RNA Sequencing of Mouse Sinoatrial Node Reveals an Upstream Regulatory Role for Islet-1 in Cardiac Pacemaker Cells

Vasanth Vedantham, Giselle Galang, Melissa Evangelista, Rahul C. Deo, Deepak Srivastava

**Rationale:** Treatment of sinus node disease with regenerative or cell-based therapies will require a detailed understanding of gene regulatory networks in cardiac pacemaker cells (PCs).

**Objective:** To characterize the transcriptome of PCs using RNA sequencing and to identify transcriptional networks responsible for PC gene expression.

**Methods and Results:** We used laser capture microdissection on a sinus node reporter mouse line to isolate RNA from PCs for RNA sequencing. Differential expression and network analysis identified novel sinoatrial node–enriched genes and predicted that the transcription factor Islet-1 is active in developing PCs. RNA sequencing on sinoatrial node tissue lacking Islet-1 established that Islet-1 is an important transcriptional regulator within the developing sinoatrial node.

**Conclusions:** (1) The PC transcriptome diverges sharply from other cardiomyocytes; (2) Islet-1 is a positive transcriptional regulator of the PC gene expression program. (Circ Res. 2015;116:797-803. DOI: 10.1161/CIRCRESAHA.116.305913.)

**Key Words:** Hcn4 protein, mouse ■ high-throughput RNA sequencing ■ Isl1 protein, mouse ■ laser capture microdissection ■ sinoatrial node ■ transcription factors

Pacemaker cells (PCs) within the sinoatrial node (SAN) generate the electric impulse that initiates each heartbeat. Sinus node dysfunction, characterized by loss of PCs and SAN fibrosis, is a common but poorly understood disease. With the advent of cellular reprogramming technology, there is increasing interest in understanding gene regulatory networks in PCs, both to improve our understanding of sinus node dysfunction and to facilitate novel therapies such as biological pacemakers and SAN regeneration.

Transcription factors important for SAN development and function have been identified, including Tbx18,1 Tbx3,4 Shox2,5 and Isl1.6,7 Efforts to reprogram non-PCs into PCs have used some of these SAN transcription factors,8–10 although their downstream transcriptional networks have not been fully elucidated. Although these efforts are highly promising, there have been no genome-wide transcriptome comparisons of putatively reprogrammed PC-like cells with bona fide PCs, largely because of the absence of PC transcriptome data. Transcriptional profiling of PCs presents special challenges because of the lack of specific molecular markers, the small size of the developing SAN, and the interdigitation of PCs with non-PCs in the SAN. At present, the lack of transcriptome data from PCs remains a barrier to further progress in understanding SAN biology and in assessing and improving the fidelity of PC reprogramming technology.

**Methods**

**Laser Capture Microdissection**

Embryos or hearts were removed intact, washed with cold PBS, and immediately embedded in optimum cutting temperature medium (OCT) and stored at −20°C until sectioning. Tissue was sectioned at a thickness of 7 μm onto membrane-coated slides (MembraneSlide NF 1.0 PEN, Zeiss Microscopy, Gottingen, Germany). For laser capture, slides were thawed to room temperature until evaporation of moisture (≈1 minute), placed on the microscope stage of a PALM Micro-Beam inverted microscope with laser capture microdissection (LCM) capability (Zeiss). The sinus node tissue was identified visually and outlined manually with the microscope user interface (Online Movie II). Laser power and catapult energy were optimized before the experiment and varied from...
experiment to experiment. After each experiment, sections were stained with 4',6-diamidino-2-phenylindole and anti-Hcn4, mounted, and visualized to confirm accurate dissection of the region of interest.

RNA-Seq data was deposited in the NCBI’s Gene Expression Omnibus (GEO series accession number GSE65658).

A detailed Methods section is available in the Online Data Supplement.

Results

LCM of Cardiac PCs for RNA Sequencing

Hcn4-green fluorescent protein (GFP) bacterial artificial chromosome (BAC) transgenic mice exhibited high-level GFP expression in the SAN and venous pole at embryonic day (E) 12.5, with further restriction of expression to the SAN at postnatal day 0 and in the adult heart (Figure 1A).11 GFP expression was complementary to Connexin-40, a marker of working atrial myocytes, and isolated adult GFP+ cells showed morphology typical of PCs and exhibited spontaneous beating (Figure 1B; Online Movie I). We isolated SAN tissue and right atrial (RA) myocardium by LCM on unfixed, unstained cryosections from flash-frozen Hcn4-GFP hearts harvested at E14.5, postnatal day 4, and postnatal day 14 (Figure 1C; Online Movie II). The SAN is heterogeneous and includes non-PC cells such as fibroblasts, endothelial cells, and atrial cardiomyocytes. To optimize enrichment for PCs, we selected regions of interest for LCM within the SAN head (near the SAN artery) that contained purer populations of PCs (Online Figure I). Such regions within the SAN, as well as neighboring RA regions, were excised from 6 consecutive embryonic heart sections and pooled to generate total RNA. We repeated the LCM experiment for 3 different embryos for each time point to generate 18 independent RNA samples, reflecting 3 biological replicates for each tissue and time point (Online Table I). Each sample was processed separately for RNA sequencing as described in Methods in the Online Data Supplement.
Differential Expression Analysis

Differential expression (DE) analysis of RNA sequencing data between SAN and RA tissue samples revealed hundreds of DE genes at each time point (Online Table II). Known SAN-associated genes Hcn4, Tbx3, Shox2, Tbx18, and Bmp4 were all enriched in the SAN tissue, whereas RA-associated genes Nkx2-5, Nppa, and Gja1 were enriched in the RA samples (Online Figure I), demonstrating the fidelity of tissue isolation. Of note, the core cardiac transcription factors Gata4, Mef2c, and Tbx5 were not differentially expressed between SAN and RA. We also found numerous genes enriched in the SAN tissue that had not been previously associated with PCs (Figure 2A; Online Table II). Gene ontology terms associated with SAN-enriched genes included Bmp and Wnt signaling pathways at E14.5 and neuronal development and function at later time points (Figure 2A; Online Table III). Conversely, RA-enriched genes were associated with gene ontology terms that included conduction, contractile apparatus, and cell junction formation (Figure 2B; Online Table III). Although a core set of genes in each tissue type exhibited DE at all time points examined (Online Table IV), there was considerable change over time within the DE gene set, highlighting the dynamic nature of...

Figure 2. Comparative expression analysis of sinoatrial node (SAN) and right atrium (RA). A, Venn diagram showing overlap of SAN-enriched genes at embryonic day (E) 14.5, postnatal day 4 (P4), and P14 using a false-discovery rate cutoff of 0.05. Selected enriched gene ontology terms with example genes are shown for each time point. B, Similar Venn diagram as (A), but for RA-enriched genes. C, Weighted gene correlation network analysis identified 11 highly correlated modules within SAN/RA data set. Module eigengene activity was computed for each module for each of the 18 SAN/RA expression time points. Color key indicates activity scale. D, Expression of SAN-enriched transcription factors within module 1 at E14.5. Transcript levels are shown on the x axis and differential expression (SAN vs RA) on the y axis.
expression during SAN development. Hierarchical clustering of SAN and RA samples revealed that biological replicates clustered together, and that, as differentiation progressed, SAN samples were more similar to each other than they were to the RA samples (Online Figure II).

Network Analysis
We used weighted gene correlation network analysis to partition the RA and SAN transcriptomes into modules exhibiting correlated gene expression (Online Figure III and Online Table V). Most time points of specific tissue were associated with ≥1 highly active module. Module 1 (M1) exhibited highest activity at E14.5, a critical period for SAN morphogenesis and PC differentiation (Figure 2C). M1 contained several previously identified SAN-enriched transcription factors, including Tbx18, Shox2, and Isl1, whereas Tbx3 clustered with Hcn4 and Hcn1 in M7. Of the factors in M1, Isl1 had the highest level of DE as well as high transcript abundance early in SAN development (Figure 2D).

Conditional Selection of Isl1 After Second Heart Field Differentiation
To test for a requirement of Isl1 in the PC gene expression program, we crossed Isl1loxp/loxp; CAG-Cre/Esr1 with Isl1loxp/loxp; Hcn4-GFP and injected intraperitoneal tamoxifen at E10.5 (Figure 3A). This strategy generated a global deletion of Isl1 after second heart field differentiation, which circumvented the early embryonic lethality associated with Isl1 loss of function. Cre+; Hcn4-GFP embryos were recovered at Mendelian ratios at E12.5, indicating that loss of Isl1 after second heart field differentiation did not lead to rapid embryonic demise. SAN Hcn4-GFP expression was reduced but readily detectable in Cre+ embryos (Figure 3B).

We then performed LCM on Cre+ and Cre− embryonic SAN at E12.5. Each sample was pooled from ≈9 SAN sections per embryo, and we used 7 different embryos (3 for Cre+ group and 4 for Cre− group; Figure 3B). Quantitative polymerase chain reaction performed on amplified cDNA before RNA
sequencing library preparation could not detect *Isl1* transcript in the samples from Cre⁺ embryos, a finding confirmed by the absence of RNA sequencing reads mapped to the floxed portion of *Isl1* (Online Figure IV).

**SAN Transcriptome in the Absence of Isl1**

A total of 590 genes exhibited DE between Cre⁺ and Cre⁻ RNA samples, of which 65% were downregulated in SAN tissue lacking *Isl1* (Figure 3C; Online Table VI). Among the genes enriched in SAN tissue versus RA at E14.5, 18% were downregulated after deletion of *Isl1* (odds ratio, 14; \(P=1.1 \times 10^{-88}\), compared with non-DE genes), indicating a positive role for *Isl1* in the PC-specific gene expression program. Of genes with established roles in SAN development and PC function, *Tbx3, Hcn4, Hcn1,* and *Cacna1g* were all significantly decreased after deletion of *Isl1*, whereas expression of other cardiac and SAN TFs including *Nkx2.5, Gata4, Mef2C, Tbx5,* and *Tbx18* was not significantly changed (Figure 3D). Conversely, 9% of RA-enriched genes (including *Nppa, Gja1, Gja5,* and *Scn5a*) were upregulated in SAN after loss of *Isl1* (odds ratio, 10; \(P=9.6 \times 10^{-55}\)), suggesting a secondary role for *Isl1* in repressing a differentiation program associated with working myocardium.

**Network and Motif Analysis on Dysregulated Genes**

Assessment of individual network activity after loss of *Isl1* function showed that M7 (containing *Hcn4* and *Tbx3*), followed by M1 (containing *Tbx18, Shox2,* and *Isl1*), were the most reduced in overall activity (Online Figure VI). Conversely, the only modules that exhibited increased activity after loss of *Isl1*, M3 and M9, were associated with increased activity in RA. To relate these global changes in expression to loss of *Isl1* function, we searched for *Isl1* binding sites within introns, untranslated regions, and 10 kb upstream of the transcriptional start sites of genes that were downregulated after loss of *Isl1* and compared the frequency of occurrence of these sites to genes that were unaffected by *Isl1* deletion. This analysis showed highly significant enrichment (\(P=0.003\) of *Isl1* binding sites among the genes downregulated after *Isl1* deletion, suggesting a role for *Isl1* as a direct transcriptional activator in the developing SAN (Figure 4A; Online Figure V). Examples of putative *Isl1* binding sites in *Tbx3* and *Hcn4* regulatory regions are shown in Online Table VII. Because of the limiting number of PCs in each embryonic heart, it was not possible to test these sites directly using chromatin immunoprecipitation. Conversely, we did not observe enrichment of *Isl1* binding sites near the genes upregulated after *Isl1* deletion, suggesting that *Isl1* may not function primarily as a direct transcriptional repressor of these genes but may suppress the atrial gene expression program indirectly (Figure 4B; Online Figure V).

**Sensitivity Analysis**

Because inadvertent contamination of a Cre⁺ SAN sample with non-PC myocytes would bias our analysis toward the observed effects, we assessed our SAN samples for the presence of outliers. We found that 1 of the 3 Cre⁺ samples consistently exhibited more dramatic changes in gene expression at more diverse loci than the other 2 samples, raising the possibility of either contamination by non-PC myocytes or simply biological variation. To insure the robustness of our findings, we repeated our DE analysis and motif analysis after excluding this outlier sample and observed no significant changes in our results (Online Table VIII and Online Figure VI).

**Discussion**

Here, we present a comprehensive transcriptional analysis of mammalian PCs by deep sequencing SAN tissue obtained
with LCM. Although PCs and RA myocytes share core transcriptional machinery, mature PCs have adopted a distinct transcriptional program, consistent with morphological and electric phenotypes exhibited by PCs. Our data set highlights dynamic changes in PC gene expression associated with differentiation and maturation and reveals a central role of Isl1 in establishing the PC gene program.

Role of Isl1 in Establishing the SAN Gene Network

Isl1 is required for second heart field development in mice and for SAN development in zebrafish. Although Isl1 expression and function in mouse PCs have been described previously, its transcriptional role in the SAN remains unknown. Our loss of function data demonstrated that SAN-enriched genes were strongly over-represented among the genes downregulated after Isl1 disruption, suggesting that Isl1 is required to generate the SAN-specific gene expression program. Furthermore, we observed significant enrichment of Isl1 binding sites within and around downregulated genes, providing strong evidence that Isl1 is a direct transcriptional activator of gene expression in PCs.

We also observed upregulation of genes associated with RA myocardium in the SAN after deletion of Isl1, including Nppa, Scn5a, Gja1, and Gja5. However, we did not observe enrichment of Isl1 binding sites near the upregulated genes, suggesting that Isl1 functions indirectly to block the RA myocyte gene expression in the SAN. Tbx3 is an established negative regulator of the atrial myocyte gene expression program and was found in our analysis to be highly downregulated after deletion of Isl1. One possibility is that upregulation of RA genes after Isl1 deletion is mediated in part by downregulation of Tbx3 (Figure 4C). This model situates Isl1 as a key transcriptional activator in PCs with a major role in establishing the SAN-specific gene expression program. As the scope of the present study is limited to expression analysis, further work will be required to delineate the functional and developmental roles of Isl1 in SAN and to validate putative direct transcriptional targets. In addition, the potential of Isl1 to augment PC derivation or reprogramming has yet to be fully explored. Our analysis suggests that multifactor reprogramming using Isl1 in combination with other cardiac and SAN transcription factors may be a fruitful approach.

References


What Is Known?

• Cardiac pacemaker cells (PCs) in the sinoatrial node (SAN) have a distinct gene expression program that underlies their unique functional properties but are difficult to isolate in sufficient quantity and with adequate purity required for expression profiling.

• Non-PC myocytes can be reprogrammed to exhibit PC-like behavior using SAN transcription factors, with varying efficiency and uncertain fidelity.

What New Information Does This Article Contribute?

• Pure SAN tissue was isolated at mid-development, the perinatal period, and after maturation for comprehensive expression profiling using RNA sequencing.

• Differential expression analysis of SAN tissue and right atrial tissue uncovered numerous SAN-enriched genes not previously associated with PCs.

• RNA sequencing of SAN tissue in the absence of the transcription factor Islet-1 (Isl1) showed that Isl1 plays an important role in establishing the PC gene expression program during heart development.

Novelty and Significance

The regeneration of SAN or the creation of biological pacemakers using cellular reprogramming are promising new approaches to the treatment of sinus node disease. Further development of these therapies will require a detailed understanding of transcriptional networks in specialized cardiac PCs. However, these cells are difficult to characterize because they are few in number, they lack of specific molecular markers, and they are interspersed among non-PC cells. In this study, we used laser capture microdissection to isolate purer populations of PCs from embryonic, perinatal, and postnatal SAN for RNA sequencing. Analysis of RNA sequencing data uncovered numerous genes expressed in SAN that were not previously associated with PCs. Based on high differential expression, high absolute expression, and a weighted gene correlation network analysis, the transcription factor Isl1 was predicted to regulate PC-specific gene expression. RNA sequencing on SAN tissue lacking Isl1 confirmed this prediction and established Isl1 as an important transcriptional regulator within the developing SAN. The results of this study could form the basis of further mechanistic investigations into the gene regulatory networks active in the SAN and future translational work on sinus node disease.
RNA Sequencing of Mouse Sinoatrial Node Reveals an Upstream Regulatory Role for Islet-1 in Cardiac Pacemaker Cells
Vasanth Vedantham, Giselle Galang, Melissa Evangelista, Rahul C. Deo and Deepak Srivastava

Circ Res. 2015;116:797-803; originally published online January 26, 2015;
doi: 10.1161/CIRCRESAHA.116.305913

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circres.ahajournals.org/content/116/5/797

Data Supplement (unedited) at:
http://circres.ahajournals.org/content/suppl/2015/01/26/CIRCRESAHA.116.305913.DC1

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation Research can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation Research is online at:
http://circres.ahajournals.org//subscriptions/
SUPPLEMENTAL MATERIAL

DETAILED METHODS

Genetically Modified Mouse Lines
Mice were maintained on mixed backgrounds. Hcn4-GFP was generated by the Gene Expression Nervous System Atlas (GENSAT) Project, and obtained from the Mutant Mouse Regional Resource Center (Stock number 030842-UCD; Tg(Hcn4-EGFP)JK158Gsat/Mmucd). CAG-cre/Esr1 mice were obtained from Jackson Labs (Stock number 004453; Tg(CAG-cre/Esr1*)5Amc/j). Isl1/loxp mice were generated by Lin Gan’s lab and were provided by John Rubenstein’s lab at University of California, San Francisco in a mixed background carrying the Ai14 reporter allele (B6.Cg-Gt(Rosa)26Sortm14(CAG-tdTomato)Hzel/J). To induce Cre recombinase activity, 0.2 mg/g intra-peritoneal tamoxifen was injected into timed pregnant female mice. Because of the presence of the Ai14 allele, Cre+;Isl1flox/flox embryos could be genotyped rapidly by visualizing live red fluorescence and processed for LCM.

Whole Mount In Situ Hybridization
Whole-mount in-situ hybridization for Hcn4 was performed as previously described.2

Immunohistochemistry and Immunocytochemistry
Embryos or hearts were fixed in 4% PFA, washed with PBS, incubated sequentially in 10%, 20% and 30% sucrose, then embedded in OCT. Sections were permeabilized with 0.3% Triton X-100, blocked with 10% normal goat serum, and incubated with primary antibody for alpha actinin (Sigma, 1:1000 dilution), Hcn4 (Alomone labs 1:200), Connexin-40 (Alpha Diagnostics, 1:200), overnight at 4°C. Slides were washed, incubated for one hour with the secondary antibody before mounting in Vectashield with DAPI (Vector Laboratories).

Isolation of Cardiac Pacemaker Cells
12 week-old adult male and female Hcn4-GFP hearts were removed and the intercaval region containing the sinoatrial node region was identified by fluorescence microscopy. The sinoatrial tissue was manually cut into 3–4 strips per heart and placed in digestion solution (in mM: 140 NaCl, 5.4 KCl, 1.0 MgCl2, 1.8 CaCl2, 1.2 KH2PO4, 5.0 HEPES, 50 Taurine, 5.5 glucose, pH 6.9) supplemented with 1 mg/mL bovine serum albumin (Sigma), 229 Units/mL collagenase type 1 (Worthington Biochemical, Lakewood, NJ), 0.9 Units/mL protease type XIV (Sigma), and 1.9 Units/ML elastase (Worthington). Tissue strips were digested for 25 minutes at 37 °C, with agitation every 5 minutes. The strips were then transferred into KB solution (in mM: 100 K-glutamate, 20 glucose, 20 taurine, 20 K-aspartate, 10 KH2PO4, 5.4 KCl, 5 creatine, 5 HEPES, 0.5 EGTA, supplemented with 1mg/mL bovine serum albumin, pH 7.2), triturated for 10 minutes, and incubated at 4 degrees for 1 hour. Drops of this solution were placed onto laminin-coated culture slides and used for live imaging and immunocytochemistry after gradual reintroduction of calcium-containing tyrode solution (in mM: 140 NaCl, 5.4 KCl, 1.0 mgCl2, 1.8 CaCl2, 5.0 glucose, 5.0 HEPES, pH 7.4).

Laser Capture Microdissection
Embryos or hearts were removed intact, washed with cold PBS, and immediately embedded in OCT and stored at -20 °C until sectioning. Tissue was sectioned at a thickness of 8 microns onto membrane-coated slides (MembraneSlide NF 1.0 PEN, Zeiss Microscopy, Gottingen, Germany). For laser capture, slides were thawed to room temperature until evaporation of moisture (approximately 1 minute), placed on the microscope stage of a PALM Micro-Beam inverted microscope with LCM capability (Zeiss). The sinus node tissue was identified visually and outlined manually with the microscope user interface (Online Video 2). Laser power and catapult energy were optimized prior to the experiment and varied from experiment-to-
experiment. After each experiment, sections were stained with DAPI and anti-Hcn4, mounted, and visualized to confirm accurate dissection of the region of interest.

**RNA Sequencing**

Total RNA was prepared from regions of interest isolated with laser capture using the PicoPure Kit (Arcturus) according to manufacturers instructions. RNA-seq libraries were prepared with the Ovalion RNA-Seq System V2 kit (NuGEN) according to manufacturers instructions. Briefly, total RNA is reverse-transcribed using a combination of random hexamers and a poly-T chimeric primer, followed by second strand synthesis using DNA polymerase. The cDNA was then amplified using single primer isothermal amplification (SPIA). Quantitative PCR for Isl1 was performed on amplified cDNA using a Taqman primer spanning the exon 2/3 junction. Libraries from the SPIA amplified cDNA were made using the Ultralow DR library kit (NuGEN) according to manufacturers instructions. The RNA-seq libraries were analyzed by Bioanalyzer (Agilent) and quantified by QPCR (KAPA). High-throughput sequencing was done using a HiSeq 2500 instrument (Illumina) with 50 million reads per sample, of which approximately half could be uniquely assigned.

**Differential Expression Analysis**

Reads were mapped to mm9 build of the mouse gene using TopHat. Differential expression analysis was performed using the edgeR package, with trimmed mean of M values (TMM) normalization. Gene Ontology (GO) enrichment was performed using Funcassociate 2.1, which uses a Fisher’s Exact Test for enrichment, and permutation based correction for multiple hypotheses. Gene lists were filtered prior to submission to the server to include only those genes annotated by one or more GO terms (18,615 genes are currently annotated). Heat maps were generated using the hmap function in the seriation package. Hierarchical clustering was performed on the 2500 most variable genes (given by coefficient of variation) in the normalized data set.

**Gene Expression Network Analysis**

Weighted Gene Correlation Network Analysis (WGCNA) was performed on the same gene set using the WGCNA package in R. For WGCNA analysis, the correlation matrix was raised to the 6th power to approximate a power law degree distribution. Module activity for each gene expression time point was computed using the factor loading of the first principal component of the matrix of expression values for genes in the module (described as the first “eigengene” in WGCNA). A few genes could not be assigned to one of the 11 modules and were placed in “module 0”. The impact of Isl1 knockout on gene expression of each module was taken as the mean fold-change (log2) computed over all genes in the module. Overlap between differential expression sets was performed using the Fisher’s Exact Test, assuming a universe of 16,496 genes, as this number was detectable by RNA-Seq in SAN or RA tissue. Odds ratios were computed for the 2x2 contingency table to provide a measure of effect size.

**Motif Analysis**

Motif enrichment for the Isl1 motif was performed twice using two different position weight matrices (PWM), including Motif 112 (Homer), depicted in Online Figure V, derived from Isl1 ChIP-Seq data and PH0088.1 (JASPAR), shown in Figure 4, derived from a universal protein binding microarray for the closely related Isl2 protein. Isl2 binds identical sequences to Isl1. A comparison was made between the median number of predicted Isl1 binding sites in candidate regulatory regions for 383 genes downregulated in the Isl1−/− mice (“diffexp set”) and a set of 221 genes with no differential expression (FDR > 0.99, “null set”). To ensure consistency of the comparison, both gene sets were limited to those genes with a minimum of 5 counts per million (estimated using the cpm function in edgeR). FASTA sequences corresponding to the introns, 5′ UTR, 3′UTR and up to 10kB upstream from the transcription start site were used to identify binding sites as previously performed. Significance of enrichment was computed using bootstrap resampling, with 10,000 samples (matching the diffexp set in number) taken from the
distribution of number of binding sites for the null set, and a median computed each time. The empirical p-value was taken as the number of iterations with a median number of binding sites greater than or equal to that observed in the diffexp set. A measure of effect size can be taken as the ratio of median number of binding sites in the diffexp and null sets.
Online Figure I. Purity of Laser Capture Microdissection. (A) RNA Sequencing revealed enrichment of previously identified SAN associated genes, while atrial genes including Gja1 and Nppa were enriched in the RA. Y-axis shows log(2) of Fold Change. Core cardiac transcription factors Gata4, Tbx5, Mef2c are not significantly changed, while Nkx2.5 is enriched in RA at E14.5, and non-statistically significantly higher in RA at P4 and P14. Green and Red indicate false discovery rate (FDR) for differential expression < 0.1 with SAN>RA and RA>SAN, respectively. Blue indicates FDR > 0.1 for differential expression. (B) A magnified view of the sinoatrial node region excised with laser capture at P4 is shown. This region corresponds to the panels depicted in Figure 1C (middle row). The vast majority of nuclei within the SAN area are GFP+ pacemaker cells.
Online Figure II. Hierarchical bi-clustering of SAN and RA samples. The top 2500 most variable genes across all samples were used for clustering. SAN and RA samples are brown and purple, respectively. Red indicates higher expression (see color scale). Each row is scaled to mean of zero with standard deviation of 1. Optimal leaf reordering was used to ensure similar samples are grouped together.
Online Figure III. Weighted gene correlation network analysis (WGCNA). WGCNA identified 11 highly correlated modules within SAN/RA data set. Module assignment for each gene is color-coded. Module activity varied across samples and time points and distinguished SAN and RA.
Online Figure IV. Deletion of Isl1 from SAN. (A) Quantitative PCR on amplified cDNA from SAN LCM in 4 Cre negative and 3 Cre⁺ samples. The primer spanned the exon 2/3 junction. Using this assay, *Isl1* was not detectable in the Cre⁺ samples. (B) The *Isl1* genomic locus is shown, with each row depicting mapped RNA sequencing reads for each sample (4 Cre⁻ samples and 3 Cre⁺ samples – sample names indicated at left). Vertical axis for each row was normalized to highest read count in that sample, indicated in the column on the right). In the *Isl1*loxp allele, loxp sites flank exon 2; Cre⁺-mediated deletion removes exon 2 and introduces a frameshift mutation that results in mRNA that is not translated into function protein. Note the absence of reads mapped to exon 2 (highlighted) in the Cre⁺ samples, indicating successful disruption of *Isl1* in these samples. Abbreviations: EX, exon. Kb, kilobase.
Online Figure V. Inferring Isl1 Function from Network and Motif Analysis (A) The heat map shows module activity after deletion of *Isl1*. Blue denotes reduced activity and red denotes increased activity. Module 7, containing *Tbx3* and *Hcn4*, was the most downregulated. Other modules active in SAN were also reduced in activity. Modules 3 and 9, which were more active in the right atrium, were increased in activity. (B) Motif analysis using a position weight matrix based on Isl1 ChIP-Seq data. Isl1 binding sites are enriched among the downregulated genes (left) but unchanged among upregulated genes (right), similar to the results of the analysis depicted in Figure 4.
Online Figure VI. Sensitivity Analysis. (A) Differential expression analysis between Isl1−/− and WT SAN samples was repeated after removal of one outlier Isl1−/− sample. Similar to the original analysis, there was highly significant overrepresentation of SAN enriched genes among the downregulated genes, and, conversely, overrepresentation of RA enriched genes among the upregulated genes. (B) Scatterplot showing downregulation of selected SAN genes after loss of Isl1 with upregulation of selected atrial genes. Some SAN genes and core cardiac transcription factors were unaffected. The results are similar to the data shown in Figure 3D for the complete dataset. (C) A motif analysis using two different position weight matrices (PH088.1 and Motif112) was performed to look for Isl1 binding sites within and 10 kb upstream of differentially expressed genes. Using both motifs, there was highly significant enrichment among the downregulated genes, but not the upregulated genes, similar to results shown in Figure 4 for the complete dataset.
SUPPLEMENTAL TABLES

Online Table I. Total RNA Yield for Individual Samples

<table>
<thead>
<tr>
<th>Tissue</th>
<th>Stage</th>
<th>Total RNA (pg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAN</td>
<td>E14.5</td>
<td>460</td>
</tr>
<tr>
<td>SAN</td>
<td>E14.5</td>
<td>730</td>
</tr>
<tr>
<td>SAN</td>
<td>E14.5</td>
<td>530</td>
</tr>
<tr>
<td>RA</td>
<td>E14.5</td>
<td>2340</td>
</tr>
<tr>
<td>RA</td>
<td>E14.5</td>
<td>2440</td>
</tr>
<tr>
<td>RA</td>
<td>E14.5</td>
<td>2400</td>
</tr>
<tr>
<td>SAN</td>
<td>P4</td>
<td>890</td>
</tr>
<tr>
<td>SAN</td>
<td>P4</td>
<td>350</td>
</tr>
<tr>
<td>SAN</td>
<td>P4</td>
<td>590</td>
</tr>
<tr>
<td>RA</td>
<td>P4</td>
<td>710</td>
</tr>
<tr>
<td>RA</td>
<td>P4</td>
<td>1770</td>
</tr>
<tr>
<td>RA</td>
<td>P4</td>
<td>2940</td>
</tr>
<tr>
<td>SAN</td>
<td>P14</td>
<td>400</td>
</tr>
<tr>
<td>SAN</td>
<td>P14</td>
<td>1090</td>
</tr>
<tr>
<td>SAN</td>
<td>P14</td>
<td>1180</td>
</tr>
<tr>
<td>RA</td>
<td>P14</td>
<td>1570</td>
</tr>
<tr>
<td>RA</td>
<td>P14</td>
<td>1600</td>
</tr>
<tr>
<td>RA</td>
<td>P14</td>
<td>2300</td>
</tr>
</tbody>
</table>

Online Tables II–VI are supplied as Excel files.

Online Table II. Differential Expression Analysis: SAN vs RA at 3 time points

Online Table III. GO analysis

Online Table IV. Lists of genes differentially expressed in SAN and RA at all three time points (FD < 0.05)

Online Table V. List of genes in each module

Online Table VI. Differential Expression Analysis: Isl1 null vs WT
Online Table VII. Putative Isl1 Binding Sites Identified by Motif Analysis

<table>
<thead>
<tr>
<th>Gene</th>
<th>Position (mm9)</th>
<th>Sequence (mm9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tbx3</td>
<td>chr5:120,126,769-120,126,783</td>
<td>AACTTAATTGATATA</td>
</tr>
<tr>
<td>Tbx3</td>
<td>chr5:120,127,165-120,127,179</td>
<td>ATAATTAAAGAC</td>
</tr>
<tr>
<td>Tbx3</td>
<td>chr5:120,131,650-120,131,664</td>
<td>GGGTTAATGGTTG</td>
</tr>
<tr>
<td>Tbx3</td>
<td>chr5:120,115,557-120,115,571</td>
<td>CAGTTAATTAGT</td>
</tr>
<tr>
<td>Tbx3</td>
<td>chr5:120,127,033-120,127,047</td>
<td>ATAATTAGATT</td>
</tr>
<tr>
<td>Hcn4</td>
<td>chr9:58,686,000-58,686,014</td>
<td>ACTCTAATGGATG</td>
</tr>
<tr>
<td>Hcn4</td>
<td>chr9:58,694,710-58,694,724</td>
<td>TTCCATAATGGCTT</td>
</tr>
<tr>
<td>Hcn4</td>
<td>chr9:58,700,568-58,700,592</td>
<td>AGATTATGT</td>
</tr>
<tr>
<td>Hcn4</td>
<td>chr9:58,700,867-58,700,882</td>
<td>CTAGTAATGGAGT</td>
</tr>
<tr>
<td>Hcn4</td>
<td>chr9:58,690,516-58,690,530</td>
<td>CTGTTAATTACTG</td>
</tr>
</tbody>
</table>

Online Table VIII (supplied as an excel file). Differential Expression Analysis, Isl1 null vs WT, after exclusion of outlier.
SUPPLEMENTAL REFERENCES


SUPPLEMENTAL VIDEO LEGENDS

**Online Video I. Spontaneously beating adult pacemaker cells.** The video shows a pair of freshly isolated calcium tolerant GFP+ cells from a digested adult SAN. GFP+ cells exhibited spontaneous beating and pacemaker-cell spindle morphologies.

**Online Video II. Laser Capture Micro-Dissection.** The video shows an example of laser capture on a P4 sinoatrial node from an Hcn4-GFP mouse cryosection. The region of interest is visualized with fluorescence microscopy and outlined manually on the user interface. The region of interest is then cut from the rest of the tissue and catapulted into a microcentrifuge tube for RNA isolation.