

Figure S1. Related to Figure 1 and 2.

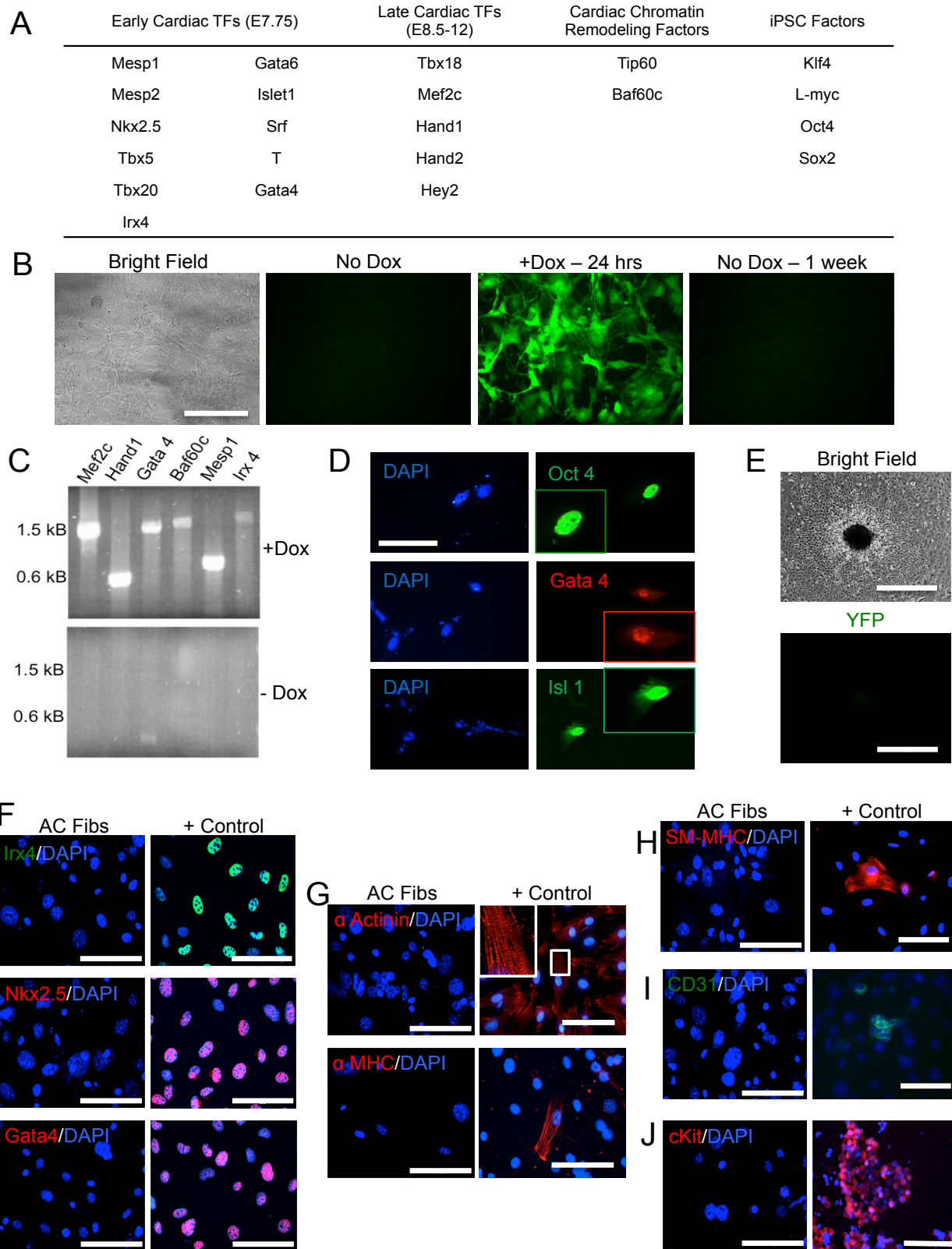


Figure S2. Related to Figure 1, 2 and 5.

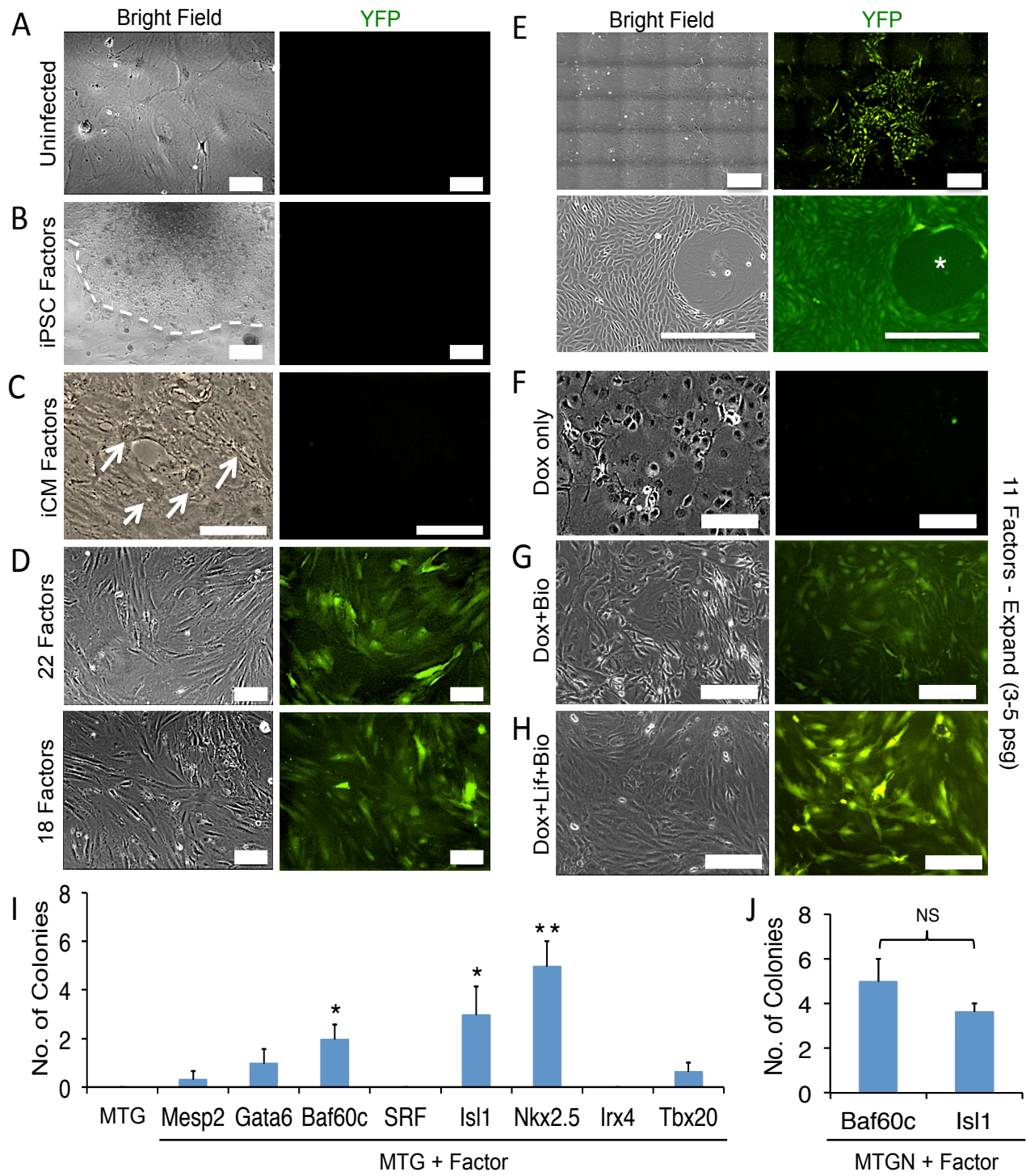


Figure S3. Related to Figure 1, 2, 3, 4 and 5.

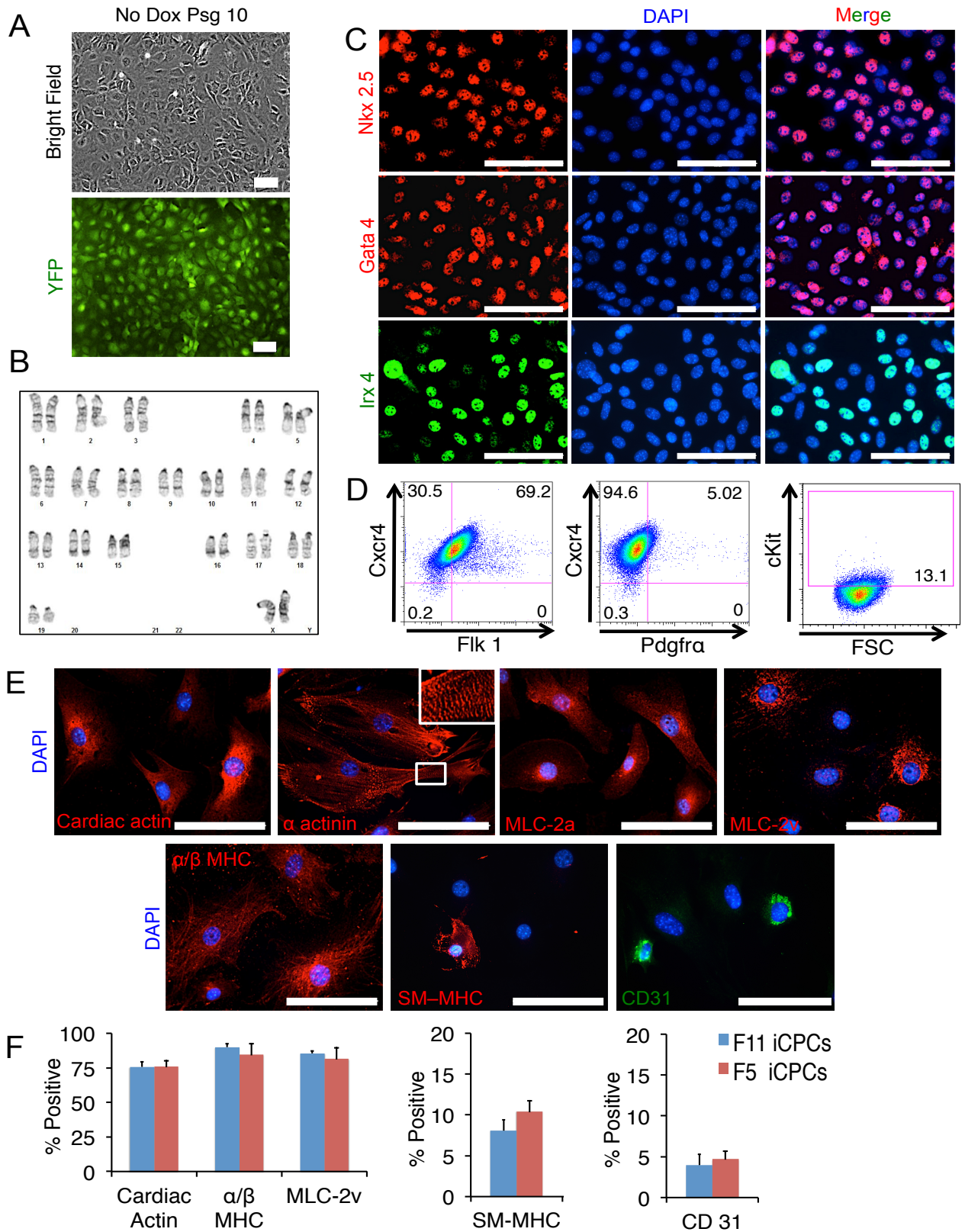


Figure S4. Related to Figure 2 and 3.

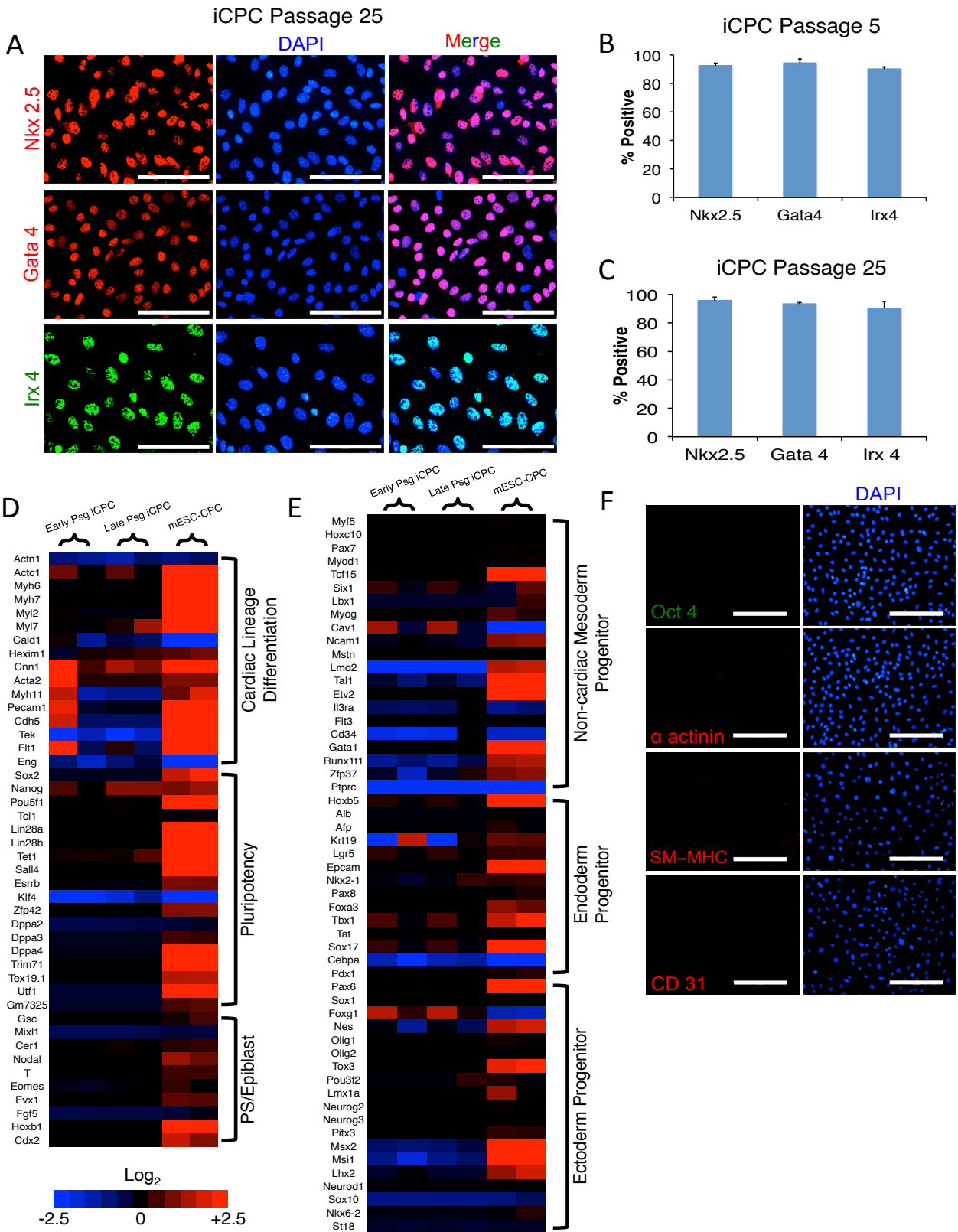


Figure S5. Related to Figure 3 and 4.

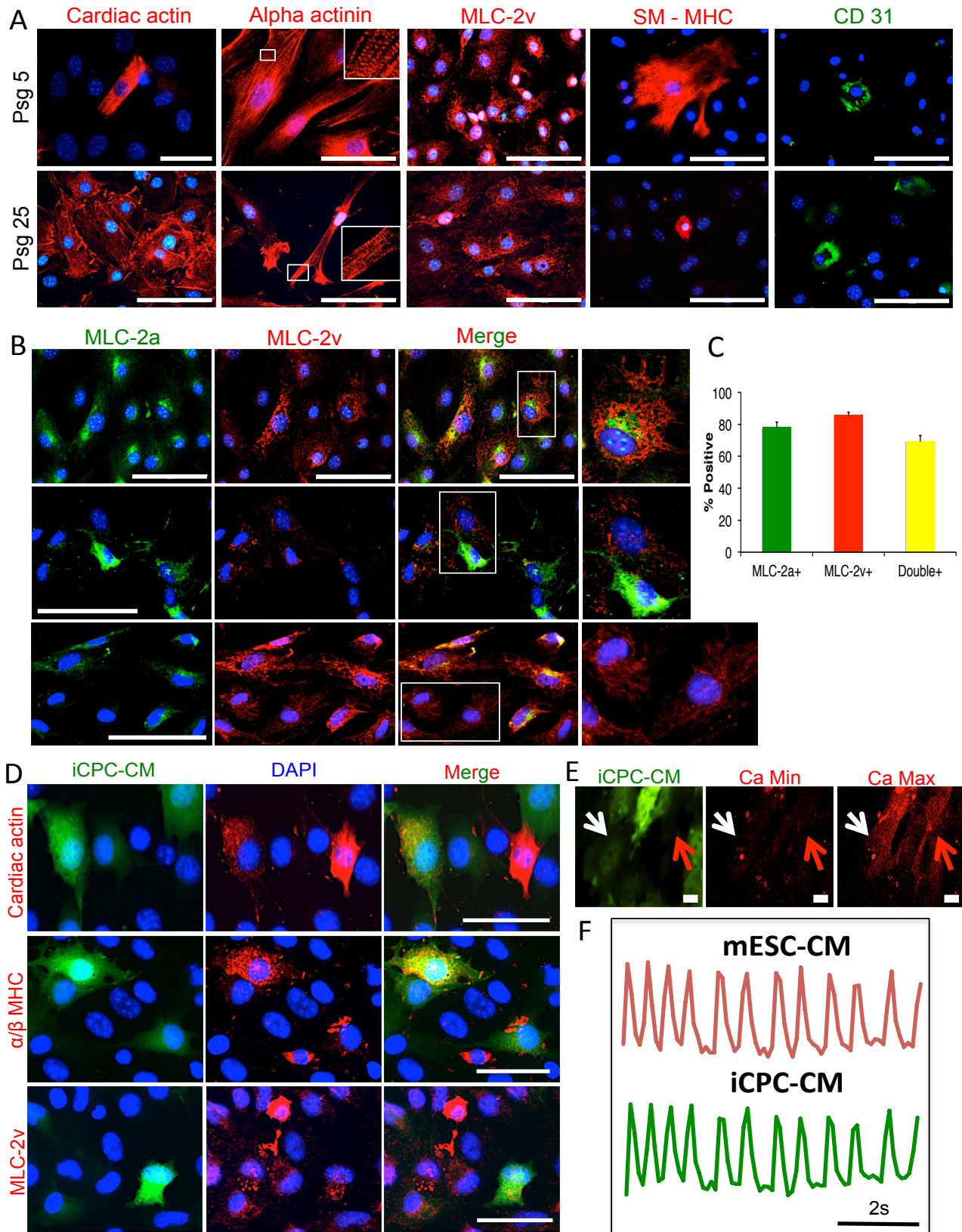


Figure S6. Related to Figure 6.

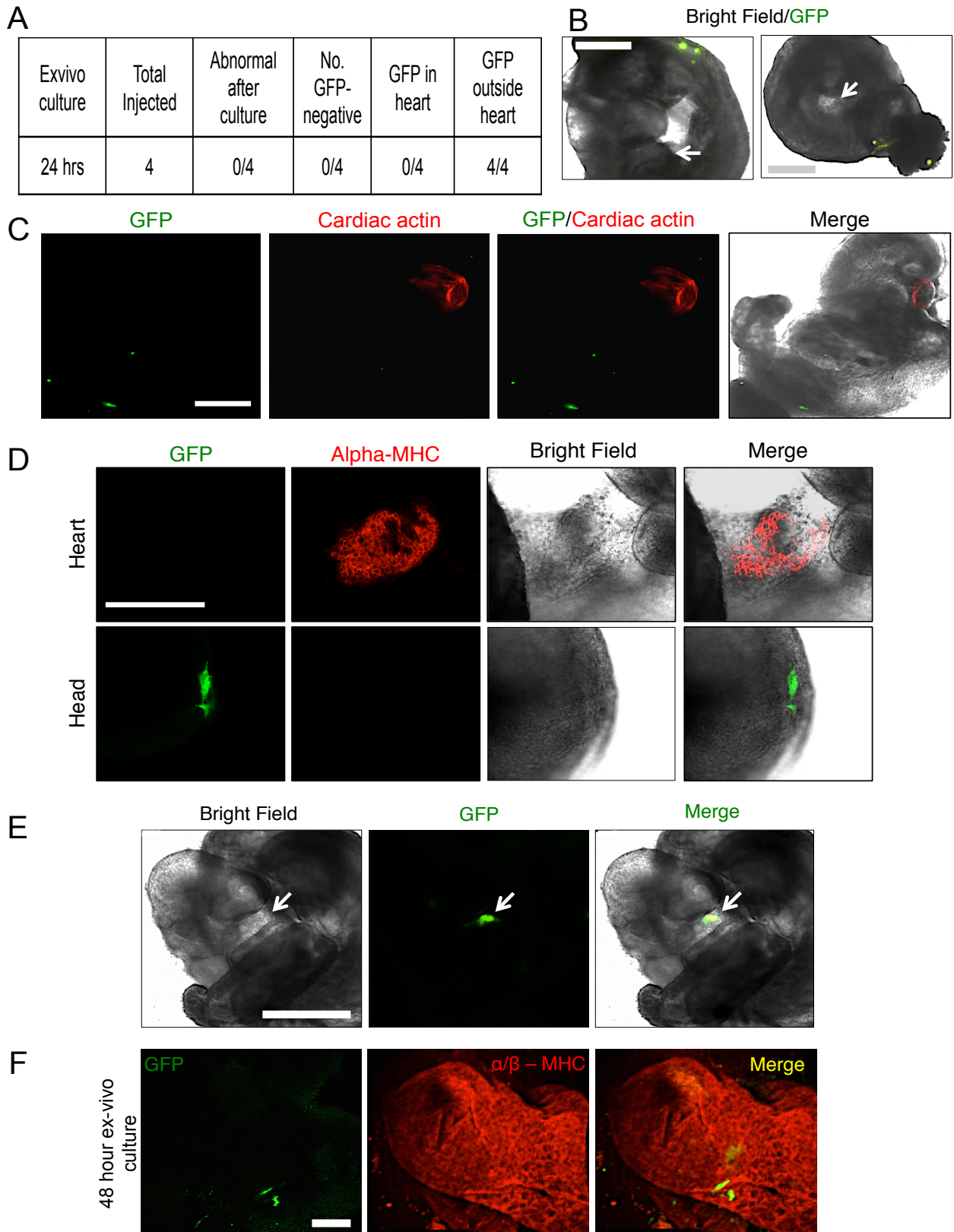
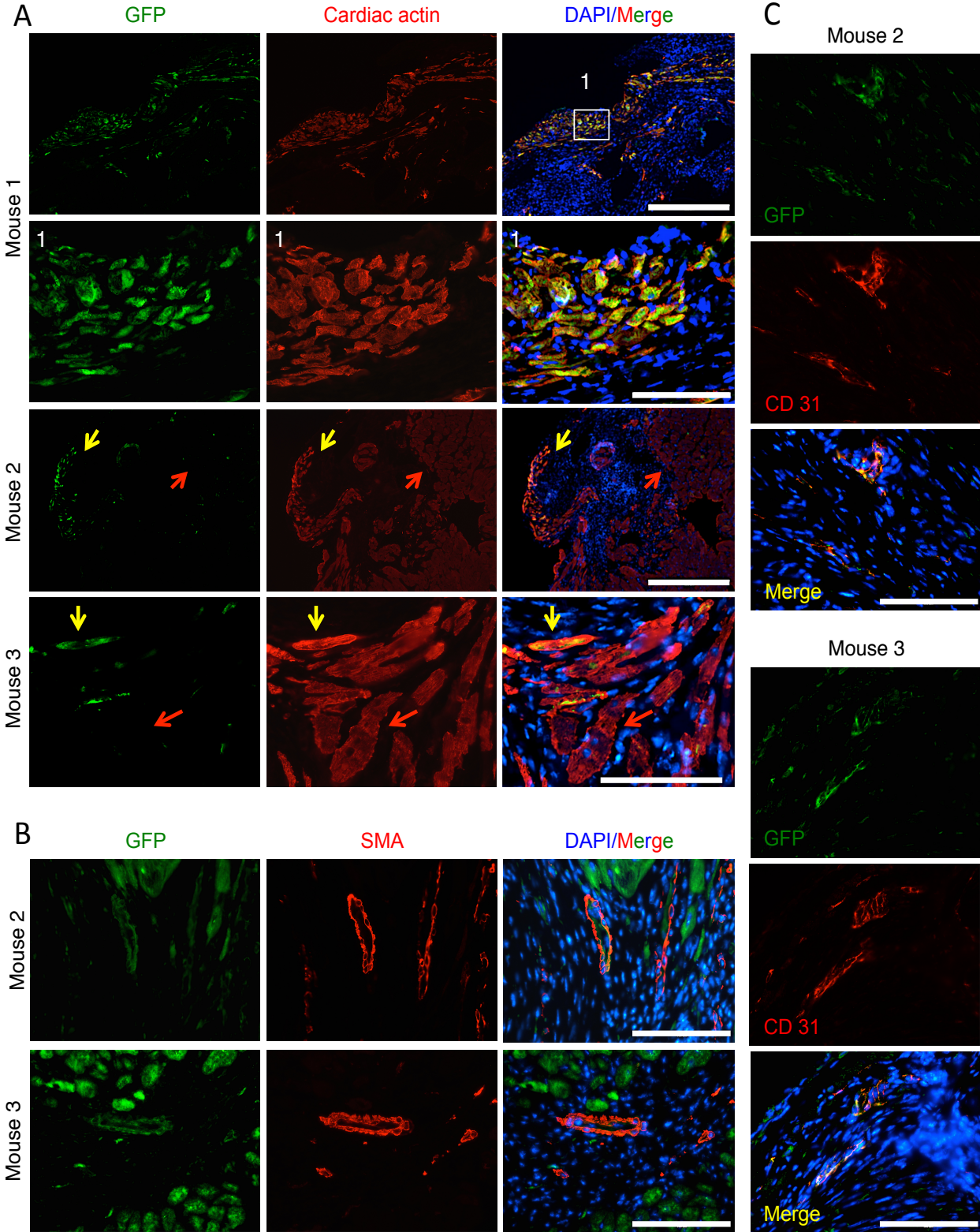


Figure S7. Related to Figure 7.



SUPPLEMENTAL FIGURE LEGENDS

Figure S1. Generation of Defined Factor Library and Characterization of Adult Cardiac Fibroblasts, Related to Figure 1 and 2

(A) Defined factors used for iCPC reprogramming screen. (B) Fibroblasts infected with dox-inducible GFP lentivirus were imaged before and 24 hrs after dox induction. GFP fluorescence was observed only after dox induction. (C) DNA gel showing RT-PCR results of fibroblasts infected with individual factors. Gene expression was detected only in the presence of dox. Bands depict full-length genes. (D) Fibroblasts infected with single factors were treated with dox for 48 hrs and immunocytochemistry was performed. Note nuclear staining for the 3 TFs tested. Uninfected fibroblasts showed no detectable immunolabeling for these TFs. (E) Adult cardiac fibroblasts (AC Fibs) were isolated by explant culture of adult heart tissue from 1-3 month old Nkx2.5-EYFP/rtTA mice. AC Fibs migrated out of tissue pieces. Note that both AC Fibs as well as heart tissue did not express EYFP. (F) AC fibs showed no staining for CPC TFs, CM markers (G), as well as smooth muscle cell (H) or endothelial cell markers (I). (J) AC Fibs showed no staining for adult cardiac stem cell marker cKit. Embryoid body differentiated mESCs were used as positive control (F-I). mESCs were used as positive control for J. Scale bar represents 400 μ m in B, 100 μ m in D,F-J, 1000 μ m in E.

Figure S2. Screening for iCPC Inducing Factors and Culture Conditions, Related to Figure 1, 2 and 5

(A) Uninfected AC Fibs showed no EYFP expression. (B) AC Fibs infected with iPS factors (Klf4, Oct4, L-myc, Sox2) showed proliferative cells that formed iPS-like colonies which were EYFP⁺ (3 weeks after adding dox). (C) Neonatal cardiac fibroblasts infected with Gata4, Mef2c, Tbx5 (iCM factors) showed spontaneous beating 25 days after dox treatment. Arrows indicate induced cardiomyocytes (iCMs). iCMs did not show EYFP expression. (D) AC Fibs infected with 22 or 18 combinations of defined factors developed into EYFP⁺ cells 3 weeks after dox induction. (E) 11-factor infected AC Fibs developed into two dimensional, proliferative colonies of EYFP⁺ cells. Images taken 3 weeks after dox treatment show a colony of reprogrammed EYFP⁺ cells, surrounded by EYFP⁻ fibroblasts. Lower images show the striking morphological difference between EYFP⁺ reprogrammed cells and EYFP⁻ fibroblasts (indicated by *). Reprogrammed cells lost parental fibroblast morphology and exhibited a high nuclear-cytoplasmic ratio. (F) AC Fibs infected with 11 factors and induced with dox developed proliferative EYFP⁺ colonies; however, upon expansion in dox only culture condition these cells lost EYFP expression and senesced within 3-5 passages. (G) Addition of BIO resulted in expandable EYFP⁺ cells. However, they became spindle-shaped and were not highly proliferative. (H) Addition of both LIF and BIO produced the brightest EYFP⁺ cells that were robustly expandable. (I & J) Number of EYFP⁺ colonies produced (per 50,000 seeded cells) after infection with respective factor combinations and culture in iCPC induction medium for 3 weeks (n=3) (* p<0.05, ** p<0.01). Scale bar represents 100 μ m in A-D, 500 μ m in E and 200 μ m in F-H. See also movie S1.

Figure S3. Five Factors Stably Reprogram Adult Cardiac Fibroblasts into Proliferative and Multipotent iCPCs. Related to Figure 1, 2, 3, 4 and 5

(A) 5F iCPCs could be stably expanded without dox. (B) 5F iCPCs showed a normal diploid karyotype. (C) Immunolabeling of F5 iCPCs showed nuclear localization of TFs Nkx2.5, Gata4, and Irx4. (D) 5 factor iCPCs expressed surface markers associated with CPCs as shown by flow cytometry

analysis. (E) F5 iCPCs were multipotent and could be differentiated into CMs (cardiac actin, α -actinin, MLC-2a, MLC-2v, α/β MHC), SMs (SM-MHC) and ECs (CD 31). Note highly organized sarcomere staining for α -actinin. (F) F11 and F5 iCPCs showed comparable cardiac lineage differentiation efficiencies. F11 and F5 iCPCs were differentiated in cardiac differentiation medium and stained for the respective cardiac lineage markers. 5-8 microscopic fields were randomly selected per coverslip per marker. Total nuclei relative to nuclei associated with particular immunolabel were counted. F11 (3 differentiation experiments, Total cells counted=1537), F5 (2 differentiation experiments, Total cells counted=1173) Scale bars = 100 μ m.

Figure S4. iCPCs are Cardiac Mesoderm Restricted Progenitors and Maintain Expression of Cardiac Progenitor TFs Across 25 Passages, Related to Figure 2 and 3. (A) F11 iCPCs passaged without dox showed nuclear expression of CPC TFs when immunostained after 5 passages (B) as well as 25 passages (A&C). (D-E) Heat map of RNA Seq data illustrating differentially expressed genes in early psg (1-3), late psg (8-10) F11 iCPCs and mESC-CPCs (Wamstad) as compared to AC Fibs. (n=2, biological replicates in each group). (D) Expression of cardiac lineage differentiation markers as well as pluripotency and epiblast associated genes remained unchanged in iCPCs as compared to AC Fibs. However, mESC-CPCs had higher expression for cardiac differentiation as well as pluripotency-associated genes as compared to AC Fibs. (E) The majority of the genes associated with non-cardiac mesoderm (blood/hepatocyte/skeletal muscle), endoderm and ectoderm progenitors remained unchanged or were downregulated in iCPCs as compared to AC Fibs. mESC-CPC samples had higher expression for some endoderm, ectoderm and non-cardiac mesoderm progenitor markers as compared to AC Fibs. (F) F11 iCPCs passaged without dox for 25 psg did not show expression for pluripotency (Oct4), CM (α actinin), smooth muscle (SM-MHC) or endothelial (CD31) markers. Data presented as mean, error bars indicate standard error of mean. Scale bar represents 100 μ m in A-C, and 200 μ m in E.

Figure S5. iCPCs are Multipotent and Differentiate into Contracting Cardiomyocytes, Smooth Muscle Cells and Endothelial Cells in vitro, Related to Figure 3 and 4. (A) F11 iCPCs maintain multipotency with extended passaging without dox and differentiate into 3 types of cardiac lineages (CMs, SMs, ECs). (B) Majority of MLC-2v positive iCPC-CMs showed staining for MLC-2a indicating they were immature CMs. However, a minority (<5%) of iCPC-CMs stained exclusively for MLC-2a or MLC-2v. Data quantified in (C). (D) F11 iCPC-CMs co-stained for GFP and CM markers, whereas mESC-CMs only stained for CM markers. (E&F) iCPC-CMs showed synchronous calcium transients with mESC-CMs 3 weeks after co-culture. White arrow = iCPC-CM, red arrow = mESC-CM. Data presented as mean, error bars indicate standard error of mean. Scale bar represents 100 μ m in A&B, 50 μ m in D, 10 μ m in E and 2 seconds in F. See also Movie S2, S3.

Figure S6. iCPCs, but not AC Fibs, Differentiate into Cardiomyocytes Upon Injected into Cardiac Crescent of Mouse Embryos. Related to Figure 6. (A) Shows the number of embryos injected with AC Fibs and their location 24 hrs after whole embryo culture. (B) AC Fibs (labeled with GFP expressing lentivirus) injected into the cardiac crescent were excluded from the developing heart tube as assessed after 24 hrs of whole embryo culture.

(C&D) Embryos injected with AC Fibs were immunostained in whole mount preparations for CM markers and GFP. AC Fibs did not differentiate into CMs, as indicated by absence of staining for CM markers. (E) iCPCs (labeled with GFP expressing lentivirus) injected into the cardiac crescent of mouse embryos colonized the developing heart tube as assessed after 24 hrs of whole embryo culture. (F) 3D reconstruction images show iCPCs differentiated into CMs, as indicated by co-expression of α/β -MHC and GFP. iCPC-CMs attained shape/size similar to native CMs after 48 hrs of ex vivo culture. Scale bar = 500 μ m in B-E, 100 μ m in F. See also Movie S4-7.

Figure S7. iCPCs Injected into the Adult Mouse Heart Post-MI Differentiate Into Cardiac Lineage Cells In vivo, Related to Figure 7. (A) Injected F5 iCPCs engrafted within the left ventricular wall and differentiated into cardiomyocytes based on organized cardiac actin immunolabeling, as assessed 28 days after injection. Note that the intensity of cardiac actin staining in iCPC-derived CMs (yellow arrow) is comparable to that in host CMs (red arrow). Scale bars - Mouse 1 = 400 μ m, 100 μ m in inset, Mouse 2 = 400 μ m, Mouse 3 = 100 μ m. (B) iCPCs differentiated into smooth muscle cells within scar tissue based on co-expression of GFP and smooth muscle actin (SMA), as assessed 28 days after injection. Scale bar = 100 μ m. (C) iCPCs differentiated into endothelial cells within scar tissue based on co-expression of GFP and CD31, as assessed 28 days after injection. Scale bar = 100 μ m.

SUPPLEMENTAL MOVIE LEGENDS

Movie S1. Neonatal Cardiac Fibroblasts are Reprogrammed to iCMs by GMT Factors, Related to Figure 1 and Figure S3. Neonatal cardiac fibroblasts were infected with GMT and treated with dox. Spontaneously contracting, EYFP⁺ iCMs were observed 4 weeks after dox induction.

Movie S2. iCPC-CMs contracted without cell fusion after co-culture with mESC-CMs, Related to Figure 4. iCPC-CMs were co-cultured with mESC-CMs. Contracting iCPC-CMs were observed two weeks after co-culture.

Movie S3. iCPC-CMs Showed Spontaneous Calcium Transients, Related to Figure 4. Cells were labeled with Rhod-2 dye and observed via epifluorescence microscope. iCPC-CMs showed spontaneous calcium oscillations that were synchronous with mESC-CMs.

Movie S4. iCPCs colonize the developing heart tube upon injection into the cardiac crescent, Related to Figure 5. GFP labeled F11 iCPCs were injected into the cardiac crescent of mouse embryos. iCPCs localized to the developing heart tube as assessed by live epifluorescence microscopy performed after 24 hrs of whole embryo culture.

Movie S5. Live imaging of 48 hour ex vivo cultured embryos demonstrates contracting heart tube with integrated iCPC-derived cells, Related to Figure 5. GFP labeled iCPCs were injected into the cardiac crescent of mouse embryos and the embryos were cultured for 48 hrs in whole embryo culture. Live imaging revealed that iCPC-derived cells appear to contract with the native heart tube.

Factor Combinations

22 Factors – *T, Mesp1, Mesp2, Tbx5, Mef2c, Isl1, Gata4, Gata6, Irx4, Nkx2.5, Hand1, Hand2, Tbx20, Tbx18, Tip60, Baf60c, SRF, Hey2, Oct4, Klf4, Sox2, L-myc.*

18 Factors - *T, Mesp1, Mesp2, Tbx5, Tbx20, Isl1, Gata4, Gata6, Irx4, Nkx2.5, Hand1, Hand2, Tbx20, Tbx18, Tip60, Baf60c, SRF, Hey2.*

11 Factors – *Mesp1, Mesp2, Gata4, Gata6, Baf60c, SRF, Isl1, Nkx2.5, Irx4, Tbx5, Tbx20.*

5 Factors – *Mesp1, Tbx5, Gata4, Nkx2.5, Baf60c.*

Cardiac Reporter Transgenic Mice

Male *Nkx2.5-EYFP* mice as previously described (Masino et al., 2004) were mated with female mice homozygous for *rtTA* (B6.Cg-Gt(ROSA)26Sortm1(rtTA*M2)Jae/J – Jackson Labs) to generate double transgenic (*Nkx2.5-EYFP/rtTA*) inducible reporter mice which were used for all subsequent studies (Figure 1A).

Production of Lentivirus Particles

HEK 293 TN cells (SBI) were used for lentivirus production. Briefly, 4.5×10^6 cells were plated in a 10cm dish on the day before transfections. Lipofectamine 2000 (Invitrogen) was used for transfections (1:2 ratio). Transfection conditions used were – 7ug lentivirus plasmid, 10ug psPAX2 (Addgene plasmid #12260 - packaging), 5ug pMD2.G (Addgene plasmid #12259 - envelope). Transfection media was incubated for 15-16 hours, after which it was replaced with 5mls of fibroblasts medium. Lentivirus supernatant was collected after 48-52 hours, filtered (0.45 uM – Millipore) and frozen down at -80C. Lentivirus supernatants were thawed in 37°C water bath immediately before infection.

Immunocytochemistry and Flow Cytometry

For immunocytochemistry and flow cytometry cells were fixed in methanol free formaldehyde (4%) for 12 min at room temperature and then permeabilized with 0.1% triton X for 6 mins at room temperature. Cells were then blocked in 2% serum (Goat or Donkey), 5% BSA in PBS for 1 hour at room temperature. Primary antibodies were incubated at 4°C overnight in blocking buffer containing 0.1% triton X. Secondary antibodies were incubated for 2 hours at room temperature in blocking buffer containing 0.1% triton X. For surface markers triton free blocking buffer was used. For immunohistochemistry heart section were fixed in 4% PFA at 4°C overnight then embedded in paraffin and stained with respective antibodies (after antigen retrieval). Primary antibodies used and their respective dilutions were – *Nkx2.5* (RD Systems – 1:100), *Gata4* (Santacruz - 1:200), *Irx4* (undiluted supernatant - Abgent), *Alpha actinin* (Sigma – 1:250), *Cardiac actin* (Sigma – 1:400), *MLC-2a* (Synaptic systems – 1:200), *MLC-2v* (Protein tech – 1:200), α/β *MHC* (Thermo Scientific – 1:200), *Cx 43* (Sigma – 1:100), *GFP* (Invitrogen – 1:200), *Cxcr4* (Biolegend 647/APC conjugate – 1:50), *Flk1* (BD Pharmingen PE conjugate – 1:20), *Pdgfr- α* (Santacruz PE conjugate – 1:20), *cKit* (BD Pharmingen PE/Cy7 conjugate -1:50), *CD31* (BD Pharmingen - 1:400), *SM-MHC* (Biomedical technologies - 1:250) *cKit* (Santacruz – 1:200). All secondary antibodies were used at 1:500 dilution (alexa 488/568/633/647).

Immunohistochemistry of Post-MI hearts

For immunohistochemistry heart sections were fixed in 4% PFA at 4°C overnight then embedded in paraffin and stained with respective antibodies (after antigen retrieval). Primary antibodies used and their respective dilutions were – Cardiac actin (Sigma – 1:50), CD31 (Abcam - 1:50), SMA (Sigma – 1:50), GFP (Invitrogen – 1:50), GFP (Abcam – 1:50). Secondary antibodies were used at 1:500 dilution (Alexa 568/633/647).

Co-culture with mESC-derived CMs

An Irx4 reporter mESCs cell line was utilized in which firefly luciferase, hph (hygromycin resistance), and tdTomato were placed downstream of endogenous Irx4 promoter. The reporter genes were separated using 2A peptide sequences to allow co-expression. CMs were obtained by embryoid body differentiation of mESCs. Hygromycin was added to select for Irx4+ CMs. All CMs expressed td-tomato. iCPC-CMs were infected with a GFP lentivirus (Addgene #17448) to trace the reprogrammed cells. iCPC-CMs and mESC-CMs were co-cultured at a ratio of 1:10 for 15-30 days in DMEM+10%FBS containing media.

Embryo Injections, Immunostaining and Imaging

For embryo injection experiments, and subsequent whole-mount immunofluorescence (IF) staining, we used F2 hybrid conceptuses from matings between inbred hybrid B6CBAF1 (Jackson laboratories, Bar Harbor, Maine). Mice were maintained on a 12-hour light/dark cycle (lights off: 13.00/lights on: 1.00). Timed matings were carried out as previously described (Champlin et al., 1973). Although 19.00 of the day of pairing was considered day 0.5, females were checked for plugs at 17.00, or 4 hours after mating. All dissections were then carried out at 15.00 eight days later and categorized according to stage, which ranged from very early headfold to 6 somite pairs (Downs and Davies, 1993). Dissections and whole mount embryo culture were carried out according to (Downs, 2006): staging of the embryos before culture was conducted in accordance with the morphological staging system described by (Downs and Davies, 1993). At the end of culture, embryos were immediately scored for heart beat and circulation, followed by closer scrutiny of extent of morphological development (heart, chorio-allantoic fusion, turning, brain, optic and otic development, and somite number) in dissection medium (Downs and Harmann, 1997). Embryos were then rinsed in 1x PBS and fixed with 4% paraformaldehyde (Sigma), prepared in 1x PBS, at 4°C, for 45 minutes with gentle agitation. Embryos were dehydrated with a gradient of increasing concentrations (25, 50, 75, & 100%) of methanol (Fisher Scientific) prepared in deionized water for storage at -20°C. For imaging we used a TE300 inverted microscope (Nikon, Tokyo, Japan) as previously described (Conklin et al., 2009) equipped with a Plan APO VC 20X (N.A. 0.75; Nikon Instruments, Tokyo, Japan) objective lens by using a mode-locked Ti:Sapphire laser (Spectra Physics Mai Tai, Mountain View, CA) using an excitation wavelength of 890 nm with either a 520 nm +/- 35 nm narrow band pass emission filter or a 580nm LP emission filter.

RNA-seq and Bioinformatics Analysis

Sequencing libraries were generated from 100 ng of total RNA using the TruSeq RNA Sample Preparation kit

(Illumina) using the full reaction volumes required by the Illumina protocol and quantitated using the Qubit fluorometer (Life Technologies) following the manufacturer's instructions. The libraries were then pooled using 45 ng/sample for samples 1, 4, and 6, and 25 ng/sample for samples 2, 3, 5, and 7. All pooling was performed for a 51+8+8 cycle Single Read high output run on a HiSeq 2500 (Illumina). The sequencer outputs were processed using Illumina's CASAVA-1.8.2 basecalling software. Demultiplexing assigned 18-34 million (M) reads per sample. ~3M were discarded for low quality or the presence of sequencing adapters in the reads. Each sample's reads were then processed using RSEM version 1.2.3 (with bowtie-0.12.9 for the alignment step). RSEM input parameters were specified such that alignments were suppressed for a read if over 200 valid alignments existed ('--bowtie-m' option), a maximum of two mismatches were allowed in the seed ('--bowtie-n' option), intermediate files were kept ('--keep-intermediate-files' option), binary BAM output file was suppressed ('--no-bam-output' option), probability of generating a read from the forward strand of the transcript was 0.5 since reads were not strand-specific ('--forward-prob' option), ten processing threads were used per sample ('-p'), and a seed length of 28 was used by the read aligner ('--seed-length' option) (Li et al., 2010). The percentage of reads that mapped to the RefSeq mm09 reference transcriptome ranged from 79% to 86%. To obtain a preliminary narrowed list of salient genes to focus on for the remainder of the experiment, differential expression analysis between samples 1, 4, and 6 was done pairwise using EBSeq version 1.5.3 (Leng et al., 2013). The EBSeq input dataset contained un-normalized expected count values for all genes as output from RSEM for each sample paired with another (sequence-independent) with condition strings as represented in the table below along with the size factors calculated by median normalization (MedianNorm function within the EBSeq package). Additional input parameters to EBSeq (specifically the EBTest function) specified five total iterations were to be run (maxround = 5), genes with similar means were to be grouped into 1000 bins (NumBin = 1000), no pooling was to be used (Pool = F), transcript variances with mean less than a variance cutoff of 10⁻¹⁰ (ApproxVal = 10⁻¹⁰) were approximated as the mean divided by (1 - 10⁻¹⁰), all model parameters (Alpha, Beta, PInput, RInput) were null so that all probabilities were estimated from the data, initial candidate genes for differential expression were taken from the 25%-75% quantile (PoolLower = .25, PoolUpper = .75), and transcripts with all zero were to be removed from the dataset (Qtrm = .99, QtrmCut = 0). The targeted false discovery rate (FDR) used for each run of EBSeq was 0.05. Transcripts per million (TPM) values were used for calculations throughout this study because this measure is preferred over RPKM (Mortazavi et al., 2008) and FPKM (Trapnell et al., 2010) measures due to its lack of dependence on the mean expressed transcript length and thus its ability to be compared across samples (Li et al., 2010).

EBSeq Run	Samples	Condition String (Same Order as in 'Samples')	Size Factors (Same Order)
1	1 and 4	"Control,Psg1"	1, 1
2	4 and 6	"Psg1,Psg5"	0.973, 1.028
3	1 and 6	"Control,Psg5"	0.960, 1.042

For heat map analysis we selected the top 5% significantly upregulated genes and performed gene ontology analysis using STRING database. 90-95% of the cardiac genes depicted represent the 'cardiovascular system development' GO category. The remaining were chosen from published cardiac reprogramming studies as well as genes known to be important in cardiac development. For fibroblast genes, we chose genes that were highly expressed in cardiac

fibroblasts and were downregulated in mESC-CPCs. Pluripotency, epiblast, non-cardiac mesoderm, endoderm and ectoderm genes were selected based on various publications characterizing the respective cell populations. Pearson sample correlation analysis was carried out on all genes' TPM values in R version 3.0.2 using the cor() function with method "pearson".

Mouse MI Model and iCPC Injections

Following induction of isoflurane anesthesia (3%), the mouse was intubated with an 18-gauge catheter and placed on a mouse ventilator at 120-130 breaths per minute with a stroke volume of 150 microliters and maintained on 2% isoflurane. A left lateral incision through the fourth intercostal space was made to expose the heart. After visualizing the left coronary artery, 7-0 prolene suture was placed through the myocardium in the anterolateral wall and secured. Coronary artery entrapment was confirmed by observing blanching of the distal circulation (ventricular apex). The lungs were over inflated and the ribs and muscle layers were closed by absorbable sutures. The skin was closed by additional suturing using 6-0 nylon or silk. The mouse was recovered from the anesthesia and extubated. Two days after the MI procedure the chest was reopened and 1-1.5 million GFP labeled iCPCs (reprogrammed using AC Fibs and 5 Factors) were injected per animals around the MI border zone (3 injections of 25 ul each, total of 75 uls). Control animals were injected with PBS. iCPC injected and control animals were recovered from surgery and monitored for survival for 4 weeks.

EXTENDED REFERENCES

Champlin, A.K., Dorr, D.L., and Gates, A.H. (1973). Determining the stage of the estrous cycle in the mouse by the appearance of the vagina. *Biology of reproduction* 8, 491-494.

Conklin, M.W., Provenzano, P.P., Eliceiri, K.W., Sullivan, R., and Keely, P.J. (2009). Fluorescence lifetime imaging of endogenous fluorophores in histopathology sections reveals differences between normal and tumor epithelium in carcinoma in situ of the breast. *Cell biochemistry and biophysics* 53, 145-157.

Downs, K.M. (2006). In vitro methods for studying vascularization of the murine allantois and allantoic union with the chorion. *Methods in molecular medicine* 121, 241-272.

Downs, K.M., and Davies, T. (1993). Staging of gastrulating mouse embryos by morphological landmarks in the dissecting microscope. *Development* 118, 1255-1266.

Downs, K.M., and Harmann, C. (1997). Developmental potency of the murine allantois. *Development* 124, 2769-2780.

Leng, N., Dawson, J.A., Thomson, J.A., Ruotti, V., Rissman, A.I., Smits, B.M., Haag, J.D., Gould, M.N., Stewart, R.M., and Kendziorski, C. (2013). EBSeq: an empirical Bayes hierarchical model for inference in RNA-seq experiments. *Bioinformatics* 29, 1035-1043.

Li, B., Ruotti, V., Stewart, R.M., Thomson, J.A., and Dewey, C.N. (2010). RNA-Seq gene expression estimation with read mapping uncertainty. *Bioinformatics* 26, 493-500.

Masino, A.M., Gallardo, T.D., Wilcox, C.A., Olson, E.N., Williams, R.S., and Garry, D.J. (2004). Transcriptional regulation of cardiac progenitor cell populations. *Circulation research* 95, 389-397.

Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature methods* 5, 621-628.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature biotechnology* 28, 511-515.