

Supplemental Material

Online Methods (pp. 2-5)

Online Tables I – XIV (pp. 6-35)

Online Figures I – VI (pp. 36-44)

References (p. 45)

Appendix on RNA-sequencing; laboratory and analytic protocols (pp. 46-65)

Online Methods

miR and mRNA expression analysis in human heart failure. The cohort of nonfailing and failing human heart samples (all Caucasian) used for these analyses has been previously described¹. Two of the 11 heart samples originally described as nonfailing had cardiac masses greater than 500 g and exhibited an altered miR signature, as previously discussed¹; these samples have been renamed ‘hypertrophied, non-failing’ in this manuscript. A further two samples clustered with failing samples rather than other nonfailing samples; these were excluded from the present data analyses. miR and mRNA expression data were calculated on the 5 remaining nonfailing (3 male, 2 female) and 13 failing (6 male, 7 female) human heart samples from which both miR and mRNA expression data were obtained (mRNA data were not available for the two ‘hypertrophied, non-failing’ hearts). Invitrogen NCode miR microarrays were used for determination of miR expression levels, while Affymetrix HuEx 1.0 exon microarrays were used to determine mRNA expression levels. Partek Genomics Suite (St. Louis, MO) software was used for array signal normalization and comparative analysis, as previously described¹.

miR and mRNA expression analysis in Gαq transgenic mice. miR microarray analysis was performed on four male nontransgenic and four male Gαq transgenic mouse hearts at LC Sciences (Houston, TX) as previously described². TaqMan microRNA assays (Applied Biosystems) for miR-499 and U6, and NCode microRNA RT-qPCR for 5S rRNA, were performed on a separate cohort of ntg and Gαq mice. RNA-sequencing for mRNA expression analysis was performed as previously described³.

Whole-genome transcription profiling by RNA sequencing. Preparation of cDNA fragments from RNA was as previously described⁴. For poly(A)+ RNA, 4 μg of total RNA was twice oligo(dT) selected using the Dynabead mRNA purification system (Invitrogen). Two hundred ng of mRNA was fragmented to ~200 nt by heating to 94 C for 2.5 min in 40 mmol/L Tris acetate pH 8.2, 100 mmol/L potassium acetate, 30 mmol/L magnesium acetate, and immediately chilled on ice. After purification on Ambion NucAway columns, 100 ng of fragmented mRNA was reverse-transcribed using SuperScript III (Invitrogen) with random hexamers as per the manufacturer’s directions (50 min, 50 C), followed by second-strand cDNA synthesis for 2 h, 16 C, in 20 mmol/L Tris.HCl pH 6.9, 90 mmol/L KCl, 4.6 mmol/L MgCl₂, 150 mmol/L β-NAD⁺, 10 mmol/L (NH₄)₂SO₄, 0.067 U/μL *E. coli* DNA ligase, 0.27 U/μL *E. coli* DNA polymerase I, and 0.013 U/mL *E. coli* RNase H (New England Biolabs).

RT-qPCR for RNA sequencing validation. SuperScript III (Invitrogen) with oligo-dT(20) primers was used to reverse-transcribe cDNA from 1 μg of total RNA. qPCR on 1/20th of the reverse-transcription product was performed in triplicate with TaqMan FAM-labeled probes and primer sets (Applied Biosystems) and Perfecta PCR supermix (Quanta Biosciences). The following TaqMan assays were used: *Actb*, Mm01205647_g1; *Gapdh*, Mm99999915_g1; *Hmbs*, Mm00660262_g1; *Irs2*, Mm03038438_m1; *Rcn2*, Mm00488777_m1; *Sos2*, Mm01265231_m1; *Tmbim6*, Mm00509864_m1.

In vivo miR target analysis by RISC sequencing. Conditions for Ago2 immunoprecipitation were adapted from Karginov *et al.*⁵, as described in⁴. Frozen mouse heart bases were homogenized in 500 μL ice-cold 50 mmol/L Tris.HCl, 5 mmol/L EDTA, 5 mmol/L EGTA, pH 7.5, with Roche Complete protease inhibitors. Yeast tRNA (Invitrogen) and SUPERnase-IN (Ambion) were added to final concentrations of 1 μg/μL and 1 U/μL, respectively, and unbroken cellular material was removed at 100g, 5’, 4 C. Nonidet P-40 was added to a final concentration of 0.5% (w/v) to solubilize proteins (15’, 4 C, rotating) and insoluble material was removed at 10000g, 15’, 4 C. The supernatant was added to 50 μL protein G-coupled Dynabeads (Invitrogen), to which 5 μg anti-Ago2 monoclonal antibody (Wako Pure, clone #2D4, lot PEM0820) had been previously bound according to the Dynabead protocol. Following 1 h rotational incubation at 4 C, the beads were washed 3x with Dynabead washing buffer, transferring the suspension to a fresh tube for the last wash. Beads were pelleted from the remaining suspension, the supernatant was

removed, and 500 μ L Trizol was added to the immunoprecipitated material to extract RNA. Ago2 immunoprecipitate-associated RNA was fragmented in acetate buffer without poly-A+ selection, purified on NucAway columns, and one half was used for cDNA synthesis using the protocol described above.

Construction of DNA-barcoded short-read RNA libraries for Illumina sequencing. Detailed methods for preparation of Illumina sequencing libraries from mouse cardiac cDNAs were recently described³. Briefly, cDNAs were end-repaired and 3' A-overhangs added. Illumina adapters with T-overhangs and customized to include three nt 'barcodes' were ligated to the cDNA at 10:1 molar excess and DNA in the 200-400 bp range was isolated via gel purification (Qiagen) on 2% low-melting agarose. One-third of the gel-purified material was amplified with 12 cycles (total myocardial cDNA) or 16 cycles (RISC-immunoprecipitated cDNA) of Phusion polymerase (New England Biolabs #F531)-mediated PCR (10 sec 98 C, 30 sec 65 C, 30 sec 72 C cycles, followed by final 5 min 72 C), using oligonucleotides complementary to Illumina sequencing adapters. The final, amplified libraries were again column-purified and quantified using PicoGreen (Quant-It, Invitrogen).

Eight barcoded libraries were combined in equimolar (10 nmol/L) amounts and diluted to 8 pmol/L for cluster formation on an Illumina HiSeq flowcell lane, followed by single-end sequencing. Library sorting by barcode and mapping to the transcriptome were performed as previously described³, using the software packages TopHat⁶ (version 1.1) and Cufflinks⁷ (version 0.9.3). Cufflinks outputs gene expression values in terms of Fragments Per Kb of exon per Million mapped reads (FPKM)⁷. When performing single-end Illumina sequencing, as we have done here, this parameter is equivalent to RPKM (Reads Per Kb of exon per Million mapped reads), in which we have estimated that an RPKM of 3 corresponds to 1 copy/cell in cardiac samples^{3,8}. We used the default options supplied with these software packages in our analyses and analyzed only those RNA elements that had expression signals in at least 50% of biological replicates.

Analysis of RNA sequencing data. Transcriptome sequencing libraries averaged 6.6×10^6 reads per sample, with 84% alignment to the mouse genome. RISCome sequencing libraries averaged 4.8×10^6 reads per sample, with 58% alignment to the mouse genome. Because they are not poly-A+ selected, RISCome sequencing libraries contained a higher proportion of ribosomal RNA (32-42%) than poly-A+ selected transcriptome libraries (3-5%). To permit comparison between the transcriptome and RISCome, measures of RNA abundance expressed as FPKM⁷ were scaled according to the proportion of mRNA in each sample. Gene symbols and FPKM values were imported into Partek Genomics Suite v6.5 (Partek, St Louis, MO) for comparison of RISCome and transcriptome expression values, computation of P-values, unsupervised hierarchical clustering with Euclidean distance, and heatmap plotting. Gene Ontology classification was performed using the BiNGO plugin to Cytoscape⁹, together with over-representation hypergeometric tests with Benjamini & Hochberg false discovery rate correction.

Analysis of RISC sequencing data. For each mRNA in each individual heart, its content in Ago2 immunoprecipitates was compared to that in the transcriptome (poly-A+ RNA) to derive a RISC score. The mean and standard error of the RISC score was calculated for each group of hearts (nontransgenic, n=7; miR-499 TG-16, n=6; and miR-499 TG-53, n=8). A RISC-enriched mRNA was defined as any mRNA with a RISC score ≥ 2 ; 63 mRNAs were found to have RISC scores in miR-499 hearts that were ≥ 1.25 -fold those in nontransgenic hearts, $P < 0.05$, and were defined as miR-499 RISC-enriched mRNAs. An additional 4 mRNAs were included, and defined as 'newly-enriched'. These had ≥ 3 -fold enrichment in Ago2-IP RNA (RISCome RNA) in transgenic compared to nontransgenic hearts, without a similar increase in transcriptome RNA, but did not meet the initial criteria. Downregulated mRNAs in the corresponding transcriptomes were defined as at least 1.1-fold downregulated compared to nontransgenic hearts, at $P < 0.025$.

*** For further detail on RNA_sequencing experimental procedures, and representative command scripts for alignment of raw sequencing reads using open-source software, see the ‘Appendix on RNA-sequencing’ at the end of the Online Information. ***

Sequence examination of RISC-targeted mRNAs for miR-499 binding sites. Bioinformatic assessment of miR-499 binding sites in mouse mRNA targets determined from RISC-sequencing was performed using Blast2seq (NCBI) to compare the wild-type miR-499 sequence against the coding region, 5’ and 3’ UTRs of these mRNAs (downloaded from the Ensembl database, www.ensembl.org/Mus_musculus, release NCBI m37, annotation v60). We used an unbiased sequence-matching approach, in which at least 2 of the 6 miR-499 canonical ‘seed’ nucleotides were required to be matched, with the additional restriction that at least 6 contiguous nucleotides needed to form Watson-Crick base pairs (with the *g:u* pair included in this definition¹⁰).

Unbiased proteomic and phosphoproteomic analysis. Two-dimensional differential in-gel electrophoresis (DiGE) was performed at Applied Biomics (Hayward, CA). 60 µg of mouse ventricular homogenates were labeled with Cy3 or Cy5 dyes (GE Healthcare), mixed, and resolved in the first dimension on a pH gradient from 4 to 9, and in the second dimension on 12% SDS-PAGE. Immediately following SDS-PAGE, proteins were visualized using a Typhoon TRIO laser scanner and analyzed by Image QuanTL software (GE Healthcare). In-gel analysis of differential protein expression was carried out with DeCyder software. Five 2D DiGE gels were analyzed, each containing nontransgenic and miR-499 transgenic heart samples. A 1.2-fold cutoff in 3 of 5 pairwise comparisons was set to determine whether differential regulation took place.

For protein identification, spots of interest were chosen using an Ettan Spot Picker (Amersham BioSciences) and digested in-gel with modified porcine trypsin protease (Trypsin Gold, Promega). Tryptic peptides were desalted and concentrated on ZipTip C18 (Millipore), eluted in 0.5 µL of matrix solution (α -cyano-4-hydroxycinnamic acid [5 mg/mL in 50% acetonitrile, 0.1% trifluoroacetic acid, 25 mmol/L ammonium bicarbonate]), and spotted on a MALDI plate (model ABI 01-192-6-AB). MALDI-TOF MS (matrix-assisted laser desorption/ionization–time-of-flight MS) and TOF/TOF tandem MS/MS were performed on an ABI 4700 mass spectrometer (Applied Biosystems, Framingham, Mass). MALDI-TOF mass spectra were acquired in reflection positive ion mode, averaging 4000 laser shots per spectrum. TOF/TOF tandem MS fragmentation spectra were acquired for each sample, averaging 4000 laser shots per fragmentation spectrum on each of the 10 most abundant ions present in each sample (excluding trypsin autolytic peptides and other known background ions). Peptide mass and associated fragmentation spectra were submitted to a GPS Explorer workstation equipped with the MASCOT search engine (Matrix Science) to search the nonredundant database of the National Center for Biotechnology Information (NCBI-nr). Searches were performed without constraining protein molecular weight or isoelectric point, with variable carbamidomethylation of cysteine and oxidation of methionine residues, and with 1 missed cleavage also allowed in the search parameters. Candidates with either a protein score confidence interval percentage or ion confidence interval percentage of >95% were considered significant.

For phosphoproteome analysis and 2D immunoblotting, 2D DiGE was performed on 150 µg of Cy5-labeled mouse ventricular homogenates. Phosphoprotein staining used the ProQ-diamond staining kit (Invitrogen). Primary antibodies for 1D and 2D immunoblots were from Cell Signaling Technology (anti-HSP90 β , #5087; anti-PP1 α , #2582). Anti-rabbit secondary antibodies were HRP-conjugated from Cell Signaling Technology (1D immunoblot) or Cy3-labeled anti-rabbit IgG antibodies from Applied Biomics.

Statistical analysis. Unless otherwise specified, all data are presented as mean \pm s.e.m and P-values were calculated using Student’s unpaired t-test (for two groups) or 1-way ANOVA (for more than 2 groups). Statistical significance was taken at P<0.05. Comparison of transcriptomic data sets, including calculation of false discovery rates, was performed using Partek Genomics Suite 6.5; significance was

taken at $P < 0.005$ ($FDR < 0.05$). The online software suite MetaCore¹¹ was used to perform gene ontology and gene set categorization, and enrichment analyses. mRNA enrichment was assigned at $P < 0.05$ in categories above the level predicted from the hypergeometric distribution.

Online Table I: Downregulated miR-499 targets in human heart failure.

Gene symbol	Gene name	Fold-change (HF vs NF)	p-value (HF vs NF)
ADAT2	adenosine deaminase, tRNA-specific 2, TAD2 homolog (<i>S. cerevisiae</i>)	-1.3	3.0E-04
AEBP2	AE binding protein 2	-1.3	2.1E-02
ALPL	alkaline phosphatase, liver/bone/kidney	-1.2	2.1E-03
ANKRD13C	ankyrin repeat domain 13C	-1.5	4.3E-02
ANKRD40	ankyrin repeat domain 40	-1.7	1.6E-03
ARHGAP23	Rho GTPase activating protein 23	-1.6	9.8E-03
ARHGAP32	Rho GTPase activating protein 32	-2.6	5.0E-03
ATF7	activating transcription factor 7	-1.3	2.9E-02
BTG1	B-cell translocation gene 1, anti-proliferative	-1.2	8.6E-03
CASK	calcium/calmodulin-dependent serine protein kinase (MAGUK family)	-1.3	2.6E-02
COMMD2	COMM domain containing 2	-1.6	3.8E-02
COPS3	COP9 constitutive photomorphogenic homolog subunit 3 (<i>Arabidopsis</i>)	-1.4	2.6E-03
CPM	carboxypeptidase M	-2.1	2.3E-07
CPSF2	cleavage and polyadenylation specific factor 2, 100kDa	-1.4	3.4E-02
DCAF12	DDB1 and CUL4 associated factor 12	-1.5	1.8E-02
DCUN1D5	DCN1, defective in cullin neddylation 1, domain containing 5 (<i>S. cerevisiae</i>)	-1.3	2.0E-03
DENND4C	DENN/MADD domain containing 4C	-1.5	4.0E-02
DNAJC15	DnaJ (Hsp40) homolog, subfamily C, member 15	-1.6	4.3E-03
DNM1L	dynamamin 1-like	-1.4	3.8E-02
DPY19L3	dpy-19-like 3 (<i>C. elegans</i>)	-1.3	1.8E-02
EDNRB	endothelin receptor type B	-2.0	8.1E-07
EFNB1	ephrin-B1	-1.2	1.8E-04
EPS8	epidermal growth factor receptor pathway substrate 8	-1.0	9.4E-04
ERLIN1	ER lipid raft associated 1	-1.2	8.5E-03
ERO1L	ERO1-like (<i>S. cerevisiae</i>)	-1.6	1.4E-03
ETNK1	ethanolamine kinase 1	-1.4	9.2E-03
FAM168A	family with sequence similarity 168, member A	-1.2	3.1E-02
FAM169A	family with sequence similarity 169, member A	-1.4	4.4E-02
FGF2	fibroblast growth factor 2 (basic)	-1.3	1.9E-02
FKBP5	FK506 binding protein 5	-2.2	5.7E-06
FNDC3A	fibronectin type III domain containing 3A	-1.5	3.6E-02
GMFB	glia maturation factor, beta	-1.7	7.5E-03
H3F3B	H3 histone, family 3B (H3.3B)	-1.2	4.4E-02
ILF3	interleukin enhancer binding factor 3, 90kDa	-1.5	4.7E-03
IRS2	insulin receptor substrate 2	-1.2	4.8E-03

Gene symbol	Gene name	Fold-change (HF vs NF)	p-value (HF vs NF)
JPH1	junctophilin 1	-1.4	2.7E-04
KATNAL1	katanin p60 subunit A-like 1	-3.3	3.3E-03
KPNA3	karyopherin alpha 3 (importin alpha 4)	-1.5	3.4E-02
LHFPL2	lipoma HMGIC fusion partner-like 2	-1.2	1.3E-03
LRRC8A	leucine rich repeat containing 8 family, member A	-1.5	8.9E-07
MAMDC2	MAM domain containing 2	-1.6	1.4E-05
MARK2	MAP/microtubule affinity-regulating kinase 2	-1.2	2.5E-02
MBNL3	muscleblind-like 3	-1.3	3.7E-03
MDM2	Mdm2 p53 binding protein homolog (mouse)	-1.4	1.7E-02
MED13L	mediator complex subunit 13-like	-1.2	4.1E-02
MON2	MON2 homolog (S. cerevisiae)	-1.5	4.2E-02
MYLK3	myosin light chain kinase 3	-1.8	3.0E-02
NAP1L1	nucleosome assembly protein 1-like 1	-1.5	3.1E-02
NIPA2	non imprinted in Prader-Willi/Angelman syndrome 2	-1.5	6.3E-03
NPLOC4	nuclear protein localization 4 homolog (S. cerevisiae)	-1.2	9.8E-03
NRIP1	nuclear receptor interacting protein 1	-1.5	1.7E-02
NUFIP2	nuclear fragile X mental retardation protein interacting protein 2	-1.6	3.7E-02
OSBPL1A	oxysterol binding protein-like 1A	-1.2	1.9E-02
P4HA1	procollagen-proline, 2-oxoglutarate 4-dioxygenase (proline 4-hydroxylase), alpha polypeptide I	-1.7	2.6E-02
PARVA	parvin, alpha	-1.4	3.6E-04
PDCD4	programmed cell death 4 (neoplastic transformation inhibitor)	-1.5	5.0E-02
PECAM1	platelet/endothelial cell adhesion molecule (CD31 antigen)	-1.6	2.4E-04
PGRMC2	progesterone receptor membrane component 2	-1.2	4.2E-02
PHF17	PHD finger protein 17	-1.2	3.9E-02
PIM1	pim-1 oncogene	-1.4	9.7E-05
PPM1D	protein phosphatase 1D magnesium-dependent, delta isoform	-1.3	1.3E-03
PPP3CA	protein phosphatase 3 (formerly 2B), catalytic subunit, alpha isoform	-1.3	4.8E-02
PRKX	protein kinase, X-linked	-1.2	2.9E-03
PTP4A1	protein tyrosine phosphatase type IVA, member 1	-1.4	2.5E-02
RAB22A	RAB22A, member RAS oncogene family	-1.5	5.6E-03
RAB5C	RAB5C, member RAS oncogene family	-1.7	7.6E-05
RAB8B	RAB8B, member RAS oncogene family	-1.6	4.8E-03
RFFL	ring finger and FYVE-like domain containing 1	-1.3	4.4E-03

Gene symbol	Gene name	Fold-change (HF vs NF)	p-value (HF vs NF)
RNF114	ring finger protein 114	-1.4	1.7E-02
ROD1	ROD1 regulator of differentiation 1 (S. pombe)	-1.9	2.0E-03
RRP15	ribosomal RNA processing 15 homolog (S. cerevisiae)	-1.6	2.9E-02
RRP1B	ribosomal RNA processing 1 homolog B (S. cerevisiae)	-1.4	1.3E-03
SAMD4B	sterile alpha motif domain containing 4B	-1.2	9.5E-03
SAMD8	sterile alpha motif domain containing 8	-1.5	3.5E-03
SEL1L	sel-1 suppressor of lin-12-like (C. elegans)	-1.6	4.7E-03
SF3B3	splicing factor 3b, subunit 3, 130kDa	-1.3	3.9E-02
SH3BGRL2	SH3 domain binding glutamic acid-rich protein like 2	-1.2	5.1E-03
SMAD4	SMAD family member 4	-1.4	2.2E-02
SOS2	son of sevenless homolog 2 (Drosophila)	-1.6	3.7E-02
SPRED1	sprouty-related, EVH1 domain containing 1	-1.4	1.6E-02
SPRY1	sprouty homolog 1, antagonist of FGF signaling (Drosophila)	-1.2	1.9E-02
STK4	serine/threonine kinase 4	-1.3	1.4E-02
SUV39H1	suppressor of variegation 3-9 homolog 1 (Drosophila)	-1.2	2.5E-02
SYAP1	synapse associated protein 1, SAP47 homolog (Drosophila)	-1.4	1.9E-02
TBC1D15	TBC1 domain family, member 15	-1.4	2.8E-02
TCF12	transcription factor 12 (HTF4, helix-loop-helix transcription factors 4)	-1.3	1.9E-02
TMBIM6	transmembrane BAX inhibitor motif containing 6	-1.4	2.0E-03
TMEM2	transmembrane protein 2	-1.6	6.6E-05
TMX1	thioredoxin-related transmembrane protein 1	-1.5	9.1E-04
TSPAN3	tetraspanin 3	-1.5	1.3E-02
UBE2V1	ubiquitin-conjugating enzyme E2 variant 1	-1.3	4.1E-02
USP47	ubiquitin specific peptidase 47	-1.6	4.5E-02
WAC	WW domain containing adaptor with coiled-coil	-1.3	3.8E-02
WNK3	WNK lysine deficient protein kinase 3	-1.4	1.2E-03
YAF2	YY1 associated factor 2	-1.4	4.8E-02
YIPF6	Yip1 domain family, member 6	-1.3	2.9E-02
ZDHHC2	zinc finger, DHHC-type containing 2	-1.6	1.5E-02
ZDHHC9	zinc finger, DHHC-type containing 9	-1.5	9.7E-04

Online Table I. Downregulated miR-499 targets, predicted by TargetScan 6, in human heart failure. mRNA expression was assessed on Affymetrix HuEx v1.0 Exon arrays and analyzed using Partek Genomics Suite. A heatmap corresponding to these data is given in **Figure 1c**. NF, nonfailing; HF, heart failure of ischemic or nonischemic origin. Downregulation was defined as a fold-change <-1.2, P<0.05.

Online Table II: Statistical comparison of LVEDD and FS% values for nontransgenic and miR-499 transgenic mice.

LVEDD						
	6 wk mean	s.e.m.	6 wk n	P-value vs ntg	P-value TG-16 vs TG-15	P-value TG-16 vs TG-53
ntg	3.5	0.04	15			
TG-16	3.6	0.05	13	P > 0.05		
TG-15	4.0	0.08	4	P < 0.01	P < 0.05	
TG-53	3.7	0.14	5	P > 0.05		P > 0.05

LVEDD						
	20 wk mean	s.e.m.	20 wk n	P-value vs ntg	P-value TG-16 vs TG-15	P-value TG-16 vs TG-53
ntg	3.6	0.04	18			
TG-16	4.0	0.06	12	P < 0.001		
TG-15	4.5	0.08	4	P < 0.001	P < 0.01	
TG-53	4.4	0.26	7	P < 0.001		P < 0.01

FS%						
	6 wk mean	s.e.m.	6 wk n	P-value vs ntg	P-value TG-16 vs TG-15	P-value TG-16 vs TG-53
ntg	73.0	0.8	15			
TG-16	65.9	1.5	13	P < 0.01		
TG-15	58.0	2.8	4	P < 0.001	P < 0.05	
TG-53	64.1	2.8	5	P < 0.01		P > 0.05

FS%						
	20 wk mean	s.e.m.	20 wk n	P-value vs ntg	P-value TG-16 vs TG-15	P-value TG-16 vs TG-53
ntg	70.6	1.7	18			
TG-16	59.4	1.3	12	P < 0.001		
TG-15	46.3	2.2	4	P < 0.001	P < 0.001	
TG-53	52.5	3.4	7	P < 0.001		P < 0.05

Online Table II. *Statistical comparison of LVEDD and FS% values for nontransgenic and miR-499 transgenic mice.* Two-way ANOVA was performed with GraphPad Prism to determine differences between miR-499 and nontransgenic mice at two different ages, and also to evaluate the interaction of genotype and time. The interaction term for genotype and time was significant for LVEDD (P=0.035), meaning that genotype did not have the same effect at both times. The interaction term for genotype and term was not significant for FS% (P=0.11). P-values for pairwise comparisons were obtained from Bonferroni post-tests in GraphPad Prism.

Online Table III: Regulated mRNAs in 4 week-old and 8 week-old miR-499 transgenic (TG-16) hearts.

Gene symbols	4-wk ntg mean	8-wk ntg mean	Fold-change 4-wk TG-16/ntg	p-value	Fold-change 8-wk TG-16/ntg	8-wk p-value
Mt2	59.6	65.0	4.6	1.4E-02	2.6	4.1E-02
Xist	8.2	3.6	2.8	4.1E-02	1.2	8.5E-01
Klhl34	3.2	2.8	2.7	1.2E-04	2.6	9.4E-05
Acta1	503.8	574.3	2.6	7.9E-04	2.4	1.0E-02
Gbp11	3.2	3.0	2.6	6.1E-03	1.7	2.3E-02
Prune2	8.6	8.0	2.3	3.9E-04	2.4	2.1E-03
Lrp11	3.9	5.7	2.2	4.4E-02	1.9	5.5E-02
Tmem52	3.6	3.7	2.1	1.3E-03	2.3	2.0E-02
Clu	182.2	196.7	2.1	1.1E-04	1.5	5.5E-02
Ntng2	5.7	6.7	2.0	3.9E-04	1.9	5.5E-04
Mt1	593.6	623.3	1.8	1.5E-03	1.5	7.6E-03
Nkain1	6.2	7.5	1.7	1.3E-03	1.6	3.1E-04
Dedd	34.0	30.1	1.7	4.7E-02	1.2	1.3E-01
Srrm4	4.2	4.8	1.7	8.4E-03	1.4	5.3E-02
Mgp	392.7	301.3	1.7	2.4E-04	1.8	1.2E-02
Scx	4.1	4.8	1.7	6.4E-03	1.3	4.8E-02
Runx1t1	3.4	3.9	1.7	6.9E-03	1.4	1.6E-01
Gsta3	4.6	7.3	1.7	4.0E-02	1.7	4.7E-02
Dnaic1	3.8	2.1	1.7	2.9E-02	1.6	1.1E-01
Fhl1	112.2	98.3	1.7	2.8E-04	1.6	1.4E-03
Mad2l2	8.2	9.3	1.7	1.8E-02	1.6	1.3E-02
Map2k6	3.7	3.1	1.7	1.7E-02	1.6	3.9E-02
Comt1	99.8	89.7	1.6	1.1E-08	1.6	4.1E-04
Cd40	3.9	4.3	1.6	1.7E-02	1.5	1.2E-01
Lrp4	3.0	2.8	1.6	1.3E-02	1.3	1.7E-01
Acot9	9.0	7.1	1.6	1.6E-02	1.4	3.2E-02
H2-Aa	17.4	13.3	1.6	9.1E-03	1.8	4.4E-02
Fam46a	7.4	5.0	1.6	6.2E-03	2.0	3.8E-04
Pik3r3	3.7	2.8	1.6	8.3E-03	2.0	1.5E-04
2310028H24Rik	8.9	9.1	1.5	1.3E-02	1.4	3.3E-03
Ephb3	4.8	4.6	1.5	1.4E-02	1.7	3.9E-03
St6galnac2	7.9	7.2	1.5	1.0E-02	1.6	4.9E-03
Lhx6	3.5	3.3	1.5	4.7E-02	1.6	1.2E-01
Tnik	10.7	11.7	1.5	6.0E-03	1.5	1.3E-03
Glns-ps1	50.0	41.1	1.5	2.2E-02	1.8	8.1E-03
Cyr61	31.2	23.7	1.4	2.8E-02	2.0	4.2E-02

Gene symbols	4-wk ntg mean	8-wk ntg mean	Fold-change 4-wk TG-16/ntg	p-value	Fold-change 8-wk TG-16/ntg	8-wk p-value
Dgat2	67.8	79.7	1.4	4.1E-03	1.8	7.1E-05
Cd74	68.8	77.8	1.4	6.8E-03	1.9	2.0E-02
Cacnb1	4.0	4.2	1.4	3.1E-02	1.5	1.6E-02
Tia1	16.5	13.5	1.4	7.6E-03	1.2	2.3E-01
Ddx49	19.1	15.6	1.4	2.4E-02	1.3	5.4E-02
Urgcp	18.2	17.5	1.4	9.3E-03	1.3	8.7E-02
Kremen1	37.0	34.0	1.4	7.5E-04	1.6	4.9E-04
Homer2	25.1	22.5	1.4	3.9E-05	1.3	2.0E-02
Prodh	27.2	25.5	1.4	7.3E-03	1.4	9.6E-03
Adprhl1	318.0	276.8	1.4	3.9E-06	1.3	1.2E-02
H2-Ab1	14.4	11.9	1.4	4.5E-02	2.0	3.1E-02
Podn	10.0	10.3	1.4	1.5E-02	1.5	1.1E-02
Cyb5rl	11.8	11.1	1.4	4.2E-02	1.2	1.6E-01
Camk2b	4.7	4.8	1.4	1.1E-02	1.9	3.5E-02
Coq4	20.0	18.2	1.4	3.9E-02	1.3	1.3E-01
Nol6	7.6	6.5	1.4	3.7E-02	1.5	5.0E-02
Odf2	27.7	23.4	1.4	4.7E-02	1.5	1.1E-05
Rps6kc1	8.9	6.8	1.4	2.3E-02	1.2	2.1E-01
Adam11	12.8	10.6	1.4	4.6E-02	1.8	6.1E-03
Icosl	3.6	3.3	1.4	1.6E-02	1.8	7.8E-02
Thrsp	19.0	21.6	1.3	3.5E-02	1.4	2.6E-02
Mrap	10.4	9.2	1.3	1.4E-02	1.2	1.9E-01
Itfg2	3.2	2.7	1.3	1.7E-03	1.2	7.2E-02
Patz1	9.4	9.5	1.3	1.4E-02	1.5	7.2E-04
Wdr6	3.8	3.1	1.3	4.2E-02	1.3	4.6E-02
Dffa	14.7	15.2	1.3	2.2E-02	1.2	2.0E-01
Tmtc1	17.6	19.7	1.3	1.0E-02	1.2	5.4E-02
Cry2	8.9	7.5	1.3	1.9E-02	1.6	3.3E-03
D630003M21Rik	3.5	3.0	1.3	3.0E-02	1.5	1.3E-01
Klhl17	5.3	4.6	1.3	3.4E-03	1.3	6.8E-02
Rab11fip3	13.5	11.7	1.3	3.5E-03	1.5	3.0E-03
E4f1	4.9	4.8	1.3	4.1E-02	1.3	2.0E-02
Golga4	66.4	57.3	1.3	1.5E-04	1.2	4.6E-02
Irak2	7.1	6.0	1.3	4.1E-02	1.4	1.2E-01
Hspb7	683.0	711.8	1.3	1.4E-03	1.3	2.2E-02
Eml3	3.7	3.2	1.3	1.2E-02	1.2	2.4E-02
Cpeb3	28.4	32.2	1.3	1.3E-02	1.2	1.3E-01
Kcnj12	12.6	11.0	1.3	5.7E-03	1.4	1.8E-03
Slc39a13	32.0	26.0	1.3	4.5E-02	1.2	8.9E-02

Gene symbols	4-wk ntg mean	8-wk ntg mean	Fold-change 4-wk TG-16/ntg	p-value	Fold-change 8-wk TG-16/ntg	8-wk p-value
Taf15	15.6	13.6	1.3	2.4E-02	1.2	7.5E-03
Mlxipl	21.0	18.5	1.3	3.7E-02	1.3	6.6E-02
Ptpn11	86.8	73.9	1.2	3.8E-04	1.4	1.5E-03
Ahnak	72.3	85.7	1.2	3.8E-03	1.4	6.2E-04
Flot1	72.2	63.3	1.2	2.0E-03	1.4	6.7E-04
Smarca2	12.2	10.7	1.2	2.7E-02	1.7	1.2E-04
Fem1a	65.4	56.5	1.2	2.1E-03	1.3	2.8E-03
Camk2a	18.0	15.3	1.2	2.9E-02	1.5	5.0E-02
Dmap1	8.6	7.2	1.2	4.6E-02	1.4	2.9E-02
Fmr1	5.4	4.9	1.2	4.1E-02	1.2	3.4E-01
Pdk1	107.9	96.9	1.2	4.1E-02	1.2	2.3E-02
Mn1	4.9	3.9	1.2	9.1E-03	2.8	8.8E-04
4931440F15Rik	9.4	8.6	1.2	3.2E-02	1.4	2.1E-03
Scand3	66.1	79.8	-1.2	2.2E-02	-1.4	1.6E-02
Eif3j	5.1	5.7	-1.2	3.2E-02	-1.3	2.3E-03
Pde4a	37.0	39.4	-1.2	4.7E-02	-1.2	1.0E-01
Mrpl34	81.5	80.7	-1.2	2.5E-03	-1.4	3.0E-04
Tnnt2	5118.2	4918.5	-1.2	7.4E-05	-1.4	3.7E-03
Atp5d	927.1	1069.8	-1.2	2.0E-02	-1.2	9.2E-04
Dnajc19	26.8	19.7	-1.2	2.9E-02	-1.6	1.1E-01
Atp5g3	1339.6	1582.0	-1.3	5.2E-04	-1.3	1.1E-03
Yaf2	5.7	4.8	-1.3	1.3E-02	-1.2	4.9E-03
Pgp	22.5	23.4	-1.3	1.2E-02	-1.2	2.3E-02
Rhot2	60.3	69.6	-1.3	2.5E-03	-1.4	6.2E-04
Mrpl17	27.4	24.2	-1.3	1.5E-02	-1.5	4.8E-05
Snrpn	14.7	12.6	-1.3	5.5E-03	-1.4	3.4E-02
Aes	721.6	860.3	-1.3	2.3E-03	-1.4	3.1E-03
Gm8325	27.7	25.0	-1.3	1.8E-04	-1.3	4.7E-03
Mpst	17.4	18.0	-1.3	3.4E-03	-1.2	1.9E-02
2010107H07Rik	64.4	71.4	-1.3	3.8E-02	-1.7	8.1E-03
Maob	5.6	4.6	-1.3	2.8E-02	-1.2	4.1E-01
Iscu	120.3	109.3	-1.4	3.0E-02	-1.4	7.0E-03
St6galnac6	81.9	76.7	-1.4	8.3E-03	-1.3	2.7E-03
Hhatl	69.7	63.6	-1.4	2.8E-05	-1.2	3.1E-02
Fbxl22	104.1	99.4	-1.4	2.6E-02	-1.4	1.5E-02
Ano10	21.4	20.9	-1.4	5.1E-04	-2.2	7.8E-03
2610528E23Rik	32.0	23.6	-1.4	3.7E-02	-1.4	6.1E-03
Rpia	4.1	3.6	-1.4	4.4E-02	-1.2	6.3E-02
Akt1	46.4	32.1	-1.5	1.6E-03	-1.3	6.4E-03

Gene symbols	4-wk ntg mean	8-wk ntg mean	Fold-change 4-wk TG-16/ntg	p-value	Fold-change 8-wk TG-16/ntg	8-wk p-value
Cish	38.5	45.3	-1.5	4.6E-02	-1.8	1.6E-02
Ppil1	31.0	33.4	-1.5	4.2E-03	-1.8	3.2E-04
Naprt1	8.7	8.8	-1.5	1.7E-03	-1.4	3.0E-02
Wnk4	5.5	4.6	-1.5	8.5E-03	-1.3	2.1E-01
Fbxo44	9.7	11.1	-1.5	6.6E-03	-2.2	1.7E-03
Fhl3	7.1	5.3	-1.6	7.6E-03	-1.2	3.9E-01
Pvrl2	10.9	10.3	-1.6	1.1E-03	-1.2	6.1E-02
Slc25a11	546.1	490.5	-1.6	1.7E-05	-1.5	5.6E-04
Metrn	13.9	16.9	-1.6	1.2E-02	-1.4	8.7E-02
Cadm4	7.3	8.2	-1.7	2.1E-03	-1.5	2.5E-03
Gm5745	6.8	10.1	-1.7	1.8E-02	-1.6	6.8E-02
Slc17a7	7.0	8.2	-1.7	1.4E-03	-3.1	1.6E-03
Spint2	53.0	58.4	-1.8	1.6E-02	-1.9	1.8E-02
Gpt	19.2	15.7	-1.9	3.7E-05	-1.7	2.9E-03
Myl4	4.4	12.5	-1.9	4.5E-02	-8.0	3.5E-01
Scn1b	17.2	16.6	-1.9	5.5E-05	-1.4	2.7E-03
Gsg11	3.8	3.6	-2.0	2.6E-03	-1.8	1.7E-03
Apex2	5.1	5.5	-2.1	4.4E-02	-1.9	4.4E-03
Art5	9.2	7.4	-2.2	1.4E-04	-1.6	2.3E-01
Ramp1	3.6	2.3	-2.3	1.2E-02	-2.1	2.0E-01
Gm11744	3.3	2.1	-2.4	3.4E-02	-1.4	1.6E-01
Gm4532	3.2	4.2	-2.8	4.6E-02	-2.8	6.0E-02
Capn3	9.5	2.8	-10.0	2.8E-02	-1.4	4.7E-01

Online Table III. *Regulated mRNAs in 4 week-old and 8 week-old miR-499 transgenic (TG-16) hearts.* mRNAs regulated in 4 week-old transgenic hearts (line TG-16) by at least 20% ($P < 0.05$), and whose regulation was maintained in 8 week-old hearts are shown, ordered by fold-change in 4 week-old transgenic hearts. mRNA content is shown as copies / cell in nontransgenic hearts, with fold-regulation and p-values for regulation in transgenic hearts.

Online Table IV: Downregulated miR-499 targets, predicted by TargetScan 6, in miR-499 transgenic mouse hearts at 4 weeks of age.

Gene symbol	Gene name	Copies/cell (ntg)	Copies/cell (TG-16)	Copies/cell (TG-53)	Fold-change (TG-16 vs ntg)	p-value (TG-16 vs ntg)	Fold-change (TG-53 vs ntg)	p-value (TG-53 vs ntg)
ARHGAP23 *	Rho GTPase activating protein 23	6.1	4.8	5.8	-1.3	1.3E-02	-1.1	4.2E-01
ARID2	AT rich interactive domain 2 (ARID, RFX-like)	4.4	3.7	4.0	-1.2	2.0E-02	-1.1	1.1E-01
CCRN4L	CCR4 carbon catabolite repression 4-like (<i>S. cerevisiae</i>)	7.7	3.4	3.7	-2.3	3.8E-08	-2.1	7.8E-10
CDH2	cadherin 2, type 1, N-cadherin (neuronal)	56.8	43.5	38.2	-1.3	6.8E-04	-1.5	7.7E-07
CDKN1A	cyclin-dependent kinase inhibitor 1A (p21, Cip1)	18.1	12.0	17.4	-1.5	1.1E-03	-1.0	6.9E-01
DCAF12 *	DDB1 and CUL4 associated factor 12	9.4	8.1	8.4	-1.2	1.6E-03	-1.1	3.1E-03
DNAJB2	DnaJ (Hsp40) homolog, subfamily B, member 2	33.0	25.0	27.2	-1.3	9.5E-06	-1.2	1.6E-03
DYRK2	dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 2	8.0	6.2	6.8	-1.3	2.3E-02	-1.2	6.5E-02
EPM2AIP1	EPM2A (laforin) interacting protein 1	4.0	3.2	2.7	-1.3	1.9E-02	-1.5	1.8E-04
ERRFI1	ERBB receptor feedback inhibitor 1	8.1	5.8	6.6	-1.4	3.9E-02	-1.2	6.4E-02
FAM168A *	family with sequence similarity 168, member A	9.4	8.1	8.5	-1.2	2.8E-02	-1.1	6.0E-02
FNIP2	folliculin interacting protein 2	9.4	6.2	6.7	-1.5	3.3E-03	-1.4	3.0E-03
FZD4	frizzled homolog 4	8.3	6.6	6.9	-1.3	3.9E-02	-1.2	3.1E-02

Gene symbol	Gene name	Copies/cell (ntg)	Copies/cell (TG-16)	Copies/cell (TG-53)	Fold-change (TG-16 vs ntg)	p-value (TG-16 vs ntg)	Fold-change (TG-53 vs ntg)	p-value (TG-53 vs ntg)
	(Drosophila)							
H2AFZ	H2A histone family, member Z	6.6	4.5	4.4	-1.5	7.9E-03	-1.5	7.2E-04
IRS2 *	insulin receptor substrate 2	3.8	2.3	4.8	-1.7	8.9E-03	1.3	6.8E-02
KLHDC5	kelch domain containing 5	1.2	1.0	1.2	-1.2	2.0E-02	-1.0	6.1E-01
LRRC8A *	leucine rich repeat containing 8 family, member A	25.7	17.1	22.7	-1.5	1.9E-02	-1.1	2.4E-01
MAPK6	mitogen-activated protein kinase 6	6.7	5.8	5.4	-1.2	1.2E-02	-1.3	3.8E-04
MAST4	microtubule associated serine/threonine kinase family member 4	6.9	5.6	6.7	-1.2	5.4E-03	-1.0	5.9E-01
PACS2	phosphofurin acidic cluster sorting protein 2	9.4	7.6	9.0	-1.2	1.0E-03	-1.0	3.3E-01
RCN2	reticulocalbin 2, EF- hand calcium binding domain	5.2	2.8	2.4	-1.8	1.3E-03	-2.2	3.3E-06
REEP1	receptor accessory protein 1	3.8	2.3	2.4	-1.7	2.5E-05	-1.6	7.7E-07
RNF44	ring finger protein 44	29.7	21.3	24.8	-1.4	2.8E-02	-1.2	2.5E-01
SLC35A4	solute carrier family 35, member A4	13.3	10.6	9.2	-1.3	8.5E-03	-1.5	1.3E-05
SNRK	SNF related kinase	65.2	50.7	62.1	-1.3	1.1E-02	-1.0	5.0E-01
SOS2 *	son of sevenless homolog 2 (Drosophila)	6.1	4.1	4.0	-1.5	6.1E-05	-1.5	3.4E-07
SOX6	SRY (sex determining	2.1	1.7	1.8	-1.2	4.7E-02	-1.2	2.6E-02

Gene symbol	Gene name	Copies/cell (ntg)	Copies/cell (TG-16)	Copies/cell (TG-53)	Fold-change (TG-16 vs ntg)	p-value (TG-16 vs ntg)	Fold-change (TG-53 vs ntg)	p-value (TG-53 vs ntg)
	region Y)-box 6							
SOX7	SRY (sex determining region Y)-box 7	11.6	8.5	9.8	-1.4	1.9E-03	-1.2	2.6E-02
TBC1D4	TBC1 domain family, member 4	16.4	11.6	8.8	-1.4	3.8E-03	-1.9	8.7E-07
TMBIM6 *	transmembrane BAX inhibitor motif containing 6	74.7	53.8	51.8	-1.4	4.0E-03	-1.4	2.0E-04
USO1	USO1 homolog, vesicle docking protein (yeast)	8.0	6.3	6.1	-1.3	1.5E-02	-1.3	4.8E-04

Online Table IV. Downregulated miR-499 targets, predicted by TargetScan 6, in miR-499 transgenic mouse hearts at 4 weeks of age. mRNA expression was assessed by RNA-sequencing and analyzed using Partek Genomics Suite. A heatmap corresponding to these data is given in **Figure 3c**. Downregulation was defined as a fold-change <-1.2, P<0.05. Seven mRNAs also downregulated in human heart failure (see **Online Table I** and **Figure 1c**) are designated with asterisks.

Online Table V: Direct cardiac targets of miR-499 identified by RISC-sequencing.

#	Gene symbol	Gene name	Fold miR-499 TG-16 / ntg	miR-499 TG-16 vs ntg, p-value
1	Rprd1b	regulation of nuclear pre-mRNA domain containing 1B	2.5	8.0E-02
2	Rora	RAR-related orphan receptor A	2.5	2.8E-02
3	Alg10b	asparagine-linked glycosylation 10 homolog B (yeast, alpha-1,2-glucosyltransferase)	2.2	7.7E-05
4	Slc5a3	solute carrier family 5 (inositol transporters), member 3	1.9	6.6E-04
5	Slc38a1	solute carrier family 38, member 1	1.8	1.0E-02
6	Pcdh19	protocadherin 19	1.7	1.3E-02
7	Rmnd5a	required for meiotic nuclear division 5 homolog A (<i>S. cerevisiae</i>)	1.7	8.4E-03
8	Lin7c	lin-7 homolog C (<i>C. elegans</i>)	1.7	2.5E-02
9	Ddi2	DDI1, DNA-damage inducible 1, homolog 2 (<i>S. cerevisiae</i>)	1.7	1.0E-02
10	Uhmk1	U2AF homology motif (UHM) kinase 1	1.7	6.2E-03
11	Klhdc5	kelch domain containing 5	1.7	1.0E-02
12	Gna13	guanine nucleotide binding protein (G protein), alpha 13	1.7	6.5E-03
13	Dcp2	DCP2 decapping enzyme homolog (<i>S. cerevisiae</i>)	1.7	8.7E-02
14	Tet3	tet oncogene family member 3	1.7	2.3E-02
15	Dnajb14	DnaJ (Hsp40) homolog, subfamily B, member 14	1.7	1.0E-02
16	Map3k2	mitogen-activated protein kinase kinase kinase 2	1.7	3.0E-02
17	Cpne3	copine III	1.7	2.0E-02
18	Sacm11	SAC1 suppressor of actin mutations 1-like (yeast)	1.6	5.9E-02
19	Atxn1l	ataxin 1-like	1.6	5.5E-02
20	Ppargc1b	peroxisome proliferator-activated receptor gamma, coactivator 1 beta	1.6	3.7E-02
21	Taok1	TAO kinase 1	1.6	3.4E-04
22	Hipk2	homeodomain interacting protein kinase 2	1.5	9.2E-02
23	Ccnj	cyclin J	1.5	3.3E-02
24	Stk35	serine/threonine kinase 35	1.5	1.5E-02
25	Otud4	OTU domain containing 4	1.5	3.0E-02
26	Slc7a2	solute carrier family 7 (cationic amino acid transporter, y system), member 2	1.5	2.5E-02
27	Mgat4a	mannosyl (alpha-1,3-)-glycoprotein beta-1,4-N-acetylglucosaminyltransferase, isozyme A	1.5	4.6E-02
28	Prrc1	proline-rich coiled-coil 1	1.5	1.0E-02
29	Skil	SKI-like oncogene	1.5	3.0E-03
30	Fam168b	family with sequence similarity 168, member B	1.5	2.7E-02
31	Trp53inp1	transformation related protein 53 inducible nuclear protein 1	1.5	8.6E-02

#	Gene symbol	Gene name	Fold miR-499 TG-16 / ntg	miR-499 TG-16 vs ntg, p-value
32	Ppp1r3b	protein phosphatase 1, regulatory (inhibitor) subunit 3B	1.5	5.0E-02
33	Hipk1	homeodomain interacting protein kinase 1	1.5	5.1E-02
34	Rbpj	recombination signal binding protein for immunoglobulin kappa J region	1.5	2.7E-02
35	Cnot6	CCR4-NOT transcription complex, subunit 6	1.4	2.6E-02
36	Foxo1	forkhead box O1	1.4	3.1E-02
37	Lfng	LFNG O-fucosylpeptide 3-beta-N- acetylglucosaminyltransferase	1.4	4.0E-02
38	Mobk11a	MOB1, Mps One Binder kinase activator-like 1A (yeast)	1.4	5.1E-02
39	Zbtb39	zinc finger and BTB domain containing 39	1.4	2.5E-02
40	Fam116a	family with sequence similarity 116, member A	1.4	6.6E-02
41	Zbtb10	zinc finger and BTB domain containing 10	1.4	3.5E-02
42	Socs4	suppressor of cytokine signaling 4	1.4	2.4E-02
43	Dicer1	dicer 1, ribonuclease type III	1.4	2.2E-02
44	Pppde1	PPPDE peptidase domain containing 1	1.4	1.6E-02
45	Blcap	bladder cancer associated protein	1.4	3.6E-02
46	Arid2	AT rich interactive domain 2 (ARID, RFX-like)	1.4	2.7E-02
47	Sox6	SRY (sex determining region Y)-box 6	1.4	2.3E-02
48	Tb11xr1	transducin (beta)-like 1 X-linked receptor 1	1.4	6.7E-02
49	Phip	pleckstrin homology domain interacting protein	1.4	8.1E-02
50	Bmpr2	bone morphogenetic protein receptor, type II (serine/threonine kinase)	1.4	2.0E-02
51	Ugcg	UDP-glucose ceramide glucosyltransferase	1.3	1.3E-02
52	Lrrc58	leucine rich repeat containing 58	1.3	2.5E-02
53	Ptplad1	protein tyrosine phosphatase-like A domain containing 1	1.3	1.3E-02
54	Bmp2k	BMP2 inducible kinase	1.3	8.5E-02
55	Prox1	prospero homeobox 1	1.3	4.1E-02
56	Mmd	monocyte to macrophage differentiation- associated	1.3	6.7E-02
57	Slc25a37	solute carrier family 25, member 37	1.3	4.5E-02
58	Rad54l2	RAD54-like 2 (<i>S. cerevisiae</i>)	1.3	4.0E-02
59	Cntn2	contactin 2 (axonal)	1.3	4.5E-02
60	Fam160b1	family with sequence similarity 160, member B1	1.3	2.1E-02
61	Jhdm1d	jumonji C domain containing histone demethylase 1 homolog D (<i>S. cerevisiae</i>)	1.3	7.6E-02
62	Zfp280b	zinc finger protein 280B	1.3	8.1E-02
63	Cmtm4	CKLF-like MARVEL transmembrane domain containing 4	1.3	2.9E-02

#	Gene symbol	Gene name	Fold miR-499 TG-16 / ntg	miR-499 TG-16 vs ntg, p-value
64	<i>Rffl</i>	ring finger and FYVE like domain containing protein	5.3	
65	<i>Med13 (Thrap1)</i>	mediator complex subunit 13	3.2	
66	<i>Fbxo3</i>	F-box protein 3	3.1	
67	<i>Gnao1</i>	guanine nucleotide binding protein, alpha O	3.1	

Online Table V. *Direct cardiac targets of miR-499 identified by RISC-sequencing.* RISC enrichment (RISC RNA / transcriptome RNA) for miR-499 hearts. Threshold for RISC enrichment was 2-fold enrichment in nontransgenic hearts, and then a further ≥ 1.25 -fold in transgenic compared to nontransgenic, $P < 0.05$. 63 mRNAs were RISC-enriched in miR-499 hearts using these criteria. *Italicized entries:* 4 ‘newly-enriched’ RISC-targeted mRNAs included on the basis of ≥ 3 -fold enrichment in RISC RNA, but not transcriptome RNA, compared to nontransgenic hearts. # numbers refer to mRNA positions, from left to right, in **Figure 5a** and **5b**.

Online Table VI: Expression levels of downregulated mRNAs directly targeted by miR-499.

Gene symbol	ntg copies/cell	miR-499 copies/cell	Fold-change, miR-499 vs ntg	p-value
Alg10b	4.6	3.8	-1.2	6.7E-04
Arid2	6.3	5.4	-1.2	2.4E-02
Cpne3	4.2	3.6	-1.2	1.7E-03
Dcp2	3.6	2.7	-1.3	1.1E-02
Ddi2	8.9	7.5	-1.2	1.9E-03
Dicer1	5.5	4.7	-1.2	4.5E-03
Fam116a	3.0	2.3	-1.3	4.7E-04
Foxo1	8.3	7.3	-1.1	1.5E-04
Hipk2	46.3	40.7	-1.1	1.9E-03
Jhdm1d	4.4	3.8	-1.2	1.1E-02
Lrrc58	11.3	8.6	-1.3	7.5E-04
Otud4	8.8	7.3	-1.2	2.1E-04
Rmnd5a	21.1	19.8	-1.1	6.3E-03
Sox6	3.1	2.5	-1.2	4.6E-03
Taok1	6.8	6.4	-1.1	2.2E-02
Uhmk1	4.6	3.9	-1.2	5.2E-03

Online Table VI. *Expression levels of downregulated RNAs directly targeted by miR-499.* Sixteen of the 67 direct mRNA targets of miR-499 were significantly downregulated at the transcriptome level. mRNA content is shown as copies / cell in nontransgenic and miR-499 TG-16 transgenic hearts. mRNAs were classified as downregulated in response to miR-499 at P<0.025 compared to nontransgenic, fold-change ≤ -1.1 .

Online Table VII is supplied as an Excel workbook.

Online Table VII. *Pathway enrichment analysis of mRNAs regulated in response to miR-499 overexpression or pressure-overload.* MetaCore¹¹ was used to classify mRNAs regulated by at least 20%, FDR < 0.05 (P<0.005) into GeneGo proprietary pathway and process networks, as well as into Gene Ontology process categories.

Online Table VIII is supplied as an Excel workbook.

Online Table VIII. *Protein function enrichment of mRNAs regulated in response to miR-499 overexpression or pressure-overload.* mRNAs regulated by at least 20%, FDR < 0.05 (P<0.005) were subjected to MetaCore 11 protein functional enrichment analysis, separating mRNAs into classes of kinases, phosphatases, enzymes, transcription factors, proteases, ligands and receptors.

Online Table IX. Regulation of kinase-coding mRNAs by miR-499 overexpression and by pressure overload.

Gene symbol	Gene name	miR-499-specific regulation		regulated by miR-499 and pressure-overload	
		Signal (miR-499 TG-16 vs nontransgenic)	p-value	Signal (miR-499 TG-16 vs nontransgenic)	p-value
Upregulated					
Map3k6	mitogen-activated protein kinase kinase kinase 6			3.6	1.2E-04
Pak3	p21 protein (Cdc42/Rac)-activated kinase 3			3.1	4.2E-04
Map4k2	mitogen-activated protein kinase kinase kinase kinase 2	2.6	4.8E-05		
Nuak1	NUAK family, SNF1-like kinase, 1			2.5	1.1E-04
Mylk4	myosin light chain kinase family, member 4	1.8	3.8E-03		
Sik1	salt inducible kinase 1			1.8	4.9E-03
Prkd1	protein kinase D1			1.7	3.0E-04
Dapk1	death associated protein kinase 1	1.6	2.6E-06		
Aak1	AP2 associated kinase 1	1.5	7.6E-03		
Alpk2	alpha-kinase 2			1.5	7.4E-05
Hipk1	homeodomain interacting protein kinase 1	1.5	4.9E-05		
Mknk2	MAP kinase-interacting serine/threonine kinase 2	1.5	1.3E-04		
Map2k3	mitogen-activated protein kinase kinase 3	1.5	1.5E-03		
Prkd2	protein kinase D2	1.4	1.6E-03		
Jak1	Janus kinase 1	1.4	6.6E-04		
Cdc42bpb	CDC42 binding protein kinase beta			1.4	4.2E-04
Dyrk1a	dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 1a	1.4	7.5E-03		

Gene symbol	Gene name	miR-499-specific regulation		regulated by miR-499 and pressure-overload	
		Signal (miR-499 TG-16 vs nontransgenic)	p-value	Signal (miR-499 TG-16 vs nontransgenic)	p-value
Syk	spleen tyrosine kinase	1.4	2.2E-03		
Snrk	SNF related kinase	1.4	1.1E-04		
Fyn	Fyn proto-oncogene	1.4	4.1E-03		
Mink1	misshapen-like kinase 1 (zebrafish)	1.4	4.2E-03		
Mapk14	mitogen-activated protein kinase 14 (p38 α)	1.4	6.9E-03		
Cdc42bpa	CDC42 binding protein kinase alpha	1.4	3.9E-04		
Map3k5	mitogen-activated protein kinase kinase kinase 5	1.4	3.9E-03		
Prkcd	protein kinase C, delta			1.4	3.8E-03
Sbk1	SH3-binding kinase 1	1.3	3.2E-03		
Map4k4	mitogen-activated protein kinase kinase kinase kinase 4	1.3	1.6E-03		
Prkce	protein kinase C, epsilon	1.3	1.3E-03		
Mylk	myosin, light polypeptide kinase			1.3	1.6E-03
Dyrk1b	dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 1b	1.3	6.1E-03		
Rps6kb2	ribosomal protein S6 kinase, polypeptide 2	1.3	5.6E-03		
Mark4	MAP/microtubule affinity-regulating kinase 4	1.3	5.1E-04		
Clk2	CDC-like kinase 2	1.2	5.0E-03		
Downregulated					
Akt1	thymoma viral proto-oncogene 1	-1.3	6.4E-03		
Camk1	calcium/calmodulin-dependent protein kinase I	-1.4	4.4E-03		

Gene symbol	Gene name	miR-499-specific regulation		regulated by miR-499 and pressure-overload	
		Signal (miR-499 TG-16 vs nontransgenic)	p-value	Signal (miR-499 TG-16 vs nontransgenic)	p-value
Camk2d	calcium/calmodulin-dependent protein kinase II, delta	-1.4	4.4E-03		
Nme2	non-metastatic cells 2, protein (NM23B) expressed in	-1.4	1.1E-03		
Mapkapk5	MAP kinase-activated protein kinase 5	-1.7	3.2E-04		
Phkg1	phosphorylase kinase gamma 1			-2.3	1.6E-04
Other Akt-related kinases and phosphatases					
Pik3c2b	phosphoinositide-3-kinase, class 2, beta polypeptide	1.1	3.5E-01		
Pik3r1	phosphatidylinositol 3-kinase, regulatory subunit, polypeptide 1 (p85 alpha)	1.4	9.2E-06		
Pik3r2	phosphatidylinositol 3-kinase, regulatory subunit, polypeptide 2 (p85 beta)	1.2	2.8E-01		
Pdpk1	3-phosphoinositide dependent protein kinase 1	1.2	1.4E-01		
Pten	phosphatase and tensin homolog	1.2	9.3E-03		
Ppp2ca	protein phosphatase 2 (formerly 2A), catalytic subunit, alpha isoform	1.1	3.1E-01		
Ppp2cb	protein phosphatase 2 (formerly 2A), catalytic subunit, beta isoform			1.2	6.7E-02
Phlpp1	PH domain and leucine rich repeat protein phosphatase 1	1.2	6.0E-02		

Online Table IX. Regulation of kinase-coding mRNAs by miR-499 overexpression and by pressure overload. Upregulated and downregulated kinase mRNAs were defined at a 20% fold-change, FDR<0.05 (P<0.005). While not all of the Akt-related kinases and phosphatases meet these criteria, they are shown for illustrative purposes.

Online Table X: Identification and regulation of proteins in miR-499 transgenic hearts.

Spot #	TG 1/ Ntg 1	TG 2/ Ntg 2	TG 3 / Ntg 3	TG 4 / Ntg 4	TG 5 / Ntg 5	Regulation?	Gene symbol
1	-1.1	-1.1	1.05	-1.22	-1.26		Mybpc3
2	1.44	2.05	-1.27	1.2	1.25	up	Mybpc3
3	-1.36	-2.59	-1.29	-1.42	-1.41	down	Serpina3k
4	-1.36	1.03	1.08	-1.37	-1.05		Hspa8
5	-1.22	1.52	1.06	-1.07	1.02		Anxa6
6	-1.17	-1.3	-1.26	-1.22	-1.22	down	Acss1
7	1.69	1.55	2.28	2.04	2.11	up	Ehd4
8	1.02	1.14	1.4	1.31	1.36		Acadv1
11	-1.16	-1.34	1.14	-2.82	-2.49	down	P4hb
12	-1.03	-1.4	1.02	-1.85	-1.76	down	P4hb
13	-1.04	-1.27	-1.25	-1.28	-1.25	down	Fgb
14	-1.38	-2.26	-1.17	-1.35	-1.41	down	Aldh4a1
16	-1.41	-1.17	-1.08	-1.59	-1.48	down	Atp5a1
17	1.02	-2.02	-1.46	1.22	1.17		Vim
18	-1.26	-1.29	-1.16	-1.43	-1.18	down	Oat
19	-1.12	-1.13	-1.19	-1.08	-1.09		Srl
20	-1.64	-1.82	-1.26	-1.46	-1.43	down	Pdha1
21	-1.42	-1.44	-1.09	-1.35	-1.28	down	Eno3
22	-1.54	-1.73	-1.43	-1.81	-1.58	down	Ckmt2
23	-1.44	-1.46	-1.14	-1.53	-1.51	down	Ckmt2
24	-1.29	-1.21	-1.57	-1.48	-1.54	down	Eca40 (Bcat2)
25	-1.57	-1.37	-1.01	-1.4	-1.4	down	Mdh2
26	-1.41	-1.26	-1.06	-1.27	-1.45	down	Ldhb
27	-1.64	-1.25	-1.12	-1.27	1	down	Mdh1
28	-1.45	-1.29	-1.1	-1.38	-1.1	down	Mdh1
29	-1.33	-1.2	-1.19	-1.6	-1.33	down	Ech1
31	-1.66	-1.44	1.03	-1.35	-1.26	down	Ldha
32	-2.12	-2.66	-1.39	-1.28	-1.29	down	Myoz2
34	1.03	2.53	1.14	-1.02	-1.08		Vdac2
35	-1.23	-1.32	-1.04	-1.24	1.03	down	Tnni3
36	1.04	1.1	1.15	1.15	1.18	up	Ywhag
41	-1.84	-1.4	-1.45	-1.23	-1.24	down	Gstm2
43	-1.66	1.01	-1.36	-1.74	-1.63	down	Mb
44	3.22	-1.03	2.16	3.24	3.7	up	Ndufaf2
45	7.43	5.44	3.85	10.08	5.43	up	Gm9234
48	-1.14	2.08	1.06	1.79	1.04		Cox5a
54	1.26	1.25	1.02	-1.14	1.11		Hspb7
55	1.72	1.11	-1.22	-1.2	-1.23	down	Pdha1
56	1.22	1.17	1.67	1.49	1.43	up	Sdha
57	1.18	1.19	1.45	-1.09	1	up	uk (unknown)
58	-1.25	-1.21	1.05	-1.03	-1.1	down	Eno1
60	1.2	1.29	1.16	1.23	1.29	up	Adprh1
62	-1.01	1.67	-1.73	-1.31	-1.31	down	Myl4
64	-1.79	1.39	-2.07	-1.11	-1.14		Myl7
66	-1.05	-1.37	1.28	-1.2	-1.18	down	Trf

Spot #	TG 1/ Ntg 1	TG 2/ Ntg 2	TG 3 / Ntg 3	TG 4/ Ntg 4	TG 5 / Ntg 5	Regulation?	Gene symbol
67	-1.28	-1.36	1.36	1.11	1.34		Hspb8
68	-1.14	-1.33	1.39	1.31	1.35	up	Hspb9

Online Table X. *Identification and regulation of proteins in miR-499 transgenic hearts.* Multiple spots were manually matched across 5 2D DiGE gels, representing 5 different nontransgenic and miR-499 transgenic hearts. 31 were identified using mass spectroscopy, while a further 15 were identified by comparison to previously published proteomic maps. Density ratios between miR-499 transgenic and nontransgenic hearts are shown for each protein. Proteins were classified as up- or downregulated if changes of at least 20% in the same direction were identified on at least 3 of 5 gels.

Online Table XI: Identification and regulation of proteins in nontransgenic and miR-499 transgenic hearts subjected to TAC.

Spot#	ntg						miR-499 TG						Gene symbol
	TAC 1 / sham 1	TAC 2 / sham 2	TAC 3 / sham 3	TAC 4 / sham 4	TAC 5 / sham 5	Regulation?	TAC 1 / sham 1	TAC 2 / sham 2	TAC 3 / sham 3	TAC 4 / sham 4	TAC 5 / sham 5	Regulation?	
1		-1.65	-1.42	1.45			-1.39	-1.29		-1.38	-1.54	down	Mybpc3
2	1.44	2.14	2.06	-2.86	-2.58	up	2.06	1.42	1.18	-1.07	-1.04	up	Mybpc3
3	1.06	1.87	2.43	3.13		up	-1.68	-1.04	2.22		2.62		Serpina3k
4	1.15	1.23	1.15	1.86	1.28	up	-1.18	-1.12	1.31		1.55		Hspa8
5	-1.15	-1.87	-1.62	1.09	-1.31	down	-1.74	1.05	-1.28	1.11	1.15		Anxa6
6	1.02	-1.06	1.19	1.08	1.04		1.07	1.06	-1.19	1.12	1.03		Acss1
7	-1.04	-1.03	-1.48	1.12	-1.08		-1.02	-1.63	1.52	2.07	2.28	up	Ehd4
8	-1.56	-1.22	-1.05	-1	1.03		-1.04	-1.28	1.04	1.52	1.38		Acadvl
11	1.03	1.48	2.01	3.67		up	1.06	-1.31	3.11	2.18			Pd4b
12	1.04	1.3	2.56	3.32	2.81	up	1.13	-1.57	2.51	2.34	2.41	up	Pd4b
13	-1.14	1.17	1.95	1.51	1.57	up	-1.53	-1.86	1.24	1.28	1.38	up	Fgb
14	1.51	1.94	2.03	2.03	1.48	up	-1.11	-1.24	1.38	1.29	1.37	up	Aldh4a1
16	1.24	1.16	-1.02	1.04	1		1.02	1.11	1.12	-1.15	-1.15		Atp5a1
17	-1.47	1.26		1.18	1.46	up	1	-4.02		1.09	-1.18		Vim
18	1.52	1.36	-1.01	1.11	1.3		-1.17	1.09	1.06	1.13	1.06		Oat
19	1.06	-1.07	1.2	1.16	1.04		1.24	1.09	-1.13	1.1	1.13		Srl
20	2.38	2.76	-1.35	1.97	2.08	up	1.21	4.48	3.54	1.91	1.68	up	Pdha1
21	1.55	1.2	1.26	1.42	1.17	up	1.15	1.17	1.42	1.27	1.26	up	Eno3
22	-1.09	-1.11	1.14	1.18	1.03		1.07	1.29	1.19	-1.44	-1.26		Ckmt2
23	1.33	1.21	1.07	1.16	1.05	up	1.03	1.35	1.05	-1.15	-1.25		Ckmt2
24	-1.08	-1.01	-1.24	-1.03	-1.25		1.05	1.05	-1.3	-1.14	-1.2		Eca40
25	1.47	1.23	1.13	1.39	1.25	up	1.06	-1	1.39	1.18	1.13		Mdh2
26	1.35	1.15	1.24	1.34	1.12		1.01	1.12	1.35	1.03	1.14		Ldhb
27	1.56	1.05	1.14	1.11	1.21	up	-1.06	1.03	-1.02	1.2	1.32		Mdh1
28	1.43	1.11	1.04	1.18	1.15		-1.05	1.14	1.03	1.09	1.12		Mdh1

Spot#	ntg						miR-499 TG						Gene symbol
	TAC 1 / sham 1	TAC 2 / sham 2	TAC 3 / sham 3	TAC 4 / sham 4	TAC 5 / sham 5	Regulation?	TAC 1 / sham 1	TAC 2 / sham 2	TAC 3 / sham 3	TAC 4 / sham 4	TAC 5 / sham 5	Regulation?	
29	1.28	1.35	1.04	1.09	1.09		1.09	1.43	-1.07	-1.45	-1.57		Ech1
31	1.33	1.29	1.18	1.39	1.18	up	-1.03	-1.1	1.64	1.44	1.76	up	Ldha
32	1.06	2.01	-1.04	1.1	1.74		1.22	-1.11	1.87	1.93	2.01	up	Myoz2
34	-2.82	-3.12	-2.28	-3.24	-2.49	down	-1.78	1.3	-2.34	-1.68	-1.74	down	Vdac2
35	-1.01	1.02		-1.22			-1.19	1.12	1.05		1.05		Tnni3
36	-1.31	-1.02	-1.2	-1.17	1.09	down	-1.26	-1.05		1.1			Ywhag
41	1.05	-1.08	1.07	1.14	1.19		1.06	-1.05	-1.21	1.31	1.08		Gstm2
43	1.16	1.51	1.91	3.16	3.49	up	2.41	-1.34	1.34	2.49		up	Mb
44	-1.38	-1.32	-2.17	-2.06	-2.62	down	-1.32	-2.5					Ndufaf2
45	1.62	-2.89	-1.4	1.23	-3.88	down	2.82	-3.14	-1.05	1.32	1.23	up	Gm9234
48	1.07	-1.2	-1.47	-1.51	-2.19	down	-1.03	1.19	-1.29	-1.36	-1.05	down	Cox5a
54	1.15	1.08	-2.35	-2.55	-2.35	down	-1.2	1.05	-1.38	-2.17	-2.34	down	Hspb7
55	-2.19	-1.97	-1.14	-1.48	-1.22	down	1.19	-1.51	-2	-1.67	-1.68	down	Pdha1
56													Sdha
58													Eno1
60													Adprh1
62	1.08	-1.22	3.1	6.59	6.77	up	-1.28	-1.79	1.21	7.39	8.16	up	Myl4
64	1.27	-1.05	2.26	8.99		up	-2.01	-2.71	1.08	26.57	35.28		Myl7
66	1.36	1.41	1.62	2.45	2.52	up	-1.19	-1.37	1.94	2.52	2.55	up	Trf
67	1.51	1.96	1.9	2.69	2.64	up	1.12	-1.37	2.17	3.12	3.59	up	Hspb8
68	1.53	1.69	1.61	2.34	2.71	up	1.16	-1.23	1.83	3.27	3.54	up	Hspb9

Online Table XI. Identification and regulation of proteins in nontransgenic and miR-499 transgenic hearts subjected to 1 week of pressure overload. Multiple spots were manually matched across 2D DiGE gels, representing 10 different nontransgenic hearts (5 sham, 5 TAC) and 10 different miR-499 transgenic hearts (5 sham, 5 TAC). One sham and one TAC heart were loaded on each gel. 31 spots were identified using mass spectroscopy, while a further 15 were identified by comparison to previously published proteomic maps. In-gel density ratios between TAC and sham treatments are shown. Proteins were classified as up- or downregulated if changes of at least 20% in the same direction were identified on at least 3 of 5 gels.

Online Table XII. Quantitation of mRNA, total protein and phosphoprotein for differentially phosphorylated proteins.

mRNA or protein	Fold-change (499TG-sham vs. ntg-sham)	p-value	Fold-change (ntg-TAC vs. ntg-sham)	p-value	Fold-change (499TG-TAC vs. 499TG-sham)	p-value	ntg-sham (copies/cell or protein density)	ntg-TAC (copies/cell or protein density)	499TG-sham (copies/cell or protein density)	499TG-TAC (copies/cell or protein density)
<i>Hsp90aa1</i>	1.1	0.32	1.5	3.3E-03	1.1	0.33	52	75	58	64
<i>Hsp90ab1</i>	1.0	0.53	1.1	0.32	1.0	0.85	478	507	496	502
Phospho Hsp90β, spot 1	1.5	>0.05	-1.5	>0.05	-2.2	>0.05	3.9	2.6	5.8	2.6
Phospho Hsp90β, spot 2	3.2	0.01	1.2	>0.05	-2.9	0.01	0.75	0.92	2.4	0.80
Total Hsp90β	1.1	>0.05	1.1	0.22	1.1	0.36				
<i>Ppp1ca</i>	-1.1	0.03	-1.1	0.27	-1.03	0.59	60	57	53	52
Phospho-PP1α	NaN	NaN	-2.5	0.006	3.7	0.01	0.16	0.065	0	0.24
Total PP1α	Spot not measurable									
<i>Pgm2</i>	1.1	0.13	1.0	0.89	-1.1	0.34	34	34	37	35
Phospho	-1.2	>0.05	-1.6	0.017	1.1	>0.05	0.56	0.34	0.48	0.53
Total Pgm2	Spot not measurable									
<i>Ndufa10</i>	-1.1	0.1	-1.2	0.014	-1.1	0.12	133	113	120	109
Phospho	2.3	0.004	1.3	0.28	2.9	0.003	0.08	0.10	0.18	0.30
Total Ndufa10	Spot not measurable									
<i>Gpd1</i>	1.1	0.57	-1.2	0.27	-1.1	0.56	17	15	19	17
Phospho	7.0	0.002	7.7	0.02	-2.4	0.18	0.07	0.52	0.47	0.22
Total Gpd1	Spot not measurable									
<i>Ehd2</i>	1.3	0.11	1.3	0.08	-1.03	0.86	30	39	40	39
Phospho	1.2	>0.05	-1.5	>0.05	1.1	>0.05	0.36	0.25	0.43	0.47
Total Ehd2	Spot not measurable									

Online Table XII. Quantitation of mRNA, total protein and phosphoprotein for differentially phosphorylated proteins. mRNA levels are expressed in copies/cell. Protein density is spot signal relative to internal phosphoprotein standard (**Online Figure V**). Phospho-protein refers to

signal intensity from ProQ-Diamond phosphoprotein staining; total protein is measured from Cy5 dye-labeling of 150 µg ventricular homogenate. The mRNA *Hsp90aa1* corresponds to the Hsp90α protein (not measured in proteomic studies) while *Hsp90ab1* corresponds to the Hsp90β protein. Other regulated phosphoproteins detected by mass spectrometry were *Ppp1ca* (PP1α), *Pgm2* (phosphoglucomutase-2), *Ndufa10* (NADH dehydrogenase (ubiquinone) 1 alpha subcomplex 10), and *Gpd1* (glycerol-3-phosphate dehydrogenase 1); *Ehd2* (EH domain-containing 2) represents an unregulated phosphoprotein. Bold text denotes significant changes. NaN, not a number (division by zero). Total protein was not measurable for several proteins due to higher sensitivity of phosphoprotein staining compared to Cy5 dye-labeling.

Online Table XIII: Comparison of previously reported *in vivo* miR-499 targets with transcriptome regulation in the current study.

Transcriptomes	Reported regulation	Fold-change (miR-499 TG-16 vs ntg, 4 wk)	p-value	Fold-change (miR-499 TG-53 vs ntg, 4 wk)	p-value	Fold-change (miR-499 TG-16 vs ntg, 8 wk)	p-value	Copies/cell (ntg, 4 wk)	Copies/cell (ntg, 8 wk)
Nat Med 2011¹²									
Ppp3ca	miR-499 target	1.1	0.26	1.2	0.12	1.5	1.9E-03	10.5	6.0
Ppp3cb	miR-499 target	-1.1	0.58	-1.3	0.01	-1.1	0.39	3.9	1.9
Dnm11	Phosphorylation status via calcineurin	-1.1	0.03	-1.3	4.4E-04	1.0	0.95	4.8	3.3
PLoS One 2011¹³									
Pah *	2.5	11.1	1.4E-05	12.3	1.6E-06	6.3	2.9E-03	0.6	1.1
Nppb	2.0	1.2	0.37	1.6	2.6E-03	1.1	0.77	143.3	38.5
Olfr872	1.9							zero FPKM, all samples	
Ppargc1a	1.9	-1.9	2.7E-05	-1.8	4.0E-05	1.6	0.03	24.5	7.7
Alcam	1.9							zero FPKM or FPKM<1	
Myh7 *	1.9	1.1	0.03	1.2	2.9E-05	2.5	0.02	475.8	359.5
Gck *	1.8	2.2	2.3E-07	4.2	1.8E-15	2.9	1.4E-03	5.0	3.8
Acta1 *	1.6	2.6	1.7E-04	4.0	1.5E-08	2.4	0.01	165.8	191.4
Rnf144b	1.6	-1.2	0.10	1.1	0.41	2.7	5.1E-03	3.8	1.2
Egr1	-2.5	-1.0	0.96	1.5	0.02	1.7	0.59	2.5	5.2
Fos	-1.9	1.1	0.92	2.0	0.04	1.9	0.44	0.8	1.6

Transcriptomes	Reported regulation	Fold-change (miR-499 TG-16 vs ntg, 4 wk)	p-value	Fold-change (miR-499 TG-53 vs ntg, 4 wk)	p-value	Fold-change (miR-499 TG-16 vs ntg, 8 wk)	p-value	Copies/cell (ntg, 4 wk)	Copies/cell (ntg, 8 wk)
Egr2	-1.6							zero FPKM or FPKM<1	
Ly75	-1.6							zero FPKM or FPKM<1	
Myl4	-1.6							zero FPKM or FPKM<3	
Myl7	-1.6							zero FPKM, all samples	
Pfkfb1 *	-1.5	1.3	0.43	-4.3	0.04	-2.1	9.2E-03	2.0	2.1
Flvcr2	-1.5							zero FPKM or FPKM<1	
Wif1	-1.5							zero FPKM or FPKM<3	
Cysltr2	-1.5							zero FPKM, all samples	

* designates P<0.05 regulation in the same direction, as in previous studies, in at least 2 of 3 cohorts from the current RNA-sequencing data.

Online Table XIII. Comparison of previously reported *in vivo* miR-499 targets with transcriptome regulation in the current study. Of the three miR-499 targets *Ppp3ca* and *Ppp3cb* (CnA α and β subunits) and *Dnm1l* (dynamin-like protein 1) reported by Wang et al.¹², only *Dnm1l* was downregulated at the transcriptome level in both miR-499 TG-16 and TG-53 transgenic mice in our studies. *Ppp3cb* was downregulated only in the higher-expressing miR-499 transgenic line. While 19 mRNAs were reported by Shieh et al.¹³ to be up- or down-regulated in miR-499 wild-type transgenic mice using microarray technology, RNA sequencing was unable to detect 9 of these (FPKM of 0, or <3). Similar regulation in RNA sequencing studies was observed for 5 of the 19 mRNAs.

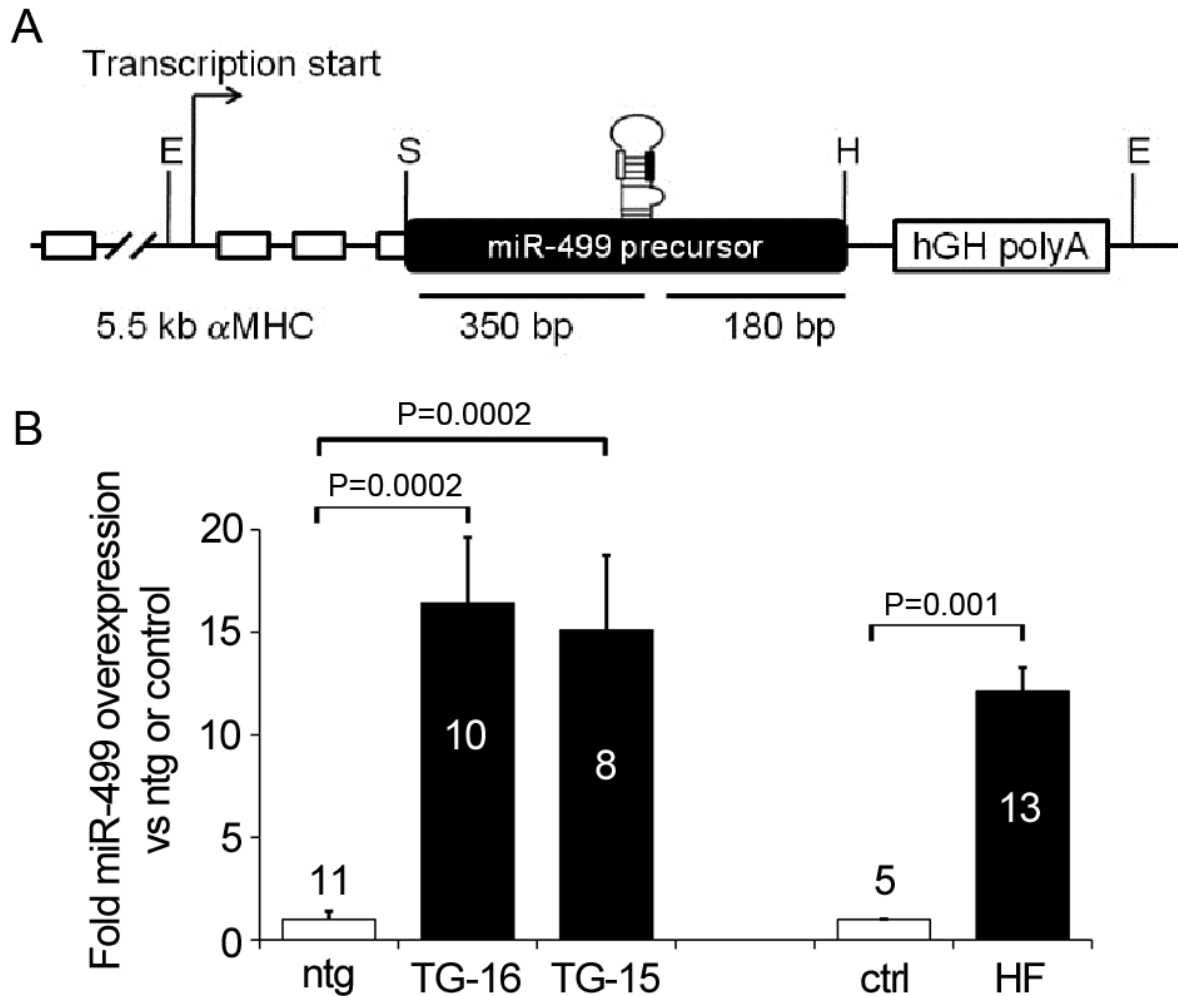
Online Table XIV: Comparison of previously reported *in vivo* miR-499 targets with RISCome regulation in the current study.

RISComes	Fold-change (miR-499 TG-16 vs ntg)	p-value (miR-499 TG-16 vs ntg)	Fold-change (miR-499 TG- 53 vs ntg)	p-value (miR-499 TG-53 vs ntg)	Mean FPKM (ntg)	Mean FPKM (miR-499 TG-16)	Mean FPKM (miR-499 TG-53)
Nat Med 2011¹²							
Ppp3ca	-1.7	0.12	1.1	0.66	82	49	90
Ppp3cb	-1.0	0.99	1.1	0.74	41	41	44
Dnm11	-1.3	0.20	-1.1	0.62	50	39	46
PLoS One 2011¹³							
Pah (up in transcriptome)					0	21	28
Nppb (up in transcriptome)	-1.2	0.61	1.9	2.5E-03	577	490	1117
Olfr872 (up in transcriptome)					zero FPKM, all samples		
Ppargc1a (up in transcriptome)	-1.6	0.06	-1.6	0.04	244	148	150
Alcam (up in transcriptome)					zero FPKM, all samples		
Myh7 (up in transcriptome)	1.0	0.83	1.0	0.98	3531	3632	3540
Gck (up in transcriptome)	2.2	0.03	4.1	4.4E-06	33	72	137
Acta1 (up in transcriptome)	2.3	0.01	3.4	8.5E-06	574	1300	1978
Rnf144b (up in transcriptome)	-1.1	0.63	1.1	0.72	51	45	55
Egr1 (down in transcriptome)	-1.0	0.93	1.4	0.11	62	60	86
Fos (down in transcriptome)					detected in 3 hearts only		
Egr2 (down in transcriptome)					detected in 2 hearts only		
Ly75 (down in transcriptome)					zero FPKM, all samples		
Myl4 (down in transcriptome)	-2.3	0.56	1.7	0.41	265	116	456

RISComes	Fold-change (miR-499 TG-16 vs ntg)	p-value (miR-499 TG-16 vs ntg)	Fold-change (miR-499 TG- 53 vs ntg)	p-value (miR-499 TG-53 vs ntg)	Mean FPKM (ntg)	Mean FPKM (miR-499 TG-16)	Mean FPKM (miR-499 TG-53)
My17 (down in transcriptome)					not in >50% of samples		
Pfkfb1 (down in transcriptome)					detected in 5 hearts only		
Flvcr2 (down in transcriptome)					detected in 1 heart only		
Wif1 (down in transcriptome)					detected in 3 hearts only		
Cysltr2 (down in transcriptome)					zero FPKM, all samples		

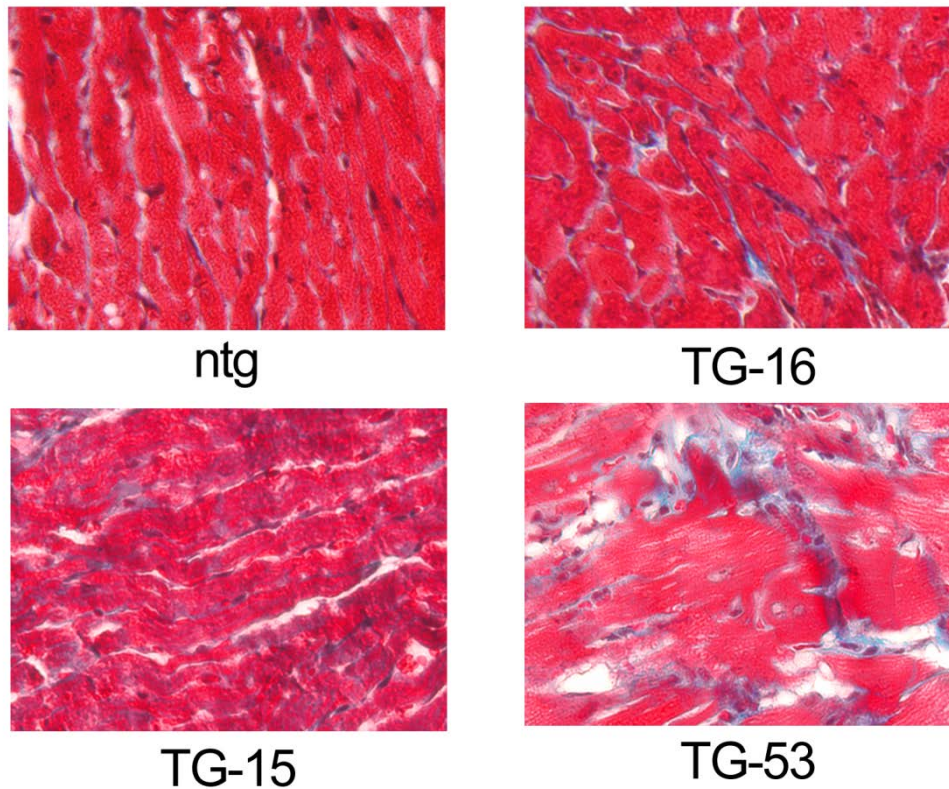
Online Table XIV. Comparison of previously reported *in vivo* miR-499 targets with RISCome regulation in the current study. None of the proposed miR-499 targets from previous studies showed significant upregulation in the RISComes of miR-499 transgenic hearts at 4 weeks of age.

Online Figure I: miR-499 levels in transgenic mice and human heart failure.



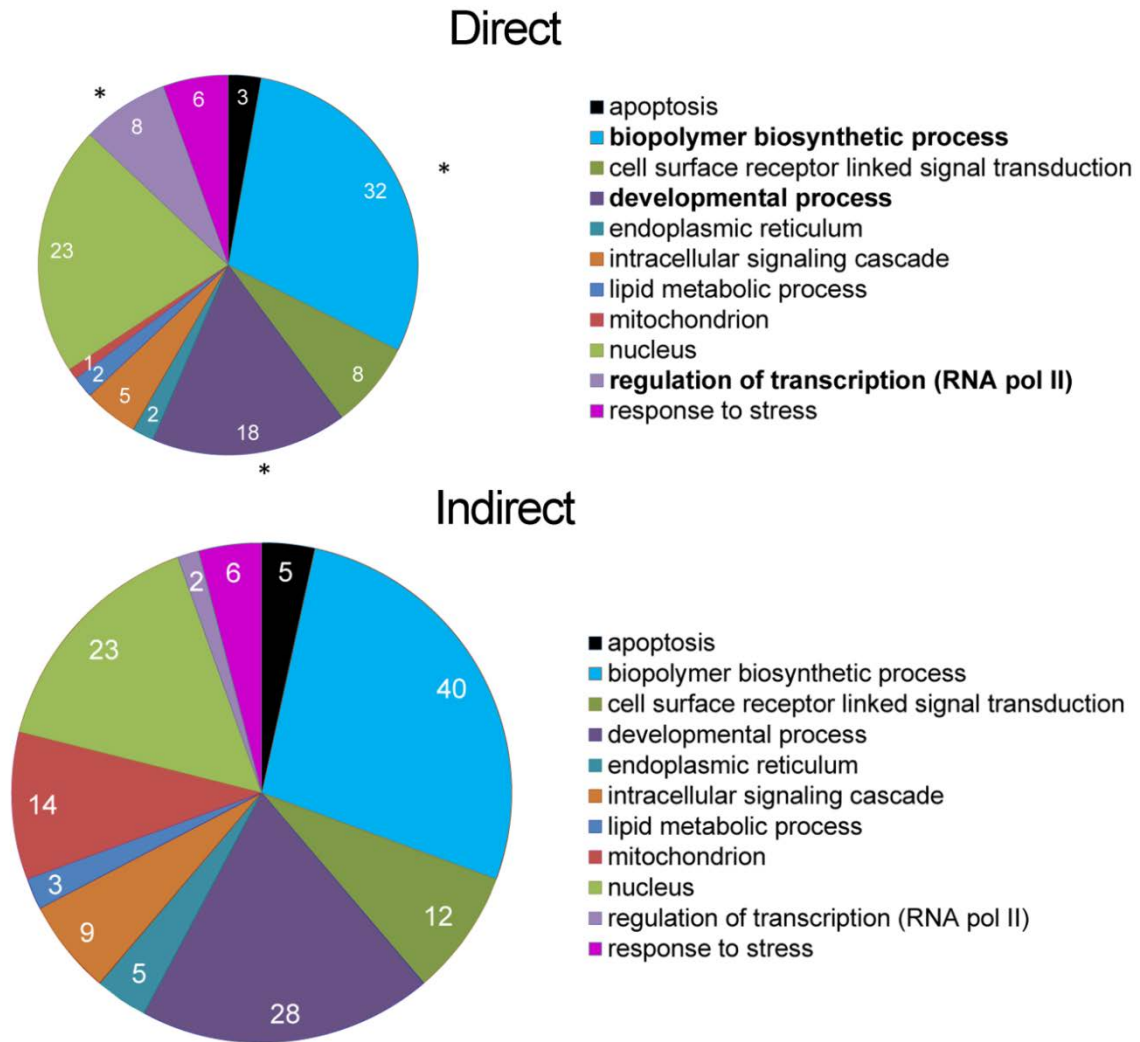
Online Figure I. *miR-499* programming of transgenic mouse hearts. **A.** A 530 bp mouse genomic fragment including the miR-499 precursor region was cloned into exon 3 of the murine *Myh6*/ α MHC transgenic expression vector. E = EcoRI, S = SalI, H = HindIII restriction sites. **B.** miR-499 expression relative to 5S rRNA was assayed using the NCode miR RT-qPCR system (Invitrogen). Relative miR-499 levels in two lines of transgenic mouse hearts (numbers in bars show number of individual hearts) compared to failing human hearts are shown.

Online Figure II: Fibrosis detection in miR-499 transgenic mouse hearts.



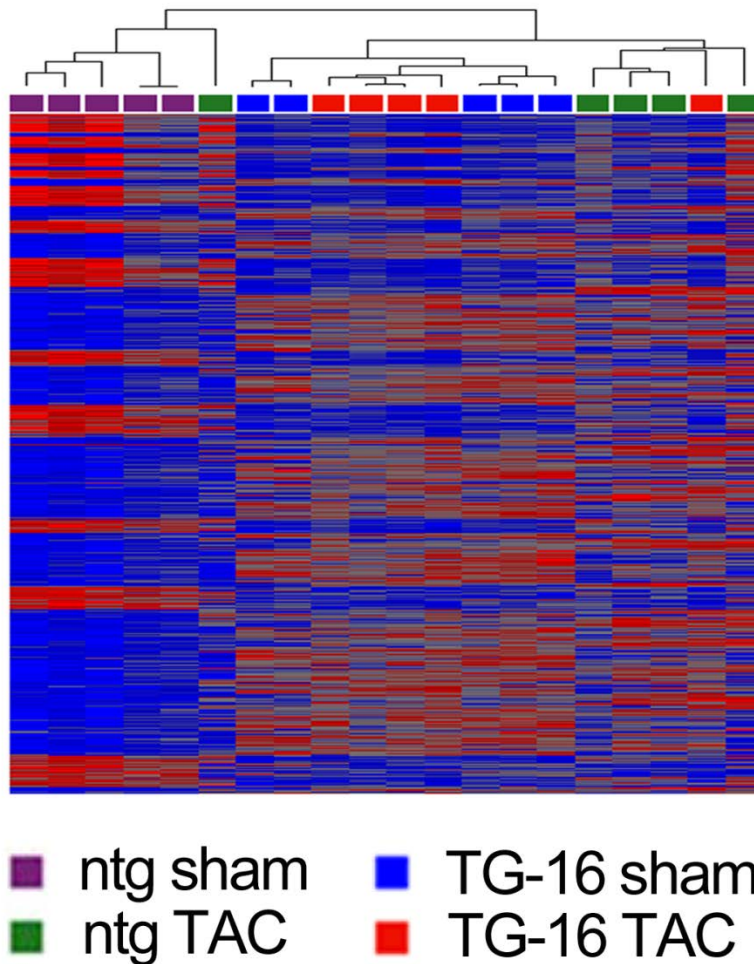
Online Figure II. *Fibrosis detection in miR-499 transgenic mouse hearts.* Masson's trichrome was used to stain formalin-fixed, paraffin-embedded sections of nontransgenic and miR-499 transgenic mouse hearts. Transgenic line designations correspond to miR-499 levels shown in **Figure 2a**.

Online Figure III: Gene Ontology classification of direct and indirect targets of miR-499.



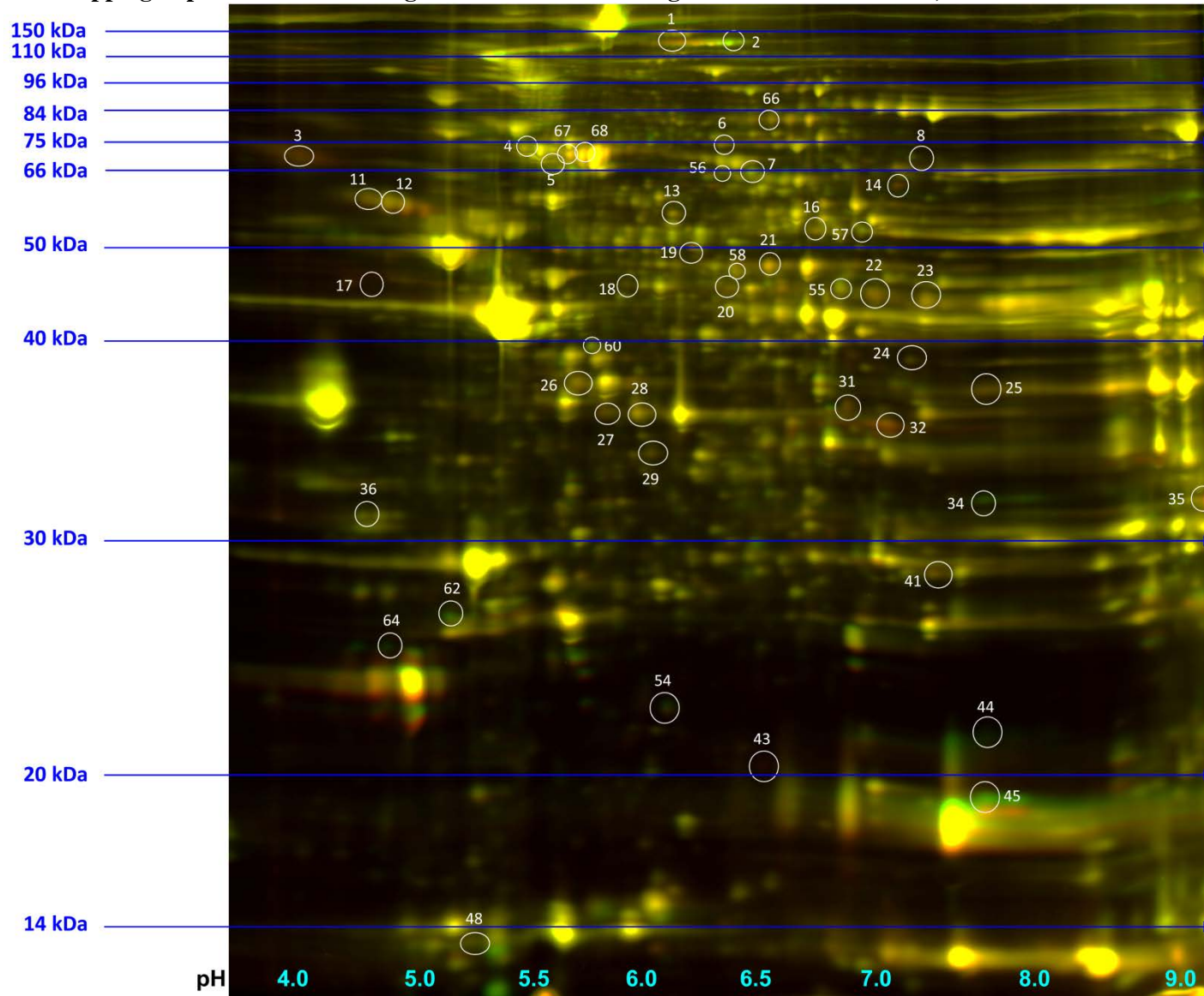
Online Figure III. Gene Ontology classification of direct and indirect targets of miR-499. Gene Ontology classification of 67 miR-499 direct mRNA targets (upper) and 136 indirect mRNA targets, regulated both at 4 weeks and 8 weeks of age (lower). Asterisks and bold type indicates over-represented categories⁹.

Online Figure IV. Unsupervised clustering of miR-499 and TAC-regulated mRNAs.

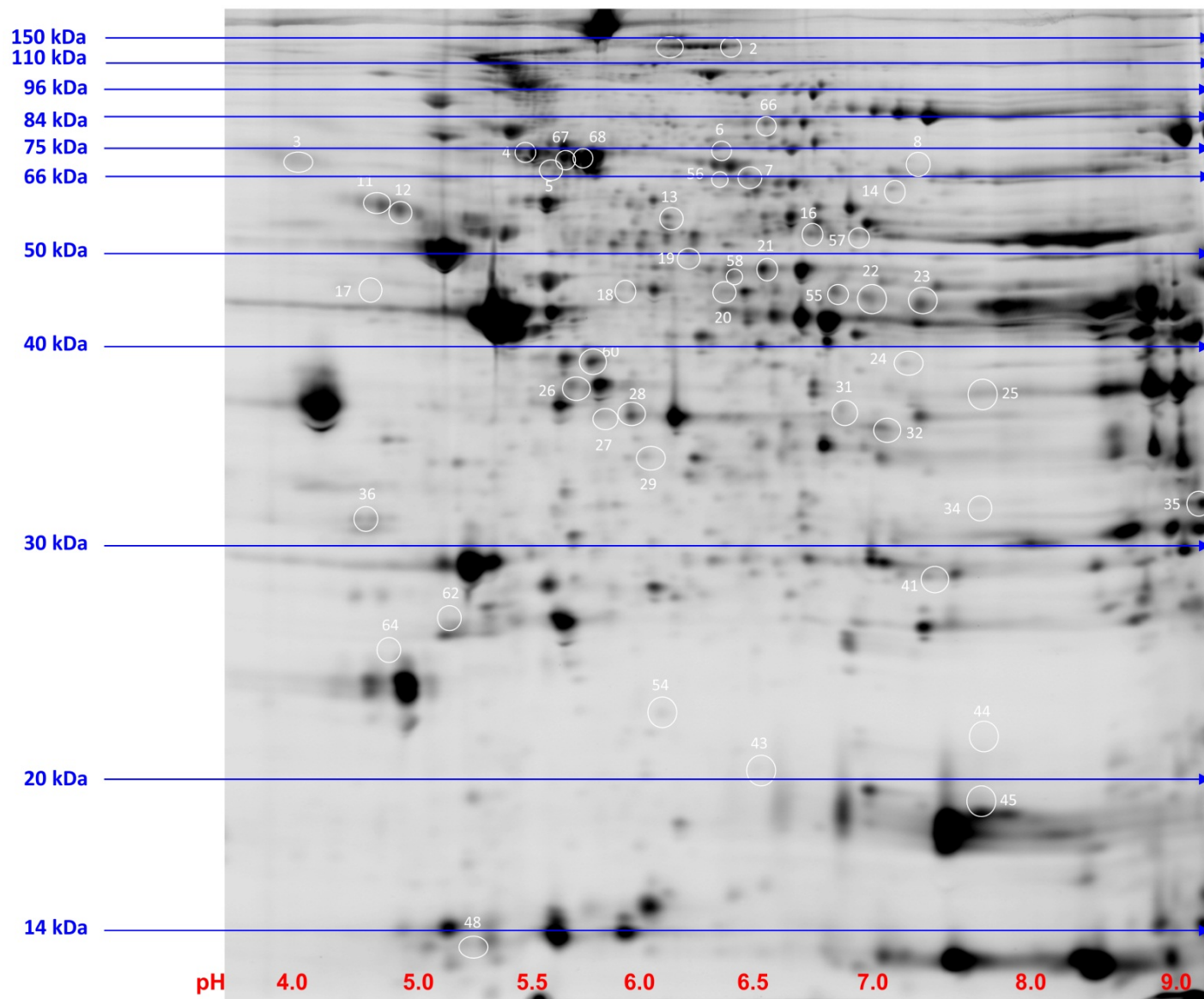


Online Figure IV. *Unsupervised clustering of miR-499 and TAC-regulated mRNAs.* Unsupervised hierarchical clustering of 1,168 miR-499-regulated and TAC-regulated mRNAs was performed with Euclidean distance and average linkage (Partek Genomics Suite). Red indicates higher expression, blue indicates lower expression.

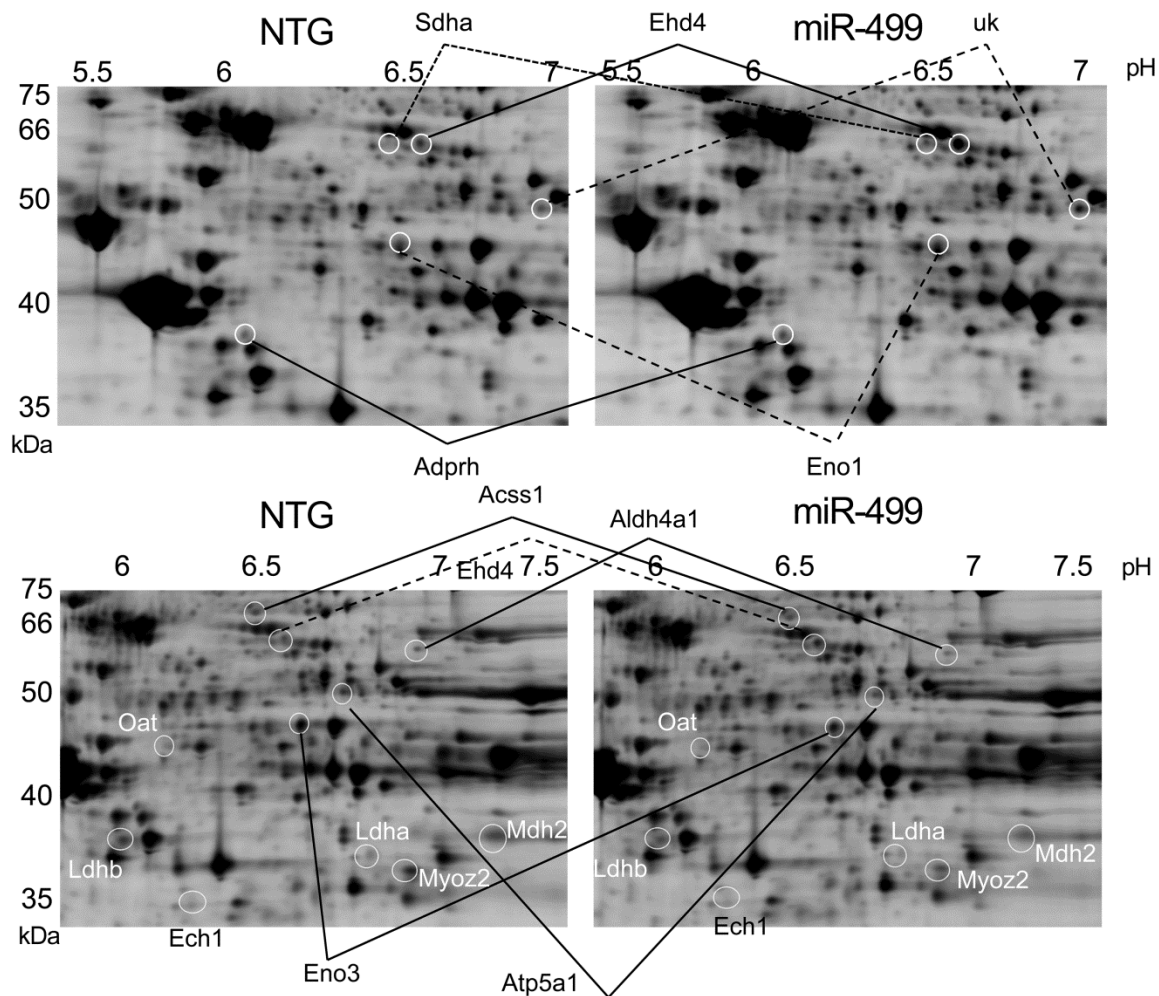
Online Figure V: Mapping of proteins in nontransgenic and miR-499 transgenic hearts on 2D DiGE, Panel A



Online Figure V, Panel B

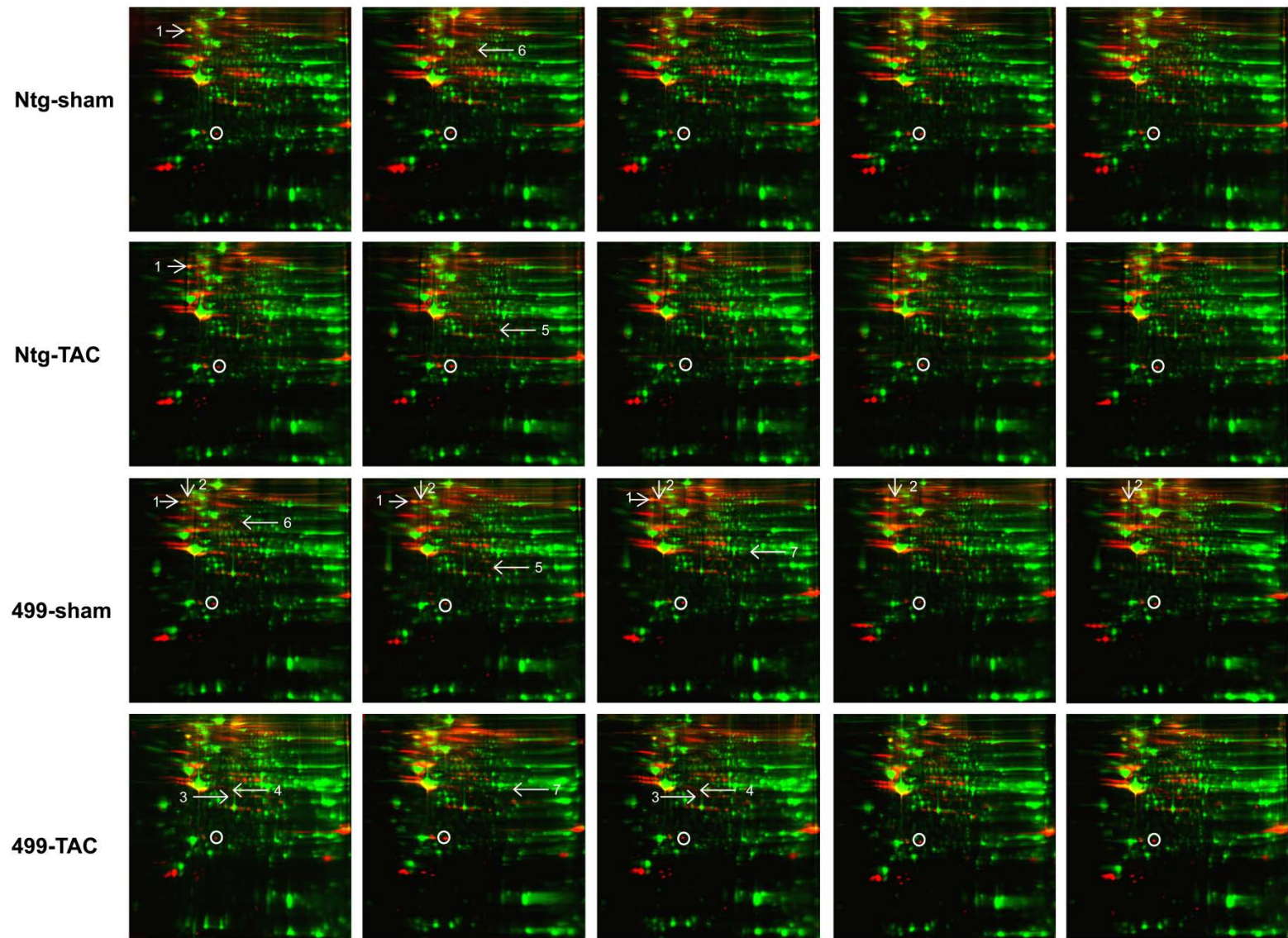


Online Figure V, Panel C



Online Figure V: Mapping of proteins in nontransgenic and miR-499 transgenic hearts on 2D DiGE. The 46 protein spots identified by mass spectrometry (n=31) or mapping to previous proteome maps (n=15), whose regulation in 5 separate pairwise comparisons between nontransgenic and miR-499 transgenic hearts is shown in **Online Table X**, are shown on a representative 2D gel. Numbers correspond to protein identifications in **Online Table X**. **Panel A:** nontransgenic protein is shown in red, while miR-499 transgenic protein is shown in green. **Panel B:** only nontransgenic proteins, from the same gel as in the upper panel, are shown. **Panel C:** Magnified views of representative 2D DiGE gel sections, showing spot comparison for a number of proteins whose expression was compared between nontransgenic and transgenic hearts.

Online Figure VI: Phosphoprotein and total protein analysis by 2D DiGE.



Online Figure VI: Phosphoprotein and total protein analysis by 2D DiGE. Five hearts from each treatment group (nontransgenic sham, nontransgenic pressure overload, miR-499 transgenic sham and miR-499 transgenic pressure overload) were analyzed using Cy5 dye staining (green) and ProQ-Diamond phosphoprotein staining (red) on the same gel. All 20 gels are shown. Spots labeled '1' and '2' correspond to phospho-HSP90 β spots; '3', phospho-PP1 α ; '4', phospho-Ndufa10; '5', phospho-Gpd1; '6', phospho-Pgm2; '7', phospho-Ehd2, quantitated in **Online Table XII**. Circled spot: internal phosphoprotein standard.

References

1. Matkovich SJ, Van Booven DJ, Youker KA, Torre-Amione G, Diwan A, Eschenbacher WH, Dorn LE, Watson MA, Margulies KB, Dorn GW, 2nd. Reciprocal regulation of myocardial microRNAs and messenger RNA in human cardiomyopathy and reversal of the microRNA signature by biomechanical support. *Circulation*. 2009;119:1263-1271.
2. Matkovich SJ, Wang W, Tu Y, Eschenbacher WH, Dorn LE, Condorelli G, Diwan A, Nerbonne JM, Dorn GW, 2nd. MicroRNA-133a protects against myocardial fibrosis and modulates electrical repolarization without affecting hypertrophy in pressure-overloaded adult hearts. *Circ Res*. 2010;106:166-175.
3. Matkovich SJ, Zhang Y, Van Booven D, Dorn GW, 2nd. Deep mRNA sequencing for in vivo functional analysis of cardiac transcriptional regulators. Application to *Gαq*. *Circ Res*. 2010;106:1459-1467.
4. Matkovich SJ, Van Booven DJ, Eschenbacher WH, Dorn GW, 2nd. RISC RNA sequencing for context-specific identification of in vivo microRNA targets. *Circ Res*. 2011;108:18-26.
5. Karginov FV, Conaco C, Xuan Z, Schmidt BH, Parker JS, Mandel G, Hannon GJ. A biochemical approach to identifying microRNA targets. *Proc Natl Acad Sci USA*. 2007;104:19291-19296.
6. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics*. 2009;25:1105-1111.
7. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L. Transcript assembly and abundance estimation from RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol*. 2010;28:511-515.
8. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods*. 2008;5:621-628.
9. Maere S, Heymans K, Kuiper M. BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics*. 2005;21:3448-3449.
10. Varani G, McClain WH. The G x U wobble base pair. A fundamental building block of RNA structure crucial to RNA function in diverse biological systems. *EMBO Rep*. 2000;1:18-23.
11. GeneGo. MetaCore. Available at: <http://www.genego.com>. Accessed 4/2/12.
12. Wang JX, Jiao JQ, Li Q, Long B, Wang K, Liu JP, Li YR, Li PF. miR-499 regulates mitochondrial dynamics by targeting calcineurin and dynamin-related protein-1. *Nat Med*. 2011;17:71-78.
13. Shieh JT, Huang Y, Gilmore J, Srivastava D. Elevated miR-499 levels blunt the cardiac stress response. *PLoS One*. 2011;6:e19481.

Appendix on RNA-sequencing

RNA-sequencing bench and analytic procedures, Matkovich+Dorn laboratories, February 2012

SAMPLE PREPARATION FOR STANDARD RNA-SEQUENCING.....	47
Initial RNA extraction	47
polyA+ RNA enrichment	47
RNA fragmentation	48
cDNA synthesis.....	48
Sequencing library preparation.....	49
ANALYSIS OF RNASEQ DATA, AND DIFFERENTIAL GENE EXPRESSION	53
Package differences for differential gene expression and statistics.....	53
Links to open-source software packages (Linux installation).....	54
Analysis procedure (tophat / cufflinks)	55
Sample bash script for RNASeq / RISCSeq analysis, giving FPKM based on all reads mapped to the <u>genome</u> and FPKM based on a 'masked' gtf (e.g. that does not include rRNAs)...	57
Sample bash scripts for RNASeq / RISCSeq analysis, giving FPKM based on all reads mapped to the <u>transcriptome</u>	59
Analysis procedure (HTSeq / DESeq / edgeR)	63
Supplemental commands: sequence sorting by index, removal of indexes, and fastq generation.....	65

Sample preparation for standard RNA-sequencing

Dorn lab, November 2009 – October 2011 (SJM)

Bench procedures are based on the following published method:

Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods. 2008;5:621-628.

Initial RNA extraction

Use Trizol (Invitrogen), following the manufacturer's directions precisely. If homogenizing tissue, keep tissue frozen (in liquid nitrogen), wash down homogenization probe with a few hundred μ L of Trizol first, add e.g. 0.5 mL Trizol directly to tube, and homogenize. For mouse heart apices, I use a total of 1 mL Trizol; the second aliquot of 0.5 mL is used to help wash tissue pieces out of the probe. This results in good extraction of the total RNA fraction. To enhance recovery of the small RNA fraction (not needed for mRNA-Seq, but useful to have for e.g. miR quantitation), increase the length of the isopropanol precipitation step at room temperature from the recommended 10 minutes to 30 minutes (advice from Invitrogen, and LC Sciences, Houston, TX).

After air-drying RNA pellet, resuspend in 30-50 μ L DEPC-treated H₂O and heat at 65 C for 5 min (as recommended by Trizol protocol).

Until completion of cDNA synthesis from RNA, use appropriate RNase-free technique at all stages!

Quantitate RNA using a standard spectrophotometer or NanoDrop.

Integrity MUST be analyzed too – e.g. 1% TBE agarose gel, using RNA dissolved in a formaldehyde- or formamide-containing solution. As an example of how to do this on native agarose (without having to deal with formaldehyde-agarose), see Novagen's RNA analysis technical bulletin TB108, p. 2-3 (this info is drawn from Maniatis et al molecular cloning manuals). Agilent BioAnalyzer chips are another option.

polyA+ RNA enrichment

1. Take e.g. 5-10 μ g total RNA, and perform 2 rounds of binding to oligo(dT)-resin.
2. I have had good success with the Dynabeads mRNA purification kit (Invitrogen #610-06), which requires the use of a Dynal magnet to permit washing of oligo(dT) beads. Use 100 μ L of beads per 5 μ g total RNA sample. This quantity of Dynabeads has a binding capacity of 1 μ g of (polyA+)RNA.
3. Follow the protocol given with the kit to perform the first round of isolation. Keep the supernatant (after initial binding of RNA to beads), and after washing, elute polyA+ RNA in 15 μ L 10 mM Tris, pH 8.2.

4. Re-prepare Dynabeads for binding by washing once in 1x 'binding buffer'; it will be necessary to prepare a dilution of the 2x buffer supplied with the kit.
5. Incubate the supernatant preserved from the initial binding step at 65 C, 2' and then chill on ice. Perform an additional round of RNA binding to beads.
6. Wash beads and elute further polyA+ RNA in another 15 uL 10 mM Tris, and add to the RNA eluted in the first isolation.
7. Quantitate polyA+ RNA using NanoDrop.

From time to time Invitrogen has had availability problems with the Dynabeads mRNA system. An alternative which has worked well for me is the 'Magnetic mRNA isolation kit' from New England Biolabs, #S1550S; the magnetic beads are compatible with the Invitrogen Dynal magnet, and the principle is much the same.

RNA fragmentation

This step 'shatters' full-length RNA transcripts into 200-500 nt fragments, permitting uniform reverse transcription during subsequent stages. The Mortazavi et al. paper documents full-length RNA+oligo(dT) reverse transcription vs fragment RNA+random primer reverse transcription, and concludes that fragmentation with random primer RT improves coverage.

1. Take 100-500 ng polyA+ RNA, and prepare a solution of 20 uL in 'fragmentation buffer' (40 mM Tris acetate, 100 mM K acetate, 30 mM Mg acetate, DEPC H₂O, pH 8.2). Recipe for 5x fragmentation buffer, 50 mL: 1.21 g Tris base, 2.45 g K acetate, 1.61 g Mg acetate tetrahydrate. Adjust pH with 1 N acetic acid, and keep aliquots frozen at -20 C.
2. Heat at 94 C, 2 min 30 s, then chill on ice.
3. Purify the fragmented RNA from the acetate buffer using e.g. Zymo RNA Clean+Concentrate columns, eluting in 8-9 uL for maximum input into Invitrogen SuperScript III 1st-strand cDNA reaction.

cDNA synthesis

1st-strand: Invitrogen SuperScript III 1st-strand synthesis system, #18080-051 (50 reactions). Each reaction can accommodate 8 uL of fragmented polyA+ RNA. Follow the directions supplied, but do not treat with supplied RNase H at the end of this procedure.

1. Combine 8 uL RNA, 1 uL 50 ng/uL **random hexamers**, 1 uL 10 mM dNTP mix.
2. Incubate at 65 C for 5 min, then place on ice for 1 min.
3. Add 10 uL 'cDNA synthesis mix', prepared as follows:

10x supplied RT buffer	2 uL
25 mM MgCl ₂	4 uL
0.1 M DTT	2 uL
RNaseOUT (40 U/uL)	1 uL
SuperScript III RT (200 U/uL)	1 uL
4. Incubate 10 min at 25 C, 50 min at 50 C, 5 min at 85 C, and chill on ice.

2nd-strand: To the 20 uL 1st-strand reaction, add on ice:

DEPC-treated H ₂ O	91 uL
5x 2 nd -strand reaction buffer	30 uL
10 mM dNTP mix	3 uL
E. coli DNA ligase (10 U/uL)	1 uL
E. coli DNA polymerase I (10 U/uL)	4 uL
E. coli RNase H (2 U/uL)	1 uL

(5x 2nd-strand reaction buffer has the following composition: 100 mM Tris.HCl pH 6.9, 450 mM KCl, 23 mM MgCl₂, 0.75 mM β-NAD⁺, 50 mM (NH₄)₂SO₄)

Vortex gently to mix, and incubate for 2 h at 16 C. Do not allow the temperature to rise above 16 C. Place tubes on ice and freeze at -20 C, or move directly to Qiagen column purification.

[An alternative to the above 1st- and 2nd-strand methods is to use Stratagene's Just cDNA synthesis kit (#200453) for up to 25 uL of fragmented RNA, and their directions with regard to time and temperature of enzymatic incubations, with the following modifications:

- Relax secondary structure at 65 C, 5 min, followed by chilling on ice, prior to adding RNA to the 1st-strand synthesis components.
- Use random primers for 1st-strand synthesis, not oligo(dT).
- I have not used radioisotopic analysis as suggested by Stratagene's protocol to monitor the efficiency of cDNA synthesis, although I have performed agarose gel electrophoresis using Stratagene's 5 ug polyA+ RNA 'test' sample supplied with the kit.
- The kit is somewhat expensive (\$60/sample) and the above recipe with 'home-made' buffer and enzymes represents a much lower cost per sample (~\$30).]

After completion of 2nd-strand synthesis, purify cDNA on Qiagen columns, eluting in 34 uL 10 mM Tris pH 8.5, and then use the EndIt end-repair kit from Epicentre Biotechnologies (details in separate protocol for preparation of Illumina sequencing libraries) to blunt cDNA fragments. Proceed with the remainder of library synthesis.

Sequencing library preparation

Continue with preparation of Illumina sequencing libraries as per Dorn lab protocol (October 2011).

For 'standard' mRNA work, it is likely that Illumina indexing adapters will be desired and 8-16 indexed samples will be combined into 1 sequencing lane of ~180 million raw 42-nt reads (indexes are read separately and do not detract from the 42-nt available for sequence alignment). As of Oct 2011 we can gain ~180 million raw 42-nt reads on 1 lane of Illumina HiSeq for \$1300.

The protocol below for sequencing library generation is based on the 'ChIP sequencing protocol' available at <http://bioinfo.mbb.yale.edu/array/resources.html> (Yale Center for Excellence in Genomic Science).

STEP 1: End Repair

Use 'End-It DNA End Repair Kit' from Epicentre Biotechnologies, Cat# ER0720

a) Combine and mix the following components in a microfuge tube

1-34 μ l DNA to be end-repaired (i.e. however much DNA was isolated)

5 μ l 10X End-Repair Buffer

5 μ l 2.5 mM dNTP Mix

5 μ l 10 mM ATP

x μ l sterile water to bring reaction volume to 49 μ l

1 μ l End-Repair Enzyme Mix

50 μ l Total reaction volume

b) Incubate at **room temperature** for 45 minutes.

c) Purify on one QIAquick column using the QIAquick PCR Purification Kit and protocol, eluting in 34 μ l of EB.

STEP 2: Addition of 'A' base to 3' Ends

Use Klenow (3'→5' exo-) from NEB Cat# M0212S

**Before starting, make up stocks of 1 mM dATP using NEB 100 mM dATP,

e.g. add 5 μ l of 100 mM dATP to 495 μ l Qiagen Buffer EB (10 mM Tris, pH 8.5); then make 50 μ l aliquots and freeze at -20C. Defrost aliquots only once.

a) Combine and mix the following components:

DNA from Step 1 34 μ l

Klenow buffer = NEB Buffer 2 5 μ l

1 mM dATP (will have to make this up) 10 μ l

Klenow fragment (3' to 5' exo minus) 1 μ l

50 μ l Total reaction volume

b) Incubate for 30 min at 37 °C.

c) Purify on one QIAquick MinElute column using the MinElute PCR Purification Kit and protocol. Elute in 10 μ l EB.

STEP 3: Adapter ligation

Use LigaFast from Promega Cat#M8221. Adapter mix is a pre-annealed blend of Illumina-Index-AdapterA and 5'-phosphorylated Illumina-Index-AdapterB (sequences can be requested from Illumina).

Annealed adapter mix is typically at 50 μ M (50 starting DNA = x g / (660 g mol⁻¹ bp⁻¹ * y bp), where bp = typical length of fragments (200-300 bp). For e.g. 1 pmol starting ds cDNA, plan to add 10 pmol annealed adapter mix (10-fold molar excess).

a) Combine and mix the following components in a microfuge tube

DNA from Step 2 10 μ l

DNA ligase buffer 15 μ l

RNase, DNase-free water 2 μ l

Adapter oligo mix >see note above 1 μ l

DNA ligase 2 μ l

30 μ l Total reaction volume

b) Incubate for 15 min at room temperature.

c) Purify on one Qiaquick MinElute column using the MinElute PCR Purification Kit and protocol. Elute in 10 μ l EB.

d) Gel-purify the DNA (using a 2% low-melting agarose gel, such as that formed using Amresco Agarose II) by cutting a gel slice that does NOT include any DNA from an adapter-adapter band migrating at \sim 120 bp. The intended ligation product may not be readily visible at this stage if the DNA input was small (0.5 μ g or less). Isolate DNA in the 150-300 bp range. If making multiple libraries be very careful to avoid cross-contamination by leaving at least two empty lanes between each sample.

e) Purify the DNA from the agarose slice using a QIAGEN Gel Extraction Kit. Perform the steps listed as 'optional' in the Qiagen booklet by adding 1 gel volume of isopropanol to the melted agarose, and an additional wash in 0.5 mL buffer QG. Elute in 30 μ l EB.

STEP 4: PCR for recovery of gel-purified material and addition of index.

Phusion DNA polymerase, NEB Cat# F-531

Illumina Index PCR 1.0

Illumina Index PCR 2.0

Illumina Index reverse primers (a different one for each library)

As for the adapter mix in step 3, sequences can be requested from Illumina.

a) Combine and mix the following components in for PCR:

DNA from Step 3 10 μ l (which means that some is still available should there be an error or further optimization required at this stage)

Phusion DNA polymerase 2x master mix 25 μ l

PCR primer 1.0, 25 μ M > 1 μ l

PCR primer 2.0, 0.5 μ M > 1 μ l

PCR index reverse primer, 25 μ M > 1 μ l

50 μ l Total reaction volume

b) Amplify using the following PCR protocol:

30 sec at 98 $^{\circ}$ C

[10 sec at 98 $^{\circ}$ C, 30 sec at 65 $^{\circ}$ C, 30 sec at 72 $^{\circ}$ C], 12 cycles total. (Use 12 cycles for 0.5 μ g initial cDNA input – increase cycle number for lower inputs).

5 min at 72 $^{\circ}$ C

Hold at 4 $^{\circ}$ C

Purify final product on Qiaquick PCR purification column and elute in 30 uL buffer EB. Run 3-5 uL on a 2% agarose gel (no need for low-melting agarose) to verify fragment sizes. Run the gel far enough, and against appropriate DNA marker ladder, to ensure there are no 100-120 bp adapter 'dimer' bands contaminating these final libraries.

STEP 5: Quantitate DNA and prepare for submission.

Quantitate final amplified, column-purified DNA using a PicoGreen assay. Dorn lab experiences during early 2009 were that PicoGreen was more accurate than NanoDrop for this step, and this was critical for appropriate formation of DNA clusters on Illumina flowcells.

Approximate fragment size will have been observed by running out some of the final product on a gel. Dilute some of the final library (or blend of indexed libraries) to 8 pM and submit to Illumina HiSeq.

Analysis of RNASeq data, and differential gene expression

Up to October 2011, the open source packages 'tophat' and 'cufflinks' have been implemented on Unix/Linux platforms to convert raw sequence reads into gene expression data (RPKM; and with later versions of cufflinks, FPKM – fragments per kilobase of exon per million sequencing reads, which is equivalent to RPKM for single-end sequencing). The number of reads/fragments mapping to each gene is thus normalized to the length of the gene, and to the total number of mapped reads. This permits comparison of expression levels between genes, and between different sequencing libraries.

→ We are now using Illumina indexing (library-specific tags on the 3' end of the read, rather than 5' barcodes). At some point, reads will be demultiplexed at the sequencing facility and converted to fastq format, eliminating the need to remove barcodes and convert Illumina's SCARF format to fastq. It will be necessary to pay attention to the quality score format (phred score) used in the new Illumina files in order to select the correct option for tophat software.

→ Partek Genomics Suite has matured to the point where it can accept .bam format files produced by 'tophat', and convert them to RPKM scores with annotated names for each gene. However, the calculations used for conversion to gene expression are less well documented than for Cufflinks.

Package differences for differential gene expression and statistics

Conversion of aligned reads to gene expression values, with or without provision for determining isoform fractions and/or new alternatively-spliced transcripts, has been the subject of much discussion. The most appropriate determination of significant differences in gene expression between treatments is also a part of the debate, with three principal groups involved (Simon Anders, Germany; Cole Trapnell & Lior Pachter, Maryland; Gordon Smyth, Australia). All have released packages to calculate expression values and associated statistics (HTSeq+DESeq, Anders; Cufflinks and Cuffdiff, Trapnell/Pachter; edgeR, Smyth).

HTSeq+DESeq, and edgeR, use the actual number of reads aligned to a given mRNA, normalized to total reads, in their gene expression and statistical calculations. They can use either local (per-gene) or global (per gene-set) dispersion (~ variance) to determine statistical significance, and rely on a negative binomial distribution. Cufflinks and Cuffdiff use FPKMs, can investigate isoform differences, but don't estimate variance in the same way.

We've typically used a more straightforward method of importing Cufflinks FPKM values into statistical software such as Partek, and calculating p-values and associated false discovery rates. This is apparently not the same as the DESeq / edgeR process, and may be giving some false positives at lower gene expression values and some false negatives at higher gene expression values in comparison to the approach used by DESeq / edgeR. However, this is still the subject of active debate.

Links to open-source software packages (Linux installation)

Tophat/Cufflinks

<http://bowtie-bio.sourceforge.net/index.shtml>

<http://tophat.cbcb.umd.edu/>

<http://cufflinks.cbcb.umd.edu/>

<http://samtools.sourceforge.net/>

HTSeq / DESeq / edgeR

<http://www-huber.embl.de/users/anders/HTSeq/doc/overview.html>

<http://www-huber.embl.de/users/anders/DESeq/>

<http://www.bioconductor.org/packages/release/bioc/html/edgeR.html>

And for general discussion of package use, advantages / disadvantages / contrasts between methods, investigate the Bioinformatics and RNA-seq forums at:

<http://seqanswers.com/forums/index.php>

Analysis procedure (tophat / cufflinks)

1. Obtain Illumina sequence read files, and split them into files pertaining to each barcode or index used to differentiate individual samples on the sequencing lane. (See Appendix at the end of this document on how to perform this operation.) Make sure the final files are in fastq format.
2. Determine the phred scale for the Illumina basecalling quality scores. Files generated with Illumina pipeline 1.3 and above (generally from 2009 onward) will require the use of an additional parameter in tophat. Files from October 2011 onward from Wash U GTAC may require the *omission* of this parameter (`--solexa1.3-quals`) due to a change to the phred scale.
3. Prepare a Bowtie index for your genome of interest. (Illumina have made several prebuilt indices available for human, mouse, rat and *Drosophila*. See link to iGenomes at <http://cufflinks.cbcb.umd.edu/igenomes.html>.)
4. Obtain a .gtf annotation file to describe the transcriptome present within your genome of interest (included within iGenomes). You may also want to prepare a 'masked' gtf, in which RNAs you don't want to consider in your analysis will have been removed, e.g. rRNAs, hnRNAs, snoRNAs. To do this, make a list of these genes (as a text file, with each gene symbol on a new line). Issue the following command at a shell prompt, within your gtf directory:

```
grep -f maskedgenefile -v all_genes.gtf > masked_genes.gtf
```

where **maskedgenefile** is your list of genes to remove and **all_genes.gtf** is the original gtf.

5. It's most convenient to gather all the fastq files you wish to analyze into the same directory. The names you give them will be used as output folder names that contain several analysis files.
6. Make the appropriate changes to variables in the prebuilt analysis script given in the following pages:

PROJECTDIR (location of your fastq files)

QSCORES (this will be `"--solexa1.3-quals"` or `""`)

BOWTIEINDEX (location of your bowtie index for genome mapping)

REFGTF (a gtf file containing information on all RNAs you would like to map)

MASKEDGTF (a gtf file without information on those RNAs you wish to ignore)

readlength (nt length of your sequencing reads)

threads (number of processing cores)

**** Note, February 2012:** tophat 1.4.0 permits mapping of reads to a defined transcriptome, which obviates the need to produce a 'masked' file in this way. Instead, you could produce a 'full' output file by mapping to the entire genome, and a 'transcriptome' output file by mapping to a defined transcriptome only. Note that different FPKM values will result from these two procedures, since the denominator for FPKM is different (either all reads mapped

to the genome, or all reads mapped to the defined transcriptome). There are two sample scripts at the end of this section; one using mapping to the full genome, and one using mapping to a defined transcriptome. **

7. Determine the total reads per indexed sample, and total aligned reads, using the output files produced from the script below.
8. Import the genes.fpkm_tracking file for each sample into Microsoft Excel (one sample per workbook tab). To make it possible to combine data from different samples into one file later on, add an additional column with the title 'Sample' and fill it all the way down with the sample ID.
9. Merge all the tables together, either in Excel or using the 'import table' facility in Microsoft Access.
10. Import the merged table into Access.
11. Make a 'crosstab query' with gene_id (the Ensembl ID) as the row heading, sample as the column heading, and FPKM as a Value (select First.)
12. Make another 'crosstab query' with gene_id as the row heading, sample as the column heading, and FPKM_status as a Value (select First.)
13. Make a 'simple query' in which the gene_id fields from both of the crosstab queries are linked together. In the query, place all of the fields from the first query together with all the fields from the second query. This will produce one table in which you have all gene FPKM values *and* their 'quality status', separated by sample.
14. Import a table into Access, from previous Access files or made in Excel, of all of the desired Ensembl IDs and gene symbols for your output (i.e. either a 'full' gene list or a 'reduced' transcriptome list). Make another simple query, linking this Ensembl ID-gene symbol table to the previous query containing all FPKM and status information.

Further analysis to designate any outlier samples, and to calculate fold-changes, p-values, may be optimally done using Partek Genomics Suite.

First, however, export your final query from step 13 as an Excel sheet. You will likely want to eliminate any mRNA entries with signals below a certain FPKM threshold, or with signals in <50% of samples. Also, match FPKM values to status values, and discard any that are not 'OK'. (Construct an IF statement in Excel to show either the FPKM or the non-OK status values.)

Once the data have been 'proofed', export them as a .csv file, and import into Partek for ANOVA (fold-changes, p-values, FDR) and heatmaps or volcano plots.

Note: heatmaps are best drawn with gene expression data in log₂ scale, so be prepared to transform your FPKM data to log₂, either in Excel or with Partek.

Sample bash script for RNASeq / RISCSeq analysis, giving FPKM based on all reads mapped to the genome.

This can be copied and pasted into a Linux-based text editor, modified according to the included comments, and made into an executable script. (Line endings will need to be changed from Windows to Unix format. Also be careful to ensure that double-dashes, hyphens and quotation marks are in Unix format, i.e. "-")

* * *

```
#!/bin/bash

# File processing script for RNASeq of Illumina sequence files
#>=November 2011 versions of tophat and cufflinks

PROJECTDIR=~/Documents/RNASeq_processing/cardiac_RNASeq/mice
FILES=$PROJECTDIR/*.fq
# If the files to be processed are present in a separate list,
# replace $PROJECTDIR/*.fq with `cat /path/to/filelist`

QSCORES="--solexa1.3-quals "

# If you are using Illumina pipeline >1.3, keep QSCORES as is.
# After October 2011 it may be necessary to change this to
# QSCORES="" as the Illumina pipelines are reverting back to the
# 'old' system

BOWTIEINDEX=~/Documents/Illumina_iGenomes/mmus_Ensembl_NCBIM37/Sequence/BowtieIndex/genome
REFGTF=~/Documents/Illumina_iGenomes/mmus_Ensembl_NCBIM37/Annotation/Genes/genes.gtf
readlength=42
threads=4

NOVELJUNCS=""

# --no-novel-juncs instructs tophat to ignore novel splice
# junctions, and instead to use only the definitions supplied in
# the reference gtf file. Replace with "--no-novel-juncs" (which
# is the default) to search for novel splice junctions. Remove
# the # from the following line to enable this facility, and add #
# prior to the line SPLICEJUNC=""

# SPLICEJUNC="G $REFGTF $NOVELJUNCS "
# Keep the space at the end of "G $REFGTF $NOVELJUNCS "; it's
# important for the loop script below

SPLICEJUNC=""
```

```

echo "This script processes fastq files generated from Illumina
sequencing through tophat and cufflinks for RNASeq."
echo "It will produce cufflinks determinations against a
complete gtf, calculating FPKM using all reads mapped to the
genome."
echo ""
echo "Ensure you have installed bowtie, tophat, samtools and
cufflinks, and added them to your PATH variable."
echo "Ensure you have supplied the correct fastq file list,
project directory, sequence read length (divisible by 2), Bowtie
index, full reference GTF, and masked reference in the first
lines of this script."
echo ""
echo "To change whether tophat searches for novel splice
variants, change parameters in the first lines of this script."
echo ""

# Don't change anything below here unless you know what you are
doing.

if [ "$readlength" -lt '50' ] ; then
seglength=$(expr "$readlength" '/' '2')
else
seglength=$readlength
fi

echo $PATH > $PROJECTDIR/program_versions.txt
touch $PROJECTDIR/fq_linenumbers.txt
echo "Divide number of lines by 4 to calculate raw read number."
>> fq_linenumbers.txt
touch $PROJECTDIR/aligned_read_counts.txt

for f in $FILES ; do
echo "Calculating number of lines for " $f
wc -l $f > lines
cat lines
cat lines >> fq_linenumbers.txt
rm lines
echo ""
tophat --segment-length $seglength --segment-mismatches 0 --no-
novel-indels -p $threads -o "$(echo $f | sed 's/\.fq$//')"
$QSCORES$SPLICEJUNC$BOWTIEINDEX $f
echo ""
cd "$(echo $f | sed 's/\.fq$//')"
echo "Calculating aligned reads for " $(echo $f | sed
's/\.fq$//')

```

```

samtools view -o accepted_hits.sam accepted_hits.bam
awk {'print $1'} accepted_hits.sam | sort | uniq | wc -l >
aligned
cat aligned
cat aligned >> $PROJECTDIR/aligned_read_counts.txt
pwd >> $PROJECTDIR/aligned_read_counts.txt
rm aligned
rm accepted_hits.sam
cd "$(echo $f | sed 's/\.fq$//')"
cufflinks -p $threads -G $REFGTF accepted_hits.bam
cd $PROJECTDIR
done

```

* * *

Sample bash scripts for RNASeq / RISCSeq analysis, giving FPKM based on all reads mapped to the transcriptome.

This can be copied and pasted into a Linux-based text editor, modified according to the included comments, and made into an executable script. (Line endings will need to be changed from Windows to Unix format. Also be careful to ensure that double-dashes, hyphens and quotation marks are in Unix format, i.e. "-")

First: tophat 1.4.0 (and above) needs to build a Bowtie index corresponding to the transcriptome (script A). This only needs to be run once for each transcriptome of interest. The new Bowtie index (in the transcriptome_data folder) is then used to process future samples (script B).

Script / command A (one continuous line):

```

tophat --segment-length 21 --segment-mismatches 0 --no-novel-
indels -p 4 --solexa1.3-quals -o mouse_heart -T -n 1 -G
~/Illumina_iGenomes/mmus_Ensembl_NCBIM37/Annotation/Current/Genes/genes.gtf --transcriptome-index=transcriptome_data/genes_only
~/Illumina_iGenomes/mmus_Ensembl_NCBIM37/Sequence/BowtieIndex/genome mouse_heart.fq

```

Script / command B:

```

#!/bin/bash

# File processing script for RNASeq of Illumina sequence files
#>=January 2012 versions of tophat and cufflinks

```

```

PROJECTDIR=~/Documents/Jan12_RNASeq
FILES=$PROJECTDIR/*.fq
# If the files to be processed are present in a separate list,
replace $PROJECTDIR/*.fq with `cat /path/to/filelist`

BOWTIEINDEX=~/Illumina_iGenomes/mmus_Ensembl_NCBIM37/Sequence/Bow
tieIndex/genome
REFGTF=~/Illumina_iGenomes/
mmus_Ensembl_NCBIM37/Annotation/Genes/genes.gtf
TRANSCRIPTOME="--transcriptome-
index=transcriptome_data/genes_only"

# Put the transcriptome data folder and bowtie index prefix name
here. This only seems to work if the transcriptome data folder
is inside the folder from which tophat is run.

readlength=42
threads=4
mismatch=1

QSCORES="--solexa1.3-quals "

# If you are using Illumina pipeline >1.3, keep QSCORES as is.
After October 2011 it may be necessary to change this to
QSCORES="" as the Illumina pipelines are reverting back to the
'old' system

NOVELJUNCS=""

# --no-novel-juncs instructs tophat to ignore novel splice
junctions, and instead to use only the definitions supplied in
the reference gtf file. Replace with "--no-novel-juncs" (which
is the default) to search for novel splice junctions. Remove
the # from the following line to enable this facility, and add #
prior to the line SPLICEJUNC=""

# SPLICEJUNC="-G $REFGTF $NOVELJUNCS "
# Keep the space at the end of "G $REFGTF $NOVELJUNCS "; it's
important for the loop script below

SPLICEJUNC=""

echo "This script processes fastq files generated from Illumina
sequencing through tophat and cufflinks for RNASeq."
echo ""
echo "Tophat 1.4.0 or above will be used to align reads to a
previously prepared transcriptome definition."

```

```

echo "Cufflinks 1.3.0 or above will calculate FPKM against a
previously prepared transcriptome gtf and associated bowtie
index; FPKM will be calculated against reads mapping to the
transcriptome, not the whole genome."
echo ""
echo "Ensure you have installed bowtie, tophat, samtools and
cufflinks, and added them to your PATH variable."
echo ""
echo "Ensure you have supplied the correct fastq file list,
project directory, sequence read length (divisible by 2), and
transcriptome-index folder in the first lines of this script."
echo ""
echo "To change whether tophat searches for novel splice
variants, change parameters in the first lines of this script."
echo ""

# Don't change anything below here unless you know what you are
doing.

if [ "$readlength" -lt '50' ] ; then
seglength=$(expr "$readlength" '/' '2')
else
seglength=$readlength
fi

echo $PATH > $PROJECTDIR/program_versions.txt
touch $PROJECTDIR/fq_linenumbers.txt
echo "Divide number of lines by 4 to calculate raw read number."
>> fq_linenumbers.txt
touch $PROJECTDIR/aligned_read_counts.txt

for f in $FILES ; do
echo "Calculating number of lines for " $f
wc -l $f > lines
cat lines
cat lines >> fq_linenumbers.txt
rm lines
echo ""
tophat --segment-length $seglength --segment-mismatches 0 --no-
novel-indels -p $threads $QSCORES$SPLICEJUNC -o "$(echo $f | sed
's/\.fq$//')" -T -n $mismatch $TRANSCRIPTOME $BOWTIEINDEX $f
echo ""
cd "$(echo $f | sed 's/\.fq$//')"
echo "Calculating aligned reads for " $(echo $f | sed
's/\.fq$//')
samtools view -o accepted_hits.sam accepted_hits.bam
awk {'print $1'} accepted_hits.sam | sort | uniq | wc -l >

```

```
aligned
cat aligned
cat aligned >> $PROJECTDIR/aligned_read_counts.txt
pwd >> $PROJECTDIR/aligned_read_counts.txt
rm aligned
rm accepted_hits.sam
cufflinks -p $threads -G $REFGTF accepted_hits.bam
cd $PROJECTDIR
done
```

Analysis procedure (HTSeq / DESeq / edgeR)

1. Obtain Illumina sequence read files, and split them into files pertaining to each barcode or index used to differentiate individual samples on the sequencing lane. (See Appendix at the end of this document on how to perform this operation.) Make sure the final files are in fastq format.
2. Determine the phred scale for the Illumina basecalling quality scores. Files generated with Illumina pipeline 1.3 and above (generally from 2009 onward) will require the use of an additional parameter in tophat. Files from October 2011 onward from Wash U GTAC may require the *omission* of this parameter (`--solexa1.3-quals`) due to a change to the phred scale.
3. Prepare a Bowtie index for your genome of interest. (Illumina have made several prebuilt indices available for human, mouse, rat and *Drosophila*. See link to iGenomes at <http://cufflinks.cbcb.umd.edu/igenomes.html>.)
4. Use tophat to align reads to the genome or transcriptome. HTSeq will count the aligned reads mapping to gene entries in the .gtf annotation file corresponding to the genome or transcriptome used by tophat. (See the example bash script.)
5. The HTSeq counts will need to be merged into a table in which columns denote samples, rows are gene entries, and individual cells are counts. Use the crosstab query instructions for Microsoft Access given in steps 9-11 of the tophat/cufflinks analysis procedure.
6. Further processing to gain differential gene expression data, p-values, and associated information is performed using the R statistical language and the modules supplied with DESeq and edgeR – detailed instructions in are included with these downloads as pdf files.

Install HTSeq and any required Python packages according to the instructions found on its installation website. To generate an example script to count raw reads, perform tophat alignment, and extract counts of aligned reads per gene using HTSeq, replace the bottom section of the previous 'script B' in this document for transcriptome processing (the entire section below the comment field stating 'Don't change anything below here unless you know what you are doing' with the following:

```
if [ "$readlength" -lt '50' ] ; then
seqlength=$(expr "$readlength" '/' '2')
else
seqlength=$readlength
fi

echo $PATH > $PROJECTDIR/program_versions.txt
touch $PROJECTDIR/fq_linenumber_counts.txt
echo "Divide number of lines by 4 to calculate raw read number."
>> fq_linenumber_counts.txt
touch $PROJECTDIR/aligned_read_counts.txt

for f in $FILES ; do
```

```

echo "Calculating number of lines for " $f
wc -l $f > lines
cat lines
cat lines >> fq_linenummer_counts.txt
rm lines
echo ""
tophat --segment-length $seglength --segment-mismatches 0 --no-
novel-indels -p $threads $QSCORES$SPLICEJUNC -o "$(echo $f | sed
's/\.fq$//')" -T -n $mismatch $TRANSCRIPTOME $BOWTIEINDEX $f
echo ""
cd "$(echo $f | sed 's/\.fq$//')"
echo "Calculating aligned reads for " $(echo $f | sed
's/\.fq$//')
samtools view -o accepted_hits.sam accepted_hits.bam
awk {'print $1'} accepted_hits.sam | sort | uniq | wc -l >
aligned
cat aligned
cat aligned >> $PROJECTDIR/aligned_read_counts.txt
pwd >> $PROJECTDIR/aligned_read_counts.txt
rm aligned
python -m HTSeq.scripts.count --stranded=no accepted_hits.sam
$REFGTF > "$(echo $f | sed 's/\.fq$/_HTSeq-count.txt/')"
rm accepted_hits.sam
cd $PROJECTDIR
done

```


Supplemental commands: sequence sorting by index, removal of indexes, and fastq generation

Use the following script:

```
#!/bin/bash

# File processing script for Illumina sequence files to remove
indexes and write out new fastq files.

SEQFILE=s_3_withinindex_sequence.txt
INDEXES=`cat s_3_indexes.list`

# Replace the SEQFILE variable with the name of the Illumina
SCARF file.  Replace the INDEXES variable with a path to a file
containing the indexes to be sorted, one per line.

echo "This script processes an Illumina sequencing SCARF file,
writing out new files separated by indexes, rendered as fastq
format."
echo ""
echo "You should have already changed the SEQFILE and INDEXES
variables to point to the raw sequence file and to the list of
indexes."
echo ""
echo "The raw sequence file is assumed to be from HiSeq single-
end, 42 nt, with index nucleotides at the 3' end."
echo ""
echo "Make sure the list of indexes contains indexes in UPPER
CASE to match the raw sequence file!"
echo ""

for i in $INDEXES ; do
echo 'Processing index '$(echo $i)
grep $(echo $i)':' $SEQFILE | sed -n s/$(echo $i)':'/'/'/p | awk
-F: '{print "@" $1 "_" $2 "_" $3 "_" $4 "_" $5 "\n" $6 "\n" "+"
"\n" substr($7,1,42) }' > 'index_$(echo $i).fq
done

# Change substr($7,1,42) to substr($7,1,x) for sequence read
lengths of x
```