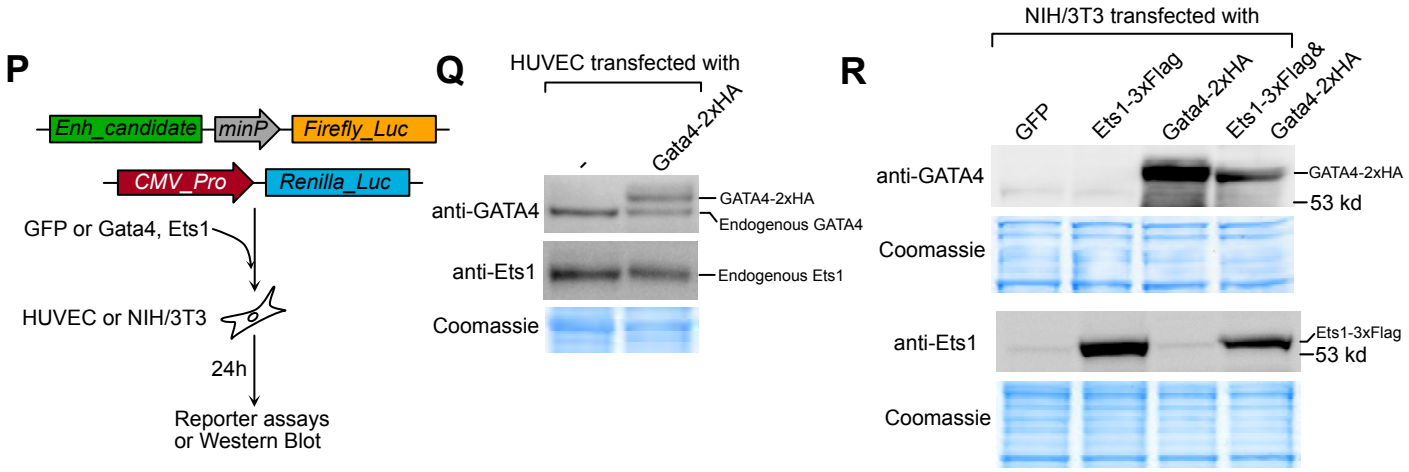


Fig. S11, continued.

I. Expression of *Gata4* and *Ets1* in single cells. **J**. Enriched KEGG pathways in Gata4 KO endocardial cells (ECCs). **K.** Distict DEGs in different cell types after *Gata4* inactivation in endothelial cells. **L**. Differentially expressed genes (DEGs) in ECCs are associated with EC-selective GATA4 regions, and are less likely to be assciated with CM-selective GATA4 regions. **M-N.** Genome browser views of Tie2 lineage-selective GATA4 regions at *Eln* and *Pdgfra* genes. Regions tested in reporter assays are highlighted. ChIP signals were RPKM normalized. **O**. Expression of *Eln* and *Pdgfra* in single cells. **P**. Schematic diagram of reporter assays in HUVEC or NIH3T3 cells. **Q-R**: Western blot shows the expression of GATA4 and ETS1 in HUVEC (P) and NIH3T3 (Q) cells.