A novel molecular signature for elevated tricuspid regurgitation velocity in sickle cell disease

Ankit A. Desai, Tong Zhou, Homaa Ahmad, Wei Zhang, Wenbo Mu, Sharon Trevino, Michael S. Wade, Nalini Raghavachari, Gregory J. Kato, Marlene H. Peters-Lawrence, Tejas Thiruvoipati, Kristin Turner, Nicole Artz, Yong Huang, Amit R. Patel, Jason X.-J. Yuan, Victor R. Gordeuk, Roberto M. Lang, Joe G. N. Garcia, Roberto F. Machado

SUPPLEMENTAL MATERIAL & METHODS

Subjects. Clinical trial registration information is provided at the following website: http://clinicaltrials.gov/ct2/show/NCT01044901 for the University of Chicago cohort. The specimens from NIH subjects were obtained under an IRB-approved protocol (current NIH protocol number 03-H-0015), which is exempt from registration requirements at ClinicalTrials.gov under the current ICMJE definition of clinical trials. The protocol is listed at the NIH Clinical Center website: http://clinicalstudies.info.nih.gov/detail/A_2003-H-0015.html

Transthoracic echocardiography (TTE). TTE was performed with the patient in the left lateral decubitus using an S5 probe connected to an iE33 ultrasound machine (Philips, Medical Imaging, Andover, MA). Digital cine loops were acquired by a licensed sonographer and subsequently reviewed off line by two independent cardiologists (>99% correlation for all PH measurements). Measurements were performed and interpreted according to the guidelines of the American Society of Echocardiography (ASE). The left ventricular ejection fraction (LV EF) was calculated using the biplanar Simpson's method of discs[1,2]. Tricuspid regurgitation was assessed in multiple views and minimum of 5 sequential complexes were recorded. Continuous-wave Doppler of the peak tricuspid regurgitant velocity (TRV) was used to estimate the right ventricular–to–right atrial pressure gradient using the modified Bernoulli's equation $(4 * [TRV]^2)$[3]. An elevated TRV was defined as a velocity $\geq$2.5 m/s. Right ventricular systolic pressure (RVSP) as a surrogate for pulmonary artery systolic pressure was quantified by adding

the Bernoulli-derived pressure gradient to the estimated mean right atrial (RA) pressure as previously described[1].

Blood collection and isolation of PBMCs. Blood was drawn via peripheral venipuncture through a 21-gauge needle into BD Vacutainer(R) CPT tubes with EDTA (Becton, Dickinson and Company, Franklin Lakes, NJ, USA). Care was taken to standardize blood sample collection and preparation. Peripheral blood mononuclear cells (PBMCs) were isolated from blood by Ficoll–Hypaque density gradient separation, and cells were lyzed (Invitrogen, Carlsbad, CA, USA) and stored at–80^0C as described previously[4].

Microarray preparation and analysis. Total RNA enriched with microRNA were isolated from these lysates using Qiagen's RNeasy plus kit as per the manufacturer's directions. Affymetrix Human Exon 1.0 ST array was employed for mRNA profiling and for miRNA profiling, we used Affymetrix GeneChip miRNA arrays. RNA quality control was performed with the use of automated electrophoresis system Experion (BioRad). Labeling reactions and hybridizations were carried out according to the standard GeneChip® WTsense target labeling protocol (Affymetrix). A 100 ng of total cellular RNA per sample was used for each labeling reaction for gene and miRNA arrays. Hybridizations were followed by binding to streptavidin-conjugated fluorescent marker. For miRNA, RNA samples were labeled using FlashTag HSR RNA labeling kit (Genisphere) and hybridized according to Affymetrix recommended protocol. Detection of bound probe was achieved following laser excitation of the fluorescent marker and scanning of the resultant emission spectra using a scanning confocal laser microscope (Affymetrix). Data acquisition was performed using Affymetrix AGCC suite. Hybridization

images were subjected to quality control with the use of Expression Console analysis tool (Affymetrix) and the use of miRNA QC tool (Affymetrix).

Exon array data were analyzed by the Affymetrix Power Tools v.1.12.0 (http://www.affymetrix.com/). Probesets with known SNPs in the dbSNP database (v129)[5] were filtered out using experimental probe masking workflow provided by the Affymetrix Power Tools. Overall, of the ~1.4 million probesets on the exon array, ~350,000 probesets were found to contain at least one probe with a SNP (~600,000 probes)[6]. We calculated the intensities of probesets by RMA algorithm[7] using "apt-probeset-summarize" program in Affymetrix Power Tools. The expression level of each the transcript clusters (gene-level) with the core set (i.e., with RefSeq-supported annotations)[8] of exons were summarized by taking averages of all annotated probesets for each transcript cluster. Adjustment for possible batch effect was conducted by COMBAT (http://jlab.byu.edu//ComBat/)[9]. We considered a transcript cluster to be reliably expressed if the P-values calculated by DABG (detection above ground) algorithm[10] in Affymetrix Power Tools were less than 0.01 in at least two third of the samples in each test. We removed genes on chromosomes X and Y to avoid the potential confounding factor of gender. We further limited our analysis to the genes with unique annotations (i.e., transcripts corresponding to unique genes) from the Affymetrix NetAffy website (www.affymetrix.com/analysis/netaffx/). For miRNA, raw data (cel files) were treated by the following workflow: background detection, BC-CG background adjustment, quantile normalization, and median polish with Affymetrix miRNA QC tool. DAVID tools[11,12] were used to identify the enriched KEGG (Kyoto Encyclopedia of Genes and Genomes, http://www.genome/jp/kegg)[13] physiological pathways among the

potentially differentially expressed genes.

Identification of a PH gene signature in SCD. The decision function of a linear SVM using a linear kernel is $f(\mathbf{x}) = sign\{(\mathbf{w} \bullet \mathbf{x}) + b\}$, where **x** is the gene expression vector of a sample, **w** is the vector of weights of the features, and *b* is a scalar offset[14-16]. The recursive feature elimination (RFE) approach recursively reduces the number of genes used in the predictor function by removing those genes with lowest weights and re-fitting the SVM algorithm using the remaining genes. In the first step, all genes are ranked by SVM according to their weights. In each of the following steps of the RFE procedure, half of the genes are eliminated from the predictor model until the gene number is smaller than 20. A five-fold cross-validation (repeated for 1,000 times) of the predictive models based on differently sized subsets of genes, as selected by RFE, was performed. After the recursive feature selection steps on each subset, we counted at each level the frequency of the features being selected among all rounds of cross-validation experiments. The top most frequently selected features are reported as the final features (signature genes)[16]. The *e1071* library of the R Statistical Package[17] was used for identifying the gene signature.

Comparison of microarray results to published PH and SCD studies. Microarray data was compared to human PBMC-based PH and SCD gene expression data (three SCD studies, no studies of PH in SCD) through a PubMed search, representing work across a variety of analyses (PAH vs PH, SCD vs control, etc). For example, one study utilized included significantly differentially expressed genes from PBMC-derived mRNA of healthy patients versus patients with SCD[18] whereas another study used differentially expressed genes between human PAH versus the combination

of secondary PH and healthy controls[19]. Common genes between the available data sets and the newly generated list of significantly differentially expressed PH genes were identified.

Pubmatrix evaluation. The relevance of newly identified candidate genes for an elevated TRV to PH was evaluated with PubMatrix (http://pubmatrix.grc.nia.nih.gov), an automated biomedical literature search engine measuring co-occurrence of two keywords in PubMed as described previously[20].

RT-PCR. Selection of three differentially expressed transcripts (arginine decarboxylase, adenosine A2b receptor, complement component 1 q subcomponent binding protein) and three miRNAs (miR21, MiR19a, and miR-15a) for confirmation and validation by RT-qPCR were based upon their ability to discriminate between patients with an elevated TRV in SCD versus those with normal TRV. These 3 were chosen based on their established roles as potential candidates involved in the pathophysiology of pulmonary hypertension and/or sickle cell disease. *GALNT13* and miR-301a were also validated by RT-PCR given their robust representation as novel candidate gens/miRNAs in elevated TRV in SCD in the current study. Results were standardized to the expression of the control gene GAPDH for gene expression and U6 snRNA for miRNA expression. We used the High Capacity cDNA Reverse Transcription Kit (Invitrogen, Laboratories, CArlsbad, CA, USA) for reverse transcription of RNA for both sets. Primers and probes for the following genes and miRNAs were obtained from Applied Biosystems (ABI) Assays on Demand using their recommended conditions (Foster City, CA, USA). The Affymetrix probe ID for the genes was utilized to select ABI primer/probe sets on the ABI website (http://www4.appliedbiosystems.com/

tools/umapit/). This provided primer/probe sets near the 3' end of the gene adjacent to the Affymetrix probe and within the coding region of the gene when possible. TaqMan (R) MGB probes and primers were FAM TM and VIC (R) dye-labeled; TaqMan (R) Universal Master Mix was obtained (ABI) for the genes and separately for the miRNAs. The BioRad CFX Real-Time PCR Detection System (Bio-Rad) was used to evaluate the three genes and three miRNAs of interests for an elevated TRV. All experimental samples and standards were run in duplicate and averaged.

Genotypic Data & Association Analysis. We obtained genotypic data on our candidate genes in a cohort of SCD patients that were genotyped using the Affymetrix SNP 6.0 platform. Common SNPs without reasonable genotyping quality (MAF >0.05, genotyping rate >90%, and HWE >0.001) were tested for conventional case-control association using PLINK, version 1.07[21]. The Cochran-Mantel-Haenszel association test was utilized to control cohort effects given that two SCD cohorts were combined as a single discovery cohort. We have also performed principal component analysis using a panel of ancestry informative markers derived for African Americans (Reich Lab, Harvard). There were no obvious outliers among our samples. Mapping of eQTLs using linear regression based on an additive genetic model (i.e., genotype AA=0, AB=1, BB=2, given allele B is has a dosage-dependent effect) was performed. Definition for cis-acting or local SNPs was based on the annotations from the Affymetrix NetAffyx website and additional SNPs on the same chromosome or other chromosomes were defined as trans-acting or distant SNPs.

References

1. Lang RM, Bierig M, Devereux RB, Flachskampf FA, Foster E, Pellikka PA, Picard MH, Roman MJ, Seward J, Shanewise JS, Solomon SD, Spencer KT, Sutton MS, Stewart WJ. Recommendations for chamber quantification: a report from the American Society of Echocardiography's Guidelines and Standards Committee and the Chamber Quantification Writing Group, developed in conjunction with the European Association of Echocardiography, a branch of the European Society of Cardiology. *J Am Soc Echocardiogr.* Dec 2005;18(12):1440-1463.

2. Lester SJ, Ryan EW, Schiller NB, Foster E. Best method in clinical practice and in research studies to determine left atrial size. *Am J Cardiol.* Oct 1 1999;84(7):829-832.

3. Quinones MA, Otto CM, Stoddard M, Waggoner A, Zoghbi WA. Recommendations for quantification of Doppler echocardiography: a report from the Doppler Quantification Task Force of the Nomenclature and Standards Committee of the American Society of Echocardiography. *J Am Soc Echocardiogr.* Feb 2002;15(2):167-184.

4. Grigoryev DN, Mathai SC, Fisher MR, Girgis RE, Zaiman AL, Housten-Harris T, Cheadle C, Gao L, Hummers LK, Champion HC, Garcia JG, Wigley FM, Tuder

RM, Barnes KC, Hassoun PM. Identification of candidate genes in scleroderma-related pulmonary arterial hypertension. *Transl Res.* Apr 2008;151(4):197-207.

5. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* Jan 1 2001;29(1):308-311.

6. Duan S, Zhang W, Bleibel WK, Cox NJ, Dolan ME. SNPinProbe_1.0: a database for filtering out probes in the Affymetrix GeneChip human exon 1.0 ST array potentially affected by SNPs. *Bioinformation.* 2008;2(10):469-470.

7. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics.* Apr 2003;4(2):249-264.

8. Pruitt KD, Tatusova T, Maglott DR. NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res.* Jan 2007;35(Database issue):D61-65.

9. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics.* Jan 2007;8(1):118-127.

10. Affymetrix. Exon array background correction. *Affymetrix Whitepaper.* 2005.

11. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc.* 2009;4(1):44-57.

12. Dennis G, Jr., Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, Lempicki RA. DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol.* 2003;4(5):P3.

13. Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M. The KEGG resource for deciphering the genome. *Nucleic Acids Res.* Jan 1 2004;32(Database issue):D277-280.

14. Guyon I, Weston J, Barnhill S. Gene selection for cancer classification using support vector machines. *Machine Learning.* 2002;46:389-422.

15. Thuerigen O, Schneeweiss A, Toedt G, Warnat P, Hahn M, Kramer H, Brors B, Rudlowski C, Benner A, Schuetz F, Tews B, Eils R, Sinn HP, Sohn C, Lichter P. Gene expression signature predicting pathologic complete response with gemcitabine, epirubicin, and docetaxel in primary breast cancer. *J Clin Oncol.* Apr 20 2006;24(12):1839-1845.

16. Zhang X, Lu X, Shi Q, Xu XQ, Leung HC, Harris LN, Iglehart JD, Miron A, Liu JS, Wong WH. Recursive SVM feature selection and sample classification for mass-spectrometry and microarray data. *BMC Bioinformatics.* 2006;7:197.

17. R_Development_Core_Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2005.

18. Jison ML, Munson PJ, Barb JJ, Suffredini AF, Talwar S, Logun C, Raghavachari N, Beigel JH, Shelhamer JH, Danner RL, Gladwin MT. Blood mononuclear cell gene expression profiles characterize the oxidant, hemolytic, and inflammatory stress of sickle cell disease. *Blood.* Jul 1 2004;104(1):270-280.

19. Rajkumar R, Konishi K, Richards TJ, Ishizawar DC, Wiechert AC, Kaminski N, Ahmad F. Genomewide RNA expression profiling in lung identifies distinct signatures in idiopathic pulmonary arterial hypertension and secondary

pulmonary hypertension. *Am J Physiol Heart Circ Physiol.* Apr 2010;298(4):H1235-1248.

20. Moreno-Vinasco L, Gomberg-Maitland M, Maitland ML, Desai AA, Singleton PA, Sammani S, Sam L, Liu Y, Husain AN, Lang RM, Ratain MJ, Lussier YA, Garcia JG. Genomic assessment of a multikinase inhibitor, sorafenib, in a rodent model of pulmonary hypertension. *Physiol Genomics.* Apr 22 2008;33(2):278-291.

21. Singer JB. Candidate gene association analysis. *Methods Mol Biol.* 2009;573:223-230.

Supplemental Tables and Figures.

Supplemental Table 1. TTE measurements of patients with SCD.

	Discovery Cohort (n=27)			Validation Cohort (n=20)			Discovery vs Validation
	Elevated TRV (n=18)	Normal TRV (n=9)	P Value	PH (n=10)	Normal TRV (n=10)	P Value	P Value**
LV EF (%)	54.9±3.6	52.1±10.5	0.42	59 ± 8.4	64 ± 3.6	0.09	0.21
RAP (mmHg)	6.4±2.9	5.5±1.7	0.43	8.0 ± 4.8	5.3 ± 1.2	0.04	0.50
RVSP (mm Hg)	33.5±7.9	21 ±4.8	<0.01	43 ± 13	26.3 ± 2.5	<0.01	0.13
TRV (m/s)	2.6±0.3	1.9±0.3	<0.01	2.9 ± 0.5	2.3 ± 0.1	<0.01	0.17

LV EF= Left ventricular ejection fraction; RAP= right atrial pressure; RVSP= right ventricular systolic pressure; TRV= maximum tricuspid regurgitation jet velocity, **Comparison between patients with an elevated TRV in the discovery cohort against patients with RHC-defined PH in the validation cohort

Supplemental Table 2. Right heart catheterization measurements of patients with PH in the validation cohort.

	Validation Cohort (n=20)			
	Total PH (n=10)	**PAH (n=5)**	**PVH (n=5)**	**P value (PAH vs PVH)**
RAP (mmHg)	9.7 ±6.7	7.7 ± 2.9	12.4 ± 9.7	0.25
PASP (mmHg)	56.3 ± 13.1	57.0 ± 15.9	55.4 ± 13.1	0.84
PADP (mmHg)	24.6 ± 8.3	23.4 ± 9.4	26.2 ± 7.2	0.59
mPAP (mmHg)	37.8 ± 8.7	38.0 ± 9.7	37.5 ± 7.9	0.93
PVR (WU)	2.1 ± 1.4	2.8 ± 1.4	1.1 ± .01	0.05
CO (L/min)	9.9 ± 1.7	10.2 ± 2.7	9.6 ± 2.7	0.72
CI (L/min/m2)	5.5 ± 1.7	5.5 ± 1.9	5.4 ± 1.7	0.95
PCWP (mmHg)	17.5 ± 6.9	12.6 ± 2.9	24.4 ± 4.3	<0.01
Avg TRV (m/s)	2.9 ± 0.4	2.8 ± 0.5	3.1 ± 0.4	0.34

Supplemental Table 3. Correlation coefficients for genes comprising the gene signature for an elevated TRV.

Genes	Gene Symbol	Correlation coefficient - RVSP	P value	Correlation coefficient - TRV	P value
Myosin light chain 10, regulatory	MYL10	-0.507	0.007	-0.491	0.009
Adenosine A2b receptor	ADORA2B	-0.601	0.001	-0.555	0.003
Complement component 1, q binding protein	C1QBP	-0.447	0.019	-0.464	0.015
Arginine decarboxylase	ADC	-0.575	0.002	-0.559	0.002
Proline/arginine-rich end leucine-rich repeat protein	PRELP	-0.479	0.011	-0.442	0.021
SAM pointed domain containing ets transcription factor	SPDEF	-0.512	0.006	-0.562	0.002
polypeptide N-acetylgalactosaminyltransferase 13	GALNT13	-0.401	0.038	-0.460	0.016
Solute carrier family 27, member 5	SLC27A5	-0.559	0.002	-0.586	0.001
N-acetyltransferase 8-like	NAT8L	-0.452	0.018	-0.429	0.026
Chrom 9 open reading frame 16	C9orf16	-0.489	0.010	-0.435	0.024

Supplemental Table 4. 50 overlapping genes between those associated with an elevated TRV in the current study and a previously published study in PAH.

Gene Name	Gene_Symbol	Cytoband
5'-nucleotidase domain containing 3	NT5DC3	12q23.3
BMI1 polycomb ring finger oncogene	BMI1	10p12.2
CKLF-like MARVEL transmembrane domain containing 8	CMTM8	3p22.3
F-box protein 3	FBXO3	11p13
G protein-coupled receptor 6	GPR6	6q21
Mdm1 nuclear protein homolog (mouse)	MDM1	12q15
NPC1 (Niemann-Pick disease, type C1, gene)-like 1	NPC1L1	7p13
PAN3 poly(A) specific ribonuclease subunit homolog (	PAN3	13q12.2
PCF11, cleavage and polyadenylation factor subunit, homolog	PCF11	11q14.1
PDZ domain containing ring finger 3	PDZRN3	3p13
PNMA-like 1	PNMAL1	19q13.32
RAB14, member RAS oncogene family	RAB14	9q33.2
RUN and FYVE domain containing 2	RUFY2	10q21.3
Ras association and DIL domains	RADIL	7p22.1
TAF15 RNA polymerase II, TATA box binding protein-associated	TAF15	17q12
TM2 domain containing 3	TM2D3	15q26.3
amyloid beta (A4) precursor-like protein 1	APLP1	19q13.12
dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 1A	DYRK1A	21q22.13
fibrillin 3	FBN3	19p13.2
fibroblast growth factor receptor substrate 2	FRS2	12q15
filamin binding LIM protein 1	FBLIM1	1p36.21
heat shock transcription factor 4	HSF4	16q22.1
helicase with zinc finger	HELZ	17q24.2
heterogeneous nuclear ribonucleoprotein H3 (2H9)	HNRNPH3	10q21.3

hypocretin (orexin) receptor 1	HCRTR1	1p35.2
isocitrate dehydrogenase 3 (NAD+) alpha	IDH3A	15q25.1
kallikrein-related peptidase 3	KLK3	19q13.3 3
karyopherin alpha 3 (importin alpha 4)	KPNA3	13q14.2
kinetochore associated 1	KNTC1	12q24.3 1
origin recognition complex, subunit 5-like (yeast)	ORC5L	7q22.1
paired-like homeodomain 3	PITX3	10q24.3 2
phosphatidylinositol glycan anchor biosynthesis, class N	PIGN	18q21.3 3
potassium voltage-gated channel, KQT-like subfamily, member 2	KCNQ2	20q13.3 3
progestin and adipoQ receptor family member VI	PAQR6	1q22
protein arginine methyltransferase 3	PRMT3	11p15.1
purine-rich element binding protein B	PURB	7p13
purinergic receptor P2X, ligand-gated ion channel, 6	P2RX6	22q11.2 1
regulation of nuclear pre-mRNA domain containing 1A	RPRD1A	18q12.2
six transmembrane epithelial antigen of the prostate 2	STEAP2	7q21.13
splicing factor, arginine/serine-rich 15	SFRS15	21q22.1 1
succinate-CoA ligase, ADP-forming, beta subunit	SUCLA2	13q14.2
transmembrane 9 superfamily member 3	TM9SF3	10q24.1
transmembrane protease, serine 6	TMPRSS6	22q12.3
transmembrane protein 37	TMEM37	2q14.2
tripartite motif-containing 22	TRIM22	11p15.4
tripartite motif-containing 46	TRIM46	1q22
tumor necrosis factor receptor superfamily, member 11a	TNFRSF11A	18q21.3 3
ubiquitin-conjugating enzyme E2G 1	UBE2G1	17p13.2
zinc finger protein 2	ZNF2	2q11.1
zinc finger, FYVE domain containing 28	ZFYVE28	4p16.3

Supplemental Table 5. Characteristics of patients with sickle cell disease in the genetics analysis

	(n=112)		
	Elevated TRV (n=49)	**Normal TRV (n=63)**	**P value**
Age (years)	37.0±10.1	35.8±12.6	0.60
Female / Male	27 / 22	35 / 28	0.96
History of ACS	24/44 (54%)	35/56 (62%)	0.55
Hemoglobin (g/dl)	8.3±1.7	8.8±2.0	0.19
NT-pro-BNP (ng/L)	828±4476	223±690	0.34
Average TRV (m/s)	2.8±0.4	2.1±0.3	<0.01

Supplemental Table 6. Correlation coefficients for the top miRNAs associated with an elevated TRV.

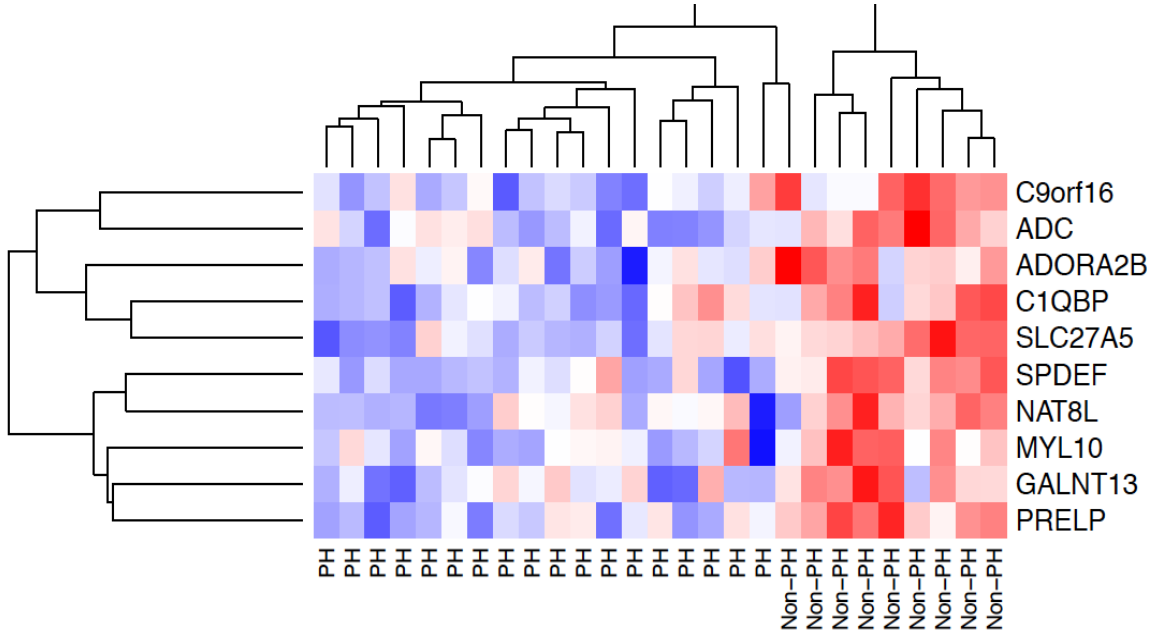
miRNA	Correlation coefficient - RVSP	P value	Correlation coefficient - TRV	P value
miR-15a	0.40	0.04	0.47	0.01
miR-19a	0.48	0.01	0.53	0.005
miR-19b	0.40	0.04	0.50	0.01
miR-21	0.42	0.03	0.50	0.01
miR-30a	0.43	0.02	0.40	0.04
miR-125b	0.36	0.07	0.48	0.01
miR-142-5p	0.50	0.009	0.51	0.007
miR-301a	0.36	0.07	0.43	0.03
miR-30e	0.38	0.05	0.42	0.03
miR-324-3p	-0.44	0.02	-0.43	0.02
miR-626	-0.39	0.04	-0.33	0.09
miR-1285	-0.33	0.10	-0.39	0.04

Supplemental Figure 1: Genomic signature of an elevated TRV in SCD (heatmap). The heatmap displays an organized hierarchical clustering of the samples with up-regulated gene expression shown in red and down-regulated gene expression shown in blue. The 10 gene signature has a 100% accuracy in discriminating echo-defined elevated TRV in patients with SCD.

Supplemental Figure 2. *ADORA2B* and *GALNT13* networks in SCD-related elevated TRV. The gene list was submitted to Ingenuity pathway analysis to identify gene association networks enriched with differentially expressed genes in SCD-related elevated TRV with focus on selected signature genes. Panel A displays a network involving *ADORA2B* among a variety of top genes correlated with an elevated TRV which are involved in carbohydrate metabolism and cell signaling. Panel B displays a network involving *GALNT13* among a variety of top genes correlated with an elevated TRV which are involved in cancer, cellular growth and proliferation and genetic disorders. Red and green color indicates gene expression changes. Red indicates upregulation, green downregulation, and no color indicates no change. Dot lines represent indirect gene interactions and solid lines, direct interactions.

Supplemental Figure 3: Heatmap of top miRNAs associated with an elevated TRV. miRNA expression profiles were compared between SCD cases with (n=17) and without (*n*=9) an elevated TRV. Heatmap analysis depicts the 12 differentially expressed miRNAs which discriminates patients with an elevated TRV from those with a normal TRV.

Supplemental Figure 1. Genomic signature of an elevated TRV in SCD.



Supplemental Figure 2. Network analysis of elevated TRV genes.

Supplemental Figure 3. Heatmap of top miRNAs associated with an elevated TRV.