Title

ATAC-Seq Reveals an Isl1 Enhancer That Regulates Sinoatrial Node Development and Function

Journal

Circulation Research, 127(12)

ISSN

0009-7330

Authors

Galang, Giselle
Mandla, Ravi
Ruan, Hongmei
et al.

Publication Date

2020-12-04

DOI

10.1161/circresaha.120.317145

Peer reviewed

HHS Public Access

Author manuscript

Circ Res. Author manuscript; available in PMC 2021 December 04.

Published in final edited form as:

Circ Res. 2020 December 04; 127(12): 1502–1518. doi:10.1161/CIRCRESAHA.120.317145.

ATAC-seq Reveals an *Isl1* Enhancer that Regulates Sinoatrial Node Development and Function

Giselle Galang[1], **Ravi Mandla**[1,*], **Hongmei Ruan**[1,*], **Catherine Jung**[1], **Tanvi Sinha**[1,2], **Nicole R. Stone**[5], **Roland S. Wu**[1,2], **Brandon J. Mannion**[6], **Prasanna K.R. Allu**[1,7], **Kevin Chang**[4], **Ashwin Rammohan**[1], **Marie B. Shi**[1], **Len A. Pennacchio**[6], **Brian L. Black**[2,3], **Vasanth Vedantham**[1,2]

[1]Cardiology Division, University of California, San Francisco, CA

[2]Cardiovascular Research Institute, University of California, San Francisco, CA

[3]Department of Biochemistry and Biophysics, University of California, San Francisco, CA

[4]School of Medicine, University of California, San Francisco, CA

[5]Gladstone Institute of Cardiovascular Disease, San Francisco, CA

[6]Environmental and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, CA, U.S. Department of Energy Joint Genome Institute, Berkeley, CA, and Comparative Biochemistry Program, University of California, Berkeley, CA

[7]Current location: Joslin Diabetes Center, Harvard Medical School, Boston, MA, 02215, USA.

Abstract

Rationale: Cardiac pacemaker cells (PCs) in the sinoatrial node (SAN) have a distinct gene expression program that allows them to fire automatically and initiate the heartbeat. Although critical SAN transcription factors, including Isl1, Tbx3, and Shox2, have been identified, the *cis*-regulatory architecture that governs PC-specific gene expression is not understood, and discrete enhancers required for gene regulation in the SAN have not been identified.

Objective: To define the epigenetic profile of PCs using comparative ATAC-seq and to identify novel enhancers involved in SAN gene regulation, development and function.

Methods and Results: We used ATAC-seq on sorted neonatal mouse SAN to compare regions of accessible chromatin in PCs and right atrial cardiomyocytes. PC-enriched ATAC-seq peaks, representing candidate SAN regulatory elements, were located near established SAN genes and were enriched for distinct sets of TF binding sites. Among several novel SAN enhancers that were experimentally validated using transgenic mice, we identified a 2.9-kb regulatory element at the *Isl1* locus that was active specifically in the cardiac inflow at E8.5 and throughout later SAN development and maturation. Deletion of this enhancer from the genome of mice resulted in SAN

Address correspondence to: Dr. Vasanth Vedantham, Smith Cardiovascular Research Building, 555 Mission Bay Blvd South, 352M, San Francisco, CA 94158, Tel: 415-476-7559, vasanth.vedantham@ucsf.edu.

*equal contribution

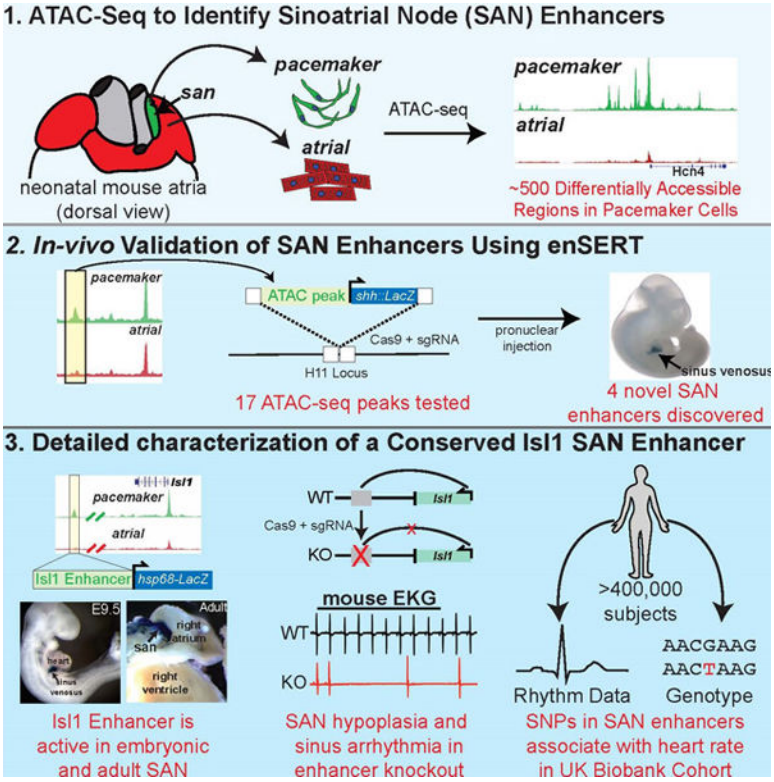
DISCLOSURES

V.V. has received consulting fees from Merck & Co. and research support for unrelated projects from Amgen.

hypoplasia and sinus arrhythmias. The mouse SAN enhancer also directed reporter activity to the inflow tract in developing zebrafish hearts, demonstrating deep conservation of its upstream regulatory network. Finally, single nucleotide polymorphisms in the human genome that occur near the region syntenic to the mouse enhancer exhibit significant associations with resting heart rate in human populations.

Conclusions: (1) PCs have distinct regions of accessible chromatin that correlate with their gene expression profile and contain novel SAN enhancers, (2) *Cis*-regulation of *Isl1* specifically in the SAN depends upon a conserved SAN enhancer that regulates PC development and SAN function, and (3) a corresponding human *ISL1* enhancer may regulate human SAN function.

Graphical Abstract



Keywords

Sinoatrial node; cis-regulatory element; cardiac conduction system; Isl1; heart rate; gene regulation; Arrhythmias; Developmental Biology; Electrophysiology; Gene Expression and Regulation; Physiology

INTRODUCTION

Cardiac pacemaker cells (PCs) within the sinoatrial node (SAN) are the leading pacemakers of the heart during normal rhythm. PCs are elongated, densely innervated cardiomyocytes that exhibit robust electrical automaticity but weak contractility[1]. Accordingly, their gene expression program diverges from that of working cardiomyocytes, with broad repression of

genes required for rapid conduction and force generation, and activation of a distinctive gene set associated with innervation and electrical automaticity[2, 3].

In recent years, the detailed transcriptome of PCs has been defined, revealing a conserved gene regulatory network that controls mammalian PC development and function, with critical roles for transcription factors (TFs) Shox2[4], Tbx3[5], and Isl1[2, 6]. However, the precise mechanisms that connect the expression of these regulators with their gene targets, and the mechanisms that control PC-specific expression of these TFs, remain incompletely understood, creating major barriers to understanding functional specialization in the cardiac conduction system in development and disease.

Epigenetic profiling has provided major insights into differentiation, functional specialization, and pathophysiology in a variety of tissues including the heart[7–9]. Partly because of the difficulty of isolating large numbers of pacemaker cells, genome-wide profiling of histone marks or TF binding events with chromatin immunoprecipitation and sequencing (ChIP-seq) has been difficult to apply to PCs. Consequently, regulatory elements that direct gene expression specifically to PCs and their progenitors have not been identified, limiting our understanding of mammalian PC development and differentiation.

The Assay for Transposase-Accessible Chromatin (ATAC-seq) is a robust method to define regions of accessible chromatin in rare cell populations[10]. ATAC-seq is more sensitive for detection of active *cis*-regulatory elements in cardiomyocytes than H3K27Ac ChIP-seq, and requires much less starting material to yield comparable signal-to-noise ratio, making it a preferred method for epigenetic profiling of PCs[11]. Here, we use ATAC-seq in purified PCs and right atrial cardiomyocytes (RACMs) to identify accessible chromatin regions that are specific to PCs. Analysis of these putative PC-specific regulatory elements identified enriched TF binding motifs, as well as several novel SAN enhancers that we experimentally validated. We then provide a detailed characterization of a 2.9-kb enhancer that activates *Isl1* specifically in the venous inflow myocardium from embryonic day (E)8.5 onwards, with progressive restriction to the SAN as development proceeds. Deletion of this enhancer resulted in SAN hypoplasia and sinus arrhythmias. The mouse *Isl1* SAN enhancer was also active in the cardiac inflow tract in transgenic zebrafish embryos, indicating deep conservation of the upstream regulatory architecture. Finally, SNPs located at the human region syntenic to the mouse *Isl1* SAN enhancer were associated with variation in resting heart rate in a large human population, providing evidence for its functional role in human SAN.

METHODS

An extended Methods section is available in the online supplement. RNA-seq and ATAC-seq datasets have been deposited in the Gene Expression Omnibus (GSE148515). Additional data supporting this study are available from the corresponding author upon reasonable request.

RESULTS

Isolation of cardiac pacemaker cells using Fluorescence Activated Cell Sorting (FACS).

To isolate pure populations of pacemaker cells (PCs) and right atrial cardiomyocytes (RACMs), *Hcn4-GFP* or *Hcn4-GFP/Myh6-mCherry* neonatal hearts were imaged under live fluorescence and the GFP+ SAN head was identified at the junction of the superior vena cava and right atrium (Figure 1A,B). For each biological replicate, SAN-RA junctions from 5 neonatal hearts were pooled for digestion to a single cell suspension for flow sorting, yielding 2–5% GFP+ cells and 25–35% mCherry+ cells. Over 99% of GFP+ cells isolated from *Hcn4-GFP/Myh6-mCherry* mice were also mCherry+, as expected (Online Figure I, A). Differential expression analysis on bulk RNA-seq data from GFP+ and mCherry+ cells (n = 3 biological replicates) demonstrated both enrichment of established SAN genes and strong depletion of known atrial genes in the GFP+ cells, with comparable expression of core cardiomyocyte TFs *Gata4*, *Tbx5*, and *Mef2C* between the two populations (Figure 1C, Online Table I, Online Figure I, B-D). Based on the magnitude of the depletion of atrial makers *Nppa, Nppb* and *Bmp10* in the GFP+ samples (Figure 1C), and on the low or undetectable expression signatures of non-myocyte cells (Online Figure I, E), we conclude that GFP+ cells are highly enriched for PCs.

Pacemaker cells demonstrate regions of differential chromatin accessibility.

Next, genomic DNA was isolated from flow sorted PCs and RACMs for ATAC-seq (~20,000 cells from 1 or 2 litters of neonates per biological replicate, Online Table II). Visualization of ATAC-seq signal at *Hcn4* and *Nppa* loci revealed discrete peaks in non-coding regions near *Hcn4* that were absent in RACMs and peaks near *Nppa* and *Nppb* that were absent in PCs (Figure 1D,E). Nonhierarchical clustering analysis on all peaks for each replicate showed that PC and RACM samples each clustered together, and that PC ATAC-seq peaks in our dataset were concordant with a previously published neonatal SAN ATAC-seq dataset[12] (Online Figure II A,B). To maximize capture of PC-specific regulatory elements, we then defined a set of consensus ATAC-seq peaks for each sample and examined the top 500 differentially accessible regions between PCs and RACMs (corresponding to a false discovery rate of 0.33, and uncorrected p < 1e-4, Figure 2A, Online Table III).

Differentially accessible regions in pacemaker cells are associated with differentially expressed genes.

In contrast to the global datasets of PC peaks and RACM peaks, which showed a distribution centered around basal promoters, the majority of the 500 differentially accessible regions that we identified were located between 50 and 500 kb from the nearest transcriptional start site (TSS), consistent with potential roles as tissue-specific regulatory elements (enhancers) (Online Figure II, C-E). To test this hypothesis, we used Genomic Regions Enrichment of Annotations Tool (GREAT) to assign differentially accessible regions to candidate genes[13]. Gene ontology (GO) analysis of this gene set (Figure 2B, Online Table IV) demonstrated a strong association with SAN development and function, and 108 out of 500 differentially accessible ATAC-seq peaks were associated with a gene that is differentially expressed in PCs vs RACMs (p < 1e-34), including *Hcn4*, *Isl1*, *Shox2*, and *Tbx3* (Figure 2B).

These data provide strong *in-silico* evidence that the set of differentially accessible ATAC-seq peaks we identified here includes regulatory elements that control gene expression in PCs.

Motif enrichment analysis of differentially accessible regions.

We then used HOMER[14] to identify enriched TF binding motifs in differentially accessible ATAC-seq peaks. In addition to over-representation of motifs for established cardiac TFs including Gata, Mef2, and T-box factors, several homeodomain TF binding sites were enriched including sites for Isl1, a critical PC TF, and Meis1, a TALE-class homeodomain protein involved in cardiac development[15] (Figure 2C, Online Table V). ATAC-seq peaks in PCs were also enriched at experimentally validated ChIP-seq peaks for Mef2C, Tbx5, Gata4, and Tead from embryonic myocardium[11], and at Nkx2.5 and Shox2 binding sites from embryonic cardiac inflow myocardium[16] (Online Figure III, A-B). Extending these analyses to the PC-enriched subset of differentially accessible ATAC-seq peaks showed robust enrichment of ATAC-seq signal at ChIP-seq peaks for Mef2C, Tbx5, and Gata4, but not as strongly for Nkx2.5 and Shox2 (Figure 2D). These data support existing models in which cardiac TFs including Tbx5, Mef2C, and Gata4 factors, but not Nkx2.5, positively regulate expression of genes that are restricted to PCs, and Shox2 represses atrial cardiomyocyte genes through competition with Nkx2.5[16].

Selected ATAC-seq peaks are sufficient to direct reporter gene expression to SAN.

To test whether differentially accessible ATAC-seq peaks are *bona fide* SAN enhancers, we used enSERT, a site-specific genomic integration method that allows high-throughput experimental validation of putative enhancers in mouse embryos[17]. Because ATAC-seq was performed at neonatal stages but enSERT-generated transgenic embryos were harvested at E11.5, we intersected differentially accessible ATAC-seq peaks with embryonic mouse heart H3K27Ac ChIP-seq datasets from ENCODE [18] (Online Figure III, C) and selected 16 regions with co-localization of PC-specific ATAC-seq peaks and H3K27Ac ChIP-seq peaks (Online Table VI). 5 out of 16 of these ATAC-seq-predicted, candidate enhancers exhibited cardiac activity in consistent patterns, of which 3 had activity in the SAN primordium (at loci for *Hcn4*, *Rgs6*, and *Ptgfr*, all PC-enriched genes), and 2 were specific for the venous inflow and PC primordium (*Rgs6* and *Ptgfr*, Figure 3A-C and Online Figure IV). These data demonstrate that a subset of PC-specific ATAC-seq peaks can function as enhancers that direct gene expression directly to the SAN, and that the entire dataset of differentially accessible ATAC-seq peaks is likely to contain many additional SAN enhancers.

Spatiotemporal activation pattern of an Isl1 locus SAN enhancer.

To identify additional enhancers at loci of critical PC TFs, we searched for differentially accessible ATAC-seq peaks within previously annotated TADs that contain *Shox2*, *Tbx3*, and *Isl1* (Online Figure V). While *Tbx3* and *Shox2* TADs each had numerous individual ATAC-seq peaks in PCs that were not present in RACMs, close examination of ATAC-seq signal over the entire *Isl1* TAD revealed a single previously uncharacterized peak located ~120 kb downstream of *Isl1* that was strongly differentially accessible between PCs and RACMs (Online Figure V, Figure 3D). To test whether the region defined by this peak might be an SAN-specific *Isl1* enhancer, we generated *Isl1 Locus SAN Enhancer (ISE)-LacZ and*

ISE-mCherry, stable transgenic mouse lines in which a 2.9-kb region encompassing the *Isl1* PC-specific ATAC-seq peak was cloned upstream of *hsp68LacZ* or *hsp68mCherry*. Of 7 transgenic mouse lines examined (2 *ISE-LacZ* and 5 *ISE-mCherry*), all exhibited restricted reporter activity in the SAN (Online Figure VI, A-C).

Onset of *ISE-LacZ* enhancer activity occurred at E8.5 in the cardiac inflow region (Figure 3E). At E9.5 enhancer activity was stronger and was restricted to the right inflow myocardium (Figure 3F and Online Figure VI, D). By E10.5 and E11.5, enhancer activity tracked the SAN primordium, with extension into the right venous valve (Figure 3G and Online Figure VI, E). By E12.5 and until neonatal stages, enhancer activity was largely restricted to the definitive SAN at the SVC-right atrial junction with extensions into the SAN tail and into the *crista terminalis* (Figure 3H-J, Online Figure VI, F). Notably, from the time of the formation of the SAN primordium and onward, ISE was active in the SAN region but not in other Isl1+ second heart field-derived tissues. Immunohistochemistry for Hcn4 and β-gal at E14.5 in *ISE-LacZ* hearts demonstrated restriction of β-gal expression to the SAN (Figure 3K). Adult hearts had persistent SAN activity (Figure 3L), with some additional low-level activity in other areas of the myocardium (Online Figure VI. G-I). Overall, our developmental time course demonstrates that *ISE* exhibits a remarkable degree of specificity for the SAN throughout development and maturation.

Isl1 Locus SAN enhancer-mCherry+ cells are functional Hcn4+ pacemaker cells.

To test whether *ISE* activity co-localized with Hcn4 expression in PCs, we crossed *ISE-mCherry* with *Hcn4-GFP*. Whole mount fluorescent imaging demonstrated co-localization of mCherry signal with GFP at E8.5 in the cardiac inflow tract (Figure 4A,A'), the sinus venosus at E10.5 (Figure 4B,B'), and the SAN at E14.5 and P2 (Figure 4C,C',D, and Online Figure VI, J-L). Whereas mCherry expression was restricted to the junction of the sinus venosus and right atrium during development, GFP was expressed more broadly, as has been previously described for Hcn4[19, 20], suggesting that *ISE* is activated by a more specific regulatory module than Hcn4. We did note a dim mCherry+/GFP- nodal extension that was visible in E14.5 embryos and in post-natal hearts (Figure 4C,C').

To test whether adult mCherry+ cells are *bona fide* cardiac pacemaker cells, we isolated single SAN cells from the intercaval regions of 5 adult *ISE-mCherry* mice. mCherry+ cells exhibited typical pacemaker cell morphologies (Figure 4E), although we also observed enhancer activity in some adult peri-nodal RACMs. However, action potentials recorded from spontaneously beating mCherry+ SAN myocytes were typical for PCs, with a depolarized maximum diastolic potential and spontaneous diastolic depolarization (Figure 4F,G). In addition, mCherry+ SAN cells exhibited robust funny current (the current carried by Hcn channels) as would be expected for PCs (Figure 4H, I).

Isl1 Locus SAN enhancer regulates Isl1 expression and SAN development.

To test directly whether *ISE* regulates *Isl1* expression in PCs and SAN development, we deleted a 2.7-kb region encompassing *ISE* in mice using CRISPR/Cas9 (Figure 5A). Homozygous *ISE* null mice (hereafter, $Isl1^{\Delta ISE/\Delta ISE}$) assessed at F2 and later generations did not demonstrate statistically significant embryonic or perinatal lethality (Figure 5B). To test

for reduction of Isl1 expression in *Isl1*$^{\Delta ISE/\Delta ISE}$ SAN, we used immunohistochemistry to detect Isl1 protein specifically in mutant PCs as compared to wild-type (WT) PCs by co-staining embryonic SAN for Isl1 and Hcn4 at E14.5. *Isl1*$^{\Delta ISE/\Delta ISE}$ Hcn4+ PCs exhibited grossly reduced Isl1 expression as compared to WT PCs (Figure 5C). To quantify this difference, we measured the ratio of nuclear Isl1 signal to background fluorescence in *Isl1*$^{\Delta ISE/\Delta ISE}$ and in WT littermates (35 PCs and 35 non-PCs from 7 sections per heart, 2 biological replicates per genotype), and found a highly significant reduction in Isl1 protein expression in PCs from *Isl1*$^{\Delta ISE/\Delta ISE}$ as compared to WT hearts, although *Isl1*$^{\Delta ISE/\Delta ISE}$ hearts did have detectable Isl1 signal above background, unlike non-PCs (Figure 5D). Importantly, these data establish that ISE regulates *Isl1* expression in the SAN *in-vivo*.

To test further whether *ISE* regulates *Isl1* transcription in *cis*, we crossed *Isl1*$^{\Delta ISE/+}$ with *Isl1*$^{3xFLAG/3xFLAG}$, a targeted mouse line generated in our lab with *3xFLAG* inserted before the *Isl1* stop codon (Online Figure VII, A). RNA was isolated by manual dissection from the inflow regions of 8 E11.5 embryos from 2 litters, yielding *Isl1*$^{3xFLAG/\Delta ISE}$ and *Isl1*$^{3xFLAG/+}$ RNA samples (n = 4 for each). To quantify the effect of *DISE* on Isl1 expression in *cis*, we designed PCR primers specific for the *Isl1–3xFLAG* mRNA (transcribed from the *Isl1* allele in *trans* to *DISE*) and WT *Isl1* mRNA (transcribed from the *Isl1* allele in *cis* to either *DISE* or WT *ISE*). While levels of *Isl1–3xFLAG* mRNA were similar between the two groups (Online Figure VII,B), expression of WT *Isl1* mRNA was reduced by 43% in E11.5 sinus venosus from *Isl1*$^{3xFLAG/\Delta ISE}$ as compared to *Isl1*$^{3xFLAG/+}$ (Online Figure VII, C), demonstrating that *ISE* regulates *Isl1* transcription via a *cis*-regulatory mechanism.

Because Isl1 positively regulates PC proliferation[6], we also assessed PC number and SAN size in *Isl1*$^{\Delta ISE/\Delta ISE}$ hearts and WT littermates at E14.5. First, the number of Hcn4+ nuclei was counted in a series of sections taken from a 200 μm long region at the SAN head at 20 βm cranial to caudal intervals (Figure 5E). The results of this analysis for 2 biological replicates showed a reduction in Hcn4+ nuclei in the SAN head region in *Isl1*$^{\Delta ISE/\Delta ISE}$ as compared to WT hearts at E14.5 (Figure 5F). In parallel, we observed a similar reduction in SAN volume as assessed by serial measurement of cross-sectional area of the Hcn4+ region in these sections (Figure 5G). These results demonstrate that *ISE* plays a role in establishing SAN cellularity and size, likely via regulation of *Isl1* expression level during a critical developmental window.

Isl1 Locus SAN enhancer regulates SAN function.

Cardiac morphology and function as determined by echocardiography were similar between adult *Isl1*$^{\Delta ISE/\Delta ISE}$ mice and their WT littermates (Online Table VII). To test for an effect of enhancer deletion specifically on heart rhythm, transmitters were implanted in 9 *Isl1*$^{\Delta ISE/\Delta ISE}$ adult mice and 11 WT littermates. Continuous electrocardiograms were recorded for 48 hours in all mice and average heart rate was determined hourly and averaged across mice within each group and sex (Figure 6A). These data demonstrated a slower average heart rate in mutant as compared to WT hearts, particularly during low activity periods (8AM – 8PM, highlighted in grey in Figure 6A). To explore this difference further, heart rates during low activity periods for each mouse were averaged over one-minute intervals and collated into 10 beat per minute bins to generate a histogram depicting the

amount of time spent at different heart rates (Figure 6B). Averaging histograms across mice within each genotype and sex showed a significant shift in mutant mice towards slower heart rates, with more time per mouse below 400 bpm in males and below 500 bpm in females ($p < 0.05$ for both). Assessment of intrinsic heart rate using atropine and propranolol, as well as maximum heart rate using isoproterenol, showed reductions for *Isl1*$^{\Delta ISE/\Delta ISE}$ versus WT littermates, but only reached statistical significance for maximum heart rate in females (Online Figure VIII). Detailed analysis of beat-to-beat variability of RR intervals in the time domain demonstrated increased variability in female *Isl1*$^{\Delta ISE/\Delta ISE}$ versus WT mice during high activity periods (Online Figure IX).

We also observed spontaneous arrhythmias in some mutant mice, including sinus pauses (defined as a two-fold or greater increase in the P-P interval from one beat to the next – Figure 6C, middle panel) and bradycardic episodes (defined as any period with heart rates less than 250 bpm—Figure 6C, lower panel). For each mouse, an investigator blinded to genotype counted the number of these events over a 24-hour period (Figure 6D). We found that while WT mice had between 0 and 2 episodes per hour, mutant mice had significantly more frequent episodes ($p < 0.01$). In addition, the cumulative duration of these episodes was also significantly greater in the mutant mice than in the WT mice (Figure 6E, $p < 0.01$), demonstrating that removal of *ISE*, in addition to affecting Isl1 expression and SAN development, also caused changes in SAN function.

To test whether episodes of SAN dysfunction in *Isl1*$^{\Delta ISE/\Delta ISE}$ could arise from abnormal cellular electrophysiology in *Isl1*$^{\Delta ISE/\Delta ISE}$ PCs, we used whole cell patch clamp to record action potentials from isolated PCs from 2 WT and 2 *Isl1*$^{\Delta ISE/\Delta ISE}$ adult mice (Online Figure X). We found no significant differences in firing rate, minimum diastolic potential, and early diastolic depolarization, suggesting that ISE likely affects SAN function through its effects on SAN morphogenesis rather than by determining functional properties of differentiated PCs.

Evolutionary conservation and upstream regulation of Isl1 locus SAN enhancer.

Using the ECR Browser[21] to visualize the degree of evolutionary conservation between mouse ISE and other mammalian species, we found that *ISE* exhibited >70% conservation in humans and opossum, suggesting that the regulatory network upstream of the enhancer may also be deeply conserved (Figure 7A, top track). Furthermore, alignment of *ISE* ATAC-seq signal with ChIP-seq datasets from embryonic myocardium[11] demonstrated that *ISE* contains experimentally validated binding sites for Gata4, Tbx5, and Tead, factors with conserved functions across vertebrate heart development (Figure 7A). Heart rhythm in zebrafish is controlled by a small population of automatically firing cells at the junction of the sinus venosus and atrium that expresses Isl1, Shox2, and Hcn4, similar to the mammalian SAN[22, 23]. To test whether the regulatory network upstream of *ISE* is present in the zebrafish SAN, we cloned mouse *ISE* into a zebrafish enhancer-reporter vector and injected this construct into fertilized zebrafish eggs derived from a *Tg(Myl7:mCherry-NTR)* zebrafish line, in which mCherry is expressed throughout the myocardium. 18 out of 160 injected embryos examined at 48 hours post fertilization (hpf) exhibited restriction of GFP

expression to the junction of the sinus venosus and atrium, consistent with deep evolutionary conservation of the regulatory network controlling enhancer activity (Figure 7B, top panel).

To test whether the regions within *ISE* containing the Gata4, Tbx5, and Tead TF binding sites are sufficient for enhancer function, we injected a 1-kb fragment containing the cardiac TF binding sites (Fr1). In 19 out of 120 embryos examined, Fr1 directed reporter expression to the entire heart and sinus venosus without restriction of enhancer activity to the SAN region (Figure 7B, bottom panel, and Figure 7C). We did not observe any consistent pattern of extra-cardiac GFP expression in zebrafish embryos injected with either construct. These results suggest a model in which the Gata4, Tbx5, and Tead may be sufficient for enhancer activation but are insufficient to confer enhancer specificity to the SAN. Finally, we tested *Fr1-hsp68-LacZ* in mouse transient transgenic assays and did not observe any consistent pattern of reporter activity, although 1 out of 3 transgenic founders had enhancer activity throughout the heart, suggesting that activation and restriction of enhancer activity in the murine SAN likely requires more than one region of the 2.9-kb full length enhancer (Online Figure XI).

SAN ATAC-seq peaks occur at sites that regulate variation in human SAN function.

It is widely believed that non-coding variants are major contributors to genetically determined variation in complex human traits such as sinus node function and disease susceptibility[17, 24]. We therefore hypothesized that single nucleotide polymorphisms occurring near sites of accessible chromatin in PCs and RACMs might exhibit associations with parameters that reflect sinus node function such as resting heart rate. To test this possibility with our ATAC-seq dataset, we first defined the set of human genomic regions that are syntenic to the murine ATAC-seq peaks we identified in RACMs and PCs, and interrogated resting heart rate data from the UK Biobank[25], a population-based cohort that included 434,532 genotyped participants meeting inclusion criteria (Online Figure XII, A). We then compiled a list of genotyped SNPs that overlap or are within 0.5 kb at either end of each of these regions of human genomic sequence (335,353 in total for all ATAC-seq peaks). After adjusting for age, sex, and body mass index, we identified SNPs associated with resting heart rate located in or near our ATAC-seq peaks, including numerous associations near genes that have been previously identified in heart rate GWAS analysis, such as *MYH7*, *TTN*, and *KIAA1755*, among many others (Figure 8A, Online Table VIII).

Next, to test whether PC-enriched enhancers might harbor associations as well, we identified 239 SNPs that were within or ±0.5 kb outside of human genomic regions syntenic to PC-enriched ATAC-seq peaks. In this analysis, we identified SNPs at the *HCN4* and *RGS6* loci, genes with established roles in heart rate regulation (Figure 8B). However, by limiting our analysis to 239 SNPs and adjusting our significance cutoff accordingly, our analysis also uncovered a SNP at the *ISL1* locus in a region of DNA syntenic to *ISE* that has a significant association with resting heart rate. To test whether other SNPs in the region surrounding the *ISE* might also exhibit associations with resting heart rate, we interrogated a broader region at the Isl1 locus, and we identified another SNP in close proximity to the *ISE* with a stronger association with resting heart rate (Figure 8C). Although the effect sizes are small (0.1 bpm/allele for rs1423611 and 0.2 bpm/allele for rs57848528, Online Figure XII B,C), the finding

of multiple SNPs in close proximity to *ISE* that associate with variation in resting heart rate in a human population provides evidence supporting the functional conservation of this enhancer. Finally, we note that in a separate ATAC-seq dataset derived from human induced pluripotent stem cells differentiated into pacemaker-like cells, there was a discrete ATAC-seq peak at the region syntenic to the *ISE*, further supporting a role for *ISE* in regulating human heart rhythm (Vincent Christoffels, personal communication).

DISCUSSION

In this work, we have defined the epigenetic profile of cardiac pacemaker cells and experimentally validated a novel set of SAN enhancers, including a deeply conserved PC-specific enhancer for *Isl1* (*ISE*) that is required for normal SAN development and function. Our findings provide new insights into transcriptional regulation in PCs and lay a foundation for future studies on PC differentiation, SAN development and SAN disease.

Comparison of epigenetic profiles of PCs and RACMs.

Our comparative ATAC-seq demonstrated that while the vast majority of ATAC-seq peaks are shared between RACMs and PCs, PCs had many more differentially accessible regions than RACMs. Notably, except for relatively few genes such as *Nppa* and *Bmp10* that are not expressed in PCs, the majority of genes associated with the working cardiomyocyte program are downregulated in PCs rather than turned off completely[2], which may not require closure of open chromatin. In contrast, the PC gene program includes many transcripts that are not expressed at all in RACMs, requiring opening of genomic regulatory regions that are otherwise closed in working cardiomyocytes. Defining the epigenetic profile of the progenitor populations for PCs and RACMs might provide further insight into the events that underlie acquisition of the distinct PC epigenetic profile.

Transcriptional regulation in PCs.

While PCs and RACMs both express cardiac TFs Tbx5, Mef2C, and Gata4 at comparable levels, PCs express lower levels of Nkx2.5 and also express Tbx3, Tbx18, Shox2, and Isl1 [26]. Whereas both Tbx3 and Shox2 repress RACM genes in PCs[5,16], loss of function experiments support a positive role for Isl1 in transcriptional regulation in PCs[6]. Our motif analysis, in which we identified enrichment of Isl1 binding sites specifically in PC-enriched ATAC-seq peaks, is also supportive of such a role. Furthermore, we identified other LIM homeodomain binding motifs (Lhx1 and Lhx3) and motifs for TALE-class homeodomain proteins (including Meis1) in PC-specific ATAC-peaks, raising the possibility of combinatorial regulation of PC gene expression by a variety of different homeodomain proteins. In this context, a role for Shox2 in activating PC-specific expression cannot be ruled out, although we observed limited co-localization of PC-enriched ATAC-seq signal at Shox2 ChIP-seq peaks. On the other hand, involvement of cardiac TFs such as Gata4, Tbx5, and Tead in activation of *ISE*, along with the strong enrichment of these and other cardiac factors such as Mef2 in our motif analysis, raises the possibility that in some instances PC-restricted expression may be achieved largely through repression in working cardiomyocytes as opposed to action of distinct sets of transcriptional activators in PCs. The epigenetic

profile presented here provides a foundation for future studies directed at testing these possibilities.

Identification of novel SAN enhancers.

In addition to *ISE*, 3 out of 16 putative enhancers we tested were sufficient to direct SAN-specific reporter activation at E11.5. All 3 (*Hcn4*, *Rgs6*, and *Ptgfr*) were located in proximity to genes enriched in PCs. Deletion of a region containing the enhancer at the *Hcn4* locus reduces Hcn4 expression and affects SAN function in mice, supporting annotation of this element as a *bona fide Hcn4* SAN enhancer[27]. The other putative enhancers that we tested were not sufficient to direct reporter expression to the SAN, suggesting that additional regions may be required for full regulatory function or that these regions are active at different developmental stages. Nevertheless, the SAN enhancers we identified are of tractable size (1–3 kb), raising the possibility of deeper exploration of upstream transcriptional regulation in PCs in future work. In addition, using a candidate gene approach, we identified regulatory elements controlling expression of the key PC-specific TFs, including Isl1, Shox2, and Tbx3. Beyond our work here with the *ISE*, several of the elements at *Shox2* and *Tbx3* loci that we identified are also present as ATAC-seq peaks within the syntenic regions in human iPS-derived cardiac pacemaker-like cells and are required for expression of Shox2 and Tbx3 in the SAN in mouse embryos (Vincent Christoffels, personal communication), providing definitive experimental validation of the regulatory elements we identified with ATAC-seq.

Modular control of Isl1 during cardiogenesis.

While *Isl1* is located within a megabase-sized gene desert that encompasses hundreds of highly conserved non-coding regions, *ISE* was the only highly differentially accessible region within the entire TAD in neonatal SAN. A broad *Isl1* second heart field enhancer was previously identified in the region immediately downstream of the *Isl1* 3'-untranslated region[28]. This enhancer was active as early as E7 in second heart field cardiac progenitors and depended on conserved forkhead TF binding sites. In contrast, activity of *ISE* was detectable at E8.5 in a restricted region near the cardiac inflow tract. Enhancer activity became progressively more restricted to the SAN as development progressed, with persistent SAN restriction in the perinatal period. Thus, the *Isl1* SAN enhancer demonstrates modular control of *Isl1* expression in different expression domains within the heart during development. Nevertheless, deletion of *ISE* did not result in complete loss of *Isl1* expression, which suggests either that other regulatory elements that we did not identify are involved in persistent *Isl1* expression in the SAN, or that there are shadow enhancers or other mechanisms that can activate *Isl1* expression in the SAN in the absence of *ISE*.

ISE was active in adult hearts at low levels in areas outside the SAN including ventricular myocardium, ventricular conduction system, and perinodal right atrial myocardium, where *Isl1* is not expressed. While this activity most likely reflects removal of *ISE* from its native genomic context in the *hsp68* vector, enhancer utilization changes with cardiac maturation, raising the possibility that *ISE* may function differently in the adult heart[11].

Role of Isl1 SAN Enhancer in SAN development and function.

Complete loss of Isl1 in the developing SAN results in reversion of PC gene expression to an atrial-like pattern[2], with early-onset severe bradycardia and SAN hypoplasia[6]. Although we observed only partial loss of *Isl1* expression in $Isl1^{\Delta ISE/\Delta ISE}$ SAN, there was still significant SAN hypoplasia, indicating that SAN growth is sensitive to Isl1 dosage. This finding fits with studies that have defined E11 to E14 (when *ISE* activity is strongest) as the period when the SAN grows most rapidly[29]. In addition, we have previously found that Isl1 expression in PCs as assessed by RNA-seq peaks at E14.5[2]. Based on these findings, we propose a model in which *ISE* generates a burst of Isl1 expression in the developing SAN to support PC proliferation during a critical window. Accordingly, loss of *ISE* results in SAN hypoplasia but does not disrupt early events leading to SAN formation.

Our findings of increased sinus arrhythmias and slower heart rate in adult $Isl1^{\Delta ISE/\Delta ISE}$ mice despite normal PC cellular electrophysiology demonstrate that developmental hypoplasia can lead to dysfunction of the adult SAN. Notably, we observed a larger heart rate decrease and greater change in beat-to-beat variation in female mice as compared to male mice. This could have resulted from sex differences in body mass leading to increased thermal stress in female mice at room temperature. Alternatively, our findings could have resulted from intrinsic sex differences in SAN structure and function that render females more sensitive to loss of *ISE*. Further investigation of sex differences, as well as more detailed analysis of heart rate variability, would be required to test these possible explanations more directly. Although the functional phenotype in $Isl1^{\Delta ISE/\Delta ISE}$ mice was mild despite loss of an important regulatory element, modest changes in SAN function under resting conditions may affect survival and fitness in mice subjected to stress or increased cardiac demand, explaining the deep evolutionary conservation of *ISE*.

Relationship of SAN enhancers identified in mice to human heart rhythm.

Because mouse hearts function at much higher rates than human hearts, sequence level conservation of a non-coding genomic element such as *ISE* does not imply conservation of function. Although a naturally occurring deletion or disruption of *ISE* co-segregating with abnormal SAN function in a human family would provide definitive evidence of a conserved role for *ISE*, recent data suggest that common sequence variation in enhancers accounts for some variation and disease susceptibility in human populations[17,30]. For example, the human region syntenic to the *Hcn4* enhancer we identified here harbors SNPs that affect susceptibility to atrial fibrillation[27]. In addition, increasingly large cohorts of patients with well-annotated clinical data have enabled GWAS-based discovery of numerous variants that have an impact on heart rhythm [31, 32]. Here we tested whether limiting the SNP analysis to a smaller search space could uncover more modest associations that might not reach statistical significance in genome-wide analysis. Our finding of 2 SNPs in the region of *ISE* that are associated with small changes in resting heart rate suggests that this approach may be fruitful for regulatory elements more broadly, and provides some evidence that *ISE* plays a role in human heart rhythm regulation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

The authors thank members of the Gladstone Histology, Genomics, and Bioinformatics core facilities, the UCSF Center for Advanced Technology, the UCSF Laboratory for Cell Analysis, and the Nikon Imaging Center at UCSF for technical assistance. The authors also thank Allan Villanueva for assistance with mouse colony maintenance. This research has been conducted using the UK Biobank Resource under Application Number 55046.

SOURCES OF FUNDING

This work was supported by U.S. National Institutes of Health grants to L.A.P. (R01HG003988), B.L.B (R01HL064658 and R01HL136182), V.V. (DP2-HL152425), and by an American Heart Association Fellowship to P.K.R.A. (18POST34030194). The UCSF Laboratory for Cell Analysis was supported by a National Cancer Institute Center Support Grant (P30CA082103). Research was conducted at the E.O. Lawrence Berkeley National Laboratory and performed under U.S. Department of Energy Contract DE-AC02-05CH11231, University of California.

Nonstandard Abbreviations and Acronyms:

SAN	Sinoatrial node
RACM	Right atrial cardiomyocyte
PC	Pacemaker cell
TF	Transcription factor
ISE	Isl1 locus SAN enhancer
ATAC-seq	Assay for transposase-accessible chromatin

REFERENCES

1. Mangoni ME, Nargeot J. Genesis and regulation of the heart automaticity. Physiol Rev. 2008;88:919–982 [PubMed: 18626064]
2. Vedantham V, Galang G, Evangelista M, Deo RC, Srivastava D. Rna sequencing of mouse sinoatrial node reveals an upstream regulatory role for islet-1 in cardiac pacemaker cells. Circ Res. 2015
3. van Eif VWW, Stefanovic S, van Duijvenboden K, Bakker M, Wakker V, de Gier-de Vries C, Zaffran S, Verkerk AO, Boukens BJ, Christoffels VM. Transcriptome analysis of mouse and human sinoatrial node cells reveals a conserved genetic program. Development. 2019;146
4. Blaschke RJ, Hahurij ND, Kuijper S, Just S, Wisse LJ, Deissler K, Maxelon T, Anastassiadis K, Spitzer J, Hardt SE, et al. Targeted mutation reveals essential functions of the homeodomain transcription factor shox2 in sinoatrial and pacemaking development. Circulation. 2007;115:1830–1838 [PubMed: 17372176]
5. Hoogaars WM, Engel A, Brons JF, Verkerk AO, de Lange FJ, Wong LY, Bakker ML, Clout DE, Wakker V, Barnett P, et al. Tbx3 controls the sinoatrial node gene program and imposes pacemaker function on the atria. Genes Dev. 2007;21:1098–1112 [PubMed: 17473172]
6. Liang X, Zhang Q, Cattaneo P, Zhuang S, Gong X, Spann NJ, Jiang C, Cao X, Zhao X, Zhang X, et al. Transcription factor isl1 is essential for pacemaker development and function. J Clin Invest. 2015;125:3256–3268 [PubMed: 26193633]
7. Gilsbach R, Schwaderer M, Preissl S, Gruning BA, Kranzhofer D, Schneider P, Nuhrenberg TG, Mulero-Navarro S, Weichenhan D, Braun C, et al. Distinct epigenetic programs regulate cardiac

myocyte development and disease in the human heart in vivo. Nat Commun. 2018;9:391 [PubMed: 29374152]

8. Gokbuget D, Blelloch R. Epigenetic control of transcriptional regulation in pluripotency and early differentiation. Development. 2019;146
9. Cui M, Wang Z, Bassel-Duby R, Olson EN. Genetic and epigenetic regulation of cardiomyocytes in development, regeneration and disease. Development. 2018;145
10. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods. 2013;10:1213–1218 [PubMed: 24097267]
11. Akerberg BN, Gu F, VanDusen NJ, Zhang X, Dong R, Li K, Zhang B, Zhou B, Sethi I, Ma Q, et al. A reference map of murine cardiac transcription factor chromatin occupancy identifies dynamic and conserved enhancers. Nat Commun. 2019;10:4907 [PubMed: 31659164]
12. Fernandez-Perez A, Sathe AA, Bhakta M, Leggett K, Xing C, Munshi NV. Hand2 selectively reorganizes chromatin accessibility to induce pacemaker-like transcriptional reprogramming. Cell Rep. 2019;27:2354–2369 e2357 [PubMed: 31116981]
13. McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM, Bejerano G. Great improves functional interpretation of cis-regulatory regions. Nat Biotechnol. 2010;28:495–501 [PubMed: 20436461]
14. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and b cell identities. Mol Cell. 2010;38:576–589 [PubMed: 20513432]
15. Mahmoud AI, Kocabas F, Muralidhar SA, Kimura W, Koura AS, Thet S, Porrello ER, Sadek HA. Meis1 regulates postnatal cardiomyocyte cell cycle arrest. Nature. 2013;497:249–253 [PubMed: 23594737]
16. Ye W, Wang J, Song Y, Yu D, Sun C, Liu C, Chen F, Zhang Y, Wang F, Harvey RP, et al. A common shox2-nkx2–5 antagonistic mechanism primes the pacemaker cell fate in the pulmonary vein myocardium and sinoatrial node. Development. 2015;142:2521–2532 [PubMed: 26138475]
17. Kvon EZ, Zhu Y, Kelman G, Novak CS, Plajzer-Frick I, Kato M, Garvin TH, Pham Q, Harrington AN, Hunter RD, et al. Comprehensive in vivo interrogation reveals phenotypic impact of human enhancer variants. Cell. 2020;180:1262–1271 e1215 [PubMed: 32169219]
18. Consortium EP. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012;489:57–74 [PubMed: 22955616]
19. Liang X, Wang G, Lin L, Lowe J, Zhang Q, Bu L, Chen Y, Chen J, Sun Y, Evans SM. Hcn4 dynamically marks the first heart field and conduction system precursors. Circ Res. 2013;113:399–407 [PubMed: 23743334]
20. Vedantham V, Evangelista M, Huang Y, Srivastava D. Spatiotemporal regulation of an hcn4 enhancer defines a role for mef2c and hdacs in cardiac electrical patterning. Dev Biol. 2013;373:149–162 [PubMed: 23085412]
21. Ovcharenko I, Nobrega MA, Loots GG, Stubbs L. Ecr browser: A tool for visualizing and accessing data from comparisons of multiple vertebrate genomes. Nucleic Acids Res. 2004;32:W280–286
22. Ren J, Han P, Ma X, Farah EN, Bloomekatz J, Zeng XI, Zhang R, Swim MM, Witty AD, Knight HG, et al. Canonical wnt5b signaling directs outlying nkx2.5+ mesoderm into pacemaker cardiomyocytes. Dev Cell. 2019;50:729–743 e725 [PubMed: 31402282]
23. Tessadori F, van Weerd JH, Burkhard SB, Verkerk AO, de Pater E, Boukens BJ, Vink A, Christoffels VM, Bakkers J. Identification and functional characterization of cardiac pacemaker cells in zebrafish. PLoS One. 2012;7:e47644
24. van Ouwerkerk AF, Bosada FM, van Duijvenboden K, Hill MC, Montefiori LE, Scholman KT, Liu J, de Vries AAF, Boukens BJ, Ellinor PT, et al. Identification of atrial fibrillation associated genes and functional non-coding variants. Nat Commun. 2019;10:4755 [PubMed: 31628324]
25. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, et al. Uk biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med. 2015;12:e1001779

26. Ye W, Song Y, Huang Z, Zhang Y, Chen Y. Genetic regulation of sinoatrial node development and pacemaker program in the venous pole. J Cardiovasc Dev Dis. 2015;2:282–298 [PubMed: 26682210]
27. van Ouwerkerk AF, Bosada F, Liu J, Zhang J, van Duijvenboden K, Chaffin M, Tucker N, Pijnappels DA, Ellinor PT, Barnett P, et al. Identification of functional variant enhancers associated with atrial fibrillation. Circ Res. 2020
28. Kang J, Nathan E, Xu SM, Tzahor E, Black BL. Isl1 is a direct transcriptional target of forkhead transcription factors in second-heart-field-derived mesoderm. Dev Biol. 2009;334:513–522 [PubMed: 19580802]
29. Challice CE, Viragh S. Origin and early differentiation of the sinus node in the mouse embryo heart. Adv Myocardiol. 1980;1:267–277 [PubMed: 7394333]
30. Albert FW, Kruglyak L. The role of regulatory variation in complex traits and disease. Nat Rev Genet. 2015;16:197–212 [PubMed: 25707927]
31. Choi SH, Jurgens SJ, Weng LC, Pirruccello JP, Roselli C, Chaffin M, Lee CJ, Hall AW, Khera AV, Lunetta KL, et al. Monogenic and polygenic contributions to atrial fibrillation risk: Results from a national biobank. Circ Res. 2020;126:200–209 [PubMed: 31691645]
32. Ramirez J, Duijvenboden SV, Ntalla I, Mifsud B, Warren HR, Tzanis E, Orini M, Tinker A, Lambiase PD, Munroe PB. Thirty loci identified for heart rate response to exercise and recovery implicate autonomic nervous system. Nat Commun. 2018;9:1947 [PubMed: 29769521]
33. Jesty SA, Steffey MA, Lee FK, Breitbach M, Hesse M, Reining S, Lee JC, Doran RM, Nikitin AY, Fleischmann BK, et al. C-kit+ precursors support postinfarction myogenesis in the neonatal, but not adult, heart. Proc Natl Acad Sci U S A. 2012;109:13380–13385
34. Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. Atac-seq: A method for assaying chromatin accessibility genome-wide. Curr Protoc Mol Biol. 2015;109:21 29 21–29
35. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for illumina sequence data. Bioinformatics. 2014;30:2114–2120 [PubMed: 24695404]
36. Ramirez F, Bhardwaj V, Arrigoni L, Lam KC, Gruning BA, Villaveces J, Habermann B, Akhtar A, Manke T. High-resolution tads reveal DNA sequences underlying genome organization in flies. Nat Commun. 2018;9:189 [PubMed: 29335486]
37. Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dundar F, Manke T. Deeptools2: A next generation web server for deep-sequencing data analysis. Nucleic Acids Res. 2016;44:W160–165
38. He A, Gu F, Hu Y, Ma Q, Ye LY, Akiyama JA, Visel A, Pennacchio LA, Pu WT. Dynamic gata4 enhancers shape the chromatin landscape central to heart development and disease. Nat Commun. 2014;5:4907 [PubMed: 25249388]
39. Zhao H, Sun Z, Wang J, Huang H, Kocher JP, Wang L. Crossmap: A versatile tool for coordinate conversion between genome assemblies. Bioinformatics. 2014;30:1006–1007 [PubMed: 24351709]
40. Fenske S, Probstle R, Auer F, Hassan S, Marks V, Pauza DH, Biel M, Wahl-Schott C. Comprehensive multilevel in vivo and in vitro analysis of heart rate fluctuations in mice by ecg telemetry and electrophysiology. Nat Protoc. 2016;11:61–86 [PubMed: 26658468]
41. Turner SD. Qqman: An r package for visualizing gwas results using q-q and manhattan plots. The Journal of Open Source Software. 2018;3:731
42. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, Boehnke M, Abecasis GR, Willer CJ. Locuszoom: Regional visualization of genome-wide association scan results. Bioinformatics. 2010;26:2336–2337 [PubMed: 20634204]

NOVELTY AND SIGNIFICANCE

What Is Known?

- Cardiac pacemaker cells (PCs) in the sinoatrial node (SAN) have a unique gene expression program that allows them to function as leading pacemakers of the heart
- Identification of tissue-specific gene regulatory elements through genome-wide profiling of chromatin accessibility has provided critical insights into how regulatory networks function in other cardiac cell types but has not yet been carried out on PCs.

What New Information Does This Article Contribute?

- PCs and right atrial cardiomyocytes (RACMs) were isolated with fluorescence activated cell sorting from neonatal mouse hearts and genome-wide maps of chromatin accessibility were generated using the assay for transposase accessible chromatin (ATAC-seq)
- Comparison of ATAC-seq peaks in PCs and RACMs revealed differentially accessible chromatin regions in PCs that functioned as tissue specific enhancers *in-vivo*, including a deeply evolutionarily conserved enhancer that regulates *Isl1* transcription in PCs and is required for normal SAN development function.
- Human single nucleotide polymorphisms (SNPs) located near regions syntenic to accessible chromatin loci in PCs are associated with resting heart rate in a large human population, including small associations for SNPs near the *Isl1* SAN enhancer.

The sinoatrial node (SAN) contains specialized cardiac pacemaker cells (PCs) that control heart rhythm though coordinated electrical automaticity, ensuring that heart rate and cardiac output are matched to metabolic demand. SAN failure is common and cannot be prevented or reversed, necessitating electronic pacemaker implantation for affected patients. Despite the importance of SAN function to cardiovascular disease, the gene regulatory networks that drive functional specialization of PCs in the SAN remain incompletely understood, limiting the development of biologically-based treatments for SAN failure. To address this knowledge gap, we compared the landscape of accessible chromatin in purified PCs with that of right atrial cardiomyocytes, revealing regions of differentially accessible chromatin that functioned as SAN enhancers *in-vivo*, including a deeply conserved regulatory element for the transcription factor *Isl1* that was required for normal SAN development and function. Regions of the human genome that are syntenic to the differentially accessible chromatin regions from mouse PCs contained numerous SNPs that were associated with human heart rate, providing evidence that regulation of gene expression in PCs has a deeply conserved architecture that is relevant to human heart rhythm. Taken together, these data provide a new foundation for investigations of transcriptional mechanisms in SAN development and disease.

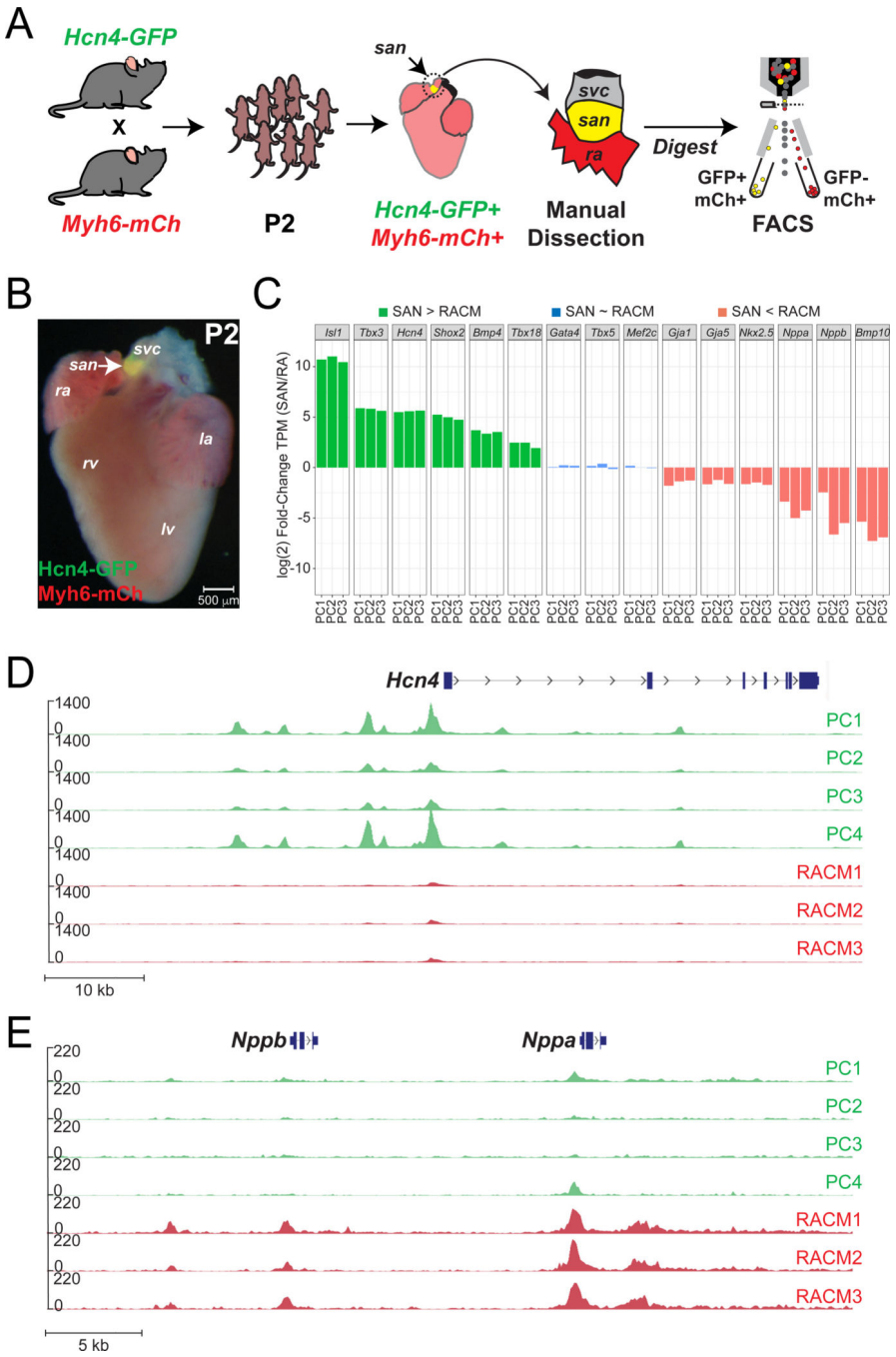


Figure 1. Isolation of Neonatal Pacemaker Cells for ATAC-seq.

(A) Isolation of neonatal pacemaker cells (PCs) and right atrial cardiomyocytes (RACMs) from the sinoatrial node (SAN) using fluorescence activated cell sorting (FACS) on samples derived from 3 biological replicates, where each biological replicate contained pooled SAN tissue from 5 neonatal mice. (B) Ventral view of an *Hcn4-GFP/Myh6-mCherry* neonatal heart with SAN positive for both fluorescent markers (yellow) (C) Enrichment of known PC genes (green bars) and depletion of known RACM genes (red bars) from bulk RNAseq on sorted PCs and RACMs (each biological replicate is shown (PC1, PC2, and PC3), data are

displayed as log(2) fold change of SAN/RACM transcripts per million (TPM)). Blue bars show genes that are not differentially expressed. (D,E) ATAC-seq signal for individual PC (n = 4) and RACM (n = 3) biological replicates at *Hcn4* locus (D) and *Nppa* locus (E) Each biological replicate consisted of pooled SAN tissue from 1–2 litters of neonatal hearts. Abbreviations: san, sinoatrial node; svc, superior vena cava; la, left atrium; ra, right atrium; lv, left ventricle; rv, right ventricle.

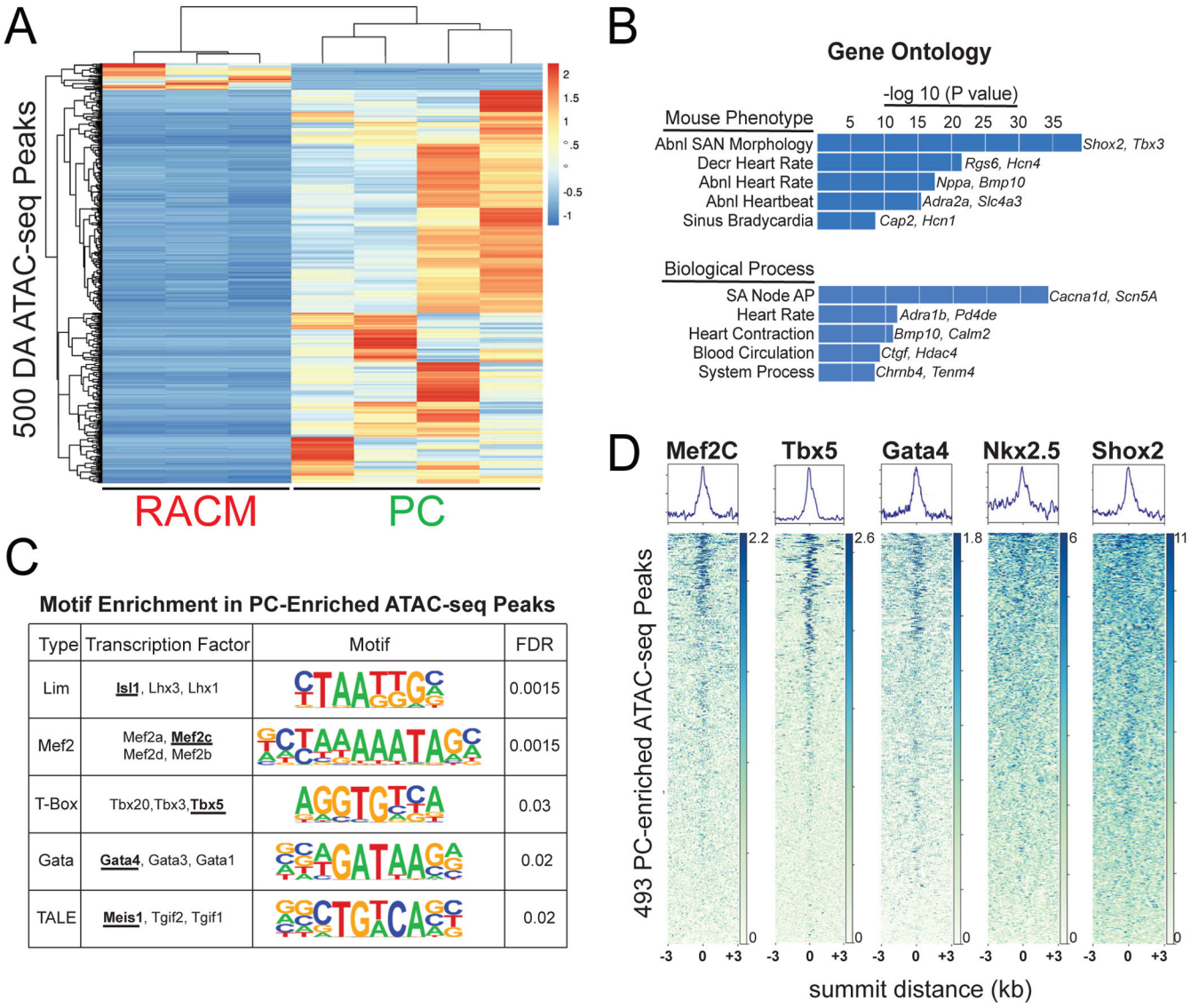


Figure 2. Proximity of Differential ATAC-seq Peaks to Differentially Expressed Genes and Enrichment of Cardiac TF Motifs and Binding Sites at ATAC-seq Peaks.

(A) Clustered heatmap with z-scores for the top 500 differentially accessible (DA) ATAC-seq peaks across all 7 samples. (B) Gene ontology (GO) terms for genes assigned to DA ATAC-seq peaks. (C) Motif enrichment in DA ATAC-seq peaks. Motifs and false discovery rates are for the underlined factors. (D) Overlap of published Mef2c, Tbx5, Gata4, Nkx2.5, and Shox2 ChIP-seq peaks with DA ATAC-seq peaks.

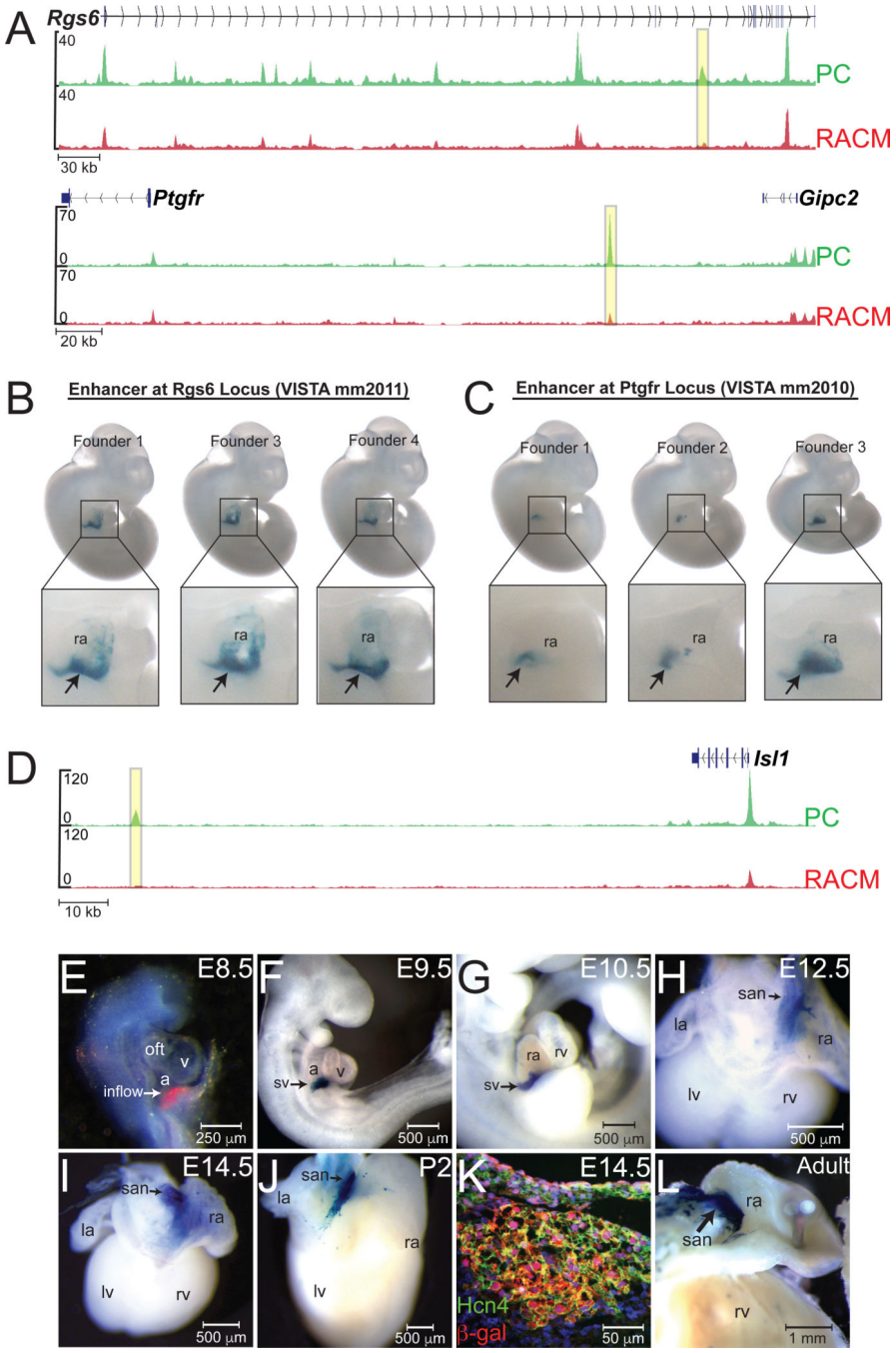


Figure 3. Selected PC-Enriched ATAC-seq Peaks Are Sufficient to Direct Reporter Activity to the Cardiac Venous Inflow and Sinoatrial Node.

(A) ATAC-seq signal in pacemaker cells (PC, green track) and right atrial cardiomyocytes (RACM, red track) at the *Rgs6* locus (top) and the *Ptgfr* locus (bottom) with differentially accessible (DA) ATAC-seq peaks highlighted in yellow. (B) 3 founders from transient transgenic enhancer reporter mice generated with enSERT for the DA ATAC-seq peak at *Rgs6* locus (VISTA ID mm2011) and (C) *Ptgfr* locus (VISTA ID mm2010), harvested at E11.5 and stained with X-gal. Lower panels show a magnified view of the venous inflow (arrow) and right atrium (ra). (D) DA ATAC-seq peak at the *Isl1* locus is highlighted in

yellow. (E-J) Whole mount images of *ISE-LacZ* embryos harvested and stained with salmon-gal at E8.5 (E) or bluo-gal at the developmental time points shown (F-J). (K) Double immunohistochemistry for Hcn4 and β-gal in the sinoatrial node (SAN) of E14.5 *ISE-LacZ* embryo. (L) Whole-mount image of an adult *ISE-LacZ* SAN stained with bluo-gal. (Abbreviations: ra, right atrium, rv, right ventricle, lv, left ventricle, la, left atrium, sv, sinus venosus, oft, outflow tract, san, sinoatrial node)

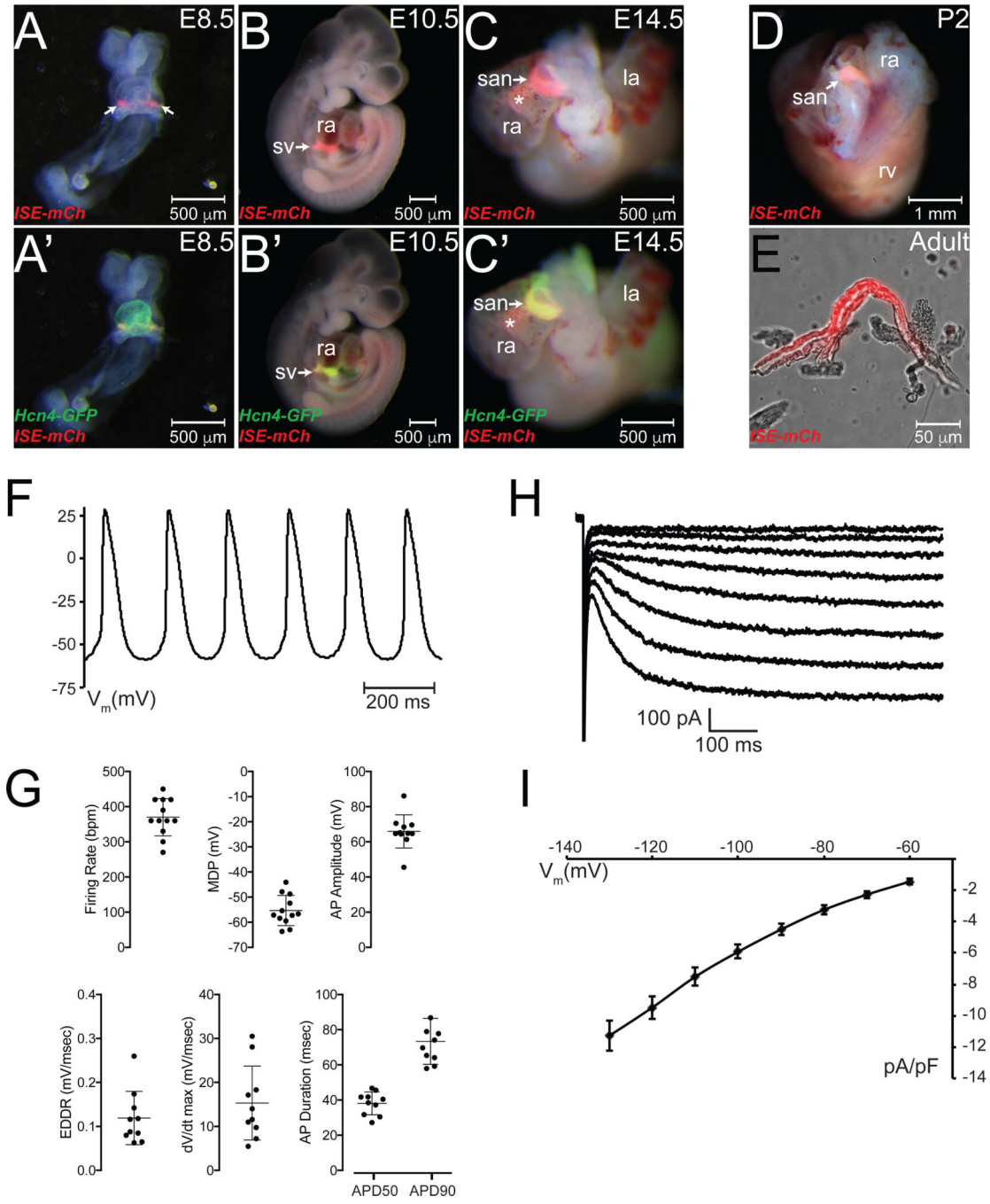


Figure 4. SAN Enhancer Activity Co-localizes with Hcn4 Expression and Functional Cardiac Pacemaker Cells.

(A,A') Whole-mount fluorescence imaging of *ISE-mCherry/Hcn4-GFP* embryos at E8.5 (arrows denote inflow tract), (B,B') E10.5, and (C,C') E14.5 imaged for mCherry alone (top row) and mCherry/GFP merge (bottom row). '*' denotes an mCherry+/GFP- SAN extension into the right atrium (D) Postnatal *ISE-mCherry* activity at P2 and (E) in adult isolated pacemaker cells (PCs) with typical PC morphology. (F) Spontaneous action potentials (APs) recorded from an adult mCherry+ isolated pacemaker cell. (G) Firing rate, maximum diastolic potential (MDP), action potential (AP) amplitude, early diastolic depolarization

rate (EDDR), upstroke velocity (dV/dt max) and AP duration at 50 percent and 90 percent repolarization in 12 isolated adult mCherry+ pacemaker cells. Error bars indicate standard deviation. (H) Whole cell voltage clamp on an mCherry+ adult cell with funny current (I(f)) elicited by a series of hyperpolarizing voltage pulses from a holding potential of −40 mV to −130 mV in 10 mV increments. (I) Mean I(f) current density at the indicated voltages in 11 mCherry+ PCs. Error bars indicate SEM. Abbreviations: sv, sinus venosus; ra, right atrium; la, left atrium; san, sinoatrial node.

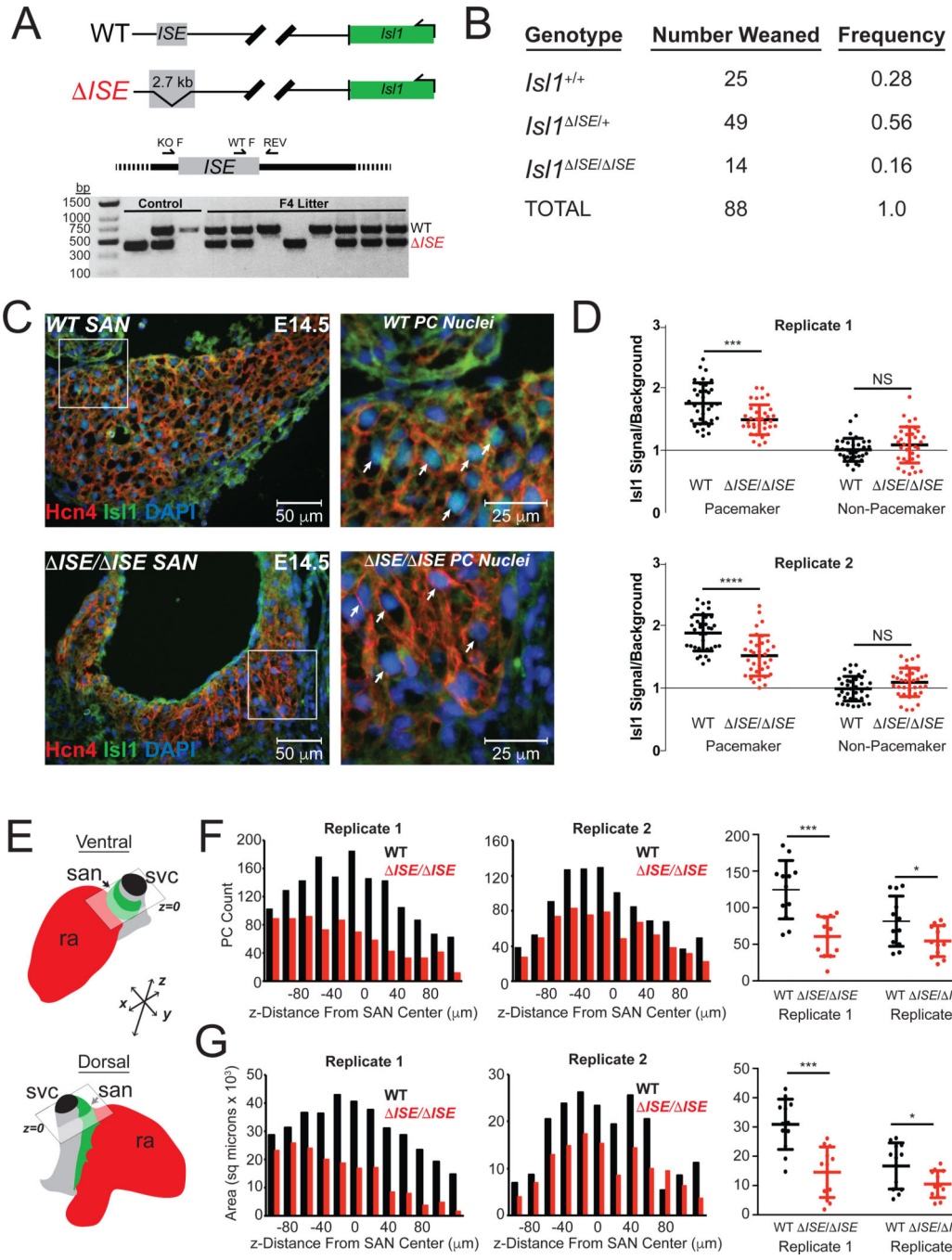


Figure 5. Deletion of the *Isl1* SAN Enhancer Leads to Reduced Isl1 Expression in SAN and Abnormal SAN Development

(A) (Top) Deletion of a 2.7-Kb genomic segment containing the Isl1 locus sinoatrial node enhancer (ISE) (Middle) Genotyping strategy to identify the mutant allele with an agarose gel (Bottom) on an F4 intercross confirming germline transmission of the deletion allele. (B) Genotypes from 88 weaned pups of $Isl1^{\Delta ISE/+}$ intercrosses with the indicated frequencies. (C) Immunohistochemistry of WT (top) and $Isl1^{\Delta ISE/\Delta ISE}$ (bottom) SAN at E14.5 stained for Hcn4 (red), Isl1 (green), and DAPI. Area of detail is shown in the right panels. Pacemaker cell nuclei in the WT are Isl1+ (turquoise, denoted by white arrows) whereas the

Isl1$^{\Delta ISE/\Delta ISE}$ has less Isl1 signal in PC nuclei. (D) Quantification of green fluorescent signal (Isl1) in 35 Hcn4+ pacemaker cell nuclei and 35 Hcn4- non-PC nuclei, normalized to tissue background green fluorescent signal, from 7 sections each from two different sets of WT and *Isl1*$^{\Delta ISE/\Delta ISE}$ littermates (denoted 'replicate 1' and 'replicate 2'). '****' denotes $p < 0.0001$ and '***' denotes $p < 0.001$. (E) Junction of the superior vena cava (svc) and right atrium (ra) showing the location of the sinoatrial node (san) and the sectioning plane used to perform pacemaker cell counts and cross sectional area measurements. Z=0 denotes the center of the sinoatrial node head. (F) Quantification of PC nuclei in 12 sections each at the indicated z-distance from the center of the SAN head. Counts were performed for 2 pairs of E14.5 littermates (denoted 'Replicate 1' and 'Replicate 2'). (G). Estimation of area of the SAN head for the sections analyzed in (F). '*' denotes $p < 0.05$.

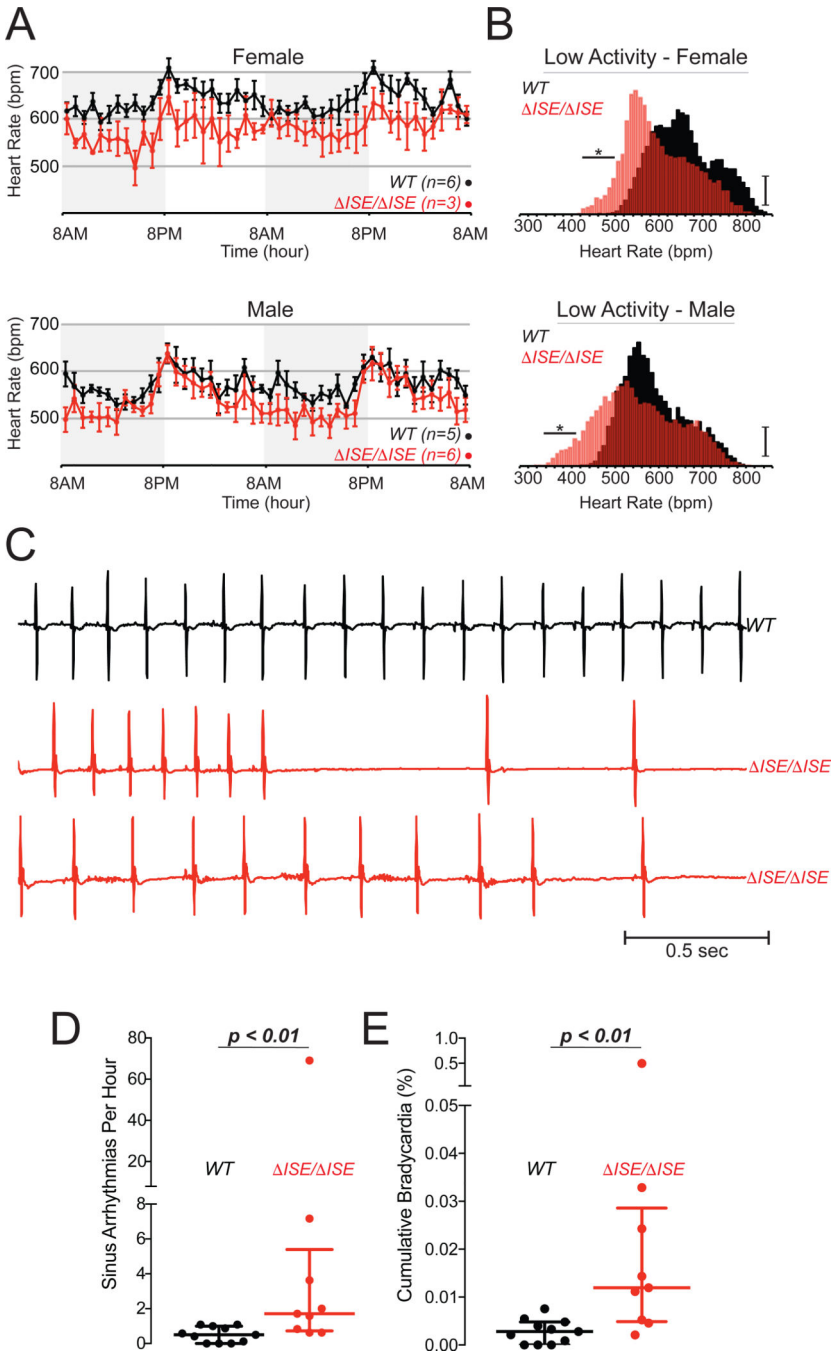


Figure 6. Abnormal SAN Function in Adult *Isl1*$^{\Delta ISE/\Delta ISE}$ Mice

(A) Average hourly heart rate in adult *Isl1*$^{\Delta ISE/\Delta ISE}$ (red, $n = 3$ for female and $n = 6$ for male) and WT littermate mice (black, $n = 6$ for female and $n = 5$ for male). Error bars denote SEM. The low activity period is highlighted in gray (B) Heart rates were averaged over 1-minute intervals for each mouse studied in (A) ($n = 11$ WT and 9 *Isl1*$^{\Delta ISE/\Delta ISE}$, analyzed separately by sex), and then binned into 10 beat per minute bins for generation of heart rate histograms for WT and *Isl1*$^{\Delta ISE/\Delta ISE}$ females (top) and males (bottom) Vertical scale bar = 25 counts. Comparison was made between WT and *Isl1*$^{\Delta ISE/\Delta ISE}$ for total counts less than 400 bpm

(male) or 500 bpm (female) for each mouse. '*' denotes $p < 0.05$ for Mann-Whitney test. (C) A 2-second electrocardiogram tracing from a WT mouse (top) and with 2 examples of sinus arrhythmias observed in 2 different *Isl1*$^{\Delta ISE/\Delta ISE}$ mice: a sinus pause (middle tracing), and an episode of bradycardia (bottom tracing) (D) Quantification of numbers of sinus arrhythmias recorded per hour over a 24-hour period in $n = 11$ WT and $n = 9$ KO mice. (E) Quantification of total percentage of time spent with heart rate less than 250 beats per minute (bradycardia) for 11 WT and 9 mutant mice. Comparison of WT and *Isl1*$^{\Delta ISE/\Delta ISE}$ distribution were made with a Mann-Whitney test with the p value shown.

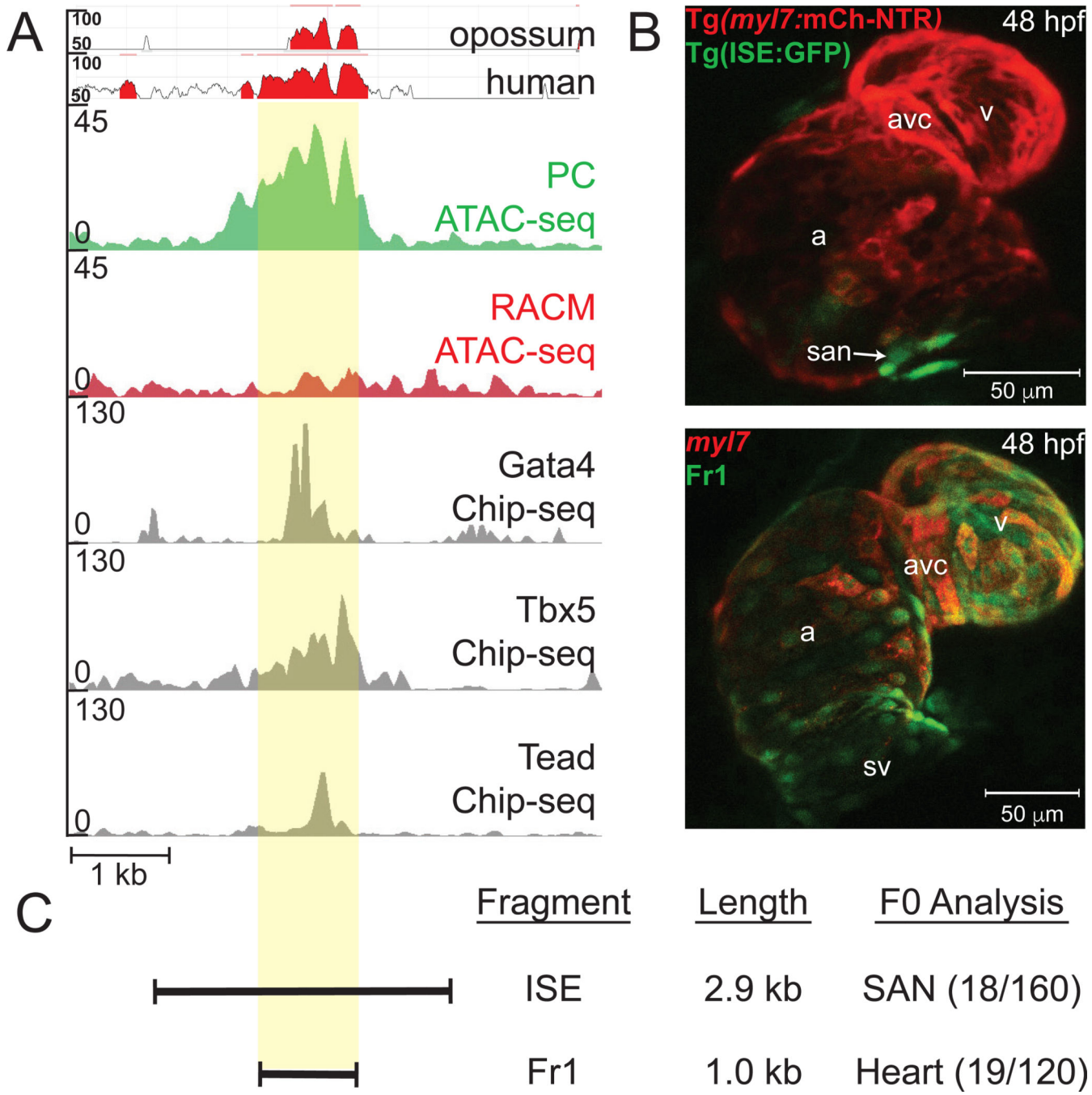


Figure 7. Evolutionary Conservation and Upstream Regulation of Isl1 SAN Enhancer Function

(A) View of *ISE* with alignment to human, mouse, and opossum genomes (above, areas in red indicate greater than 70% conservation, scale from 50% to 100%) and previously published embryonic heart ChIP-seq data for Gata4, Tbx5, and Tead. (B) Maximum intensity projections of confocal z-stacks of 48 hpf Tg(*myl7:mCherry-NTR)* zebrafish hearts after injection of *ISE:GFP* reporter construct (top panel), or a 1.0-kb deletion fragment (Fr1, lower panel). Abbreviations: san, sinoatrial node, sv, sinus venosus, a, atrium, avc, av canal, v, ventricle. (C) Enhancer-reporter analysis of founders from injections of ISE and a 1-kb

fragment (Fr1) in 48 hpf zebrafish hearts with the indicated patterns of reporter activity. Numbers in parentheses indicate the embryos observed to have the indicated expression pattern/total embryos injected that survived to 48 hpf.

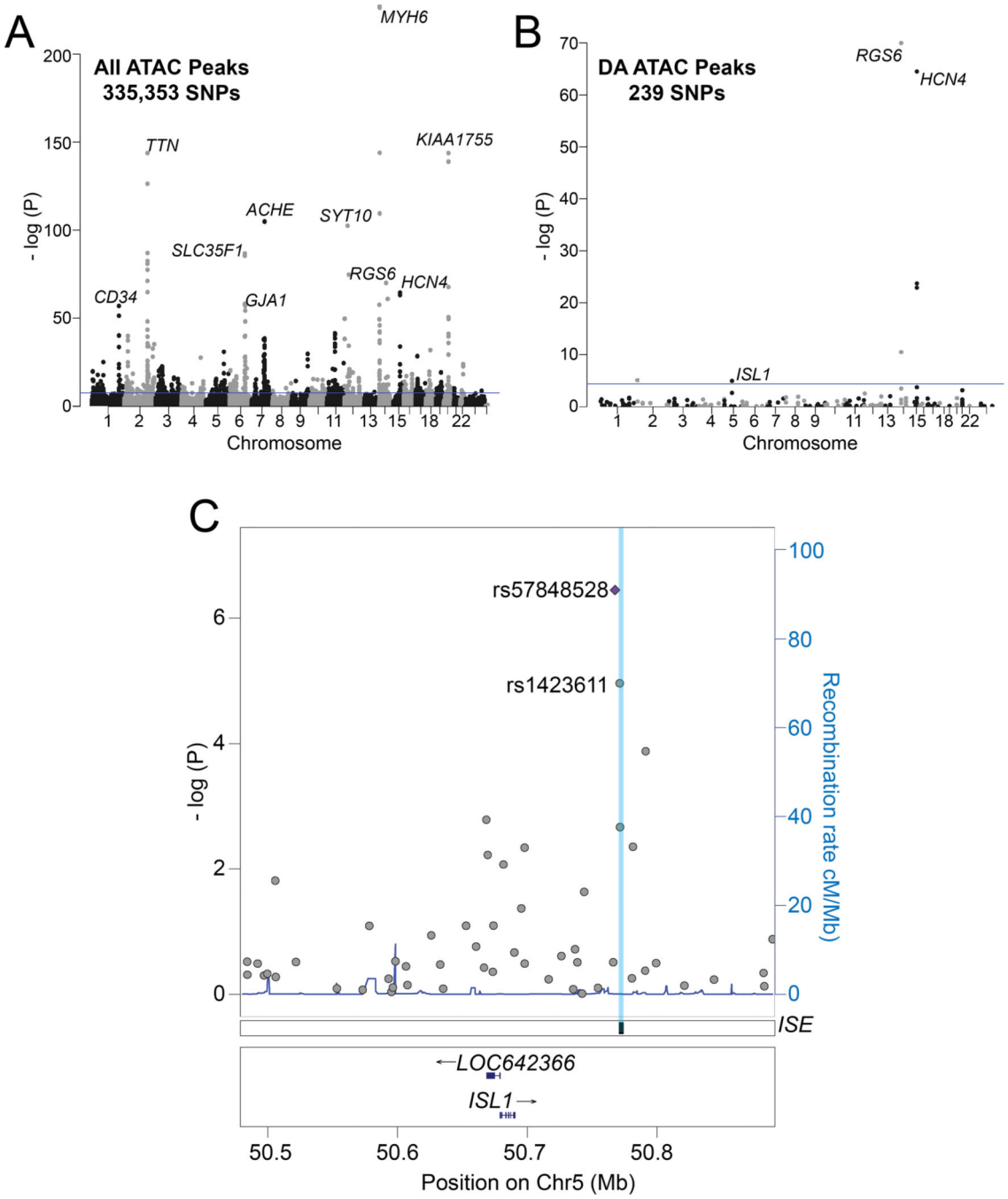


Figure 8. Association of Pacemaker Cell ATAC-seq Peaks with Resting Heart Rate Associated Single Nucleotide Polymorphisms.

(A) Manhattan plot showing association of 335,353 SNPs within or ± 0.5 kb away from a region syntenic to any mouse ATAC-seq peak with human resting heart rate ascertained from the UK Biobank cohort. Blue line indicates significance threshold of Bonferroni adjusted *p* value of 0.01 (corresponding to uncorrected *p* value 1.2×10^{-8}) (B) Manhattan plot similar to (A) including 239 SNPs within or ± 0.5 kb away from the top 500 differentially accessible (DA) ATAC-seq peaks. Blue line indicates Bonferroni adjusted *p* value of 0.01 (corresponding to uncorrected *p* value 4.2×10^{-5}). (C) Manhattan plot showing association of

SNPs at the *ISL1* locus with resting heart rate. The human genomic region syntenic to the mouse *Isl1* SAN enhancer is highlighted.