SCIENTIFIC DATA

OPEN

Data Descriptor: Transcriptomic analyses of murine ventricular cardiomyocytes

Morgan Chevalier[1,*], Sarah H. Vermij[1,*], Kurt Wyler[2], Ludovic Gillet[3,4], Irene Keller[5] & Hugues Abriel[1]

Received: 19 October 2017
Accepted: 14 June 2018
Published: 21 August 2018

Mice are used universally as model organisms for studying heart physiology, and a plethora of genetically modified mouse models exist to study cardiac disease. Transcriptomic data for whole-heart tissue are available, but not yet for isolated ventricular cardiomyocytes. Our lab therefore collected comprehensive RNA-seq data from wildtype murine ventricular cardiomyocytes as well as from knockout models of the ion channel regulators CASK, dystrophin, and SAP97. We also elucidate ion channel expression from wild-type cells to help forward the debate about which ion channels are expressed in cardiomyocytes. Researchers studying the heart, and especially cardiac arrhythmias, may benefit from these cardiomyocyte-specific transcriptomic data to assess expression of genes of interest.

Design Type(s)	parallel group design • gene expression analysis objective • sequence analysis objective
Measurement Type(s)	transcription profiling assay
Technology Type(s)	RNA sequencing
Factor Type(s)	genotype • phenotype • biological replicate • candidate_gene
Sample Characteristic(s)	Mus musculus • cardiac muscle fiber

[1]Ion Channels and Channelopathies Laboratory, Institute for Biochemistry and Molecular Medicine, University of Bern, 3012 Bern, Switzerland. [2]Interfaculty Bioinformatics Unit and Swiss Institute of Bioinformatics, University of Bern, 3012 Bern, Switzerland. [3]Pain Center, Department of Anesthesiology, University Hospital Center (CHUV) and Faculty of Biology and Medicine (FBM), University of Lausanne, 1011 Lausanne, Switzerland. [4]Department of Fundamental Neurosciences, Faculty of Biology and Medicine (FBM), University of Lausanne, 1005 Lausanne, Switzerland. [5]Department for BioMedical Research and Swiss Institute of Bioinformatics, University of Bern, 3007 Bern, Switzerland. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to H.A. (email: hugues.abriel@ibmm.unibe.ch).

Background & Summary

In this study, we present next-generation RNA sequencing (RNA-seq) data of murine ventricular cardiomyocytes (CMC). To date, only whole-heart RNA-seq data have been published[1–3], in which a variety of cell types, such as fibroblasts, endothelial cells, and atrial and ventricular cardiomyocytes, are pooled. We endeavoured to provide RNA-seq data of isolated CMCs for several reasons. Firstly, since the pump function of the heart relies on proper CMC function, CMCs are the most thoroughly studied cardiac cell type. Researchers studying CMCs may benefit from CMC-specific RNA-seq data from which expression of genes of interest can be extracted. Secondly, because of the crucial role of ion channels in cardiac electrical excitability and arrhythmogenesis, researchers that study cardiac arrhythmias have debated the question of which ion channels are expressed in CMCs. However, existing ion channel expression data are low-throughput, often contradictory[4–6], fragmented[7], or expression is assessed in the whole heart. The present work reveals the expression of the more than 350 ion channel family members, including pore-forming and auxiliary subunits, in CMCs (see Fig. 1 and Tables 1–3 (available online only)). We therefore believe that these data will be valuable for ion channel researchers attempting to resolve the ongoing debate.

We have also included cardiac-specific knockout models of the ion channel regulators dystrophin, synapse-associated protein-97 (SAP97), and calmodulin-activated serine kinase (CASK). They interact with ion channels and modify their cell biological properties, such as membrane localization[3,8–11]. Notably, CASK provides a direct link between ion channel function and gene expression. It regulates transcription factors (TFs) in the nucleus, such as Tbr-1, and induces transcription of T-element-containing genes[12]. CASK also regulates TFs of the basic helix-loop-helix family, which bind E-box elements in promoter regions, by modulating the inhibitor of the DNA-binding-1 TF[13]. Additionally, CASK and SAP97 directly interact with each other[11]. For these reasons, we include CASK, SAP97, and dystrophin knockout mice to investigate whether these three proteins have a similar effect on gene expression, which may suggest their involvement in similar pathways. However, research beyond the scope of this paper would be needed to determine whether CASK-dependent TF regulation caused the differential expression that we observed.

To date, mutations in approximately 27 ion channel genes have been associated with cardiac arrhythmias, such as congenital short- and long-QT syndrome (SQTS and LQTS), Brugada syndrome (BrS), and conduction disorders (see http://omim.org)[14–16]. Notably, our ion channel expression data, as presented in Fig. 1 and Tables 1–3 (available online only), reveal that several arrhythmia-associated ion channel genes are not or are scarcely expressed in murine ventricular CMCs (including Kcne2, Kcne3, Scn2b, and Scn3b). Although murine and human ion channel expression may differ, we are presently unaware of any available transcriptome of human CMCs[17,18]. We are also unable to either exclude or assess the effect of enzymatic isolation on the transcriptome. Finally, other cardiac cell types such as (myo)fibroblasts may express these ion channels and therefore may be important for arrhythmogenesis. Indeed, many ion channel genes that are not expressed in cardiomyocytes have been reported in murine whole-heart tissue[2]. These include Scn1a, Scn3b, 10 voltage-gated Ca^{2+} channels, 10 K_v channels, and four two-pore K^+ channels. Conversely, all ion channel genes expressed in CMCs are also reported in whole-heart expression data.

In sum, this study presents RNA-seq data from wildtype murine ventricular CMCs, as well as from SAP97, CASK, and dystrophin knockouts and controls (see Fig. 2 for a schematic overview of study design). We performed differential gene expression analysis to compare the knockouts to their controls, and we extracted wildtype ion channel gene expression data (Tables 1–3 (available online only), Fig. 1). We believe that these data will be valuable for researchers studying cardiomyocytes and ion channels to assess expression of genes of interest.

Methods

Mouse models

All animal experiments conformed to the *Guide to the Care and Use of Laboratory Animals* (US National Institutes of Health, publication No. 85-23, revised 1996); have been approved by the Cantonal Veterinary Administration, Bern, Switzerland; and have complied with the Swiss Federal Animal Protection Law. Mice were kept on a 12-hour light/dark cycle. Lights were on from 6:30 AM to 6:30 PM. To avoid the influence of circadian rhythm, mice were sacrificed between 10:00 AM and 1:00 PM. Mice were all male and were between the ages of 8 and 15 weeks.

MHC-Cre. The cardiac-specific murine alpha-myosin heavy chain (μMHC) promoter drives the expression of Cre recombinase, which, in turn, can recombine LoxP sequences. The μMHC-Cre strain was generated as previously described[19] and acquired from the Jackson Laboratory (stock #011038).

CASK and SAP97 knockout mice. CASK KO and SAP97/Dlg1 KO mice were generated as previously described[9,20]. Both the CASK and SAP97 mouse lines were on mixed backgrounds. The appropriate control mice were selected in accordance with the publications that characterized both mouse lines[9,20]. CASK control mice express Cre while the first CASK exon is not floxed. SAP97 control mice are Cre-negative and the first SAP97 gene was floxed.

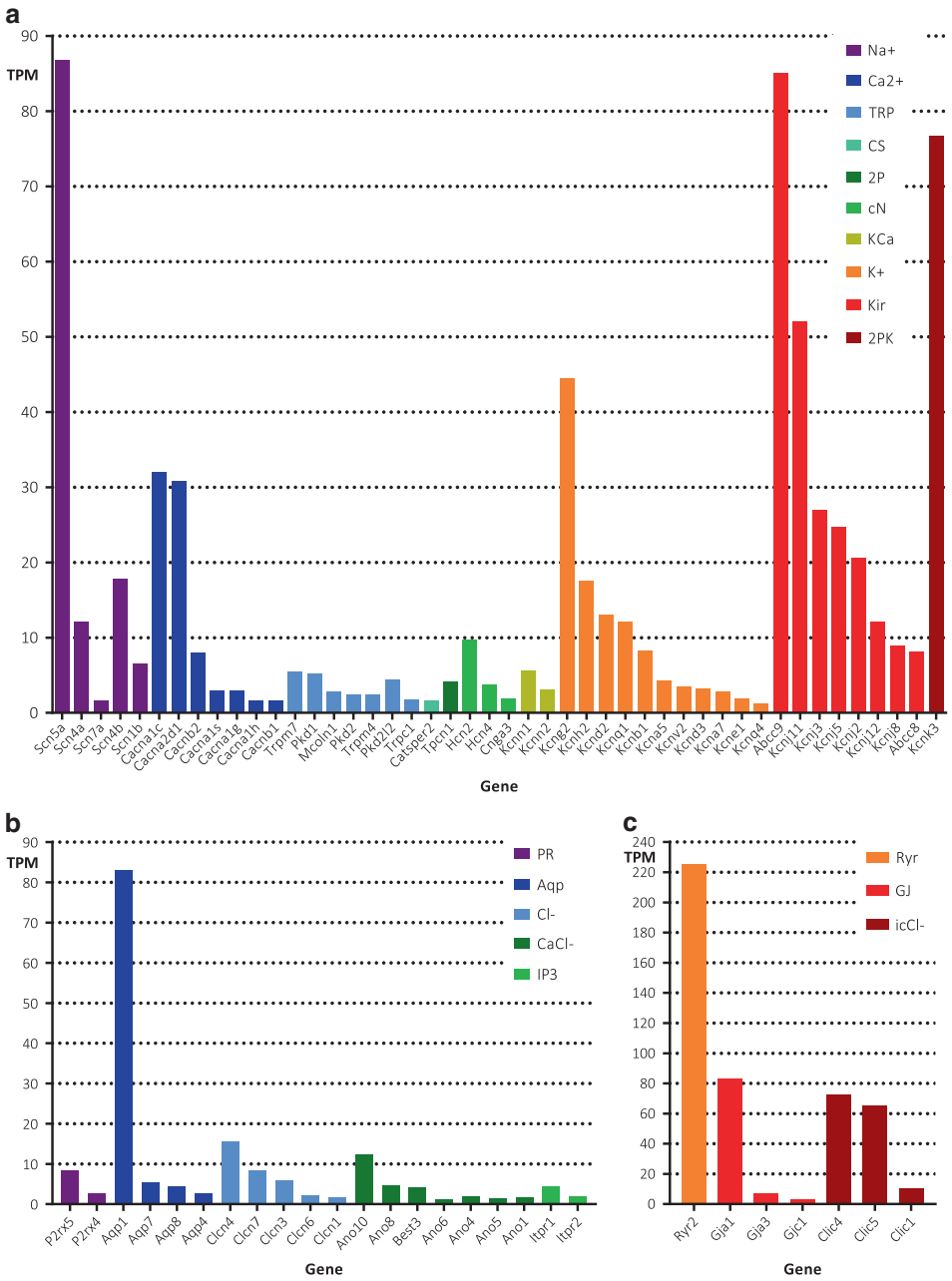

Figure 1. Gene expression of ion channels in murine ventricular cardiomyocytes. (**a**) Expression levels of voltage-gated ion channel genes: voltage-gated sodium channels (Na+; purple), voltage-gated calcium channels (Ca2+; blue), transient receptor potential cation channels (TRP; light blue), CS, CatSper channels (aqua), two-pore channels (2P; green), cyclic-nucleotide-regulated channels (cN; light green), calcium-activated potassium channels (KCa; ochre), voltage-gated potassium channels (K+; orange), inwardly rectifying potassium channels (Kir; red) and two-pore potassium channels (2PK; burgundy). (**b**) Expression levels of the ligand-gated purinergic receptor gene (PR; purple) and of ion channel genes from the "other" category: aquaporins (Aqp; blue), voltage-sensitive chloride channels (Cl-; light blue), calcium-activated chloride channels (CaCl-; green) and inositol triphosphate receptors (IP3; light green). (**c**) Expression levels of more ion channel genes from the "other" category: ryanodine receptors (Ryr; orange), gap junction proteins (GJ; red) and chloride intracellular channels (icCl-; burgundy). All expression levels are average TPM values of WT samples ($n = 5$). Shown are genes with more than 75 reads per gene (normalized for gene length, prior to conversion to TPM) from Tables 1–3 (available online only).

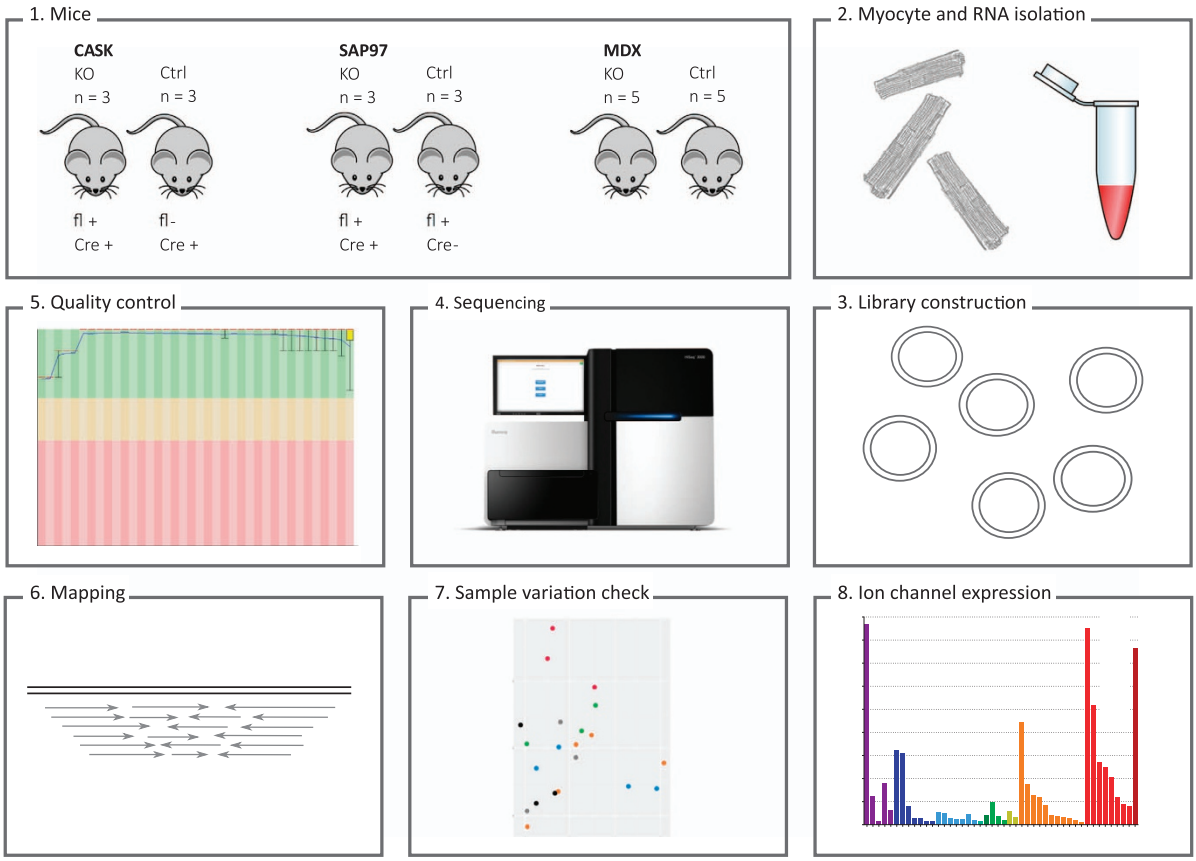

Figure 2. Experimental design and workflow. (1) 22 mice with six different genetic backgrounds (CASK KO and control, SAP97 KO and control, and MDX and control) were used. fl+, first exon of gene is floxed; Cre+, Cre recombinase is expressed. (2) Cardiomyocytes were isolated on a Langendorff system and RNA was isolated with a FFPE Clear RNAready kit. (3) Libraries were constructed with 1 µg RNA per sample using a TrueSeq Stranded Total RNA protocol and (4) sequenced on an Illumina HiSeq3000 machine. (5) Quality of the reads was assessed with FastQC, and (6) reads were mapped to the *Mus musculus* reference genome (GRCm38.83) with Tophat. (7) To assess sample variation within each group, we performed principle component analyses (PCA) (see Fig. 3). (8) Lastly, ion channel expression was determined.

Dystrophin knockout (MDX-5CV) mice. The MDX-5CV strain demonstrates total deletion of the dystrophin protein. It was created as previously described[21], and acquired from the Jackson laboratory (stock #002379). MDX mice were on pure Bl6/Ros backgrounds. Control mice were on pure Bl6/J background, except for MDX_Ct5 and MDX_5, which were Bl6/Ros mice backcrossed three times on Bl6/J.

Cardiomyocyte isolation

Mice ($n = 3$–5 per genotype, male, age 10–15 weeks) were heparinized (intraperitoneal injection of 100 µL heparin (5000 U/mL; Biochrom AG)) and killed by cervical dislocation. Hearts were excised, and the aortas were cannulated in ice-cold phosphate-buffered saline (PBS). Subsequently, hearts were perfused on a Langendorff system in a retrograde manner at 37 °C with 5 mL perfusion buffer (1.5 mL/min; in mM: 135 NaCl, 4 KCl, 1.2 NaH_2PO_4, 1.2 $MgCl_2$, 10 HEPES, 11 glucose), followed by the application of type II collagenase (Worthington CLS2; 25 mL of 1 mg/mL in perfusion buffer with 50 µM $CaCl_2$). Left and right ventricles were triturated in PBS to dissociate individual ventricular cardiomyocytes and then filtered through a 100 µm filter.

RNA extraction and sequencing

RNA-seq was performed by the Next Generation Sequencing Platform at the University of Bern. Total RNA was isolated from freshly dissociated cardiomyocytes with an FFPE Clear RNAready kit (AmpTec, Germany), which included a DNase treatment step. RNA quality was assessed with Qubit and Bioanalyzer, and RNA quantity was checked with Qubit.

Sample ID	Genotype	# read pairs total	# non-rRNA read pairs	% of total	Insert size	# read pairs mapping to a gene	% of total	# no-feature read pairs	% of total	# ambiguous read pairs	% of total
CASK_Ct1	WT+Cre	47,543,799	47,343,548	99.58	492	34,076,980	71.67	2,721,942	5.73	8,287,647	17.43
CASK_Ct2	WT+Cre	45,437,500	45,287,356	99.67	476	33,988,592	74.8	1,440,229	3.17	7,578,641	16.68
CASK_Ct3	WT+Cre	55,117,414	54,944,721	99.69	479	40,469,641	73.42	1,790,829	3.25	11,381,005	20.65
CASK_KO1	CASK_fl+Cre	45,685,573	45,565,815	99.74	472	32,765,612	71.72	5,738,568	12.56	5,504,670	12.05
CASK_KO2	CASK_fl+Cre	55,895,769	55,607,105	99.48	511	39,344,558	70.39	2,372,238	4.24	12,476,403	22.32
CASK_KO3	CASK_fl+Cre	56,437,329	56,008,449	99.24	499	42,159,185	74.7	2,804,256	4.97	9,655,232	17.11
MDX_1	MDX	17,485,935	17,320,513	99.05	380	*(sample exluded)*					
MDX_2	MDX	39,536,744	39,037,330	98.74	447	27,475,113	69.49	3,258,092	8.24	6,543,268	16.55
MDX_3	MDX	39,626,959	39,432,254	99.51	455	27,841,146	70.26	2,169,924	5.48	7,327,584	18.49
MDX_4	MDX	42,406,246	40,919,896	96.49	488	29,805,158	70.28	2,905,415	6.85	6,497,990	15.32
MDX_5	MDX	50,934,677	47,518,076	93.29	484	34,233,480	67.21	1,864,210	3.66	10,028,518	19.69
MDX_Ct1	WT	48,311,563	46,288,106	95.81	380	32,827,800	67.95	4,353,181	9.01	7,779,264	16.1
MDX_Ct2	WT	47,283,192	46,988,962	99.38	446	32,237,279	68.18	2,304,142	4.87	10,939,883	23.14
MDX_Ct3	WT	35,275,617	34,938,284	99.04	427	24,235,276	68.7	3,631,537	10.29	4,922,208	13.95
MDX_Ct4	WT	33,977,175	32,900,815	96.83	515	25,298,933	74.46	1,713,065	5.04	4,619,558	13.6
MDX_Ct5	WT	49,379,536	45,499,492	92.14	485	32,227,570	65.27	1,210,708	2.45	10,698,976	21.67
SAP_Ct1	WT+Cre	47,930,112	47,715,719	99.55	461	34,192,649	71.34	1,965,652	4.1	9,896,590	20.65
SAP_Ct2	WT+Cre	44,934,245	44,566,395	99.18	444	30,350,071	67.54	4,879,732	10.86	7,491,483	16.67
SAP_Ct3	WT+Cre	43,586,968	43,382,766	99.53	451	29,836,839	68.45	1,332,267	3.06	8,751,881	20.08
SAP_KO1	SAP_fl+Cre	44,319,566	44,146,526	99.61	452	34,155,235	77.07	2,606,959	5.88	5,090,692	11.49
SAP_KO2	SAP_fl+Cre	41,547,517	41,397,099	99.64	469	28,697,320	69.07	3,765,768	9.06	7,431,842	17.89
SAP_KO3	SAP_fl+Cre	46,143,349	45,812,985	99.28	443	30,710,635	66.55	5,476,238	11.87	8,174,538	17.72

Table 4. RNA-seq raw data and mapping metrics. Total and non-ribosomal RNA read pairs, average RNA fragment size (bp), and mapping metrics, including absolute number and percentages of read pairs mapping to all annotated exons of the mouse reference genome, and no-feature and ambiguous reads, per sample. Note the low number of read pairs in MDX_1, which is therefore excluded from further analysis. CASK KO and Ctrl, SAP97 KO and Ctrl *n*s = 3, MDX KO $n = 4$, MDX Ctrl $n = 5$.

To allow sequencing of long non-coding RNA (lncRNA), libraries were constructed with 1 µg RNA using the TruSeq Stranded Total RNA kit after Ribo-Zero Gold (Illumina) treatment for rRNA depletion. Library molecules with inserts < 300 base pairs (bp) were removed. Paired-end libraries (2x150 bp) were sequenced on an Illumina HiSeq3000 machine.

RNA-seq data analysis

Between 17.5 and 56.4 million read pairs were obtained per sample and the quality of the reads was assessed using FastQC v.0.11.2 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Ribosomal RNA (rRNA) was removed by mapping the reads with Bowtie2 v.2.2.1 (ref. 22) to a collection of rRNA sequences (NR_003279.1, NR_003278.3 and NR_003280.2) downloaded from NCBI (www.ncbi.nlm.nih.gov). No quality trimming was required.The remaining reads were mapped to the *Mus musculus* reference genome (GRCm38.83) with Tophat v.2.0.13 (ref. 23). We used htseq-count v.0.6.1 (ref. 24) to count the number of reads overlapping with each gene, as specified in the Ensembl annotation (GRCm38.83). Detailed information about the genes including the Entrez Gene ID, the MGI symbol and the description of the gene was obtained using the Bioconductor package BioMart v.2.26.1 (ref. 25).

Raw reads were corrected for gene length and TPM (transcripts per million) values were calculated to compare the expression levels among samples. Gene lengths for the latter step were retrieved from the Ensembl annotation (GRCm38.83) as the total sum of all exons.

Principal component analysis (PCA) plots were done in DESeq2 v.1.10.1 (ref. 26) (https://bioconductor.org/packages/release/bioc/html/DESeq2.html) using the 500 genes with the most variable expression across samples. A regularized log transformation was applied to the counts before performing the PCA.

Statistics

To assess differential gene expression between genotypes, a Wald test was performed with the Bioconductor package DESeq2 v.1.10.1 (ref. 26). We considered p values of up to 0.01, accounting for a Benjamini-Hochberg false discovery rate adjustment, to indicate significant difference. Statistical tools used included DESeq2, R-3.2.5 (https://cran.r-project.org), and Biomart_2.26.1 (www.biomart.org).

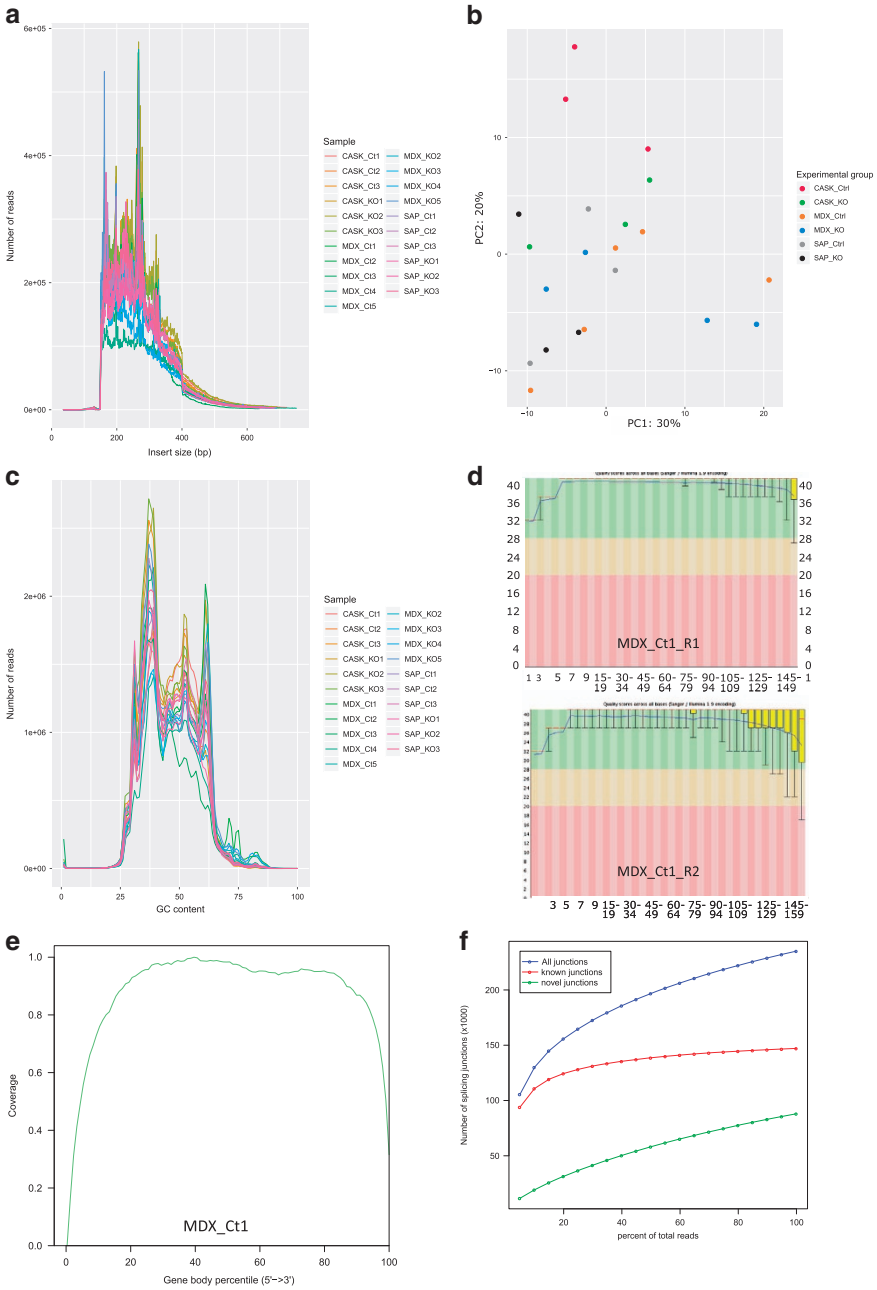

Figure 3. Quality control. (**a**) Histogram of inferred insert size for each sample, which represents distance between the two reads of one RNA fragment. (**b**) Principle component analyses (PCA) plots were performed to assess variability of samples within and between groups. Plot of the first two axes from a PCA based on the 500 genes with the most variable expression across all samples except MDX_1. CASK control (red, $n=3$) and KO (green, $n=3$); MDX control (orange, $n=5$) and KO (blue, $n=4$); SAP97 control (grey, $n=3$) and KO (black, $n=3$). (**c**) Distribution of GC content of the reads for each sample. (**d**) Base quality (Phred scores) along the length of the reads in each FastQC file of MDX_Ct1 as representative sample. The box plots are drawn as follows: red line, median; yellow box, range between upper and lower quartiles; whiskers, range between 10 and 90% quantiles. The blue line shows the mean quality. Y-axis represents quality scores across all bases. X-axis represents position in read (bp). (**e**) Gene body coverage. Distribution of reads along the length of the genes (5'-end on the left, 3'-end on the right). Shown image of sample MDX_Ct1 is representative for all samples. (**f**) Saturation report, depicting the number of splice junctions detected using different subsets of the data from 5 to 100% of all reads. Red, known junction based on the provided genome annotation; green, novel junctions; blue, all junctions. The red line reaches a plateau where adding more data does not increase the number of detected junctions, indicating that the sequencing depth suffices for performing alternative splicing analysis.

Data Records

The data were submitted to NCBI Gene Expression Omnibus (GEO) (Data Citation 1). This GEO project contains raw data and TPM values from all samples, and differential gene expression analysis between knockout and control samples.

Technical Validation

RNA metrics

RNA-seq yielded 1.0 billion read pairs in total, with an average of 44.5 million read pairs per sample (standard deviation 8.4 million). The number of read pairs (in millions) was 306 for CASK KO and Ctrl, 268 for SAP97 KO and Ctrl, and 404 for MDX and Ctrl (see Table 4 for an overview of RNA-seq metrics, including mapping rates). One sample (MDX_1) yielded few reads and was therefore excluded from further analyses. The proportion of reads mapping to annotated exons ranged from 65 to 77%. Mapping, no-feature (2–13%), and ambiguous (11–23%) read pairs together accounted for 89–97% of the total number of RNA reads (Table 4). Read pairs covered 49,671 genes of the *Mus musculus* reference genome (GRCm83.38).

Quality assessment

The quality of all samples was assessed with FastQC. Except for MDX_1, all samples were of high quality. Where applicable, a representative example (MDX_Ct1) is shown. Firstly, the insert size histogram (Fig. 3a) shows that the inferred insert size of each sample exceeded 150 base pairs, demonstating that the sequencing was not contaminated by adapter sequences. Secondly, the GC content plot (Fig. 3c) ideally shows a roughly normal distribution centred around the average GC content of the genome, which varies between species. The peaks observed in Fig. 3c are likely caused by sequences that are detected at high copy numbers, and should not pose problems for downstream analyses. Furthermore, Phred scores (Fig. 3d) are well within the green area of the graph indicating good base quality along the length of reads. As well, the gene coverage graph (Fig. 3e) of sample MDX_Ct1 shows that reads are distributed evenly along the length of the gene body. Because the gene coverage for all other samples is highly comparable to that of MDX_Ct1, only one example is shown. Lastly, the saturation report (Fig. 3f) represents the number of splice junctions detected using different subsets of the data from 5 to 100% of all reads. At sequencing depths sufficient to perform alternative splicing analysis, at least the red line, representing known junctions, should reach a plateau where adding more data does not much increase the number of detected junctions. Only MDX_1 does not reach this plateau.

Gene expression variation of biological replicates

We performed Principle Component Analyses (PCA) to assess whether samples from the same experimental group have similar gene expression profiles (Fig. 3b). Of note, samples within each sample group still show considerable variation. The mixed genetic background of most sample groups may explain this variation; only the MDX control mice are on a pure Bl6/J background. The variation seen in MDX control mice is likely due to a batch effect, as two rounds of samples were sequenced. However, considering that PCA plots are based on the 500 genes with the highest variability in one sample, our genes of interest, including all ion channel genes, show similar expression levels throughout all samples.

Ion channel expression

Based on the list of ion channel genes from HUGO Gene Nomenclature Committee (https://www.genenames.org/cgi-bin/genefamilies/set/177), we distilled ion channel expression from WT mice expressed as TPM (Tables 1–3 (available online only), Fig. 1).

References

1. Cerrone, M. *et al.* Plakophilin-2 is required for transcription of genes that control calcium cycling and cardiac rhythm. *Nat Commun* **8,** 106 (2017).
2. Harrell, M. D., Harbi, S., Hoffman, J. F., Zavadil, J. & Coetzee, W. A. Large-scale analysis of ion channel gene expression in the mouse heart during perinatal development. *Physiol. Genomics* **28,** 273–283 (2007).
3. Petitprez, S. *et al.* SAP97 and dystrophin macromolecular complexes determine two pools of cardiac sodium channels Nav1.5 in cardiomyocytes. *Circ. Res.* **108,** 294–304 (2011).
4. Zimmer, T., Haufe, V. & Blechschmidt, S. Voltage-gated sodium channels in the mammalian heart. *Glob. Cardiol. Sci. Pract* **4,** 449–463 (2014).
5. Balse, E. *et al.* Dynamic of ion channel expression at the plasma membrane of cardiomyocytes. *Physiol. Rev.* **92,** 1317–1358 (2012).
6. Westenbroek, R. E. *et al.* Localization of sodium channel subtypes in mouse ventricular myocytes using quantitative immunocytochemistry. *J. Mol. Cell. Cardiol.* **64,** 69–78 (2013).
7. Stroud, D. M. *et al.* Contrasting Nav1.8 Activity in Scn10a-/- Ventricular Myocytes and the Intact Heart. *J. Am. Heart Assoc* **5,** 1–10 (2016).
8. Leonoudakis, D., Conti, L. R., Radeke, C. M., McGuire, L. M. & Vandenberg, C. A. A multiprotein trafficking complex composed of SAP97, CASK, Veli, and Mint1 is associated with inward rectifier Kir2 potassium channels. *J. Biol. Chem.* **279,** 19051–19063 (2004).
9. Eichel, C. A. *et al.* Lateral membrane-specific MAGUK CASK down-regulates NaV1.5 channel in cardiac myocytes. *Circ. Res.* **119,** 544–556 (2016).
10. Abriel, H., Rougier, J. S. & Jalife, J. Ion channel macromolecular complexes in cardiomyocytes: roles in sudden cardiac death. *Circ. Res.* **116,** 1971–1988 (2015).

11. Nafzger, S. & Rougier, J. S. Calcium/calmodulin-dependent serine protein kinase CASK modulates the L-type calcium current. *Cell Calcium* **61,** 10–21 (2017).
12. Hsueh, Y. P., Wang, T. F., Yang, F. C. & Sheng, M. Nuclear translocation and transcription regulation by the membrane-associated guanylate kinase CASK/LIN-2. *Nature* **404,** 298–302 (2000).
13. Sun, R. *et al.* Human calcium/calmodulin-dependent serine protein kinase regulates the expression of p21 via the E2A transcription factor. *Biochem. J.* **419,** 457–466 (2009).
14. Ashcroft, F. M. From molecule to malady. *Nature* **440,** 440–447 (2006).
15. Rickert-Sperling, S., Kelly, R. G. & Driscoll, D. J. *Congenital Heart Diseases: The Broken Heart* 721–736 (Springer-Verlag, 2016).
16. Lahrouchi, N., Behr, E. R. & Bezzina, C. R. Next-generation sequencing in post-mortem genetic testing of young sudden cardiac death cases. *Front. Cardiovasc. Med* **3,** 13 (2016).
17. Heinig, M. A. *et al.* Natural genetic variation of the cardiac transcriptome in non-diseased donors and patients with dilated cardiomyopathy. *Genome Biol.* **18,** 170 (2017).
18. van den Berg, C. W. *et al.* Transcriptome of human foetal heart compared with cardiomyocytes from pluripotent stem cells. *Development* **142,** 3231–3238 (2015).
19. Agah, R. *et al.* Gene recombination in postmitotic cells. Targeted expression of Cre recombinase provokes cardiac-restricted, site-specific rearrangement in adult ventricular muscle in vivo. *J. Clin. Invest.* **100,** 169–179 (1997).
20. Gillet, L. *et al.* Cardiac-specific ablation of synapse-associated protein SAP97 in mice decreases potassium currents but not sodium current. *Heart Rhythm* **12,** 181–192 (2015).
21. Chapman, V. M., Miller, D. R., Armstrong, D. & Caskey, C. T. Recovery of induced mutations for X chromosome-linked muscular dystrophy in mice. *Proc. Natl. Acad. Sci. USA* **86,** 1292–1296 (1989).
22. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9,** 357–359 (2012).
23. Kim, D. *et al.* TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14,** R36 (2013).
24. Anders, S., Pyl, P. T. & Huber, W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31,** 166–169 (2015).
25. Durinck, S. *et al.* BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics* **21,** 3439–3440 (2005).
26. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15,** 550 (2014).

Data Citation

1. *Gene Expression Omnibus* GSE102772 (2017).

Acknowledgements

The research was financially supported by Swiss National Science Foundation Grant 310030_165741 (HA). We thank the Next Generation Sequencing Platform of the University of Bern, in particular Muriel Fragnière, for performing the high-throughput sequencing experiments; prof. Rolf Jaggi for excellent technical advice on RNA isolation from cardiomyocytes; and Joseph Allan for language editing.

Author Contributions

M.C. planned and performed experiments on KO and control samples, and submitted data to the repositories. S.H.V. wrote the Data Descriptor, analyzed the data, and made tables and figures. K.W. performed bioinformatic analyses, and made tables and figures. I.K. performed bioinformatic analyses, made tables and figures, and co-supervised the work. L.G. performed experiments on WT samples. H.A. planned the experiments, obtained funding, and supervised the project

Additional Information

Tables 1, 2 and 3 are available only in the online version of this paper.

Competing interests: The authors declare no competing interests.

How to cite this article: Chevalier, M. *et al*, Transcriptomic analyses of murine ventricular cardiomyocytes. *Sci. Data* 5:180170 doi: 10.1038/sdata.2018.170 (2018).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

SCIENTIFIC DATA

Corrigendum: Transcriptomic analyses of murine ventricular cardiomyocytes

Published: 9 October 2018

Morgan Chevalier, Sarah H. Vermij, Kurt Wyler, Ludovic Gillet, Irene Keller & Hugues Abriel

Correction to: *Scientific Data* https://doi.org/10.1038/sdata.2018.170, published online 21 August 2018.

In the CASK and SAP97 knockout mice sub-section of the Methods in this Data Descriptor it is incorrectly stated that CASK KO mice were generated as previously published in references 9 and 20. This is incorrect. Instead, mice in which the first coding exon of the CASK gene is flanked by *loxP* sites ($CASK^{tm1Sud}$, purchased from the Jackson Laboratory, stock #006382) were crossed with αMHC-Cre mice.

In the same section of the Data Descriptor it is incorrectly stated that in SAP97 mice the first SAP97 gene was floxed. In actuality the first three coding exons were floxed.

In Table 4 of the Data Descriptor, the genotype of samples SAP_Ct1, SAP_Ct2 and SAP_Ct3 is incorrectly listed as WT+Cre. The correct genotype is WT_fl, consistent with Figure 2 and the Mouse models sub-section of the Methods.

Correspondence and requests for materials should be addressed to H.A. (email: hugues.abriel@ibmm.unibe.ch)