Open Access

Repository of the Max Delbrück Center for Molecular Medicine (MDC) in the Helmholtz Association

http://edoc.mdc-berlin.de/16360

A comprehensive TALEN-based knockout library for generating human induced pluripotent stem cell-based models for cardiovascular diseases

Karakikes, I. and Termglinchan, V. and Cepeda, D.A. and Lee, J. and Diecke, S. and Hendel, A. and Itzhaki, I. and Ameen, M. and Shrestha, R. and Wu, H. and Ma, N. and Shao, N.Y. and Seeger, T. and Woo, N.A. and Wilson, K.D. and Matsa, E. and Porteus, M.H. and Sebastiano, V. and Wu, J.C.

This is the final version of the accepted manuscript. The original article has been published in final edited form in:

Circulation Research
2017 MAY 12 ; 120(10): 1561-1571
2017 FEB 28 (originally published online)
doi: 10.1161/CIRCRESAHA.116.309948

Publisher: American Heart Association
© 2017 American Heart Association, Inc.

NEW METHODS

A Comprehensive TALEN-Based Knockout Library for Generating Human Induced Pluripotent Stem Cell-Based Models for Cardiovascular Diseases

Ioannis Karakikes[1,2], Vittavat Termglinchan[1,3], Diana A. Cepeda[4], Jaecheol Lee[1], Sebastian Diecke[1,8,9], Ayal Hendel[5], Ilanit Itzhaki[1], Mohamed Ameen[1], Rajani Shrestha[1], Haodi Wu[1], Ning Ma[1], Ning-Yi Shao[1], Timon Seeger[1], Nicole Woo[1], Kitchener D. Wilson[1,6], Elena Matsa[1], Matthew H. Porteus[5], Vittorio Sebastiano[4,7], Joseph C. Wu[1,3,4]

[1]Stanford Cardiovascular Institute; [2]Department of Cardiothoracic Surgery; [3]Department of Medicine, Division of Cardiovascular Medicine; [4]Institute of Stem Cell Biology and Regenerative Medicine; [5]Department of Pediatrics; [6]Department of Pathology; [7]Department of Obstetrics and Gynecology, Stanford University School of Medicine, Stanford, CA 94305, USA; [8]Berlin Institute of Health, Germany; [9]Max Delbrueck Center, Berlin, Germany

I.K., V.T., and D.C. contributed equally to this study.

Running title: Cardiomyopathy TALEN Knockout Library

Subject Terms:
Basic Science Research
Cardiomyopathy
Congenital Heart Disease
Genetically Altered and Transgenic Models

Address correspondence to:

Dr. Joseph C. Wu
265 Campus Drive G1120B
Stanford, CA 94305-5454
joewu@stanford.edu

Dr. Vittorio Sebastiano
265 Campus Drive G1167
Stanford, CA 94305-5454
vsebast@stanford.edu

This manuscript was sent to Roger J. Hajjar, Consulting Editor, for review by expert referees, editorial decision, and final disposition.

In January 2016, the average time from submission to first decision for all original research papers submitted to *Circulation Research* was 13.77 days.

ABSTRACT

Rationale: Targeted genetic engineering using programmable nucleases such as transcription activator–like effector nucleases (TALENs) is a valuable tool for precise, site-specific genetic modification in the human genome.

Objective: The emergence of novel technologies such as human induced pluripotent stem cells (iPSCs) and nuclease-mediated genome editing represent a unique opportunity for studying cardiovascular diseases in vitro.

Methods and Results: By incorporating extensive literature and database searches, we designed a collection of TALEN constructs to knockout (KO) eighty-eight human genes that are associated with cardiomyopathies and congenital heart diseases. The TALEN pairs were designed to induce double-strand DNA break near the starting codon of each gene that either disrupted the start codon or introduced a frameshift mutation in the early coding region, ensuring faithful gene KO. We observed that all the constructs were active and disrupted the target locus at high frequencies. To illustrate the general utility of the TALEN-mediated KO technique, six individual genes (*TNNT2, LMNA/C, TBX5, MYH7, ANKRD1,* and *NKX2.5*) were knocked out with high efficiency and specificity in human iPSCs. By selectively targeting a dilated cardiomyopathy (DCM)-causing mutation (*TNNT2 p.R173W*) in patient-specific iPSC-derived cardiac myocytes (iPSC-CMs), we demonstrated that the KO strategy ameliorates the DCM phenotype in vitro. In addition, we modeled the Holt-Oram syndrome (HOS) in iPSC-CMs in vitro and uncovered novel pathways regulated by *TBX5* in human cardiac myocyte development.

Conclusion: Collectively, our study illustrates the powerful combination of iPSCs and genome editing technology for understanding the biological function of genes and the pathological significance of genetic variants in human cardiovascular diseases. The methods, strategies, constructs and iPSC lines developed in this study provide a validated, readily available resource for cardiovascular research.

Keywords:
Genome editing, iPSCs, gene knockout, dilated cardiomyopathy, Holt-Oram syndrome, stem cell, cardiac, gene targeting, disease modeling.

Nonstandard Abbreviations and Acronyms:

cTAL	cardiomyopathy TALEN-based
DSB	Double-strand break
ECM	Extracellular matrix
HOS	Holt-Oram syndrome
iPSCs	induced pluripotent stem cells
iPSC-CMs	iPSC-derived cardiac myocytes
NHEJ	non-homologous end joining
SMRT	single-molecule real-time
TALENs	transcription activator–like effector nucleases
TSS	transcription starting sites
EADs	Early after depolarizations
SCVI	Stanford Cardiovascular Institute

INTRODUCTION

Cardiovascular disease is a major cause of morbidity and mortality around the world. In recent years, exciting progress has been made in defining the etiology of congenital heart disease (CHD)[1] and inherited cardiomyopathies,[2] including hypertrophic cardiomyopathy (HCM), dilated cardiomyopathy (DCM), left ventricular non-compaction (LVNC), and arrhythmogenic right ventricular cardiomyopathy/dysplasia (ARVC/D). Recent advances in genomics and molecular medicine have identified genetic mutations in plethora of genes that are implicated in the pathogenesis of inherited cardiomyopathies. Although the molecular analysis efforts have revealed important insights regarding the role of genetics in cardiomyopathies, the underlying molecular mechanisms remain poorly understood and the genotype-phenotype relationship from the ever-growing number of disease-associated gene mutations remains to be established.

Recent advances in technologies for genome editing using site-specific nucleases,[3] such as zinc finger nucleases (ZFNs), transcription activator–like effector nucleases (TALENs), and the clustered regularly interspaced short palindromic repeat/CRISPR-associated protein 9 (CRISPR/Cas9) system, offer a powerful tool for reverse genetics, genome engineering, and targeted transgene integration experiments to be performed in a precise and predictable manner. The use of engineered nucleases to make targeted, permanent changes to genes have revolutionized molecular genetics and present an alternative to the more established method of RNA interference (RNAi)-mediated knockdown using short hairpin RNA (shRNA) or short interfering RNA (siRNA). However, the RNAi-mediated post-transcriptional down-regulation of gene expression without changing the genetic code does not completely shut off the gene of interest.[4] In most cases, functional RNA or protein remains and is translated albeit at lower levels. Thus, the gene function is reduced, but not eliminated. By contrast, genome editing changes the genetic code and typically causes a functional "knockout" (KO), or complete elimination of the gene function. The nucleases cut both DNA strands of the targeted locus generating a double-strand break (DSB) in the chromosome, which is then repaired by the non-homologous end joining (NHEJ) mechanism that re-ligates the two free chromosome ends. However, NHEJ is error-prone, often resulting in small insertions or deletions that can disrupt, or knockout, the gene of interest.

Over the past decade, the advent of the human induced pluripotent stem cell (iPSC) technology, and the improvements in the differentiation method of iPSCs into specific cell types, such as cardiac myocytes (iPSC-CMs)[5], endothelial cells (iPSC-ECs),[6] and smooth muscle cells (iPSC-SMCs),[7] provide an unprecedented opportunity for the generation of patient-specific in vitro models for disease modeling. Combining genome editing and iPSC technologies can successfully create human-based cell knockout models in vitro. Such models could improve our understanding of the underlying pathological mechanisms, and potentially lead to novel therapies.[8]

In this study, we describe the design, construction, and validation of a cardiomyopathy TALEN-based (cTAL) panel to knock out a comprehensive set of genes associated with cardiovascular diseases. We demonstrated the utility of this panel, and presented two case studies that provided novel insights into the pathogenesis of genetic cardiovascular disease. The readily available cTAL panel will allow researchers to fast-track projects by providing a validated panel of TALEN constructs for gene KO genome editing. This approach could provide novel insights into gene function, disease mechanisms, and ultimately disease pathogenesis.

METHODS

TALEN construction.

TALEN genomic binding sites were designed using the TAL Effector Nucleotide Targeter 2.0,[9] with the following constraints: (i) having a repeat array length of 15 repeat variable di-residue domains, and (ii) having a spacer length of 14–18 nucleotides. A preceding T base in position "0" anchored each binding site as has been shown to be optimal for naturally occurring TAL proteins.[10,11] Each custom TALEN was generated from a library of 832 plasmids through a five-piece subcloning ligation: three sequence-specific tetramer-recognition pieces, one trimer-recognition piece, and an expression vector backbone (pTAL) as previously described.[12] Briefly, the tetramer or trimer TAL repeats were digested out of library plasmids with the restriction enzyme BsmBI (NEB), gel purified, and subcloned into the pTAL vectors. The forward and reverse TALENs were subcloned into the pTAL_GFP and pTAL_RFP backbones, respectively. The sequences of all constructs used in this study are provided in the Supplemental Information. The TALEN plasmids will be available from Addgene. The cell lines are available upon request from the Stanford CVI iPSC Biobank (http://med.stanford.edu/scvibiobank.html).

Culture and cardiac differentiation of iPSCs.

The human iPSC lines (SCVI-15, SCVI-114, and SCVI-19) were obtained from the Stanford CVI iPSC Biobank. The iPSCs were maintained under feeder-free conditions in defined E8 media (Life Technologies) on tissue culture plates coated with hESC-qualified Matrigel (BD Biosciences) in 5% CO_2/5% O_2/90% N_2 environment at 37°C. Human iPSCs were differentiated toward cardiac myocytes using a small molecule mediated directed differentiation protocol.[13] Briefly, cardiac differentiation was initiated by treatment with recombinant BMP4 and Activin A (Day 0-3), followed by treatment with 5 µM IWR-1 for 72 hr (day 4 to day 6).

TALEN transfection.

Human iPSCs were enzymatically dissociated with Accutase (Sigma) and plated on Matrigel coated dishes at 1:3 ratio in E8 supplemented with 10 um Y-27632 (Selleck Chemicals). 24 – 48 hr later, human iPSCs were dissociated with Accutase into single cells. ~2x10^6 cells were transfected with a pair of TALENs (1.0 µg of each TALEN) by nucleofection using the Amaxa 4D Nucleofector system (Lonza) with the P3 Primary Cell Nucleofector Kit and program CM-150 per manufacturer's instructions (Lonza). Following nucleofection, iPSCs were re-suspended in 1 ml pre-warmed E8 supplemented with 5 µM Thiazovivin and then plated in 6-well plates pre-coated with Matrigel and allowed to recover for 48 hr.

SMRT sequencing.

Genome-editing outcomes at the endogenous loci were quantified using single-molecule real-time (SMRT) DNA sequencing as previously described.[14] Genomic DNA was extracted from TALEN-transfected iPSCs at 72 hr post-nucleofection without enrichment for transfected cells, and used as a template for PCR amplification using primer pairs designed to amplify a ~500 bp fragment surrounding the TALEN targeted loci. The PCR amplicons were purified using the nucleotide removal kit (Qiagen) and the sequencing libraries were constructed using the DNA Template Prep Kit 1.0 (Pacific Biosciences). SMRTbell libraries contained amplicons that were pooled together with different barcodes appended to allow multiplex analysis. Purified, closed circular SMRTbell libraries were annealed with a sequencing primer complementary to a portion of the single-stranded region of the hairpin. For all SMRTbell libraries, annealing was performed at a final template concentration between 30 and 60 nM, with a 20-fold molar excess of sequencing primer. All annealing reactions were carried out at 80°C for 2 min, with a slow cool to 25°C at a rate of 0.1°C/s. Annealed templates were stored at −20°C until polymerase binding. DNA polymerase enzymes were stably bound to the primed sites of the annealed SMRTbell templates using the DNA Polymerase Binding Kit 2.0 (Pacific Biosciences). SMRTbell templates (3 nM) were incubated with 6 nM of polymerase in the presence of phospholinked nucleotides at 30°C for 2 hr. Following incubation, samples were stored at 4°C. Sequencing was performed within 72 hr of binding using a final concentration

of 0.3 nM. Each sample was sequenced using the DNA Sequencing Kit 2.0 (Pacific Biosciences). Sequencing data collection was performed on the PacBio RS (Pacific Biosciences) using C2/C2 chemistry and movies of 55 min in each case. The SMRT Sequencing Analysis pipeline was implemented in Strawberry Perl and utilizes the NCBI BLAST software as well as the mEmboss Needleman-Wunsch pairwise alignment algorithm.

Isolation of targeted clonal cell populations.

TALEN-transfected iPSCs were washed once with PBS and enzymatically dissociated with Accutase for 3-5 min at 37°C followed by gently pipetting to ensure single cell suspension. The cells were washed once in PBS and re-suspended in E8 supplemented with Y-27632 (10 um). Double GFP^{+}/RFP^{+} cells were then sorted by fluorescence activated cell sorter (FACSAria II; BD Biosciences), plated on 6-well plates at a clonal density of 1,000 cells/well and allowed to recover. After 7-10 days, putative single cell- derived clones were manually picked, expanded, and maintained in standard conditions.

RNA-sequencing.

Total RNA was isolated with the RNeasy Isolation kit with on-column DNase I treatment (Qiagen), and the quality of the RNA samples was assessed using the Agilent Bioanalyzer 2100 (Agilent). ERCC spike-in synthetic transcripts were added at manufacturer's recommended amounts (Life Technologies) and 1 μg of each RNA was enriched for poly-A RNA using the Dynabeads® mRNA Direct Kit (Life Technologies) per manufacturer's protocol. Whole transcriptome library preparation was performed using 5-10 ng of fragmented enriched poly-A RNA according to the manufacturer's protocol (Ion Total RNA-Seq Kit V2 protocol; Life Technologies), followed by purification with AMPure beads (Beckman-Coulter Genomics). The quality and quantity of the libraries was assessed using the Agilent Bioanalyzer High Sensitivity Chip (Agilent). Each library concentration was adjusted to 100 pM and 70 ul were used for Ion Template preparation in the automated Ion Chef system and loaded on the Ion PI Chip Kit v2 (Life Technologies). Sequencing was performed in the Ion Proton sequencing platform using the Ion PI™ Sequencing 200 Kit v3 per manufacturer's protocol (Life Technologies). Base calls were collected with Ion Torrent Suite software (Life Technologies).

***Allelic discrimination by digital PCR*.**

Total RNA was extracted from iPSC-CMs at day 30 post-differentiation using RNeasy Mini Kit (Qiagen), and complementary DNA (cDNA) preparation was carried out using the iScript cDNA Synthesis Kit (Bio-Rad Laboratories). The concentration of cDNA was reduced to about 0.2 ng/μl RNA equivalent and 1 ng (5 μl of 0.2 ng/μl) of RNA-equivalent cDNA was mixed with primers, probes and ddPCR Supermix reaction (total volume 20 μl). The final concentrations of the primers and the probe were 900 nM and 500 nM, respectively. The following primers and probes were used for discriminating the expression of the R173W and the WT *TNNT2* alleles; Fw: GAGGAGGAGAACAGGAG and Rv: GCATCATGTTGGACAAAGCC. wt-probe: [FAM]AGGATGAGGCCCGGAAGAAGA[BHQ] and mt-probe [HEX]AGGATGAGGCCTGGAAGAAGA[BHQ]. Droplet formation was carried out using a QX100 droplet generator. A rubber gasket is placed over the cartridge and loaded into the droplet generator. The emulsion (35 μL in volume) was then slowly transferred using a multichannel pipette to a 96-Well twintec™ PCR Plates (Eppendorf). The plate was heat-sealed with foil and the emulsion was cycled to end point per the manufacturer's protocol with an annealing temperature at 61°C. Finally, the samples were analyzed using a BioRad QX100 reader. The expression of TNNT2 was quantified by Real-Time qPCR (Applied Biosystems) using a custom TaqMan probe designed to detect the wild type transcript after TALEN-mediated KO (Fw: AGACGCCTCCAGGATCTGT, Rv: GCTTCTTCCTGCTCCTCCTC, Probe: [FAM]CAGACATGGTCTCTGCTCTCCCTC[BHQ].

TALEN off-target detection.

Genomic DNA was extracted from genome edited iPSC clones using the DNeasy Blood & Tissue Kit (Qiagen). The potential TALEN off-target sites were predicted *in silico* based on sequence homology using

the bioinformatics tool PROGNOS.[15] The primers designed by PROGNOS were used to amplify the genomic regions of putative off-target sites by PCR. The PCR products were analyzed by Sanger sequencing.

ChIP-seq analysis.

The raw Fastq files of ChIP-seq were aligned to human genome (hg19) by TMAP (https://github.com/iontorrent/TS/tree/master/Analysis/TMAP), and then all duplicate reads aligned to same loci were removed.[16] Peak calling was applied by HOMER,[17] and the parameters are: style "factor", genome "hg19", fold-change cutoff 4.0 of DNA input, fold-change cutoff of peak calling 2.0, and p-value cutoff 0.0001. Peaks were annotated by HOMER, and the nearest genes were assigned as the genes of the peaks. All sequences around coding promoters (upstream 400 bp, downstream 100 bp) were extracted and motif enrichment analysis was performed using HOMER. Then KEGG enrichment analysis was performed using the GeneAnswers package (http://www.bioconductor.org/packages/release/bioc/html/GeneAnswers.html), and adjusted p-value cutoff was 0.1. All alignment bam files were processed by IGVTools, and loaded to IGV genome browser[18] for the visualization of specific genes, all tracks normalized to 1 million reads.

Whole-cell patch-clamp recordings.

Contracting monolayer iPSC-CMs were enzymatically dispersed (Accutase, Sigma) and attached to Matrigel-coated glass coverslips (Warner, USA) for whole-cell patch clamp recordings. These recordings were conducted using an EPC-10 patch clamp amplifier (HEKA, Germany). 3-4 MΩ glass pipettes were prepared using thin-wall borosilicate glass (A-M System, USA) with a micropipette puller (Sutter Instrument, P-97, USA). Action potentials (APs) were recorded from iPSC-CMs suffused with Tyrode's solution at 37°C. The Tyrode's solution consisted of NaCl (140 mM), KCl (5.4 mM), $CaCl_2$ (1.8 mM), $MgCl_2$ (1 mM), HEPES (10 mM), and glucose (10 mM); pH was adjusted to 7.4 with NaOH. The pipette solution consisted of KCl (120 mM), $MgCl_2$ (1 mM), Mg-ATP (3 mM), HEPES (10 mM), and EGTA (10 mM); pH was adjusted to 7.2 with KOH. Data were acquired using PatchMaster software (HEKA, Germany) and digitized at 1.0 kHz. Data were analyzed using a custom-written MATLAB program.

Statistical analysis.

Unpaired two-tailed Student's *t* tests were used to determine the significance between two groups, assuming significance at $P < 0.05$. The one-way analysis of variance (ANOVA) was used to determine whether there are any statistically significant differences among the means of three or more groups, with $P < 0.05$ considered statistically significant. All values are expressed as the mean ± SEM.

RESULTS

Design, construction, and characterization of TALEN constructs.

We selected 88 genes associated with cardiomyopathies and congenital heart diseases (Figure 1a and Online Table I), including genes implicated in syndromes for which clinical diagnosis may be challenging, such as CHARGE syndrome (chromodomain helicase DNA binding protein 7 (*CHD7*) mutation), Leigh syndrome (*SURF1* mutations), Holt-Oram syndrome (TBX5 mutations), Noonan syndrome, LEOPARD syndrome, Raf-1 proto-oncogene, serine/threonine kinase (*RAF1*) and protein tyrosine phosphatase, non-receptor type 11 (*PTPN11*) mutations. To knock out these genes in the human genome, we designed TALENs that target sequences located around the start codon, ATG, of each gene (Figure 1b). We constructed one TALEN pair construct for each gene using a library of pre-assembled tetramers/trimers through a five-piece subcloning ligation.[12] The details of the TALEN design for each gene and the respective target site are shown in Online Table I. To validate the genome editing activities of the TALEN library in human iPSCs, we quantified the level of NHEJ using the SMRT technology.[14] Every TALEN pair tested was active and efficiently induced small deletions, insertions, or both at the target sites (Online Table II). The individual TALEN pairs induced mutations with a frequency ranging from 0.5% to 50% (Figure 1c), and the majority of the TALEN-mediated NHEJ outcomes were deletions of variable lengths within the spacer region, while insertion mutations were only observed in a few instances (Online Table II).

To illustrate the general utility of the TALEN-mediated NHEJ technique, we next targeted six individual genes (*TNNT2, TBX5,* lamin A/C *(LMNA/C),* myosin, heavy chain 7, cardiac muscle, beta *(MYH7),* ankyrin repeat domain 1 (cardiac muscle) *(ANKRD1),* and NK2 homeobox 5 *(NKX2.5)*) in human iPSCs. After TALEN transfection and FACS sorting, we screened single cell-derived clones for NHEJ events. We observed that the targeted loci were disrupted at high efficiency, with indels occurring in 33% to 100% of the clones screened (Table 1). These results indicate that all of our TALEN constructs are highly active and can be used for gene KO experiments.

Targeted disruption of the cardiac troponin T gene causes sarcomere disassembly.

Mutations associated with cardiomyopathies are commonly inherited in an autosomal dominant manner. Mutant proteins are thought to act through a dominant-negative mode that either interfere with normal function or assume a new function. In some instances, the mutant allele is inactivated, resulting in haploinsufficiency whereby a single functional copy of the gene is insufficient to maintain the normal phenotype. Although mutations in the cardiac troponin T (*TNNT2*) gene are commonly implicated in familial HCM, distinct mutations can also lead to DCM.[19] To address whether haploinsufficiency of *TNNT2* is responsible for HCM or DCM, we ablated either one or both *TNNT2* alleles in human iPSCs by TALEN-mediated gene KO in a single round of TALEN targeting. We generated both monoallelic (heterozygous) KO (*TNNT2*$^{+/-}$) and biallelic (homozygous) KO (*TNNT2*$^{-/-}$) iPSC lines (Figure 2a). These TNNT2-KO iPSC lines retained their pluripotency as assessed by immunostainining and gene expression assays of pluripotency markers (Online Figure I). Upon differentiation, the cardiac Troponin T protein (cTnT) was not detected in *TNNT2*$^{-/-}$ iPSC-CMs, while comparable levels of cTnT were observed in wild-type and *TNNT2*$^{+/-}$ iPSC-CMs (Figure 2b). At the mRNA level, *TNNT2*$^{+/-}$ iPSC-CMs had reduced expression of the non-targeted transcript compared to the parental iPSC-CMs (Figure 2c), suggesting that the cTnT protein levels are not regulated at the transcription level. Most likely a post-transcriptional mechanism, such as an increase in ribosome translational kinetics or lower protein turnover rates, is responsible for the comparable levels of cTnT protein expression in the *TNNT2*$^{+/-}$ and WT iPSC-CMs. At the functional level, we observed that *TNNT2*$^{-/-}$ iPSC-CMs displayed severe sarcomeric disarray (Figure 2d) and exhibited impaired intracellular Ca^{2+} cycling (Online Figure II). In contrast, *TNNT2*$^{+/-}$ iPSC-CMs showed no functional or structural abnormalities, suggesting that one TNNT2 allele is sufficient to maintain normal cTnT protein

expression and cardiac myocyte structure and function (Figure 2 and Online Figure II). These results suggest that haploinsufficiency is unlikely to explain the pathogenesis of cardiomyopathies associated with *TNNT2* mutations.

Phenotypic rescue of DCM by targeted allelic-specific KO in vitro.

To test this hypothesis, we next disrupted the starting codon of *TNNT2* gene in a patient-specific iPSC line harboring a missense mutation in exon 12 of the *TNNT2* gene (NM_001001430.2: *c.517 C>T; p.R173W*) (Figure 3a).[20] We screened the TALEN-targeted clones for NHEJ events, and identified an iPSC clone with a disruption of the starting codon of the mutant *TNNT2 p.R173W* allele (hereafter referred to as DCM-KO) and without any detectable off-target mutations (Figure 3a and Online Table III). This isogenic KO line retained pluripotency as assessed by both immunostaining and gene expression assays of pluripotency markers (Online Figure III). We differentiated the isogenic iPSC lines to iPSC-CMs and observed that the DCM-KO iPSC-CMs had undetectable (<10%) mRNA expression of the mutant *TNNT2* allele when compared to the parental line, consistent with the activation of the nonsense-mediated mRNA decay mechanism following the NHEJ repair process (Online Figure IV).[21] In addition, we observed that the loss of the mutant allele ameliorated the DCM phenotype in vitro, including sarcomere disarray (Figure 3b-c) and Ca^{+2} cycling parameters (Figure 3d-e). Taken together, these data suggest that the *TNNT2 p.R173W* is a dominant negative mutation, and allelic-specific KO could ameliorate the DCM phenotype in vitro.

Modeling Holt-Oram syndrome in vitro.

Cardiac development is a critical and complex embryologic process requiring the integration of cell commitment, growth, looping, septation, and chamber specification.[22] Multiple transcription factors, including *NKX2.5*, *GATA4*, and *TBX5* play important roles in cardiac development, and genetic studies have implicated dominant mutations in these genes in human CHD. *TBX5* is a T-box-containing transcription factor, which like other T-box family members, has been implicated in vertebrate tissue patterning and differentiation.[23-25] *TBX5* represents one of the few genes which, when mutated, is known to cause CHD.[23] *TBX5* haploinsufficiency is associated with Holt-Oram syndrome (HOS), a congenital disorder characterized by structural cardiac and limb abnormalities.[26] *Tbx5* heterozygous null ($Tbx5^{-/+}$) mice recapitulated the CHD seen in HOS patients, whereas homozygous null mice ($Tbx5^{-/-}$) are growth arrested at E9.0 and die in utero by E10.5 due to severe heart defects.[26] Although the expression of many genes such as *NPPA, GJA5, IRX4, MYL2, GATA4, NKX2.5,* and *HEY2* was reduced in TBX5-null hearts,[26] little is known about their downstream targets and hence the molecular basis of HOS is poorly understood.

As a proof-of-principle experiment for creating CHD models, we generated a human cell-based HOS in vitro model by utilizing TALEN-mediated NHEJ to knockout the TBX5 gene in iPSCs. In humans, TBX5 is highly regulated through alternative splicing and several transcript variants encoding different isoforms have been described for TBX5. Based on RNA-seq data of iPSC-CMs, the transcript variant 4 (NM_181486) is the predominant TBX5 isoform that is expressed in iPSC-CMs. Of note, the presence of this transcript was also reported in the initial identification of TBX5 as the HOS gene.[27] The isoforms 1 (NM_000192) and 3 (NM_080717) were also detected in iPSC-CMs, albeit at very low levels (Online Figure V). Hence, we designed a TALEN pair and targeted the starting codon at exon 1 of the major isoform 4 and isoform 1 (Figure 4a). We identified an iPSC clone carrying a homozygous deletion, which resulted in frameshift mutations and an early termination of the TBX5 gene (hereafter referred to as TBX5-KO) (Figure 4b). The isogenic TBX5-KO iPSCs retained their pluripotency as assessed by immunostaining and gene expression assays of pluripotency markers (Online Figure VI). In order to check the specificity, we assessed potential off-target cutting sites in the edited clones using *in silico* prediction algorithms and did not detect any mutations in the 25 most likely off-target sites, suggesting a high specificity of the TBX5 TALEN pair (Online Table IV). We then differentiated the isogenic iPSC clones into iPSC-CMs and

confirmed that the TBX5 (isoforms 1 and 4) was not expressed at the protein level (Figure 4c). The directed differentiation protocol yielded cultures enriched (70%–85%) in cTnT (+) beating iPSC-CMs in both WT and TBX5-KO iPSC lines at day 15 post-differentiation (Online Figure VII) that displayed a typical sarcomeric morphology (Figure 4b). As HOS is associated with electrophysiological abnormalities,[26,28] we next characterized the action potential (APs) of the isogenic iPSC-CMs. Both TBX5-KO and WT iPSC-CMs displayed typical AP morphologies, including ventricular-, atrial-, and nodal-like subtypes (Figure 4d and Online Table V). However, we observed that 35% of TBX5-KO iPSC-CMs exhibited marked proarrhythmic activity characterized by the development of depolarizing humps during phase 2 and 3 of the action AP that resemble early after-depolarizations (EADs) when compared to the parental iPSC-CMs (Figure 4e).

Identification of novel TBX5 target genes.

To identify downstream targets and TBX5-dependent molecular networks, we next performed chromatin immunoprecipitation coupled to massively parallel sequencing (ChIP-seq) together with RNA-seq analyses. RNA-seq analysis of isogenic iPSC-CMs revealed profound changes in global gene expression. We identified 349 up-regulated and 645 down-regulated gene transcripts in TBX5-KO when compared to the parental WT iPSC-CMs at a false discovery rate (FDR) of 5%. Analysis of a representative subset of these genes by qRT-PCR in independent experiments validated our findings (Online Figure VIII). Of note, the most significant down-regulated gene was *NPPA* (Figure 5a and Figure 5b), a known direct target of TBX5.[29] As available antibodies for TBX5 are not suitable for genome-wide ChIP-seq, we used a lentivirus to express a FLAG-tagged TBX5 in WT iPSC-CMs. We performed FLAG-mediated ChIP-seq to define the binding sites of TBX5 genome-wide. We identified 4,518 TBX5-bound peaks that were significantly enriched in the TBX5-FLAG sample compared with the control sample (FDR < 0.01). To validate the ChIP-seq peaks, we next performed *de novo* motif analysis to investigate the predominant motifs enriched in TBX5 binding sites. As expected, the identified peaks were highly enriched for the previously experimentally discovered motif of TBX5 (Online Figure XI).[29]

Next, to define the direct TBX5 gene regulatory networks, we correlated *TBX5* binding and *TBX5*-mediated gene regulation by combining the gene set containing TBX5 peaks with the genes differentially expressed between the TBX5KO and WT iPSC-CMs. We annotated the TBX5-bound regions to the nearest transcription-starting site (TSS) and identified 341 candidate *TBX5* direct target genes (118 up- and 223 down-regulated genes) (Figure 5c). To further refine the identification of TBX5 target genes, we analyzed the 223 downregulated gene set andrevealed important genes associated with cardiac myocyte function, such as cardiac myosin-binding protein C (*MYBPC3*), titin (*TTN*), calsequestrin (*CASQ2*), natriuretic peptide type A (*NPPA*), connexin 43 (*GJA5),* and sodium voltage-gated channel alpha subunit 5 (*SCN5A*). Remarkably, we found that 40% of the TBX5 candidate target genes were enriched in unexpected pathways ostensibly unrelated to processes associated with heart function. These pathways included extracellular matrix (ECM)-receptor interaction, focal adhesion, and protein digestion and absorption (Figure 5d). We found that the *TBX5* was bound to promoter regions of key components of the embryonic provisional matrix, including perlecan (HSPG2),[30] fibronectin (FN1),[31,32] fibulin-1 (*FBLN1*),[33] collagen XIV (*COL14A3*),[34] versican (*VCAN*),[35-37] and versican-degrading protease *ADMTS9*.[34] These ECM components play essential roles in cardiac development and are indispensable for normal heart development by regulating heart tube segmentation, chamber specification, endocardial cushion formation, interventricular septal formation, and cardiac myocyte differentiation.[38] Taken together, these data suggest that genes encoding embryonic ECM components are direct TBX5 targets and represent potential novel candidate genes associated with HOS and CHD.

DISCUSSION

In the past decade, advances in cardiovascular genetics have uncovered a plethora of genes associated with inherited cardiomyopathies. Delineating the role of cardiomyopathy-associated genes and variants could provide a better understanding to the underlying pathogenic mechanisms, and provide potential targets for therapeutic interventions. The advent of new technologies, including iPSC and genome editing with designer nucleases, has provided an unprecedented opportunity for disease modeling in vitro. Since the development of a highly active TALEN architecture[39] and simplified engineering platforms[12], TALEN-mediated genome editing has been demonstrated in diverse cell types, including pluripotent stem cells.[12,40-42] The relatively unconstrained target site requirements[43] and the high degree of specificity of TALENs, provide a valuable tool for genome editing.

In principle, a TALEN pair can be targeted to any site in a genome, allowing more freedom and flexibility in target site selection with minimal off-target mutagenesis when compared to newer technologies such as CRISPR/Cas9.[44-46] In this study, we designed, constructed, and validated TALEN vectors as an effective tool for gene KO in human iPSCs. The cTAL panel consists of 88 TALEN pairs that are designed to knockout genes that are associated with cardiomyopathies and CHD. Every TALEN pair was individually validated in human iPSCs and found to be active at the targeted locus. Furthermore, we have established that the target sites needs to be carefully chosen as TALEN pairs that target either the start codon (ATG) or regions immediately after are more effective in disrupting the open reading frame of the targeted gene. In contrast, indels at the 5-end UTR are inefficient in modifying the open reading frame. It should also be noted that even though the start codon is deleted, there might be a downstream translation starting sites that could function alternatively.

An important issue in cardiovascular genetics is determining whether putative mutations are causative of the disease, and establishing causality for putative disease causing variants is becoming increasingly clinically relevant. As a proof-of-concept, we showed that the DCM phenotype in iPSC-CMs was ameliorated by selectively disrupting the starting codon of the DCM-causing *TNNT2* allele in a patient-specific iPSCs. In addition, using a similar strategy, we created a CHD model of HOS in vitro and identified a number of novel genes that are associated with TBX5 haploinsufficiency, providing an entry point to understanding the complex phenotypes caused by TBX5 haploinsufficiency and the pathogenesis of HOS. Taken together, these results demonstrated that TALEN-mediated gene KO strategies in iPSCs could be used to interrogate disease-causing mutations in a wide range of diseases and cell types as well as to model complex diseases in vitro.

In summary, combining iPSC and genome editing technologies holds great promise for advancing fundamental knowledge of the pathogenesis of inherited cardiomyopathies and CHD. The methods, strategies, and constructs developed in this study provide a validated, readily available resource for cardiovascular research that simplifies the custom generation of iPSC knock-out cell lines, and will therefore have a broad applicability for the generation of iPSC-based disease models and functional studies.

SOURCES OF FUNDING

This study was funded by the NIH 4R00 HL104002-03, AHA 15BGIA22730027, and Stanford CVI Seed Grant (I.K.); Prince Mahidol Award Foundation, Thailand (V.T.); the German Research Foundation (T.S.); U54 DH068158-01 (Pilot Project) (VS), California Institute of Regenerative Medicine (CIRM) RB3-2209 (VS); and AHA 13EIA14420025, NIH R01 HL126527, NIH R01 HL130020, NIH R01 HL123968, NIH R01 HL132875, NIH R01 HL133272, and NIH R24 HL117756 (JCW).

DISCLOSURE

None.

REFERENCES

1. Fahed AC, Gelb BD, Seidman JG, Seidman CE. Genetics of congenital heart disease: The glass half empty. *Circ Res*. 2013;112:707-720
2. Watkins H, Ashrafian H, Redwood C. Inherited cardiomyopathies. *N Engl J Med*. 2011;364:1643-1656
3. Kim H, Kim JS. A guide to genome engineering with programmable nucleases. *Nat Rev Genet*. 2014;15:321-334
4. Mittal V. Improving the efficiency of rna interference in mammals. *Nat Rev Genet*. 2004;5:355-365
5. Karakikes I, Ameen M, Termglinchan V, Wu JC. Human induced pluripotent stem cell-derived cardiac myocytes: Insights into molecular, cellular, and functional phenotypes. *Circ Res*. 2015;117:80-88
6. Adams WJ, Zhang Y, Cloutier J, Kuchimanchi P, Newton G, Sehrawat S, Aird WC, Mayadas TN, Luscinskas FW, Garcia-Cardena G. Functional vascular endothelium derived from human induced pluripotent stem cells. *Stem Cell Reports*. 2013;1:105-113
7. Lee TH, Song SH, Kim KL, Yi JY, Shin GH, Kim JY, Kim J, Han YM, Lee SH, Lee SH, Shim SH, Suh W. Functional recapitulation of smooth muscle cells via induced pluripotent stem cells from human aortic smooth muscle cells. *Circ Res*. 2010;106:120-128
8. Hockemeyer D, Jaenisch R. Induced pluripotent stem cells meet genome editing. *Cell Stem Cell*. 2016;18:573-586
9. Doyle EL, Booher NJ, Standage DS, Voytas DF, Brendel VP, Vandyk JK, Bogdanove AJ. Tal effector-nucleotide targeter (tale-nt) 2.0: Tools for tal effector design and target prediction. *Nucleic Acids Res*. 2012;40:W117-122
10. Moscou MJ, Bogdanove AJ. A simple cipher governs DNA recognition by tal effectors. *Science*. 2009;326:1501
11. Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, Lahaye T, Nickstadt A, Bonas U. Breaking the code of DNA binding specificity of tal-type iii effectors. *Science*. 2009;326:1509-1512
12. Ding Q, Lee YK, Schaefer EA, Peters DT, Veres A, Kim K, Kuperwasser N, Motola DL, Meissner TB, Hendriks WT, Trevisan M, Gupta RM, Moisan A, Banks E, Friesen M, Schinzel RT, Xia F, Tang A, Xia Y, Figueroa E, Wann A, Ahfeldt T, Daheron L, Zhang F, Rubin LL, Peng LF, Chung RT, Musunuru K, Cowan CA. A talen genome-editing system for generating human stem cell-based disease models. *Cell Stem Cell*. 2013;12:238-251
13. Karakikes I, Senyei GD, Hansen J, Kong CW, Azeloglu EU, Stillitano F, Lieu DK, Wang J, Ren L, Hulot JS, Iyengar R, Li RA, Hajjar RJ. Small molecule-mediated directed differentiation of human embryonic stem cells toward ventricular cardiac myocytes. *Stem Cells Transl Med*. 2014;3:18-31
14. Hendel A, Kildebeck EJ, Fine EJ, Clark JT, Punjya N, Sebastiano V, Bao G, Porteus MH. Quantifying genome-editing outcomes at endogenous loci with smrt sequencing. *Cell Rep*. 2014;7:293-305
15. Fine EJ, Cradick TJ, Zhao CL, Lin Y, Bao G. An online bioinformatics tool predicts zinc finger and tale nuclease off-target cleavage. *Nucleic Acids Res*. 2014;42:e42
16. Maze I, Shen L, Zhang B, Garcia BA, Shao N, Mitchell A, Sun H, Akbarian S, Allis CD, Nestler EJ. Analytical tools and current challenges in the modern era of neuroepigenomics. *Nat Neurosci*. 2014;17:1476-1490
17. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and b cell identities. *Mol Cell*. 2010;38:576-589
18. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. Integrative genomics viewer. *Nat Biotechnol*. 2011;29:24-26

19. Hershberger RE, Pinto JR, Parks SB, Kushner JD, Li D, Ludwigsen S, Cowan J, Morales A, Parvatiyar MS, Potter JD. Clinical and functional characterization of tnnt2 mutations identified in patients with dilated cardiomyopathy. *Circ Cardiovasc Genet*. 2009;2:306-313
20. Sun N, Yazawa M, Liu J, Han L, Sanchez-Freire V, Abilez OJ, Navarrete EG, Hu S, Wang L, Lee A, Pavlovic A, Lin S, Chen R, Hajjar RJ, Snyder MP, Dolmetsch RE, Butte MJ, Ashley EA, Longaker MT, Robbins RC, Wu JC. Patient-specific induced pluripotent stem cells as a model for familial dilated cardiomyopathy. *Sci Transl Med*. 2012;4:130ra147
21. Wen J, Brogna S. Nonsense-mediated mrna decay. *Biochem Soc Trans*. 2008;36:514-516
22. Epstein JA. Franklin h. Epstein lecture. Cardiac development and implications for heart disease. *N Engl J Med*. 2010;363:1638-1647
23. Papaioannou VE, Silver LM. The t-box gene family. *Bioessays*. 1998;20:9-19
24. Bruneau BG, Nemer G, Schmitt JP, Charron F, Robitaille L, Caron S, Conner DA, Gessler M, Nemer M, Seidman CE, Seidman JG. A murine model of holt-oram syndrome defines roles of the t-box transcription factor tbx5 in cardiogenesis and disease. *Cell*. 2001;106:709-721
25. Rodriguez-Esteban C, Tsukui T, Yonei S, Magallon J, Tamura K, Izpisua Belmonte JC. The t-box genes tbx4 and tbx5 regulate limb outgrowth and identity. *Nature*. 1999;398:814-818
26. Basson CT, Cowley GS, Solomon SD, Weissman B, Poznanski AK, Traill TA, Seidman JG, Seidman CE. The clinical and genetic spectrum of the holt-oram syndrome (heart-hand syndrome). *N Engl J Med*. 1994;330:885-891
27. Basson CT, Bachinsky DR, Lin RC, Levi T, Elkins JA, Soults J, Grayzel D, Kroumpouzou E, Traill TA, Leblanc-Straceski J, Renault B, Kucherlapati R, Seidman JG, Seidman CE. Mutations in human tbx5 [corrected] cause limb and cardiac malformation in holt-oram syndrome. *Nat Genet*. 1997;15:30-35
28. Nadadur RD, Broman MT, Boukens B, Mazurek SR, Yang X, van den Boogaard M, Bekeny J, Gadek M, Ward T, Zhang M, Qiao Y, Martin JF, Seidman CE, Seidman J, Christoffels V, Efimov IR, McNally EM, Weber CR, Moskowitz IP. Pitx2 modulates a tbx5-dependent gene regulatory network to maintain atrial rhythm. *Sci Transl Med*. 2016;8:354ra115
29. Mori AD, Zhu Y, Vahora I, Nieman B, Koshiba-Takeuchi K, Davidson L, Pizard A, Seidman JG, Seidman CE, Chen XJ, Henkelman RM, Bruneau BG. Tbx5-dependent rheostatic control of cardiac gene expression and morphogenesis. *Dev Biol*. 2006;297:566-586
30. Costell M, Carmona R, Gustafsson E, Gonzalez-Iriarte M, Fassler R, Munoz-Chapuli R. Hyperplastic conotruncal endocardial cushions and transposition of great arteries in perlecan-null mice. *Circ Res*. 2002;91:158-164
31. George EL, Georges-Labouesse EN, Patel-King RS, Rayburn H, Hynes RO. Defects in mesoderm, neural tube and vascular development in mouse embryos lacking fibronectin. *Development*. 1993;119:1079-1091
32. Georges-Labouesse EN, George EL, Rayburn H, Hynes RO. Mesodermal development in mouse embryos mutant for fibronectin. *Dev Dyn*. 1996;207:145-156
33. Cooley MA, Kern CB, Fresco VM, Wessels A, Thompson RP, McQuinn TC, Twal WO, Mjaatvedt CH, Drake CJ, Argraves WS. Fibulin-1 is required for morphogenesis of neural crest-derived structures. *Dev Biol*. 2008;319:336-345
34. Tao G, Levay AK, Peacock JD, Huk DJ, Both SN, Purcell NH, Pinto JR, Galantowicz ML, Koch M, Lucchesi PA, Birk DE, Lincoln J. Collagen xiv is important for growth and structural integrity of the myocardium. *J Mol Cell Cardiol*. 2012;53:626-638
35. Mjaatvedt CH, Yamamura H, Capehart AA, Turner D, Markwald RR. The cspg2 gene, disrupted in the hdf mutant, is required for right cardiac chamber and endocardial cushion formation. *Dev Biol*. 1998;202:56-66
36. Yamamura H, Zhang M, Markwald RR, Mjaatvedt CH. A heart segmental defect in the anterior-posterior axis of a transgenic mutant mouse. *Dev Biol*. 1997;186:58-72
37. Henderson DJ, Copp AJ. Versican expression is associated with chamber specification, septation, and valvulogenesis in the developing mouse heart. *Circ Res*. 1998;83:523-532

38. Rienks M, Papageorgiou AP, Frangogiannis NG, Heymans S. Myocardial extracellular matrix: An ever-changing and diverse entity. *Circ Res*. 2014;114:872-888
39. Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF, Meng X, Paschon DE, Leung E, Hinkley SJ, Dulay GP, Hua KL, Ankoudinova I, Cost GJ, Urnov FD, Zhang HS, Holmes MC, Zhang L, Gregory PD, Rebar EJ. A tale nuclease architecture for efficient genome editing. *Nat Biotechnol*. 2011;29:143-148
40. Hockemeyer D, Wang H, Kiani S, Lai CS, Gao Q, Cassady JP, Cost GJ, Zhang L, Santiago Y, Miller JC, Zeitler B, Cherone JM, Meng X, Hinkley SJ, Rebar EJ, Gregory PD, Urnov FD, Jaenisch R. Genetic engineering of human pluripotent cells using tale nucleases. *Nat Biotechnol*. 2011;29:731-734
41. Karakikes I, Stillitano F, Nonnenmacher M, Tzimas C, Sanoudou D, Termglinchan V, Kong CW, Rushing S, Hansen J, Ceholski D, Kolokathis F, Kremastinos D, Katoulis A, Ren L, Cohen N, Gho JM, Tsiapras D, Vink A, Wu JC, Asselbergs FW, Li RA, Hulot JS, Kranias EG, Hajjar RJ. Correction of human phospholamban r14del mutation associated with cardiomyopathy using targeted nucleases and combination therapy. *Nat Commun*. 2015;6:6955
42. Park CY, Kim J, Kweon J, Son JS, Lee JS, Yoo JE, Cho SR, Kim JH, Kim JS, Kim DW. Targeted inversion and reversion of the blood coagulation factor 8 gene in human ips cells using talens. *Proc Natl Acad Sci U S A*. 2014;111:9253-9258
43. Reyon D, Tsai SQ, Khayter C, Foden JA, Sander JD, Joung JK. Flash assembly of talens for high-throughput genome editing. *Nat Biotechnol*. 2012;30:460-465
44. Kuscu C, Arslan S, Singh R, Thorpe J, Adli M. Genome-wide analysis reveals characteristics of off-target sites bound by the cas9 endonuclease. *Nat Biotechnol*. 2014;32:677-683
45. Guilinger JP, Thompson DB, Liu DR. Fusion of catalytically inactive cas9 to foki nuclease improves the specificity of genome modification. *Nat Biotechnol*. 2014;32:577-582
46. Tsai SQ, Wyvekens N, Khayter C, Foden JA, Thapar V, Reyon D, Goodwin MJ, Aryee MJ, Joung JK. Dimeric crispr rna-guided foki nucleases for highly specific genome editing. *Nat Biotechnol*. 2014;32:569-576

FIGURE LEGENDS

Figure 1. The cTAL-KO panel. **a)** Genes associated with cardiomyopathies and congenital heart diseases included in the panel. **b)** Schematic representation of the gene KO strategy. **c)** Frequency distribution of the TALEN-mediated mutagenesis in human iPSCs as assessed by single-molecule real-time (SMRT) technology. The DNA fragments surrounding the TALEN target site was amplified and sequenced by PacBio RS as described in the "Materials and Methods" section. The mutation frequency of each TALEN pair was calculated as follows: mutation frequency (%) = number of reads containing a different length of deletion mutations/total number of reads harboring deletion mutation in the target locus × 100. HCM, hypertrophic cardiomyopathy; DCM, dilated cardiomyopathy; LVNC, left ventricular non-compaction; ARVD, arrhythmogenic right ventricular dysplasia; RC, restrictive cardiomyopathy.

Figure 2. Generation of *TNNT2* knockout iPSC clones. **a**) Schematic representation of TNNT2 gene structure. TALENs were designed to target the translation initiation site (ATG) at exon 2 of TNNT2 gene. Boxes indicate the TALEN binding sites. Deletions in the two alleles of each clone are indicated. **b**) Expression of cardiac troponin-T protein in isogenic wild-type (WT), heterozygous ($TNNT2^{+/-}$), and homozygous ($TNNT2^{-/-}$) knockout iPSC-CMs. Representative blots of the protein expression and densitometric analysis of TNNT2 protein expression levels normalized to α-sarcomeric actinin (ACTN2) in isogenic iPSC-CMs as indicated. Data represent mean ± SEM of three independent differentiation experiments, * $P < 0.05$. **c**) mRNA expression of the WT allele in the $TNNT2^{+/-}$ and WT iPSC lines. A qPCR probe was designed to distinguish between the non-edited (WT) and the TALEN-mutated mRNA of the $TNNT2^{+/-}$ iPSC-CMs. Gene expression levels were normalized to cardiac specific gene *ACTN2*. Data represent mean ± SEM of three independent differentiation experiments, *$P < 0.05$. **d**) Representative immunofluorescence images of iPSC-CMs stained for the cardiac myocyte-specific markers cardiac troponin-T (TNNT2, red) and α-sarcomeric actinin (ACTN2, green). DNA was counterstained with DAPI (blue). Scale bar = 20 μm. All the assays were performed at 30 days post-differentiation with one isogenic pair.

Figure 3. TNNT2 R173W is a dominant, causal DCM mutation. **a**) Generation of allelic-specific *TNNT2* knockout iPSC clones. TALENs were designed to target the translation initiation site (ATG) at exon 2 of TNNT2 gene. Boxes indicate the TALEN binding sites. The nucleotide in red indicates the missense mutation for R173W. A deletion in the TNNT2 allele (R173W) allele is indicated. **b**) Representative immunofluorescence images of iPSC-CMs stained for the cardiac myocyte-specific markers cardiac troponin-T (TNNT2, red) and α-sarcomeric actinin (ACTN2, green). DNA was counterstained with DAPI (blue). Scale bar = 20 μm. **c**) Quantification of disorganized sarcomeric staining pattern in WT, isogenic DCM, and DCM-KO iPSC-CMs. Data represent mean ± SEM (n=150 iPSC-CMs per iPSC line), *$P < 0.05$. **d**) Representative Ca^{2+} transients of iPSC-CMs as indicated. **e**) Quantification of calcium handling parameters in WT, isogenic DCM, and DCM-KO iPSC-CMs. Data represent mean ± SEM (n=30 iPSC-CMs per line), *$P < 0.05$. All the assays were performed at 30 days post-differentiation with one isogenic pair.

Figure 4. Modeling HOS in human iPSCs. **a**) Schematic representation of TBX5 gene structure. TALENs were designed to target the translation initiation site (ATG) at exon 1 of TBX5 gene. Boxes indicate the TALEN binding sites. The TALEN-mediated deletions in the two alleles of the iPSC clone are shown. **b**) Representative immunofluorescence images of iPSC-CMs stained for the cardiac myocyte-specific marker cardiac troponin-T (cTnT). DNA was counterstained with DAPI (blue). **c**) Assessment of TBX5 protein expression in isogenic iPSC-CMs by western blot analysis; cTnT was used as a loading control. **d)** AP characterization in isogenic iPSC-CMs. **d)** TBX5-KO iPSC-CMs exhibit a proarrhythmia phenotype manifested as early after-depolarizations (EADs) during phase 2 and 3 of the AP waveform.

<u>Figure 5.</u> TBX5 regulates extracellular matrix (ECM) genes in iPSC-CMs. **a**) Top 20 differentially expressed genes between isogenic TBX5-KO and WT iPSC-CMs as assessed by RNA-seq. Blue bars represent up-regulated genes; red bars represent down-regulated genes. **b)** Representative browser tracks of *NPPA* gene expression in isogenic WT and TBX5-KO iPSC-CMs, and ChIP-seq footprint shows that TBX5 binds to the TSS of the *NPPA* gene. **c**) Intersection with ChiP-seq and transcriptional profiling identified 341 candidate TBX5 direct target genes. Blue circles represent up-regulated genes and red circles represent down-regulated genes (TBX5KO/WT); green circle represent TBX5-bound regions. **d**) A significant enrichment of extracellular matrix (ECM) components were observed in *TBX5* direct target genes. The extracellular matrix (ECM)-receptor interaction and focal adhesion were the two most significant gene-sets over-represented among the 223 down-regulated (TBX5KO/WT) TBX5-bound genes.

NOVELTY AND SIGNIFICANCE

What Is Known?

- Advances in cardiovascular genetics have uncovered many genes associated with inherited cardiomyopathies.
- The use of human induced pluripotent stem cell-derived cardiac myocytes (iPSC-CMs) provides an unprecedented opportunity for the generation of human cell-based disease models to study genetic cardiomyopathies.

What New Information Does This Article Contribute?

- Transcription activator–like effector nucleases (TALENs) facilitate gene knockout (KO) with high efficiency, precision and accuracy.
- Successful creation of human-based KO cell models in vitro by combining genome editing and iPSC-CM technologies.
- TALEN-mediated allele-specific KO ameliorate dilated cardiomyopathy (DCM)-associated phenotypes in iPSC-CMs in vitro.
- Modeling Holt-Oram syndrome (HOS) in iPSC-CMs in vitro uncovered novel genes and pathways regulated by *TBX5*.

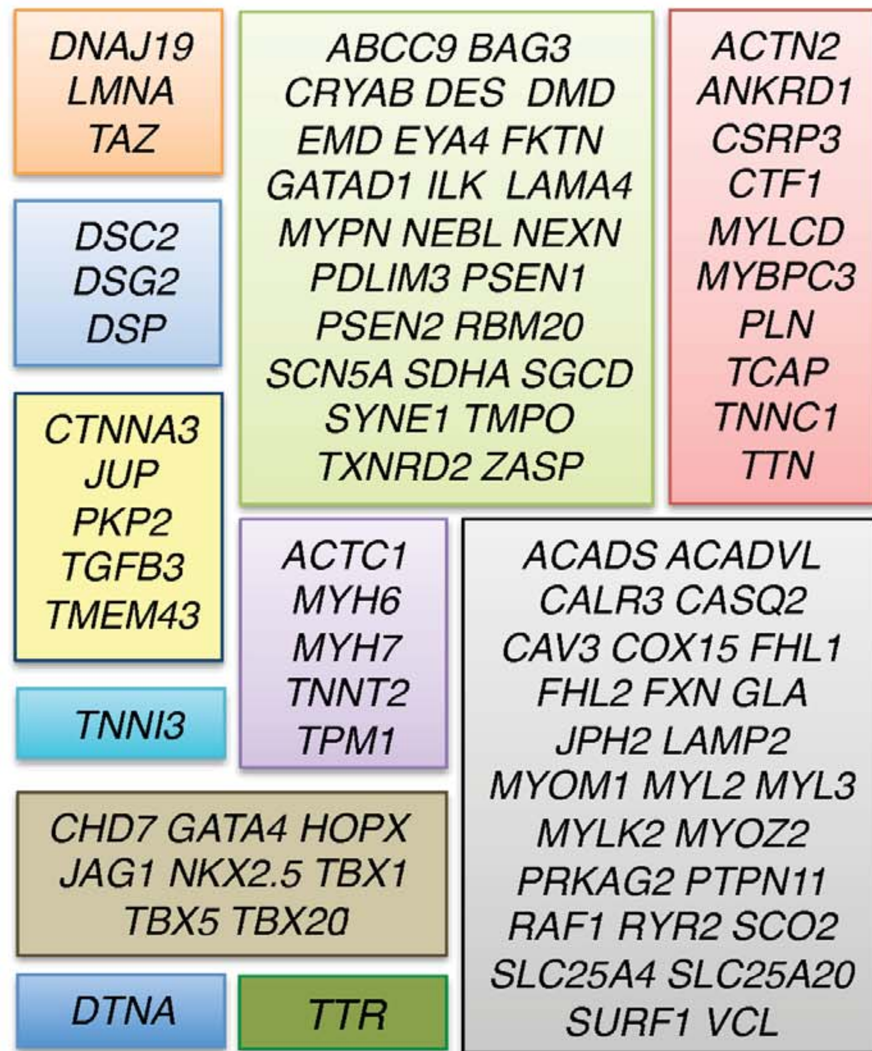
The advent of human iPSC technology and an increasingly refined capacity to differentiate iPSCs into disease-relevant cell types, such as iPSC-CMs, provide an unprecedented opportunity for the generation of human cell-based disease models to study genetic cardiomyopathies. Genome editing can be used to change the DNA in iPSCs to aid the understanding of the biology of cardiomyopathy-associated genes and how they work. We can now make changes (or 'edits') to the DNA in specific location in the genome using an 'engineered nuclease', an enzyme that can be tailored to cut the genome in a specific place. Here we harnessed this technology to generate iPSC-based KO models of genetic cardiomyopathies to study the underlying pathogenic phenotypes and mechanisms, as well as to genetically correct the disease in vitro. Implementation of this unique and clinically relevant model system presents a significant advantage in cardiovascular research as it can circumvent complications in translating data from models across different species and biological characteristics. Ultimately, a better understanding of molecular mechanism(s) of genetic cardiomyopathies could provide opportunities for diagnosis and prognosis as well as enable the development of personalized therapeutic interventions.

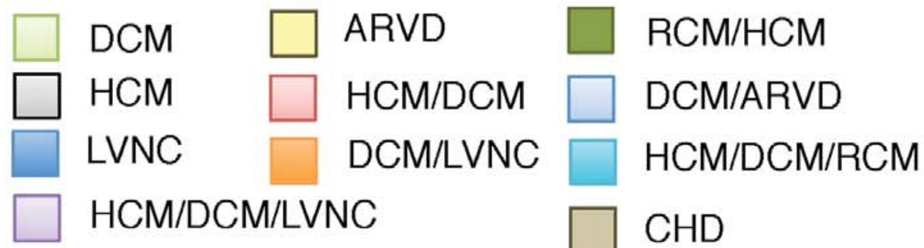
Table 1. Efficiency of TALEN-Mediated Gene KO in iPSCs

Targeted Gene	NHEJ (%)	Clones Screened	Mutants Clones	Efficiency (%)
TNNT2	13.1	22	11	50.0
LMNA	12.5	12	8	66.7
MYH7	50.2	24	24	100
ANKRD1	6.7	24	11	45.8
TBX5	48.5	32	26	81.3
NKX2.5	9.4	26	20	76.9

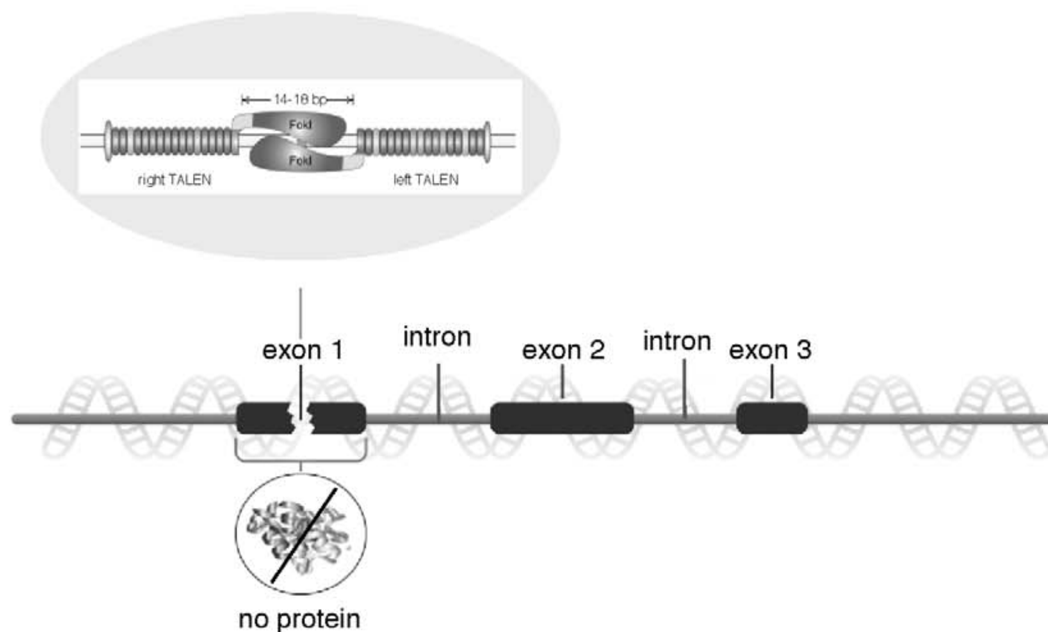
Figure 1

A





B



C

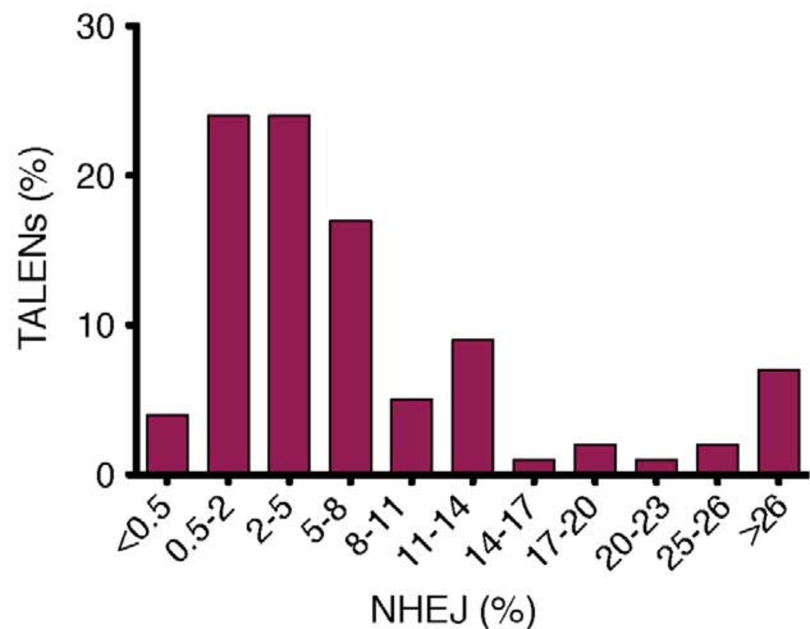
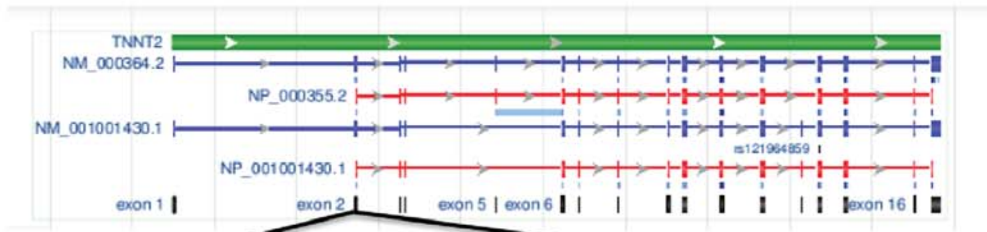


Figure 2

A

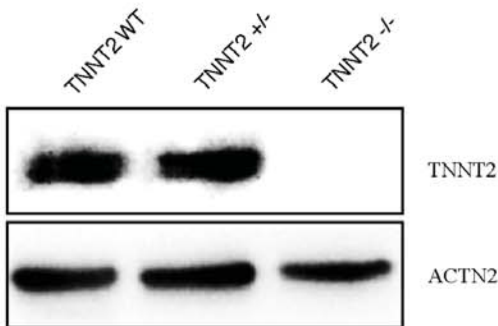


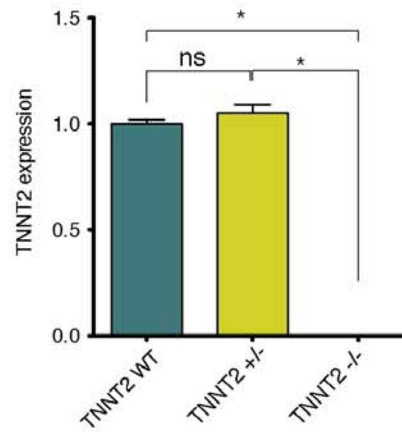
WT T T T T G G A G G G A G A G C A G A G A C C **A T G** T C T G A C A T A G A A G A G G T G G T G G A A
A A A A C C T C C C T C T C G T C T C T G G **T A C** A G A C T G T A T C T T C T C C A C C A C C T T

C-2
T T T T G G A G G G A G A G C A G A G A C C _ _ _ _ _ T G A C A T A G A A G A G G T G G T G G A A 5bp del
T T T T G G A G G G A G A G C A G A G A C C **A T** _ _ C T G A C A T A G A A G A G G T G G T G G A A 2bp del

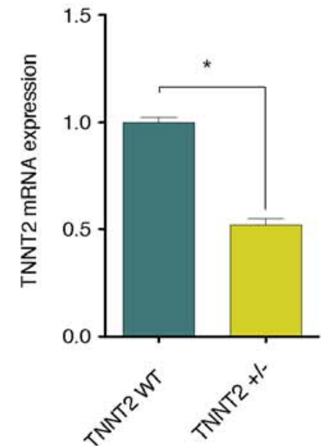
C-3
T T T T G G _ A G A A G A G G T G G T G G A A 27bp del
T T T T G G A G G G A G A G C A G A G A C C **A T G** T C T G A C A T A G A A G A G G T G G T G G A A wt

B





C



D

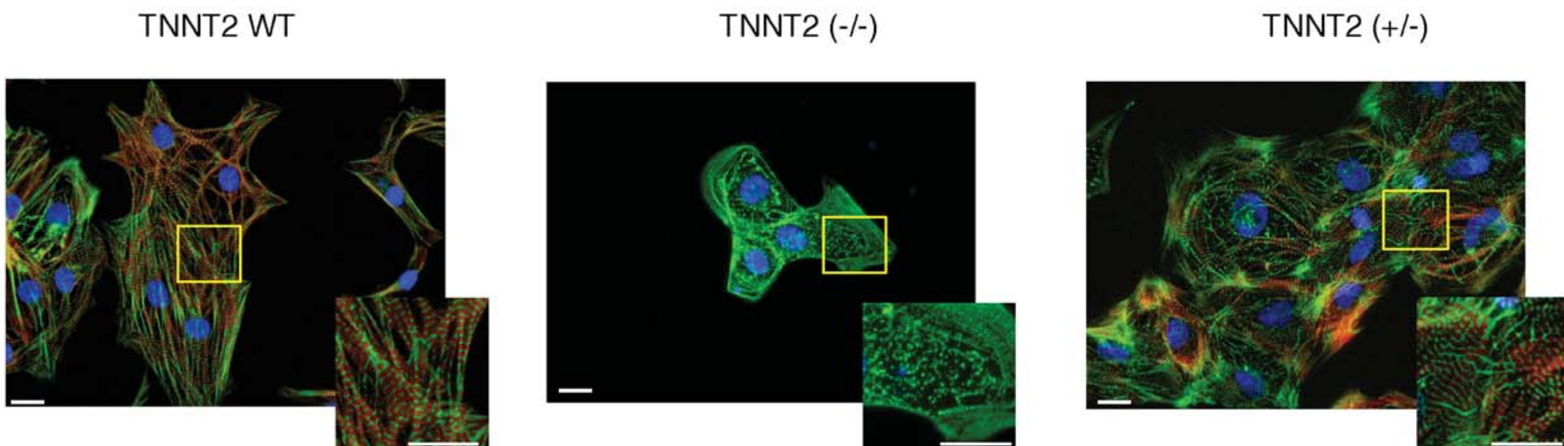


Figure 3

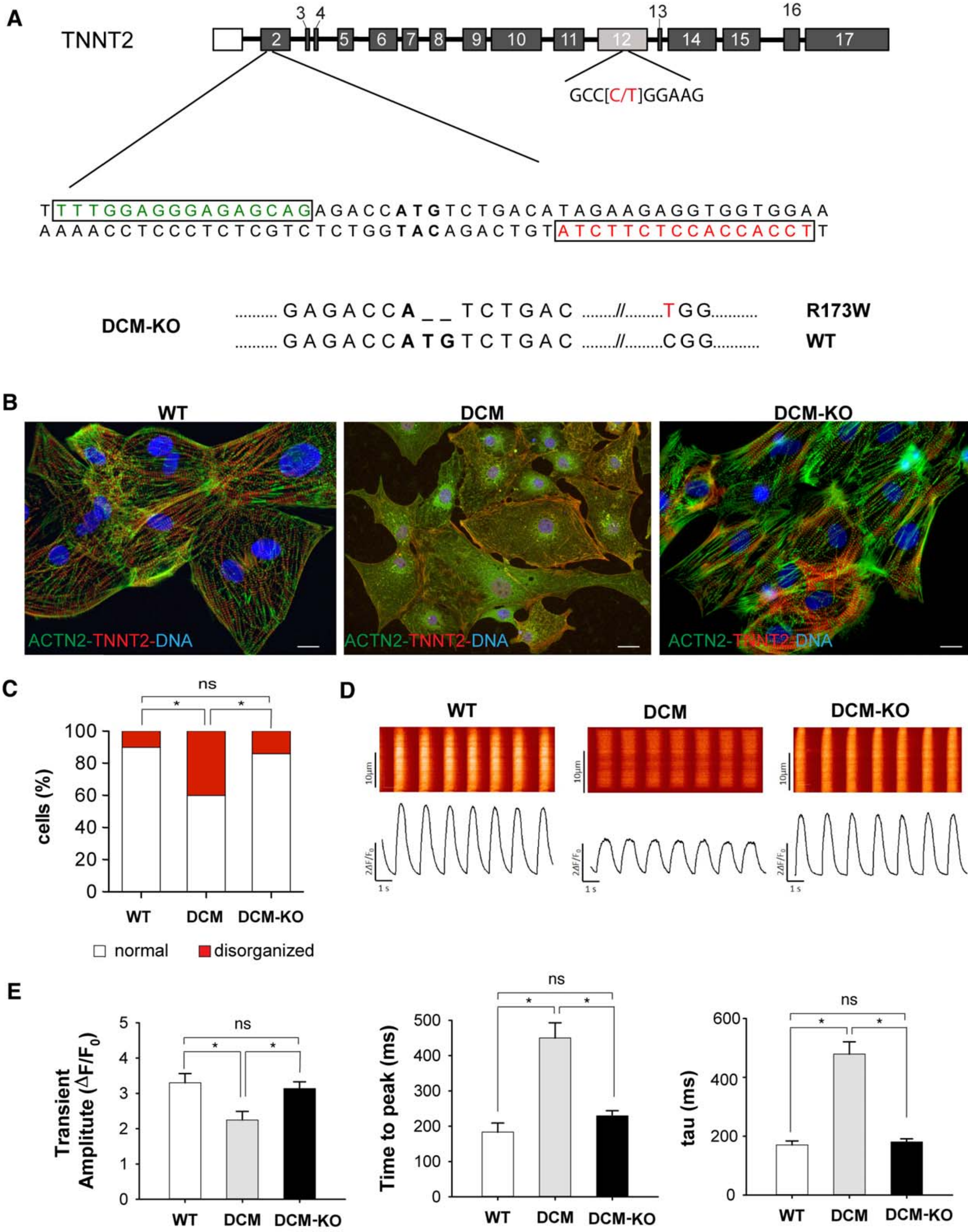
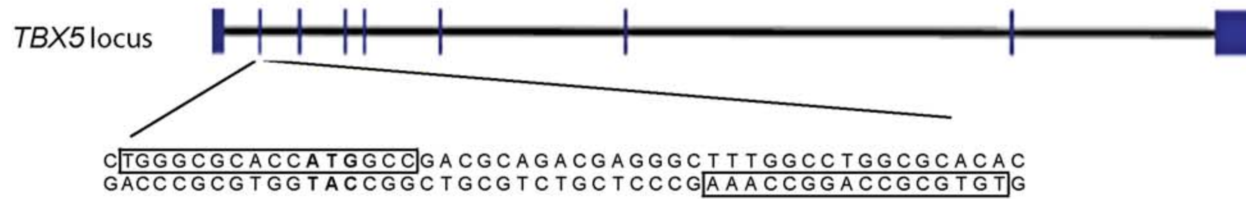


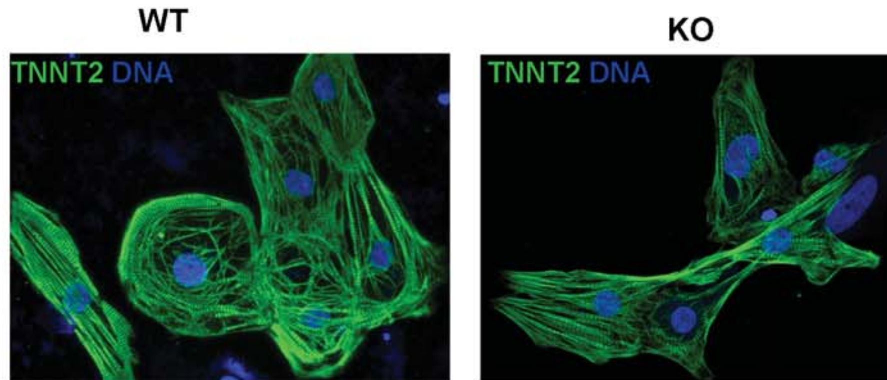
Figure 4

A

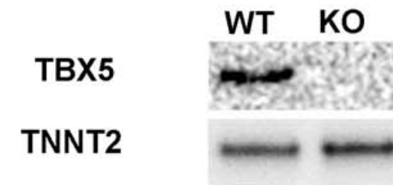


Alelle A CTGGGCGCACCATGGCCGACGCAGACG__________CCTGGCGCACAC 10bp del
Alelle B CTGGGCGCACCATGGCC______________________TGGCGCACAC 22bp del

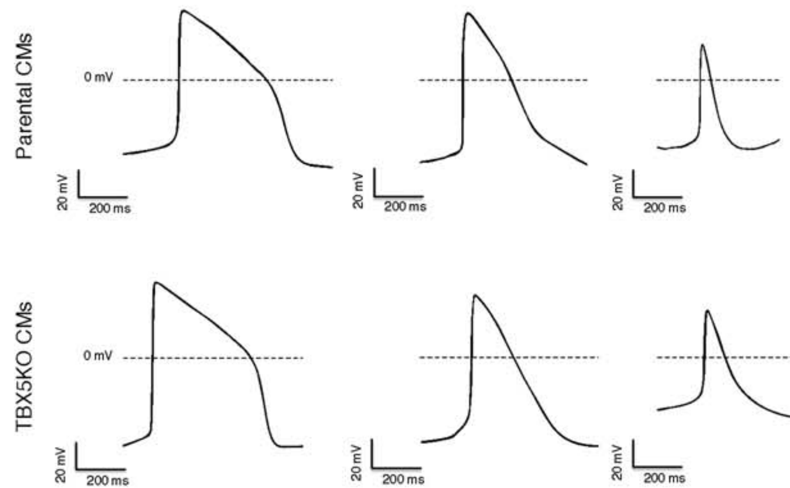
B



C



D



E

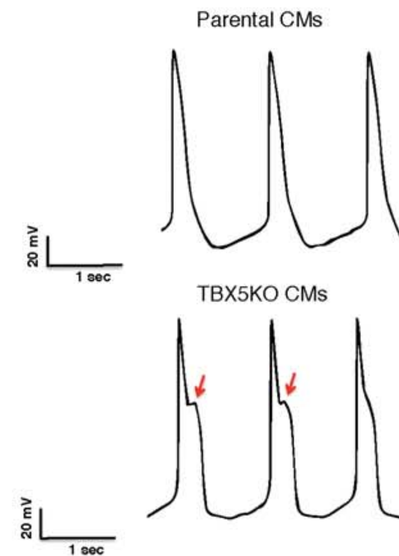
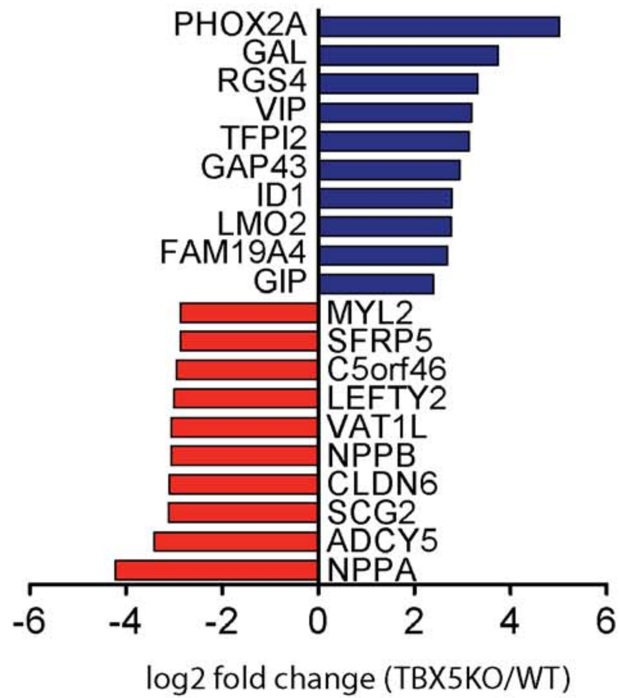
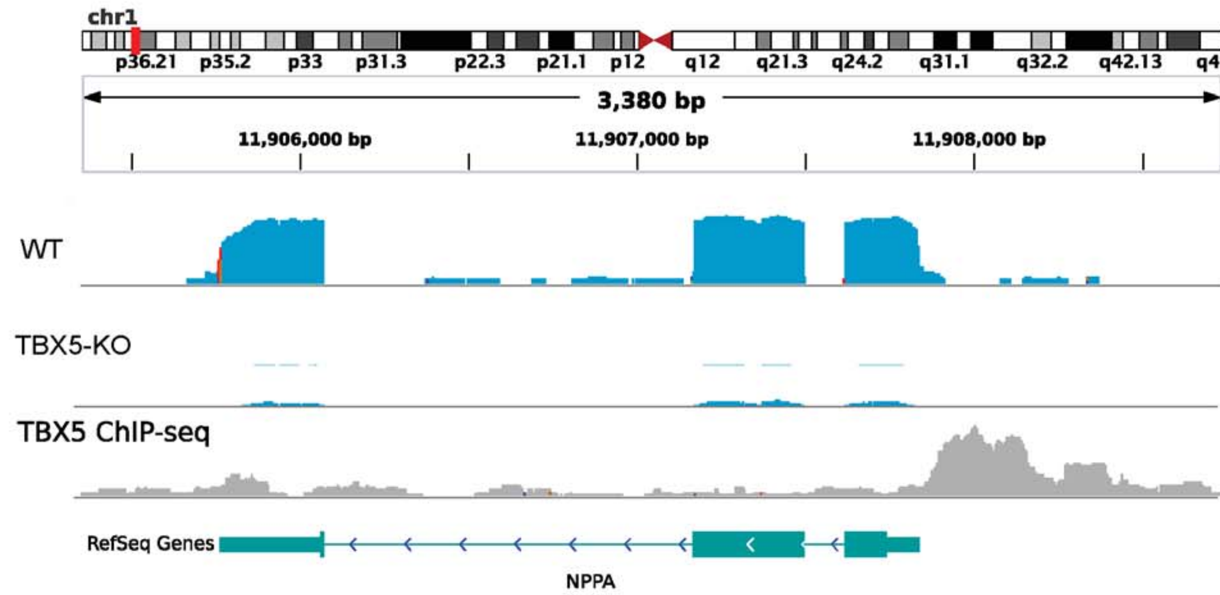
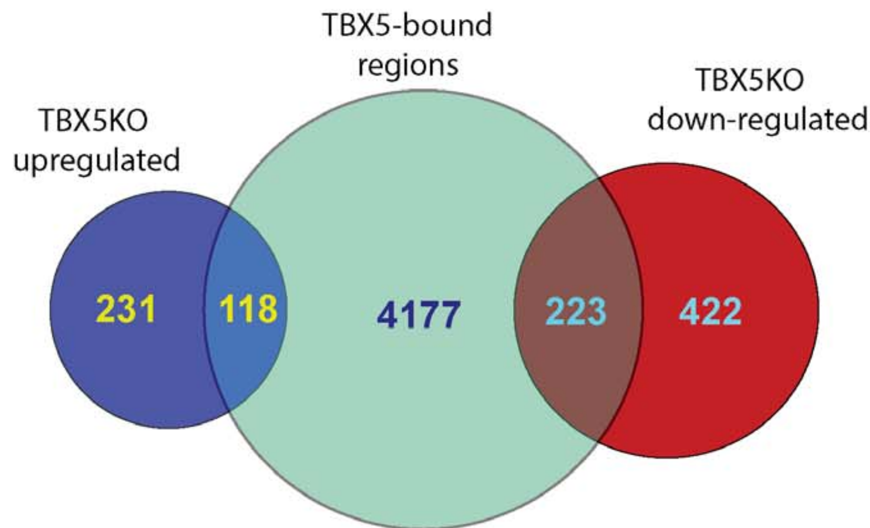
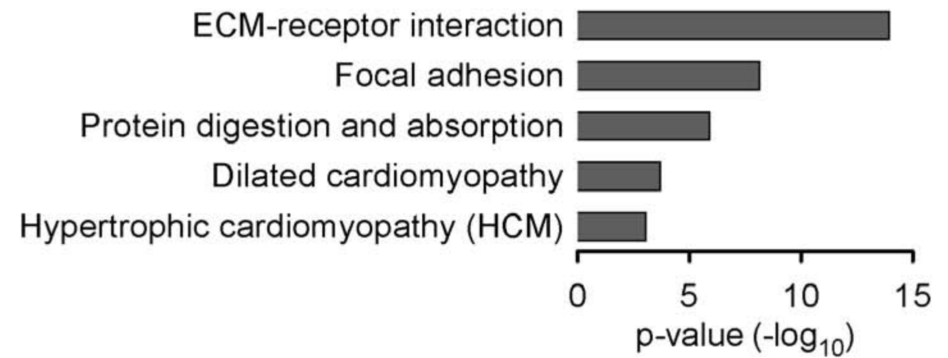


Figure 5









Circulation Research
JOURNAL OF THE AMERICAN HEART ASSOCIATION

A Comprehensive TALEN-Based Knockout Library for Generating Human Induced Pluripotent Stem Cell-Based Models for Cardiovascular Diseases

Ioannis Karakikes, Vittavat Termglinchan, Diana A Cepeda, Jaecheol Lee, Sebastian Diecke, Ayal Hendel, Ilanit Itzhaki, Mohamed Ameen, Rajani Shrestha, Haodi Wu, Ning Ma, Ning-Yi Shao, Timon Seeger, Nicole A Woo, Kitchener D Wilson, Elena Matsa, Matthew H Porteus, Vittorio Sebastiano and Joseph C Wu

Circ Res. published online February 28, 2017;
Circulation Research is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2017 American Heart Association, Inc. All rights reserved.
Print ISSN: 0009-7330. Online ISSN: 1524-4571

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circres.ahajournals.org/content/early/2017/02/28/CIRCRESAHA.116.309948

Data Supplement (unedited) at:
http://circres.ahajournals.org/content/suppl/2017/02/28/CIRCRESAHA.116.309948.DC1

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in *Circulation Research* can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to *Circulation Research* is online at:
http://circres.ahajournals.org//subscriptions/

SUPPLEMENTAL MATERIAL

A Comprehensive TALEN-Based Knockout Library for Generating Human Induced Pluripotent Stem Cell-Based Models for Cardiovascular Diseases

Ioannis Karakikes[1,2]*, Vittavat Termglinchan[1,3]*, Diana Cepeda[4]*, Jaecheol Lee[1], Sebastian Diecke[1,8,9], Ayal Hendel[5], Ilanit Itzhaki[1], Mohamed Ameen[1], Rajani Shrestha[1], Haodi Wu[1], Ning Ma[1], Ning-Yi Shao[1], Timon Seeger[1], Nicole Woo[1], Kitchener D. Wilson[1,6], Elena Matsa[1], Matthew H. Porteus[5], Vittorio Sebastiano[4,7], Joseph C. Wu[1,3,4]

[1]Stanford Cardiovascular Institute; [2]Department of Cardiothoracic Surgery; [3]Department of Medicine, Division of Cardiovascular Medicine; [4]Institute of Stem Cell Biology and Regenerative Medicine; [5]Department of Pediatrics; [6]Department of Pathology; [7]Department of Obstetrics and Gynecology, Stanford University School of Medicine, Stanford, CA 94305, USA; [8]Berlin Institute of Health, Germany; [9]Max Delbrueck Center, Berlin, Germany

*equal contribution

SUPPLEMENTAL METHODS

Genotyping iPSC clones. Genomic DNA was extracted from iPSC clones using the DNeasy Blood & Tissue Kit (Qiagen). Genotyping at the TALEN target site was then performed for each sample by PCR amplification using the PrimeSTAR GXL DNA Polymerase (Clonetech) with a primer pair designed to amplify a ~500 bp fragment surrounding the TALEN targeted site. The PCR amplicons were purified with the QIAquick PCR Purification Kit (Qiagen) and blunt-end cloned with the StrataClone Blunt PCR Cloning Kit (Stratagene) per manufacturer's protocol. The cloning reactions mixture (2 ul) was transformed into competent cells and plated on agar containing ampicillin (50 ug/ml) treated with 40 µl of 2% X-gal for blue-white color screening. After overnight incubation, white colonies were picked and grown for 16 hr at 37°C in ampicillin-containing LB broth. Plasmid DNA was extracted using the QIAprep Spin Miniprep Kit (Qiagen) and digested with EcoRI (Fermentas) to identify PCR insert-containing plasmids. Ten putative insert-containing plasmids were sequenced by Sanger to confirm presence of the mutant allele(s).

Immunocytochemistry. iPSCs were cultured on Matrigel-coated coverslips, fixed in 4% paraformaldehyde (10 min at room temperature), and permeabilized in blocking/permeabilization buffer (2% BSA / 2% FBS / 0.01 % Triton-X in PBS) for 45 min at room temperature and incubated with the indicated primary antibodies re-suspended in PBS / 2% BSA / 2% FBS. Following an overnight incubation at 4°C, the cells were washed three times in PBS-0.1% Tween-20 and incubated with an Alexa-conjugated secondary antibody (Life Technologies) diluted in blocking/permeabilization buffer (1:750). Finally, after washing three times in PBS / 0.1% Tween-20, the cells were counterstained with DAPI (Life Technologies). The following

antibodies were used: mouse monoclonal anti-OCT4 (1:100, Santa Cruz; sc-5279), goat polyclonal anti-NANOG (1:100, R&D systems; AF1997), mouse monoclonal anti-SOX2 (1:100, R&D systems; MAB2018), and mouse monoclonal anti-SSEA-4 (1:100, R&D systems; MAB1435). Similarly, iPSC-CMs were dissociated and cultured on Matrigel-coated coverslips for 4-5 days, then fixed in 4% paraformaldehyde and permeabilized in blocking/permeabilization buffer for 45 min. The cells were incubated with Alexa-conjugated primary antibodies overnight at 4°C, washed in PBS, and counterstained with DAPI. The following primary antibodies were used: mouse monoclonal anti-cardiac troponin T (1:200, Thermo Fisher Scientific; MS-295-P1) and mouse monoclonal anti-alpha actinin (1:200, Abcam; ab9465). For double staining experiments, the monoclonal antibodies were fluorescently labeled using the Zenon antibody labeling kit (Life Technologies), then applied directly to the samples. Immunofluorescence images were acquired using a Nikon epifluorescence microscope.

Western blot analysis. Cells were lysed in RIPA buffer (Sigma) supplemented with protease and phosphatase inhibitors (Roche) for 30 min on ice. Following lysis, cells were sonicated for 10 sec and then centrifuged (12,000g) for 10 min at 4°C. The protein concentration of the lysate was quantified using the Pierce BCA Protein Assay Kit (Thermo Fisher Scientific) and 30 μg of protein lysate was used in SDS polyacrylamide gel electrophoresis and followed by blotting. The blots were probed with antibodies against cardiac Troponin T (Thermo Fisher Scientific; MS-295-P1), alpha-sarcomeric actin (Abcam; ab28052), and TBX5 (Abgent; AP14687a).

Chromatin immunoprecipitation. Differentiated iPSC-CMs (2.5 x10^7) were infected (MOI = 1) with a lentivirus expressing a FLAG-epitope tagged TBX5 (TBX5-FLAG in pLX303 was a

gift from William Hahn; Addgene plasmid # 42563). After seven days, the cells were fixed with 1% formaldehyde for 10 min to generate protein-protein and protein-DNA crosslinks. The cross-linking reaction was stopped by adding 2.5 M glycine and incubated for 10 min at room temperature, washed twice with cold PBS. Cells were then scraped, mechanically sheared using sonication, and centrifuged at 10,000g for 30 min at 4°C. The supernatant was incubated overnight at 4°C with 10 µl of either anti-FLAG (F1804, Sigma-Aldrich) or mouse IgG (sc-2027, Santa Cruz Biotechnology) that were covalently conjugated to Dynabeads® Protein A/G (Life Technologies). A small portion of the crosslinked, sheared chromatin was saved and served as the 'Input' negative control DNA. The next day, the beads were rinsed with sonication buffer (50 mM Hepes pH 7.9, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Na-deoxycholate, 0.1% SDS, 0.5 mM PMSF), high salt buffer (50 mM Hepes pH 7.9, 500 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Na-deoxycholate, 0.1% SDS, 0.5 mM PMSF), and LiCl buffer (20 mM Tris, pH 8.0, 1 mM EDTA, 250 mM LiCl, 0.5% NP-40, 0.5% Na-deoxycholate, 0.5 mM PMSF). The washed beads were incubated with elution buffer (50 mM Tris, pH 8.0, 1 mM EDTA, 1% SDS, 50 mM $NaHCO_3$) for 1 hr at 65°C and then reverse cross-linked by adding 5 M NaCl and incubated overnight at 65°C. The immunoprecipitated DNA was treated with Rnase A and Proteinase K, and finally purified using the ChIP DNA clean and concentrator kit following the manufacturer's protocol (Zymo Research). Twenty ng of ChIP DNA or 'input' DNA was used for library preparation using the IonXpress Plus Fragment Library Kit according to the manufacturer's protocol (Publication Number 4473623 Revision B; Life Technologies). Briefly, the DNA was end-repaired and purified. The end-repaired DNA was ligated to Ion-compatible adapters, followed by nick repair to complete the linkage between barcode adapters and DNA inserts. The library was amplified by PCR and purified with two rounds of AMPure® XP

(Beckam-Coultier) bead capture to size-select fragments for downstream template preparation using the automated Ion Chef system. Sequencing was performed using the Ion PI Sequencing IC Kit and the Ion PI Chip v2 on the Ion Proton sequencer (Life Technologies).

Lentivirus production. The day prior to transfection, $5x10^6$ HEK293T cells (Life Technologies) were plated in 10 cm dish in DMEM media supplemented with 10% FBS. A transfection cocktail containing 2 µg FLAG-TBX5 (Addgene #42563) plasmid, 1.5 µg pMD2.G envelope plasmid (Addgene #12259), and 0.5 µg psPAX2 packaging plasmid (Addgene #12260) was prepared in 50 µl serum-free Opti-MEM (Life Technologies) and mixed with 12 µl Lipofectamine 2000 (Life Technologies) diluted in 50 µl serum-free Opti-MEM. After 10 min incubation at room temperature, the transfection mixture was added to the cells and incubated overnight at 37^oC and 5% CO_2. The next day, the media was replaced with serum-free OPTI-MEM and the transfected HEK293T cells were cultured for an additional 72 hr, and the supernatant was collected every 24 hr. The combined virus containing supernatant was centrifuged at 3000g for 15 min to remove the cell debris, followed by concentration by PEG-it according to the manufacturer's protocol (System Biosciences). The infectious viral titer in the concentrated supernatant was estimated by transfection of HEK293T cells with 10-fold serial dilutions (10^{-1} to 10^{-6}), followed by quantifying the number of FLAG-expressing cells or colonies of cells at 72 hr post-infection.

SNP karyotyping. SNP karyotype analysis was performed on the Illumina's CytoSNP-850K genotyping microarrays, which measure approximately 850,000 SNPs across the genome. All genomic DNA was isolated from iPSC clones according to the manufacturer's protocol (Qiagen). Input genomic DNA (500 ng) was processed, hybridized to the array, and scanned on an Illumina

HiScan according to the manufacturer's instructions. CNVs were identified using the cnvPartition Pluginv.3.2.0 in GenomeStudio (Illumina) by assessing both the B-allele-frequency and Log R ratios.

Ca^{2+} imaging. Dissociated iPSC-CMs were reseeded in Matrigel-coated 8-well Lab Tek II chambers (Nalge Nunc International). Cells were recovered for 3 days and were loaded with 5 μM Fluo-4 AM with 0.02% Pluronic F-127 (Molecular Probes) in Tyrode's solution for 15 min at 37°C, and were washed with Tyrode's solution afterwards. Ca^{2+} imaging was conducted using a Zeiss LSM 510Meta confocal microscope (Carl Zeiss AG, Göttingen, Germany). Spontaneous Ca^{2+} transients of single beating iPSC-CMs were obtained using a time-lapse line scanning recording mode (512 pixels x 1920 lines) under 40X objective (Plan Apochromat, 0.95 NA) at 37°C, and the raw data was analyzed using customized Interactive Digital Language (IDL) script. Ca^{2+} signal was normalized to the intracellular basal line (F_0), and transient amplitude was expressed as $\Delta F/F_0$.

Validation of RNA-seq data by qPCR Total RNAs were isolated from iPSC-CMs using the miRNeasy Mini kit (QIAGEN). 1 μg of RNA was used to synthesize cDNA using the iScript™ cDNA Synthesis kit (Bio-Rad). 0.25 μl of the reaction was used to quantify gene expression by qPCR using TaqMan probes and TaqMan Universal PCR Master Mix. Expression values were normalized to the average expression of housekeeping gene 18s.

ONLINE FIGURE LEGENDS

Online Figure I. A) Representative immunofluorescence images of isogenic TNNT2-KO iPSC colonies stained for the pluripotency-associated markers OCT-4, NANOG, SOX-2 and SSEA-4, as indicated. **B)** Relative mRNA expression of pluripotency-associated genes NANOG, OCT-3/4 and SOX-2. Expression levels are expressed relative to the parental iPSC line. Values represent mean ± SEM (n=3).

Online Figure II. Intracellular calcium cycling analysis. **A)** Representative line-scan images and spontaneous Ca^{2+} transients for isogenic wild-type (WT), heterozygous (*TNNT2*$^{+/}$), and homozygous (*TNNT2*$^{-/-}$) knockout iPSC-CMs. **B)** Comparison of tangential amplitude, time to peak, and decay tau of calcium imaging between each isogenic group. Data represents mean ± SEM of n = 25 single iPSC-CMs per line. Unpaired two–tailed t–test with **P < 0.01, n.s. = not significant.

Online Figure III. A) Representative immunofluorescence images of isogenic DCM-KO iPSC colonies stained for the pluripotency-associated markers OCT-4, NANOG, SOX-2, and SSEA-4, as indicated. **B)** Relative mRNA expression of pluripotency-associated genes NANOG, OCT-3/4, and SOX-2. Expression levels are expressed relative to the parental iPSC line. Values represent mean ± SEM (n=3). **C)** Digital karyotype analysis of the parental iPSC clone.

Online Figure IV. A) ddPCR for the TNNT2 R173W mutant and wild-type allelic discrimination from the parental- and DCM-KO iPSC-CMs. Green and blue dots represent droplets containing the mutant and the wild-type alleles, respectively. Pink line indicates the

detection threshold. **B)** Quantification of ddPCR shows the average frequency of the WT and mutant alleles in the iPSC-CMs as indicated. Values represent mean ± SEM (n=3).

Online Figure V. RNA-seq analysis of TBX5 gene isoforms in iPSC-CMs derived from the indicated iPSC lines generated by the Stanford CVI iPSC Biobank.

Online Figure VI. A) Representative immunofluorescence images of isogenic TBX5-KO iPSC colonies stained for the pluripotency-associated markers OCT-4, NANOG, SOX-2 and SSEA-4, as indicated. **B)** Relative mRNA expression of pluripotency-associated genes NANOG, OCT-3/4, and SOX-2. Expression levels are expressed relative to the parental iPSC line. Values represent mean ± SEM (n=3). **c)** Digital karyotype analysis of the parental iPSC clone.

Online Figure VII. Quantification of the cardiomyocyte differentiation efficiency. Flow cytometry analysis of the differentiation efficiency of isogenic TBX5-KO and parental WT iPSC lines at 15 days after differentiation. Representative contour plots of iPSC-CMs immunolabeled with isotype control antibody (IgG-Alexa-488) or cardiac troponin T antibody (cTnT-Alexa-488) in isogenic iPSC-CMs as indicated.

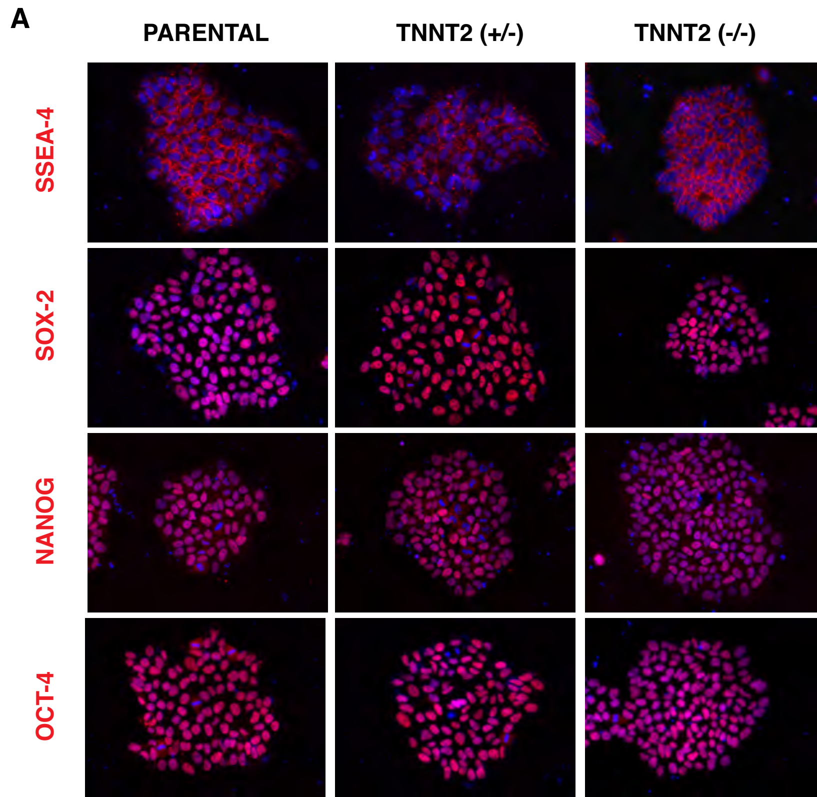
Online Figure VIII. Validation of RNA-seq data by qPCR. Quantitative real-time PCR of selective differentially expressed genes. Gene expressions were normalized to 18s and expressed as fold-change relative to parental WT iPSC-CMs. **A)** Upregulated genes and **B)** downregulated genes from RNA-seq data. Values represent mean ± SEM (n=3).

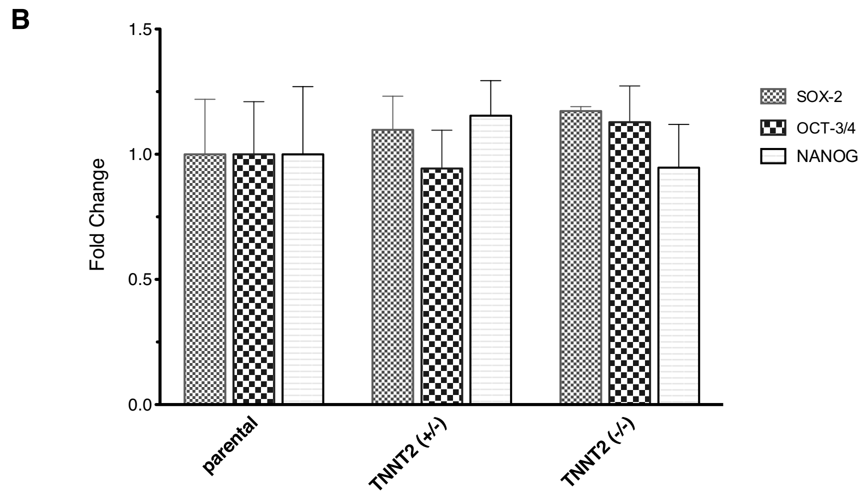
Online Figure XI. *In vitro* TBX5 binding motifs. De novo motif discovery of *in vitro* motif by HOMER using the TBX5 peaks of the ChIP-seq data. Motifs found by *de novo* discovery were compared with available consensus and optimal *in vitro* motifs from the indicated reference.

SUPPLEMENTAL REFERENCES

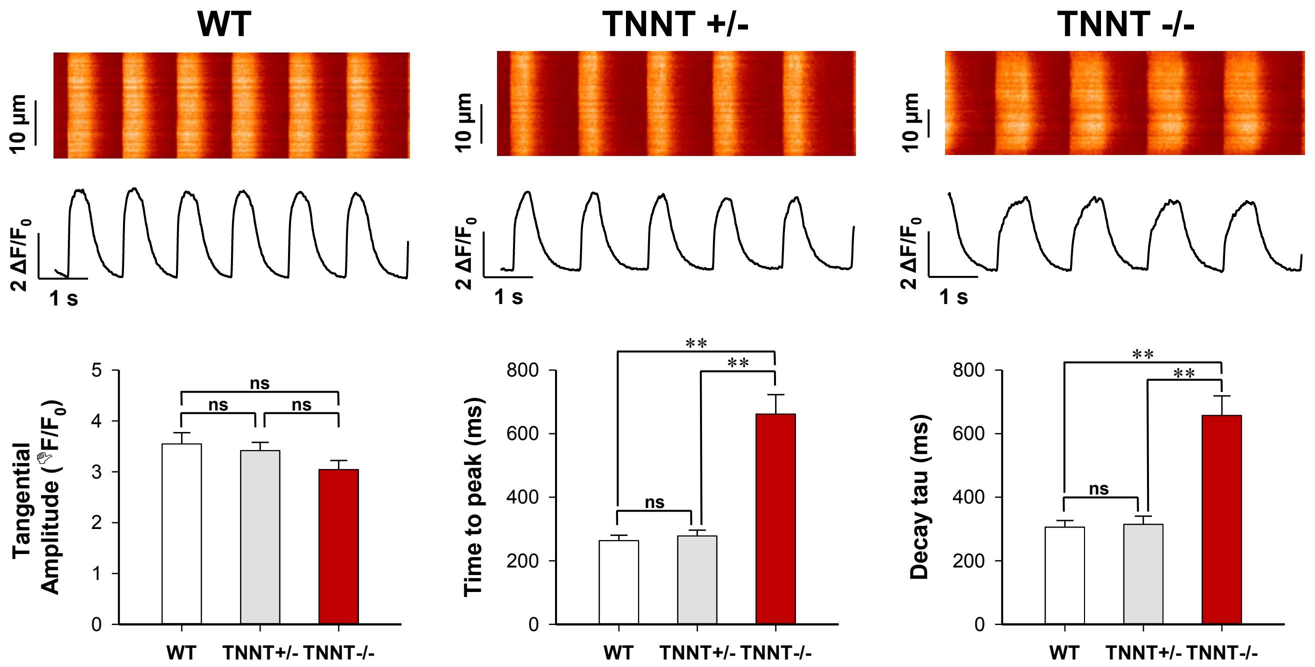
1. Mori, A.D.*, et al.* Tbx5-dependent rheostatic control of cardiac gene expression and morphogenesis. *Developmental Biology* 297, 566-586 (2006).
2. He, A., Kong, S.W., Ma, Q. & Pu, W.T. Co-occupancy by multiple cardiac transcription factors identifies transcriptional enhancers active in heart. *Proceedings of the National Academy of Sciences of the United States of America* 108, 5632-5637 (2011).

Online Figure I

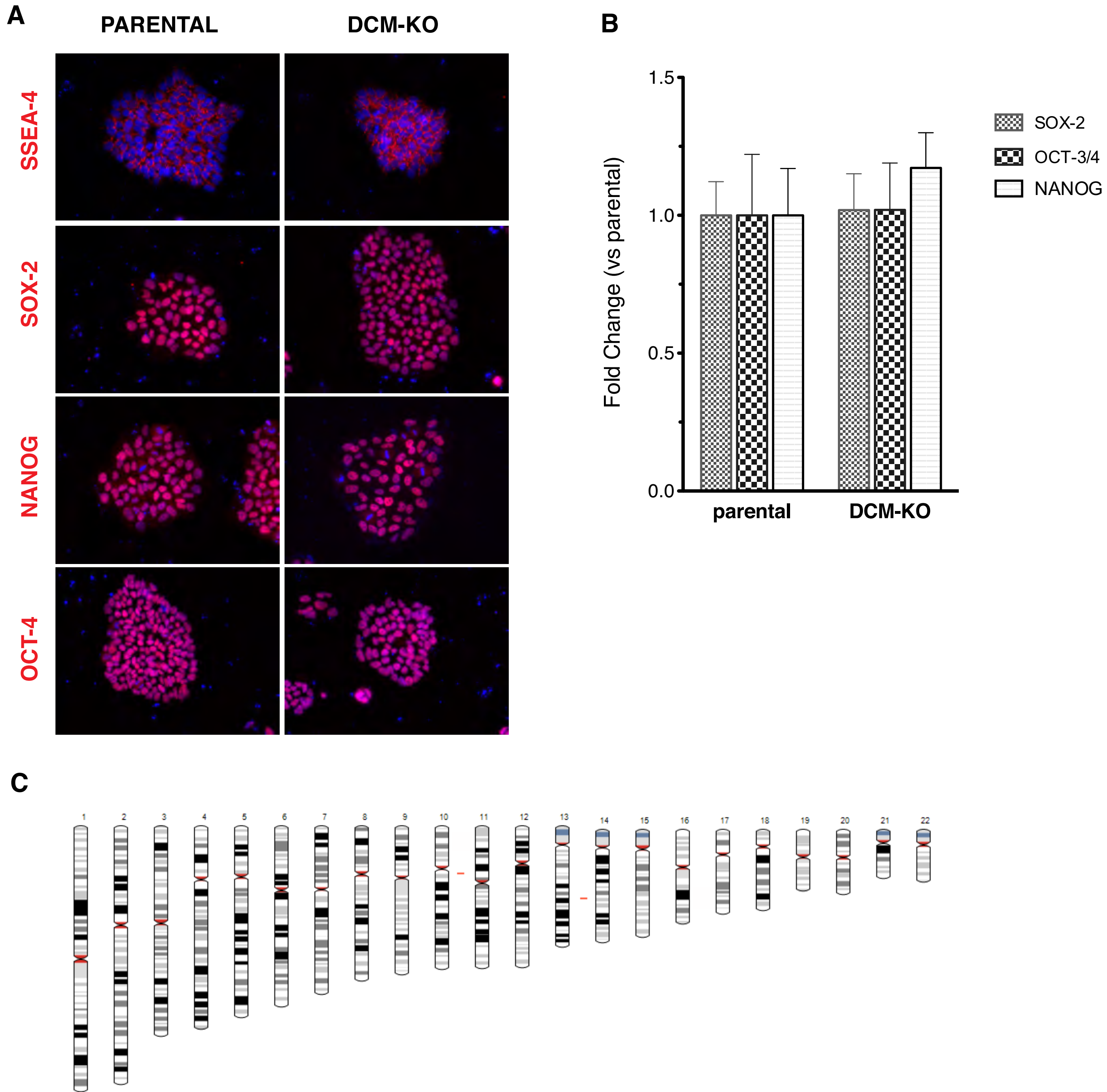




Online Figure II

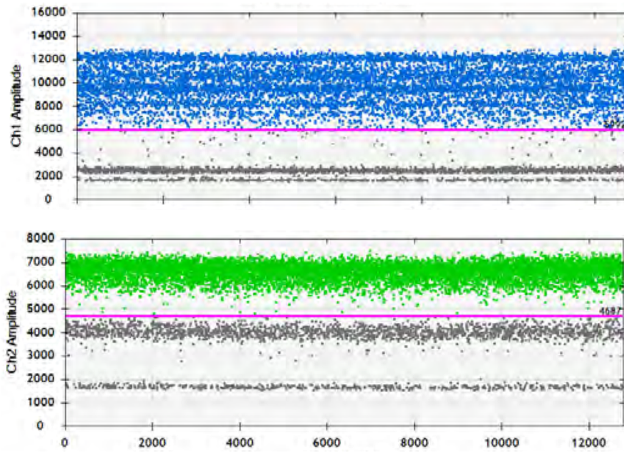


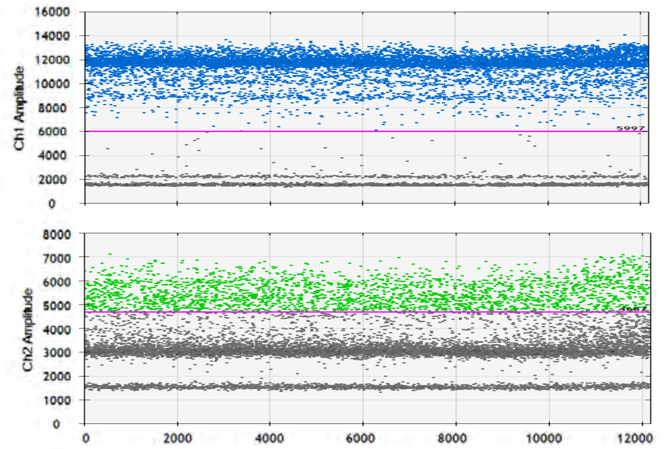
Online Figure III



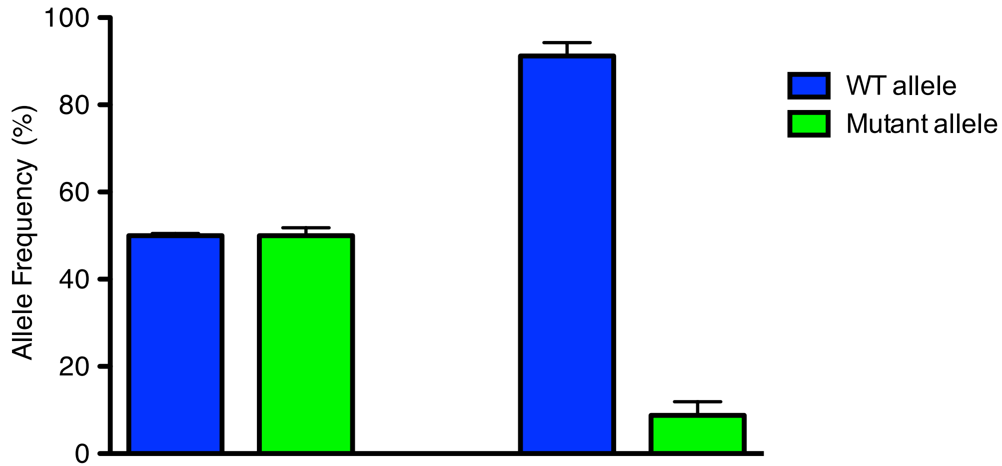
Online Figure IV

A

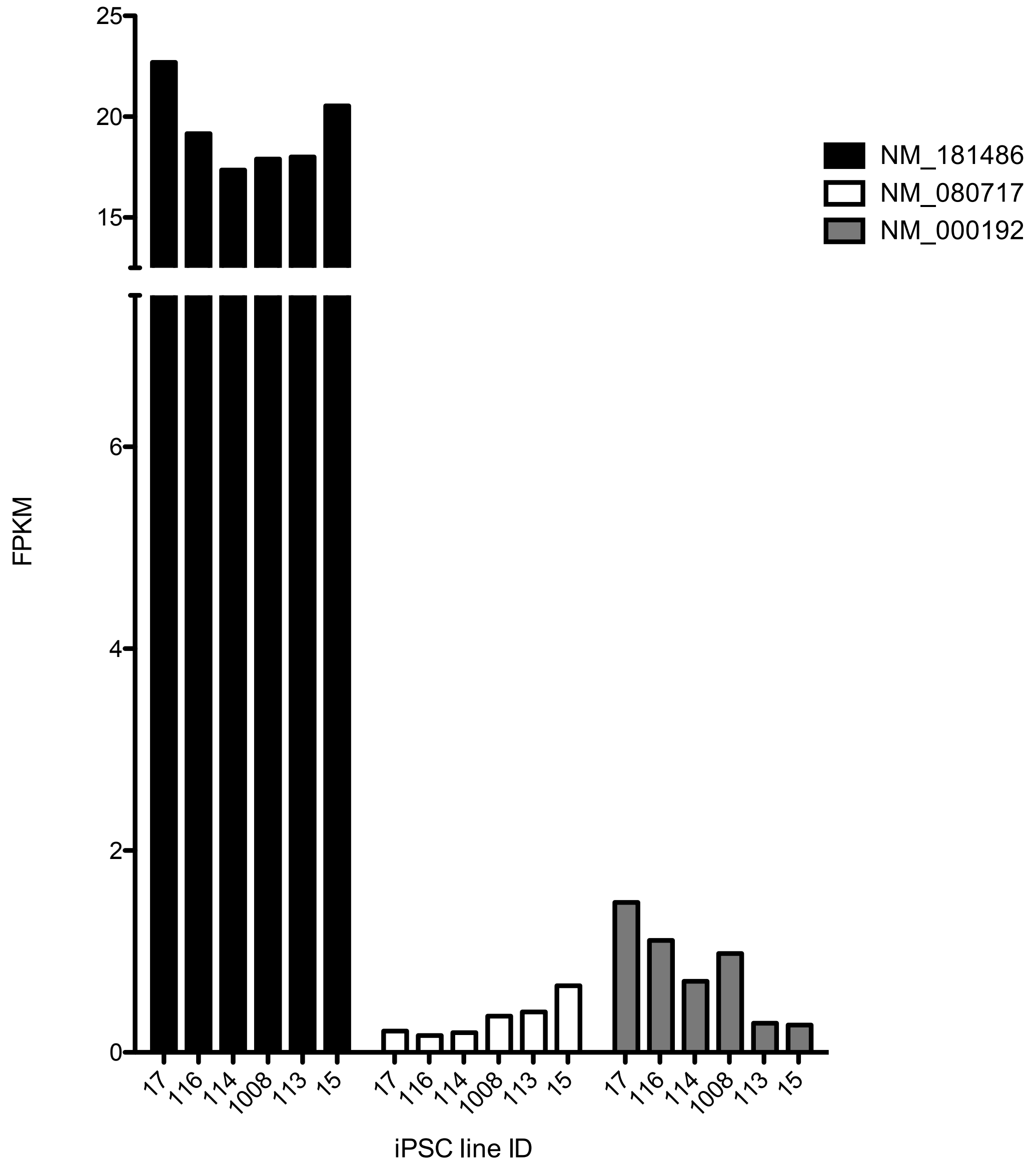




B

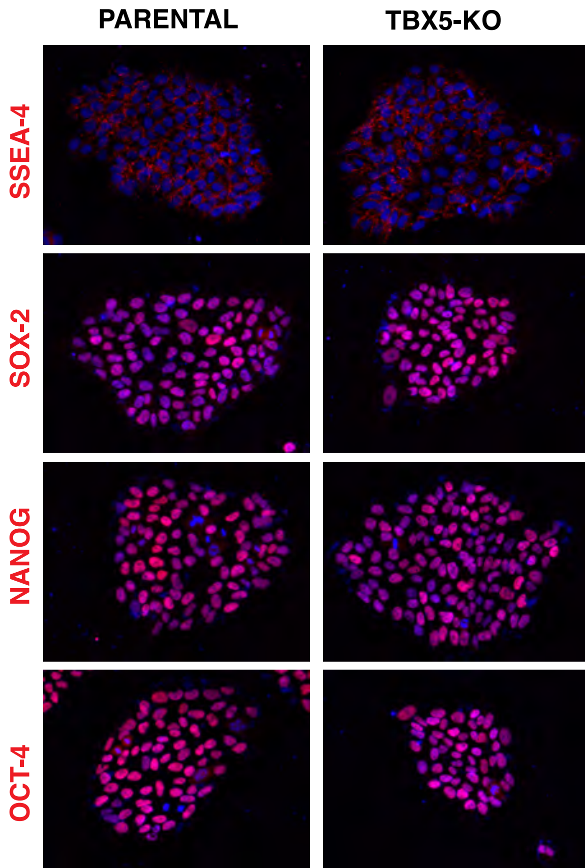


Online Figure V

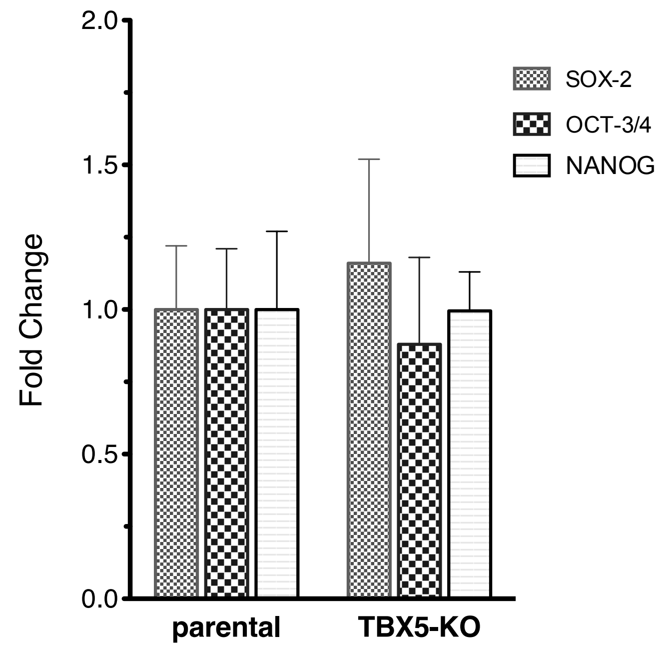


Online Figure VI

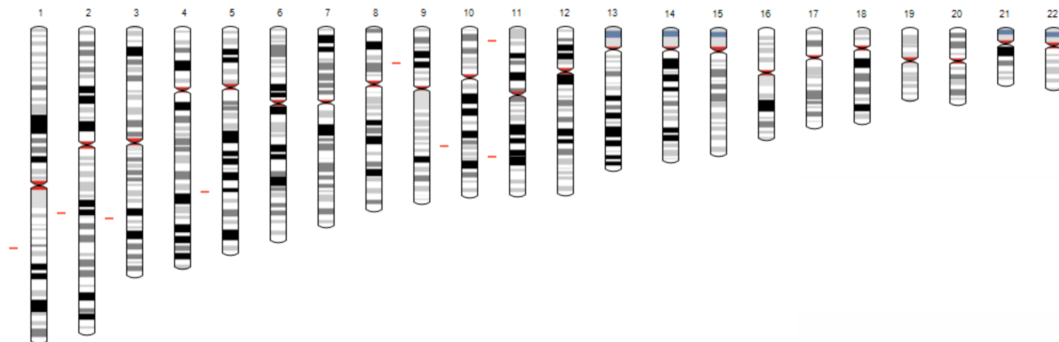
A



B

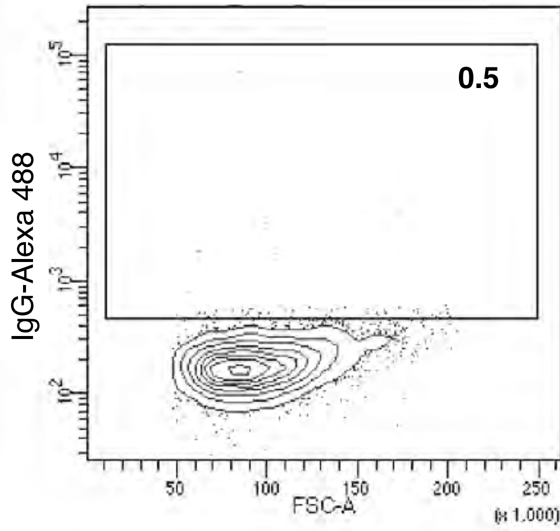


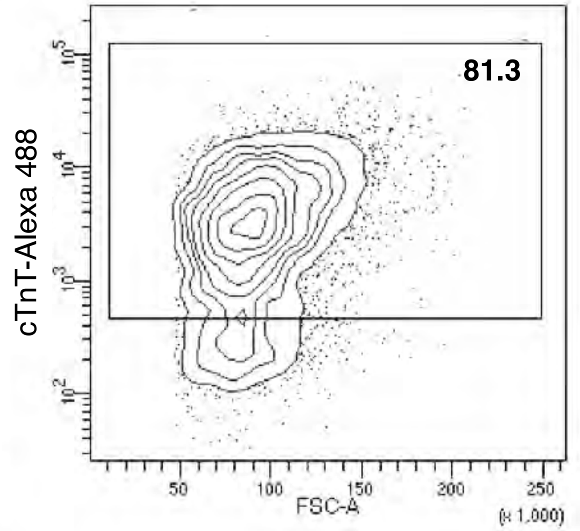
C



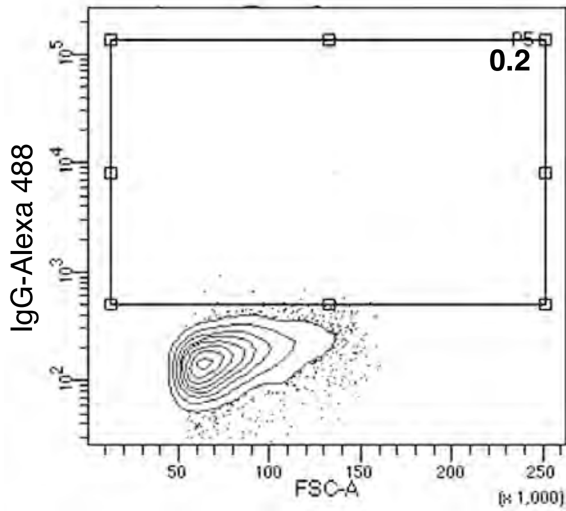
Online Figure VII

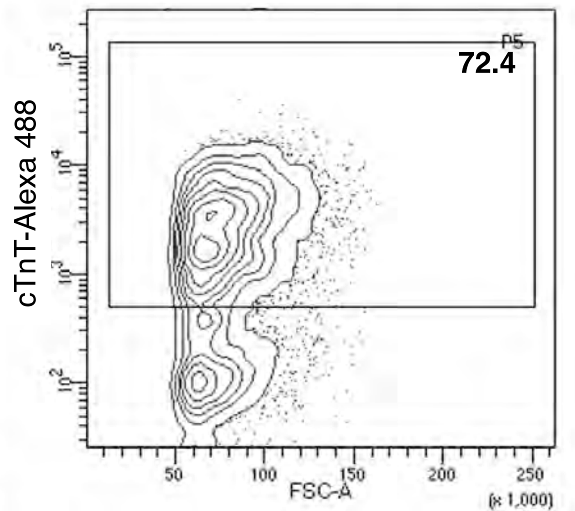
TBX5-KO



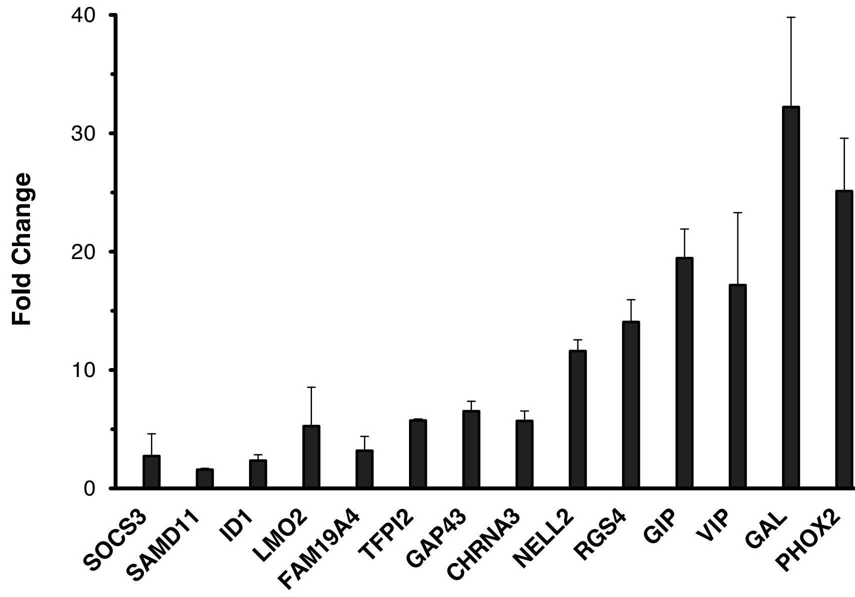


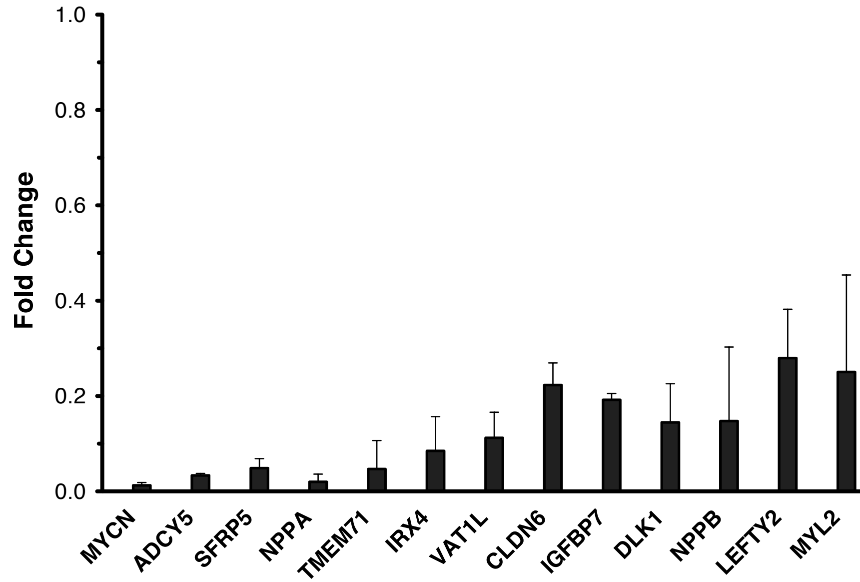
Parental





Online Figure VIII





Online Figure IX

Consensus motif

Mori, *et al*

Known motif

He, *et al*

ChIP Homer

Online Table I. Mutagenesis Efficiency of TALEN Pairs in human iPSCs as assessed by single-molecule real time (SMRT) technology. NHEJ, Non-homologous end joining.

Clone ID	Gene	TALEN Plus Strand Target Sequence	NHEJ %
DC47B, DC48	ABCC9	**T AAGAAGAAATGAGCC tttcattttgtggta ACAACATTTCTTCAT A**	6%
DC49B, DC50B	ACADS	**T GGCCGCCGCGCTGCT cgcccgggcctcgggcc CTGCCCGCAGAGGTG A**	44.80%
DC73B, DC74B	ACADVL	**T CGAGCCAGCGGCGCC cggagagattcggag ATGCAGGCGGCTCGG A**	42.67%
DC51, DC52	ACTC1	**T GCAGAACCCCCTGAA gctgtgccaagatgtgt GACGACGAGGAGACC A**	13.37%
DC53, DC54	ACTN2	**T CGCGCCCCGCCGCAG ccccggccaaccgagcg CCATGAACCAGATAG A**	2.82%
DC59, DC60	ANKRD1	**T CCTTCAGCCAACATG atggtactgaaagtagag GAACTGGTATGTAAG A**	6.73%
DC65B, DC66B	BAG3	**T GAGCGCCGCCACCCA ctcgcccatgatgc AGGTGGCGTCCGGCA A**	8.40%
DC67B, DC68B	CALR3	**T GCACACCCCCATGGC ccgggctttggtccag CTCTGGGCCATATGC A**	3.46%
DC69, DC70	CASQ2	**T GGGAACGAGAAACAA aagttttcccaaatgaag AGAACTCACTTGTTT A**	0.23%
DC71C, DC72B	CAV3	**T GGATCCCCCCAGCTC tgcgatgatggcagaag AGCACACAGATCTCG A**	10.34%
DC75B, DC76B	CHD7	**T GGTTTGGAGGAGCCG tgtgttggaagaagatg GCAGATCCAGGAATG A**	0.33%
DC77C, DC78	COX15	**T GTCATCAGTATGCAG cgattgctctttccg CCGTTGAGGGCCTTG A**	10.21%
DC79B, DC80	CRYAB	**T CACACTCACCTAGCC accatggacatcgcc ATCCACCACCCCTGG A**	11.49%
DC81B, DC82B	CSRP3	**T GACCTTGACCAGATA gtcttcaagatgccaaac TGGGGCGGAGGCGCA A**	1.61%
DC83C, DC84B	CTF1	**T GAAGGGAGCCGGGAT cagccaggggccagcat GAGCCGGAGGGAGGG A**	19.70%
DC209, DC210	CTNNA3	**T GTTTGTGCACAGGCA gcatgtcagctgaa ACACCAATCACATTG A**	7.72%
DC85C, DC86B	DES	**T CACCATGAGCCAGGC ctactcgtccagcc AGCGCGTGTCCTCCT A**	33.75%
DC133, DC134	DMD	**T ATCGCTGCCTTGATA tacacttttcaaaatgct TTGGTGGGAAGAAGT A**	0.25%
DC135, DC136	DNAJC19	**T GGTGAGTGCGGCCTT ccggtcttcttggcgacc TCCGGCCCAGGCCTC A**	1.13%
DC87, DC88	DSC2	**T GCCCCGAGCCCTCTC catggaggcagcccgc CCCTCCGGCTCCTGG A**	28%
DC89B, DC90B:	DSG2	**T GCGGCGGCGGGAGGC ggaggcgagggtgcgat GGCGCGGAGCCCGGG A**	1.78%
DC91B, DC92B	DSP	**T GCCCGCCGACATGAG ctgcaacggaggct CCCACCCGCGGATCA A**	4.86%
DC137, DC138	DTNA	**T ACACATTGTAACTAT tttgtctcatagaatgat TGAAGATAGTGGGAA A**	4.67%
DC93, DC94	EMD	**T CCGCCTGAGCCCGCA cccgccatggacaact ACGCAGATCTTTCGG A**	49.67%
DC95, DC96	EYA4	**T GAGAAAACCACATGG aagactcccaggattt AAATGAACAATCAGT A**	21.71%
DC139, DC140	FHL1	**T CCAGCTACAAGGTGG gcaccatggcggaga AGTTTGACTGCCACT A**	5.80%
DC141, DC142	FHL2	**T TGCTGAAAAGCCAGG agtcaaaatgactgagc GCTTTGACTGCCACC A**	3.55%
DC97, DC98	FKTN	**T CAAAAGACAACCAAG tgagcagcacagacta ATGAGTAGAATCAAT A**	2.92%

DC99, DC100	FXN	**T GTGGACTCTCGGGCG ccgcgcagtagccggc CTCCTGGCGTCACCC A**	6.15%
DC101, DC102	GATA4	**T ATCAGAGCTTGGCCA tggccgccaaccacgggc CGCCCCCCGGTGCCT A**	0.49%
DC103, DC104	GATAD1	**T CTGCGCCCGCGGGGG ccgcccgagccggccacc ATGCCGCTGGGCCTG A**	5.66%
DC105, DC106	GLA	**T ATGCTGTCCGGTCAC cgtgacaatgcagct GAGGAACCCAGAACT A**	12.75%
DC211, DC212	HOPX	**T GCCCCGCAGCGCGCA gggaccatgtcggcggag ACCGCGAGCGGCCCC A**	3%
DC107, DC108	ILK	**T CGGCGCCGGGACGCT gctatggacgacattt TCACTCAGTGCCGGG A**	5.44%
DC109, DC110	JAG1	**T CCCGAGTGCCCGCGG cgcgcggcgcagcg ATGCGTTCCCCACGG A**	7.99%
DC111, DC112	JPH2	**T TGTCAGGGGCTATGA tgagatgagtgggggcc GCTTCGACTTTGATG A**	1.56%
DC113, DC114	JUP	**T CCTTTGTGCCCCCAG tagccacgatggaggtga TGAACCTGATGGAGC A**	0.83%
DC115B, DC116	LAMA4	**T TGAGCTCAGCCTGGC gctcggttctgcct CTGTGGCTCCTCTGG A**	4.51%
DC117, DC118	LAMP2	**T CTGCGGGGTCATGGT gtgcttccgcctctt CCCGGTTCCGGGCTC A**	12.14%
DC45&DC46:	LMNA	**T CCGGGACCCCTGCCC cgcgggcagcgctgcca ACCTGCCGGCCATGG A**	12.54%
DC35 & DC36	LMNA	**T GCCAACCTGCCGGCC atggagaccccgtcccag CGGCGCGCCACCCGC A**	18.70%
DC119, DC120	MYLCD	**T CGGCAGCTGTTGTGG ggcaccatgcgaggc TTCGGGCCAGGCTTG A**	26.77%
DC121, DC122	MYBPC3	**T CGTGCCTGGTGTGAC gtctctcaggatgcctga GCCGGGGAAGAAGCC A**	0.69%
DC123, DC124	MYH6	**T CTCTGACCCAGGGGA agcaccaagatgaccg ATGCCCAGATGGCTG A**	9.32%
DC43 & DC44	MYH7	**T GGCAGTCTTTGGGGC tgccgccccctacc TGCGCAAGTCAGAGA A**	6.08%
DC41 & DC42	MYH7	**T TCGGAGATGGCAGTC tttggggctgccgcccc CTACCTGCGCAAGTC A**	50.22%
DC125, DC126	MYL2	**T GCTGGGTCCTTTCCA ccatggtgagtacaaggg CTCCAGGAGGTGATG A**	0.51%
DC127, DC128	MYL3	**T GTACTTACAGCCCCC aatggcccccaaaaagc CAGAGCCCAAGAAGG A**	7.78%
DC129, DC130	MYLK2	**T CCCTACCTCATGGCG acagaaaatggagcagtt GAGCTGGGAATTCAG A**	5.98%
DC131, DC132	MYOM1	**T CCTTCAAGGGGCACA ggatgtctttgccttttt ATCAGAGGTGCCACC A**	0.84%
DC143, DC144	MYOZ2	**T AATACTATGATGAAG cagagaaaacagcaa GCAACAGCCATCATG A**	1.44%
DC145, DC146	MYPN	**T TTGTGACAGCATGCA agacgacagcataga AGCTTCTACTTCCAT A**	24.84%
DC213, DC214	NEBL	**T GAGGGTCCCTGTATT tgaggatataaaagat GAAACTGAAGAAGAA A**	1.27%
DC147, DC148	NEXN	**T AGAGCAAACATGAAT gatatttcccaaaag GCTGAGGTAAGTCTC A**	11.87%
DC57B & DC58	NKX2.5	**T GAGACTGGCGCTGCC accatgttccccagc CCTGCTCTCACGCCC A**	9.40%
DC149, DC150	PDLIM3	**T CAGAGCCCGGTGGGC gggaggaaggcggc ATGCCCCAGACGGTG A**	1.19%
DC151, DC152	PKP2	**T CGGTCGCCCCCACCG gccccatggcagcccccg GCGCCCCAGCTGAGT A**	4.43%
DC215, DC216	PLN	**T TCCTGTCCTGCTGGT atcatggagaaagtcca ATACCTCACTCGCTC A**	0.73%

DC153, DC154	PRKAG2	**T CAACTTCTGGTTAGA gttatgggaagcgcggtt ATGGACACCAAGAAG A**	4.81%
DC155, DC156	PSEN1	**T CTATACAGTTGCTCC aatgacagagttac CTGCACCGTTGTCCT A**	2.31%
DC157, DC158	PSEN2	**T CCAGGTGCTTCCAGA ggcagggctatgctca CATTCATGGCCTCTG A**	11.91%
DC159, DC160	PTPN11	**T CGCGGAGCCGGAGGG cgggaggaacatgac ATCGCGGAGGTGAGG A**	3.87%
DC161, DC162	RAF1	**T AAGCTGCATCAATGG agcacatacaggga GCTTGGAAGACGATC A**	4.06%
DC163, DC164	RBM20	**T CCCGGGCGGGTCTCG ccccgcatggtgctgg CAGCAGCCATGAGCC A**	3.60%
DC165, DC166	RYR2	**T GGCCGATGGGGGCGA gggcgaagacgagatcca GTTCCTGCGAACTGT A**	0.83%
DC167, DC168	SCN5A	**T GAGAAGATGGCAAAC ttcctattacctcggggc ACCAGCAGCTTCCGC A**	1.94%
DC217, DC218	SCO2	**T GTTTCCAGGAGCATC agatccatgctgctgct GACTCGGAGCCCCAC A**	2.43%
DC169, DC170	SDHA	**T CCGGGGCCTGTCGCG gctgctgagcgctcgg CGCCTGGCGCTGGCC A**	3.64%
DC171, DC172	SGCD	**T GAGTGAAGGGACCAG gtggagatggtgag TAATTCCCGGGAGCG A**	0.32%
DC173, DC174	SLC25A20	**T GACAGACGGAGTGAC agacggactgacca TGGCCGACCAGCCAA A**	2.21%
DC219, DC220	SLC25A4	**T GAGAGCGTCGAGCTG tcaccatgggtgatca CGCTTGGAGCTTCCT A**	7.63%
DC175, DC176	SURF1	**T GGCGGCGGTGGCTGC gttgcagctggggctgcg GGCGGCGGGGCTGGG A**	3.87%
DC177, DC178	SYNE1	**T CCGGAGGGACCATGG caacctccagaggggcct CCCGGTGTCCTCGGG A**	0.48%
DC179, DC180	TAZ	**T GGGAGCGCCGGCCGC gggccgggtgggga TGCCTCTGCACGTGA A**	0.91%
DC181, DC182	TBX1	**T GCCAGGATCCCCGGC agggatgcacttca GCACCGTCACCAGGG A**	5.24%
DC183, DC184	TBX20	**T GGCCAGGACCGCGTG ctggggaccatggagt TCACGGCGTCCCCCA A**	15.26%
DC61B & DC62	TBX5	**T GGGCGCACCATGGCC gacgcagacgagggc TTTGGCCTGGCGCAC A**	48.45%
DC185, DC186	TCAP	**T GAGGAGTGATCATGG ctacctcagagctgagct GCGAGGTGTCGGAGG A**	1.32%
DC187, DC188	TGFB3	**T CCCCCTGGCCTCTCT tcccagctcacacatg AAGATGCACTTGCAA A**	13.38%
DC189, DC190	TMEM43	**T CCCACCATGGCCGCG aatgtgagtatccccg GGCCAGCCGGGCCAC A**	2.43%
DC191B, DC192B	TMPO	**T GGGGAGGGGGCTTCG cagatccccgagatgc CGGAGTTCCTGGAAG A**	5.05%
DC193, DC194	TNNC1	**T CCTGTGAGCCGCCAG catggatgacatctaca AGGCTGCGGTGAGGG A**	7.25%
DC195, DC196	TNNI3	**T CCCGGCCTGAGTCTC agcatggcggatgggtga GTGATGCCCCAAGGC A**	1.70%
DC39 & DC40	TNNT2	**T TTGGAGGGAGAGCAG agaccatgtctgaca TAGAAGAGGTGGTGG A**	2.79%
DC37 & DC38	TNNT2	**T TTTCTCCTTTTGGAG ggagagcagagacca TGTCTGACATAGAAG A**	13.14%
DC197, DC198	TPM1	**T CGCCGCCGCCACCAT ggacgccatcaagaag AAGATGCAGATGCTG A**	6.45%
DC199, DC200	TTN	**T TTTCAGAGTGCCTAG aaagatgacaactcaag CACCGACGTTTACGC A**	0.72%
DC201, DC202	TTR	**T TGGCAGGATGGCTTC tcatcgtctgctcct CCTCTGCCTTGCTGG A**	2.86%

DC203, DC204	TXNRD2	**T GGCGGTGGCGCTGCG gggattaggagggcgct TCCGGTGGCGGACGC A**	0.57%
DC205, DC206	VCL	**T TCGCCGCCCCGCTCG ccgccgcgatgccagtg TTTCATACGCGCACG A**	1.25%
DC207, DC208	ZASP	**T GCAGAGGCGGCCGCT gacagcaccagcatgtct TACAGTGTGACCCTG A**	3.09%

Online Table II. Frequency and position of TALEN-mediated mutagenesis in human iPSCs. Deletions and insertions of the top 5 variants are shown.

ABCC9 Mutations in 134 of 2738 sequences ≈ 4.9%

```
AGAAGAAATGAGCCTTTCATTTTGTGGTAACAACATTTCTTCATATAATATCAACGATGGTGTACTACAAAATTCCTGCTTTGTGGAT  WT
AGAAGAAATGAGCCTTTCATTTTGTGGTAACAACATTTCT----------TCAACGATGGTGTACTACAAAATTCCTGCTTTGTGGAT  Δ10 x15
AGAAGAAATGAGCCTTTCATTTTGTGGTAACAACATTTCTTC----AATATCAACGATGGTGTACTACAAAATTCCTGCTTTGTGGAT  Δ4 x7
AGAAGAAATGAGCCTTTCATTTTGTGGTAA---------------------CAACGATGGTGTACTACAAAATTCCTGCTTTGTGGAT  Δ21 x6
AGAAGAAATGAGCCTTTCATTTTGTGGTAACAAC------------AATATCAACGATGGTGTACTACAAAATTCCTGCTTTGTGGAT  Δ12 x4
AGAAGAAATGAGCCTTTCATTTTGTGGTAAC---------------AATATCAACGATGGTGTACTACAAAATTCCTGCTTTGTGGAT  Δ15 x4
```

ACADS Mutations in 888 of 1982 sequences ≈ 44.8%

```
GGGACTGTGTCTGTCGCCCATGGCCGCCGCGCTGCTCGCCCGGGCCTCGGGCCCTGCCCGCAGAGGTGAGTGCGCTGGGGATCCGTAC  WT
GGGACTGTGTCTGTCGCCCATGGCCGCCGCGCTGCTCGCC-------CGGGCCCTGCCCGCAGAGGTGAGTGCGCTGGGGATCCGTAC  Δ7 x49
GGGACTGTGTCTGTCGCCCATGGCCGCCGCGCTGCTC------------------GCCCGCAGAGGTGAGTGCGCTGGGGATCCGTAC  Δ18 x44
GGGACTGTGTCTGTCGCCCATGGCCGCCGCGCTGCTC-------------GCCCTGCCCGCAGAGGTGAGTGCGCTGGGGATCCGTAC  Δ13 x41
GGGACTGTGTCTGTCGCCCATGGCCGCCGCG----------------------CTGCCCGCAGAGGTGAGTGCGCTGGGGATCCGTAC  Δ22 x12
GGGACTGTGTCTGTCGCCCATGGCCGCCGCGCT-----------------GCCCTGCCCGCAGAGGTGAGTGCGCTGGGGATCCGTAC  Δ17 x11
```

ACADVL Mutations in 1092 of 2559 sequences ≈ 42.7%

```
CGCCAGAGCTGGGTCAGAGCTCGAGCCAGCGGCGCCCGGAGAGATTCGGAGATGCAGGCGGCTCGGATGGCCGCGAGCTTGGGGCGGC  WT
CGCCAGAGCTGGGTCAGAGCTCGAGCCAGCGGCGCC----------CGGAGATGCAGGCGGCTCGGATGGCCGCGAGCTTGGGGCGGC  Δ10 x418
CGCCAGAGCTGGGTCAGAGCTCGAGCCAGCGGCGCC----------CGGAGATGCAGGCGGCTCGGATGGCCGCGAGCTT-GGGCGGC  Δ11 x45
CGCCAGAGCTGGGTCAGAGCTCGAGCCAGCGGCGC-----------CGGAGATGCAGGCGGCTCGGATGGCCGCGAGCTTGGGGCGGC  Δ11 x45
CGCCAGAGCTGGGTCAGAGCTCGAGCCAGCGGCGCCCGGA--------GAGATGCAGGCGGCTCGGATGGCCGCGAGCTTGGGGCGGC  Δ8 x31
CGCCAGAGCTGGGTCAGAGC-------------------------TCGGAGATGCAGGCGGCTCGGATGGCCGCGAGCTTGGGGCGGC  Δ25 x18
```

ACTC1 Mutations in 451 of 3372 sequences ≈ 13.4%

```
CGCCCTCCCCTCCTCAACCTGCAGAACCCCCTGAAGCTGTGCCAAGATGTGTGACGACGAGGAGACCACCGCCCTGGTGTGCGACAAC  WT
CGCCCTCCCCTCCTCAACCTGCAGAACCCCCTGAAGC----------TGTGTGACGACGAGGAGACCACCGCCCTGGTGTGCGACAAC  Δ10 x77
CGCCCTCCCCTCCTCAACCTGCAGAACCCCCTGAAGC------------TGTGACGACGAGGAGACCACCGCCCTGGTGTGCGACAAC  Δ12 x12
CGCCCTCCCCTCCTCAACCTGCAGAACCCCCTGAAGCTGTGCcaaCAAGATGTGTGACGACGAGGAGACCACCGCCCTGGTGTGCGAC  +3 x11
CGCCCTCCCCTCCTCAACCTGCAGAACCCCCTGAAGCTG---CAAGATGTGTGACGACGAGGAGACCACCGCCCTGGTGTGCGACAAC  Δ3 x10
CGCCCTCCCCTCCTCAACCTGCAGAACCCCCTG----------AAGATGTGTGACGACGAGGAGACCACCGCCCTGGTGTGCGACAAC  Δ10 x10
```

ACTN2 Mutations in 44 of 1563 sequences ≈ 2.8%

```
GCCCGTGCGTCCGAGCCCCTCGCGCCCCGCCGCAGCCCCGGCCAACCGAGCGCCATGAACCAGATAGAGCCCGGCGTGCAGTACAACT  WT
GCCCGTGCGTCCGAGCCCCTCGCGCCCCGCCGCAGCC---------------CCATGAACCAGATAGAGCCCGGCGTGCAGTACAACT  Δ15 x4
CCCG-------------------/ /--------------------CGAGCGCCATGAACCAGATAGAGCCCGGCGTGCAGTACAACT  Δ94 x2
GCCCGTGCGTCCGAGCCCCTCGCGCCCCGCCGCAGCCCCG---------GCGCCATGAACCAGATAGAGCCCGGCGTGCAGTACAACT  Δ9 x2
GCCCGTGCGTCCGAGCCCCTCGCGCCCCGCCGCAGCC--------CCGAGCGCCATGAACCAGATAGAGCCCGGCGTGCAGTACAACT  Δ8 x2
GCCCGTGCGTCCGAGCCCCTCGCGCCCCGCCGtAGCCCCG---------GCGCCATGAACCAGATAGAGCCCGGCGTGCAGTACAACT  Δ9 (Δ10+1) x2
```

ANKRD1 Mutations in 81 of 1203 sequences ≈ 6.7%

```
ACAGAAAAACATACAAGACTCCTTCAGCCAACATGATGGTACTGAAAGTAGAGGAACTGGTATGTAAGATGCATTAATTTTATAAAAT  WT
AGAC-------------------/ /--------------------AGTAGAGGAACTGGTATGTAAGATGCATTAATTTTATAAAAT  Δ211 x9
ACAGAAAAACATACAAGACTCCTTCAGCCAACATGATG---------GTAGAGGAACTGGTATGTAAGATGCATTAATTTTATAAAAT  Δ9 x6
ACAGAAAAACATACAAGACTCCTTCAGCCAACATGATGGT---------------ACTGGTATGTAAGATGCATTAATTTTATAAAAT  Δ15 x4
ACAGAAAAACATACAAGACTCCTTCAGCCAACATGATGGTACTG-AAGTA---GAACTGGTATGTAAGATGCATTAATTTTATAAAAT  Δ4 x2
ACAGAAAAACATACAAGACTCCTTCAGCCAACATGATGGTACTG-AAGTA---GgACTGGTATGTAAGATGCATTAATTTTATAAAAT  Δ4 (Δ5 +1) x2
```

BAG3 Mutations in 290 of 3453 sequences ≈ 8.4%

```
CGGGCAGACCCCAACCCAGCATGAGCGCCGCCACCCACTCGCCCATGATGCAGGTGGCGTCCGGCAACGGTGACCGCGACCCTTTGCC  WT
CGGGCAGACCCCAACCCAGCATGAGCGCCGCCACCCACTC---------GCAGGTGGCGTCCGGCAACGGTGACCGCGACCCTTTGCC  Δ9 x70
CGGGCAGACCCCAACCCAGCATGAGCGCCGCCACCCACTC------GATGCAGGTGGCGTCCGGCAACGGTGACCGCGACCCTTTGCC  Δ6 x8
CGGGCAGACCCCAACCCAGCATGAGCGCCGCCACCCACTCGCC---GtTGCAGGTGGCGTCCGGCAACGGTGACCGCGACCCTTTGCC  Δ3 (Δ4+1) x6
CGGGCAGACCCCAACCCAGCATGAGCGCCGCCACCCACTCG-----GATGCAGGTGGCGTCCGGCAACGGTGACCGCGACCCTTTGCC  Δ5 x5
CGGGCAGACCCCAACCCAGCATGAGCGCCGCCACCCACTCGC----------aGTGGCGTCCGGCAACGGTGACCGCGACCCTTTGCC  Δ10 (Δ11+1) x4
```

CALR3 Mutations in 96 of 2778 sequences ≈ 3.5%

```
GGCGGCGACCGGAAGCGCAGTGCACACCCCCATGGCCCGGGCTTTGGTCCAGCTCTGGGCCATATGCATGCTGCGAGTGGCGCTGGCT  WT
GGCGGCGACCGGAAGCGCAGTGCACACCCCCA------------TGGTCCAGCTCTGGGCCATATGCATGCTGCGAGTGGCGCTGGCT  Δ12 x10
GGCGGCGACCGGAAGCGCAGTGCACACCCCCATGGCCCG------GGTCCAGCTCTGGGCCATATGCATGCTGCGAGTGGCGCTGGCT  Δ6 x7
GGCGGCGACCGGAAGCGCAGTGCACACCCCCATGGCCCGGG-------CCAGCTCTGGGCCATATGCATGCTGCGAGTGGCGCTGGCT  Δ7 x5
GGCGGCGACCGGAAGCGCAGTGCACACCCCCATGGCCCGG-----------GCTCTGGGCCATATGCATGCTGCGAGTGGCGCTGGCT  Δ11 x5
GGCGGCGACCGGAAGCGCAGTGCACACCCCCATGGCCC------------------GGGCCATATGCATGCTGCGAGTGGCGCTGGCT  Δ18 x5
```

CASQ2 Mutations in 13 of 5713 sequences ≈ 0.2%

```
ATTCTGCACACGGCATATTTGGGAACGAGAAACAAAAGTTTTCCCAAATGAAGAGAACTCACTTGTTTATTGTGGGGATTTATTTTCT  WT
ATTCTGCACACGGCATATTTGGGAACGAGAAAC-AAAGT---------TGAAGAGAACTCACTTGTTTATTGTGGGGATTTATTTTCT  Δ10 x2
ATTCTGCACACGGCATATTTGGGAACGAGAAACAAAAGTTTTCCCaaatgAAATGAAGAGAACTCACTTGTTTATTGT-GGGATTTAT  +4 (Δ1 +5) x1
ATTCTGCACACGGCATATTTGGGAACGAGAAACAAAAG-TTTCCCaaatgaAAATGAAGAGAACTCACTTGTTTATTGTGGGGATTTA  +5 (Δ1 +6) x1
ATTCTGCACACGGCATATTTGGGAACGAGAAACAAAAG-TTTCtCActttgTGAAGtGAACTCACTTGTTTATTGTGGGGATTTATTT  +2 (Δ5 +7) x1
ATTCTGCACACGGCATATTTGGGAACGAGAAACAAAAG--TTCCCaaatgaAAATGAAGAGAACTCACTTGTTTATTGTGGGGATTTA  +4 (Δ2 +6) x1
```

CAV3 Mutations in 465 of 4495 sequences ≈ 10.3%

```
CAGCTCGGATCTCCTCCTGTGGATCCCCCCAGCTCTGCGATGATGGCAGAAGAGCACACAGATCTCGAGGCCCAGATCGTCAAGGATA  WT
CAGCTCGGATCTCCTCCTGTGGATCCCCCCAGCTCTGCGATGAT----GcAGAGCACACAGATCTCGAGGCCCAGATCGTCAAGGATA  Δ4 (Δ5 +1) x15
CAGCTCGGATCTCCTCCTGTGGATCCCCCCAGCTCTGC---GATGGCAGAAGAGCACACAGATCTCGAGGCCCAGATCGTCAAGGATA  Δ3 x14
CAGCTCGGATCTCCTCCTGTGGAT-CCCCCAGCTCTGCGATGAT----GcAGAGCACACAGATCTCGAGGCCCAGATCGTCAAGGATA  Δ5 (Δ6 +1) x12
CAGCTCGGATCTCCTCCTGTGGATCCCCCCAGCTCT---------GCAGAAGAGCACACAGATCTCGAGGCCCAGATCGTCAAGGATA  Δ9 x8
CAGCTCGGATCTCCTCCTGTGGATCCCCCCAGCTCTGCGAT-------GAAGAGCACACAGATCTCGAGGCCCAGATCGTCAAGGATA  Δ7 x7
```

CHD7 Mutations in 7 of 2146 sequences ≈ 0.3%

```
CAGGCAAGCTCCTGAGCTGTGGTTTGGAGGAGCCGTGTGTTGGAAGAAGATGGCAGATCCAGGAATGATGAGTCTTTTTGGCGAGGAT  WT
CAGGCAAGCTCCTGAGCTGTGGTTTG-------------------GAAGATGGCAGATCCAGGAATGATGAGTCTTTTTGGCGAGGAT  Δ19 x3
CAGGCAAGCTCCTGAGCTGTGGTTTGGAGGAGCCGTGTGTTGGA----------AGATCCAGGAATGATGAGTCTTTTTGGCGAGGAT  Δ10 x2
CAGGCAAGCTCCTGAGCTGTGGTTTGGAGGAGCCGTGTG------------------TCCAGGAATGATGAGTCTTTTTGGCGAGGAT  Δ18 x1
CAGGCAAGCTCCTGAGCTGT-GTTTGGAGGAGCCGTGTGTTG---GAAGATGGCAGATCCAGGAATGATGAGTCTTTTTGGCGAGGAT  Δ4 x1
```

COX15 Mutations in 322 of 3153 sequences ≈ 10.2%

```
TGGAAGAGGTGGCTGTTCCCTGTCATCAGTATGCAGCGATTGCTCTTTCCGCCGTTGAGGGCCTTGAAGGGGAGGCAGTATCTGCCGC  WT
TGGAAGAGGTGGCTGTTCCCTGTCATCAGTATGCAGCGA---------------TTGAGGGCCTTGAAGGGGAGGCAGTATCTGCCGC  Δ15 x13
TGGAAGAGGTGGCTGTTCCCTGTCATCAGTATGCAG----------------CGTTGAGGGCCTTGAAGGGGAGGCAGTATCTGCCGC  Δ16 x5
TGGAAGAGGTGGCTGTTCCCTGTCATCAGTATGCAGCGATTG-----aggGCCGTTGAGGGCCTTGAAGGGGAGGCAGTATCTGCCGC  Δ5 (Δ8+3) x4
TGGAAGAGGTGGCTGTTCCCTGTCATCAGTATGCAGCGATT----cTTCCGCCGTTGAGGGCCTTGAAGGGGAGGCAGTATCTGCCGC  Δ4 (Δ5+1) x4
TGGAAGAGGTGGCTGTTCCCTGTCATCAGTATGCAGCGAT---TCTTTCCGCCGTTGAGGGCCTTGAAGGGGAGGCAGTATCTGCCGC  Δ3 x4
```

CRYAB Mutations in 319 of 2776 sequences ≈ 11.5%

```
CTGACCAGCCAGCTGACCCCTCACACTCACCTAGCCACCATGGACATCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTT  WT
CTGACCAGCCAGCTGACCCCTCACACTCACCTAG-------------------CCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTT  Δ19 x26
CTGACCAGCCAGCTGACCCCTCACACTCACCTA---------------GCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTT  Δ15 x21
CTGACCAGCCAGCTGACCCCTCACACTCACCTAGCCA----------------CCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTT  Δ16 x17
CTGACCAGCCAGCTGACCCCTCACACTCACCTAGCCA------------CCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTT  Δ12 x11
CTGACCAGCCAGCTGACCCCTCACAC--------------------TCGCCATCCACCACCCCTGGATCCGCCGCCCCTTCTTTCCTT  Δ20 x9
```

CSRP3 Mutations in 43 of 2679 sequences ≈ 1.6%

```
CTTTATGTCCCCTTAGACTTGACCTTGACCAGATAGTCTTCAAGATGCCAAACTGGGGCGGAGGCGCAAAATGTGGAGCCTGTGAAAA  WT
CTTTATGTCCCCTTAGACTTGACCTTGACCAGATAGTCTcC-----GCCAAACTGGGGCGGAGGCGCAAAATGTGGAGCCTGTGAAAA  Δ5 (Δ6 +1) x3
CTTTATGTCCCCTTAGACTTGACCTTGACCAGATAGTCTT--------CAAACTGGGGCGGAGGCGCAAAATGTGGAGCCTGTGAAAA  Δ8 x2
CTTTATGTCCCCTTAGACTTGACCTTGACCAGATA-------------------GGGGCGGAGGCGCAAAATGTGGAGCCTGTGAAAA  Δ19 x2
CTTTATGTCCCCTTAGACTTGACCTTGACCAGATAGTCTTCAAGAT--------------GAGGCGCAAAATGTGGAGCCTGTGAAAA  Δ14 x1
CTTTATGTCCCCTTAGACTTGACCTTGACCAGATAGTCTTCAAGActatccGCCAAACTGGGGCGGAGGCGCAAAATGTGGAGCCTGT  +5 (Δ1 +6) x1
```

CTF1 Mutations in 771 of 3914 sequences ≈ 19.7%

```
CCCCCTCGAAAGGGGGGCGTGAAGGGAGCCGGGATCAGCCAGGGGCCAGCATGAGCCGGAGGGAGGGAAGTCTGGGTAAGGGGCTGAG  WT
CCCCCTCGAAAGGGGGGCGTGAAGGGAGCCGGGATCA-------GCCAGCATGAGCCGGAGGGAGGGAAGTCTGGGTAAGGGGCTGAG  Δ7 x253
CCCCCTCGAAAGGGGGGCGTGAAGGGAGCCGGGAT-----------CAGCATGAGCCGGAGGGAGGGAAGTCTGGGTAAGGGGCTGAG  Δ11 x50
CCCCCTCGAAAGGGGGGCGTGAAGGGAGCCGGGATCAGCCA---GCCAGCATGAGCCGGAGGGAGGGAAGTCTGGGTAAGGGGCTGAG  Δ3 x42
CCCCCTCGAAA-GGGGGCGTGAAGGGAGCCGGGATCA-------GCCAGCATGAGCCGGAGGGAGGGAAGTCTGGGTAAGGGGCTGAG  Δ8 x19
CCCCCTCGAAAGGGGGGCGTGAAGGGAGCCGGGATCAGCCAGG---CAGCATGAGCCGGAGGGAGGGAAGTCTGGGTAAGGGGCTGAG  Δ3 x18
```

CTNNA3 Mutations in 230 of 2979 sequences ≈ 7.7%

```
TTATTAATAAGCATCCTTTTGTGTTTGTGCACAGGCAGCATGTCAGCTGAAACACCAATCACATTGAATATCGATCCTCAGGATCTGC  WT
TTATTAATAAGCATCCTTTTGTGTTTGTGCACAGG--------CAGCTGAAACACCAATCACATTGAATATCGATCCTCAGGATCTGC  Δ8 x38
TTATTAATAAGCATCCTTTTGTGTTTGTGCACAGGCAGCA-------TGAAACACCAATCACATTGAATATCGATCCTCAGGATCTGC  Δ7 x28
TTATTAATAAGCATCCTTTTGTGTTTGTGCACAGGCAGCATG-----TGAAACACCAATCACATTGAATATCGATCCTCAGGATCTGC  Δ5 x5
TTATTAATAAGCATCCTTTTGTGTTTGTGCACAGGCAG-----CAGCTGAAACACCAATCACATTGAATATCGATCCTCAGGATCTGC  Δ5 x5
TTATTAATAAGCATCCTTTTGTGTTTGTGCACAG-----------GCTGAAACACCAATCACATTGAATATCGATCCTCAGGATCTGC  Δ11 x5
```

DES Mutations in 1056 of 3129 sequences ≈ 33.7%

```
CCGCCAGCCTCGCCCGCGCCGTCACCATGAGCCAGGCCTACTCGTCCAGCCAGCGCGTGTCCTCCTACCGCCGCACCTTCGGCGGGGC  WT
CCGCCAGCCTCGCCCGCGCCGTCACCATGAG--------------CCAGCCAGCGCGTGTCCTCCTACCGCCGCACCTTCGGCGGGGC  Δ14 x88
CCGCCAGCCTCGCCCGCGCCGTCACCATG------------------AGCCAGCGCGTGTCCTCCTACCGCCGCACCTTCGGCGGGGC  Δ18 x70
CCGCCAGCCTCGCCCGCGCCGTCACCATGAGCCAG-------------GCCAGCGCGTGTCCTCCTACCGCCGCACCTTCGGCGGGGC  Δ13 x27
CCGCCAGCCTCGCCCGCGCCGTCACCATG------------------AGCCAGCGCGTGTCCTCCTACCGCCGCACCTTCGGCGGGG-  Δ19 x20
CCGCCAGCCTCGCCCGCGCCGTCACCATGAG--------------CCAGCCAGCGCGTGTCCTCCTACCGCCGCACCTTCGGCGGGG-  Δ15 x18
```

DMD Mutations in 4 of 1589 sequences ≈ 0.3%

```
AACTTTTACCAGGTTTTTTTTATCGCTGCCTTGATATACACTTTTCAAAATGCTTTGGTGGGAAGAAGTAGAGGACTGTTGTAAGTAC  WT
AACTTTTACCAGGTTTTTTTTATCGCTGCCTTGATATACA------------CTTTGGTGGGAAGAAGTAGAGGACTGTTGTAAGTAC  Δ12 x2
AACTTTTACCAGGTTTTTTTTATCGCTGCCTTGATATACACTT-------gGCTTTGGTGGGAAGAAGTAGAGGACTGTTGTAAGTAC  Δ7 (Δ8 +1) x1
AACTTTTACCAGGTTTTTTTTATCGCTGCCTTGATATACAC------AAATGCTTTGGTGGGAAGAAGTAGAGGACTGTTGTAAGTAC  Δ6 x1
```

DNAJC19 Mutations in 42 of 3727 sequences ≈ 1.1%

```
GGGAGCCCAGCCGGAGCCATGGTGAGTGCGGCCTTCCGGTCTTCTTGGCGACCTCCGGCCCAGGCCTCAACCTCAGCTCCCCGCCTCG  WT
GGGAGCCCAGCCGGAGCCATGGTGAGT--------------------GCGACCTCCGGCCCAGGCCTCAACCTCAGCTCCCCGCCTCG  Δ20 x6
GGGAGCCCAGCCGGAGCCATGGTGAGTGCGGCCTTCCGGTC----TGGCGACCTCCGGCCCAGGCCTCAACCTCAGCTCCCCGCCTCG  Δ4 x2
GGGAGCCCAGCCGGAGCCATGGTGAGTGCGGCCTTCCGGTCTT----GCGACCTCCGGCCCAGGCCTCAACCTCAGCTCCCCGCCTCG  Δ4 x1
GGGAGCCCAGCCGGAGCCATGGTGAGTGCGGCCTTCCGGT-------------------CCAGGCCTCAACCTCAGCTCCCCGCCTCG  Δ19 x1
GGGAGCCCAGCCGGAGCCATGGTGAGTGCGGCCTTCCGGT-------------------CCAGGCCTCAACCTCAGCT-CCCGCCTCG  Δ20 x1
```

DSC2 Mutations in 1200 of 4286 sequences ≈ 28%

```
CCCGACGCTCGGCCCGCGACCTGCCCCGAGCCCTCTCCATGGAGGCAGCCCGCCCCTCCGGCTCCTGGAACGGAGCCCTCTGCCGGCT  WT
```

```
CCCGACGCTCGGCCCGCGACCTGCCCCGAGCCCTCTCCA--------------------------TGGAACGGAGCCCTCTGCCGGCT  Δ26 x87
CCCGACGCTCGGCCCGCGACCTGCCCCGAGCCCT---------------------------CTCCTGGAACGGAGCCCTCTGCCGGCT  Δ27 x43
CCCGACGCTCGGCCCGCGACCTG-CCCGAGCCCTCTCCA--------------------------TGGAACGGAGCCCTCTGCCGGCT  Δ27 x18
CCCGACGCTCGGCCCGCGACCTGCCCCGAGCCCT---------------------CTCCGGCTCCTGGAACGGAGCCCTCTGCCGGCT  Δ21 x17
CCCGACGCTCGGCCCGCGACCTGCCCCGAGCCCTCTCCATGGAGgcaGCAGCCCGCCCCTCCGGCTCCTGGAACGGAGCCCTCTGCCG  +3 x14
```

DSG2 Mutations in 36 of 2027 sequences ≈ 1.8%

```
AGGCGGCGGCGCGGAGCGGTGCGGCGGCGGGAGGCGGAGGCGAGGGTGCGATGGCGCGGAGCCCGGGACGCGCGTACGCCCTGCTGCT  WT
AGGCGGCGGCGCGGAGCGGTGCGGCGGCGGGAGGCGGAG--------GCGATGGCGCGGAGCCCGGGACGCGCGTACGCCCTGCTGCT  Δ8 x3
AGGCGGCGGCGCGGAGCGGTGCGGCGGCGGGAGGCGGA--------------GGCGCGGAGCCCGGGACGCGCGTACGCCCTGCTGCT  Δ14 x2
AGGCGGCGGCGCGGAGCGGTGCGGCGGCGGGAG----------------------GCGGAGCCCGGGACGCGCGTACGCCCTGCTGCT  Δ22 x2
AGGCGGCGGCGCGGAGCGGTGCGGCGGCGGGAGGCGGAGGCGAGG----GATGGCGCGGAGCCCGGGACGCGCGTACGCCCTGCTGCT  Δ4 x1
AGGCGGCGGCGCGGAGCGGTGCGGCGGCGGGAGGCGGAGGCGAGGtgtacTGCGATGGCGCGGAGCCCGGGACGCGCGTACGCCCTGC  +4 (Δ1 +5) x1
AGGCGGCGGCGCGGAGCGGTGCGGCGGCGGGAGGCGGAGGCGAG---GCGATGGCGCGGAGCCCGGGACGCGCGTACGCCCTGCTGCT  Δ3 x1
```

DSP Mutations in 81 of 1668 sequences ≈ 4.9%

```
GCGCTGAGCCGCTCTCCCGATTGCCCGCCGACATGAGCTGCAACGGAGGCTCCCACCCGCGGATCAACACTCTGGGCCGCATGATCCG  WT
GCGCTGAGCCGCTCTCCCGATTGCCCGCCGACATGAGCT---------GCTCCCACCCGCGGATCAACACTCTGGGCCGCATGATCCG  Δ9 x3
GCGCTGAGCCGCTCTCCCGATTGCCCGCCGACATGA------------GCTCCCACCCGCGGATCAACACTCTGGGCCGCATGATCCG  Δ12 x3
GCGCTGAGCCGCTCTCCCGATTGCCCGCCGACAT-----------GAGGCTCCCACCCGCGGATCAACACTCTGGGCCGCATGATCCG  Δ11 x3
GCGCTGAGCCGCTCTCCCGATTGCCCGCCGACATGAGCTGCA------GCTCCCACCCGCGGATCAACACTCTGGGCCGCATGATCCG  Δ6 x2
GCGCTGAGCCGCTCTCCCGATTGCCCGCCGACATGAGCTGC-------------------------ACACTCTGGGCCGCATGATCCG  Δ25 x2
```

DTNA Mutations in 222 of 4749 sequences ≈ 4.7%

```
CCTCAATAGCGTGAGGATAATACACATTGTAACTATTTTGTCTCATAGAATGATTGAAGATAGTGGGAAAAGAGGAAATACCATGGCA  WT
CCTCAATAGCGTGAGGATAATACACATTGTAACTATTTTGTCTcataCATAGAATGATTGAAGATAGTGGGAAAAGAGGAAATACCAT  +4 x7
CCTCAATAGCGTGAGGATAATACACATTGTAACTATTTTGTCtcaTCATAGAATGATTGAAGATAGTGGGAAAAGAGGAAATACCATG  +3 x6
CCTCAATAGCGTGAGGATAATACACATTGTAACTATTTT--------GAATGATTGAAGATAGTGGGAAAAGAGGAAATACCATGGCA  Δ8 x5
CCTCAATAGCGTGAGGATAATACACATTGTAACTATTT------------TGATTGAAGATAGTGGGAAAAGAGGAAATACCATGGCA  Δ12 x5
CCTCAATAGCGTGAGGATAATACACATTGTAACTATT----------------TTGAAGATAGTGGGAAAAGAGGAAATACCATGGCA  Δ16 x4
```

EMD Mutations in 1053 of 2120 sequences ≈ 49.7%

```
GGCCCGGGCCGCCGCCAGGCCTCCGCCTGAGCCCGCACCCGCCATGGACAACTACGCAGATCTTTCGGATACCGAGCTGACCACCTTG  WT
GGCCCGGGCCGCCGCCAGGCCTCCGCCTGAGCCCGCACC---------------CGCAGATCTTTCGGATACCGAGCTGACCACCTTG  Δ15 x171
GGCCCGGGCCGCCGCCAGGCCTCCGCCTGAGCCCGC--------------ACTACGCAGATCTTTCGGATACCGAGCTGACCACCTTG  Δ14 x54
GGCCCGGGCCGCCGCCAGGCCTCCGCCTGAGCCCG-------------CAACTACGCAGATCTTTCGGATACCGAGCTGACCACCTTG  Δ13 x33
GGCCCGGGCCGCCGCCAGGCCTCCGCCTGAGCC---------------------CGCAGATCTTTCGGATACCGAGCTGACCACCTTG  Δ21 x26
GGCCCGGGCCGCCGCCAGGCCTCCGCCTGAGCCCGCACCCGC--------------CAGATCTTTCGGATACCGAGCTGACCACCTTG  Δ14 x25
```

EYA4 Mutations in 66 of 304 sequences ≈ 21.7%

```
CTTGGGAGTGGCAGGAGAAGTGAGAAAACCACATGGAAGACTCCCAGGATTTAAATGAACAATCAGTAAGTCTTCATTCTCAGTTTTG  WT
TCTT--------------------/ /--------------------GATTTAAATGAACAATCAGTAAGTCTTCATTCTCAGTTTTG  Δ197 x5
CTTGGGAGTGGCAGGAGAAGTGAGAAAACCACATGGAA---------GATTTAAATGAACAATCAGTAAGTCTTCATTCTCAGTTTTG  Δ9 x4
CTTGGGAGTGGCAGGAGAAGTGAGAAAACCAC----------------ATTTAAATGAACAATCAGTAAGTCTTCATTCTCAGTTTTG  Δ16 x4
AGAG-----------------------/ /-----------------------AATGAACAATCAGTAAGTCTTCATTCTCAGTTTTG  Δ198 x2
CTTGGGAGTGGCAGGAGAAGTGAGAAAACCACATGGAAGACT---AGGATTTAAATGAACAATCAGTAAGTCTTCATTCTCAGTTTTG  Δ3 x2
```

FHL1 Mutations in 196 of 3382 sequences ≈ 5.8%

```
TGCTTGCCCCCGCAGGTCCCTCCAGCTACAAGGTGGGCACCATGGCGGAGAAGTTTGACTGCCACTACTGCAGGGATCCCTTGCAGGG  WT
TGCTTGCCCCCGCAGGTCCCTCCAGCTACAAGGTG-----------------------/ /------------------------GGC  Δ241 x6
GGGA--------------/ /---------------GCACC----------------------------------------TGCAGGG  Δ124 x5
TGCTTGCCCCCGCAGGTCCCTCCAGCTACAAGGTGGGCACCATG-------------------/ /-------------------GCA  Δ223 x5
TGCTTGCCCCCGCAGGTCCCTCCAGCTACAAGG------------------------/ /-------------------------TGG  Δ242 x5
TGCTTGCCCCCGCAGGTCCCTCCAGCTACAA---------------GGAGAAGTTTGACTGCCACTACTGCAGGGATCCCTTGCAGGG  Δ15 x5
```

FHL2 Mutations in 168 of 4734 sequences ≈ 3.5%

```
TTCTTTTCTTTTTGATAGGTTGCTGAAAAGCCAGGAGTCAAAATGACTGAGCGCTTTGACTGCCACCATTGCAACGAATCTCTCTTTG  WT
TTCTTTTCTTTTTGATAGGTTGCTGAAAAGCCAGGAGTCAAAatgATGACTGAGCGCTTTGACTGCCACCATTGCAACGAATCTCTCT  +3 x14
TTCTTTTCTTTTTGATAGGTTGCTGAAAAGCCAGGAGTCA-----ACTGAGCGCTTTGACTGCCACCATTGCAACGAATCTCTCTTTG  Δ5 x6
TTCTTTTCTTTTTGATAGGTTGCTGAAAAGCCAGGAGT----------------------/ /----------------------CAA  Δ114 x5
TTCTTTTCTTTTTGATAGGTTGCTGAAAAGCCAGGAGTCAAA--------------------/ /--------------------GCG  Δ175 x4
TTCTTTTCTTTTTGATAGGTTGCTGAAAAGCCAGGAGTCAAaatgAATGACTGAGCGCTTTGACTGCCACCATTGCAACGAATCTCTC  +4 x4
```

FKTN Mutations in 4 of 137 sequences ≈ 2.9%

```
ATGAAAACGACTGAGATACTTTCAAAAGACAACCAAGTGAGCAGCACAGACTAATGAGTAGAATCAATAAGAACGTGGTTTTGGCCCT  WT
ATGAAAACGACTGAGATACTTTCAAAAGACAACCAAGTGAGC----CAGAC-AATGAGT--AATCAATAAGAACGTGGTTTTGG-CCT  Δ8 x1
ATGAAAACGACTGAGATACTTTCAAAAGACAACCAAGTGAG-----CAGACTAATGAGTAGAATCAATAAGAACGTGGTTTTGGCCCT  Δ5 x1
ATGAAAACGACTGAGATACTTTCAAAAGACAACCAAGTGAG---------CTAATGAGTAGAATCAATAAGAACGTGGTTTTGGCCCT  Δ9 x1
ATGAAAACGACTGAGATACTTTC-AAAGACAA-CAAG-----AGCACAGAC--ATGA-TAGAATCAATAAGAACGTGGTTTTGGCCCT  Δ10 x1
```

FXN Mutations in 156 of 2537 sequences ≈ 6.1%

```
GGCGGCAGACCCGGAGCAGCATGTGGACTCTCGGGCGCCGCGCAGTAGCCGGCCTCCTGGCGTCACCCAGCCCGGCCCAGGCCCAGAC  WT
GGCGGCAGACCCGGAGCAGCATGTGGACTCTCGGGC-----------GCCGGCCTCCTGGCGTCACCCAGCCCGGCCCAGGCCCAGAC  Δ11 x7
GGCGGCAGACCCGGAGCAGCATGTGGA--------------------------CTCCTGGCGTCACCCAGCCCGGCCCAGGCCCAGAC  Δ26 x5
GGCGGCAGACCCGGAGCAGCATGTGGACTCTCGGGCGCCGCGC---AGCCGGCCTCCTGGCGTCACCCAGCCCGGCCCAGGCCCAGAC  Δ3 x4
GGCGGCAGACCCGGAGCAGCATGTGGACTCTCGGGC---------------GCCTCCTGGCGTCACCCAGCCCGGCCCAGGCCCAGAC  Δ15 x3
GGCGGCAGACCCGGAGCAGCATGTGGACTCTCGG---------------------------GTCACCCAGCCCGGCCCAGGCCCAGAC  Δ27 x3
```

GATA4 Mutations in 14 of 2857 sequences ≈ 0.5%

```
GGGAGCTCGCAGGGACCATGTATCAGAGCTTGGCCATGGCCGCCAACCACGGGCCGCCCCCCGGTGCCTACGAGGCGGGCGGCCCCGG  WT
GGGAGCTCGCAGGGACCATGTATCAGAGCTTGGCCATGGCCGCCAACCACGGG----CCCCCGGTGCCgTACGAGGCGGGCGGCCCCG  Δ3 (Δ4 +1) x1
GGGAGCTCGCAGGGACCATGTATCAGAGCTTGGCCATGGCCGCCAA----------------------------------GGCCCCgG  Δ33 (Δ34 +1) x1
GGGAGCTCGCAGGGACCATGTATCAGAGCTTGGCCATGGCCGCCA--CACG----GCCCCCCGGTGCCTACGAGGCGGGCGGCCCCGG  Δ6 x1
GGGAGCTCGCAGGGACCATGTATCAGAGCTTGGCCATGGCCGCCAaccACCACGGGCCGCCCCCCGGTGCCTACGAGGCGGGCGGCCC  +3 x1
GGGAGCTCGCAGGGACCATGTATCAGAGCTTGGCCATGGCCGCcaaCAACCACGGGCCGCCCCCCGGTGCCTACGAGGCGGGCGGCCC  +3 x1
```

GATAD1 Mutations in 217 of 3832 sequences ≈ 5.7%

```
CCGTCCGCCATTCCCGTGTCTCTGCGCCCGCGGGGGCCGCCCGAGCCGGCCACCATGCCGCTGGGCCTGAAGCCCACCTGCAGCGTAT  WT
CCGTCCGCCATTCCCGTGTCTCTGCGCCCGCGG-----------------------------GGGCCTGAAGCCCACCTGCAGCGTAT  Δ29 x14
CCGTCCGCCATTCCCGTGTCTCTGCGCCCGCGGGGGCCGC-----------------CCGCTGGGCCTGAAGCCCACCTGCAGCGTAT  Δ17 x8
CCGTCCGCCATTCCCGTGTCTCTGCGCCCGCG------------------------------GGGCCTGAAGCCCACCTGCAGCGTAT  Δ30 x6
CCGTCCGCCATTCCCGTGTCTCTGCGCCCGCGGGGGCCGCCCGA---GCCACCATGCCGCTGGGCCTGAAGCCCACCTGCAGCGTAT  Δ3 x5
CCGTCCGCCATTCCCGTGTCTCTGCGCCCGCGGG-------------GGCCACCATGCCGCTGGGCCTGAAGCCCACCTGCAGCGTAT  Δ13 x5
```

GLA Mutations in 553 of 4338 sequences ≈ 12.7%

```
CTGAGGAACCCAGAACTACATCTGGGCTGCGCGCTTGCGCTTCGCTTCCTGGCCCTCGTTTCCTGGGACATCCCTGGGGCTAGAGCAC  WT
CTGAGGAACCCAGAACTACATCTGGGCTGCGCGCTTG-----CGCTTCCTGGCCCTCGTTTCCTGGGACATCCCTGGGGCTAGAGCAC  Δ5 x41
CTGAGGAACCCAGAACTACAT---------------------------CTGGCCCTCGTTTCCTGGGACATCCCTGGGGCTAGAGCAC  Δ27 x18
CTGAGGAACCCAGAACTACATCTGGGCTGCG-----------CGCTTCCTGGCCCTCGTTTCCTGGGACATCCCTGGGGCTAGAGCAC  Δ11 x14
CTGAGGAACCCAGAACTACATCTGGG----------------------CTGGCCCTCGTTTCCTGGGACATCCCTGGGGCTAGAGCAC  Δ22 x11
CTGAGGAACCCAGAACTACATCTGGGCTGCGCGCTTGCG---CGCTTCCTGGCCCTCGTTTCCTGGGACATCCCTGGGGCTAGAGCAC  Δ3 x6
```

HOPX Mutations in 75 of 2504 sequences ≈ 3%

```
CACCGCCGCCGCTTCTCCCTGCCCCGCAGCGCGCAGGGACCATGTCGGCGGAGACCGCGAGCGGCCCCACAGAGGACCAGGTGGAAAT  WT
CACCGCCGCCGCTTCTCCCTGCCCCGCAGCGCGCAGG---------------GACCGCGAGCGGCCCCACAGAGGACCAGGTGGAAAT  Δ15 x5
CACCGCCGCCGCTTCTCCCTGCCCCGCAGCGC---------------GCGGAGACCGCGAGCGGCCCCACAGAGGACCAGGTGGAAAT  Δ15 x3
CACCGCCGCCGCTTCTCCCTGCCCCGCAGCGCGCAGGGACCATGT-------GACCGCGAGCGGCCCCACAGAGGACCAGGTGGAAAT  Δ7 x2
CACCGCCGCCGCTTCTCCCTGCCCCGCAG--------------------------------------CACAGAGGACCAGGTGGAAAT  Δ38 x2
CACCG-------------------------------------------------CCGCGAGCGGCCCCACAGAGGACCAGGTGGAAAT  Δ49 x2
```

ILK Mutations in 193 of 3547 sequences ≈ 5.4%

```
GGCTTCCCCAATCCAGGGGACTCGGCGCCGGGACGCTGCTATGGACGACATTTTCACTCAGTGCCGGGAGGGCAACGCAGTCGCCGTT  WT
GGCTTCCCCAATCCAGGGGACTCGGCGCCGGGACGCTGCTATG---GACATTTTCACTCAGTGCCGGGAGGGCAACGCAGTCGCCGTT  Δ3 x28
GGCTTCCCCAATCCAGGGGACTCGGCGCCGG---------------GACATTTTCACTCAGTGCCGGGAGGGCAACGCAGTCGCCGTT  Δ15 x7
GGCTTCCCCAATCCAGGGGACTCGGCGCCGGGACGCTGCTAT----GACATTTTCACTCAGTGCCGGGAGGGCAACGCAGTCGCCGTT  Δ4 x6
GGCTTCCCCAATCCAGGGGACTCGGCGCCGGGACGCTGCT---------ATTTTCACTCAGTGCCGGGAGGGCAACGCAGTCGCCGTT  Δ9 x6
GGCTTCCCCAATCCAGGGGACTCGGCGCCGGGACGCTGCTATGgacGACGACATTTTCACTCAGTGCCGGGAGGGCAACGCAGTCGCC  +3 x4
```

JAG1 Mutations in 125 of 2479 sequences ≈ 5%

```
CCCCACGGACGCGCGGCCGGTCCGGGCGCCCCCTAAGCCTCCTGCTCGCCCTGCTCTGTGCCCTGCGAGCCAAGGTAGGAGCCCTTCT  WT
CCCCACGGACGCGCGGCCGGTCCGGGCGCCCCCTAA-----------GCCCTGCTCTGTGCCCTGCGAGCCAAGGTAGGAGCCCTTCT  Δ11 x8
CCCCACGGACGCGCGGCCGGTCCGGGCGCCCCCTAAGCCTCctgCTGCTCGCCCTGCTCTGTGCCCTGCGAGCCAAGGTAGGAGCCCT  +3 x7
CCCCACGGACGCGCGGCCGGTCCGGGCGCCCCCTAAGCCTCCTG------CTGCTCTGTGCCCTGCGAGCCAAGGTAGGAGCCCTTCT  Δ6 x3
CCCCACGGACGCGCGGCCGGTCCGGGCGCCCCCTAAGCCT---------CCTGCTCTGTGCCCTGCGAGCCAAGGTAGGAGCCCTTCT  Δ9 x31
CCCCACGGACGCGCGGCCGGTCCGGGCGCCCCCTAAGCCTCCTGCTC-----GCTCTGTGCCCTGCGAGCCAAGGTAGGAGCCCTTCT  Δ5 x2
```

JPH2 Mutations in 41 of 2629 sequences ≈ 1.6%

```
ACGCTGGAGGACGGGGAGGTTGTCAGGGGCTATGATGAGATGAGTGGGGGCCGCTTCGACTTTGATGATGGAGGGGCGTACTGCGGGG  WT
ACGCTGGAGGACGGGGAGGTTGTCAGGGGCTATGATGAGATGAGT-----CCGCTTCGACTTTGATGATGGAGGGGCGTACTGCGGGG  Δ5 x2
ACGCTGGAGGACGGGGAGGTTGTC----------------------------------ACTTTGATGATGGAGGGGCGTACTGCGGGG  Δ34 x2
ACGCTGGAGGACGGGGAGGTTGTCAGGGGCTATGATGAGATGAGTGGG---CGCTTCGACTTTGATGATGGAGGGGCGTACTGCGGGG  Δ3 x1
ACGCTGGAGGACGGGGAGGTTGTCAGGGGCTATGATGAGATGAGTGG----CGCTTCGACTTTGATGATGGAGGGGCGTACTGCGGGG  Δ4 x1
ACGCTGGAGGACGGGGAGGTTGTCAGGGGCTATGATGAGATGAGTG---------cTCGACTTTGATGATGGAGGGGCGTACTGCGGGG  Δ8 (Δ9 +1) x1
```

JUP Mutations in 34 of 4100 sequences ≈ 0.8%

```
TTCCTGCTTCCTGACTTCCTCCTTTGTGCCCCCAGTAGCCACGATGGAGGTGATGAACCTGATGGAGCAGCCTATCAAGGTGACTGAG  WT
TTCCTGCTTCCTGACTTCCTCCT------------------------------TGAACCTGATGGAGCAGCCTATCAAGGTGACTGAG  Δ30 x3
TTCCTGCTTCCTGACTTCCTCCTTTGTGCCCCCAGTAGCCAtC-----AGGTGATGAACCTGATGGAGCAGCCTATCAAGGTGACTGA  Δ4 (Δ5+1) x2
TTCCTGCTTCCTGACTTCCTCCTTTGTGCCCCCAGTAGCCacgACGATGGAGGTGATGAACCTGATGGAGCAGCCTATCAAGGTGACT  +3 x2
TTCCTGCTTCCTGACTTCCTCCTTTGTGCCCCCA-----------GGAGGTGATGAACCTGATGGAGCAGCCTATCAAGGTGACTGAG  Δ11 x2
TTCCTGCTTCCTGACTTCCTCCTTTGTGCCCCCAGTAGCCACGATG--------GAACCTGATGGAGCAGCCTATCAAGGTGACTGAG  Δ8 x1
```

LAMA4 Mutations in 57 of 1263 sequences ≈ 4.5%

```
GATGTCAGCGGAGAAATGGCTTTGAGCTCAGCCTGGCGCTCGGTTCTGCCTCTGTGGCTCCTCTGGAGCGCTGCCTGCTCCCGCGCCG  WT
GATGTCAGCGGAGAAATGGCTTTGAGCTCAGC-------------CTGCCTCTGTGGCTCCTCTGGAGCGCTGCCTGCTCCCGCGCCG  Δ13 x4
GATGTCAGCGGAGAAATGGCTTTGAGCTCAGCCTGGCGCTCGGT------TCTGTGGCTCCTCTGGAGCGCTGCCTGCTCCCGCGCCG  Δ6 x3
GATGTCAGCGGAGAAATGGCTTTGAGCTCAGCCTGGCGC-----TCTGCCTCTGTGGCTCCTCTGGAGCGCTGCCTGCTCCCGCGCCG  Δ5 x2
GATGTCAGCGGAGAAATGGCTTTGAGCTCAGCCTGGCG-------CTGCCTCTGTGGCTCCTCTGGAGCGCTGCCTGCTCCCGCGCCG  Δ7 x2
CACG--------------------/ /------------------CTGCCTCTGTGGCTCCTCTGGAGCGCTGCCTGCTCCCGCGCCG  Δ96 x1
```

LAMP2 Mutations in 254 of 2092 sequences ≈ 12.1%

```
TCGCCGCCGTCGCCGCCTGCTCTGCGGGGTCATGGTGTGCTTCCGCCTCTTCCCGGTTCCGGGCTCAGGGCTCGTTCTGGTCTGCCTA  WT
TCGCCGCCGTCGCCGCCTGCTCTGCGGGGTCATGGTGTG-------CTCTTCCCGGTTCCGGGCTCAGGGCTCGTTCTGGTCTGCCTA  Δ7 x26
TCGCCGCCGTCGCCGCCTGCTCTGCGGGGTCATGGTGTG---------CTTCCCGGTTCCGGGCTCAGGGCTCGTTCTGGTCTGCCTA  Δ9 x18
TCGCCGCCGTCGCCGCCTGCTCTGCGGGGTCATGGTGT------------------------GCTCAGGGCTCGTTCTGGTCTGCCTA  Δ24 x14
TCGCCGCCGTCGCCGCCTGCTCTGCGGGGTCATGGTGT------GCCTCTTCCCGGTTCCGGGCTCAGGGCTCGTTCTGGTCTGCCTA  Δ6 x4
```

```
TCGCCGCCGTCGCCGCCTGCTCTGCGGGGTCATGGTGTGCTT---------------------CTCAGGGCTCGTTCTGGTCTGCCTA  Δ21 x3
```

LMNA Mutations in 183 of 1459 sequences ≈ 12.5%

```
CGCTGCCAACCTGCCGGCCATGGAGACCCCGTCCCAGCGGCGCGCCACCCGCAGCGGGGCGCAGGCCAGCTCCACTCCGCTGTCGCCC  WT
CGCTGCCAACCTGCCGGCCATGGAGACCCCGTC--------------------------------CCAGCTCCACTCCGCTGTCGCCC  Δ32 x18
CGCTGCCAACCTGCCGGCCATGGAGACCCCGTCCCAGC-------------------GGCGCAGGCCAGCTCCACTCCGCTGTCGCCC  Δ19 x10
CGCTGCCAACCTGCCGGCCATGGAGACCCCGTCCCAGCGGCGcgcCGCCACCCGCAGCGGGGCGCAGGCCAGCTCCACTCCGCTGTCG  +3 x7
CGCTGCCAACCTGCCGGCCATGGAGACCCCGTCC-----------------CAGCGGGGCGCAGGCCAGCTCCACTCCGCTGTCGCCC  Δ17 x6
CGCTGCCAACCTGCCGGCCATGGAGACCCCGTCCCAGCGGCGcgccCGCCACCCGCAGCGGGGCGCAGGCCAGCTCCACTCCGCTGTC  +4 x5
```

MLYCD Mutations in 611 of 2282 sequences ≈ 26.8%

```
AGCGGCGGCGGCGCTCCCCCTCGGCAGCTGTTGTGGGGCACCATGCGAGGCTTCGGGCCAGGCTTGACGGCCAGGCGTCTCCTCCCGC  WT
AGCGGCGGCGGCGCTCCCCCTCGGCAGCTGTTGTGG------------GGCTTCGGGCCAGGCTTGACGGCCAGGCGTCTCCTCCCGC  Δ12 x50
AGCGGCGGCGGCGCTCCCCCTCGGCAGCTGTTGT------------GAGGCTTCGGGCCAGGCTTGACGGCCAGGCGTCTCCTCCCGC  Δ12 x31
AGCGGCGGCGGCGCTCCCCCTCGGCAGCTGTTGTG-------------------GGGCCAGGCTTGACGGCCAGGCGTCTCCTCCCGC  Δ19 x19
AGCGGCGGCGGCGCTCCCCCTCGGCAGCTGT------------TGCGAGGCTTCGGGCCAGGCTTGACGGCCAGGCGTCTCCTCCCGC  Δ12 x14
AGCGGCGGCGGCGCTCCCCCTCGGCAGCTGTTGT--------------------GGGCCAGGCTTGACGGCCAGGCGTCTCCTCCCGC  Δ20 x13
```

MYBPC3 Mutations in 13 of 1895 sequences ≈ 0.7%

```
TGGGTGACCTGTGCCTGCTTCGTGCCTGGTGTGACGTCTCTCAGGATGCCTGAGCCGGGGAAGAAGCCAGGTAGCTTTAGGACTGGGG  WT
TGGGTGACCTGTGCCTGCTTCGTGCCTGGTG---------------TGCCTGAGCCGGGGAAGAAGCCAGGTAGCTTTAGGACTGGGG  Δ15 x3
TGGGTGACCTGTGCCTGCTTCGTGCCTGGTGT------------GATGCCTGAGCCGGGGAAGAAGCCAGGTAGCTTTAGGACTGGGG  Δ12 x2
TGGGTGACCTGTGCCTGCTTCGTGCCTGGTGTGACGTCTCTCaggAGGATGCCTGAGCCGGGGAAGAAGCCAGGTAGCTTTAGGACTG  +3 x1
TGGGTGACCTGTGCCTGCTTCGTGCCTGGTGTGACGTC--------TGCCTGAGCCGGGGAAGAAGCCAGGTAGCTTTAGGACTGGGG  Δ8 x1
TGGGTGACCTGTGCCTGCTTCGTGCCTGGTGTGACGT---TCAGGAT-CCTGAGCCGGGGAAGAAGCCAGGTAGCTTTAGGACTGGGG  Δ4 x1
```

MYH6 Mutations in 172 of 1845 sequences ≈ 9.3%

```
GGAGTAACATAGCCCTCCTGTCTCTGACCCAGGGGAAGCACCAAGATGACCGATGCCCAGATGGCTGACTTTGGGGCAGCGGCCCAGT  WT
GGAGTAACATAGCCCTCCTGTCTCTGACCCAGGGGAAGC---------ACCGATGCCCAGATGGCTGACTTTGGGGCAGCGGCCCAGT  Δ9 x20
AGCC----------------------------------------/ /----------------------------------------CTG  Δ262 x9
GGAGTAACATAGCCCTCCTGTCTCTGACCCA-GGGAAGC---------ACCGATGCCCAGATGGCTGACTTTGGGGCAGCGGCCCAGT  Δ10 x5
GGAGTAACATAGCCCTCCTGTCTC----------------------TGACCGATGCCCAGATGGCTGACTTTGGGGCAGCGGCCCAGT  Δ22 x5
GCCA--------------------/ /---------------------ACCGATGCCCAGATGGCTGACTTTGGGGCAGCGGCCCAGT  Δ255 x4
```

MYH7 Mutations in 176 of 2894 sequences ≈ 6.1%

```
CCAGGCACAGCCATGGGAGATTCGGAGATGGCAGTCTTTGGGGCTGCCGCCCCCTACCTGCGCAAGTCAGAGAAGGAGCGGCTAGAAG  WT
CCAGGCACAGCCATGGGAGATTCGGAGATGGCAGTCTTTGGGGCTGC-------------CGCAAGTCAGAGAAGGAGCGGCTAGAAG  Δ13 x49
CCAGGCACAGCCATGGGAGATTCGGAGATGGCAGTCTTT-GGGCTGC-------------CGCAAGTCAGAGAAGGAGCGGCTAGAAG  Δ14 x10
CCAGGCACAGCCATGGGAGATTCGGAGATGGCAGTCTTTGGGGCTG----------CCTGCGCAAGTCAGAGAAGGAGCGGCTAGAAG  Δ10 x5
CCAGGCACAGCCATGGGAGATTCGGAGATGGCAGTCTT-------------CCCTACCTGCGCAAGTCAGAGAAGGAGCGGCTAGAAG  Δ13 x4
CCAGGCACAGCCATGGGAGATTCGGAGATGGCAGTCTTTGGGGCTG-----CgCTA-CTGCGCAAGTCAGAGAAGGAGCGGCTAGAAG  Δ6 (Δ7 +1) x3
```

MYH7 Mutations in 1143 of 2276 sequences ≈ 50.2%

```
CAGCCATGGGAGATTCGGAGATGGCAGTCTTTGGGGCTGCCGCCCCCTACCTGCGCAAGTCAGAGAAGGAGCGGCTAGAAGCGCAGAC  WT
CAGCCATGGGAGATTCGGAGATGGCAGTCTTTGGGGCTGC-------------CGCAAGTCAGAGAAGGAGCGGCTAGAAGCGCAGAC  Δ13 x488
CAGCCATGGGAGATTCGGAGATGGCAGTCTTT-GGGCTGC-------------CGCAAGTCAGAGAAGGAGCGGCTAGAAGCGCAGAC  Δ14 x137
CAGCCATGGGAGATTCGGAGATGGCAGTCTT-------------CCCTACCTGCGCAAGTCAGAGAAGGAGCGGCTAGAAGCGCAGAC  Δ13 x115
CAGCCATGGGAGATTCGGAGATGGCAGTCTTTGGGGCTG----------CCTGCGCAAGTCAGAGAAGGAGCGGCTAGAAGCGCAGAC  Δ10 x15
CAGCCATGGGAGATTCGGAGATGGCAGTCTTTGGGG--------------CTGCGCAAGTCAGAGAAGGAGCGGCTAGAAGCGCAGAC  Δ14 x10
```

MYL2 Mutations in 20 of 3949 sequences ≈ 0.5%

```
AATTCTTCTCGGGAGGCAGTGCTGGGTCCTTTCCACCATGGTGAGTACAAGGGCTCCAGGAGGTGATGATGCCGGGTGGGCGAGGAGA  WT
AATTCTTCTCGGGAGGCAGTGCTGGGTCCTTTCCACCA---TGtGTACAAGGGCTCCAGGAGGTGATGATGCCGGGTGGGCGAGGAGA  Δ3 (Δ4 +1) x2
AATTCTTCTCGGGAGGCAGTGCTGGGTCCTTTCCACCATGGTGAGT---AaGGCTCCAGGAGGTGATGATGCCGGGTGGGCGAGGAGA  Δ3 (Δ4 +1) x1
AATTCTTCTCGGGAGGCAGTGCTGGGTCCTTTCCACCATGGTGAGTgaaagacACAAGGGCTCCAGGAGGTGATGATGCCGGGTGGGC  +7 x1
AATTCTTCTCGGGAGGCAGTGCTGGGTCCTTTCCACCATGGTGAGTggaaaggaccACAAGGGCTCCAGGAGGTGATGATGCCGGGTG  +10 x1
AATTCTTCTCGGGAGGCAGTGCTGGGTCCTTTCCACCATGGTGagtAGTACAAGGGCTCCAGGAGGTGATGATGCCGGGTGGGCGAGG  +3 x1
```

MYL3 Mutations in 171 of 2199 sequences ≈ 7.8%

```
TTCTCTCCACATCCCTCTCTGTACTTACAGCCCCCAATGGCCCCCAAAAAGCCAGAGCCCAAGAAGGATGATGCCAAGGCAGCCCCCA  WT
TTCTCTCCACATCCCTCTCTGTACTTACAGCCCCCAATG-----------GCCAGAGCCCAAGAAGGATGATGCCAAGGCAGCCCCCA  Δ11 x18
TTCTCTCCACATCCCTCTCTGTACTTACAG-CCCCAATG-----------GCCAGAGCCCAAGAAGGATGATGCCAAGGCAGCCCCCA  Δ12 x5
TTCTCTCCACATCCCTCTCTGTACTTACAGCCCCCAATG-----------GCCAGAGCCCAAGAAGGATGATGCCAAGGCAG-CCCCA  Δ12 x4
TTCTCTCCACATCCCTCTCTGTACTTACA----------GCCCCCAAAAAGCCAGAGCCCAAGAAGGATGATGCCAAGGCAGCCCCCA  Δ10 x4
TTCTCTCCACATCCCTCTCTGTACTTACAGCCCCCAATGGCCC---AAAAAGCCAGAGCCCAAGAAGGATGATGCCAAGGCAGCCCCCA  Δ3 x2
```

MYLK2 Mutations in 254 of 4248 sequences ≈ 6%

```
ACAAGCAGCAGCACACGCCTCCCTACCTCATGGCGACAGAAAATGGAGCAGTTGAGCTGGGAATTCAGAACCCATCAACAGGTGCCAA  WT
ACAAGCAGCAGCACACGCCTCCCTACCTCATGGCGAC---------AGCAGTTGAGCTGGGAATTCAGAACCCATCAACAGGTGCCAA  Δ9 x22
ACAAGCAGCAGCACACGCCTCCCTACCTCATGGCGA------------CAGTTGAGCTGGGAATTCAGAACCCATCAACAGGTGCCAA  Δ12 x12
ACAAGCAGCAGCACACGCCTCCCTACCTCATGGCGACA-------GAGCAGTTGAGCTGGGAATTCAGAACCCATCAACAGGTGCCAA  Δ7 x10
ACAAGCAGCAGCACACGCCTCCCTACCTCATGGCGACAGAAAAtggTGGAGCAGTTGAGCTGGGAATTCAGAACCCATCAACAGGTGC  +3 x5
ACAAGCAGCAGCACACGCCTCCCTACCTCATGGCGACAGAA--------AGTTGAGCTGGGAATTCAGAACCCATCAACAGGTGCCAA  Δ8 x5
ACAAGCAGCAGCACACGCCTCCCTACCTCATGGCGACA---------------GAGCTGGGAATTCAGAACCCATCAACAGGTGCCAA  Δ15 x5
```

MYOM1 Mutations in 35 of 4149 sequences ≈ 0.8%

```
TTCCTTCAGGTGGCCCGGTTCCTTCAAGGGGCACAGGATGTCTTTGCCTTTTTATCAGAGGTGCCACCAGCACTATGATCTCAGCTAC  WT
TTCCTTCAGGTGGCCCGGTTCCTTCAAGaGGCACAGGATGTCTTTG-----TTATCAGAGGTGCCACCAGCACTATGATCTCAGCTAC  Δ5 (Δ6 +1) x2
```

```
TTCCTTCAGGTGGCCCGGTTCCTTCAAGaGGCACAGGATGTCTttgTTGCCTTTTTATCAGAGGTGCCACCAGCACTATGATCTCAGC  +3 (Δ1 +4) x2
TTCCTTCAGGTGGCCCGGTTCCTTCAAGGGGCACAGGATGTCTTTGC---TTTA------GTG----CAcCACTATGATCTCAGCTAC  Δ13 (Δ14 +1) x1
TTCCTTCAGGTGGCCCGGTTCCTTCAAGGGGCACAGGATGTCTTTGC-----TATCAGAGGTGCCACCAGCACTATGATCTCAGCTAC  Δ5 x1
TTCCTTCAGGTGGCCCGGTTCCTTCAAGGGGCACAGGATGTCTttgTTGCCTTTTTATCAGAGGTGCCACCAGCACTATGATCTCAGC  +3 x1
```

MYOZ2 Mutations in 67 of 4653 sequences ≈ 1.4%

```
AAAAAAACCATGCTATCACATAATACTATGATGAAGCAGAGAAAACAGCAAGCAACAGCCATCATGAAGGAAGTCCATGGAAATGGTA  WT
AAAAAAACCATGCTATCACATAATACTATGATGAAGCAG-----------AGCAACAGCCATCATGAAGGAAGTCCATGGAAATGGTA  Δ11 x10
AAAAAAACCATGCTATCACATAATACTATGATGAAGCAG-------AGCAAGCAACAGCCATCATGAAGGAAGTCCATGGAAATGGTA  Δ7 x4
AAAAAAACCATGCTATCACATAATACTATGATG----------------AAGCAACAGCCATCATGAAGGAAGTCCATGGAAATGGTA  Δ16 x4
AAAAAAACCATGCTATCACATAATACTATGATGAAGCAGAGA----AGCAAGCAACAGCCATCATGAAGGAAGTCCATGGAAATGGTA  Δ4 x2
AAAAAAACCATGCTATCACATAATACTATGATGAAGCAGAGA-------cAGCAACAGCCATCATGAAGGAAGTCCATGGAAATGGTA  Δ7 (Δ8+1) x2
```

MYPN Mutations in 345 of 1389 sequences ≈ 24.8%

```
AAACTTTTTGTTATTATTATTTTGTGACAGCATGCAAGACGACAGCATAGAAGCTTCTACTTCCATATCTCAGCTTCTAAGAGAGAGC  WT
AAACTTTTTGTTATTATTATTTTGTGACAGCATGCA-----------------------/ /-----------------------AGA  Δ210 x12
AAACTTTTTGTTATTATTATTTTGTGACAGCAT-----------GCATAGAAGCTTCTACTTCCATATCTCAGCTTCTAAGAGAGAGC  Δ11 x11
AAACTTTTTGTTATTATTATTTTGTGACAGCATGCAA---GACAGCATAGAAGCTTCTACTTCCATATCTCAGCTTCTAAGAGAGAGC  Δ3 x9
AAACTTTTTGTTATTATTATTTTGTGACAGCATGCA----GACAGCATAGAAGCTTCTACTTCCATATCTCAGCTTCTAAGAGAGAGC  Δ4 x5
AAACTTTTTGTTATTATTATTTTGTGACAGCATGCAAGACGAC--------AGCTTCTACTTCCATATCTCAGCTTCTAAGAGAGAGC  Δ8 x4
```

NEBL Mutations in 44 of 3474 sequences ≈ 1.3%

```
AATATTTTAAAGGGTAAAAATGAGGGTCCCTGTATTTGAGGATATAAAAGATGAAACTGAAGAAGAAAAGATAGGGGAAGAAGAAAAT  WT
AATATTTTAAAGGGTAAAAATGAGGGTCCCTGTATTTGAGGAtatTATAAAAGATGAAACTGAAGAAGAAAAGATAGGGGAAGAAGAA  +3 x2
AATATTTTAAAGGGTAAAAATGAGGGTCCCTGTATTTGAGG-----AAAGATGAAACTGAAGAAGAAAAGATAGGGGAAGAAGAAAAT  Δ5 x2
AATATTTTAAAGGGTAAAAATGAGGGTCCCTGT---------TATAAAAGATGAAACTGAAGAAGAAAAGATAGGGGAAGAAGAAAAT  Δ9 x2
AATATTTTAAAGGGTAAAAATGAGGGTCCCTGT----------------------ACTGAAGAAGAAAAGATAGGGGAAGAAGAAAAT  Δ22 x2
AATATTTTAAAGGGTAAAAATGAGGGTCCC---------------------TGAAACTGAAGAAGAAAAGATAGGGGAAGAAGAAAAT  Δ21 x2
```

NEXN Mutations in 179 of 1508 sequences ≈ 11.9%

```
ATAATCAGCCCAAGACCACATAGAGCAAACATGAATGATATTTCCCAAAAGGCTGAGGTAAGTCTCAAAAGTAAAAATAAAAATAAAA  WT
ATAATCAGCCCAAGACCACATAGAGCAAA----------------CAAAAGGCTGAGGTAAGTCTCAAAAGTAAAAATAAAAATAAAA  Δ16 x9
ATAATCAGCCCAAGACCACATAGAGCAAACATGAA------------------TGAGGTAAGTCTCAAAAGTAAAAATAAAAATAAAA  Δ18 x8
ATAATCAGCCCAAGACCACATAGAGCAAACA----------------------TGAGGTAAGTCTCAAAAGTAAAAATAAAAATAAAA  Δ22 x4
ATAATCAGCCCAAGACCACATAGAGCAAACATGAATGATATTTCC-----GGCTGAGGTAAGTCTCAAAAGTAAAAATAAAAATAAAA  Δ5 x3
ATAATCAGCCCAAGACCACATAGAGCAAACATGAATGATATT---CAAAAGGCTGAGGTAAGTCTCAAAAGTAAAAATAAAAATAAAA  Δ3 x3
```

NKX2-5 Mutations in 132 of 1404 sequences ≈ 9.4%

```
CTGCCGCCCACCTGGCGCTGTGAGACTGGCGCTGCCACCATGTTCCCCAGCCCTGCTCTCACGCCCACGCCCTTCTCAGTCAAAGACA  WT
CTGCCGCCCACCTGGCGCTGTGAGACTGGCGCTG------------CCAGCCCTGCTCTCACGCCCACGCCCTTCTCAGTCAAAGACA  Δ12 x12
CTGCCGCCCACCTGGCGCTGTGAGACTGGCGCTGCCA---------CCAGCCCTGCTCTCACGCCCACGCCCTTCTCAGTCAAAGACA  Δ9 x7
CTGCCGCCCACCTGGCGCTGTGAGACTGGCGCTGCCACCATGTT-----------CTCTCACGCCCACGCCCTTCTCAGTCAAAGACA  Δ11 x3
CTGCCGCCCACCTGGCGCTGTGAGACTGGCGCTGCCACCATG----CCAGCCCTGCTCTCACGCCCACGCCCTTCTCAGTCAAAGACA  Δ4 x3
CTGCCGCCCACCTGGCGCTGTGAGACTGGCGCTGCCACCATG---CCCAGCCCTGCTCTCACGCCCACGCCCTTCTCAGTCAAAGACA  Δ3 x2
```

PDLIM3 Mutations in 29 of 2445 sequences ≈ 1.2%

```
GGCTGCCCTGCGCGGGGACACTCAGAGCCCGGTGGGCGGGAGGAAGGCGGCATGCCCCAGACGGTGATCCTCCCGGGCCCTGCGCCCT  WT
GGAC----------------------------------/ /-------------------------------------GCCCTGCGCCCT  Δ93 x3
GGCTGCCC---------------------------------------------------------TGATCCTCCCGGGCCCTGCGCCCT  Δ56 x3
GGCTGCCCTGCGCGGGGACACTCAGAGCCCGGTGGGC------AAGGCGGCATGCCCCAGACGGTGATCCTCCCGGGCCCTGCGCCCT  Δ6 x2
GGCTGCCCTGCGCG-----------------------------------------------GGTGATCCTCCCGGGCCCTGCGCCCT  Δ48 x2
ACGC-----------------------------/ /----------------------------CGGTGATCCTCCCGGGCCCTGCGCCCT  Δ76 x1
```

PKP2 Mutations in 132 of 2979 sequences ≈ 4.4%

```
CCAGAGGCAGGCGAGCAGCTCGGTCGCCCCCACCGGCCCCATGGCAGCCCCCGGCGCCCCAGCTGAGTACGGCTACATCCGGACCGTC  WT
CCAGAGGCAGGCGAGCAGCTCGGTCGCCCCCACCG-----------GCCCCCGGCGCCCCAGCTGAGTACGGCTACATCCGGACCGTC  Δ11 x12
CCAGAGGCAGGCGAGCAGCTCGGTCGCCCCCA------------------CCGGCGCCCCAGCTGAGTACGGCTACATCCGGACCGTC  Δ18 x10
CCAGAGGCAGGCGAGCAGCTCGGTCGCCCCCACCG--------------------GCCCCAGCTGAGTACGGCTACATCCGGACCGTC  Δ20 x8
CCAGAGGCAGGCGAGCAGCTCGGTCGCCCCCACCGG------------CCCCGGCGCCCCAGCTGAGTACGGCTACATCCGGACCGTC  Δ12 x7
CCAGAGGCAGGCGAGCAGCTCGGTCGCC-------------------CCCCCGGCGCCCCAGCTGAGTACGGCTACATCCGGACCGTC  Δ19 x5
```

PLN Mutations in 8 of 1097 sequences ≈ 0.7%

```
GACCACTTAAAACTTCAGACTTCCTGTCCTGCTGGTATCATGGAGAAAGTCCAATACCTCACTCGCTCAGCTATAAGAAGAGCCTCAA  WT
GACCACTTAAAACTTCAGACTTCCTGTCCTGCTGGTATCATGGctggtatcatAG-gAGTCCAATACCTCACTCGCTCAGCTATAAGA  +9 (Δ2 +11) x1
GACCACTTAAAACTTCAGACTTCCTGTCCTGCTGGTATCATG----AAaTCC-ATACCTCACTCGCTCAGCTATAAGAAGAGCCTCAA  Δ5 (Δ6 +1) x1
GACCACTTAAAACTTCAGACTTCCTGTCCTGCTGGTATCATG-----AGTCCAATACCTCACTCGCTCAGCTATAAGAAGAGCCTCAA  Δ5 x1
GACCACTTAAAACTTCAGACTTCCTGTCCTGCTGGTATCAT---GAAAGTCCAATACCTCACTCGCTCAGCTATAAGAAGAGCCTCAA  Δ3 x1
GACCACTTAAAACTTCAGACTTCCTGTCCTGCT-GTA-CATG--------------CaTCACTCGCTCAGCTATAAGAAGAGCCTCAA  Δ16 (Δ17 +1) x1
```

PRKAG2 Mutations in 117 of 2433 sequences ≈ 4.8%

```
CCCGAGGAGTTTCGCAGAATCAACTTCTGGTTAGAGTTATGGGAAGCGCGGTTATGGACACCAAGAAGAAAAAAGATGTTTCCAGCCC  WT
CCCGAGGAGTTTCGCAGAATCAACTTCTGGTTAGA---------------GTTATGGACACCAAGAAGAAAAAAGATGTTTCCAGCCC  Δ15 x22
CCCGAGGAGTTTCGCAGAATCAACTTCTGGTTAGAGTTAT---------------------/ /---------------------GGc  Δ178 x20
CCCGAGGAGTTTCGCAGAATCAACTTCTGGTTAGA---------------GTTATGGACACCAAGAAG-AAAAAGATGTTTCCAGCCC  Δ16 x4
CCCGAGGAGTTTCGCAGAATCAACTTCTGGTTAGA---------------GTTATGGACACCAAGAAGAAAAAAGATGTTTCCAG-CC  Δ16 x3
CCCGAGGAGTTTCGCAGAATCAACTTCTGGTTAGAGTTATGGgaaGAAGCGCGGTTATGGACACCAAGAAG-AAAAAGATGTTTCCAG  +2 (Δ1 +3) x2
```

PSEN1 Mutations in 88 of 3807 sequences ≈ 2.3%

```
TGTTTTCTGTGAAACAGTATTTCTATACAGTTGCTCCAATGACAGAGTTACCTGCACCGTTGTCCTACTTCCAGAATGCACAGATGTC  WT
TGTTTTCTGTGAAACAGTATTTCTATACAGTTGCTCCAATgacaGACAGAGTTACCTGCACCGTTGTCCTACTTCCAGAATGCACAGA  +4 x5
TGTTTTCTGTGAAACAGTATTTCTATACAGTTGCTCCAATG--------ACCTGCACCGTTGTCCTACTTCCAGAATGCACAGATGTC  Δ8 x4
TGTTTTCTGTGAAACAGTATTTCTATACAGTTGCTCCAATG--------------ACCGTTGTCCTACTTCCAGAATGCACAGATGTC  Δ14 x4
TGTTTTCTGTGAAACAGTATTTCTATACAGTTGCTCCAAT------GTTACCTGCACCGTTGTCCTACTTCCAGAATGCACAGATGTC  Δ6 x4
```

PSEN2 Mutations in 344 of 2888 sequences ≈ 11.9%

```
AAGGTCCTTGTGCTCCTTTTTCCAGGTGCTTCCAGAGGCAGGGCTATGCTCACATTCATGGCCTCTGACAGCGAGGAAGAAGTGTGTG  WT
AAGGTCCTTGTGCTCCTTTTTCCAGGTGCTTCCAGAGGCAGG----TGCTCACATTCATGGCCTCTGACAGCGAGGAAGAAGTGTGTG  Δ4 x12
AAGGTCCTTGTGCTCCTTTTTCCAGGTGCTTCCAGAG----------GCTCACATTCATGGCCTCTGACAGCGAGGAAGAAGTGTGTG  Δ10 x11
AAGGTCCTTGTGCTCCTTTTTCCAGGTGCTTCCAGAGGCA-------GCTCACATTCATGGCCTCTGACAGCGAGGAAGAAGTGTGTG  Δ7 x9
AAGGTCCTTGTGCTCCTTTTTCCAGGTGCTTCCAGAGGCAG------GCTCACATTCATGGCCTCTGACAGCGAGGAAGAAGTGTGTG  Δ6 x8
AAGGTCCTTGTGCTCCTTTTTCCAGGTGCTTCCAGA-----------------------/ /-----------------------GGC  Δ172 x7
```

PTPN11 Mutations in 58 of 1498 sequences ≈ 3.9%

```
CCTGAGCAAGGAGCGGGTCCGTCGCGGAGCCGGAGGGCGGGAGGAACATGACATCGCGGAGGTGAGGAGCCCCGAGGGGCCCGGCGCG  WT
CCTGAGCAAGGAGCGGGTCCGTCGCGGAGCCGGAGGGCGGGAGgaacGAACATGACATCGCGGAGGTGAGGAGCCCCGAGGGGCCCGG  +4 x2
CCTGAGCAAGGAGCGGGTCCGTCGCGGAGCCGGAGGGCGGGA-------GACATCGCGGAGGTGAGGAGCCCCGAGGGGCCCGGCGCG  Δ7 x2
CCTGAGCAAGGAGCGGGTCCGTCGCGGAGCCGGAGGGCG---GGAACATGACATCGCGGAGGTGAGGAGCCCCGAGGGGCCCGGCGCG  Δ3 x2
CCTGAGCAAGGAGCGGGTCCGTCGCGGAGC---------GGAGGAACATGACATCGCGGAGGTGAGGAGCCCCGAGGGGCCCGGCGCG  Δ9 x2
CCTGAGCAAGGAGCGGGTCCGTCGCGGAGCCGGAGGGCGGGAGGA-----------------TGAGGAGCCCCGAGGGGCCCGGCGCG  Δ17 x1
```

RAF1 Mutations in 70 of 1725 sequences ≈ 4.1%

```
TTACCAGGTTTAAGAATTGTTTAAGCTGCATCAATGGAGCACATACAGGGAGCTTGGAAGACGATCAGCAATGGTTTTGGATTCAAAG  WT
TTACCAGGTTTAAGAATTGTTTAAGCTGCATCAAT-------------GGAGCTTGGAAGACGATCAGCAATGGTTTTGGATTCAAAG  Δ13 x6
TTACCAGGTTTAAGAATTGTTTAAGCTGCATCAAT------------GGGAGCTTGGAAGACGATCAGCAATGGTTTTGGATTCAAAG  Δ12 x2
TTACCAGGTTTAAGAATTGTTTAAGCTGCATCA-------------AGGGAGCTTGGAAGACGATCAGCAATGGTTTTGGATTCAAAG  Δ13 x2
TTACCAGGTTTAAGAATTGTTTAAGCTGCAT--------------CAGGGAGCTTGGAAGACGATCAGCAATGGTTTTGGATTCAAAG  Δ14 x2
TTACCAGGTTTAAGAATTGTTTAAGCTGCATCAATGGAGCACAT--------------AGACGATCAGCAATGGtTTTGGATTCAAA  Δ13 (Δ14 +1) x1
```

RBM20 Mutations in 152 of 4228 sequences ≈ 3.6%

```
CCTTGAGCTCTCTCGCCGCGATCCCGGGCGGGTCTCGCCCCGCATGGTGCTGGCAGCAGCCATGAGCCAGGACGCGGACCCCAGCGGT  WT
CCTTGAGCTCTCTCGCCGCGATCCCGGGCGGGTCTC------------GCTGGCAGCAGCCATGAGCCAGGACGCGGACCCCAGCGGT  Δ12 x5
CCTTGAGCTCTCTCGCCGCGATCCCGG---------------------GCTGGCAGCAGCCATGAGCCAGGACGCGGACCCCAGCGGT  Δ21 x5
CCTTGAGCTCTCTCGCCGCGATCCCGGGCGGGTCTCGCCCCGCA------TGGCAGCAGCCATGAGCCAGGACGCGGACCCCAGCGGT  Δ6 x4
CCTTGAGCTCTCTCGCCGCGATCCCGGGCGGGTCTC----------------GCAGCAGCCATGAGCCAGGACGCGGACCCCAGCGGT  Δ16 x4
CCTTGAGCTCTCTCGCCGCGATCCCG-------------------------GGCAGCAGCCATGAGCCAGGACGCGGACCCCAGCGGT  Δ25 x4
```

RYR2 Mutations in 11 of 1322 sequences ≈ 0.8%

```
GGCGAGGAGGCGCGGAACCATGGCCGATGGGGGCGAGGGCGAAGACGAGATCCAGTTCCTGCGAACTGTAAGCGCCGTGCGTCGCGTG  WT
GGCGAGGAGGCGCGGAACCATGGCCGATGGGGGCGAGGGCGAAgacGACGAGATCCAGTTCCTGCGAACTGTAAGCGCCGTGCGTCGC  +3 x3
GGCGAGGAGGCGCGGAACCATGGCCGATGGGGGCGAGGGCGAAGACGA----CAGT-----gGAACTGTAAGCGCCGTGCGTCGCGTG  Δ9 (Δ10 +1) x1
GGCGAGGAGGCGCGGAACCATGGCCGATGGGGGCGAGGGCGAActgtaagcgatgGcCGAtggggcGATCCAGTTCCTGCGAACTGT  +19 (Δ1 +20) x1
GGCGAGGAGGCGCGGAACCATGGCCGATGGGGGCGAGGGCGA----GcGATCCAGTTCCTGCGAACTGTAAGCGCCGTGCGTCGCGTG  Δ4 (Δ5 +1) x1
GGCGAGGAGGCGCGGAACCATGGCCGATGGGGGCGAGGGCGA-----AGATCCAGTTCCTGCGAACTGTAAGCGCCGTGCGTCGCGTG  Δ5 x1
```

SCN5A Mutations in 87 of 4481 sequences ≈ 1.9%

```
CCTGTGCCCAGAAGCAGGATGAGAAGATGGCAAACTTCCTATTACCTCGGGGCACCAGCAGCTTCCGCAGGTTCACACGGGAGTCCCT  WT
CCTGTGCCCAGAAGCAGGATGAGAAGATGGCAAACTT-------CCTCGGGGCACCAGCAGCTTCCGCAGGTTCACACGGGAGTCCCT  Δ7 x5
CCTGTGCCCAGAAGCAGGATGAGAAGATGGCAAACT----------TCGGGGCACCAGCAGCTTCCGCAGGTTCACACGGGAGTCCCT  Δ10 x3
CCTGTGCCCAGAAGCAGGATGAGAAGATGGCA-----CC----AgCTCGGGGCACCAGCAGCTTCCGCAGGTTCACACGGGAGTCCCT  Δ9 (Δ10+1) x3
CCTGTGCCCAGAAGCAGGATGAGAAGATGGCAAACTTCCTATtacTACCTCGGGGCACCAGCAGCTTCCGCAGGTTCACACGGGAGTC  +3 x2
CCTGTGCCCAGAAGCAGGATGAGAAGATGGCAAACTTC----------------CCAGCAGCTTCCGCAGGTTCACACGGGAGTCCCT  Δ16 x2
```

SCO2 Mutations in 84 of 3459 sequences ≈ 2.4%

```
GGCTCCTGACGCCTGTGCTTGTTTCCAGGAGCATCAGATCCATGCTGCTGCTGACTCGGAGCCCCACAGCTTGGCACAGGCTCTCTCA  WT
GGCTCCTGACGCCTGTGCTTGTTTCCAGGAGCATCAGATCCA---TGCTGCTGACTCGGAGCCCCACAGCTTGGCACAGGCTCTCTCA  Δ3 x15
GGCTCCTGACGCCTGTGCTTGTTTCCAGG---ATCA-------------------TCGGAGCCCCACAGCTTGGCACAGGCTCTCTCA  Δ22 x3
GGCTCCTGACGCCTGTGCTTGTTTCCAGGAGCATCAGATCCtggtgctgactcggagATcaGaTGCTGCTGACTCGGAGCCCCACAGC  +18 (Δ1 +19) x2
GGCTCCTGACGCCTGTGCTTGTTTCCAGGAGCATC-------TGCTGCTGCTGACTCGGAGCCCCACAGCTTGGCACAGGCTCTCTCA  Δ7 x2
GGCTCCTGACGCCTGTGCTTGTTTCCAGGAGCATC------------------------AGCCCCACAGCTTGGCACAGGCTCTCTCA  Δ24 x2
```

SDHA Mutations in 66 of 1815 sequences ≈ 3.6%

```
AACAGCAGACATGTCGGGGGTCCGGGGCCTGTCGCGGCTGCTGAGCGCTCGGCGCCTGGCGCTGGCCAAGGCGGTGAGTCCGTGCCGC  WT
AACAGCAGACATGTCGGGGGTCCGGGGCCTGTCGCGGCTGCTGAGCG--------CTGGCGCTGGCCAAGGCGGTGAGTCCGTGCCGC  Δ8 x2
AACAGCAGACATGTCGGGGGTCCGGGGCCTGTCGCGGCT-------GCTCGGCGCCTGGCGCTGGCCAAGGCGGTGAGTCCGTGCCGC  Δ7 x2
AACAGCAGACATGTCGGGGGTCCGGGGCCTGTCGCGGCTGCTGAGCGCT---CGCCTGGCGCTGGCCAAGGCGGTGAGTCCGTGCCGC  Δ3 x1
AACAGCAGACATGTCGGGGGTCCGGGGCCTGTCGCGGCTGCTGAGCGCT---------GCGCTGGCCAAGGCGGTGAGTCCGTGCCGC  Δ9 x1
AACAGCAGACATGTCGGGGGTCCGGGGCCTGTCGCGGCTGCTGAGC-CT---CGCCTGGCGCTGGCCAAGGCGGTGAGTCCGTGCCGC  Δ4 x1
```

SGCD Mutations in 7 of 2203 sequences ≈ 0.3%

```
AGACATTACTGCCGGGAGTGTTGAGTGAAGGGACCAGGTGGAGATGGTGAGTAATTCCCGGGAGCGAAGCTTGTTCAAGGCCCTGCTC  WT
AGACATTACTGCCGGGAGTGTTGAGTGAAGGGACCAGGTGGAgatGATGGTGAGTAATTCCCGGGAGCGAAGCTTGTTCAAGGCCCTG  +3 x2
AGACATTACTGCCGGGAGTGTTGAGTGAAGGGACCAGGT---GAT-GTGAGTAATTCCCGGGAGCGAAGCTTGTTCAAGGCCCTGCTC  Δ4 x1
AGACATTACTGCCGGGAGTGTTGAGTGAAGGGACCAG-----GATGGTGAGTAATTCCCGGGAGCGAAGCTTGTTCAAGGCCCTGCTC  Δ5 x1
AGACATTACTGCCGGGAGTGTTGAGTGAAGGGACCA---------GGTGAGTAATTCCCGGGAGCGAAGCTTGTTCAAGGCCCTGCTC  Δ9 x1
AGACATTACTGCCGGGAGTGTTGAGTGAAGGGACCA---------GGTGAGTAATT-CCGGGAGCGAAGCTTGTTCAAGGCCCTGCTC  Δ10 x1
```

SLC25A20 Mutations in 12 of 542 sequences ≈ 2.2%

```
AAGCCAGGACGGCCCGAGAACTGACAGACGGAGTGACAGACGGACTGACCATGGCCGACCAGCCAAAACCCATCAGCCCGCTCAAGAA  WT
AAGCCAGGACGGCCCGAGAACTGACAGACGGAGTGACA--------GACCATGGCCGACCAGCCAAAACCCATCAGCCCGCTCAAGAA  Δ8 x3
AAGCCAGGACGGCCCGAGAACTGACAGACGGAGTGACAGACG----GACCATGGCCGACCAGCCAAAA-CCATCAGCCCGCTCAAGAA  Δ5 x1
AAGCCAGGACGGCCCGAGAACTGACAGACGGAGTGACAGACG--------------GACCAGCCAAAACCCATCAGCCCGCTCAAGAA  Δ14 x1
AAGCCAGGACGGCCCGAGAACTGACAGACGGAGTGACAGAC-GACTG-----------CCAcaCAAAACCCATCAGCCCGCTCAAGAA  Δ12 (Δ14 +2) x1
AAGCCAGGACGGCCCGAGAACTGACAGACGGAGTGACAGtCGG-----CCATGGCCGACCAGCCAAAACCCATCAGCCCGCTCAAGAA  Δ5 (Δ6 +1) x1
```

SLC25A4 Mutations in 85 of 1114 sequences ≈ 7.6%

```
CGAACGGGCTGCCTGCGGGCTGAGAGCGTCGAGCTGTCACCATGGGTGATCACGCTTGGAGCTTCCTAAAGGACTTCCTGGCCGGGGG  WT
CGAACGGGCTGCCTGCGGGCTGAGAGCGTCGAGC------------TGATCACGCTTGGAGCTTCCTAAAGGACTTCCTGGCCGGGGG  Δ12 x9
CGAACGGGCTGCCTGCGGGCTGAGAGCGTCGAGCTG-------------TCACGCTTGGAGCTTCCTAAAGGACTTCCTGGCCGGGGG  Δ13 x7
CGAACGGGCTGCCTGCGGGCTGAGAGCGTCGAGCTGTCACCAtgggTGGGTGATCACGCTTGGAGCTTCCTAAAGGACTTCCTGGCCG  +4 x4
CGAACGGGCTGCCTGCGGGCTGAGAGCGTCGAGCTGTCACCA--------------TGGAGCTTCCTAAAGGACTTCCTGGCCGGGGG  Δ14 x3
CGAACGGGCTGCCTGCGGGCTGAGAGCGTCGA---------------------GCTTGGAGCTTCCTAAAGGACTTCCTGGCCGGGGG  Δ21 x3
```

SURF1 Mutations in 12 of 310 sequences ≈ 3.9%

```
CCCGCGGGGCCGGGTGCGATGGCGGCGGTGGCTGCGTTGCAGCTGGGGCTGCGGGCGGCGGGGCTGGGACGGGTGAGCGCCGGGTGCG  WT
CGGG----------------------/ /---------------------TGCG----------------------------GGGTGCG  Δ141 x2
CCCGCGGGGCCGGGTGCGATGGCGGCGGTGGCTGCGTTGCAGCT-GGGCT----GCGGCGGGGCTGGGACGGGTGAGCGCCGGGTGCG  Δ5 x1
CCCGCGGGGCCGGGTGCGATGGCGGCGGTGGCTGCGTTGCAGCT--------GGGCtGCGcGGCTGGGACGGGTGAGCGCCGGGTGCG  Δ8 (Δ10 +2) x1
CCCGCGGGGCCGGGTGCGATGGCGGCGGTGGCTGCGTTGCAggtgtagtgGCTtgcaggtgcaggtgtagGtGGCTGCGGGCGGCGGG  +26 (Δ1 +27) x1
CCCGCGGGGCCGGGTGCGATGGCGGCGGTGGCTGCG-TGCAGCTG-----GCGGGCGGC---GCTGGGACGGGTGAGCGCCGGGTGCG  Δ9 x1
```

SYNE1 Mutations in 15 of 3122 sequences ≈ 0.5%

```
TTGGTGTTGGCTTCGTGCTTCCGGAGGGACCATGGCAACCTCCAGAGGGGCCTCCCGGTGTCCTCGGGATATCGCCAATGTGATGCAG  WT
TTGGTGTTGGCTTCGTGCTTCCGGAGGGACCATGGCAA------------CCTCCCGGTGTCCTCGGGATATCGCCAATGTGATGCAG  Δ12 x2
TTGGTGTTGGCTTCGTGCTTCCGGAGGGACCATGGCAACCTCCAGAGG---CTCCCGGTGTCCTCGGGATATCGCCAATGTGATGCAG  Δ3 x1
TTGGTGTTGGCTTCGTGCTTCCGGAGGGACCATGGCAACCTCCAGA---GCCTcCCGGTGTCCTCGGGATATCGCCAATGTGATGCA  Δ2 (Δ3 +1) x1
TTGGTGTTGGCTTCGTGCTTCCGGAGGGACCATGGCAACCTCCA------------GGTGTCCTCGGGATATCGCCAATGTGATGCAG  Δ12 x1
TTGGTGTTGGCTTCGTGCTTCCGGAGGGACCATGGCAA-aTC--------aCTCCCGGTGTCCTCGGGATATCGCCAATGTGATGCAG  Δ9 (Δ11 +2) x1
```

TAZ Mutations in 38 of 4160 sequences ≈ 0.9%

```
CCACAGGCCGGCCCGGGGCGCTGGGAGCGCCGGCCGCGGGCCGGGTGGGGATGCCTCTGCACGTGAAGTGGCCGTTCCCCGCGGTGCC  WT
CCACAGGCCGGCCCGGGGCGCTGGGA--------------------------GCCTCTGCACGTGAAGTGGCCGTTCCCCGCGGTGCC  Δ26 x3
CCACAGGCCGGCCC--------------------------------GGGGATGCCTCTGCACGTGAAGTGGCCGTTCCCCGCGGTGCC  Δ32 x3
CCACAGGCCGGCCCGGGGCGCTGGGAGCGCCGGCCGCGGGCCGGGT----ATGCCTCTGCACGTGAAGTGGCCGTTCCCCGCGGTGCC  Δ4 x2
CCACAGGCCGGCCCGGGGCGCTGGGAGCGCCGGCCGCGGGCCGGG---------CTCTGCACGTGAAGTGGCCGTTCCCCGCGGTGCC  Δ9 x2
CCACAGGCCGGCCCGGGGCGCTGGG-GaGCtG-----GGGC---GctGGGATGCCTCTGCACGTGAAGTGGCCGTTCCCCGCGGTGCC  Δ9 (Δ13 +4) x2
```

TBX1 Mutations in 254 of 4843 sequences ≈ 5.2%

```
ACCGGGTGAAGCTTCGCTGGCTGCCAGGATCCCCGGCAGGGATGCACTTCAGCACCGTCACCAGGGACATGGAAGGTGAGCCTCCAGG  WT
ACCGGGTGAAGCTTCGCTGGCTGCCAGGATCCCCG----------------GCACCGTCACCAGGGACATGGAAGGTGAGCCTCCAGG  Δ16 x16
ACCGGGTGAAGCTTCGCTGGCTGCCAGGATCC----------------------CCGTCACCAGGGACATGGAAGGTGAGCCTCCAGG  Δ22 x12
ACCGGGTGAAGCTTCGCTGGCTGCCAGGATCCCCGG-------------CAGCACCGTCACCAGGGACATGGAAGGTGAGCCTCCAGG  Δ13 x11
ACCGGGTGAAGCTTCGCTGGCTGCCAGGATCCCCGGCAGGGATgcaGCACTTCAGCACCGTCACCAGGGACATGGAAGGTGAGCCTCC  +3 x10
ACCGGGTGAAGCTTCGCTGGCTGCCAGGATCCCCGG----------------------CACCAGGGACATGGAAGGTGAGCCTCCAGG  Δ22 x7
```

TBX20 Mutations in 553 of 3624 sequences ≈ 15.3%

```
AGTTCGGACGACCCCGTCCCTGGCCAGGACCGCGTGCTGGGGACCATGGAGTTCACGGCGTCCCCCAAGCCCCAACTCTCCTCCCGGG  WT
AGTTCGGACGACCCCGTCCCTGGCCAGGACCGCGTGC---------TGGAGTTCACGGCGTCCCCCAAGCCCCAACTCTCCTCCCGGG  Δ9 x48
AGTTCGGACGACCCCGTCCCTGGCCAGGACCGCGTGCTGGGGaccACCATGGAGTTCACGGCGTCCCCCAAGCCCCAACTCTCCTCCC  +3 x29
AGTTCGGACGACCCCGTCCCTGGCCAGGACCGCGTGCTGG-------GGAGTTCACGGCGTCCCCCAAGCCCCAACTCTCCTCCCGGG  Δ7 x21
AGTTCGGACGACCCCGTCCCTGGC-----------------------------/ /-----------------------------CAA  Δ244 x20
AGTTCGGACGACCCCGTCCCTGGCCAGGACCGCGTGC---------TGGAGTTCACGGCGT-CCCCAAGCCCCAACTCTCCTCCCGGG  Δ10 x14
```

TBX5 Mutations in 266 of 549 sequences ≈ 48.5%

```
CCTTGCGCGGGCACAGGGCCCTGGGCGCACCATGGCCGACGCAGACGAGGGCTTTGGCCTGGCGCACACGCCTCTGGAGCCTGACGCA  WT
CCTTGCGCGGGCACAGG--------------------------------/ /---------------------------------GCC  Δ337 x20
CCTTGCGCGGGCACAGGGCCCTGGGCGCA--------------------------/ /---------------------------CCA  Δ313 x17
CCTTGCGCGGGCACAGGGCCCTGGGCGCA--------------------------/ /---------------------------CCT  Δ292 x17
CCCA---------------------------------------/ /----------------------------------------TCT  Δ380 x14
CCTTGCGCGGGCACAGGGCCCTGGGCGCACCATGGCCGACGCAGACG-----------------/ /------------------ATC  Δ265 x14
```

TCAP Mutations in 40 of 3026 sequences ≈ 1.3%

```
CCTGGGAGGGGAGAGAGAATGAGGAGTGATCATGGCTACCTCAGAGCTGAGCTGCGAGGTGTCGGAGGAGAACTGTGAGCGCCGGGAG  WT
CCTGGGAGGGGAGAGAGAATGAGGAGTGATCATGGCTAC-------CTGAGCTGCGAGGTGTCGGAGGAGAACTGTGAGCGCCGGGAG  Δ7 x11
CCTGGGAGGGGAGAGAGAATGAGGAGTGATCATGGCTACCTCAGAGCTG----GCGAGGTGTCGGAGGAGAACTGTGAGCGCCGGGAG  Δ4 x2
CCTGGGAGGGGAGAGAGAATGAGGAGTGATCATGGCTACCTCAG----GAGCTGCGAGGTGTCGGAGGAGAACTGTGAGCGCCGGGAG  Δ4 x1
CCTGGGAGGGGAGAGAGAATGAGGAGTGATCATGGCTACCTCA-----GAGCTGCGAGGTGTCGGAGGAGAACTGTGAGCGCCGGGAG  Δ5 x1
CCTGGGAGGGGAGAGAGAATGAGGAGTGATCATGGCTACCTCagaAGAGCTGAGCTGCGAGGTGTCGGAGGAGAACTGTGAGCGCCGG  +3 x1
```

TGFB3 Mutations in 446 of 3334 sequences ≈ 13.4%

```
TTCCTCTCCAGGCCTTGCCGTCCCCCTGGCCTCTCTTCCCAGCTCACACATGAAGATGCACTTGCAAAGGGCTCTGGTGGTCCTGGCC  WT
TTCCTCTCCAGGCCTTGCCGTCCCCCTGGCCTCTCTTCCCAGctcaCTCACACATGAAGATGCACTTGCAAAGGGCTCTGGTGGTCCT  +4 x16
TTCCTCTCCAGGCCTTGCCGTCCCCCTGGCCTCTCTTCC-----CACACATGAAGATGCACTTGCAAAGGGCTCTGGTGGTCCTGGCC  Δ5 x15
TTCCTCTCCAGGCCTTGCCGTCCCCCTGGCCTCTCTTCC---------CATGAAGATGCACTTGCAAAGGGCTCTGGTGGTCCTGGCC  Δ9 x15
TTCCTCTCCAGGCCTTGCCGTCCCCCTGGCCTCTCTTCC-------CACATGAAGATGCACTTGCAAAGGGCTCTGGTGGTCCTGGCC  Δ7 x11
```

```
TTCCTCTCCAGGCCTTGCCGTCCCCCTGGCCTCTCTTCCCAGCtcaTCACACATGAAGATGCACTTGCAAAGGGCTCTGGTGGTCCTG  +3 x7
```

TMEM43 Mutations in 37 of 1524 sequences ≈ 2.4%

```
GGCGGCGGCAGCGAGCCGGGTCCCACCATGGCCGCGAATGTGAGTATCCCCGGGCCAGCCGGGCCACACCCAGGCTTCCCCGTCGCCC  WT
GGCGGCGGCAGCGAGCCGGGTCCCACCATGGCCGCGAATGTGAGTATCCCC------GCCGGGCCACACCCAGGCTTCCCCGTCGCCC  Δ6 x1
GGCGGCGGCAGCGAGCCGGGTCCCACCATGGCCGCGAATGTGAGTATC----GGCCAGCCGGGCCACACCCAGGCTTCCCCGTCGCCC  Δ4 x1
GGCGGCGGCAGCGAGCCGGGTCCCACCATGGCCGCGAATGTGAGT-TCC----GgCAGCCGG--CACACCCAGGCTTCCCCGTCGCCC  Δ7 (Δ8 +1) x1
GGCGGCGGCAGCGAGCCGGGTCCCACCATGGCCGCGAATGT---------------AtCCGGGCCgACACCCAGGCTTCCCCGTCGCCC  Δ14 (Δ16 +2) x1
GGCGGCGGCAGCGAGCCGGGTCCCACCATGGCCGCGAATG-----AT---------tGCCGGGCCACACCCAGGCTTCCCCGTCGCCC  Δ14 (Δ15 +1) x1
```

TMPO Mutations in 118 of 2337 sequences ≈ 5%

```
GTGGGGAGGGGGCTTCGCAGATCCCCGAGATGCCGGAGTTCCTGGAAGACCCCTCGGTCCTGACAAAAGACAAGTTGAAGAGTGAGTT  WT
GTGGGGAGGGGGCTTCGCAGATCCCCGAGATGCCGGAGTT----------CCCTCGGTCCTGACAAAAGACAAGTTGAAGAGTGAGTT  Δ10 x5
GTGGGGAGGGGGCTTCGCAGATCCCCGAGATGCCGGAGTTCCT----GAaCCCTCGGTCCTGACAAAAGACAAGTTGAAGAGTGAGTT  Δ4 (Δ5 +1) x3
GTGGGGAGGGGGCTTCGCAGAT---------------------------CCCCTCGGTCCTGACAAAAGACAAGTTGAAGAGTGAGTT  Δ27 x3
GTGGGGAGGGGGCTTCGCAGATCCCCGAGATGCCGGAGTTCCTG-----aCCCTCGGTCCTGACAAAAGACAAGTTGAAGAGTGAGTT  Δ5 (Δ6 +1) x2
GTGGGGAGGGGGCTTCGCAGATCCCCGAGATGCCGGAGTTC-----tGAaCCCTCGGTCCTGACAAAAGACAAGTTGAAGAGTGAGTT  Δ5 (Δ7 +2) x2
```

TNNC1 Mutations in 243 of 3354 sequences ≈ 7.2%

```
TGGCAACCCCAGCAAGCTGTCCTGTGAGCCGCCAGCATGGATGACATCTACAAGGCTGCGGTGAGGGACAGGGCTGGGTAGGGCTGGG  WT
TGGCAACCCCAGCAAGCTGTCCTGTGAGCCGCCAGCAT---------------GGCTGCGGTGAGGGACAGGGCTGGGTAGGGCTGGG  Δ15 x15
AGCC----------------------/ /-----------------------AGGCTGCGGTGAGGGACAGGGCTGGGTAGGGCTGGG  Δ218 x9
TGGCAACCCCAGCAAGCTGTCCTGTGAGCCGCCAG---------------CAAGGCTGCGGTGAGGGACAGGGCTGGGTAGGGCTGGG  Δ15 x8
TGGCAACCCCAGCAAGCTGTCCTGTGAGCCGCCAGCATGGatgacATGACATCTACAAGGCTGCGGTGAGGGACAGGGCTGGGTAGGG  +5 x5
TGGCAACCCCAGCAAGCTGTCCTGTGAGCCGCCAGCAT----------------GCTGCGGTGAGGGACAGGGCTGGGTAGGGCTGGG  Δ16 x4
```

TNNI3 Mutations in 47 of 2759 sequences ≈ 1.7%

```
TCGCCCTGCCTCCTGCCATTCCCGGCCTGAGTCTCAGCATGGCGGATGGGTGAGTGATGCCCCAAGGCAGTGGGAGTTGGGGGCGACC  WT
TCGCCCTGCCTCCTGCCATTCCCGGCCTGAGTCTCAGC-------ATGGGTGAGTGATGCCCCAAGGCAGTGGGAGTTGGGGGCGACC  Δ7 x10
TCGCCCTGCCTCCTGCCATTCCCGGCCTGAGTCTCAGCATGGcggaCGGATGGGTGAGTGATGCCCCAAGGCAGTGGGAGTTGGGGGC  +4 x3
TCGCCCTGCCTCCTGCCATTCCCGGCCTGAGTCTCAGCATG------GGaTGAGTGATGCCCCAAGGCAGTGGGAGTTGGGGGCGACC  Δ6 (Δ7 +1) x3
TCGCCCTGCCTCCTGCCATTCCCGGCCTGAGTCTCAGC-------ATGGGTGAGTGATG-CCCAAGGCAGTGGGAGTTGGGGGCGACC  Δ8 x3
TCGCCCTGCCTCCTGCCATTCCCGGCCTGAGTCTCAGCAT--------GGTGAGTGATGCCCCAAGGCAGTGGGAGTTGGGGGCGACC  Δ8 x2
```

TNNT2 Mutations in 67 of 510 sequences ≈ 13.1%

```
CCTTTGTACCTGCACTGACTTTTTTCTCCTTTTGGAGGGAGAGCAGAGACCATGTCTGACATAGAAGAGGTGGTGGAAGAGTACGAGG  WT
CCTTTGTACCTGCACTGACTTTTTTCTCCTTTTGGAGGGAGagcAGCAGAGACCATGTCTGACATAGAAGAGGTGGTGGAAGAGTACG  +3 x5
CCTTTGTACCTGCACTGACTTTTTTCTCCTTTTGGAGGGAGAGCAGA----------GACATAGAAGAGGTGGTGGAAGAGTACGAGG  Δ10 x3
CCTTTGTACCTGCACTGACTTTTTTCTCCTTTTGGAGGGAGAGCAGA--------CTGACATAGAAGAGGTGGTGGAAGAGTACGAGG  Δ8 x2
CCTTTGTACCTGCACTGACTTTTTTCTCCTTTTGGAGGGAG----cAGACCATGTCTGACATAGAAGAGGTGGTGGAAGAGTACGAGG  Δ4 (Δ5 +1) x2
CCTTTGTACCTGCACTGACTTTTTTCTCCTTTTGGAGG-------GAGACCATGTCTGACATAGAAGAGGTGGTGGAAGAGTACGAGG  Δ7 x2
```

TPM1 Mutations in 188 of 2913 sequences ≈ 6.5%

```
TGCTGCAGCCCCAGGGCCCCTCGCCGCCGCCACCATGGACGCCATCAAGAAGAAGATGCAGATGCTGAAGCTCGACAAGGAGAACGCC  WT
TGCTGCAGCCCCAGGGCCCCTCGCCGCCGCCACCATG--------------GAAGATGCAGATGCTGAAGCTCGACAAGGAGAACGCC  Δ14 x6
TGCTGCAGCCCCAGGGCCCCTCGCCGCCGCCACCATG-----------------GATGCAGATGCTGAAGCTCGACAAGGAGAACGCC  Δ17 x5
TGCTGCAGCCCCAGGGCCCCTCGCCGCCGCCAC------------CAAGAAGAAGATGCAGATGCTGAAGCTCGACAAGGAGAACGCC  Δ12 x5
TGCTGCAGCCCCAGGGCCCCTCGCCGCCGCCACCATGGAC-----------GAAGATGCAGATGCTGAAGCTCGACAAGGAGAACGCC  Δ11 x4
TGCTGCAGCCCCAGGGCCCCTCGCCGCCGCCACCATG-----------GAAGAAGATGCAGATGCTGAAGCTCGACAAGGAGAACGCC  Δ11 x4
```

TTN Mutations in 28 of 3899 sequences ≈ 0.7%

```
CTAATTTATTTTCTCTTCTTTTTCAGAGTGCCTAGAAAGATGACAACTCAAGCACCGACGTTTACGCAGCCGTTACAAAGCGTTGTGG  WT
CTAATTTATTTTCTCTTCTTTTTCAGAGTGCCTAGAAAGATG----tTCAAGCACCGACGTTTACGCAGCCGTTACAAAGCGTTGTGG  Δ4 (Δ5 +1) x3
CTAATTTATTTTCTCTTCTTTTTCAGAGTGCCTAGAAAGATGacaACAACTCAAGCACCGACGTTTACGCAGCCGTTACAAAGCGTTG  +3 x3
CTAATTTATTTTCTCTTCTTTTTCAGAGTGCCTAGA-------------AAGCACCGACGTTTACGCAGCCGTTACAAAGCGTTGTGG  Δ13 x2
CTAA--------------------------------------ACAACTCAAGCACCGACGTTTACGCAGCCGTTACAAAGCGTTGTGG  Δ38 x2
CTAATTTATTTTCTCTTCTTTTTCAGAGTGCCTAGAAAGATGACA---CcAGCACCGACGTTTACGCAGCCGTTACAAAGCGTTGTGG  Δ3 (Δ4 +1) x1
```

TTR Mutations in 187 of 6549 sequences ≈ 2.9%

```
TCACAGAAGTCCACTCATTCTTGGCAGGATGGCTTCTCATCGTCTGCTCCTCCTCTGCCTTGCTGGACTGGTATTTGTGTCTGAGGCT  WT
TCACAGAAGTCCACTCATTCTTGGCAGGATGGCTT-----------CTCCTCCTCTGCCTTGCTGGACTGGTATTTGTGTCTGAGGCT  Δ11 x32
TCACAGAAGTCCACTCATTCTTGGCAGGATG--------------GCTCCTCCTCTGCCTTGCTGGACTGGTATTTGTGTCTGAGGCT  Δ14 x27
TCACAGAAGTCCACTCATTCTTGGCAGGATGGCTTCTCAT------CTCCTCCTCTGCCTTGCTGGACTGGTATTTGTGTCTGAGGCT  Δ6 x6
TCACAGAAGTCCACTCATTCTTGGCAGGAT----------------CTCCTCCTCTGCCTTGCTGGACTGGTATTTGTGTCTGAGGCT  Δ16 x5
GGCA---------------/ /----------------AT---------------------------CTGGTATTTGTGTCTGAGGCT  Δ124 x3
```

TXNRD2 Mutations in 6 of 1044 sequences ≈ 0.6%

```
CCCCACGACGATGGCGGCAATGGCGGTGGCGCTGCGGGGATTAGGAGGGCGCTTCCGGTGGCGGACGCAGGCCGTGGCGGGCGGGGTG  WT
CCCCACGACGATGGCGGCAATGGCGGTGGCGCTGC--GGAT-----aGGCGCTTCCGGTGGCGGACGCAGGCCGTGGCGGGCGGGGTG  Δ7 (Δ8 +1) x1
CCCCACGACGATGGCGGCAATGGCGGTGGCGCTGC--GGAT-----aGaCGCTTCCGGTGGCGGACGCAGGCCGTGGCGGGCGGGGTG  Δ7 (Δ9 +2) x1
CCCCACGACGATGGCGGCAATGGCGGTGGCGCTGC---GATTAG------aCTTCCGGTGGCGGACGCAGGCCGTGGCGGGCGGGGTG  Δ9 (Δ10 +1) x1
CCCCACGACGATGGCGGCAATGGCGGTGGCGCTGC---GAT-----aGaCGCTTCCGGTGGCGGACGCAGGCCGTGGCGGGCGGGGTG  Δ8 (Δ10 +2) x1
CCCCACGACGATGGCGGCAATGGCGGTGGCGCT---------------GCGCTTCCGGTGGCGGACGCAGGCCGTGGCGGGCGGGGTG  Δ15 x1
```

VCL Mutations in 52 of 4159 sequences ≈ 1.3%

```
ACTTCTCTGTCGCCCGCGGTTCGCCGCCCCGCTCGCCGCCGCGATGCCAGTGTTTCATACGCGCACGATCGAGAGCATCCTGGAGCCG  WT
ACTTCTCTGTCGCCCGCGGTTCGCCGCCCCGCTCGCC--------GCCAGTGTTTCATACGCGCACGATCGAGAGCATCCTGGAGCCG  Δ8 x7
ACTTCTCTGTCGCCCGCGGTTCGCCGCCCCGCTCGCCGCC-----GCCAGTGTTTCATACGCGCACGATCGAGAGCATCCTGGAGCCG  Δ5 x4
ACTTCTCTGTCGCCCGCGGTTCGCCGCCCCGCTCGCCGCCG------------aTCATACGCGCACGATCGAGAGCATCCTGGAGCCG  Δ12 (Δ13 +1) x3
```

```
ACTTCTCTGTCGCCCGCGGTTCGCCGCCCCGCTC-----------GCCAGTGTTTCATACGCGCACGATCGAGAGCATCCTGGAGCCG  Δ11 x2
ACTTCTCTGTCGCCCGCGGTTCGCCGCC---------------ATGCCAGTGTTTCATACGCGCACGATCGAGAGCATCCTGGAGCCG  Δ15 x2
```

ZASP Mutations in 14 of 453 sequences ≈ 3.1%

```
ACCCTCTCTACCCTTTGTCTGCAGAGGCGGCCGCTGACAGCACCAGCATGTCTTACAGTGTGACCCTGACTGGGCCCGGGCCCTGGGG  WT
ACCCTCTCTACCCTTTGTCTGCAGAGGCGGCCGCTGA------CAGCATGTCTTACAGTGTGACCCTGACTGGGCCCGGGCCCTGGGG  Δ6 x4
ACCCTCTCTACCCTTTGTCTGCAGAGGCGGCCGCTGACAGCAC---CATGTCTTACAGTGTGACCCTGACTGGGCCCGGGCCCT-GGG  Δ4 x1
ACCCTCTCTACCCTTTGTCTGCAGAGGCGGCCGCTGACAG--CCAGCATGT-------aGTGACCCTGACTGGGCCCGGGCCCTGGGG  Δ9 (Δ10 +1) x1
ACCCTCTCTACCCTTTGTCTGCAGAGGCGGCCGCT-----------------------/ /------------------------GGA  Δ230 x1
ACCCTCTCTACCCTTTGTCTGCAGAGGCGGCCG------------------CTTACAGTGTGACCCTGACTGGGCCCGGGCCCTGGGG  Δ18 x1
```

Online Table III. Predicted off-target loci in TNNT2-KO and DCM-KO iPSC clones

GENE		PCR PRIMERS	AMPLICON (bp)
LOC286094	FW: RV:	GTGGCACAGCAGACTTACAGG GCAGCCTGATATATCCCCTTCC	331
ZNF10	FW: RV:	GCCTTCATCAGAGATTTGACCCC GAGGCAGAGAACCTCCAGATAAAG	345
ORC4	FW: RV:	GCCAGACAGTGAGAAAGATGCAG GGAAGCCTTGCTGGTAACATAGTC	528
CDC20B	FW: RV:	CCCCATGGTTAGCTGGAAAATGGA GTAGGGGAACTGGTGTGCTACT	520
VAMP2	FW: RV:	GGATGCGCCACAGAATTGGG TCTCCAGGACTATTGAGCCCAG	326
CABLES1	FW: RV:	AGAAGGGGTGCAGGTGTACTC ACTGTCGCGATACGGCAGCA	328
FAT3	FW: RV:	CGCCTTGTGAGATTTTTTCCCTGG CCTTGCTCAAGGTCAGCTGTATC	332
C12ORF51	FW: RV:	TCCTTCCCGGCCTTGCTGTA TGCCCATTCAGAGCAGACGC	328
PCDH15	FW: RV:	CAGGCATCAAGTTGGTCGTGCA TCCTTCTGCCTGTCCCCTTC	341
ATXN7L2	FW: RV:	CATCCATGCCTCCAACCCAC CGTTAAATGCAAGTGCGGGGAATG	332
PCID2	FW: RV:	CATAATGGGGACTTCCGTGGG GCTTGACCTTCGAGTGTTTGCC	492
PVRL1	FW: RV:	GGAGAGCGAGACTCTGTCTCA GCTGGGGAGGCAATAGGTATG	332
SDCCAG3	FW: RV:	GTAAAGCTGGCTCCTGTGGC GACTTCCTGCCAGCATGGTG	338
KCNN3	FW: RV:	GGGAGAAGAAGAAGAGGAAGAGG CTCCTCCTCCTTTTCATCATCGTC	355
FBL	FW: RV:	TTTCAGGTCCTGGAGCTTTTGGG TAGGAGATGGTGTGGTGGACCAG	328
WFDC2	FW: RV:	AGGTGGTAAGTGGAGGGGGA GGCTCAGAGAGGTAAACAACATGC	339
RNASEH2B	FW: RV:	CTGCCTCTCTGAGTGTTAACTTCC CTGGTGAAACGACGTGGTAGC	351
ZNF667	FW: RV:	CAGTTGACCCTTGAGCCATGTTAG GGTCACCTTACTGACACCTAAGC	365
PZP	FW: RV:	GGGCATGAGGCTTGTTGTTCTTTG GGCCAAAGCGCAGAAAGCAG	377
CASP12	FW: RV:	GCATGGCAGTATAGAATTCCTGGG GAGAAGCTGAAGGATGCAGGG	377
BARX2	FW: RV:	GGAGCCAGCGAGAATTAAAAGGG ACAGGCAGGCTTTCCAGGCA	340
ERMN	FW: RV:	CGTGTGCCATGTTCATGCTTTCC CTCAGCCTGTTTCTTCCAGTGC	371
CNTN4	FW: RV:	CTGCTATGCCTGTAGGGGTTG CCAGCCATCCCATTACTGGGT	359
NEFM	FW: RV:	GGTTTTTGGGGGACTACATGCAC GGTACCCCCCCAAATTTAAAGAGG	335

Online Table IV. Predicted off-target loci in TBX5-KO clones

GENE		PCR PRIMERS	AMPLICON (bp)
PRKCE	FW:	CAAACCAGCTTCGCTTGGTTCTGA	418
	RV:	CAACCTTGAGCTCGGACCAAAAGA	
RMND5A	FW:	CTGTGCTAGCTAATCCAGTCTGC	412
	RV:	CCAGTTGAGAAAGGTTCCTCCAAG	
SNAR-E	FW:	GAAGGGCTGGGATTACAGGC	325
	RV:	TGACCATGTGATCCATCATGGGG	
TBPL1	FW:	CTAACGCCAGGGGCTTCTGA	377
	RV:	AAGGATGGGAGTGGGAGAGG	
PTPRU	FW:	CAGCAGGAACAAAGAGGCTAAGG	326
	RV:	GAAAAGGGTGAGCTGGCCTG	
MCF2L	FW:	TAGGCAGGGACCCTCCATAC	346
	RV:	ACCCTCAGGCTCTCAGAGTC	
ZC3H3	FW:	GCCCATCAACTGAGGTGGAG	326
	RV:	GGCTGTGGCTGATTCCAGCA	
ZC3H3	FW:	CCCATCAACTGAGGTGGAGAC	326
	RV:	TGGCTGTGGCTGATTCCAGCA	
ARHGEF10	FW:	ACAGAGCCTCTCCCTAGGTG	326
	RV:	CAGAACCCAGCCATTCAGCTGAAG	
TOP3B	FW:	AGCTCTTGAGCCACGGGTGA	331
	RV:	TCAGCATCTTGTGCCCAGCG	
ZNF692	FW:	ATACTTGCTGTCTCCACTCTGCC	327
	RV:	ATGGGTGGTGTTTAGAGCCATGAG	
TRPM1	FW:	CAATGCCTGGCAGACAGCCT	336
	RV:	AGAATTCCGGCCACGTAGCAC	
ASIC2	FW:	CAGGATGATCTCCATCTCCTGAC	330
	RV:	CAAGCCTCAGTTTCCTCGTGTG	
TSPEAR	FW:	GAAGCAAGGCTCTGGGAGGA	357
	RV:	TTCCTCCCAGAGCCCTGCTT	
DAGLA	FW:	CACTGTGCTCCTTCAGACGG	328
	RV:	AGTTAAGGGTGGGGTGGTGG	
SFMBT2	FW:	TTTTGCAGGGGATGGAAAGGGAG	328
	RV:	TCTTGGCCTCTTCTTTGCCCTG	
ANGPT1	FW:	CACCTGGTATTCATAGAGGCCC	406
	RV:	GGAAGTTATCCTGGCAGTGCTAG	
C11orf87	FW:	CCCCCGAAAAGGCAACACAC	367
	RV:	GCCTTGGGCCCAATTCAATTCC	
ABRACL	FW:	GGCTGAAGTTCAGTGGCATGATC	350
	RV:	GGGTTCAAGCAATTCTTCTGCCTC	
MSI2	FW:	TCTCTGTGGATTGGGTGAGAGG	325
	RV:	ATAGGATCTCACCGTGTTAGCCAG	
LY86	FW:	GGCCTTGCTAGGATTAGAACTCAC	330
	RV:	GGGAGCATGTTAGACTCAGCG	
ADAM20P1	FW:	GGAAACTGCCAAGGCTTGGG	417
	RV:	GGTCTCAGATGGAGATGAGGAAC	
BCKDHB	FW:	AAGCCTCTCCCTCTCAGCCT	375
	RV:	AACTGGCTTATCTCTTCTCCCTCC	
NTHL1	FW:	CAGCACCTGTCTCTGAGTGG	345
	RV:	CCCTGTCTTTCAGAGCAAGGTG	

Online Table V. Characterization of action-potentials recordings from isogenic WT and TBX5KO iPSCs-derived cardiomyocytes. Results are provided as mean ± SD. Maximal diastolic potential (MDP; mV), action potential amplitude (APA; mV), overshoot (mV), upstroke velocity (V/sec), and action potential duration (APD)50, APD70 and APD90 (the time intervals required to reach 50%, 70% and 90% of repolarization).

	WT-CMs		**TBX5KO-CMs**	
	Ventricular (n=18)	Atrial (n=4)	Ventricular (n=16)	Atrial (n=3)
MDP (mV)	-62.5 ± 6.1	-60.3 ± 3.8	-63.1 ± 4	-60.2 ± 4.7
APA(mV)	113 ± 7.9	103.8 ± 13.2	112.6 ± 6.8	102.2 ± 6.9
Overshoot (mV)	50.5 ± 7.1	39.2 ± 10	49.4 ± 7.3	42 ± 3.7
Upstroke Velocity (mV)	13.6 ± 3.8	19.3 ± 8.2	13 ± 3.5	12.2 ± 4.1
APD50 (mV)	263.4 ± 80.7	159 ± 58.7	297.6 ± 99.9	153.8 ± 35.2
APD70 (mV)	308.2 ± 96.5	190.6 ± 70.7	353.4 ± 111.2	223.6 ± 43.4
APD90 (mV)	337.5 ± 103.6	226.2 ± 88.5	384.4 ± 114.9	270 ± 49.8
Cycles per minute	53.9 ± 18.5	55.8 ±33.4	51 ± 20.3	51.6 ± 8.3