SUPPLEMENTAL APPENDIX

Supplemental Methods

Human cardiac fibroblast (cFib) isolation and culture

Human left ventricle (LV) tissue samples from male 17yo, male 23yo and a female 27yo non-failing donors, were provided by Sean Lal and Cris dos Remedios at the Sydney Heart Bank, University of Sydney. The collection of human tissue from human heart transplant patients and donor hearts was approved by the University of Sydney Human Research and Ethics Committee by St. Vincent's Hospital and by the Australian Red Cross Blood Service. Heart tissue was collected only when the donor was declared brain-dead by the transplant coordinator. All donors were de-identified and their associated clinical data were stored securely. Cardiac tissue samples were then fragmented into approximate $5mm^2$ pieces in liquid nitrogen. Fragmented samples were washed in pre-cooled PBS, and further minced into $1mm^2$ samples on ice. Tissue pieces were then plated onto 6-well plates that were pre-coated with 1% gelatin for 1 hr at room temperature, cultured in minimal essential medium eagle alpha modification (MEMα) (#12561-056, Gibco) complete medium (supplemented with 20% FBS (#S-FBS-AU-015, Serana), 4mM L-glutamine (#G7513, Sigma), 100U Penicillin/Streptomycin (#10378016, Life Technologies) and incubated at 37°C, 5% CO_2. Explants were cultured during this time without passage. First outgrowth of small fibroblast-like cells was harvested with 2X Trypsin-EDTA (Sigma #59418C), diluted in PBS, and discarded due to residual cellular debris. Second and third outgrowths of fibroblasts were harvested and seeded into T25 flasks for subculture in complete MEMα culture medium. Cells were

cultured in continuous monolayer culture, and passaged at 70%-80% confluence using trypsin-EDTA for 5 min at 37°C. Equal volume of complete medium was added to flasks to deactivate trypsin. Cells were pelleted at 250g for 5 min at room temperature to remove trypsin.

Human cardiac stromal cell (cSC) isolation and culture

cSCs were isolated by explant culture and purified by cell sorting as previously described.(1) Briefly, cFib were cultured for ~3 weeks, changing the culture media every 2 days. cFib were dissociated and stained with anti-PDGFRα-APC (1:10, #FAB1264A, R&D), anti-CD31-PE (1:20, #12-0319, eBioscience) and anti-CD90-FITC (1:20, #328108, BioLegend). The cells were then sorted for PDGFRα+/CD90+/$CD31^-$ fraction using Influx machine (BD Biosciences), as described previously.(1) cSC were then cultured in MEMα complete medium, 4mM L-glutamine, 100U Penicillin/Streptomycin and incubated at 37°C, 5% CO_2.

Fibroblast-myofibroblast differentiation assay

For downstream RNA or protein extraction, cFib or cSCs were seeded into T150 flasks at a density of 1.31 x 10^4 cells/cm^2 in complete MEMα culture medium for 12 hr. For immunocytochemistry (ICC) 2.5 x 10^4 cFib or cSC were seeded onto 12mm glass coverslips laid into 24 well plates in complete MEMα culture medium for 12 hr. Cells were then washed in PBS and cultured in serum starved (SS, 0% FCS) MEMα culture medium supplemented with 4mM L-glutamine and 100U Penicillin/Streptomycin. At 24 hr, media was replaced with SS medium with or without transforming growth factor-beta (TGF-β, 10ng/mL Biolegend,

#580704), PDGF-AA (10ng/mL, #100-13A, Peprotech), PDGF-AB (10ng/mL, #100-00AB, Peprotech) or PDGF-BB (10ng/mL, #100-14B, Peprotech). Media was replaced every 24 hr for 72 hr before cells were either collected for RNA extraction, protein extraction or fixed in 10% formaldehyde for 10 mins for subsequent immunocytochemistry (ICC).

Human cardiac fibroblast/stromal cell (cFib/cSC) immunocytochemistry (ICC)

All steps were carried out at room temperature unless otherwise stated. Fixed cFib or cSC from fibroblast-myofibroblast differentiation assays were washed in PBS 3 times for 5 min, permeabilised in 0.1% triton-X100 (#BP151-100, Fisher) in PBS for 10 minutes, washed in PBS 3 times for 5 min, blocked in 0.01%triton-X100 plus 10% normal goat serum (Sigma, #G9023) in PBS for 1 hour and incubated in either mouse anti-αSMA (αSMA, 1:500, #M0851, Dako), mouse anti-αSMA plus rabbit anti-vimentin (1:100, #5741S, Cell Signalling) or rabbit anti-Ki67 (1μg/mL, #ab15580, Abcam) primary antibodies overnight at 4°C. Following this, coverslips were washed in PBS 3 times for 5 mins and incubated in goat anti-mouse Alexafluor488 (1:500, #A11029, Invitrogen) and goat anti-rabbit Alexafluor594 (1:500, #A11037, Invitrogen) for 1 hour at room temperature. Coverslips were then washed in PBS 3 times for 5 min before mounting on slides in 50:50 PBS:glycerol. Slides were imaged on an Olympus VS120 slide scanner, analysed using the Binary Reconstruct FIJI plugin and plotted using GraphPad Prism 8.

Western blotting

cFib from fibroblast-myofibroblast differentiation assays were detached with 2x Trypsin-EDTA for 5 mins and using a cell-scraper to scrape remaining cells off the petri dish. Trypsin-EDTA was deactivated by adding double volume complete medium to trypsin-EDTA. Cell suspension was transferred to a 50mL falcon tube and centrifuged at 250 x g for 5 mins at room temperature to pellet cells and remove trypsin. Supernatant was discarded and dry cell pellets were frozen at -80°C for storage. Cell pellets were then treated with RIPA buffer (#R0287, Sigma) whilst petri dishes are kept on ice, and a cell scraper is used to remove and lyse remaining cells. Lysate was transferred to Eppendorf tubes, and then stored at -80°C. Protein lysate obtained above were tested for protein concentrations using a bicinchoninic acid (BCA) assay (#23225, Thermo Scientific) according to the manufacturer's instructions. Western blotting was carried out as described (Mahmood and Yang, 2012). Briefly, 20µg of protein was added to Laemmli sample buffer (#1610747, Bio-Rad) with β-mercaptoethanol (#0482-100ML, VMR) and proteins were denatured at 95°C for 5 minutes prior to SDS-PAGE. Proteins were separated using SDS-PAGE in 25 mM Tris-Base, 192 mM glycine, 0.1% SDS in Milli-Q water at 100V for 1 hr and transferred to a nitrocellulose membrane via a wet transfer in 25 mM Tris-Base, 192 mM glycine, 20% (v/v) methanol in Milli-Q water at 100V for 90 mins. Membranes were washed 4 x 5 mins in PBST, blocked in 5% skim milk-PBST for 1 hr at RT followed by overnight incubation on a rocker at 4°C with mouse anti-αSMA antibody (1:5000, #M0851, Dako) in 5% skim milk-PBST. Membranes were then washed 4 times for 5 mins in PBST and then incubated in anti-mouse horse radish peroxidase (HRP, 1:10000, #31430, ThermoFisher) in 5% skimmed milk/PBST for 1 hour at RT. Membranes were then washed 4 times for 5 mins, developed using SuperSignal West Femto Maximum Sensitivity Substrate

(#34095, ThermoFisher) and imaged on the Bio-Rad ChemiDoc imaging platform. Blots were then stripped (#ab282569, Abcam) and reprobed for GAPDH (1:5000, #PA1-987, ThermoFisher). Densitometry analysis was performed using FIJI with αSMA protein expression normalised to GAPDH.

Quantitative polymerase chain reaction (qPCR)

cDNA was synthesised using 1μg RNA and 250ng/μL random primers (#C1181, Promega) in dH_2O in a 15μL reaction and incubated for 5min at 70°C followed by incubation for 1min on ice. Then 5x MMLV RT buffer (#M531A, Promega), 10mM dNTP (#N0447S, New England Biolabs), 25 units Rnase Inhibitor (#N2111, Promega), 200 units M-MLV RT (#M1701, Promega) and dH_2O to a total of 10μL/reaction was added to the previous 15μL RNA reaction (total 25μL) and incubated in a thermocycler (MiniAmp Plus Thermal Cycler, Thermo Fisher) for 10 min at 25°C, 60 min at 37°C, 15 min at 70°C and held at 4°C. cDNA was then diluted to 40ng/μL in dH_2O and stored at -20°C. For qPCR, a reaction of SensiFAST SYBR No-ROX Mix (#BIO-86020, Bioline), 10μM forward primer (see below), 10μM reverse primer (see below), 1μL cDNA and dH_2O to a total of 10μL/reaction. Using a Bio-Rad CFX384 qPCR Machine, this reaction was incubated for 2 mins at 95°C, then cycled 40 times for 5 sec at 95°C, 10 sec at 60°C and 15 sec at 72°C before holding at 4°C for collection. Analysis was performed with Bio-Rad CFX Manager 3.1 software. Relative gene expression was calculated using the ΔΔCt method, normalised to GAPDH housekeeping gene.(2) Log base 2-fold change data are presented relative to GAPDH. Analysis was visualised using GraphPad Prism 8.

qPCR Primers

ACTA2 forward - TTCAATGTCCCAGCCATGTA

ACTA2 reverse - GAAGGAATAGCCACGCTCAG

LOX forward - CAGAGGAGAGTGGCTGAGG

LOX reverse - CCAGGTAGCTGGGGTTTACA

Col1a1 forward - GTGCTAAAGGTGCCAATGGT

Col1a1 reverse - CTCCTCGCTTTCCTTCCTCT

Col1a2 forward - CTGCAAGAACAGCATTGCAT

Col1a2 reverse - GGCGTGATGGCTTATTTGTT

TNC forward -TTCACTGGAGCTGACTGTGG

TNC reverse - TAGGGCAGCTCATGTCACTG

GAPDH forward - ACCCACTCCTCCACCTTTG

GAPDH reverse - CTCTTGTGCTCTTGCTGGG

cFib RNAseq

Cells were then detached, pelleted and RNA was then extracted using an RNeasy mini kit (#74104, Qiagen) according to the manufacturer's instructions. Library preparation was carried out using TruSeq stranded mRNA library prep kit and single indexes (#20020594, #20020492, #20020493, Illumina) and carried out according to the manufacturer's instructions. TapeStation gels were run to confirm high quality (>8.0 RINe score) input RNA and for final library quality controls. Illumina sequencing library quality control with Agilent ScreenTape assay and qPCR (#6291, AGRF) followed by RNAseq was carried (Fig S1) out by

the Australian Genome Research Facility (AGRF) on an Illumina NovaSeq 6000 S1 100 cycle lane (#6427, AGRF). FastQC (v0.11.8) was used for quality control of sequenced reads. Samples sequenced on separate lanes were merged before alignment to the human genome GRCh38.p12 from GENCODE using STAR RNA-seq aligner (v2.7.3a). Normalized counts per millions (CPM), obtained using the edgeR package (v3.32.1), were used for principal component analysis with the PCAtools (v2.2.0) R package. Differential expression analysis was performed using DESeq2. Heatmaps were generaterd using DE genes (≤0.05 adjusted p-value, log_2fold change of either ≤-0.5 or ≥0.5) and the online Morpheus heatmap tool (Morpheus, https://software.broadinstitute.org/morpheus) with hierarchical clustering using one-minus Pearson method. Volcano plots of DE genes were generated using the EnhancedVolcano RStudio package. Venn diagrams of overlapping DE genes were generated using the ggVennDiagram RStudio package. For Reactome pathway analyses, DE genes in PDGF-AB versus SS control samples were inputted into the online Reactome or Reactfoam pathway analysis software (https://reactome.org/). Pathway tree analyses were carried out with the online (integrated Differential Expression and Pathway analysis (iDEP) software(3) (http://bioinformatics.sdstate.edu/idep/) using the GAGE method, a minimum gene set size of 15 and a pathway significance cutoff (FDR) of 0.2 to display gene ontology (GO) biological process gene sets. Pathway analysis was carried out using online the Database for Annotation, Visualization and Integrated Discovery (DAVID) tool (https://david.ncifcrf.gov/tools.jsp) and visualised using the RStudio package ggplot2. Gene Set Enrichment Analysis (GSEA) was carried out using the Broad Institute GSEA 4.1.0 software with a preranked gene list and 1000 permutations.

Transwell migration assays

1.5 x 10^4 cells were seeded onto the Boyden chambers (#3422, Costar) in 100µl of serum starved MEMα culture medium. 600µl of serum starved MEMα culture medium was added to the base of each well and cells were treated with the following treatment conditions, Serum Starved only, PDGF-AB (10ng/ml), TGF-β (10ng/ml), PDGF-AB (10ng/ml) + TGF-β(10ng/ml). Cells were cultured for 24 hours, media aspirated and membranes are PBS washed, then fixed in cold 70% ethanol for 30 minutes. Membranes were then PBS washed incubated in 1µg/ml DAPI-PBS for 10 minutes then PBS washed. Membrane seeding surfaces were then scraped to remove remaining non-migrated cells, dissected from transwells and mounted on microscope slides using ProLong Gold Antifade Mountant (#P10144, ThermoFisher).

Live cell imaging

cFib from fibroblast-myofibroblast differentiation assays were imaged every 20 minutes for 24 hr on a Leica Live Cell Imaging Microscope (Leica, #DMI6000B).

Mouse myocardial infarction surgery

All animal experiments were approved by the St Vincent's Hospital and Garvan Institute of Medical Research Animal Ethics Committee. Mice were bred and housed in the BioCORE facility of the Victor Chang Cardiac Institute. Rooms were temperature and light/dark cycle

controlled. Standard food was provided *ad libitum*. Soggy and/or high nutrient food was provided after surgery. Male mice aged 8-12 weeks (young) were used.

Mice were anaesthetised by intraperitoneal injection of ketamine (75mg/kg) and xylazine (20mg/kg), and intratracheally intubated. PDGF-AB ligand (R&D, 20µg/ml) or vehicle (PBS) was administered systemically by a subcutaneous minipump (Alzet), inserted to the right external jugular vein via 13mm of 28G catheter (Alzet) through a 1mm incision. The ligand/vehicle was continuously released at a concentration of 0.5µl/hr for 5 days. The hearts were exposed via a left intercostal incision, the left anterior descending coronary arteries were ligated just before the first diagonal branching point. The wounds suture closed and lungs inflated. Sham operations were performed without the ligation of the coronary artery. The mice were given intra-muscular injection of Buprenorphine for twice a day for 3 days. 5 days post-MI mice were euthanised via injection of 1M KCL (#P933, Sigma) and hearts were perfused with 25ml PBS. Hearts were sliced in half and embedded in optimal cutting temperature compound (OCT, #4583, Tissue-Tek) in plastic moulds (#4565, Tissue-Tek). Hearts were then immediately frozen in a freezing pan containing isopentane LN_2-cooled to the point of near freezing. Samples were then stored at -80°C until cryosectioning.

Mouse imaging mass cytometry (IMC) sample preparation

Two consecutive 7µm cryosections were taken from each OCT embedded sample using a Microm HM 505 E Cryostat, collected onto Superfrost plus gold slides (#K5800AMNZ72, ThermoScientific), air dried for 10 mins and fixed in pre-cooled (-20°C) acetone (#179973, Sigma-Aldrich) for 10 mins. Acetone was then poured off and allowed to evaporate until slides

were completely dry. Circles were then drawn around the samples with a PAP pen (#PPM, Hurst Scientific) and 75μL of 10% NGS-0.01% Triton-X-PBS (blocking buffer) was added to each sample for 1 hr. Following this 1:100 anti-αSMA-FITC (#F3777, Sigma) and 1:300 anti-col1α1 (#91144S, Cell Signalling) conjugated to Cy5 using a sulfo-Cyanine5 antibody labelling kit (#3321-10rxn, Lumiprobe) according to the manufacturer's instructions, hereafter referred to as anti-col1α1-Cy5, was diluted in blocking buffer and added to each sample for 1 hr. 1 of 2 cryosections was then washed in PBS 3 times for 5 mins. 1μg/mL DAPI was then added to this section for 10 mins, followed by washing in PBS 3 times for 5 mins and mounting in PBS:glycerol (1:1). The entire IHC section was then imaged on an Olympus VS120 slide scanner to identify the infarct zone (IZ), border zone (BZ) and remote zone (RZ), enabling regions of interest (ROI) selection on the IMC at a later timepoint.

IMC data acquisition

IMC metal conjugated antibodies were conjugated using the Maxpar® X8 Multimetal Labeling Kit (#201300, Fluidigm) kit according to the manufacturer's instructions. IMC metal conjugated antibodies consisted of: vimentin-113 (clone D21H3), Ki67-115 (clone B56), Ly6G-141 (clone 1A8), CD11c-142 (clone N418), CD4-143 (clone RM4-5), FITC-144 (clone FIT-22), CD45-147 (clone 30-F11), CD11b-148 (clone M1/70), fibronectin-149 (clone TV-1), Ly6C-150 (clone HK1.4), CD25-151 (clone 3C7), CD3e-152 (clone 145-2C11), phospho-Akt-154 (clone J1.233.371), CD31-155 (clone MEC13.1), CD274-156 (clone B7-H1), FoxP3-158 (clone FJK-16s), Cx43-159 (polyclonal), T-bet-160 (clone 4B10), FXIIIa-161 (polyclonal), TER119-162 (clone TER-119), F4/80-163 (clone BM8), cyclin-B1-164 (clone GNS-1), NaKATPase-165 (clone EP1845Y), IL6-167 (clone MP5-20F3), CD8-168 (clone 53-6.7), Ly6B.2-169 (clone 7/4), NK1.1-

170 (clone PK136), CD44-171 (clone IM7), Cy5-172 (Cy5-15), CD115-173 (clone AFS98), IA/IE-174 (clone M5/114.15.2) and αActinin-176 (clone EP2529Y). Following anti-αSMA-FITC and anti-col1α1-Cy5 incubation, the second cryosection was washed in PBS 3 times for 5 mins. 75μL of IMC metal-conjugated antibodies diluted in blocking buffer was then added and incubated at 4°C overnight. Samples were then washed in PBS 3 times for 5 mins, 4% paraformaldehyde was added, incubated for 20 mins at RT and again washed in PBS 3 times for 5 mins. Iridium DNA intercalator diluted in PBS (1:300 for heart sections and 1:250 for spleen sections) was added and incubated for 30 mins at RT. Samples were then washed in PBS 3 times for 5 mins then dipped in ultra-pure H_2O 3 times, air dried, then stored in an air tight container before imaging on an IMC with a Hyperion Imaging System (Fluidigm). For data acquisition, the IMC was initially tuned ensuring the resolution (mass 1) >400, transients cross talk (1) <0.15, transients cross talk (2) <0.05, mean duals for 175Lu per laser shot >700. Using a combination of a panoramic bright field image taken on the Hyperion with the previously mentioned IHC image on the adjacent section was used for ROI selection of IZ, BZ and RZ. An energy test was then performed, determining the optimal ablation energy to be 1 dB. ROIs were then ablated and data acquired for downstream analyses. For individual channel visualisation the Histocat++ software was used and channels were pseudo-coloured blue, red and green.

IMC multicut cell segmentation

For multicut cell segmentation, individual channels from ROIs within MCD files were converted to a tif file format using the Histocat++ software. Raw tif files were then processed using the RStudio package Spectre and converted into HDF5 files. Ilastik machine learning

software used HDF5 files for pixel classification of CM cell borders and nuclei, non-CM cell borders and nuclei and background areas to create probability files. Next boundary-based segmentation with multicut was performed on Ilastik using ROI HDF5 files and their associated probability files using a 0.5 pre-smooth before seeds threshold for water shedding. Training and multicut was then performed on every ROI in the dataset to produce associated multicut segmentation files. Next object classification was performed on Ilastik to identify cell and non-cell objects using ROI and multicut segmentation HDF5 files. Both object predictions and object identities were exported as tif files for each ROI. Ilastik pixel classification was again performed with ROI HDF5 files for region classification of scar and non-scar (myocardium) areas and exported as simple segmentation tif files.

IMC cellular analysis

All cellular analysis was done using the RStudio package Spectre unless otherwise stated. First, tif files were read into Spectre and spatial objects were created. Next cell mask (object identities), cell type (object predicitons) and region (simple segmentation tif files from Ilastik multicut cell segmentation were read into Spectre followed by polygons and outlines being generated. Initial mask quality control plots were generated to ensure that each mask for each ROI was suitable. Single cell expression plots were then generated for each ROI for each marker, the data extracted for each cell, annotated as 'cell' or 'non-cell' to identify segmented areas without cells and the area of different regions calculated. Data was then saved as qs, csv and fcs files. For cellular analysis previously generated qs and csv files were read into Spectre on RStudio. An inverse hyperbolic sine function (ArcSinh) was then applied to expression data before rescaling. Non-cells, background and any objects $<5\mu m^2$ were filtered

out of analysis. Then clustering and dimensionality reduction were performed using 20 metaclusters and a perplexity of 200. tSNE plots and an expression heatmap were then plotted. Clusters were then annotated according to the heatmap expression levels for each cluster. Data was then saved in csv and fcs file formats. At this point fcs files were taken into FlowJo analysis software and gated to compare treatment groups. Data was then plotted using GraphPad Prism 8.

IMC spatial analysis

All csv and qs spatial data files were read into Spectre followed by running spatial analysis and clustered heatmap generation. Saved output csv files were then used to compare treatment groups with one another. These data were then used to generate volcano plots using the EnhancedVolcano RStudio package to display all statistically significant ($p<0.05$) spatial comparisons that also met a threshold of log_2fold change of <-0.5 or >0.5. Some of these selected spatial comparisons were then plotted on GraphPad Prism 8.

Pig myocardial infarction surgery

All experiments were conducted in accordance with local guidelines and regulations, and all study protocols were approved by the Animal Ethics Committee of the Western Sydney Local Health District. Authors complied with the ARRIVE guidelines for conducting animal research.(4) Experiments were conducted using adult, female Landrace swine which were housed individually and receiving standard care. All animals were 2 months old and were acclimatised for ≥1 week before commencing experimental procedures. All procedures

below are for 11 day pig studies, as 28 day histology samples were taken from our previous study.(5)

After 12 hours fasting, animals were premedicated with intramuscular ketamine (10 mg/kg), methadone (0.3 mg/kg), and midazolam (0.3 mg/kg). Animals were intubated and mechanically ventilated (Anestar-S, Datascope) at 10 ml/kg tidal volume with supplemental oxygen to maintain an end-tidal CO_2 of 35 to 45 mmHg. General anesthesia was induced with intravenous Propofol (2 to 5 mg/kg) and maintained with 2% inhaled isoflurane. Anaesthetic depth was monitored by cessation of movement, eye position, loss of muscular tone, and absence of palpebral and pedal reflexes. Animals were instrumented with a peripheral 24-gauge intravenous catheter in an ear vein, surface electrocardiography electrodes, a right femoral artery (6F) vascular access sheaths as well as left femoral artery (5F) access for invasive hemodynamic monitoring. All vascular access was obtained under ultrasound guidance. 5000 units of unfractionated heparin was given after vascular access was established followed by maintenance boluses to maintain an activated clotting time (ACT) between 250 and 350 seconds. A bolus of IV Amiodarone 150mg was given as a slow bolus at this time. This was followed by a bolus of Lignocaine (1mg/kg) and continued as an infusion. The left coronary artery was engaged percutaneously via the right femoral artery using a 6F Hockeystick guiding catheter (Medtronic). A 0.36-mm coronary guidewire (Asahi Sion Blue) was delivered into the LAD. Myocardial infarction was induced by inflation with a 2.5-mm angioplasty balloon (Boston Scientific) for 90 min distal to the first major diagonal. Cessation of flow distal to the occlusion was confirmed angiographically. Occurrence of ventricular arrhythmias during this period were managed with IV Metoprolol and DC cardioversion as needed. Hemodynamic changes were managed with IV Metaraminol to maintain a mean

arterial pressure above 55 mmHg. Coronary angiography was performed after reperfusion to confirm vessel patency. PDGF-AB assigned animals went on to have a mini-pump insertion prior to proceeding to recovery. Control animals proceeded to recovery.

Pig minipump insertion surgery

A line was drawn between the tip of the manubrium and the angle of the right Ramus of mandible, about 5-6cm long. A 50/50 mix of bupivacaine and lignocaine was injected through the skin and into the deeper tissues along the length of this incision. An incision was made through the skin and subcutaneous tissue followed by blunt dissection through the underlying muscle in the direction of the muscle fibres. Haemostasis was strictly maintained. The right internal jugular vein was dissected from the sheath. A pocket was created, using blunt dissection, for the mini pump cranial to the vein to allow for it to sit flush with the vein. A 5-0 monofilament was used as a purse string to aid control of the vein. A 16 gauge intravenous cannula was passed into the vein to enable passage of the minipump tubing. A basket suture was tied around the minipump with 3-0 non-absorbable monofilament and the minipump was attached to the tubing (5mm overlap). The minipump was placed beside the vein and tied in place to the medial muscle sheath. Muscle layer was closed followed by closure of the subcutaneous later along the intradermal suture line. The skin was sutured and a dressing placed in peri-operative period to keep the wound clean during recovery. rhPDGF-AB (#100-00AB, Peprotech) or vehicle was delivered at a fixed infusion rate of 65 μg/kg for 7 days via a 2-ml ALZET (Durect) osmotic mini-pump inserted into the right external jugular vein.

Cardiac magnetic resonance imaging (cMRI)

After 12 hours fasting, 2 days or 9 days post-MI, animals were premedicated, intubated, ventilated, anesthetized and instrumented as described for infarct surgery.

All CMR examinations were performed on a Siemens 3T Prisma (Siemens Medical systems, Erlangen, Germany) utilising an 18 channel body array anteriorly together with posterior spine array coils and 4 lead ECG gating. All sequences were performed with breath-holding technique in end-expiration with one signal average. Axial and coronal TrueFISP (true fast imaging with steady state free precession) and axial HASTE (half Fourier Acquired-Single Shot spin echo) sequences through the chest were acquired to plan preliminary 2-chamber, 4-chamber and short axis stack (SAX) single slice images. True 4-chamber, 2-chamber and 3-chamber single slice TrueFISP cines were then acquired, followed by 16 contiguous 8mm SAX cine slices, starting just below the apex and extending into the pulmonary artery, planned from the 2 and 4 chamber cine images in end-diastole. TrueFISP cines were acquired with the following parameters - TR: R-R interval of individual animal, TE: 1.3ms, FOV: 320mm, slice thickness: 8mm, in plane resolution: 1.4mm x 1.4mm, flip angle: 70 degrees, retrospective reconstruction with 25 calculated phases, scan time around 10 seconds. Late Gadolinium Enhancement (LGE) or Delayed enhancement (DE) images were acquired for tissue characterisation using a segmented inversion recovery fast gradient echo sequence in SAX, LVLA and 4-chamber planes. TI (time of inversion) scout was performed at 9 minutes post contrast to select appropriate inversion time (usually around 300ms) for DE acquisitions which began at around 11 minutes post-injection. The SAX was run with 30 contiguous slices of 4mm thickness, covering the entire heart with in-plane resolution of 1.5mm x 1.5mm

whereas for the 2-ch and 4-ch sequences, single 8mm thick slices were acquired with in-plane resolution of 1.4mm x 1.4mm. Other DE parameters were TR: R-R interval of individual animal, TE: 1.65ms, flip angle: 20 degrees. Assessment of volumes was done according to Society for Cardiovascular Magentic Resonance guidelines.(6) Manual endocardial and epicardial contouring was performed on the short axis stacks. Left ventricle (LV) end-diastolic image was chosen as frame 1. LV end-systolic image was chosen based on the smallest blood pool within the LV cavity. LV outflow tract blood volume was included in the volume calculation. Papillary muscles were excluded from mass calculations.

Pig cardiac tissue collection

Animals were euthanized with 75-150mg/kg of IV potassium chloride. The heart was washed in normal saline before being sectioned. Sectioning was performed from the apex to the base at 5mm intervals. The first section was termed “Apex” and each subsequent section was termed “L1” onwards. “L2” and “L4” were immediately taken for snap freezing (Fig S2). The remaining sections were fixed in 10% neutral-buffered formalin. Fixation was performed for 48 hours with formalin being changed at 24 hours. Samples were subsequently stored in 70% ethanol. Samples to be frozen were cut along the width of the LV wall. Samples were taken from the infarct zone (IZ), border zone (BZ) and remote zone (RZ) as well as the right ventricle (RV).

Pig RNAseq sample preparation, library preparation and sequencing run

Frozen pig cardiac tissue from level 1 of the IZ, BZ, RZ or RV was placed in a liquid nitrogen (LN_2) cooled metal plate and pulverised by hitting a LN_2-cooled rod with a hammer onto the tissue. Both the plate and rod were made in-house. Tissue was then further minced using a scalpel before collecting the powder into a LN_2-cooled 1.5mL tube and homogenised for 5 seconds using a benchtop vortex. 1g of tissue powder was then taken for RNA extraction using an RNeasy Maxi Kit (Quiagen, #75162) according to the manufacturer's instructions and with a 50mL syringe to homogenise. RNA quality, library preparation and quality controls were carried the same as cFib RNAseq (Fig S3). RNAseq was carried out by AGRF on an Illumina NovaSeq S4 Lane, 300 cycle (#NGS-SEQ-011, AGRF).

RNAseq was also performed separately on *in vitro* pig cardiac fibroblast (pFib) samples to improve genome coverage for *de novo* library assembly. pFib (P6049, Cell Biologics) were expanded until passage 2 allowing their differentiation in vitro. RNA extraction was performed using RNeasy mini-Kit (Qiagen, 74106). Common myofibroblast differentiation markers were tested before sending RNA for sequencing. Libraries were prepared from 500 ng of total RNA with the Illumina TruSeq Stranded mRNA reagents using a unique dual indexing strategy, and following the official protocol automated on a Sciclone liquid handling robot (PerkinElmer). Libraries were quantified by a fluorometric method (QubIT, Life Technologies) and their quality assessed on a Fragment Analyzer (Agilent Technologies). Cluster generation was performed with 2nM of an equimolar pool from the resulting libraries using the Illumina HiSeq 3000/4000 PE Cluster kit reagents and four samples (no differentiated and differentiated pFib in duplicates) were sequenced on the Illumina HiSeq

4000 using HiSeq 3000/4000 SBS Kit reagents for 300 cycles (paired end), around 150 million reads/sample.

Pig RNAseq QC and mapping of sequencing reads

For post-MI pig cardiac tissue (40 samples in 2 sequencing lanes) Illumina adapter sequences were removed using Trimmomatic v0.39(7) and reads were mapped to SusScrofa 11.1 reference genome using Hisat2,(8) with –dta and –summary-file options. Only reads with MAPQ (mapping quality score) of at least 2 were kept for the analysis.

For *in vitro* pFib (4 samples) read quality was individually assessed for in vitro fibroblast samples using FastQC. No significant issues in relation to read quality were identified for all in vitro samples. Reads were mapped to SusScrofa 11.1 RefSeq genome using STAR aligner,(9) with options to specify BAM as the output format (--outSAMtype BAM SortedByCoordinate) and the strand information for spliced alignments for compatibility with Cuffmerge (--outSAMstrandField intronMotif).(10) Only uniquely mapped reads were kept for the analysis.

Pig RNAseq *de novo* transcriptome assembly

Transcriptomes were assembled individually for all *in vivo* and *in vitro* pig samples using StringTie denovo mode (i.e. without specifying -G option).(11) For each dataset (i.e. in vivo or in vitro), we merged individual transcriptomes into a single transcriptome using Cuffmerge, with default options. This was followed by merging in vivo and in vitro transcriptomes into a

single final non-redundant transcriptome using SusScrofa 11.1 genome as the reference (specified in both -s and -g options).

Pig RNAseq classification of assembled genes

We classified assembled genes into 3 different functional categories; 1. Known protein-coding genes, 2. Known non-coding genes, or 3. Novel genes. By known genes, we mean genes that have at least one exon overlapping with exons of either NCBI RefSeq or GenCode (v32) (https://www.gencodegenes.org/human/). With genes with multiple transcripts, genes were classified as known if any exons of the transcripts overlap with corresponding annotations in the reference. For known genes, RefSeq annotations NM- and NR- were used for known mRNAs and ncRNAs respectively. For genes that do not overlap with either known genes were classified as novel genes. To be qualified as being overlapped, exons must be on the same strand and have at least one nucleotide overlapped. We used bedtools intersect program for the overlap analysis.(12)

Pig RNAseq Identification of novel lncRNAs

As the novel genes can include both novel protein-coding or non-coding genes, we applied a further classification step. In this step, we characterised identified novel genes based on their 1. gene structure and 2. coding potential. First, genes with single exons or size of less than 200 nucleotides were filtered out. Second, we predicted coding potential of genes using geneID v1.4. GeneID builds Markov chain models to infer sequence-based coding scores derived from multiple species.(13) We used param.default (i.e. human model) as the model

for the prediction, with -s and -G options as recommended setting for non-human mammals. To set up a threshold score that distinguishes non-coding genes from coding counterparts, we identified an intersection of interpolated coding scores from all known mRNA and known ncRNA genes. Intuitively, this intersection value (coding potential=10.75) indicates the coding potential above which predominantly represents PCGs. Accordingly, we removed novel genes with a coding score higher than this intersection value. Non-coding status of more than 99% of the novel lncRNAs were also confirmed by another coding potential tool CNCI.(14) Only novel genes with a coding potential of below this threshold were kept as novel lncRNAs.

Pig RNAseq DE gene analysis

To perform DE gene analysis, we first quantified read counts of all identified genes using htseq-count with --stranded=no (i.e. to count reads regardless of the strand as this information is not available) and -m union (i.e. to include reads for all overlapping genes) options. Genes were identified as DE if the gene was either up- or down-regulated in PDGF-AB samples compared to the control (at least 2-fold change), and this difference was statistically significant (Bonferroni-adjusted $p<0.05$). DESeq2 R library was used for DE analysis.(15)

To obtain biological insight of differentially expressed novel lncRNAs, we linked lncRNAs to nearest known protein-coding genes using Bedops closest-features program,(16) then performed gene ontology (GO) enrichment analysis for these protein coding genes.(17) We used TopGO R library for the GO enrichment analysis using Fisher's exact test (one-sided) between PDGF-AB treated vs control samples.(18)

Differential transcript usage analysis

Salmon(19) was used to quantify transcript abundance. DRIMseq was then used to perform differential transcript usage analysis. First, transcripts were retained for DRIMseq analysis according to the filtering criteria followed. We first define n to be the total number of samples, and n.small to be the sample size of the smallest group. We then used 3 filters for a transcript to be retained in the dataset. (1) It has a count of at least 10 in at least n.small samples, (2) it has a relative abundance proportion of at least 0.1 in at least n.small samples and (3) the total count of the corresponding gene is at least 10 in all n samples. 11415 out off 33867 transcripts were retained. Dirichlet-multinomial model was then fitted to detect genes that showed evidence of differential transcript usage (adjusted P-value <= 0.05).

Histology

Formalin-fixed heart slices (apex, level 1 and level 3) were cut into 4 segments and infiltrated with paraffin wax overnight on a tissue processor. 4µm paraffin sections were cut, dried in an oven at 60°C for 2 hr. Slides were deparaffinized in xylene and rehydrated with gradient ethanol (100%, 100%, 100%, 70%) then water. For picrosirius red with fast green staining, slides were placed in 0.1% Sirius red in picric acid with 0.1% fast green (#SIRIUSFAST500, POCD) for 4 hr and washed in running water. Slides were then dipped in 100% ethanol 5 times, fresh 100% ethanol 10 times, incubated in xylene 6 times for 1 min each and mounted in permanent mounting media (#FNNII065C, Froninecat) before cover-slipping. All slides were then imaged on a slide scanning microscope using a 20X lens (Olympus VS120).

Pig immunohistochemistry (IHC)

For IHC, sections were cut and dried as above, followed by deparaffinization in xylene and rehydrated with gradient ethanol (100%, 100%, 100%, 95%, 70%, 50%) then water followed by subsequent epitope retrieval in pH6.0 5mM Sodium Citrate for 20 min at 95°C using a pressure cooker (Decloaking Chamber TM NxGen, Biocare Medical). Tissues were then cooled in water, PBS, then subsequently permeabilized in 0.1% Triton X-100 (#BP151-100, Fisher Scientific)-PBS for 5 min and placed in PBS again after. Tissues were then blocked in 5% normal goat serum - 5% normal donkey serum - 0.01% Triton X-100-PBS (blocking buffer) for 1 hr[#].

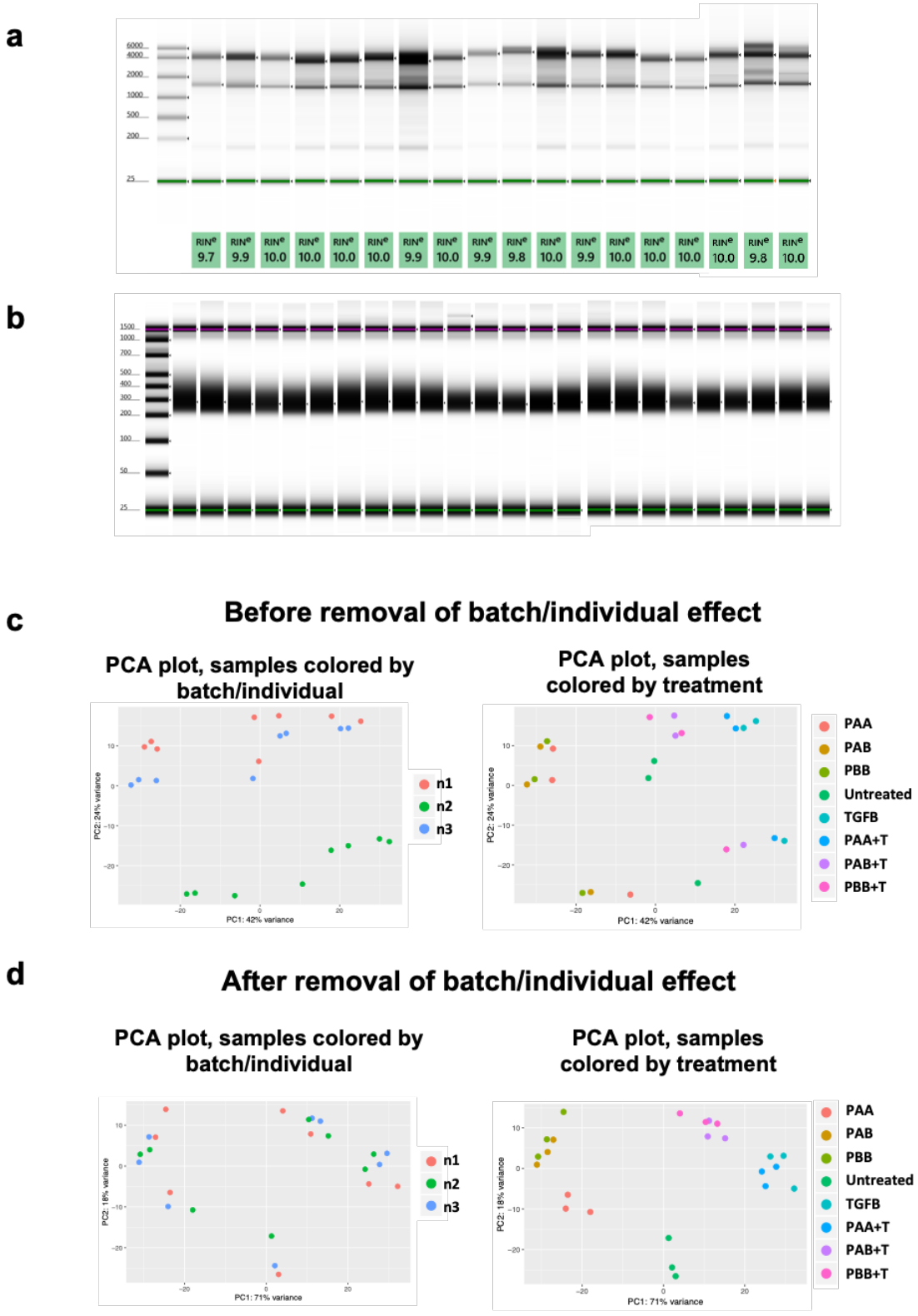
[#]After blocking, for fibroblast/myofibroblast/proliferation, immunostaining was then performed with antibodies directed against chicken anti-vimentin (1:100, #ab24525, Abcam), rabbit anti-Ki67 (5μg/mL, #ab15580, Abcam) and mouse anti-α-smooth muscle actin (αSMA, 1:300, #M0851, Dako) diluted in blocking buffer and incubated overnight at 4°C. Slides were then washed 3 times for 5 min in PBS, followed by adding secondary antibodies goat anti-chicken Alexafluor488 (1:500, #ab150173, Abcam), goat anti-rabbit AlexaFluor555 (1:500, #A21430, Invitrogen) and goat anti-mouse Alexafluor647 (1:500, #A21236, Invitrogen). Slides were then washed 3 times in PBS for 5 min each, followed by the addition of DAPI (1ug/mL) diluted in PBS for 10 min and 3 PBS washes for 5 min each. Sections were then mounted in PBS:glycerol (1:1) and stored at 4°C before imaging on a slide scanning microscope (Olympus VS120) and converted using VS desktop software (Olympus). Images were analysed using FIJI and the BinaryReconstruct plugin.

[#]After blocking, for collagen/immune cell immunostaining, antibodies directed against rabbit anti-CD3 (1:50, #ab5690, Abcam), sheep anti-FXIIIa (1:50, #SAF13A-AP, Affinity Biologicals) and mouse anti-αSMA were diluted in blocking buffer and incubated overnight at 4°C. Slides were then washed 3 times for 5 min in PBS, followed by adding secondary antibodies donkey anti-sheep AlexaFluor488 (1:500, #A11015, Thermofisher), goat anti-rabbit AlexaFluor555 and donkey anti-mouse DyLight755 (1:100, #SA5-10171, Invitrogen) and incubated at room temperature for 1 hr. Slides were then washed 5 times in PBS and conjugated antibody col1α1-Cy5 (1:200) was added for 1 hour. Slides were then washed 3 times in PBS for 5 min each, followed by the addition of DAPI (1ug/mL) diluted in PBS for 10 min and 3 PBS washes for 5 min each. Sections were then mounted in PBS:glycerol (1:1) and stored at 4°C before imaging on a slide scanning microscope (Olympus VS120) and converted using VS desktop software (Olympus). Images were then deconvolved using Huygens Professional software, segmented using CellProfiler pixel expansion and converted into fcs files using RStudio. These files were then analysed on FlowJo to perform imaging cytometry. GraphPad Prism 8 was then used for statistical analyses (one-way ANOVA with a Tukey's multiple comparisons test, *$p<0.05$, **$p<0.01$, ***$p<0.001$).

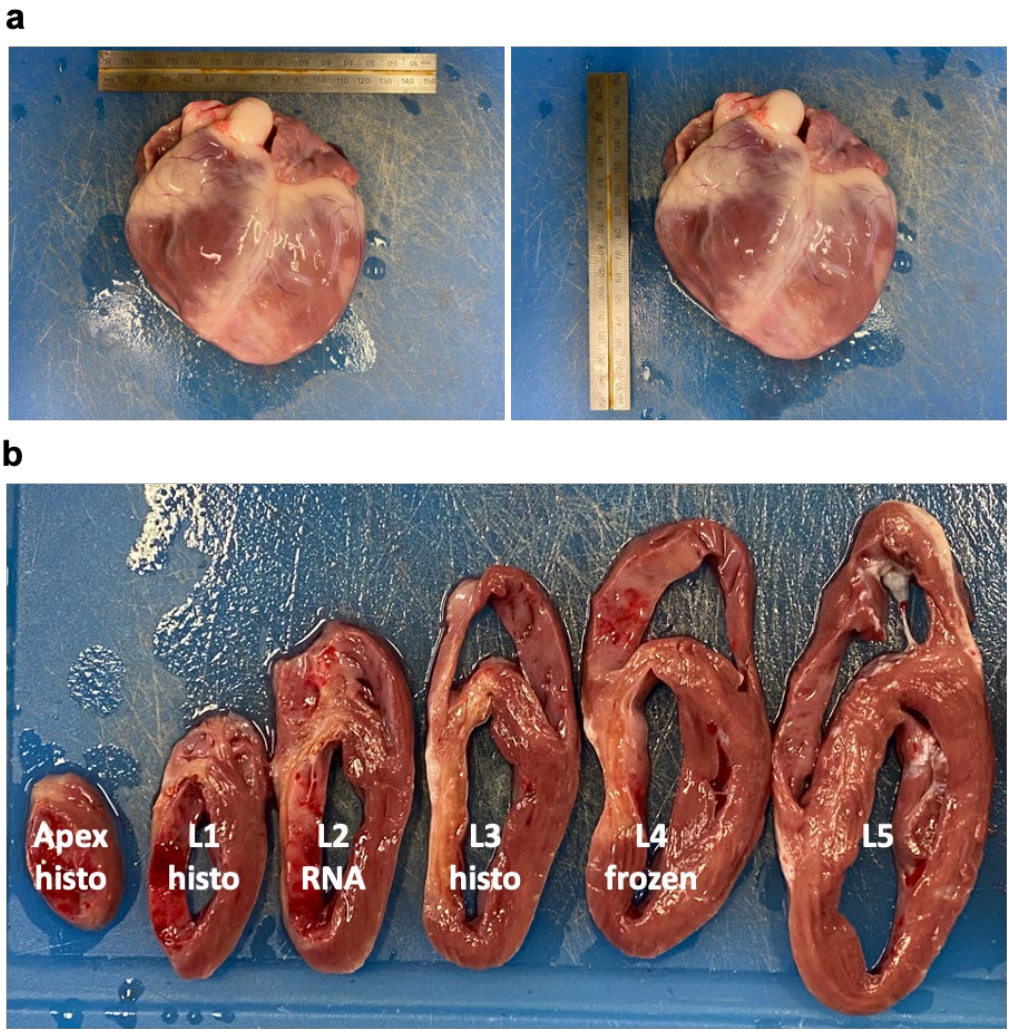
Supplemental Table 1: All significant IMC spatial comparisons for PDGF-AB IZ vs control IZ

Spatial comparison	p-value	log2fc
Dist. Fibroblast 5-Fibroblast 5	0.001	-4.545
Dist. Fibroblast 5-Immune 1	0.001	-3.166
Dist. Cluster 12-Cluster 16	0.005	-0.133
Dist. CM 2-Immune 1	0.012	-2.487
Dist. Fibroblast 5-Fibroblast 3	0.018	-2.367
% of cells Fibroblast 1	0.018	-2.765
Dist. Fibroblast 5-CM 3	0.019	-2.600
Dist. Cluster 16-Cluster 16	0.027	-0.389
Dist. Fibroblast 5-Fibroblast 1	0.031	-2.214
Dist. Fibroblast 5-Cluster 16	0.033	-2.180
Neigh. Fibroblast 3-Fibroblast 1	0.036	-1.398
Neigh. Cluster 14-Fibroblast 1	0.036	-2.757
Dist. Immune 1-Immune 1	0.037	-1.613
Cells per 100 sqr. pixels - Fibroblast 1	0.037	-2.943
Dist. CM 3-Immune 1	0.041	-1.638
Neigh. Fibroblast 1-Fibroblast 1	0.047	-1.429
Neigh. Immune 1-Cluster 12	0.048	-6.645

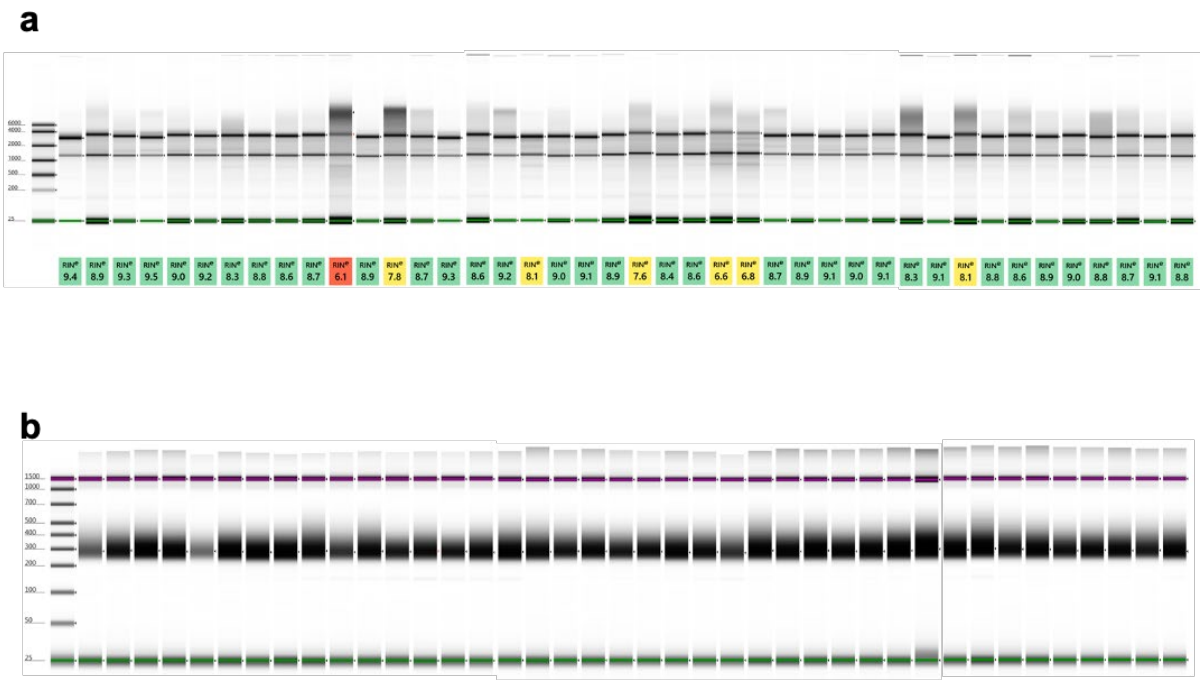
Supplemental Figure 1. Quality controls for human cardiac fibroblast (cFib) RNAseq. a) TapeStation analysis of cFib RNA quality with associated RIN^e scores. b) TapeStation cFib RNAseq library preparation quality control. c,d) Principal component analysis (PCA) of cFib RNAseq before (c) and after (d) batch/individual effects.



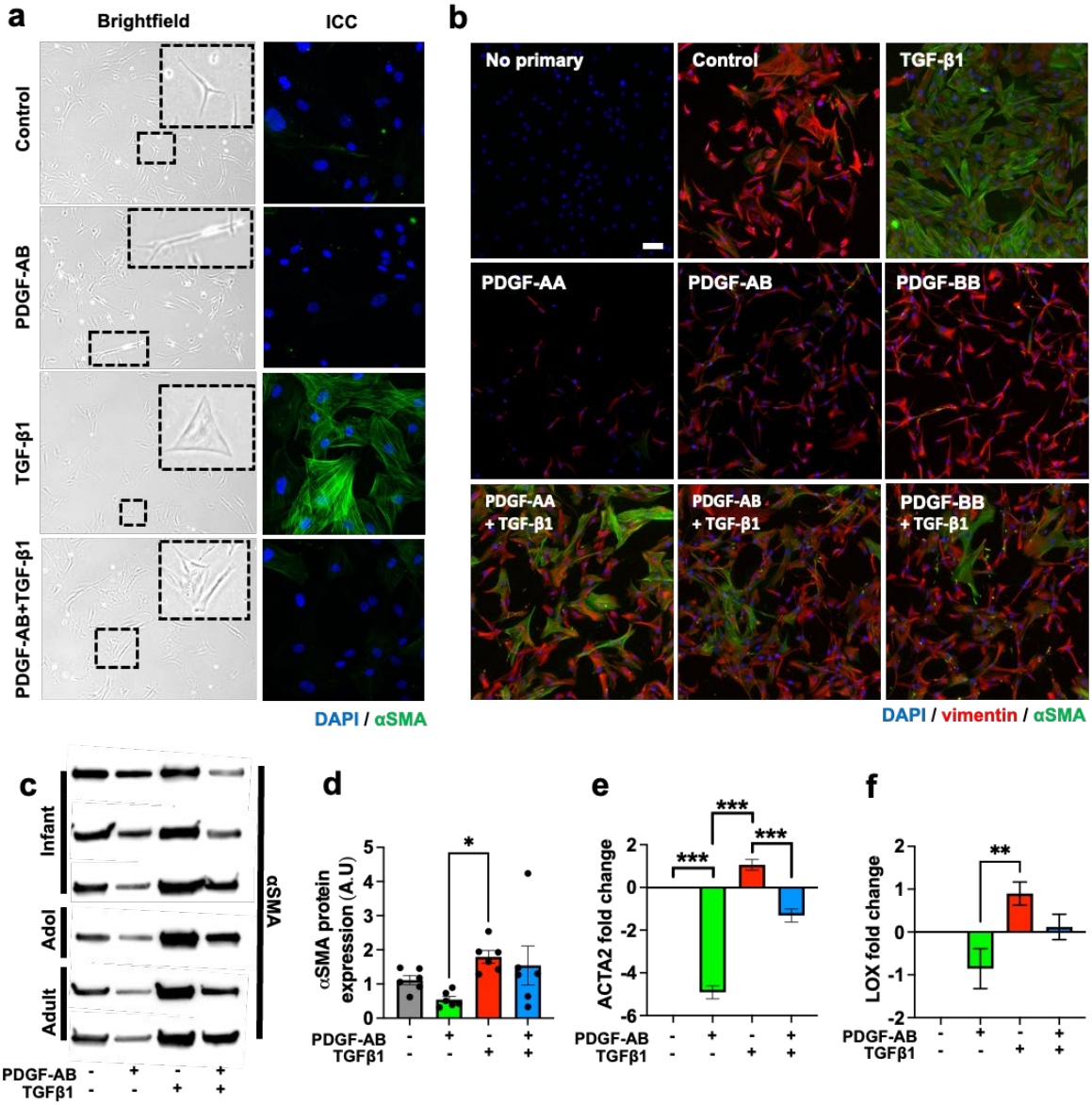
Supplemental Figure 2. Pig cardiac tissue slices for histology and RNA extraction. a) Representative picture of pig heart 11 days post-MI. b) Pictures of pig heart slice levels (L) and whether samples from each level were fixed for histology (histo) or frozen for RNA extraction.



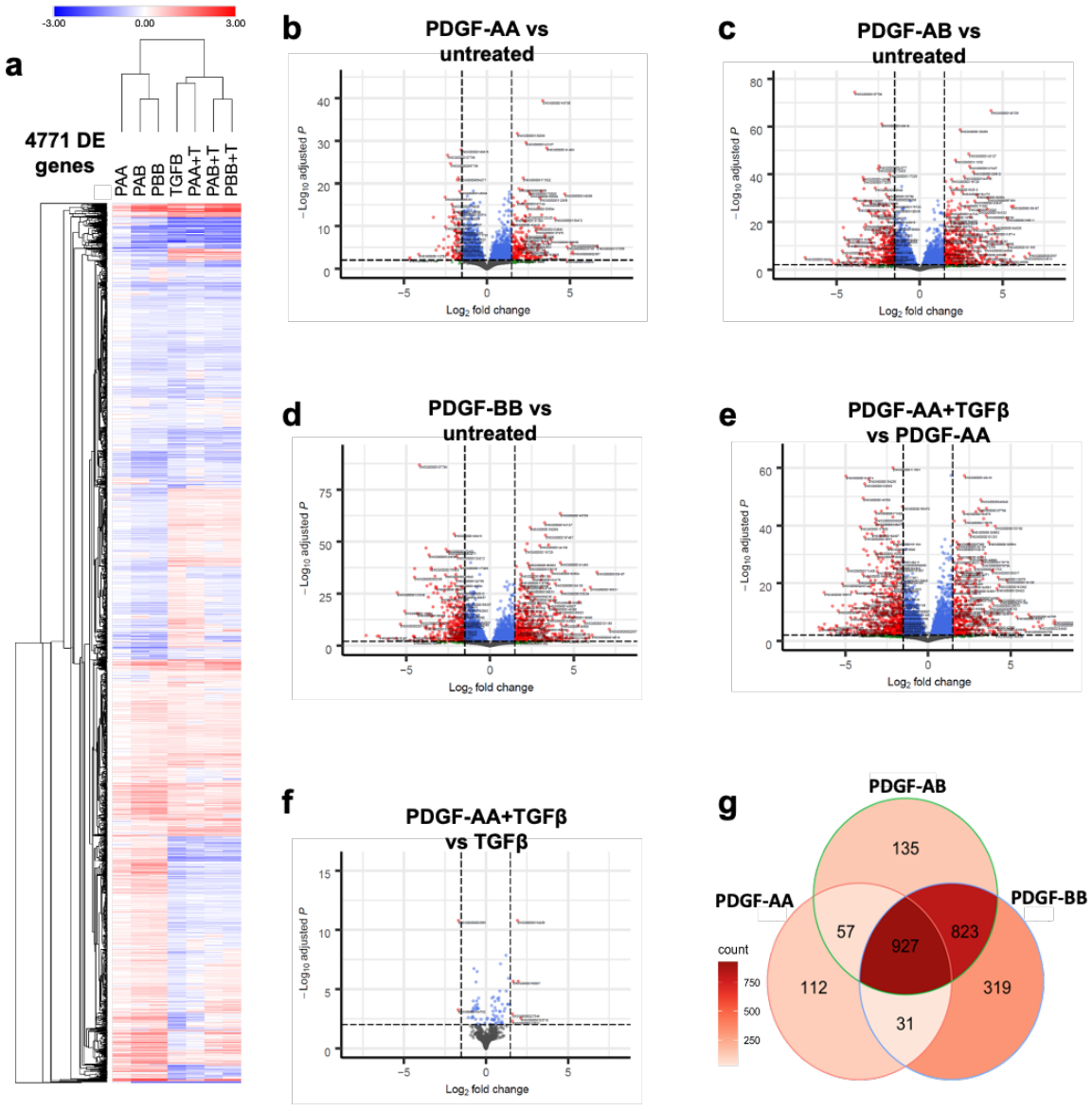
Supplemental Figure 3. Quality controls for post-MI pig cardiac tissue RNAseq. a) TapeStation analysis of pig RNA quality with associated RIN^e scores. b) TapeStation pig RNAseq library preparation quality control.



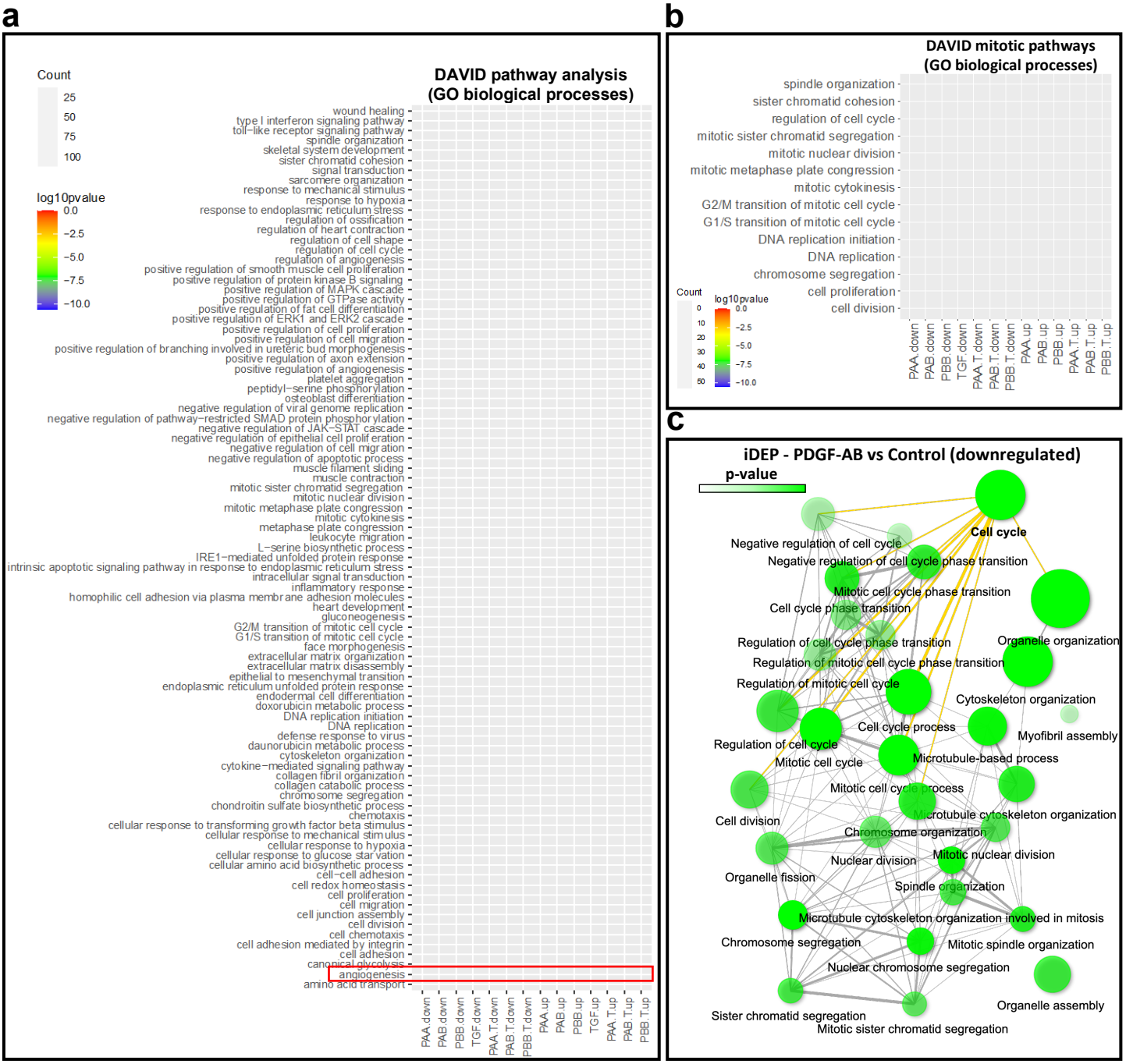
Supplemental Figure 4. PDGFs reduce fibroblast-myofibroblast differentiation in human cardiac stromal cells (cSC). cSC treated with PDGF-AA (PAA), PDGF-AB (PAB), PDGF-BB (PBB) all +/- TGF-β1 were a) imaged using brightfield microscopy and immunostained with α-smooth muscle actin (αSMA, green), b) immunostained for vimentin (red) and (αSMA, green), c) protein extracted, run on western blots, probed for αSMA and d) analysed for densitometry, d,e) RNA extracted and run on qPCR for ACTA2 or LOX and normalised to GAPDH. Statistical analyses were performed with a one-way ANOVA with a Tukey's multiple comparisons test, *p<0.05, **p<0.01, ***p<0.001).



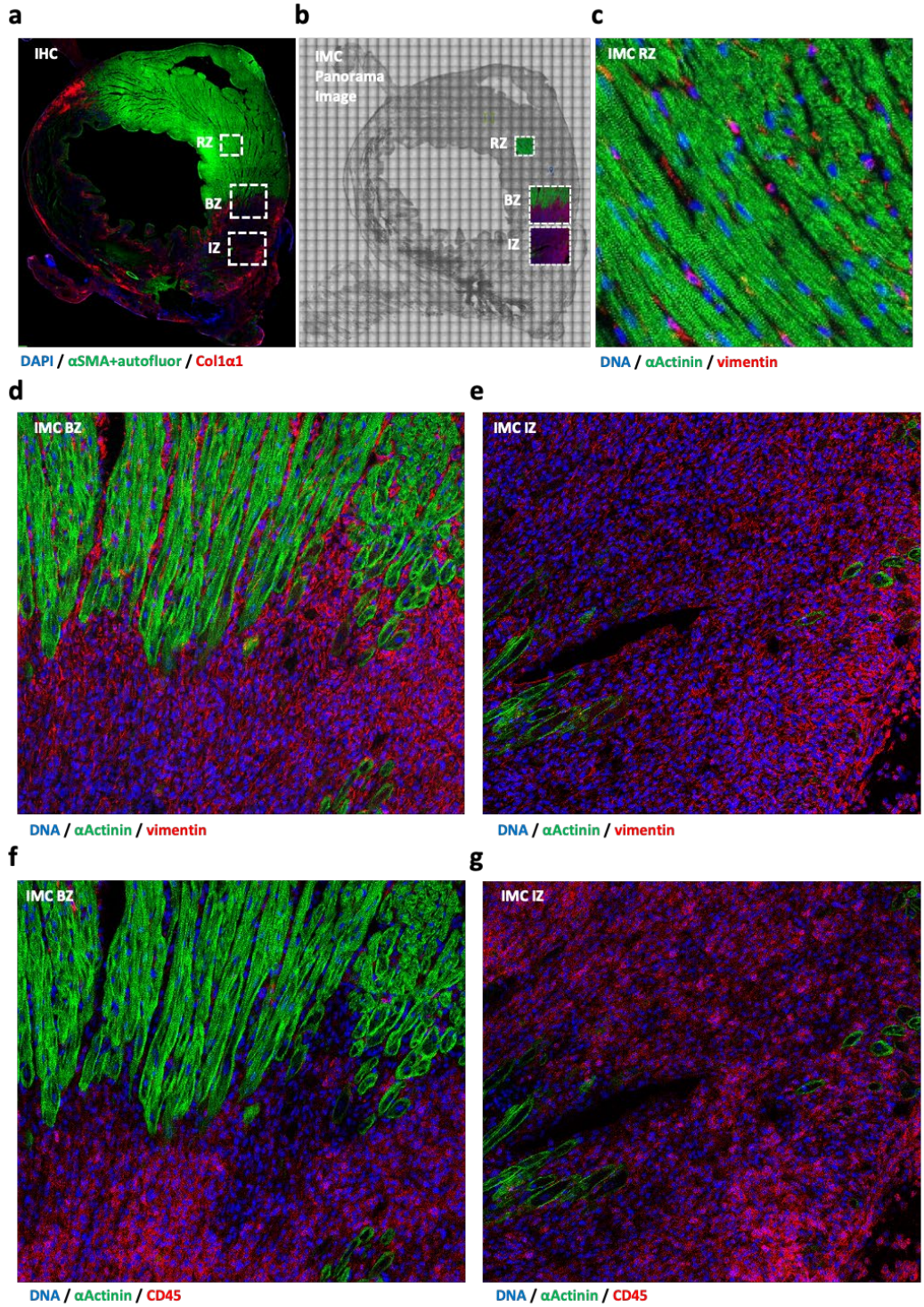
Supplemental Figure 5. PDGF-AB and PDGF-BB have similar effect on cFib gene expression that are distinct from PDGF-AA. cFib were treated with PDGF-AA (PAA), PDGF-AB (PAB) or PDGF-BB (PBB) +/- TGF-β1 (T) with RNAseq and differential gene expression analysis. a) Heatmap with hierarchical clustering of cFib treatment groups showing differentially expressed (DE) genes (scale bar=log_2fold change). b-f) Volcano plots showing DE genes within individual comparisons, red dots represent genes that met both p-value and fold change cut-offs. g) Venn diagram of overlapping DE genes in PDGF treated cFib.



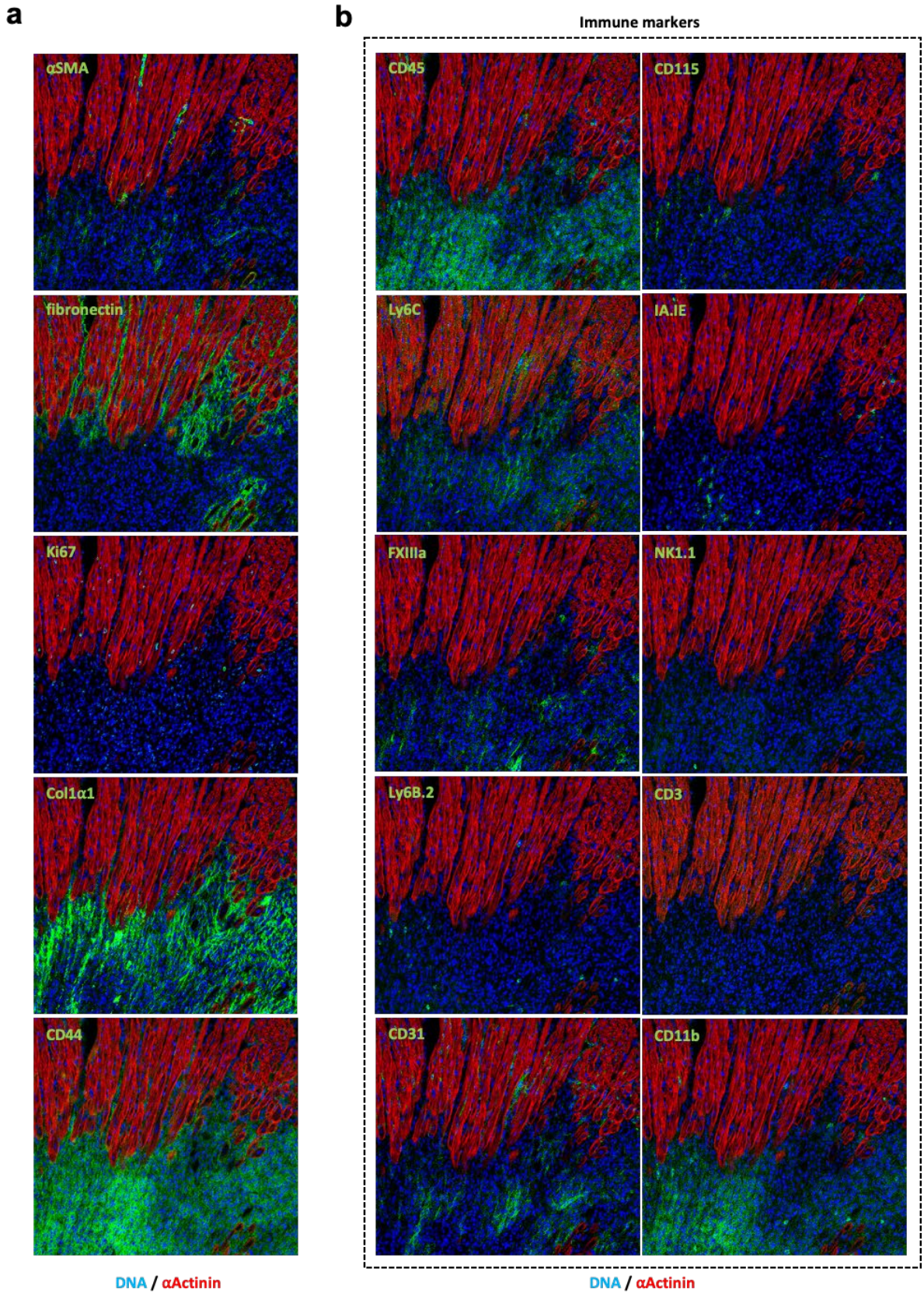
Supplemental Figure 6. DAVID and iDEP pathway analyses of human cardiac fibroblasts (cFib) RNAseq. RNAseq of cFib treated with PDGF-AA (PAA), PDGF-AB (PAB), PDGF-BB (PBB) +/- TGF-β1 (T) with (a) DAVID analysis for GO biological process terms (down=downregulated, up=upregulated), (b) DAVID analysis for GO biological process terms related to mitosis pathways and (c) iDEP pathway tree analysis of cFib RNAseq showing downregulated pathways for PDGF-AB samples.



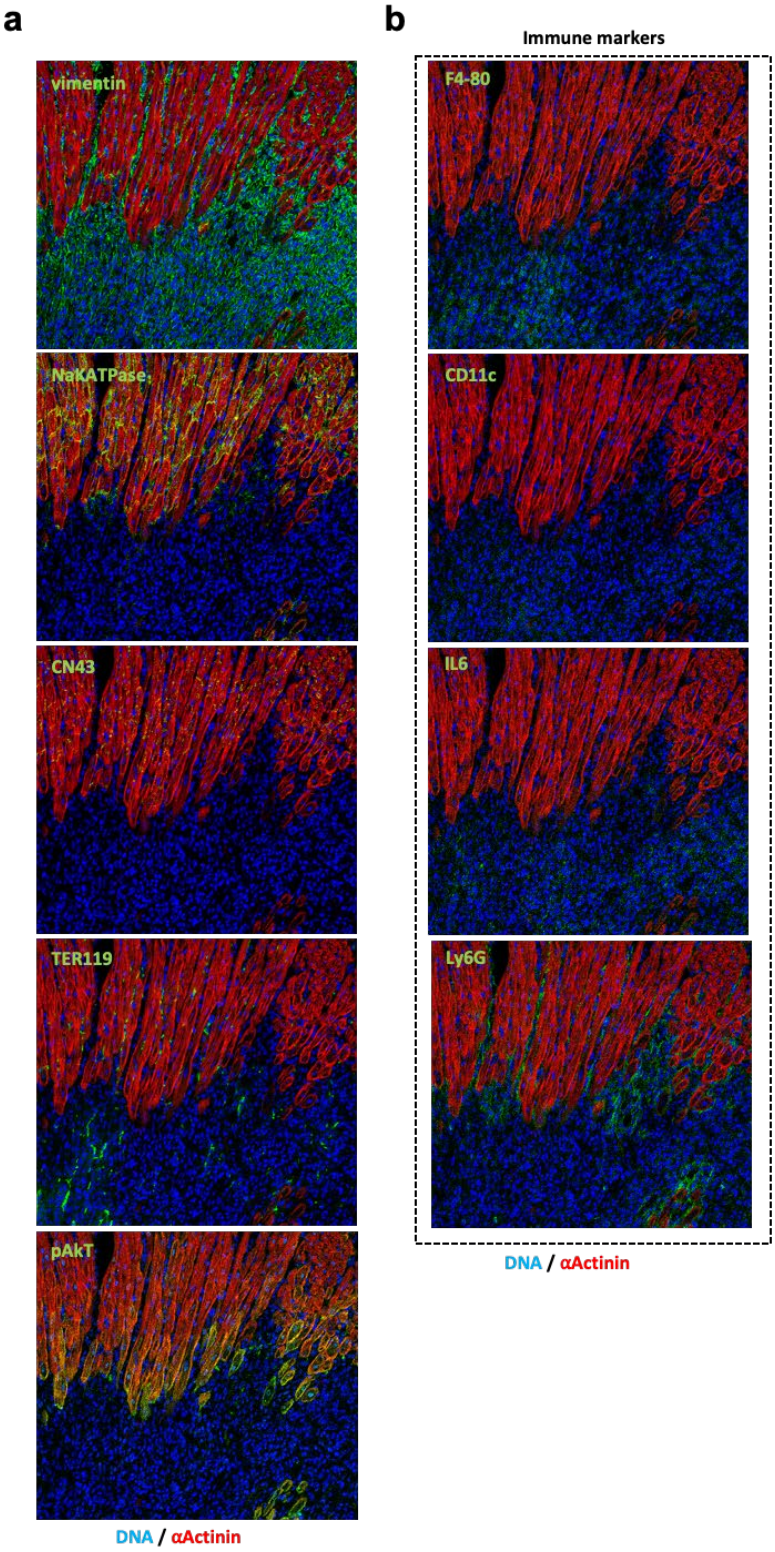
Supplemental Figure 7. Area selection for imaging mass cytometry (IMC). Post-MI mouse cardiac frozen tissue sections were (a) immunostained with αSMA (green) and col1α1 (red) or (b-g) stained with IMC metal-conjugated antibody panel. b) Brightfield panorama of cardiac section from IMC imaging system. c-e) IMC image of αActinin (green) and vimentin (red) in post-MI (c) remote zone (RZ), (d) border zone (BZ) and (e) infarct zone (IZ). f-g) IMC image of αActinin (green) and CD45 (red) in post-MI (f) BZ and (g) IZ.



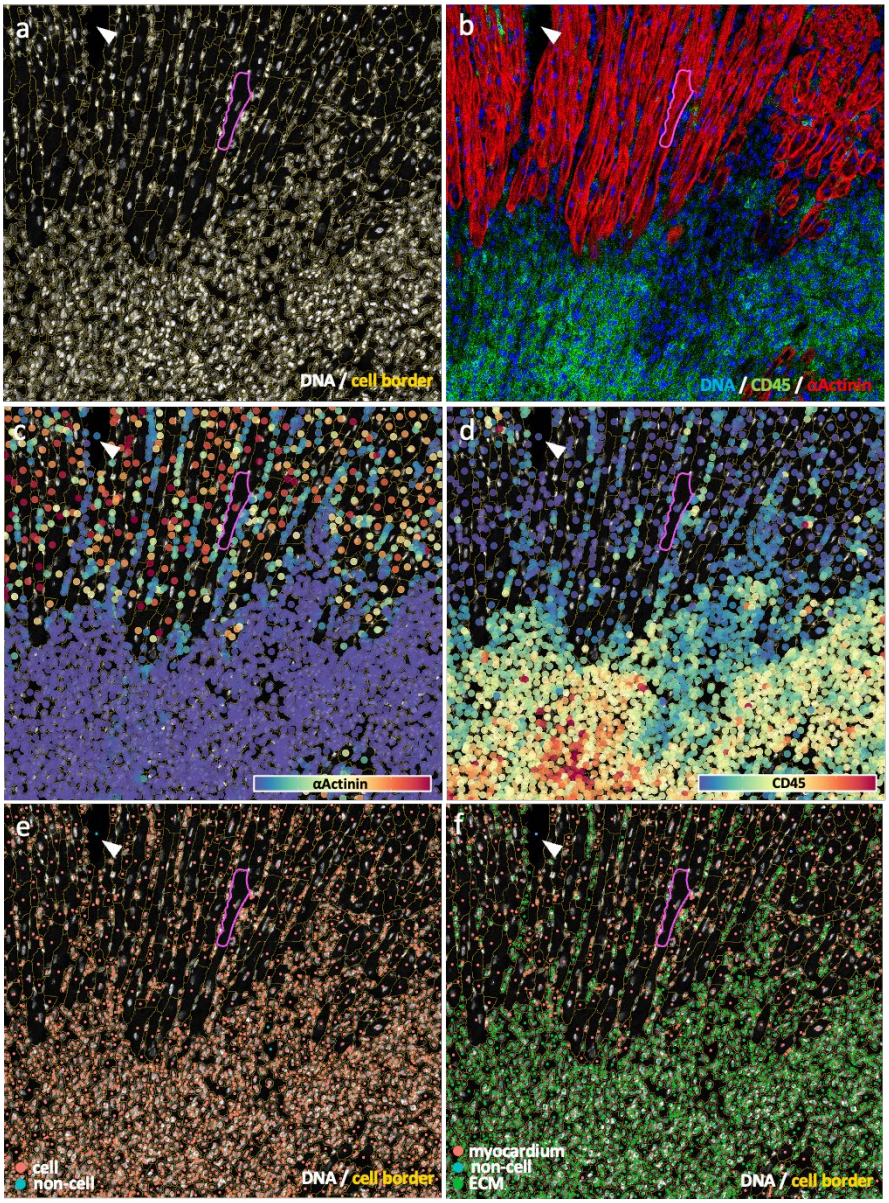
Supplemental Figure 8. IMC antibody panel (1/2). Post-MI mouse border zone (BZ) image showing (a) non-immune and (b) immune antibody markers (green) with αActinin (red) and DNA (blue).



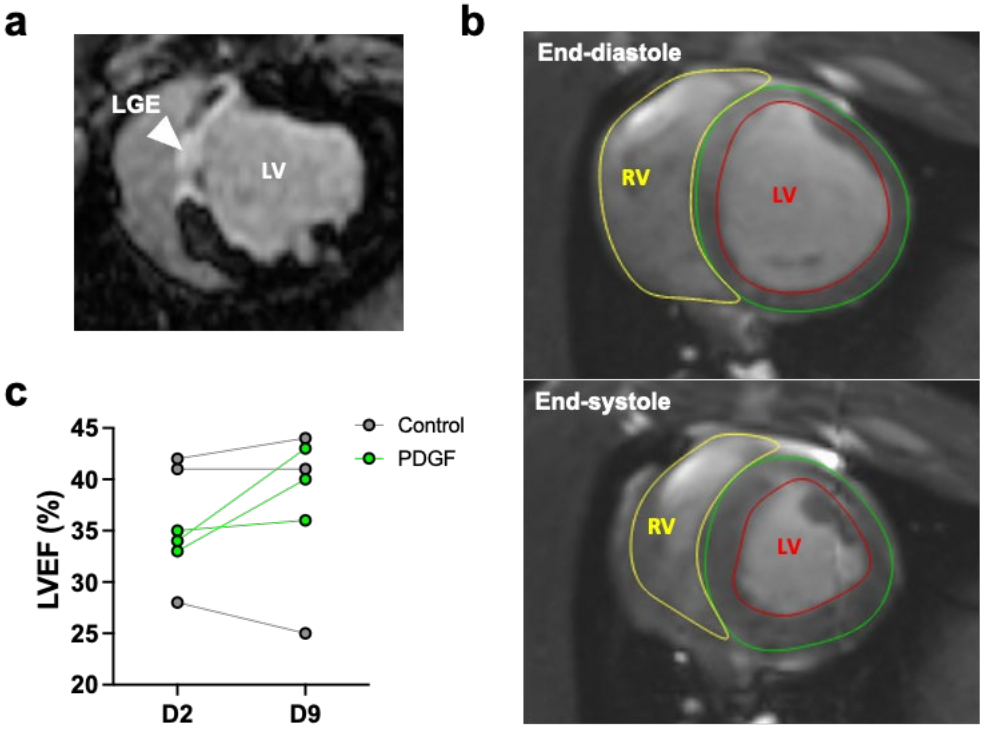
Supplemental Figure 9. IMC antibody panel (2/2). Post-MI mouse border zone (BZ) (IMC) image showing (a) non-immune and (b) immune antibody markers (green) with αActinin (red) and DNA (blue).



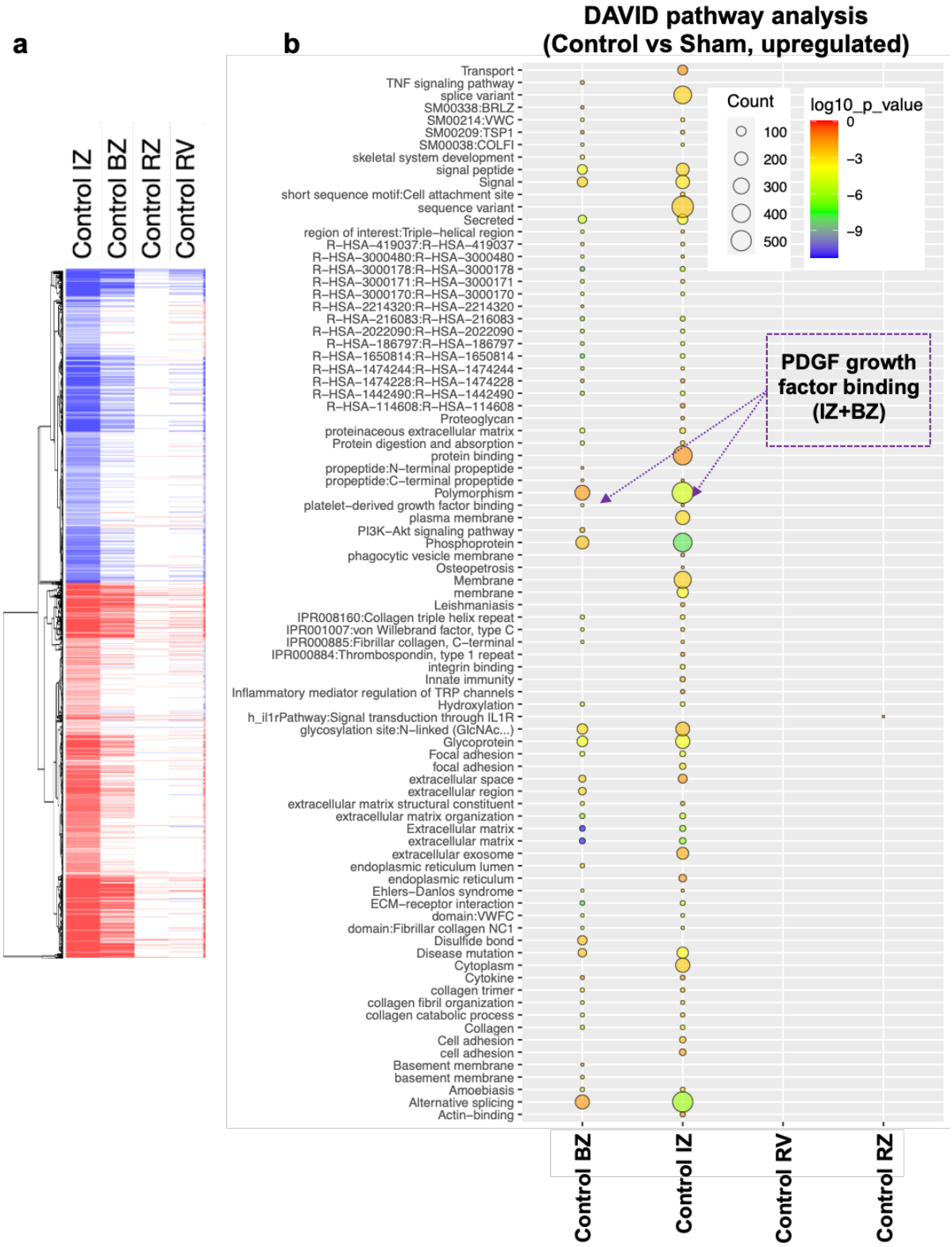
Supplemental Figure 10. IMC multi-cut cell segmentation successfully segments individual cardiomyocytes (CMs) and non-CMs. Post-MI mouse border zone (BZ) IMC image with (a) multi-cut cell segmentation in yellow, (b) DNA, CD45 and αActinin in blue, green and red, respectively, (c) αActinin level for each cell, with blue showing low expression and red showing high expression, (d) CD45 level for each cell, with blue showing low expression and red showing high expression, (e) background identification with cells in red and non-cells in blue, (d) area identification with myocardium in red, ECM/scar in green and non-cells in blue. Arrowhead in all images shows an area of background with no cells. Pink outline in all images shows multi-cut segmentation border for a cardiomyocyte.



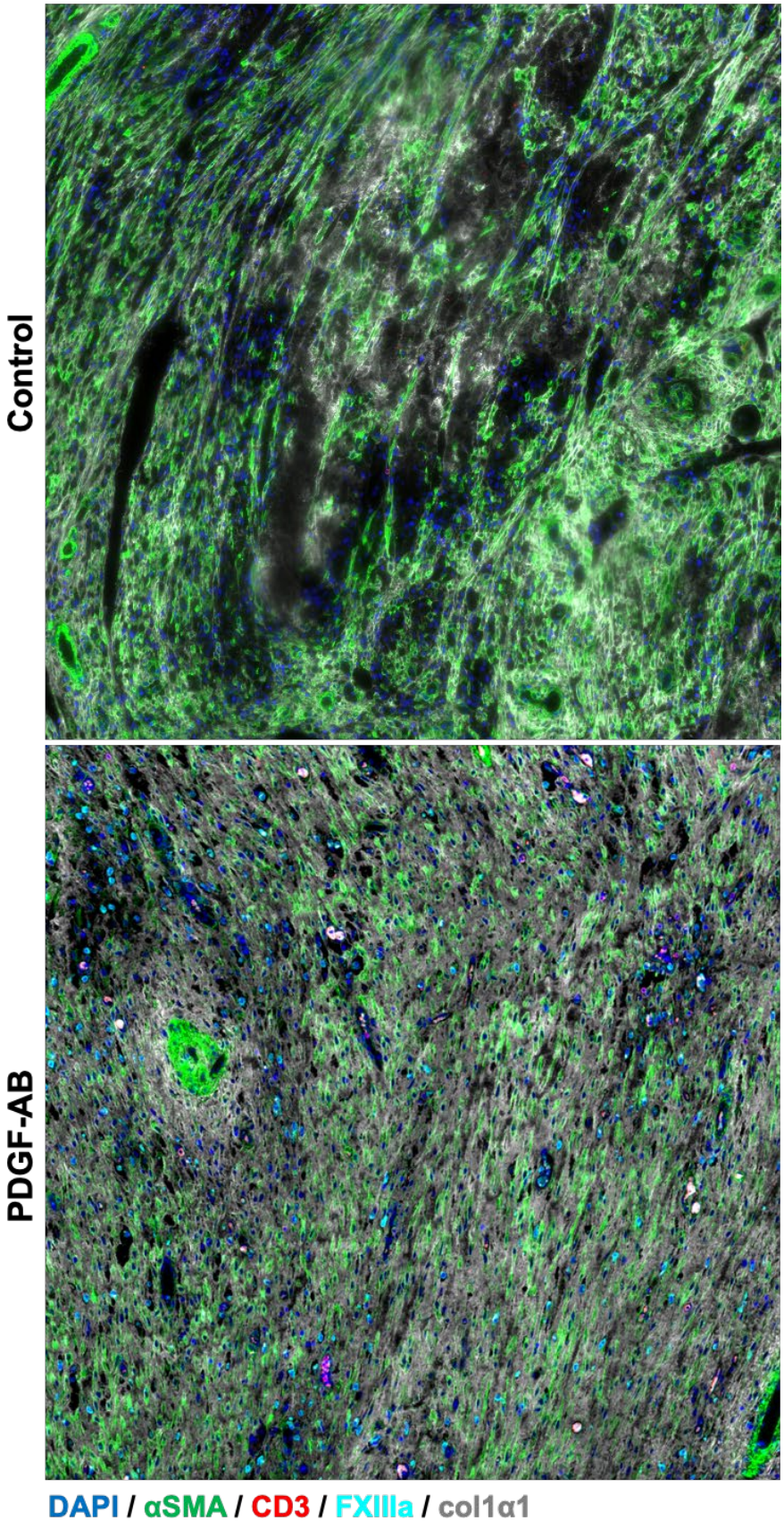
Supplemental Figure 11. Late-gadolinium enhancement (LGE) of left ventricle (LV) IZ, cardiac MRI (cMRI) contouring and individual left ventricular ejection fraction (LVEF) values. Short axis cMRI of pig LV 9 days post-MI. a) LGE shown by lighter area within the LV (arrowhead) indicating infarct area. b) Representative cMRI contour tracing of the LV and right ventricle (RV) at end-diastole and end-systole. c) Individual pig cMRI data for LVEF.



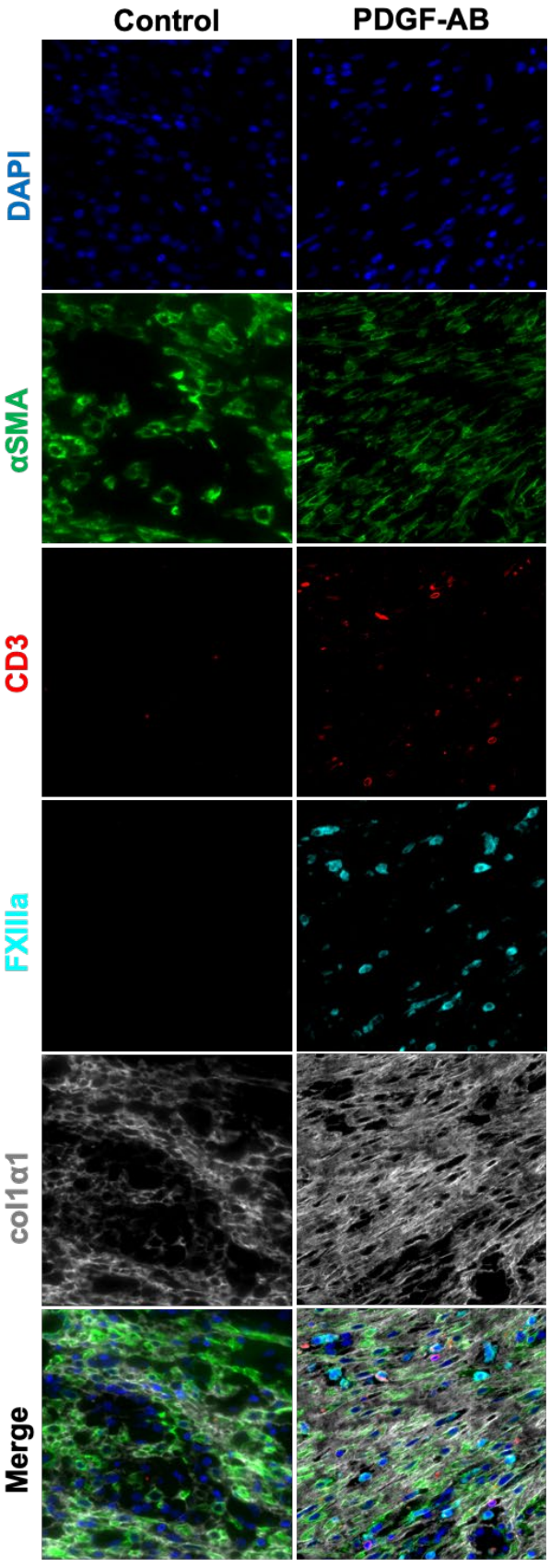
Supplemental Figure 12. *De novo* library assembly reveals novel genes, transcripts and pig infarct biology. a) Heatmap of DE genes in control infarcted vs sham pigs. b) DAVID pathway analysis for upregulated DE genes in control pigs (vs shams).



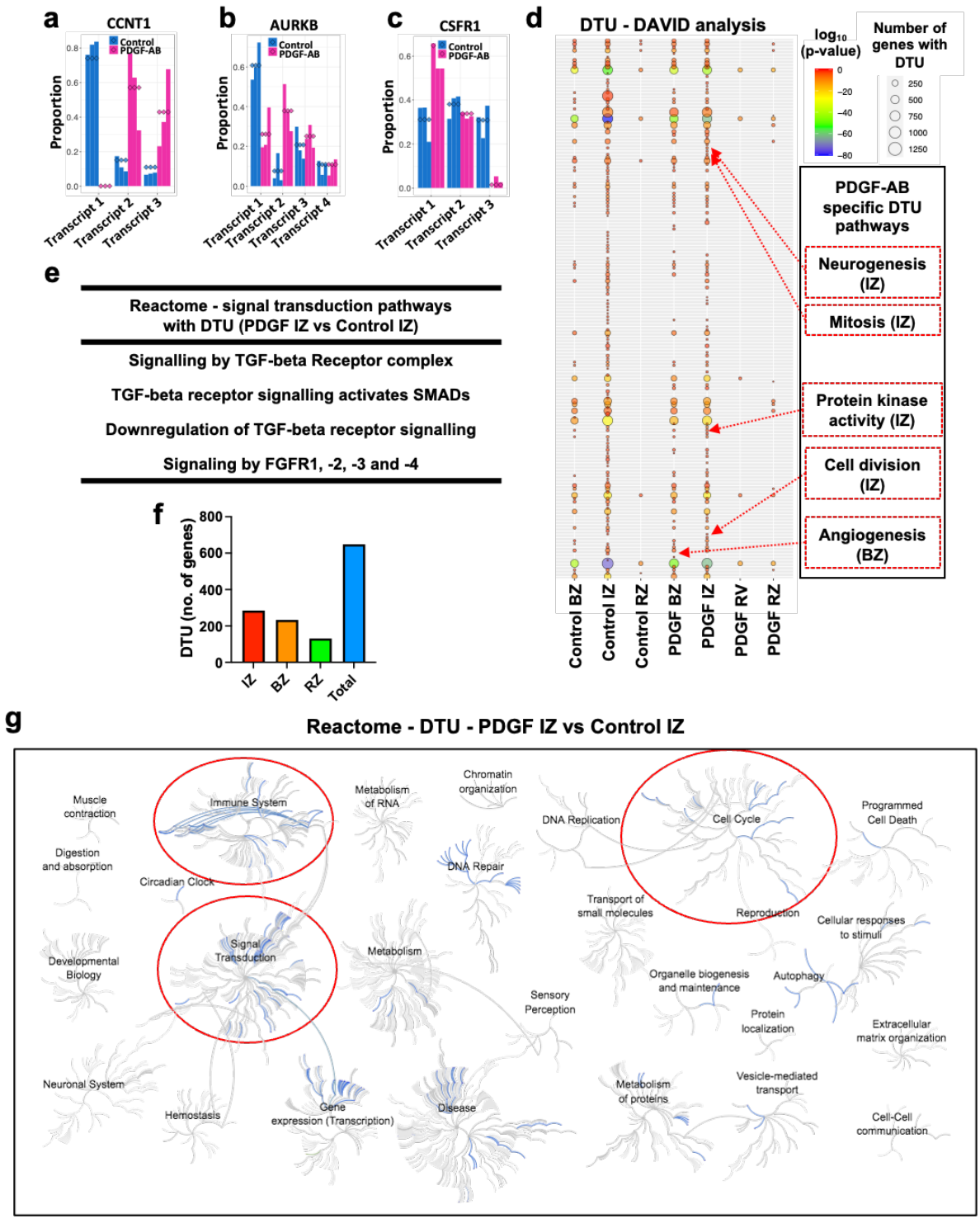
Supplemental Figure 13. PDGF-AB accelerates post-MI scar formation in pigs. IHC in pig IZ 11 days post-MI showing col1α1 scar organization.



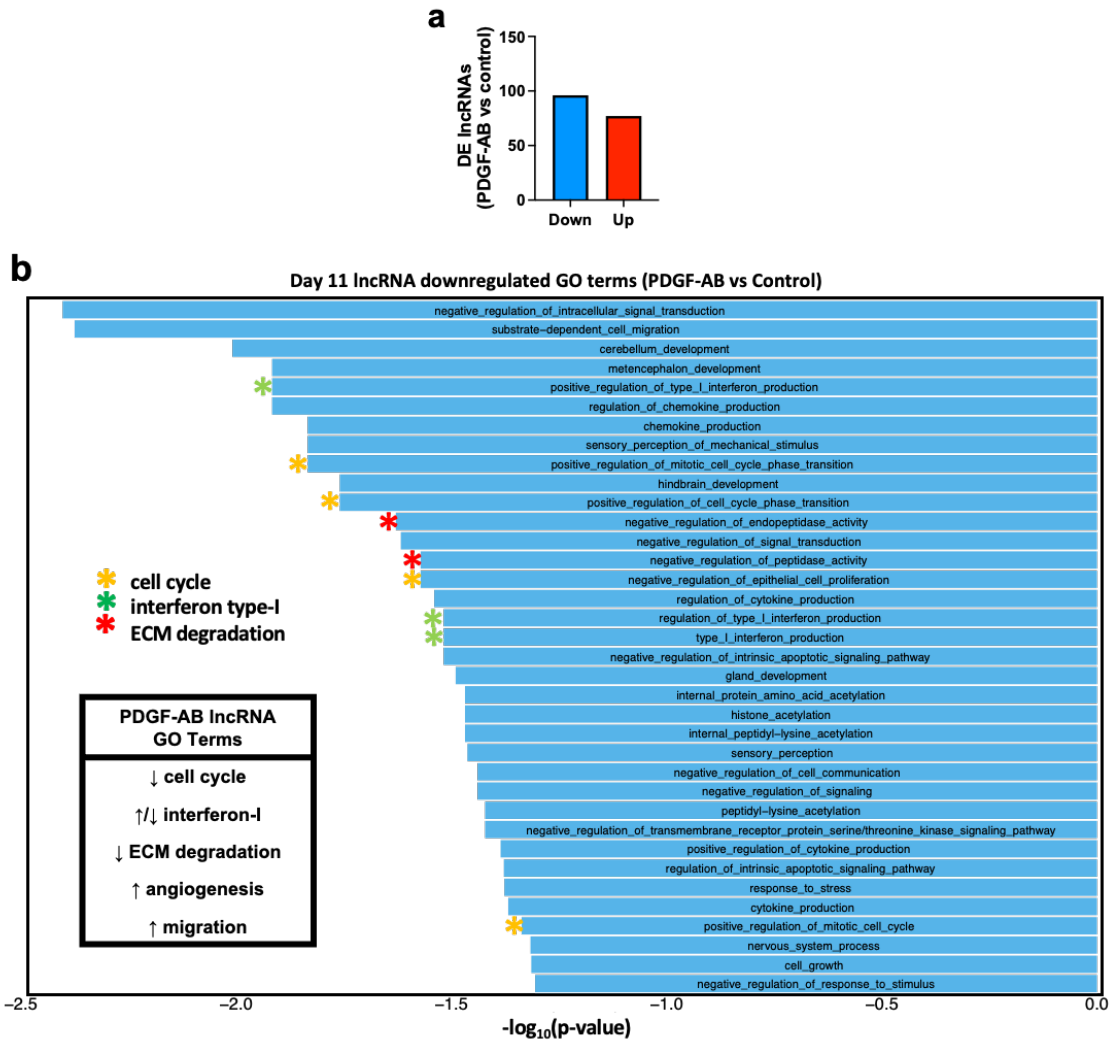
Supplemental Figure 14. PDGF-AB increases FXIIIa^{+} macrophages and collagen-1α1 in pigs 11 days post-MI. Individual channels from IHC of pig IZs 11 days post-MI.



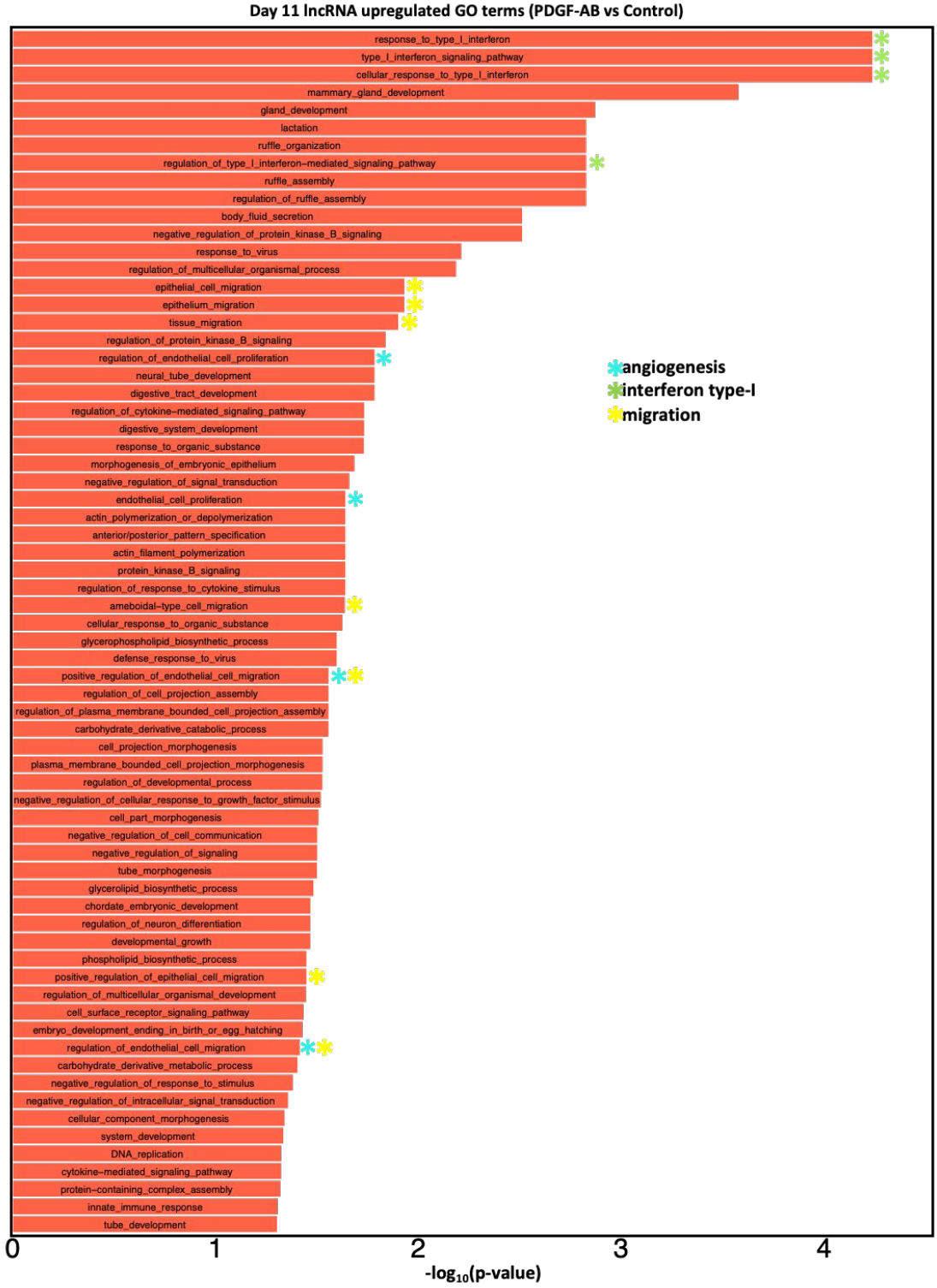
Supplemental Figure 15. PDGF-AB effects differential transcript usage (DTU) in genes within cell cycle, immune system and signal transduction pathways. a-e) Differential transcript usage (DTU) analysis of pig RNAseq. a-c) DTU analysis plots showing the proportion (y-axis) of each transcript variant (x-axis) for CCNT1 (a, cell cycle), AURKB (b, cell cycle) and CSFR1 (c, immune system) genes. d) DAVID analysis of genes with DTU for control or PDGF-AB-treated pigs versus shams. e) Reactome analysis of genes with DTU for signal transduction pathways in PDGF-AB-treated pig IZs. f) Number of genes with DTU in PDGF-AB-treated pigs. g) Reactome analysis of genes with DTU, blue represents significantly affected pathways.



Supplemental Figure 16: PDGF-AB impacts transcript variant and lncRNA expression within cell cycle and TGF-β pathways in pigs post-MI. a) Number novel differentially expressed (DE) long non-coding RNAs (lncRNAs) in PDGF-AB-treated pigs. b) GO term analysis for genes proximal to downregulated lncRNAs in PDGF-AB-treated pigs (table summarizes both downregulated and upregulated GO terms).



Supplemental Figure 17. lncRNAs are upregulated proximal to genes in migration, angiogenesis and interferon type-I pathways. GO term analysis of genes proximal to upregulated lncRNAs**.**



Supplemental References

1. Le TYL, Pickett HA, Dos Remedios CG, Barbaro PM, Kizana E, Chong JJH. Platelet-Derived Growth Factor Receptor-Alpha Expressing Cardiac Progenitor Cells Can Be Derived from Previously Cryopreserved Human Heart Samples. Stem Cells Dev. 2018;27(3):184–98.

2. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 2001;25(4):402–8.

3. Ge SX, Son EW, Yao R. iDEP: an integrated web application for differential expression and pathway analysis of RNA-Seq data. BMC Bioinformatics. 2018;19(1).

4. du Sert NP, Hurst V, Ahluwalia A et al. The ARRIVE guidelines 2.0: Updated guidelines for reporting animal research. PLoS Biol. 2020;18(7).

5. Thavapalachandran S, Grieve SM, Hume RD et al. Platelet-derived growth factor-AB improves scar mechanics and vascularity after myocardial infarction. Sci Transl Med. 2020;12(524).

6. Schulz-Menger J, Bluemke DA, Bremerich J et al. Standardized image interpretation and post-processing in cardiovascular magnetic resonance - 2020 update : Society for Cardiovascular Magnetic Resonance (SCMR): Board of Trustees Task Force on Standardized Post-Processing. J Cardiovasc Magn Reson. 2020;22(1).

7. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114–20.

8. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods. 2015;12(4):357–60.

9. Dobin A, Davis CA, Schlesinger F et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15–21.

10. Trapnell C, Williams BA, Pertea G et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol. 2010;28(5):511–5.

11. Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol. 2015;33(3):290–5.

12. Quinlan AR. BEDTools: The Swiss-Army Tool for Genome Feature Analysis. Curr Protoc Bioinforma. 2014;47:11.12.1-11.12.34.

13. Parra G, Blanco E, Guigó R. GeneID in Drosophila. Genome Res. 2000;10(4):511–5.

14. Sun L, Luo H, Bu D et al. Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. Nucleic Acids Res. 2013;41(17).

15. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12).

16. Neph S, Kuehn MS, Reynolds AP et al. BEDOPS: high-performance genomic feature operations. Bioinformatics. 2012;28(14):1919–20.

17. Harris MA, Clark J, Ireland A et al. The Gene Ontology (GO) database and informatics resource. Nucleic Acids Res. 2004;32(Database issue).

18. Alexa A, Rahnenführer J, Lengauer T. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. Bioinformatics. 2006;22(13):1600–7.

19. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods. 2017;14(4):417–9.