Tomo-Seq Identifies SOX9 as a Key Regulator of Cardiac Fibrosis During Ischemic Injury

Editorial, see p 1410

Grégory P.A. Lacraz, MSc, PhD*
Jan Philipp Junker, MSc, PhD*
Monika M. Gladka, MSc, PhD
Bas Molenaar, MSc
Koen T. Scholman, MSc
Marta Vigil-Garcia, MSc
Danielle Versteeg, BS
Hesther de Ruiter, BS
Marit W. Vermunt, MSc, PhD
Menno P. Creyghton, MSc, PhD
Manon M.H. Huibers, MSc, PhD
Nicolaas de Jonge, MD
Alexander van Oudenaarden, MSc, PhD
Eva van Rooij, MSc, PhD

BACKGROUND: Cardiac ischemic injury induces a pathological remodeling response, which can ultimately lead to heart failure. Detailed mechanistic insights into molecular signaling pathways relevant for different aspects of cardiac remodeling will support the identification of novel therapeutic targets.

METHODS: Although genome-wide transcriptome analysis on diseased tissues has greatly advanced our understanding of the regulatory networks that drive pathological changes in the heart, this approach has been disadvantaged by the fact that the signals are derived from tissue homogenates. Here we used tomo-seq to obtain a genome-wide gene expression signature with high spatial resolution spanning from the infarcted area to the remote to identify new regulators of cardiac remodeling. Cardiac tissue samples from patients suffering from ischemic heart disease were used to validate our findings.

RESULTS: Tracing transcriptional differences with a high spatial resolution across the infarcted heart enabled us to identify gene clusters that share a comparable expression profile. The spatial distribution patterns indicated a separation of expressional changes for genes involved in specific aspects of cardiac remodeling, such as fibrosis, cardiomyocyte hypertrophy, and calcium handling (*Col1a2*, *Nppa*, and *Serca2*). Subsequent correlation analysis allowed for the identification of novel factors that share a comparable transcriptional regulation pattern across the infarcted tissue. The strong correlation between the expression levels of these known marker genes and the expression of the coregulated genes could be confirmed in human ischemic cardiac tissue samples. Follow-up analysis identified SOX9 as common transcriptional regulator of a large portion of the fibrosis-related genes that become activated under conditions of ischemic injury. Lineage-tracing experiments indicated that the majority of COL1-positive fibroblasts stem from a pool of SOX9-expressing cells, and in vivo loss of *Sox9* blunted the cardiac fibrotic response on ischemic injury. The colocalization between SOX9 and COL1 could also be confirmed in patients suffering from ischemic heart disease.

CONCLUSIONS: Based on the exact local expression cues, tomo-seq can serve to reveal novel genes and key transcription factors involved in specific aspects of cardiac remodeling. Using tomo-seq, we were able to unveil the unknown relevance of SOX9 as a key regulator of cardiac fibrosis, pointing to SOX9 as a potential therapeutic target for cardiac fibrosis.

*Drs. Lacraz and Junker contributed equally.

Correspondence to: Eva van Rooij, MSc, PhD, Hubrecht Institute, KNAW Department of Cardiology, University Medical Center Utrecht, Uppsalalaan 8, 3584CT Utrecht, The Netherlands. E-mail e.vanrooij@hubrecht.eu

Sources of Funding, see page 1408

Key Words: fibrosis ■ myocardial ischemia ■ SOX9 transcription ■ ventricular remodeling

Clinical Perspective

What Is New?

- SOX9 is a key regulator of cardiac fibrosis after ischemic injury in mice by regulating the expression of many extracellular matrix-related proteins.
- SOX9 is induced in cardiac tissue from patients suffering from ischemic heart disease and colocalizes with COL1 expression.
- Reduced levels of SOX9 lead to less cardiac fibrosis after ischemic injury in mice.
- Tomo-seq can be used to identify new players in cardiac biology and disease.

What Are the Clinical Implications?

- Our data suggest that therapeutic inhibition of SOX9 in the diseased heart could lead to a reduction in cardiac fibrosis.

Ischemic heart disease induces a heterogeneous remodeling response across the damaged area that involves fibroblast activation, cardiomyocyte hypertrophy, and changes in calcium handling, all of which are eventually detrimental for cardiac function.[1,2] Fibroblast activation and cardiomyocyte hypertrophy occur as a direct effect of the local stress signals caused by the loss of viable tissue in the infarcted area. Subsequently, a decline in contractility of the surviving cardiomyocytes occurs, caused by a change in metabolism and calcium-handling genes.[3]

Genome-wide transcriptome analysis on extracts from diseased tissues has significantly enhanced our understanding of the gene regulatory networks that drive these pathological changes in the heart.[4,5] However, to date, these approaches have been disadvantaged by the fact that the signals are derived from tissue homogenates, which inherently causes the loss of spatial information and dilutes out more localized expression signatures. Recent developments in RNA amplification strategies provide the opportunity to use small amounts of input RNA for genome-wide sequencing. Here we use tomo-seq[6] to obtain a genome-wide gene expression signature with high spatial resolution spanning from the infarcted area to the remote. Tracing transcriptional differences across the infarcted heart enabled us to identify clusters of genes with a comparable gene expression profile. In these individual clusters, we recognized genes with well-known functions in specific aspects of heart remodeling, such as *Col1a2* for fibrosis, *Nppa* for cardiomyocyte hypertrophy, or *Serca2* for contractility. Correlation analyses using the spatial distribution patterns of these marker genes allowed for the identification of novel factors that share a comparable transcriptional regulation pattern across the infarcted tissue. Subsequent functional annotation analysis indicated that these genes could be linked to the known gene function of their reference gene. The strong correlation between the expression levels of the marker genes and the expression of the coregulated genes could be confirmed in human ischemic tissue samples.

Our data show that the high spatial resolution in gene expression signatures obtained by tomo-seq reveals new regulators, genetic pathways, and transcription factors that are active in well-defined regions of the heart and potentially involved in specific aspects of heart disease. Using this technique, we identified SOX9 as a potent regulator of many of the *Col1a2* coregulated genes. In vivo loss of *Sox9* reduced the expression of many extracellular matrix (ECM) genes, which coincided with a blunted cardiac fibrotic response on ischemic injury. These data unveil the currently unknown relevance of SOX9 as a key regulator of cardiac fibrosis and underscores that tomo-seq can be used to increase our mechanistic insights into cardiac remodeling to help guide the identification of novel therapeutic candidates.

METHODS

An expanded Methods section is available in the online-only Data Supplement. All reported gene expression data are available in the online-only Data Supplement.

Ischemia Reperfusion Model

Animal experiments were performed in accordance with the institutional review committee at the Hubrecht Institute. Mice were randomly subjected to either sham or ischemia reperfusion surgery as previously described.[7] Two weeks after surgery, cardiac tissues were collected for further analysis.

Tomo-Seq

Tomo-seq experiments were performed as described elsewhere.[6] In short, 2.5-mm-wide portions of cardiac mouse tissue spanning from the infarct toward the remote region of the left ventricular anterior wall were embedded in tissue-freezing medium, frozen on dry ice, and cryosectioned into 48 slices of 80 μm thickness. We extracted RNA from individual slices and prepared barcoded Illumina sequencing libraries according to the CEL-seq protocol.[8] Paired-end reads obtained by Illumina sequencing were aligned to the transcriptome using Burrows-Wheeler Aligner's Smith-Waterman alignment.[9] The 5′ mate of each pair was used for mapping, discarding all reads that mapped equally well to multiple loci. The 3′ mate was used for barcode information. Reads counts were normalized to the same number of total reads per section. Tomo-seq data analysis was performed in MATLAB (MathWorks) using custom-written code. For data analysis, we used an expression cutoff of >4 reads in >1 section. In differential expression analysis (Figure 1C), we determined the boundary between remote and infarcted zone based on the spatial partitioning detected by pairwise

Figure 1. High-resolution gene expression atlas of the infarcted heart by tomo-seq.
A, Schematic representation of a mouse heart after sham surgery (Sham) and 14 days after ischemia reperfusion (14 dpIR). **B**, Pairwise correlation for all sections across all genes showing clusters of correlated sections 14 dpIR in 1 biological replicate. (*Continued*)

comparison of sections across all genes in 1 biological replicate (Figure 1B). For the infarcted zone, we used sections 1 to 26, and for the remote zone we used sections 29 to 47. The border zone (sections 27–29) was omitted to reduce ambiguity in assignment of sections to zones. We then compared the sections within and outside the infarcted zone and assessed statistical significance with Wilcoxon rank sum test. For this analysis, each section was considered as an independent measurement. Furthermore, filtering was applied for genes that showed a ≥2-fold expression difference between the remote and infarcted zones. For this analysis, the mean expression levels for each gene in the two zones were calculated. Concerning the hierarchical clustering, expression traces of the genes that passed the differential expression filter in Figure 1C were used for analysis. The data were standardized by *Z*-score normalization (along rows of data) so that the mean expression is 0 and the standard deviation is 1 to remove differences in the expression level between genes. Euclidean distance was used as a distance metric. The assignment of genes to clusters I through III (Figure 1D) was determined manually considering the similarity in gene expression pattern across the ischemic heart.

SOX9 Animal Models

Sox9 (*Sox9*$^{fl/fl}$) mutant mice harboring 2 *loxP* sites flanking exons 2–3[10] were crossed with *Rosa26-CreERT2* mice (*R26*^{CreERT2}) to obtain an inducible *Sox9* loss of function model (*Sox9*$^{fl/+}$;*R26*^{CreERT2}). For lineage-tracing studies, mice expressing *CreERT2* under the control of the *Sox9* promoter[11] were bred with the *Rosa26-tdTomato* reporter mouse (*R26R*^{TdT}) to obtain *Sox9*^{CreERT2};*R26R*^{TdT} mice. To induce the CreERT2 protein, *Sox9*$^{fl/+}$;*R26*^{CreERT2} and *Sox9*^{CreERT2};*R26R*^{TdT} mice were injected with Tamoxifen (corn oil/ethanol) intraperitoneally (2 mg at the day of surgery and 2 and 4 days after injury). Control mice (referred to as *Sox9*$^{fl/+}$;*R26*^{CreERT2} vehicle) received an equal volume of the vehicle that was used to deliver Tamoxifen.

Pathway and Transcription Factor Binding Site Enrichment

To investigate whether genes share a similar biological function, we searched for overrepresentation in the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway using DAVID (Database for Annotation, Visualization and Integrated Discovery).[12] The enriched genes in the KEGG pathway are shown as *P* values corrected for multiple hypothesis testing using the Benjamini-Hochberg method.

Table. Top 30 Genes Showing the Most Similar Expression Pattern to *Col1a2*

Col1a2 Similar Genes	SOX9 Predicted Binding Site	SOX9 Validated Binding Site
Col1a2	×	×
Sparc	×	
Fstl1	×	×
Serping1	×	
Pdgfrl	×	
Tmem45a	×	
Col3a1	×	×
Sfrp1	×	×
Lox	×	×
Ecrg4		
Dkk3	×	
Col1a1	×	
Itgbl1	×	
Fn1	×	×
Thbs2		
Cthrc1	×	
Col8a1	×	
Col5a2	×	
Lum	×	×
Fbln2	×	
Gas1	×	×
Antxr1	×	×
Thbs1	×	×
Ogn	×	
Col16a1	×	×
Vim	×	×
Cxcl16	×	
Timp1	×	×
Rnase4	×	×
Ddah1	×	×

Genes that contain a predicted/validated SOX9 binding site in their promoter region are marked.

Detection of overrepresented conserved transcription factor binding sites in the set of genes spatially coregulated to *Col1a2* was determined using single-site analysis in oPOSSUM 3.0 (online software). The enrichment of SOX9 binding sites

Figure 1 Continued. C, Pairwise correlation for all sections across genes exhibiting ≥2-fold and statistically significant differential expression between the infarct and remote zones 14 dpIR. **D**, Hierarchical clustering of expression traces for all genes that were found to be differentially expressed in **C**. **E**, Spatial expression pattern of 3 reference genes *Col1a2*, *Nppa*, and *Serca2* in the hearts from Sham, 1 dpIR, and 14 dpIR mice. **F**, Validation of the expression pattern by in situ hybridization (ISH) 14 dpIR. Four-chamber view (**left**) and higher magnification (**right**) are shown. Scale bars, 1 mm (**left**) and 200 μm (**right**). **G**, Spatial expression traces of 10 coregulated genes 14 dpIR. Reference genes are shown in red, and the 10 most similar genes are shown in grey. Black bold traces show other known markers involved in fibrosis, hypertrophy, and contractility (*Sparc*, *Nppb*, and *Pln*, respectively). **H**, Validation of the coexpression pattern of *Col1a2*/*Sparc*, *Nppa*/*Nppb*, and *Serca2*/*Pln* by ISH. Scale bars, 200 μm. **I** through **K**, KEGG analysis showing the enriched pathways of the top 150 genes in which *Col1a2* (**I**), *Nppa* (**J**), and *Serca2* (**K**) coregulated genes are involved.

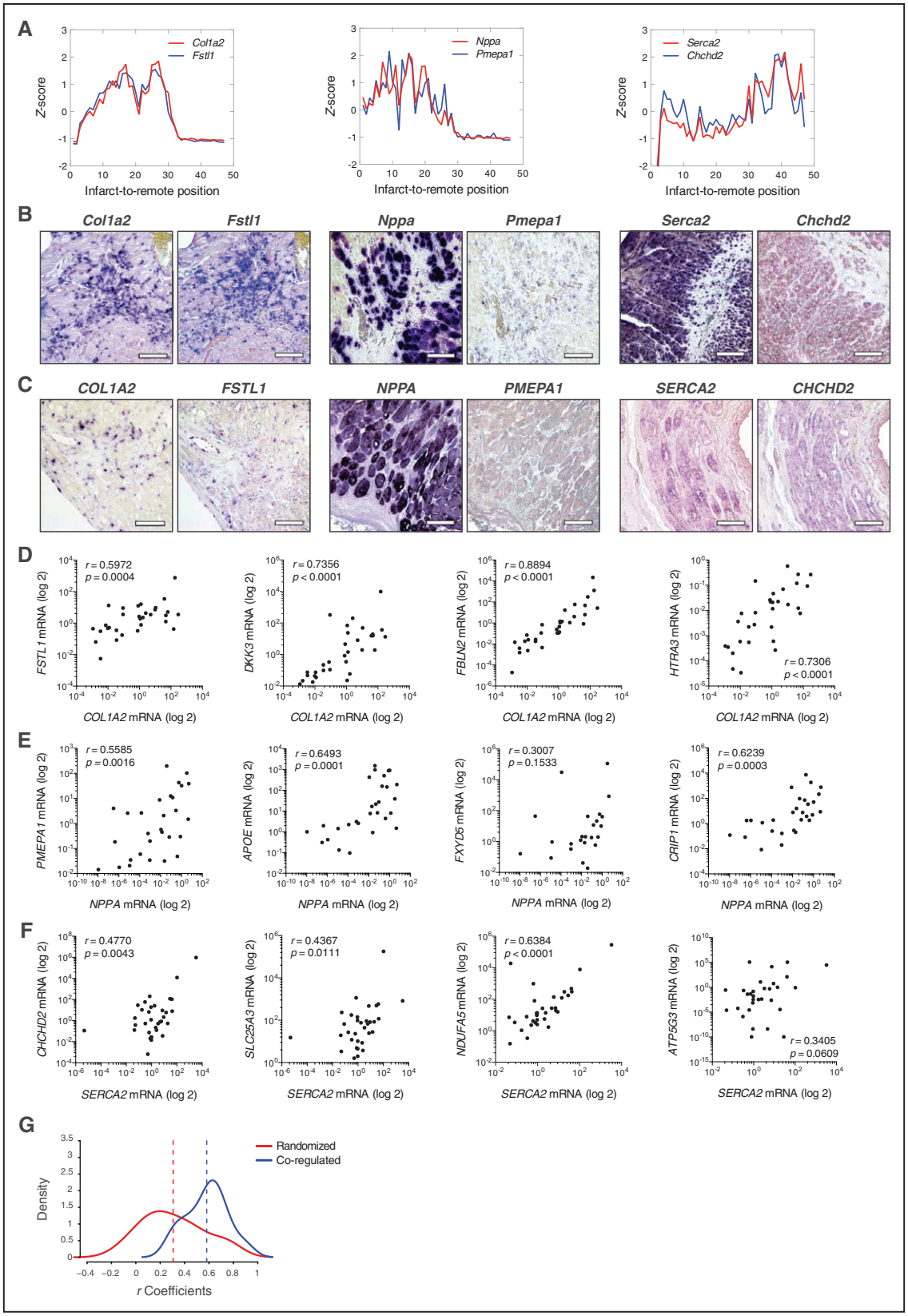


Figure 2. Identification of novel genes involved in remodeling and function of the ischemic heart.
A, Spatial expression traces of 3 selected novel genes coregulated with *Col1a2* (*Fstl1*), *Nppa* (*Pmepa1*), or *Serca2* (*Chchd2*) in mice 14 dpIR in 1 biological replicate. Expression traces were normalized by *Z*-score transformation. **B**, Validation of coexpression of *Col1a2/Fstl1*, *Nppa/Pmepa1*, and *Serca2/Chchd2* by in situ hybridization (ISH) in mice 14 dpIR. Scale bars, 200 μm. **C**, Validation of the coexpression pattern of *COL1A2/FSTL1*, *NPPA/PMEPA1*, and *SERCA2/CHCHD2* by ISH on human ischemic heart tissue. (*Continued*)

was determined using the *Z*-score, which uses the normal approximation to the binomial distribution to compare the occurrence rate of a transcription factor binding site in the set of target genes to the expected rate estimated from the precomputed background set.

Human Heart Samples

Approval for studies on human tissue samples was obtained from the Medical Ethics Committee of the University Medical Center Utrecht, The Netherlands (12#387). Written informed consent was obtained or in certain cases waived by the ethics committee when obtaining informed consent was not possible because of patient death. In this study, we included tissue from the left ventricular free wall of patients with end-stage heart failure secondary to ischemic heart disease. This end-stage heart failure tissue was obtained at explanation of the failing heart during heart transplantation or autopsy. For each case, 3 areas of the infarcted heart tissue were included: (1) infarct zone, (2) border zone, and (3) remote area. For in situ hybridization (ISH) analysis, 3 patients were included. From these patients, the border zone of the infarcted hearts was used for ISH to verify tomo-seq. Gene expression values in infarct zone, border zone, and remote area obtained by real-time polymerase chain reaction (PCR) were plotted for correlation analysis. Left ventricular free wall of nonfailing donor hearts that could not be transplanted for technical reasons were used for comparison. In these cases, neither donor patient histories nor echocardiography revealed signs of heart disease.

Statistical Analysis

Values are presented as mean±standard error of the mean. Previous studies were used to predetermine sample size. Statistical analyses between 2 groups were conducted using the 2-tailed unpaired or paired Student *t* test or a Mann-Whitney test when the normality assumption was not met. Comparison among >2 groups was performed using a 2-way analysis of variance with Bonferroni's post hoc test. Pearson's correlation coefficients were used to calculate gene pair correlation based on gene expression in human samples. KEGG pathways are ranked by their respective *P* value corrected for multiple hypotheses testing using the Benjamini-Hochberg method. *P* value <0.05 was interpreted to denote statistical significance. Prism 6 (GraphPad Software, Inc.) was used for statistical analyses.

RESULTS

Tomo-Seq Performed on the Infarcted Mouse Heart

To obtain precise spatial information on local gene expression changes occurring in the heart in response to ischemic injury, we collected cardiac tissue from infarcted mice exposed to 1 hour of ischemia, followed by either 1 or 14 days of ischemia reperfusion (1 and 14 dpIR) and harvested tissues from sham-operated mice as control (Sham) (Figure 1A and Figure I in the online-only Data Supplement).[7] Histological and molecular analysis confirmed a classical cardiac remodeling response in our model of ischemic injury, as exemplified by cardiac hypertrophy (hematoxylin and eosin staining), fibrosis (Sirius Red staining), and a change in expression of cardiac markers (Figure I in the online-only Data Supplement).[7] Using microdissection, a small portion of the anterior wall of the left ventricle spanning from the infarct toward the remote (2.5 mm wide and 4.0 mm long) was processed into ≈50 consecutive cryosections with a thickness of 80 μm (Figure 1A). Subsequent RNA extraction from individual slices, followed by RNA amplification, barcoding strategies, and RNA sequencing,[6] provided genome-wide data about the spatial distribution in gene expression across the injured heart (Databases I through III in the online-only Data Supplement). A spatial partitioning between infarcted and remote area was visible at 1 and 14 dpIR but not in the sham-operated samples when performing pairwise comparison of sections across all expressed genes (Figure 1B and Figure II in the online-only Data Supplement). The spatial separation became considerably more pronounced after filtering for genes that showed a ≥2-fold and statistically significant differential expression between the infarct and remote zone by tomo-seq (Figure 1C and Figure IIB in the online-only Data Supplement). The number of regulated genes was found to be the highest 14 dpIR, which included 2357 coding genes and 134 noncoding transcripts (Figure 1D and Figure IIC and Databases IV through VI in the online-only Data Supplement. KEGG analysis on these regulated genes showed enrichment for inflammatory pathway activation at day 1 after injury, whereas pathways involved in ECM, disease, and cardiomyocyte remodeling were found to be regulated 14 dpIR (Figure IID and Tables I and II in the online-only Data Supplement).

Gene Expression Patterns Reveal Localized Remodeling Responses

Tracing transcriptional differences across the infarcted heart enabled us to identify clusters of genes with a

Figure 2 Continued. Scale bars, 100 μm. **D–F**, Real-time polymerase chain reaction (PCR) analysis of genes that are spatially coregulated in mice 14 dpIR (see Database VII in the online-only Data Supplement) using human cardiac tissues from patients with ischemic heart disease. Control hearts and remote, border-zone, and infarct zones from ischemic hearts are plotted. Data are presented as log 2-transformed values. Pearson correlation (*r*) and significance of coregulated gene expression are shown (*n*=27–34; *P*<0.05 is considered significant). **G**, Kernel density plot of Pearson *r* values of the correlation in expression among the 4 corresponding *COL1A2*, *NPPA*, and *SERCA2*coregulated genes (coregulated; *n*=12) versus genes that are not coregulated (ie, genes cross-referenced from different lists) (randomized; *n*=24) (see Figure VI in the online-only Data Supplement). Dotted lines depict the mean of the *r* values of all correlated and noncorrelated genes.

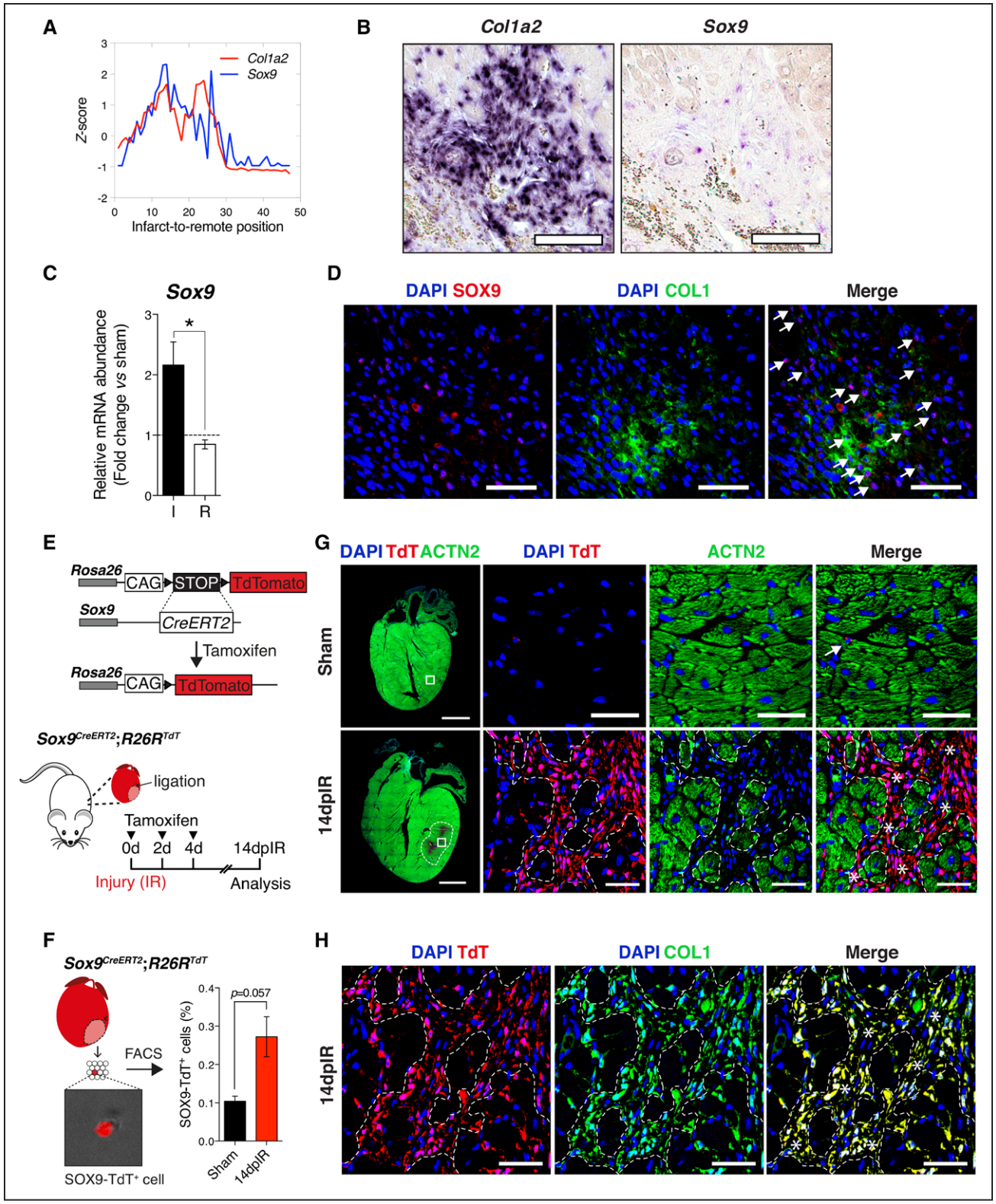


Figure 3. Identification of *Sox9* as a key regulator of fibrosis-related genes.
A, Spatial coexpression of *Col1a2* and *Sox9* determined by tomo-seq in the heart 14 dpIR in 1 biological replicate. **B**, Validation of the coexpression of *Sox9* and *Col1a2* in mice determined by in situ hybridization (ISH) 14 dpIR. Scale bars, 100 μm. **C**, polymerase chain reaction (PCR) analysis of *Sox9* expression in infarct (I) and remote (R) cardiac regions. Data are presented as fold change over Sham-operated control hearts (n=5–6; *P<0.05 versus sham). **D**, Validation of the coexpression of SOX9 and COL1 in the infarct/border zone in mice determined by coimmunostaining 14 dpIR. Nuclei were (*Continued*)

comparable differential regulation throughout the infarcted heart at 14 dpIR (Figure 1D). The individual clusters contained well-known marker genes for specific aspects of heart remodeling, Collagen type I alpha 2 (*Col1a2*) (identified in cluster I), Natriuretic peptide A (*Nppa*) (identified in cluster II), and sarco- and endoplasmic reticulum Ca^{2+}-ATPase (*Serca2*) (located in cluster III). *Col1a2* is expressed in activated fibroblasts and important for cardiac fibrosis,[13] whereas *Nppa* is a cardiomyocyte-specific stress marker involved in myocyte hypertrophy.[14] Cardiomyocyte contractility is regulated by calcium fluxes to and from the sarcoplasmic reticulum and is impaired during heart disease. *Serca2* is a key regulator of Ca^{2+} transfer into the sarcoplasmic reticulum in muscle cells that is decreased during heart failure, which contributes to the decline in function.[3] The expression traces for *Col1a2, Nppa*, and *Serca2* confirmed a gene-specific differential regulation from the infarcted area to the remote (Figure 1E). As expected, *Col1a2* and *Nppa* were more abundantly expressed in the infarcted region 14 dpIR, whereas *Serca2* actually showed a decrease in expression toward the infarcted region (Figure 1E). ISH on cardiac tissue 14 dpIR confirmed the *Col1a2* expression to originate from activated fibroblasts, whereas the transcriptional peaks for *Nppa* stemmed from the stressed, hypertrophic cardiomyocytes immediately flanking the fibrotic regions (Figure 1F). We observed a decline in *Serca2* expression more toward the injured area, which is likely because of both a loss in cardiomyocytes and a decrease in transcriptional activation because the *Nppa* signals clearly indicate the presence of viable myocytes in this region (Figure 1F). The reproducibility of the obtained gene expression profiles was confirmed on a second set of samples (Figure III in the online-only Data Supplement).

Tomo-Seq Identifies Potential New Players for Cardiac Remodeling and Function

An important advantage of tomo-seq over genome-wide sequencing techniques on tissue homogenates is that the local information on gene regulation allows for correlation analysis to identify genes with a comparable spatial distribution in transcriptional regulation.[6] Because we observed a gene-specific expression profile throughout the infarcted tissue for *Col1a2*, *Nppa*, and *Serca2*, we used the Euclidean distance of *Z*-score-transformed spatial expression traces[6] to measure pattern similarity between genes 14 dpIR using *Col1a2*, *Nppa*, and *Serca2* as reference genes. In doing so, we obtained a gene list that showed the greatest similarity in expressional differences across the infarcted tissue with our reference genes (Table, Figure 1G, and Databases VII and VIII in the online-only Data Supplement), a majority of which could be identified within the corresponding gene cluster identified in Figure 1D. It is interesting to note that next to *Col1a2*, *Nppa*, and *Serca2*, these lists also contained other well-known genes related to the biological function of the reference genes. Among the *Col1a2* coexpressed genes, we recognized additional genes known for their function in ECM deposition (eg, *Sparc* and *Col3a1*),[2] whereas many of the genes coregulated with *Nppa* encode for proteins involved in cardiomyocyte hypertrophy (*Nppb* and *Myh7*). The gene list for *Serca*2 contained *Pln* and *Ryr2*, both well known for their importance in cardiac calcium handling and contractility (Table and Databases VII and VIII in the online-only Data Supplement).[3] This coexpression of genes could be confirmed by ISH and indicated the signals to stem from the same cell population (Figure 1H). The connection between the spatially coregulated genes and their biological function was underscored by KEGG pathway analysis. The 150 genes with the highest similarity in expressional changes with either *Col1a2*, *Nppa*, or *Serca2* throughout the infarcted heart at 14 dpIR indicated enrichment for the cellular function known to be associated with the reference genes (Figure 1I through 1K and Tables III through V and Databases VII and VIII in the online-only Data Supplement). The known biological link of several of the listed genes and the functional connection based on gene ontology analysis suggest that the correlation analysis can serve to identify genes that are functionally related to the biological function of the reference genes.

RNA sequencing on whole-tissue homogenates from the infarcted area from 3 independent mice 14 dpIR showed a comparable directional regulation in gene expression, with the *Col1a2*- and *Nppa*-related genes going up after infarct, whereas the *Serca2*-related genes are going down compared with sham-operated mice (Figure IV in the online-only Data Supplement).

Figure 3 Continued. counterstained with 4',6-diamidino-2-phenylindole (DAPI). White arrows point to cells expressing SOX9 in their nuclei (purple). Scale bars, 50 μm. **E**, Schematic representation of the lineage tracing strategy of *Sox9*-expressing cells. Reporter mice conditionally expressing TdTomato driven by the *Sox9* promotor (*Sox9*^{CreERT2};*R26R*^{TdT}) were subjected to Sham surgery or IR, injected with Tamoxifen at days 0, 2, and 4 after surgery, and analyzed after 14 days. **F**, Fluorescence-activated cell sorting (FACS) quantification of cardiac SOX9-TdT^{+} cells in the hearts from Sham and 14 dpIR mice (n=3–4; *P<0.05 versus healthy Sham control hearts). **G** and **H**, Coimmunostaining against TdTomato (TdT) and α-actinin-2 (ACTN2) (**G**) or COL1 (**H**) in the hearts of Sham-operated mice and 14 dpIR. White stars in the Merge field indicate SOX9-TdT-positive regions. Scale bars, 1 mm (4-chamber view) and 50 μm (higher magnification).

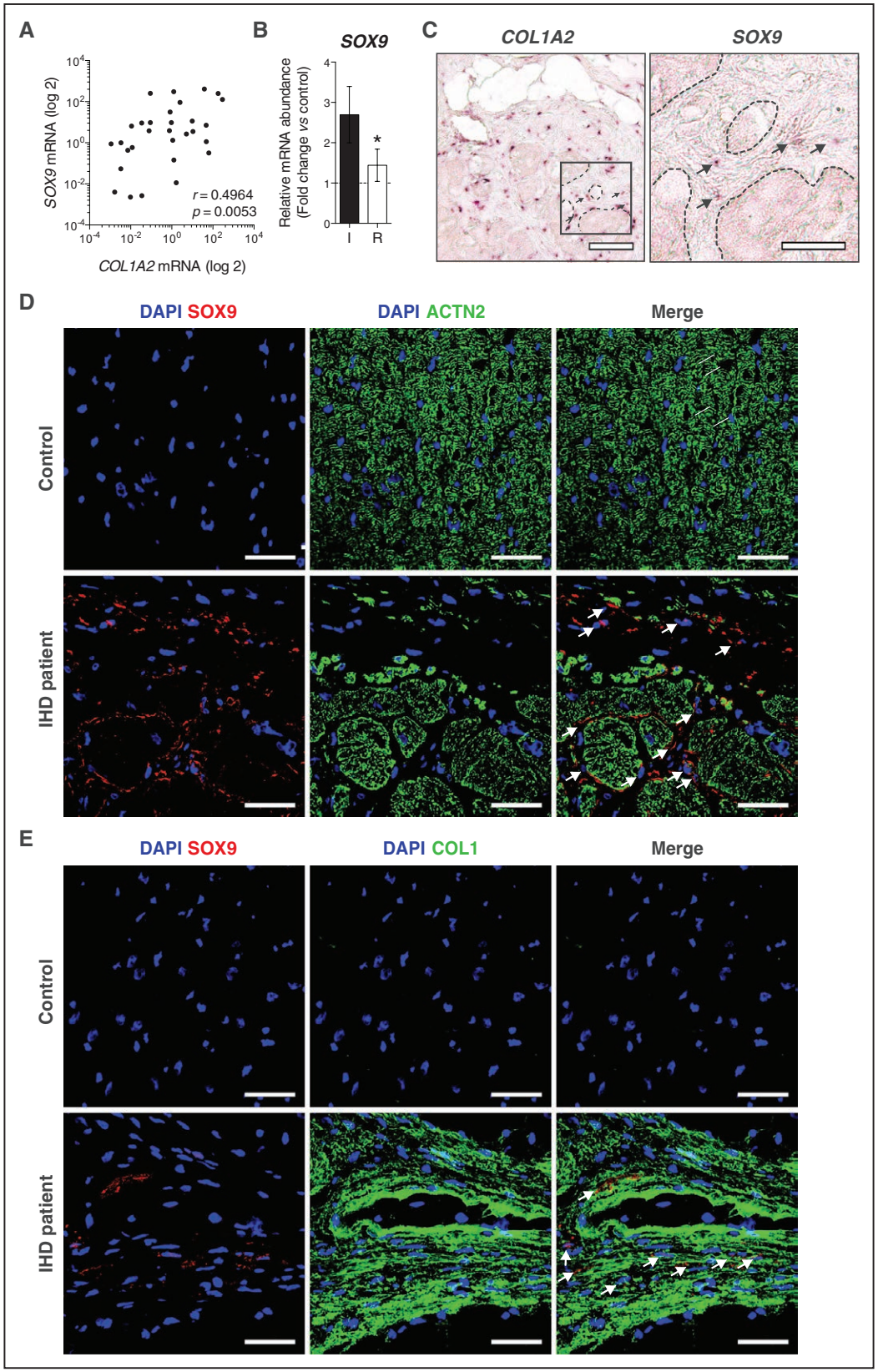


Figure 4. SOX9 is expressed in the fibrotic region in human cardiac tissue.

A, Pearson correlation of *SOX9* and *COL1A2* expression determined by real-time polymerase chain reaction (PCR) analysis on cardiac patient tissue (*n*=30). Data are presented as log 2-transformed values. **B**, Real-time PCR analysis of *SOX9* expression in infarct (I) and remote (R) cardiac regions of ischemic tissue samples. Data are presented as fold change over (*Continued*)

However, in contrast with the data obtained by tomo-seq, the changes observed by RNA sequencing on tissue homogenates failed to provide spatial information on coexpression of genes and showed smaller changes with a high interanimal variation (Figure IV in the online-only Data Supplement). Because tomo-seq analysis is based on the correlation in gene expression within a single sample, the variation among animals is of lesser importance.

Tomo-seq analysis for lncRNAs specifically showed localized expression changes, albeit far less pronounced and specific than for coding genes (Figure V in the online-only Data Supplement), which is likely because of the low abundance of lncRNA transcripts.

The Correlation in Expression of Novel Genes Linked to Cardiac Remodeling and Function Is Conserved in Humans

Although multiple well-known markers of fibrosis, hypertrophy, and calcium handling could be identified among the genes with a similar transcriptional activation pattern, we also found multiple ill-studied genes that so far have not been linked to aspects of cardiac remodeling (Databases VII and VIII in the online-only Data Supplement). To confirm the correlation in transcriptional activation, we randomly chose 1 candidate from each list to explore in more detail. The *Z*-score-transformed expression traces at 14 dpIR indicated a close correlation in expressional regulation among *Col1a2* with *Fstl1*, *Nppa* with *Pmepa1*, and *Serca2* with *Chchd2* (Figure 2A), which could be confirmed by ISH on murine cardiac tissue 14 dpIR (Figure 2B). Further confirmation for a correlation in expression of these novel factors with *COL1A2*, *NPPA*, and *SERCA2* was obtained by ISH on ischemic human heart tissues (Figure 2C).

The validity of using tomo-seq to identify genes that are expressionally linked was strengthened by the observation that real-time PCR analysis on cardiac tissue from patients suffering from ischemic heart disease confirmed the correlation between the expression levels of *COL1A2*, *NPPA*, *SERCA2*, and the newly identified genes (Figure 2D through 2F). The correlation was strongly reduced when we cross-referenced genes from different lists (Figure VI in the online-only Data Supplement). The density plot for the cumulative Pearson correlation coefficients validates the shift toward a higher correlation among genes that belong to the same list (coregulated) compared with genes that were not shown to be coregulated (randomized) by tomo-seq (Figure 2G).

It remains to be determined which of the newly defined genes are relevant for cardiac remodeling. Yet the functional link between the coregulated genes and the fact that we can validate the coregulation in both mice and human implies that tomo-seq allows for the identification of novel genes that are potentially relevant for specific aspects of pathological remodeling of the infarcted heart.

Tomo-Seq Identifies SOX9 as a Key Transcription Factor for Cardiac Fibrosis

The overlap in differential expression throughout the infarcted heart triggered us to explore whether a common transcription factor could be responsible for the synchrony in transcriptional regulation of the different gene clusters. Using an *in silico* approach, we searched for transcription factors (using oPOSSUM 3.0) that contain ≥1 predicted binding site(s) in the promoter regions of the top 30 *Col1a2* coregulated genes (Table and Table VI in the online-only Data Supplement). Among these factors, we identified SOX9 as a potential candidate. SOX9 is a transcription factor recognized for its role in chondrocyte differentiation.[10]

Although so far unstudied in the adult heart, previous work showed that SOX9 has a potent function in fibrosis.[15] Expression trace analysis for *Sox9* revealed a strong spatial correlation with *Col1a2* (Figure 3A). ISH indicated *Sox9* to be expressed in the same region of the infarcted area as *Col1a2* although at a much lower level (Figure 3B). Real-time PCR on tissues from infarcted mouse heart further confirmed *Sox9* upregulation in the infarct zone (Figure 3C). Based on the predicted binding site(s) in the promoter regions of multiple *Col1a2* coregulated genes, its proposed function in liver fibrosis, and the overlap in transcriptional regulation with *Col1a2* in the infarcted heart, we decided to further pursue SOX9 in cardiac fibrosis. The induction in *Sox9* expression was only observed 14 dpIR and restricted to the infarcted area (Figure VIIA and VIIB in the online-only Data Supplement). Staining for both SOX9 and COL1 indicated SOX9 protein to be detectable in the same region as COL1 (Figure 3D).

To start exploring the fate of SOX9-expressing cells in the infarcted heart, we used a lineage-tracing approach using a TdTomato reporter mouse model driven by the promoter of the *Sox9* gene ($Sox9^{CreERT2}$;$R26R^{TdT}$) (Figure 3E). Fluorescence-activated cell sorting performed

Figure 4 Continued. healthy control hearts (n=3–10; *P<0.05). **C**, Validation of the coexpression pattern of *SOX9*/*COL1A2* by in situ hybridization in human ischemic cardiac tissue. Scale bars, 100 μm (**left**) and 50 μm (**right**). **D** and **E**, Coimmunostaining against SOX9 and α-actinin-2 (ACTN2) (**D**) or COL1 (**E**) in the hearts of control individuals or patients suffering from ischemic heart disease (IHD). White arrows point to cells expressing SOX9. Scale bars, 50 μm.

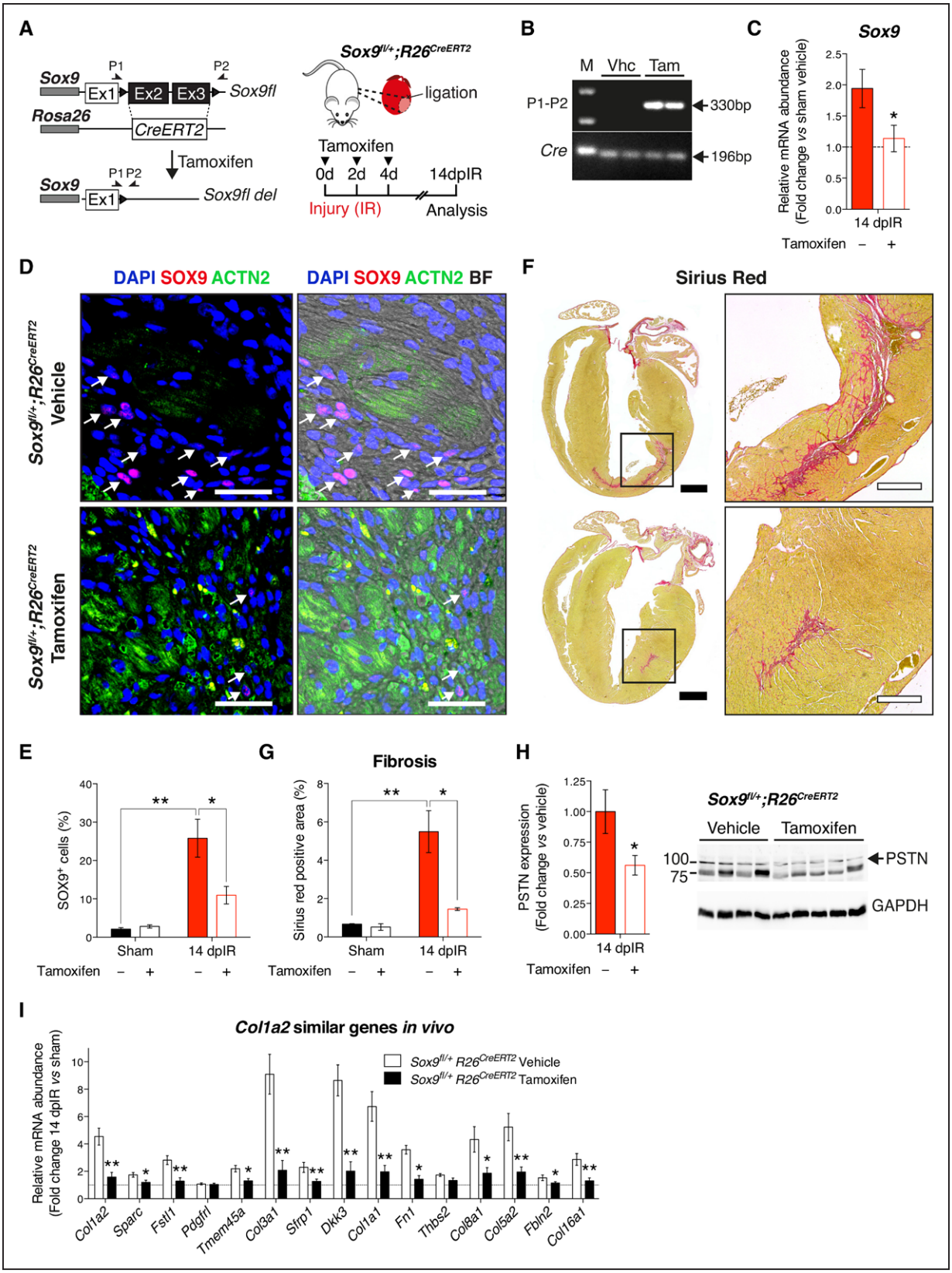


Figure 5. Loss of *Sox9* in mice protects against cardiac fibrosis.
A, Schematic representation of the targeting strategy for conditional *Sox9* deletion. *Sox9*$^{fl/+}$;*R26*^{CreERT2} were subjected to Sham surgery or IR, injected with Tamoxifen at days 0, 2, and 4 after surgery, and analyzed after 14 days. **B**, Polymerase chain reaction (PCR) genotyping for *Sox9* floxed-deleted allele (*Sox9fl del*) and *Cre* transgene. Genomic (*Continued*)

on single cells isolated from the left ventricle indicated a significant elevation of the SOX9-TdT+ cell population 14 dpIR compared with Sham (Figure 3F). Immunostaining clearly showed a colocalization between SOX9-TdT+- and COL1-expressing cells, which were surrounded by cardiomyocytes (marked by α-actinin-2) (Figure 3G and 3H). A similar overlap in expression was observed between SOX9-TdT+ and cells labeled with 2 other fibroblast markers: periostin and vimentin (Figure VIIIA and VIIIB in the online-only Data Supplement, respectively).[16,17] These data demonstrated that SOX9 is predominantly active in the fibroblast population that repopulates the infarcted area after injury.

To further explore whether SOX9 is involved in the transcriptional activation of the *Col1a2*-coregulated genes, we treated fibroblasts with transforming growth factor β1 after we exposed them to either an siRNA against *Sox9* or a control. In addition to a strong repressive effect on *Sox9* expression, we also observed a significant repressive effect on 6 out of 15 potential SOX9 targets listed as *Col1a2*-coregulated genes, with a general downward trend for the remaining genes (Figure VIIC in the online-only Data Supplement). This was also true for additional fibrosis-related genes (Figure VIID in the online-only Data Supplement), indicating a global function for SOX9 in fibroblast activation.

Real-time PCR analysis in tissue samples from human ischemic hearts showed a significant correlation between the levels of expression of *COL1A2* and *SOX9* (Figure 4A). In agreement with our mouse data, real-time PCR indicated an expressional increase in *SOX9* expression toward the infarcted area (Figure 4B), where the majority of the fibrosis is located. ISH showed that a subpopulation of *COL1A2*-positive cells was also positive for *SOX9* in human ischemic hearts (Figure 4C). Immunostaining further confirmed that SOX9-expressing cells were also positive for COL1 (Figure 4D and 4E).

SOX9 Regulates Cardiac Fibrosis During Ischemia Reperfusion Injury

To examine the effect of SOX9 in vivo, we generated inducible *Sox9* heterozygous knockout mice (*Sox9*$^{fl/+}$;*R26*^{CreERT2}) (Figure 5A). Tamoxifen injection at the day of surgery and 2 and 4 days after injury resulted in a disruption of the *Sox9* allele as confirmed by PCR on genomic DNA (Figure 5B) and further quantified by real-time PCR and immunofluorescence after IR in the infarcted region (Figure 5C through 5E). *Sox9* loss of function was accompanied with a profound reduction in fibrosis, which was quantified by the amount of Sirius Red staining in the infarcted region (Figure 5F and 5G). Periostin, a protein marking activated fibroblasts,[17] was also reduced in the infarcted *Sox9*$^{fl/+}$;*R26*^{CreERT2} mice treated with Tamoxifen, further confirming the importance of SOX9 as a key driver for fibrosis in the ischemic heart (Figure 5H). Expression analysis for 15 randomly selected *Col1a2*-coexpressed genes showed an increase in expression in response to ischemic injury. Loss of *Sox9* resulted in a significant reduction in expression for 13 out of 15 genes 14 dpIR compared with control animals (Figure 5I and Database VII in the online-only Data Supplement, highlighted in yellow).

High expression levels of SOX9 have previously been described in chondrocytes, and publically available SOX9 ChIP-seq data in this cell type[18] showed that 15 out of the 30 genes that were coexpressed with *Col1a2* in our study have ≥1 of their predicted binding sites directly occupied by SOX9 (eg, *Col1a2*, *Fn1*, *Lum*, and *Vim*; Table and Figure IX in the online-only Data Supplement). It is important to note that these sites were found enriched for the histone mark H3K27ac in the adult mouse heart (ENCODE dataset), which further demonstrates that these regions are active and open for transcription factors such as SOX9 in vivo. Altogether these data demonstrate that SOX9 has the ability to occupy the promoter region of ECM-related genes and may actively regulate these genes in the heart.

DISCUSSION

Here we applied tomo-seq to obtain a genome-wide gene expression profile with a high spatial resolution throughout the mammalian heart after ischemic injury. Cardiac ischemia reperfusion damage induces a heterogeneous remodeling response that involves

Figure 5 Continued. DNA isolated from *Sox9*$^{fl/+}$;*R26*^{CreERT2} treated with vehicle or Tamoxifen. Used forward (P1) and reverse (P2) primers are indicated as demi arrowheads in **A**. M indicates marker. **C**, Real-time PCR analysis of *Sox9* in the hearts (infarct zone) from *Sox9*$^{fl/+}$;*R26*^{CreERT2} mice 14 dpIR injected with either vehicle or Tamoxifen (n=4–7; *P<0.05 versus the indicated groups). **D**, Coimmunostaining against SOX9 and α-actinin-2 (ACTN2) on corresponding infarcted heart tissue 14 dpIR. **Right,** Same section including bright field (BF). Scale bars, 50 μm. **E**, Quantification of SOX9 expression 14 dpIR in the fibrotic region of the left ventricle or the corresponding region in Sham mice (n=3; **P<0.01; *P<0.05 versus Sham). **F**, Histological sections of infarcted hearts stained for Sirius Red (collagen) 14 dpIR. Scale bars, 1 mm (**left**) and 400 μm (**right**). **G**, Quantification of Sirius Red-positive area in infarcted area or corresponding Sham region (n=3–7; **P<0.01; *P<0.05 versus corresponding Sham). **H**, Western blot analysis of the fibrotic protein periostin (PSTN) in the hearts of *Sox9*$^{fl/+}$;*R26*^{CreERT2} mice 14 dpIR. Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) was used as a loading control (n=4–5; *P<0.05 versus myocardial infarction [MI] vehicle-injected mice). **I**, Real-time PCR analysis of *Col1a2*-coexpressed genes in the mouse hearts 14 dpIR (n=6; **P<0.01; *P<0.05 versus Sham injected with vehicle).

several key processes, such as cardiac fibrosis, cardiomyocyte hypertrophy, and a change in calcium handling within the heart muscle cells.[2,3,13,14] Localized expressional differences of well-known marker genes for these remodeling processes allowed us to uncover novel genes that showed a comparable transcriptional regulation and are linked to specific aspects of cardiac remodeling. Using this dataset, we identified SOX9 as a key transcriptional regulator of ECM-related genes and showed that in vivo loss of *Sox9* after myocardial infarction blunted the cardiac fibrotic response on ischemic injury.

Although RNA sequencing techniques on tissue samples have been instrumental in defining genes relevant for cardiac remodeling and repair,[4,5] so far these approaches have been disadvantaged by the fact that the signals are derived from tissue homogenates, which inherently causes the loss of spatial information and dilutes out more localized expression signatures. In addition, conventional methods for defining localized changes in genes expression, such as ISH or immunohistochemistry, are limited to a defined set of candidate genes and do not allow for genome-wide screening for novel relevant gene candidates. Recent developments in RNA amplification strategies provide the opportunity to use small amounts of input RNA for genome-wide sequencing, as exemplified by tomo-seq.[6] Although recent studies showed this method to provide insightful data for the developing and injured zebrafish heart,[6,19] our study shows for the first time the relevance for the mammalian heart after ischemic injury. Especially the transcriptional differences, introduced by the localized heterogeneity in remodeling throughout an individual infarcted heart, appeared to be valuable for the identification of clusters of genes that showed a comparable regulation in expression. For this study, we focused on genes that showed an equivalent transcriptional regulation pattern across the infarcted tissue as well-known functions in fibrosis, cardiomyocyte hypertrophy, or contractility (*Col1a2*, *Nppa*, or *Serca2*).[2,3,13,14] Based on subsequent functional annotation analysis, expressional confirmation in human ischemic tissue samples, and functional in vitro and in vivo assays, we conclude that the high spatial resolution in gene expression signatures obtained by tomo-seq allows for the identification of new relevant factors for specific aspects of heart disease. While we were preparing our manuscript, it was also reported that *Fstl1*, 1 of our top *Col1a2*-coregulated genes, is important for cardiac fibroblast activation,[20] which further underscores the relevance of our approach for identifying new players in specific cardiac remodeling responses.

Using our tomo-seq data, we identified SOX9 as common transcription factor able to regulate the expression of the majority of the *Col1a2*-coregulated genes. SOX9 is a transcription factor essential for chrondrogenesis by the activation of many ECM genes.[21] In the heart, SOX9 is highly expressed during development, where it promotes epithelial-to-mesenchymal transition and ECM organization during heart valve development.[22,23] In the adult heart, SOX9 has been shown to play a role in valve calcification.[24,25] Although SOX9 has been implicated in the fibrotic response of the liver,[15] so far it was unknown to play a role in cardiac fibrosis. We show that SOX9 is induced in response to ischemic injury and that in vivo loss of SOX9 after myocardial infarction blunts the cardiac fibrotic response on damage, revealing a previously unknown function for SOX9 in cardiac fibrosis. In addition, we show that SOX9 is mainly active in the fibroblast population that repopulates the infarcted area after injury.

In our efficacy studies, we make use of a reduction in SOX9 levels instead of complete deletion, which is sufficient to cause an effect on cardiac fibrosis after injury. An equally profound phenotype in the heart has been reported by others on heterozygous deletion of *Rock1* and *Klf6*, 2 other key regulators of fibrosis.[26,27] This finding suggests that the molecular mechanisms that drive cardiac fibrosis are sensitive to small perturbations in gene expression. Because therapeutically targeting of SOX9 would also moderately lower expression levels, we think this genetic model offers a good representation of what would happen when using an inhibitor of SOX9 in the clinic as a therapy for cardiac fibrosis.

Here we show that the high spatial resolution in gene expression signatures obtained by tomo-seq reveals new regulators, genetic pathways, and transcription factors that are active in well-defined regions of the heart and potentially involved in specific aspects of heart disease. This knowledge increases our mechanistic insights into cardiac remodeling and function and will help guide the identification of novel therapeutic candidates. However, the applicability of this approach is far greater than ischemic heart disease and the remodeling aspects we now focused on, and it can serve to identify new relevant factors for many different biological processes and disease states.

ACKNOWLEDGMENTS

The authors thank Jeroen Korving, Reinier van der Linden, Stefan van der Elst, and Harry Begthel for technical assistance.

SOURCES OF FUNDING

This work was supported by grants from the European Research Council (AdG 294325 GeneNoiseControl and CoG 615708 MICARUS) and a grant from the Leducq Foundation.

DISCLOSURES

None.

AFFILIATIONS

From Hubrecht Institute, KNAW (G.P.A.L., J.P.J., M.M.G., B.M., K.T.S., M.V.-G., D.V., H.d.R., M.W.V., M.P.C., A.v.O., E.v.R.), Department of Pathology (M.M.H.H.), Department of Cardiology (N.d.J., E.v.R.), University Medical Center Utrecht, The Netherlands; and Berlin Institute for Medical Systems Biology, Max Delbrück Center for Molecular Medicine, Berlin, Germany (J.P.J.).

FOOTNOTES

Received February 10, 2017; accepted July 6, 2017.

The online-only Data Supplement is available with this article at http://circ.ahajournals.org/lookup/suppl/doi:10.1161/CIRCULATIONAHA.117.027832/-/DC1.

Circulation is available at http://circ.ahajournals.org.

REFERENCES

1. Tarone G, Balligand JL, Bauersachs J, Clerk A, De Windt L, Heymans S, Hilfiker-Kleiner D, Hirsch E, Iaccarino G, Knöll R, Leite-Moreira AF, Lourenço AP, Mayr M, Thum T, Tocchetti CG. Targeting myocardial remodelling to develop novel therapies for heart failure: a position paper from the Working Group on Myocardial Function of the European Society of Cardiology. *Eur J Heart Fail*. 2014;16:494–508. doi: 10.1002/ejhf.62.
2. van Berlo JH, Maillet M, Molkentin JD. Signaling effectors underlying pathologic growth and remodeling of the heart. *J Clin Invest*. 2013;123:37–45. doi: 10.1172/JCI62839.
3. Lehnart SE, Maier LS, Hasenfuss G. Abnormalities of calcium metabolism and myocardial contractility depression in the failing heart. *Heart Fail Rev*. 2009;14:213–224. doi: 10.1007/s10741-009-9146-x.
4. Herrer I, Roselló-Lletí E, Rivera M, Molina-Navarro MM, Tarazón E, Ortega A, Martínez-Dolz L, Triviño JC, Lago F, González-Juanatey JR, Bertomeu V, Montero JA, Portolés M. RNA-sequencing analysis reveals new alterations in cardiomyocyte cytoskeletal genes in patients with heart failure. *Lab Invest*. 2014;94:645–653. doi: 10.1038/labinvest.2014.54.
5. Liu Y, Morley M, Brandimarto J, Hannenhalli S, Hu Y, Ashley EA, Tang WH, Moravec CS, Margulies KB, Cappola TP, Li M; MAGNet consortium. RNA-Seq identifies novel myocardial gene expression signatures of heart failure. *Genomics*. 2015;105:83–89. doi: 10.1016/j.ygeno.2014.12.002.
6. Junker JP, Noël ES, Guryev V, Peterson KA, Shah G, Huisken J, McMahon AP, Berezikov E, Bakkers J, van Oudenaarden A. Genome-wide RNA Tomography in the zebrafish embryo. *Cell*. 2014;159:662–675. doi: 10.1016/j.cell.2014.09.038.
7. Hullinger TG, Montgomery RL, Seto AG, Dickinson BA, Semus HM, Lynch JM, Dalby CM, Robinson K, Stack C, Latimer PA, Hare JM, Olson EN, van Rooij E. Inhibition of miR-15 protects against cardiac ischemic injury. *Circ Res*. 2012;110:71–81. doi: 10.1161/CIRCRESAHA.111.244442.
8. Hashimshony T, Wagner F, Sher N, Yanai I. CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. *Cell Rep*. 2012;2:666–673. doi: 10.1016/j.celrep.2012.08.003.
9. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2010;26:589–595. doi: 10.1093/bioinformatics/btp698.
10. Akiyama H, Chaboissier MC, Martin JF, Schedl A, de Crombrugghe B. The transcription factor Sox9 has essential roles in successive steps of the chondrocyte differentiation pathway and is required for expression of Sox5 and Sox6. *Genes Dev*. 2002;16:2813–2828. doi: 10.1101/gad.1017802.
11. Kopp JL, Dubois CL, Schaffer AE, Hao E, Shih HP, Seymour PA, Ma J, Sander M. Sox9+ ductal cells are multipotent progenitors throughout development but do not produce new endocrine cells in the normal or injured adult pancreas. *Development*. 2011;138:653–665. doi: 10.1242/dev.056499.
12. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009;4:44–57. doi: 10.1038/nprot.2008.211.
13. Li AH, Liu PP, Villarreal FJ, Garcia RA. Dynamic changes in myocardial matrix and relevance to disease: translational perspectives. *Circ Res*. 2014;114:916–927. doi: 10.1161/CIRCRESAHA.114.302819.
14. Sergeeva IA, Hooijkaas IB, Van Der Made I, Jong WM, Creemers EE, Christoffels VM. A transgenic mouse model for the simultaneous monitoring of ANF and BNP gene activity during heart development and disease. *Cardiovasc Res*. 2014;101:78–86. doi: 10.1093/cvr/cvt228.
15. Hanley KP, Oakley F, Sugden S, Wilson DI, Mann DA, Hanley NA. Ectopic SOX9 mediates extracellular matrix deposition characteristic of organ fibrosis. *J Biol Chem*. 2008;283:14063–14071. doi: 10.1074/jbc.M707390200.
16. Chang HY, Chi JT, Dudoit S, Bondre C, van de Rijn M, Botstein D, Brown PO. Diversity, topographic differentiation, and positional memory in human fibroblasts. *Proc Natl Acad Sci U S A*. 2002;99:12877–12882. doi: 10.1073/pnas.162488599.
17. Snider P, Standley KN, Wang J, Azhar M, Doetschman T, Conway SJ. Origin of cardiac fibroblasts and the role of periostin. *Circ Res*. 2009;105:934–947. doi: 10.1161/CIRCRESAHA.109.201400.
18. Ohba S, He X, Hojo H, McMahon AP. Distinct transcriptional programs underlie Sox9 regulation of the mammalian chondrocyte. *Cell Rep*. 2015;12:229–243. doi: 10.1016/j.celrep.2015.06.013.
19. Wu CC, Kruse F, Vasudevarao MD, Junker JP, Zebrowski DC, Fischer K, Noël ES, Grün D, Berezikov E, Engel FB, van Oudenaarden A, Weidinger G, Bakkers J. Spatially resolved genome-wide transcriptional profiling identifies BMP signaling as essential regulator of zebrafish cardiomyocyte regeneration. *Dev Cell*. 2016;36:36–49. doi: 10.1016/j.devcel.2015.12.010.
20. Maruyama S, Nakamura K, Papanicolaou KN, Sano S, Shimizu I, Asaumi Y, van den Hoff MJ, Ouchi N, Recchia FA, Walsh K. Follistatin-like 1 promotes cardiac fibroblast activation and protects the heart from rupture. *EMBO Mol Med*. 2016;8:949–966. doi: 10.15252/emmm.201506151.
21. Ikeda T, Kamekura S, Mabuchi A, Kou I, Seki S, Takato T, Nakamura K, Kawaguchi H, Ikegawa S, Chung UI. The combination of SOX5, SOX6, and SOX9 (the SOX trio) provides signals sufficient for induction of permanent cartilage. *Arthritis Rheum*. 2004;50:3561–3573. doi: 10.1002/art.20611.
22. Akiyama H, Chaboissier MC, Behringer RR, Rowitch DH, Schedl A, Epstein JA, de Crombrugghe B. Essential role of Sox9 in the pathway that controls formation of cardiac valves and septa. *Proc Natl Acad Sci U S A*. 2004;101:6502–6507. doi: 10.1073/pnas.0401711101.
23. Lincoln J, Kist R, Scherer G, Yutzey KE. Sox9 is required for precursor cell expansion and extracellular matrix organization during mouse heart valve development. *Dev Biol*. 2007;305:120–132. doi: 10.1016/j.ydbio.2007.02.002.
24. Peacock JD, Levay AK, Gillaspie DB, Tao G, Lincoln J. Reduced sox9 function promotes heart valve calcification phenotypes in vivo. *Circ Res*. 2010;106:712–719. doi: 10.1161/CIRCRESAHA.109.213702.
25. Huk DJ, Austin BF, Horne TE, Hinton RB, Ray WC, Heistad DD, Lincoln J. Valve endothelial cell-derived Tgfβ1 signaling promotes nuclear localization of Sox9 in interstitial cells associated with attenuated calcification. *Arterioscler Thromb Vasc Biol*. 2016;36:328–338. doi: 10.1161/ATVBAHA.115.306091.
26. Rikitake Y, Oyama N, Wang CY, Noma K, Satoh M, Kim HH, Liao JK. Decreased perivascular fibrosis but not cardiac hypertrophy in ROCK1+/– haploinsufficient mice. *Circulation*. 2005;112:2959–2965. doi: 10.1161/CIRCULATIONAHA.105.584623.
27. Sawaki D, Hou L, Tomida S, Sun J, Zhan H, Aizawa K, Son BK, Kariya T, Takimoto E, Otsu K, Conway SJ, Manabe I, Komuro I, Friedman SL, Nagai R, Suzuki T. Modulation of cardiac fibrosis by Krüppel-like factor 6 through transcriptional control of thrombospondin 4 in cardiomyocytes. *Cardiovasc Res*. 2015;107:420–430. doi: 10.1093/cvr/cvv155.

Circulation

Tomo-Seq Identifies SOX9 as a Key Regulator of Cardiac Fibrosis During Ischemic Injury

Grégory P.A. Lacraz, Jan Philipp Junker, Monika M. Gladka, Bas Molenaar, Koen T. Scholman, Marta Vigil-Garcia, Danielle Versteeg, Hesther de Ruiter, Marit W. Vermunt, Menno P. Creyghton, Manon M.H. Huibers, Nicolaas de Jonge, Alexander van Oudenaarden and Eva van Rooij

Circulation. 2017;136:1396-1409; originally published online July 19, 2017;
doi: 10.1161/CIRCULATIONAHA.117.027832
Circulation is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231

Print ISSN: 0009-7322. Online ISSN: 1524-4539

SUPPLEMENTAL MATERIAL

Tomo-seq identifies SOX9 as a key regulator of cardiac fibrosis during ischemic injury

Grégory P.A. Lacraz, MSc, PhD; Jan Philipp Junker, MSc, PhD; Monika M. Gladka, MSc, PhD; Bas Molenaar, MSc; Koen T. Scholman, MSc; Marta Vigil-Garcia, MSc; Danielle Versteeg, BS; Hesther de Ruiter, BS; Marit W. Vermunt, MSc, PhD; Menno P. Creyghton, MSc, PhD; Manon M.H. Huibers, MSc, PhD; Nicolaas de Jonge, MD; Alexander van Oudenaarden, MSc, PhD and Eva van Rooij, MSc, PhD

Supplemental Methods

Ischemia-Reperfusion (IR) injury

Mice were weighed and injected IP with Ketamine/xylazine after which they were ventilated with O_2 containing 1.5–2.5% isoflurane using a micro-ventilator (UMV-03, UNO BV) for anesthesia. The body temperature was maintained at 37°C with a heating pad. After removal of the hair the skin was disinfected with iodine and 10% ethanol and incised on the left of the midline and the heart was accessed via the third intercostal space. A branch of the left anterior descending (LAD) coronary artery was identified and after placing a PE tubing the LAD was ligated with a 7-0 silk suture. Ligation was confirmed by distal cyanosis. Following 1 hr of ischemia, the PE tubing was removed and allowed for the removal of the ligature and subsequent reperfusion. After closing the rib cage with a 5-0 silk suture and the skin with a wound clip, the animals were disconnected from the ventilator and placed unrestrained on a nose cone with 100% oxygen in a warm recovery cage until fully ambulatory, at which point the oxygen was turned off. To alleviate pain or distress, buprenorphine was injected subcutaneously as analgesic (0.05–0.1 mg/kg for mice), given once at completion of surgery. Sham animals underwent an identical operation without induction of the infarct. At the indicated time points after ischemia reperfusion or sham surgery, the animals were killed by an overdose of isoflurane followed by cervical dislocation and tissues were snap-frozen or collected in 4% paraformaldehyde for subsequent analysis.

Tomo-seq

Tomo-seq experiments were performed as described elsewhere. [1] In short, 2.5 mm wide portions of cardiac mouse tissue spanning from the infarct towards the remote region of the left ventricular anterior wall were embedded in tissue freezing medium, frozen on dry ice, and

cryosectioned into 48 slices of 80 μm thickness. We extracted RNA from individual slices and prepared barcoded Illumina sequencing libraries according to the CEL-seq protocol.[2] Paired-end reads obtained by Illumina sequencing were aligned to the transcriptome using BWA.[3] The 5' mate of each pair was used for mapping, discarding all reads that mapped equally well to multiple loci. The 3' mate was used for barcode information. Read counts were normalized to the same number of total reads per section. Tomo-seq data analysis was performed in MATLAB (MathWorks) using custom-written code. For data analysis we used an expression cut-off of >4 reads in >1 section. In differential expression analysis (Figure 1C), we determined the boundary between remote and infarcted zone based on the spatial partitioning detected by pairwise comparison of sections across all genes in one biological replicate (Figure 1B). For the infarcted zone, we used sections 1-26, and for the remote zone we used sections 29-47. The border zone (sections 27-29) was omitted in order to reduce ambiguity in assignment of sections to zones. We then compared the sections within and outside the infarcted zone and assessed statistical significance with Wilcoxon rank sum test. For this analysis, each section was considered as an independent measurement. Furthermore, filtering was applied for genes that showed at least a two-fold expression difference between remote and infarcted zone. For this analysis, the mean expression levels for each gene in the two zones was calculated. Concerning the hierarchical clustering, expression traces of the genes that passed the differential expression filter in Figure 1C were used for analysis. The data was standardized by *Z*-score normalization (along rows of data) so that the mean expression is zero and the standard deviation is 1 in order to remove differences in expression level between genes. Euclidean distance was used as distance metric. The assignment of genes to clusters I-III (Figure 1D) was determined manually considering the similarity in gene expression pattern across the ischemic heart.

***In situ* hybridization**

ISH probe constructs were generated by PCR amplification of partial coding sequence from cDNA of mouse and human heart tissue, and subsequent ligation into either pSPT18 or pSPT19 vectors (Roche). Primers used for partial coding sequence amplification can be found in Supplemental Table 7 online. Probe plasmids were linearized and riboprobes were transcribed from linearized template in the presence of digoxigenin-11-UTP. Mouse hearts were fixed in formaldehyde overnight, and dehydrated stepwise into ethanol. Dehydrated hearts were incubated in xylene for 2 hrs and overnight in paraffin at 55°C. Hearts were embedded in paraffin blocks, which were solidified at room temperature. Hearts were sectioned transversally using a microtome (Leica) and the sections (10 μm thick) were adhered to Starfrost slides. Paraffin was dissolved in xylene and slides were rehydrated stepwise in ethanol. The slides were prepared for hybridization by the following incubations: 0.2 M HCl for 15 min; 30 μg/ml Proteinase K for 15/20 min for mouse/human slides at 37°C. Proteinase K was stopped by rinsing the slides in 0.2% glycine. The slides were post-fixed for 10 min in 4% PFA, then in acetic anhydride 5 min, and washed in SSC buffer. Slides were prehybridized in hybridization buffer for 1 hr at 65°C. Probes were diluted in 1 μg/ml hybridization solution and denatured by heating at 95°C followed cooling on ice. Hybridization buffer and probes were placed on the slides and incubated for 24 hrs at 65°C, after which the hybridization buffer was removed and slides were incubated at 65°C in 50% formamide/2× SSC buffer three times for 15 min. Slides were washed in TBS-T followed by maleic acid buffer. Slides were blocked for at least 30 min at room temperature in blocking solution and incubated with 1:1500 anti-DIG-AP in blocking solution for 3–4 hrs at room temperature. Slides were washed in maleid acid, TBS-T and NTM buffers. Probe detection was carried out by incubation in NBT/BCIP substrate for 4–24 hrs in the dark. As soon as a signal could be detected the reaction was stopped by washing

the slides in NTM buffer and demi-water subsequently, after which they were mounted with aquatex.

Pathway and transcription factor binding site enrichment

To investigate whether genes share a similar biological function, we searched for over-representation in the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway using DAVID.[4] The enriched genes in the KEGG pathway are shown as *p* values corrected for multiple hypothesis testing using the Benjamini-Hochberg method.

Detection of over-represented conserved transcription factor binding sites in the set of genes spatially co-regulated to *Col1a2* was determined using single site analysis in oPOSSUM 3.0 (online software). The enrichment of SOX9 binding sites was determined using the *Z*-score, which uses the normal approximation to the binomial distribution to compare the rate of occurrence of a TFBS in the target set of genes to the expected rate estimated from the pre-computed background set.

Human heart samples

Approval for studies on human tissue samples was obtained from the Medical Ethics Committee of the University Medical Center Utrecht, The Netherlands (12#387). Written informed consent was obtained or in certain cases waived by the ethics committee when obtaining informed consent was not possible due to death of the patient. In this study, we included tissue from the left ventricular free wall of patients with end-stage heart failure secondary to ischemic heart disease. This end-stage heart failure tissue was obtained at explanation of the failing heart during heart transplantation or at autopsy. For each case three areas of the infarcted heart tissue were included; 1) infarct zone, 2) border zone and 3) remote area. For ISH analysis, three patients were included. From these patients the border zone of the

infarcted hearts was used for ISH to verify tomo-seq and real-time PCR analysis. Left ventricular free wall of non-failing donor hearts, that could not be transplanted for technical reasons, were used for comparison. In these cases, neither donor patient histories, nor echocardiography revealed signs of heart disease.

mRNA analysis

Total RNA from the infarct (containing the infarct and the peri-infarct zone) and the remote zone (classified as myocardial tissue unaffected by the IR), or their equivalent tissue fraction from sham operated hearts, or cultured cells was isolated using Trizol reagent (Invitrogen) according to the manufacturer's instructions and quantified by optical density at 260 nm using a Nanodrop 1000 spectrophotometer (Thermoscientific). For quantitative detection of mRNA levels, complementary DNAs were synthesized by using iScript cDNA Synthesis Kit (Bio-Rad). Real-time PCR analysis was carried out in a CFX96 Connect detection system (Bio-Rad) using gene specific primers (provided in Supplemental Table 8 online) and iQ SYBR Green Supermix (Bio-Rad). Reverse transcribed RNA (10 ng) was used as template for each reaction. All reactions were run in duplicate with no template control. The PCR conditions were: 95ºC for 15 min, followed by 40 cycles at 95ºC for 15 s, 60ºC for 30 s and 72ºC for 30 s. Glyceraldehyde 3-phosphate dehydrogenase (*Gapdh*) was used as house-keeping gene.

Histology and immunofluorescence

Heart tissues from C57BL/6 and mutant mice or human patients were embedded either in paraffin after fixation in 4% formalin for two consecutive days, or in OCT after fixation in PLP buffer containing 0.1 M phosphate buffer containing 0.2 M L-lysine, pH 7.4, 2.12 mg/ml NaIO4, and 4% paraformaldehyde overnight, washed in phosphate buffer, and dehydrated in 30% sucrose in phosphate buffer overnight. Hematoxylin and eosin (H&E) and Sirius Red

stains were performed using routine histology protocols. Images were acquired using a Leica DM4000 M (Leica Microsystems) microscope. The Sirius Red-positive area was determined as a percentage of the infarcted area using ImageJ. $SOX9^+$ cells were quantitated in three different fields (magnification × 20) of the fibrotic area 14dpIR (~1300 cells/heart) or the corresponding region in Sham (~850 cells/heart) (n=3 mouse per group). For immunofluorescence staining, paraffin sections were dewaxed in xylene for 5 min, rehydrated in ethanol (2× 5 min in 100% ethanol, 2× 5 min in 96% ethanol, 2× 5 min in 70% ethanol), and rinsed three times with demi water. This was followed by antigen retrieval boiling with citrate buffer for 20 min. Sections were then washed in PBS, incubated for 45 min at room temperature in a blocking solution containing 1% BSA, 0.01% Triton X-100, and 0.01% Tween20, and then incubated overnight at 4℃ with a primary antibody specific for SOX9 (rabbit, 1:250, Millipore #ab5535), type 1 collagen (mouse, 1:500, Abcam #ab6308), alpha-actinin (ACTN2, mouse, 1:400, Sigma #A7732), and TdTomato (goat, 1:1000, Sicgen #AB8181-200). Sections were incubated for 60–90 min at room temperature with species specific secondary antibodies conjugated with Alexa 488 (1:400, Life Technologies #A11001) or Alexa Fluor 568 (1:300, Life Technologies #A10042 and #A11057). Nuclear counterstaining was performed using DAPI prior to mounting (Prolong Gold antifade reagent, Life Technologies #P36934). Images were acquired using a SPE (Leica Microsystems) microscope.

3T3-L1 cell culture

3T3-L1 cells were purchased from ATCC and cultured in Dulbecco's Minimum GlutaMAX (low glucose) (Gibco) supplemented with 10% FBS (Life technologies). Cells were serum-starved for 16 hrs after which they were transfected for 24 hrs with 10 nM *Sox9* siRNA (Thermoscientific) or scrambled siRNA as control, using Lipofectamin 2000 reagent (Invitrogen) in Opti-MEM media (Gibco). TGFβ1 (5 ng/ml; Preprotech) was then added and

the cells were harvested at the indicated time points for RNA extraction and subsequent real-time PCR analysis.

SOX9 animals models

Sox9 (*Sox9*$^{fl/fl}$) mutant mice harboring two *loxP* sites flanking the exons 2–3 [6] were crossed with *Rosa26-CreERT2* mice (*R26*^{CreERT2}) to obtain an inducible *Sox9* loss-of-functional model (*Sox9*$^{fl/+}$;*R26*^{CreERT2}). For lineage tracing studies, mice expressing *CreERT2* under the control of the *Sox9* promoter [7] were bred with the *Rosa26-tdTomato* reporter mouse (*R26R*^{TdT}) to obtain *Sox9*^{CreERT2};*R26R*^{TdT} mice. To induce the CreERT2 protein, *Sox9*$^{fl/+}$;*R26*^{CreERT2} and *Sox9*^{CreERT2};*R26R*^{TdT} mice were injected with Tamoxifen (corn oil/ethanol) intraperitoneally (2 mg per day per mouse) at the day of surgery and 2 and 4 days after injury. Control mice (referred to as *Sox9*$^{fl/+}$;*R26*^{CreERT2} Vehicle) received an equal volume of the vehicle that was used to deliver Tamoxifen. Primers used for genotyping are provided in Supplemental Table 8 online.

FACS analysis

Hearts (left ventricle) were digested by a medium containing 20 µg/ml DNAse I (Worthington Biochemical #LK003172) and 26 U/ml Liberase™ (Roche #5401020001) for 15 min at 37°C. The dissociated single cells were collected by filtration through a 100-µm nylon mesh. Expression of TdTomato was analysed by flow cytometry using a FACSAria III flow cytometer (BD Biosciences) and the FlowJo software (Tree Star).

Western Blotting

Standard Western blot analysis was performed on homogenates from infarcted hearts of *Sox9*$^{fl/+}$;*R26*^{CreERT2} mice using Periostin antibody (Pierce #PA5-34641) and GAPDH (Merck

#mab374) as a loading control. Densitometry of the Western blots was performed using ImageQuant TL Software (GE Healthcare Life Sciences software).

Statistical analysis

Values are presented as mean ± s.e.m. Previous studies were used to predetermine sample size. Statistical analyses between two groups were conducted using the two-tailed unpaired or paired Student's *t*-test or a Mann-Whitney test when the normality assumption was not met. Comparison among groups was performed using a two-way ANOVA with Bonferroni's post-hoc test. Pearson's correlation coefficients were used to calculate gene pair correlation based on gene expression in human samples. KEGG pathways are ranked by their respective p value corrected for multiple hypothesis testing using the Benjamini-Hochberg method. p value <0.05 was interpreted to denote statistical significance. Prism 6 (GraphPad Software, Inc.) was used for statistical analyses.

Supplemental Tables

Supplemental Table 1. Functional annotation chart of predicted pathways activated 1 dpIR using the list of differentially regulated genes determined by tomo-seq.

KEGG term	Genes	Fold Enrich ment	P-value
Biosynthesis of unsaturated fatty acids	*Acot1, Acot2, Acot9, Hadha, Hsd17b12, Elovl6, Scd2*	3.6	1.7E-1
Fc gamma R-mediated phagocytosis	*Fcgr1, Fcgr2b, Limk1, Rac2, Was, Cfl1, Dock2, Gsn, Hck, Inpp5d, Map2k1, Ncf1, Pak1, Pik3r1, Pla2g6, Arpc5, Ptprc, Asap1, Vasp, Vav1*	2.8	1.7E-3
Leukocyte transendothelial migration	*Rac2, Rock1, Actb, Actn4, Actn1, Cyba, Itga4, Itgal, Itgam, Itgb2, Mmp9, Mapk11, Myl9, Ncf1, Ncf2, Ncf4, Pxn, Pik3r1, Actg1, Thy1, Vcam1, Vasp, Vav1*	2.7	1.6E-3
Hematopoietic cell lineage	*Cd24a, Cd33, Cd44, Cd9, Fcgr1, Csf2ra, Gp1bb, H2-Eb1, Itga4, Itgam, Il1b, Il1r1, Il1r2, Il11, Il4ra, Csf3r*	2.7	2.2E-2
Regulation of actin cytoskeleton	*Fgd3, Iqgap1, Limk1, Nckap1l, Rac2, Rock1, Was, Actb, Actn4, Actn1, Baiap2, Cfl1, Cyfip1, Cyfip2, Diap1, Enah, Fn1, Gsn, Itga4, Itgal, Itgam, Itgax, Itgb2, Map2k1, Myh14, Myh9, Myl9, Pak1, Pak2, Pxn, Pik3r1, Pdgfra, Pdgfa, Arpc5, Actg1, Rras2, Apc, Mylk3, Tmsb4x, Vav1*	2.6	7.9E-6
Adherens junction	*Crebbp, Iqgap1, Rac2, Was, Actb, Actn4, Actn1, Baiap2, Insr, Actg-ps1, Ptpn1, Ptpn6, Snai1, Tgfbr1*	2.6	5.3E-2
Chemokine signaling pathway	*Fgr, Rac2, Rock1, Was, Arrb2, Xcr1, Ccl2, Ccl21a, Ccl3, Ccl4, Ccl6, Ccl7, Ccl9, Ccr1, Ccr8, Cxcl16, Cxcr6, Dock2, Elmo1, Hck, Cxcr2, Map2k1, Ncf1, Nfkbib, Pak1, Pxn, Pik3r1, Plcb1, Ccl27a, Ppbp, Cxcl5, Vav1*	2.5	3.3E-4
Pyrimidine metabolism	*Prim1, Entpd5, Pola1, Pole3, Pold4, Polr1c, Polr3a, Tk1, Rrm1, Txnrd1, Tyms, Umps, Upp1, Uck1, Uck2*	2.2	1.4E-1
Cytokine-cytokine receptor interaction	*Relt, Xcr1, Ccl2, Ccl21a, Ccl3, Ccl4, Ccl6, Ccl7, Ccl9, Ccr1, Ccr8, Cxcl16, Ccr6, Csf2ra, Csf2rb, Csf2rb2, Crlf2, Il1b, Il1r1, Il1r2, Il10ra, Il11, Il15, Il17ra, Il4ra, Il8ra, Pdgfra, Pdgfa, Csf3r, Il2rg, Ccl27a, Ppbp, Cxcl5, Tgfb2, Tgfbr1, Tnfrsf12a*	2.1	2.0E-3
Focal adhesion	*Elk1, Rac2, Rock1, Vwf, Actb, Actn4, Actn1, Diap1, Fn1, Flnc, Itga4, Map2k1, Myl9, Pak1, Pak2, Parva, Pxn, Pik3r1, Pgf, Pdgfra, Pdgfa, Actg-ps1, Mylk3, Tnc, Thbs1, Vasp, Vav1, Zyx*	2.0	1.9E-2

1 **Supplemental Table 2. Functional annotation chart of predicted pathways activated 14**
2 **dpIR using the list of differentially regulated genes determined by tomo-seq.**
3

KEGG term	Genes	Fold Enrich ment	P-value
Valine, leucine and isoleucine degradation	*Hmgcs2, Hibch, Oxct1, Auh, Acat1, Acaa2, Acadm, Acadsb, Aldh6a1, Dbt, Dld, Ehhadh, Hadh, Ivd, Mccc1, Mccc2, Mcee, Mut, Hibadh, Hadhb, Pcaa*	3.5	1.0E-5
ECM-receptor interaction	*Cd36, Agrn, Comp, Chad, Col1a1, Col1a2, Col3a1, Col4a2, Col5a1, Col5a2, Col6a1, Col6a2, Col11a1, Fn1, Itga11, Itga2b, Itga6, Itgav, Itgb5, Itgb8, Lamc1, Sdc2, Sdc3, Sdc4, Tnc, Tnxb, Thsb1, Thsb2, Thsb3, Thsb4, Vtn*	3.1	7.9E-8
Hypertrophic cardiomyopathy (HCM)	*Serca2, Actc1, Actb, Ace, Cacna1c, Cacna2d1, Cacnb2, Des, Igf1, Itga11, Itga2b, Itga6, Itgav, Itgb5, Itgb6, Itgb8, Lmna, Myh7, Prkaa2, Prkag1, Ryr2, Sgcg, Ttn, Tgfb2, Tgfb3, Tnni3, Tnnt2*	2.5	2.8E-4
Lysosome	*Abca2, Atp6v0b, Atp6v0d1, Atp6ap1, Cd63, Cd68, Gm2a, Asah1, Ap1s1, Ap4s1, Naglu, Arsa, Ctsa, Ctsd, Ctsf, Ctsk, Ctss, Ctsz, Clta, Fuca1, Glb1, Gns, Gba, Gusb, Hexa, Hexb, Lgmn, Lipa, Lamp1, Laptm5, Man2b1, Manba, Neu1, Pla2g15, Ap3s1, Slc11a1, Smpd1*	2.4	6.9E-6
Oxidative phosphorylation	*Atp5b, Atp5j, Atp5f1, Atp5g3, Atp6v0b, Atp6v0d1, Atp6ap1, Ndufa1, Ndufa2, Ndufa4, Ndufa4l2, Ndufa5, Ndufa6, Ndufb3, Ndufb4, Ndufb5, Ndufb6, Ndufab1, Ndufc2, Ndufs1, Ndufs4, Ndufs5, Cox7b, Cox7a2, Atp5l, Atp5g2, Atp5h, Atp5g1, Uqcrh, Uqcrb, Atp5k, Ndufc1, Gm5457, Cox6c, Ndufs3, Sdha, Sdhb, Sdhd, Uqcrc2*	2.3	2.0E-5
Focal adhesion	*Bcl2, Flt4, Rac2, Src, Actb, Actn4, Actn1, Bcar1, Figf, Comp, Chad, Col1a1, Col1a2, Col3a1, Col4a2, Col5a1, Col5a2, Col6a1, Col6a2, Col11a1, Egf, Egfr, Fn1, Flnc, Flna, Flnb, Igf1, Itga11, Itga2b, Itga6, Itgav, Itgb6, Itgb8, Lama4, Lamb2, Lamc1, Mapk3, Myl9, Parva, Pik3cd, Pik3r2, Pip5k1c, Pdgfra, Pdgfrb, Myl12b, Prkcb, Ppp1ca, Tln1, Tnc, Tnxb, Thsb1, Thsb2, Thsb3, Thsb4, Vasp, Vav1, Vtn, Zyx*	2.3	1.4E-7
Dilated cardiomyopathy	*Actc1, Actb, Adcy4, Adcy7, Cacna1c, Cacna2d1, Cacnb2, des, Igf1, Itga11, Itga2b, Itga6, Itgav, Itgb5, Itgb6, Itgb8, Lmna, Pln, Ryr2, Sgcg, Tnt, Tgfb2, Tgfb3, Tpm2, Tnni3, Tnnt2*	2.2	1.3E-3
Alzheimer's disease	*Atp5b, Atp5j, Atp5f1, Atp5g3, Bad, Ndufa1, Ndufa2, Ndufa4, Ndufa4l2, Ndufa5, Ndufa6, Ndufb3, Ndufb4, Ndufb5, Ndufb6, Ndufab1, Ndufc2, Ndufs1, Ndufs4, Ndufs5, App, Apoe, Bace2, Cacna1c, Capn1, Casp9, Cdk5, Cox7b, Cox7a2, Eif2ak3, Gnaq, Itpr1, Lpl, Mapk3, Cycs, Atp5g2, Atp5h, Atp5g1, Uqcrh, Uqcrb, Ndufc1, Gapdh, Gm5457, Cox6c, Ndufs3, Sdha, Sdhb, Sdhd, Serca2, Tnfrsf1a, Uqcrc2*	2.2	2.6E-6
Parkinson's disease	*Atp5b, Atp5j, Atp5f1, Atp5g1, Atp5g2, Atp5g3, Atp5h, Ndufa1, Ndufa2, Ndufa4, Ndufa4l2, Ndufa5, Ndufa6, Ndufb3, Ndufb4, Ndufb5, Ndufb6, Ndufab1, Ndufc2, Ndufs1, Ndufs4, Ndufs5, Casp9, Cox7b, Cox7a2, Cycs, Uqcrb, Uqcrh, Ndufc1, Gm5457, Cox6c, Ndufs3, Slc25a4, Sdha, Sdhb, Sdhd, Uqcrc2, Uchl1, Ube2l6*	2.2	3.2E-5
Huntington's disease	*Atp5b, Atp5j, Atp5f1, Atp5g3, Ndufa1, Ndufa2, Ndufa4, Ndufa4l2, Ndufa5, Ndufa6, Ndufb3, Ndufb4, Ndufb5, Ndufb6, Ndufab1, Ndufc2, Ndufs1, Ndufs4, Ndufs5, Ap2b1, Ap2s1, Casp9, Clta, Cox7b, Cox7a2, Dctn1, Dctn4, Gnaq, Ap2a2,*	2.1	6.5E-6

	Itpr1, Ppargc1a, Polr2f, Cycs, Atp5g2, Atp5h, Atp5g1, Uqcrh, Uqcrb, Ndufc1, Gm5457, Cox6c, Gm7511, Ap2m1, Ndufs3, Hdac1, Slc25a4, Sdha, Sdhb, Sdhd, Tfam, Uqcrc2		
Regulation of actin cytoskeleton	*Cd14, Rras, Iqgap1, Nckap1l, Rac2, Arhgef6, Arhgef1, Arhgef7, Tiam1, Was, Arpc1b, Arpc4, Actb, Actn4, Actn1, Baiap2, Bcar1, Csk, Chrm2, Egf, Egfr, Fgf1, Fgf13, Fgfr1, Fn1, Itga11, Itga2b, Itga6, Itgam, Itgav, Itgb2, Itgb5, Itgb6, Itgb8, Mapk3, Mras, Myh9, Myl9, Pik3cd, Pik3r2, Pip5k1b, Pip5k1c, Pdgfra, Pdgfrb, Myl12b, Arpc2, Gm5637, Ppp1ca, Nras, Slc9a1, Vav1*	1.8	5.5E-4

1

Supplemental Table 3. Functional annotation chart of predicted pathways activated 14 dpIR using the one hundred fifty most similar genes to the cardiac stress gene *Col1a2* determined by tomo-seq.

KEGG term	Genes	Fold Enrichment	P-value
ECM-receptor interaction	*Col1a1, Col1a2, Col3a1, Col5a1, Col5a2, Fn1, Itgb1, Thbs1, Thbs4*	16.9	2.0E-7
TGF-beta signaling pathway	*Dcn, Id2, Id3, Thbs1, Thbs4,Tgfb3*	11.3	5.0E-4
Focal adhesion	*Actn1, Col1a1, Col1a2, Col3a1, Col5a1, Col5a2, Fn1, Igf1, Itgb1, Thbs1, Thbs2, Thbs4*	8.5	1.7E-6
Antigen processing and presentation	*Ctsl, Ctss, H2-Ab1, H2-Eb1, Lgmn*	7.7	5.0E-2

Supplemental Table 4. Functional annotation chart of predicted pathways activated 14 dpIR using the one hundred fifty most similar genes to the cardiac stress gene *Nppa* determined by tomo-seq.

KEGG term	Genes	Fold Enrichment	P-value
Prion diseases	*C1qa, C1qb, Lamc1*	10.9	2.6E-1
Ribosome	*Rpl11, Rps21, Rps14, Rpsa, RplP0, Fau, Rpl37*	10.0	2.7E-3
Hypertrophic cardiomyopathy	*Actb, Itgb5, Lmna, Myh7*	6.1	2.9E-1
ECM-receptor interaction	*Comp, Itgb5, Lamb2, Lamc1*	6.1	3.6E-1
Dilated cardiomyopathy	*Actb, Itgb5, Lmna, Myh7*	5.5	2.5E-1
Focal adhesion	*Actb, Comp, Flna, Itgb5, Lamb2, Lamc1, Rhoa*	4.5	9.2E-2
Tight junction	*Rras, Actb, Myh7, Rhoa*	3.8	4.8E-1

Supplemental Table 5. Functional annotation chart of predicted pathways activated 14 dpIR using the one hundred fifty most similar genes to the cardiac stress gene *Serca2* determined by tomo-seq.

KEGG term	Genes	Fold Enrichment	P-value
Fatty acid elongation in mitochondria	*Acaa2, Hadh, Hadhb*	32.1	1.6E-2
Valine, leucine and isoleucine degradation	*Oxct1, Auh, Acat1, Acaa2, Acadm, Dbt, Dld, Hadh, Ivd, Hibadh, Hadhb*	20.5	6.0E-10
Oxidative phosphorylation	*Atp5b, Atp5j, Atp5f1, Atp5g3, Ndufa2, Ndufa5, Ndufa6, Ndufb3, Ndufb5, Ndufb6, Ndufc2, Ndufs1, Ndufs4, Cox7b, Cox7a2, Atp5h, Uqcrh, Uqcrb, Atp5k, Ndufc1, Cox6c, Ndufs3, Sdha, Sdhb, Sdhd, Uqcrc2*	17.1	1.7E-23
Parkinson's disease	*Atp5b, Atp5j, Atp5f1, Atp5g3, Ndufa2, Ndufa5, Ndufa6, Ndufb3, Ndufb5, Ndufb6, Ndufc2, Ndufs1, Ndufs4, Cox7b, Cox7a2, Atp5h, Uqcrh, Uqcrb, Ndufc1, Cox6c, Ndufs3, Slc25a4, Sdha, Sdhb, Sdhd, Uqcrc2*	16.7	1.5E-23
Alzheimer's disease	*Atp5b, Atp5j, Atp5f1, Atp5g3, Ndufa2, Ndufa5, Ndufa6, Ndufb5, Ndufb6, Ndufc2, Ndufs1, Ndufs4, Cox7b, Cox7a2, Atp5h, Uqcrh, Uqcrb, Ndufc1, Cox6c, Ndufs3, Slc25a4, Sdha, Sdhb, Sdhd, Uqcrc2*	12.2	6.8E-23
Huntington's disease	*Atp5b, Atp5j, Atp5f1, Atp5g3, Ndufa2, Ndufa5, Ndufa6, Ndufb5, Ndufb6, Ndufc2, Ndufs1, Ndufs4, Cox7b, Cox7a2, Atp5h, Uqcrh, Uqcrb, Ndufc1, Cox6c, Ndufs3, Slc25a4, Sdha, Sdhb, Sdhd, Uqcrc2*	12.2	3.1E-20
Fatty acid metabolism	*Acat1, Acaa2, Acsl1, Acadl, Acadm, Hadh, Hadhb*	13.3	6.7E-5
Citrate cycle (TCA cycle)	*Cs, Dlat, Dld, Fh1, Idh2, Mdh1, Pdha1, Sdha, Sdhb, Sdhd, Sucla2*	11.0	1.1E-8
Cardiac muscle contraction	*Serca2, Actc1, Cacna1c, Cox7b, Cox7a2, Uqcrh, Uqcrb, Cox6c, Ryr2, Uqcrc2*	11.0	1.4E-6
PPAR signaling pathway	*Cd36, Acsl1, Acadl, Acadm, Fabp3, Fabp4, Lpl*	7.6	1.5E-3

Supplemental Table 6. Conserved transcription factor binding sites (TFBS) for the set of 30 genes showing the greatest similarity in expressional regulation to *Col1a2* 14 dpIR in mice.

Transcription factor	Class	Family	Target gene hits	*Z*-score
NFATC2	Ig-fold	Rel	28	15.631
RUNX1	Ig-fold	Runt	28	11.582
FOXO3	Winged Helix-Turn-Helix	Forkhead	27	6.786
Nkx2-5	Helix-Turn-Helix	Homeo	28	6.461
YY1	Zinc-coordinating	BetaBetaAlpha-zinc finger	28	5.285
KLF4	Zinc-coordinating	BetaBetaAlpha-zinc finger	25	5.200
SOX17	Other Alpha-Helix	High Mobility Group	28	4.838
SOX9	**Other Alpha-Helix**	**High Mobility Group**	**27**	**4.227**
Arnt::Ahr	Zipper-Type	Helix-Loop-Helix	26	3.659
PRXX2	Helix-Turn-Helix	Homeo	27	3.022
SPI1	Winged Helix-Turn-Helix	Ets	29	2.876
EBF1	Zipper-Type	Helix-Loop-Helix	25	2.842
TBP	Beta-sheet	TATA-binding	26	1.089
ELK1	Winged Helix-Turn-Helix	Ets	25	-2.576
ZEB1	Zinc-coordinating	BetaBetaAlpha-zinc finger	29	-3.273

Results are sorted by *Z*-score for the detected transcription factors regulating ≥25 *Col1a2* co-regulated genes.

1 **Supplemental Table 7. List of primer sequences used to create *in situ* hybridization**
2 **probes.**
3

Gene	Organism	Sequence
Chchd2	Mus Musculus	Fw SalI_GTCGACCTTATTGGTTGCCCGGTGAG Rv BamH1_GGATCCCAAGGAGGCCAAACACTTCC
Col1a2	Mus Musculus	Fw XbaI_GCTCTAGACCTTGACATTGCACCTCTGG Rv SacI_GCGAGCTCATGAAGAGAGAACCTGGGGC
Fstl1	Mus Musculus	Fw XbaI_GCTCTAGAGACGCCCTCATTGAACTGTC Rv SacI_GCGAGCTCCTCCATGGCGCTTGAAGTAC
Nppa	Mus Musculus	Fw XbaI_GCTCTAGAAGCAAACATCAGATCGTGCC Rv SacI_GCGAGCTCATATGCAGAGTGGGAGAGGC
Nppb	Mus Musculus	Fw XbaI_GCTCTAGAAGGAAATGGCCCAGAGACAG Rv SacI_GCGAGCTCCAAAAGCAGGAAATACGCTATGT
Pln	Mus Musculus	Fw XbaI_GCTCTAGAAGTGCAATACCTCACTCGCT Rv SacI_GCGAGCTCCATGTACACAGCCCTTGAGC
Pmepa1	Mus Musculus	Fw SalI_GCGTCGACCTCTTCCCCTTTCCATCTCC Rv BamH1_GCGGATCCCTCACCAAGCTAGGCACCTC
Serca2	Mus Musculus	Fw XbaI_GCTCTAGAGGTTCGAGGAAGGGGAAGAA Rv SacI_GCGAGCTCTGTGGTAAGTGTGCCTGTCT
Sox9	Mus Musculus	Fw XbaI_GCTCTAGAACATGGAGGACGATTGGGAGA Rv SacI_GCGAGCTCAGGGTGATAGTCTGAGCAGGC
Sparc	Mus Musculus	Fw XbaI_GCTCTAGACATAAGCTCACCGTCCACAAG Rv SacI_GCGAGCTCGCAAGAACCTGAAAGCCCAA
CHCHD2	Homo sapiens	Fw XbaI_GCTCTAGAGCTTAGCTCTTCGGTGGTTG Rv SacI_ GCGAGCTCTTTGCCCAGTCCCAGATTCT
COL1A2	Homo sapiens	Fw XbaI_GCTCTAGACAAGGTTTCCAAGGACCTGC Rv SacI_GCGAGCTCGAAGACCACGAGAACCAGGA
FSTL1	Homo sapiens	Fw XbaI_GCTCTAGACCCCACTCACTACCCTGTTT Rv SacI_GCGAGCTCACATTCAGACTGGTCCACGT
NPPA	Homo sapiens	Fw XbaI_GCTCTAGAGACAGACGTAGGCCAAGAGA Rv_KpnI_GCGGTACCTAATGCATGGGGTGGGAGAG
PMEPA1	Homo sapiens	Fw Sal1_GCGTCGACCTCTTCCCCTTTCCATCTCC Rv Bamh1_GCGGATCCCTCACCAAGCTAGGCACCTC
SERCA2	Homo sapiens	Fw XbaI_GCTCTAGAATGGTGGTTCATTGCTGCTG Rv SacI_GCGAGCTCTGGAGAGGGATCTGGCTACT
SOX9	Homo sapiens	Fw XbaI_GCTCTAGAAGGAAGTCGGTGAAGAACGG Rv SacI_GCGAGCTCCCGTAGCTGCCCGTGTAG
SPARC	Homo sapiens	Fw XbaI_GCTCTAGAAGTACATCGCCCTGGATGAG Rv SacI_GCGAGCTCGTCCGTGCTCCCAAAAGTTT

4

1 **Supplemental Table 8. List of primer sequences used for real-time and genotyping PCR.**

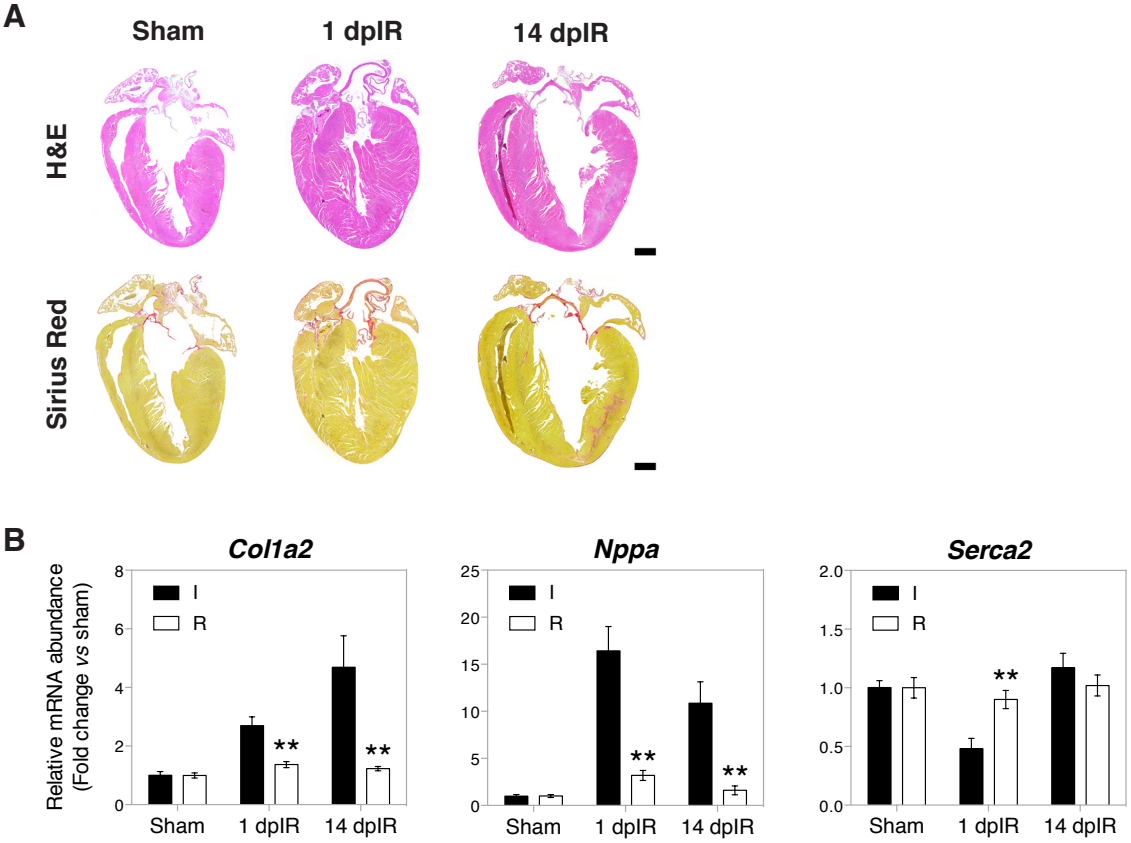
Gene	Organism	Accession number		Sequence
Acta2	Mus Musculus	NM_007392.2	Forward Reverse	ACCCACCCAGAGTGGAGAAG AGCATCATCACCAGCGAAG
Cdrap	Mus Musculus	NM_019394.3	Forward Reverse	GGCCAAGTGGTGTATGTCTTC GCTGCCAGGTCTCCATAGTAAC
Col1a1	Mus Musculus	NM_007742.3	Forward Reverse	AATGCAATGAAGAACTGGACTG CCCTCGACTCCTACATCTTCTG
Col1a2	Mus Musculus	NM_007743.2	Forward Reverse	CAAGGACCTGCTGGTGAAC TGGTCCAACGACTCCTCTC
Col2a1	Mus Musculus	NM_031163.3	Forward Reverse	GCAAAGATGGCTCTAATGGAA GACTTCCAGGGGGACCAA
Col3a1	Mus Musculus	NM_009930.2	Forward Reverse	GATGGCAAAGATGGATCACCTGG GACCCTTTTCTCCTGGGATGC
Col4a1	Mus Musculus	NM_009931.2	Forward Reverse	GAGAGACAGGACCCTTTGGAC GGTCATCTGTTGTCTGACTATGC
Col5a2	Mus Musculus	NM_007737.2	Forward Reverse	GAAAGGCTGGTGATCAAGGT TTTCTCCCCGAGGTCCTAAT
Col8a1	Mus Musculus	NM_007739.2	Forward Reverse	GCCAGCCAAGCCTAAATGT CAGAGTTCAGGGAAATGATGAA
Col11a1	Mus Musculus	NM_007702.2	Forward Reverse	CAGATTGTGACTTAACATCCAAGG CTCGATTATATCCTCAGGTGCAT
Col11a2	Mus Musculus	NM_009926.1	Forward Reverse	CAGCCTAGCAGATGGCAAAT CAGTCCACAATGAGAGTGACAGA
Col16a1	Mus Musculus	NM_028266.5	Forward Reverse	GCATTGCAGGAGAAAATGGT CCATCTTGCCATAACCTGGA
Ctgf	Mus Musculus	NM_010217.2	Forward Reverse	TGACCTGGAGGAAAACATTAAGA AGCCCTGTATGTCTTCACACTG
Dkk3	Mus Musculus	NM_015814.2	Forward Reverse	TGTGTTGTGCCTTCCAAAGA GTTCCCAGGTGATGAGATCC
Ecm1	Mus Musculus	NM_007899.2	Forward Reverse	AAGTGGAAGGGTCCTTAGCAA TCGATGAAGGCCAGTTTCTC
Fbln1	Mus Musculus	NM_010180.2	Forward Reverse	TGCATCAATACAGTGGGCTCT CCAGTCTCACATTCGTCAATATCT
Fbln2	Mus Musculus	NM_007992.2	Forward Reverse	CAGGTGGCCTCTAACACCAT TTGCAGGGTCCATTGTCTTT
Fn1	Mus Musculus	NM_010233.1	Forward Reverse	CCACTGTGGAGTACGTGGTTAG AAGCAATTTTGATGGAATCGAC
Fstl1	Mus Musculus	NM_008047.5	Forward Reverse	CAGCCATCAACATCACCACT ATGAGGGCGTCAACACAGA
Gapdh	Mus Musculus	NM_008084.2	Forward Reverse	TGTCGTGGAGTCTACTGGTG ACACCCATCACAAACATGG
Htra3	Mus Musculus	NM_030127.2	Forward Reverse	TTGCCACGATTGTAATCCAC GGTGTTCTGCAGGGCAAA

Igf1	Mus Musculus	NM_001111276.1	Forward Reverse	CACCTTCACCAGCTCCACCAC CAGTCTCCTCAGATCACAGCTC
Nppa	Mus Musculus	NM_008725.2	Forward Reverse	GGTAGGATTGACAGGATTGGAG GCTTAGGATCTTTTGCGATCTG
Pdgfrl	Mus Musculus	NM_026840.3	Forward Reverse	TGCAGAGACCTCAAAAGGTG CCTGATCTTCCTCCCAGAAA
Serca2	Mus Musculus	NM_009722.3	Forward Reverse	TGGTGATATAGTGGAAATTGCTG GAGTTGTAGACTTGATGGATGTCAA
Sfrp1	Mus Musculus	NM_013834.3	Forward Reverse	GTGGTTCAAGATGTGCTCCA TCAGAGCAGCCAACATGC
Sox9	Mus Musculus	NM_011448.4	Forward Reverse	TATCTTCAAGGCGCTGCAA TCGGTTTTGGGAGTGGTG
Sparc	Mus Musculus	NM_022316.2	Forward Reverse	CCATTGGCGAGTTTGAGAA GTGCACTTGGTGGCAAAGA
Spp1	Mus Musculus	NM_001204201.1	Forward Reverse	CCCGGTGAAAGTGACTGATT TTCTTCAGAGGACACAGCATTC
Tcf4	Mus Musculus	NM_013685.2	Forward Reverse	CATATTTGTGGCCATTGAAGG CAGCTCTTTGTCCGTCCCTA
Tgfb3	Mus Musculus	NM_009368.3	Forward Reverse	CCCTGGACACCAATTACTGC TCAATATAAAGGGGGCGTACA
Thbs2	Mus Musculus	NM_011581.3	Forward Reverse	TCGGACCTCAAGTATGAGTGC TCTAAGAAGGGGTGTTTGCAG
Tmem45a	Mus Musculus	NM_019631.3	Forward Reverse	AACATTATTACGTCGGACAGAGATT AGATAAACTGTTCACCAGCTATACCA
APOE	Homo sapiens	NM_000041.2	Forward Reverse	GGTCGCTTTTGGGATTACCT CATGGTCTCGTCCATCAGC
ATP5G3	Homo sapiens	NM_001689.4	Forward Reverse	GGTGATATTGTAGAAATTGCTGTTG TGACTGGTCAACTCTTAGTGTGGT
COL1A2	Homo sapiens	NM_000089.3	Forward Reverse	TGATGGAAAAGGAGTTGGACTT CAGGTCCTTGGAAACCTTGA
CHCHD2	Homo sapiens	NM_016139.2	Forward Reverse	GCTGAAACAGTGCCGACTT AACTTAGTTATGAGAGCTGATTTTCCA
CRIP1	Homo sapiens	NM_001311.4	Forward Reverse	GCAACAAGGAGGTGTACTTCG CACATTTCTCGCACTTCAGG
DKK3	Homo sapiens	NM_015881.5	Forward Reverse	GAGGACACGCAGCACAAAT TTTGCCAGGTTCACTTCTGA
FBLN2	Homo sapiens	NM_001004019.1	Forward Reverse	CAGGTGGCCTCTAACACCAT TTGCAGGGTCCATTGTCTTT
FSTL1	Homo sapiens	NM_007085.4	Forward Reverse	GCCATCAATATTACAACGTATCCA TCAATGAGAGCATCAACACAGA
FXYD5	Homo sapiens	NM_144779.2	Forward Reverse	TCAGCAGACTCAACTATCATGGA GGGCTGGAGTTCTGTGTAGACT
GAPDH	Homo sapiens	NM_002046.3	Forward Reverse	GGGTCATCATCTCTGCCCC GGTCATGAGTCCTTCCACGA
HTRA3	Homo sapiens	NM_053044.3	Forward Reverse	AGCTACAGAATGGGGACTCCT AGCAACAACACAGGGAGCTT

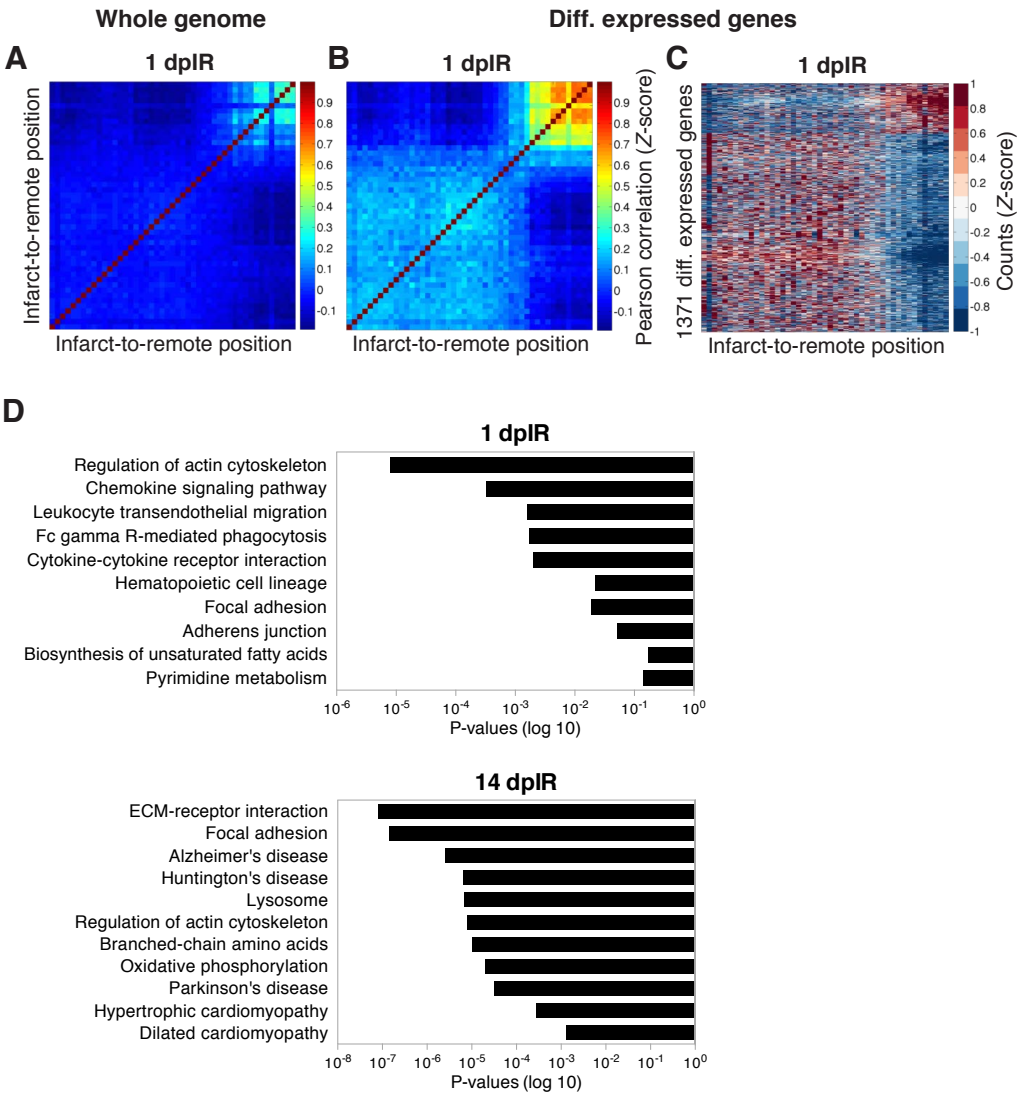
NDUFA5	Homo sapiens	NM_005000.2	Forward Reverse	GGTGTGCTGAAGAAGACCACT TTGTGTACAATATTCTTAGCCTCTCG
NPPA	Homo sapiens	NM_006172.3	Forward Reverse	CCGTGAGCTTCCTCCTTTTA CCAAATGGTCCAGCAAATTC
PMEPA1	Homo sapiens	NM_020182.3	Forward Reverse	CTGCACGGTCCTTCATCAG TTGCCTGACACTGTGCTCTC
SERCA2	Homo sapiens	NM_001681.3	Forward Reverse	GGTGATATTGTAGAAATTGCTGTTG TGACTGGTCAACTCTTAGTGTGGT
SLC25A3	Homo sapiens	NM_213611.2	Forward Reverse	TCTTGTATAGCAATATGCTTGGAGA GGGCAATGTCAGCAAAGAAT
SOX9	Homo sapiens	NM_000346.3	Forward Reverse	GTACCCGCACTTGCACAAC TCTCGCTCTCGTTCAGAAGTC
Sox9fl	Transgene	Genotyping primers	Forward Reverse	GTCATATTCACGCCCCCATT AGACTCTGGGCAAGCTCTGG
Sox9fl del	Transgene	Genotyping primers	Forward Reverse	TGGTAATGAGTCATACACAGTAC GTCAAGCGACCCATGAACGC
Sox9creERT2	Transgene	Genotyping primers	Forward Reverse	GCGGTCTGGCAGTAAAAACTATC GTGAAACAGCATTGCTGTCACTT
Cre	Transgene	Genotyping primers	Forward Reverse	GAAGCAACTCATCGATTGATTTACG CACTATCCAGGTTACGGATATAGTTC
TdTomato	Transgene	Genotyping primers	Forward Reverse	CTGTTCCTGTACGGCATGG GGCATTAAAGCAGCGTATCC

1

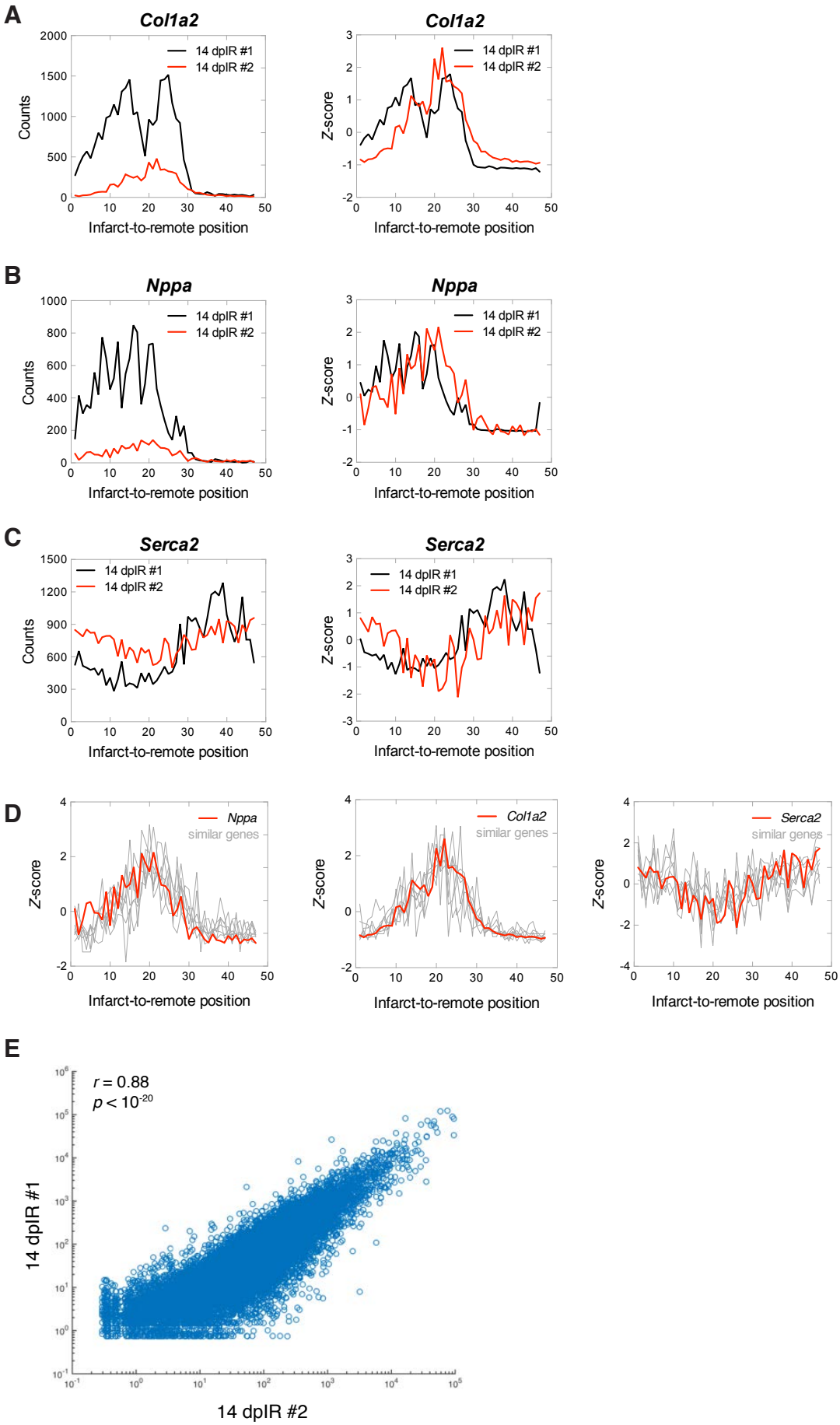
Supplemental Figure 1



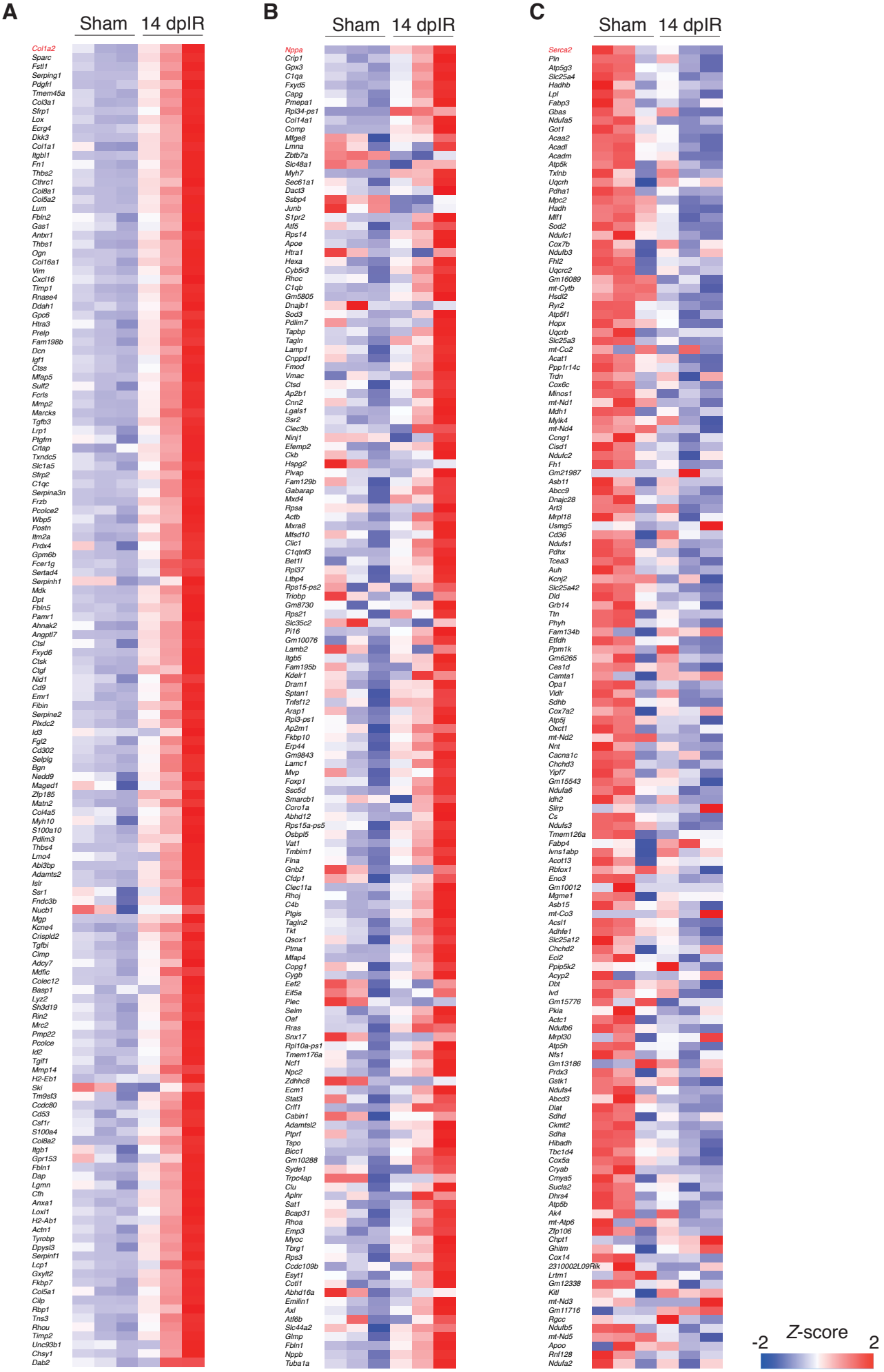
Supplemental Figure 2



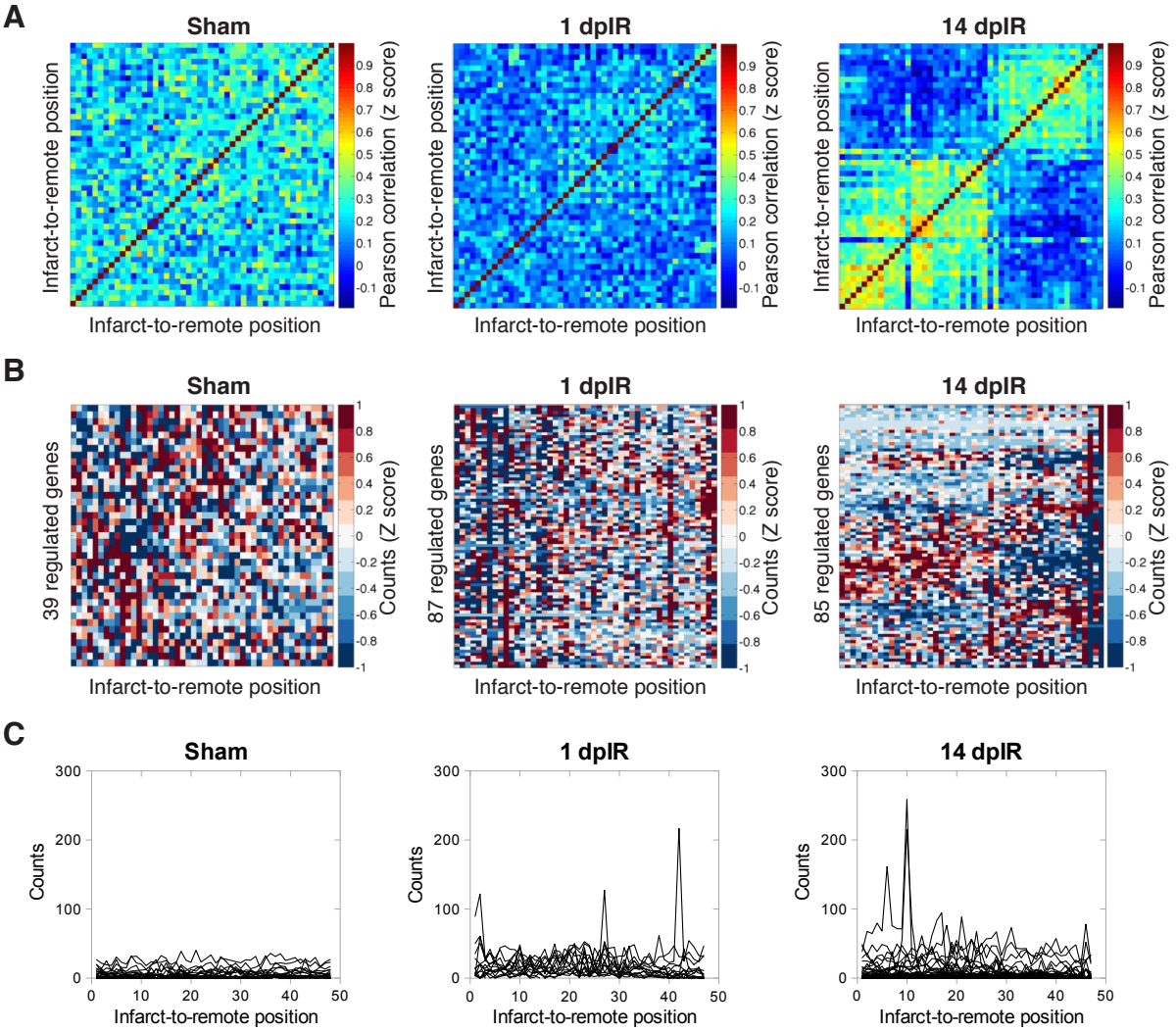
Supplemental Figure 3



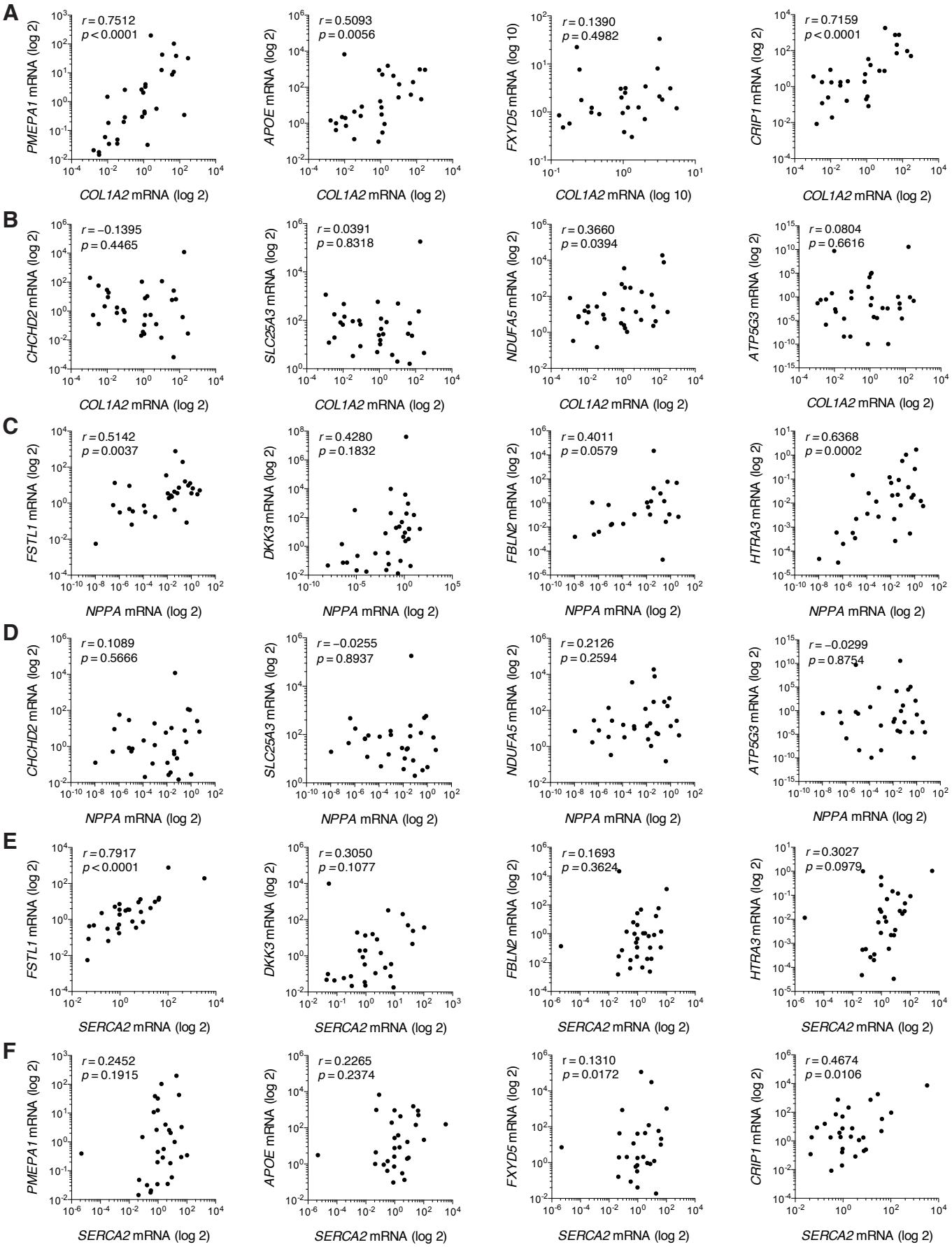
Supplemental Figure 4



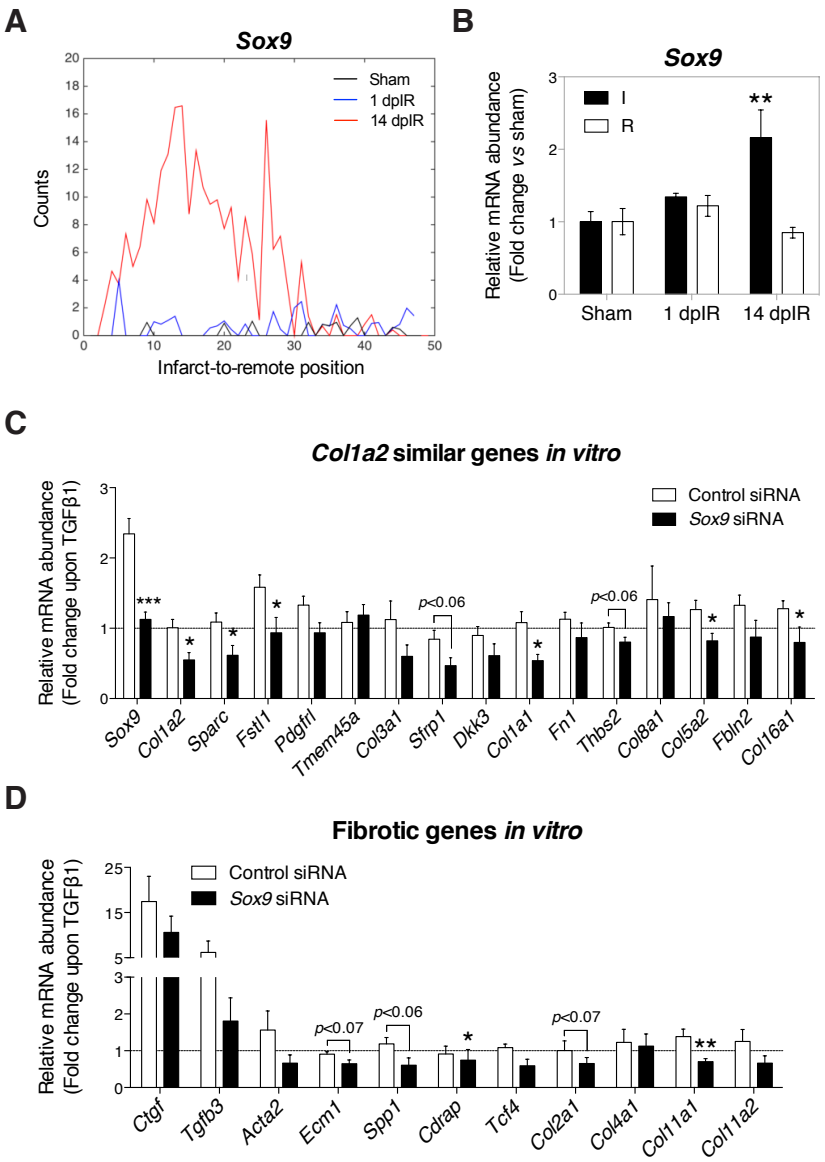
Supplemental Figure 5



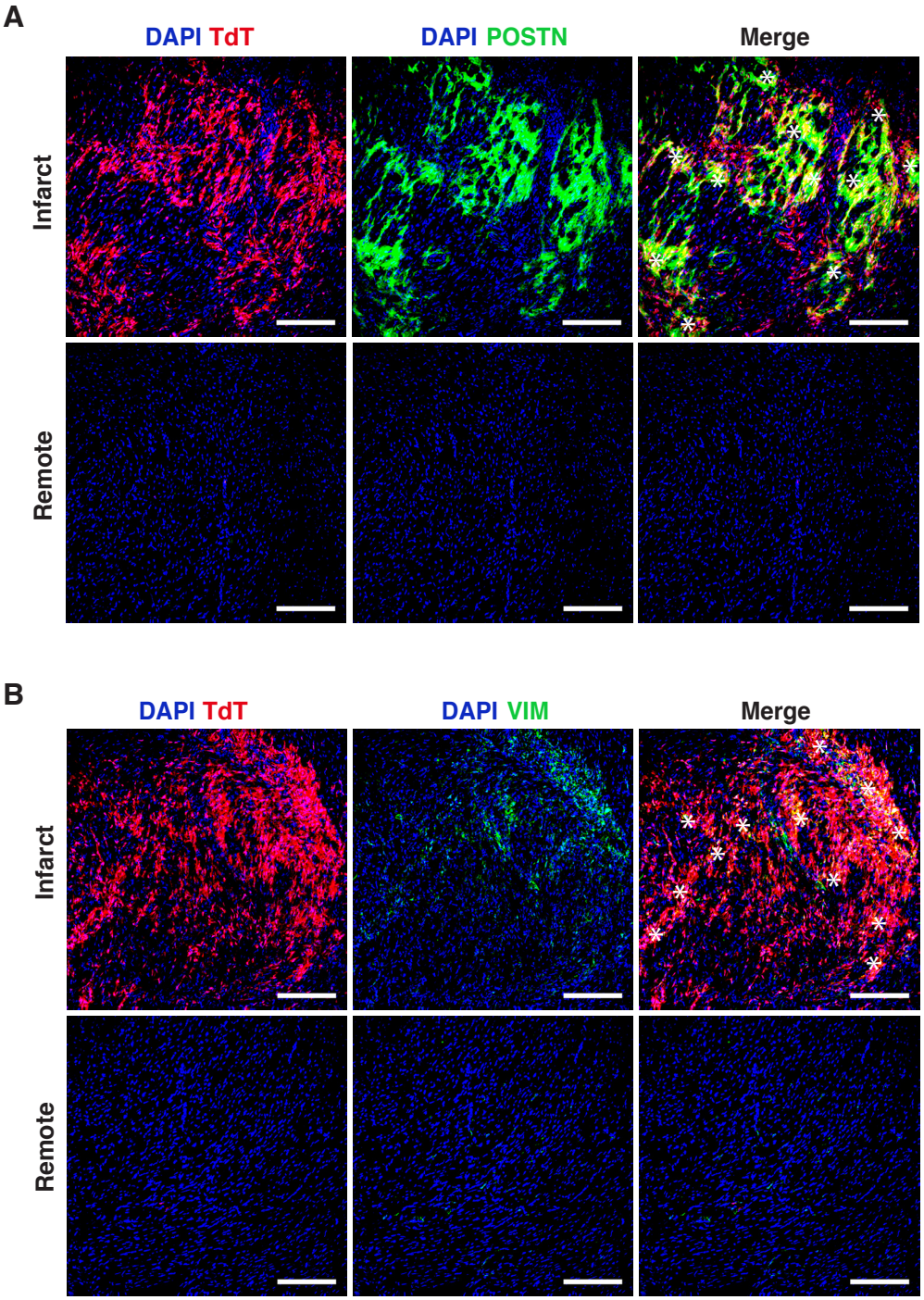
Supplemental Figure 6



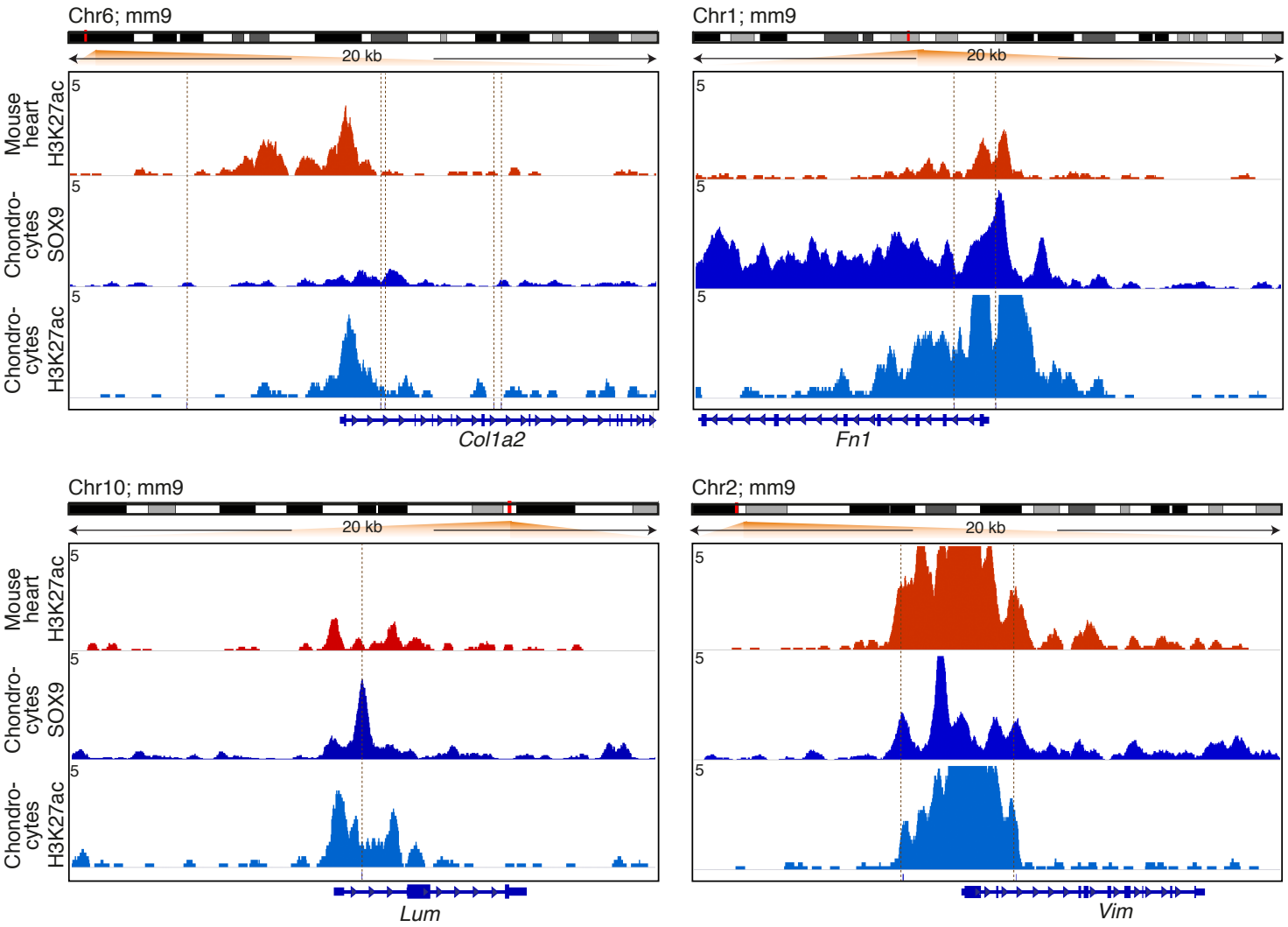
Supplemental Figure 7



Supplemental Figure S8



Supplemental Figure 9



Supplemental Figure Legends

Supplemental Figure 1. Cardiac ischemia-reperfusion injury. A, Histological sections of hearts stained for Hematoxylin and Eosin (H&E) or Sirius Red from Sham, 1, and 14 dpIR mice. Scale bars, 1 mm. **B**, Real-time PCR analysis of three reference genes associated to cardiac fibrosis (*Col1a2*), remodeling (*Nppa*), and contractility/calcium handling (*Serca2*) in infarct (I) and remote (R) zones in Sham, 1, and 14 dpIR mice (n=6 per group; $^{**}p<0.01$ *vs* infarct).

Supplemental Figure 2. High resolution gene expression atlas of the infarcted heart by tomo-seq. A, Pairwise correlation for all sections across all genes expressed at >4 reads in >1 section analysed by tomo-seq 1 dpIR in one biological replicate. **B**, Pairwise correlation for all sections across genes exhibiting at least two-fold and statistically significant differential expression between the infarct and remote zones by tomo-seq at 1 dpIR. **C**, Hierarchical clustering of expression traces for all genes exhibiting at least two-fold and statistically significant differential expression between the infarct and remote zones, using Euclidean distance as metric at 1 dpIR. **D**, KEGG analysis showing the enriched pathways the genes exhibiting a two-fold differential expression between the infarct and remote zones at 1 or 14 dpIR are involved in. Pathways are ranked by their respective p value corrected by the Benjamini-Hochberg method.

Supplemental Figure 3. Confirmation of tomo-seq data. A through **C**, Spatial expression pattern of *Col1a2* (**A**), *Nppa* (**B**), and *Serca2* (**C**) determined by tomo-seq 14 dpIR in two distinct biological replicates (14 dpIR #1 and #2). Traces show gene expression as the number of read counts (*left*) or after *Z*-score normalization (*right*). *Z*-score normalization removes overall expression differences between genes by setting the mean to 0 and the standard

deviation to 1. **D**, Spatial expression traces of ten co-regulated genes 14 dpIR in the second biological replicate based on the data obtained in the first biological replicate. Reference genes are shown in red, and ten most similar genes are shown in grey. **E**, Spearman correlation between the two sets of samples. Data are shown as total number of reads (summed over all sections) for each gene.

Supplemental Figure 4. Validation of tomo-seq by RNA-seq. Validation of the differential gene expression for the one hundred thirty most similar genes to *Col1a2* (**A**), *Nppa* (**B**), and *Serca2* (**C**) between Sham and 14 dpIR hearts by RNA-seq performed on the complete injured area or a corresponding region in the Sham hearts (*n*=3 per group).

Supplemental Figure 5. High resolution lncRNA expression atlas of the infarcted heart by tomo-seq. A, Pairwise correlation for all sections across all non-coding genes detected at greater than four reads in more than one section by tomo-seq in the hearts from Sham, 1, and 14 dpIR mice. Especially at 14 dpIR, anatomical regions of the infarcted heart i.e., infarct and remote zones are detectable as clusters of correlated sections. **B**, Hierarchical clustering of expression traces for all genes exhibiting a peak of expression using Euclidean distance as metric. **C**, Spatial expression traces of detected non-coding genes by tomo-seq in the hearts of Sham, 1, and 14 dpIR mice. Traces are shown as read counts.

Supplemental Figure 6. Lack of correlation between genes from different reference lists. Real-time PCR analysis on human ischemic heart tissue indicating the lack of correlation when cross-referencing genes from different lists. **A**, *COL1A2* expression correlation with the expression of *NPPA* co-regulated genes. **B**, *COL1A2* expression correlation with the expression of *SERCA2* co-regulated genes. **C**, *NPPA* expression correlation with the expression of

COL1A2 co-regulated genes. **D**, *NPPA* expression correlation with the expression of *SERCA2* co-regulated genes. **E**, *SERCA2* expression correlation with the expression of *COL1A2* co-regulated genes. **F**, *SERCA2* expression correlation with the expression of *NPPA* co-regulated genes. Control hearts and remote, border-zone and infarct zones from ischemic hearts are plotted. Data are presented as log 2 transformed values. Pearson correlation (r) and significance of co-regulated gene expression is shown (n=27–34; p<0.05 is considered as significant).

Supplemental Figure 7. Regulation of ECM-related genes. A, Spatial expression pattern of *Sox9* determined by tomo-seq in the hearts from Sham, 1, and 14 dpIR mice. **B**, Real-time PCR analysis of *Sox9* in the infarct (I) and remote (R) zones in the hearts from Sham, 1, and 14 dpIR mice. Data are displayed as fold change *vs* matched Sham hearts (n=5–6 per group; **p<0.01). **C** and **D**, Real-time PCR analysis of *Col1a2* co-expressed genes and other fibrotic genes including SOX9 target genes in 3T3-L1 fibroblasts activated with TGFβ1 (5 ng/ml) for 3 hrs after knock-down of *Sox9* using a siRNA (10 nM) or a control siRNA (n=5 in two independent experiments; *p<0.05; **p<0.01; ***p<0.001 *vs* control siRNA).

Supplemental Figure 8. Lineage tracing of SOX9-expressing cells. Co-immunostaining for TdTomato (TdT) and Periostin (POSTN) (**A**) or Vimentin (VIM) (**B**) in the hearts from sham-operated mice and 14 dpIR. White stars in the Merge field indicate SOX9-TdT-positive regions. Scale bars, 200 μm.

Supplemental Figure 9. Peak intensity plots of ChIP-seq reads for SOX9 and histone modification marks (H3K27ac) in four ECM related genes induced in the heart 14 dpIR. Normalized read signals from each ChIP-seq were plotted on a 20-kb window flanking the SOX9 binding sites predicted by oPOSSUM 6.0 (cf. Supplemental Table 6). ChIP-seq signals

1 for H3K27Ac obtained in the heart (ENCODE dataset) were mapped to the corresponding

2 SOX9 peak regions obtained in chondrocytes.[8]

3

Supplemental References

1. Junker JP, Noel ES, Guryev V, Peterson KA, Shah G, Huisken J, McMahon AP, Berezikov E, Bakkers J and van Oudenaarden A. Genome-wide RNA Tomography in the zebrafish embryo. *Cell*. 2014;159:662-675.
2. Hashimshony T, Wagner F, Sher N and Yanai I. CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. *Cell Rep*. 2012;2:666-673.
3. Li H and Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2010;26:589-595.
4. Huang da W, Sherman BT and Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009;4:44-57.
5. Pritchett J, Harvey E, Athwal V, Berry A, Rowe C, Oakley F, Moles A, Mann DA, Bobola N, Sharrocks AD, Thomson BJ, Zaitoun AM, Irving WL, Guha IN, Hanley NA and Hanley KP. Osteopontin is a novel downstream target of SOX9 with diagnostic implications for progression of liver fibrosis in humans. *Hepatology*. 2012;56:1108-1116.
6. Akiyama H, Chaboissier MC, Martin JF, Schedl A and de Crombrugghe B. The transcription factor Sox9 has essential roles in successive steps of the chondrocyte differentiation pathway and is required for expression of Sox5 and Sox6. *Genes Dev*. 2002;16:2813-2828.
7. Kopp JL, Dubois CL, Schaffer AE, Hao E, Shih HP, Seymour PA, Ma J and Sander M. Sox9+ ductal cells are multipotent progenitors throughout development but do not produce new endocrine cells in the normal or injured adult pancreas. *Development*. 2011;138:653-665.
8. Ohba S, He X, Hojo H and McMahon AP. Distinct Transcriptional Programs Underlie Sox9 Regulation of the Mammalian Chondrocyte. *Cell Rep*. 2015;12:229-243.