Journal Pre-proof

Genome Sequencing in a Genetically Elusive Multi-Generational Long QT Syndrome Pedigree Identifies a Novel LQT2-Causative Deeply Intronic *KCNH2* Variant

Kathryn E. Tobert, BA, David J. Tester, BS, Wei Zhou, MD, Carla M. Haglund-Turnquist, John R. Giudicessi, MD, PhD, Michael J. Ackerman, MD, PhD

PII: S1547-5271(22)00121-7

DOI: https://doi.org/10.1016/j.hrthm.2022.02.004

Reference: HRTHM 9146

To appear in: *Heart Rhythm*

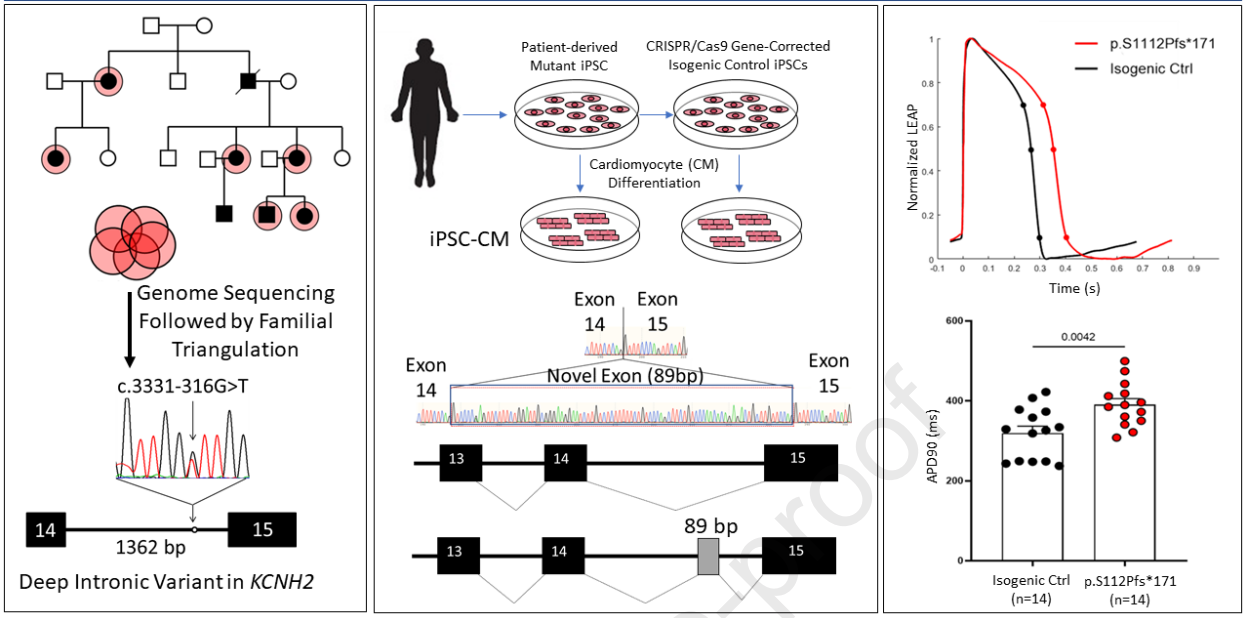
Received Date: 8 December 2021

Revised Date: 31 January 2022

Accepted Date: 1 February 2022

Please cite this article as: Tobert KE, Tester DJ, Zhou W, Haglund-Turnquist CM, Giudicessi JR, Ackerman MJ, Genome Sequencing in a Genetically Elusive Multi-Generational Long QT Syndrome Pedigree Identifies a Novel LQT2-Causative Deeply Intronic *KCNH2* Variant, *Heart Rhythm* (2022), doi: https://doi.org/10.1016/j.hrthm.2022.02.004.

Genome Sequencing in a Genetically Elusive Multi-Generational Long QT Syndrome Pedigree Identifies a Novel LQT2-Causative Deeply Intronic *KCNH2* Variant



Genome Sequencing in a Genetically Elusive Multi-Generational Long QT Syndrome Pedigree Identifies a Novel LQT2-Causative Deeply Intronic *KCNH2* Variant

Kathryn E. Tobert, BA[1]*, David J. Tester, BS[1]*, Wei Zhou, MD[1], Carla M. Haglund-Turnquist,[1] John R. Giudicessi, MD, PhD[1,2], and Michael J. Ackerman, MD, PhD[1]

***Ms. Tobert and Mr. Tester are co-equal first authors.**

Running Title: Tobert LQT2-Causative Deep Intronic *KCNH2* Variant

Institutional affiliations: [1]Departments of Cardiovascular Medicine (Division of Heart Rhythm Services, Windland Smith Rice Genetic Heart Rhythm Clinic), Pediatric and Adolescent Medicine (Division of Pediatric Cardiology), and Molecular Pharmacology & Experimental Therapeutics (Windland Smith Rice Sudden Death Genomics Laboratory), Mayo Clinic, Rochester, MN. [2]Departments of Cardiovascular Medicine (Clinician-Investigator Training Program), Mayo Clinic, Rochester, MN.

DISCLOSURES

Dr. Ackerman is a consultant for Abbott, ARMGO Pharma, Boston Scientific, Bristol Myers Squibb, Daiichi Sankyo, Invitae, LQT Therapeutics, Medtronic, and UpToDate. Dr. Ackerman and Mayo Clinic are involved in an equity/royalty relationship with AliveCor, Anumana, and Pfizer. However, none of these entities contributed to this study in any manner. The other authors report no conflicts.

Correspondence:
Michael J. Ackerman, M.D., Ph.D.
Mayo Clinic Windland Smith Rice Sudden Death Genomics Laboratory
Guggenheim 501, Mayo Clinic, Rochester, MN 55905
ackerman.michael@mayo.edu
@MJAckermanMDPhD

Word count of manuscript: 4752 words

ABSTRACT

Background: Most of long QT syndrome (LQTS) stems from pathogenic variants in *KCNQ1*, *KCNH2*, or *SCN5A*. However, ~10-20% of LQTS index cases remain genotype-negative.

Objective: Here, we identified and characterized functionally a novel LQTS genetic substrate in a multi-generational, "genotype-negative" LQTS pedigree.

Methods: The patient was a 40-year-old female with a history of syncope, seizures, ventricular fibrillation, and a family history of LQTS and sudden death. Commercial genetic testing of all LQTS-causative genes was negative. Genome sequencing was performed on 6 affected family members. Patient-specific and CRISPR/Cas9 "gene-corrected" isogenic control induced pluripotent stem cell-derived cardiomyocytes (iPSC-CMs) were generated.

Results: No ultra-rare, nonsynonymous heterozygous variants co-segregated among the 6 LQTS phenotype-positive individuals. Instead, a deep intronic *KCNH2* variant (c.3331-316G>T) was present in all affected individuals. RT-PCR analysis of patient-specific iPSC-CM-derived RNA revealed that c.3331-316G>T creates a novel 89 base-pair exon that results in a frame-shift variant (p.S1112Pfs*171). The action potential duration (APD_{90}) was significantly longer in p.S1112Pfs*171-iPSC-CMs (602.4 ± 12.2 ms, n=70) compared to isogenic control iPSC-CMs (425.7 ± 9.3 ms, n=61, p<0.0001). Further, the field potential duration (FPD) was significantly longer in p.S1112Pfs*171-iPSC-CMs (358.9 ± 7.7 ms, n=65) compared to isogenic control iPSC-CMs (282.2 ± 10.8 ms, n=51, p<0.0001).

Conclusions: A novel deep intronic *KCNH2* variant was identified in a multi-generational, genetically elusive LQTS pedigree. The iPSC-CMs establish that the variant is the monogenetic

cause for this family's LQTS. Deep intronic variants within the two most common LQTS-susceptibility genes should be considered in patients with seemingly, genetically elusive LQTS.

Key Words: CRISPR/Cas9, Genetic testing, iPSC-cardiomyocytes, HERG/Kv11.1, Long QT syndrome, Potassium channel

INTRODUCTION

Long QT syndrome (LQTS) is typically an autosomal dominant genetic heart disease characterized by a prolonged QT interval on the 12-lead electrocardiogram (ECG). Patients with LQTS can present with arrhythmic syncope/seizures, sudden cardiac arrest (SCA), or sudden cardiac death (SCD) and often these events occur following a precipitating event like exercise, auditory trigger, or extreme emotion[1]. LQTS occurs in ~1 in 2000 people[2] and for those who are untreated, it is estimated that there is a 50% 10-year mortality in the highest-risk subset[3].

Approximately 90% of LQTS cases stem from pathogenic variants in 17 known LQTS-susceptibility genes, with almost all of those variants occurring in 3 major genes, *KCNQ1*, *KCNH2*, and *SCN5A*[4]. Loss-of-function (LOF) in either *KCNQ1*-encoded I_{Ks} ($K_v7.1$) potassium channel (LQT1, ~35-40%) and or *KCNH2*-encoded I_{Kr} ($K_v11.1$) potassium channel (LQT2, ~30-35%), and gain-of-function variants in *SCN5A*-encoded I_{Na} ($Na_v1.5$) sodium channel (LQT3, ~5-10%) underlie the pathological prolongation of the ventricular cardiomyocyte's action potential duration (APD)[4]. Variants in minor LQTS-causative genes are responsible for < 5% of cases[4]. However, some of these previously implicated LQTS-causative genes have been demoted recently to 'limited evidence' or 'refuted evidence' status in terms of their disease-gene association[4, 5]. Following comprehensive LQTS genetic testing, ~10% of patients remain genetically elusive as to the cause of their LQTS and are classified as "genotype-negative" LQTS.

In recent years, exome or genome sequencing (GS) has been utilized for individuals and families who have a clinically irrefutable phenotype with a negative genetic test result in order to reveal either a novel causative gene or potentially a novel, previously unexplored pathogenic substrate within one of the already known disease-causative genes[6-11]. Here, we performed GS

on a multigenerational, genetically elusive pedigree to identify their specific genetic substrate and explain their autosomal dominant LQT2-like phenotype. Patient-specific induced pluripotent stem cell-derived cardiomyocytes (iPSC-CMs) were developed and used to perform disease modeling studies alongside CRISPR/Cas9-gene variant corrected, isogenic control iPSC-CMs.

METHODS

Pedigree with Autosomal Dominant LQTS

A family with autosomal dominant LQTS was referred to the Mayo Clinic Windland Smith Rice Sudden Death Genomics Laboratory after commercially available genetic testing for LQTS was negative (**Figure 1**). Following written informed consent for this Mayo Clinic IRB (1216-97) approved study, blood was collected from participating family members and genomic DNA was isolated.

Genome Sequencing (GS)

Genomic DNA, derived from 6 phenotype-positive family members, was subjected to genome sequencing (GS, **Figure 2**) at Mayo Clinic's Medical Genome Facility in Rochester, Minnesota as previously described[9, 10]. Following GS, single nucleotide variants (SNVs) and insertion/deletions (INDELs) were filtered to identify variants which followed an autosomal dominant inheritance pattern using Ingenuity Variant Software (Qiagen, Redwood City, CA). All variants were first filtered for a call quality score ≥ 20 and a read-depth of ≥ 10. Only rare (minor allele frequency [MAF] < $5x10^{-5}$ in gnomAD[12]) non-synonymous (NS) variants present in all 6 affected family members were considered initially. However, deep intronic variants within the three major LQTS genes (*KCNQ1*, *KCNH2*, and *SCN5A*) were also considered. Deep intronic variants were assessed for their ability to create novel splice-sites using the Alamut Visual software (Interactive Biosoftware, version 2.15) splicing prediction module that incorporates

SpliceSiteFinder-like[13], MaxEntScan[14], GeneSplicer[15], and NNSPLICE[16]. Additionally, the freely available tool, SpliceAI, was used to determine the predicted effect of potential splice-altering variants[17]. The population frequency threshold of $< 5x10^{-5}$ was based on an LQTS disease prevalence of 1 in 2000, 50% penetrance, and 1% maximum allelic contribution in LQTS[18]. The candidate variant was confirmed by Sanger sequencing.

Reprogramming Peripheral Blood Mononuclear Cells (PBMCs) into Induced Pluripotent Stem Cells (iPSCs) and Quality Control

PBMCs were reprogrammed into iPSCs, which then underwent quality control studies, including karyotyping, pluripotent marker staining, and confirmatory sequencing using previously published methods[19, 20].

Mutation Correction of *KCNH2's* Deep Intronic Variant Using CRISPR/Cas9 Technology

Genome editing of cell lines was conducted under contract by Applied Stem Cell. The c.3331-316G>T-*KCNH2* mutation in KCNH2-p.S1112Pfs*171-derived iPSC cells was corrected using CRISPR/Cas9 technology. Briefly, two guide RNAs (gRNAs) were designed, validated *in vitro*, and one candidate gRNA (**Supplemental Table 1**) was chosen for genome editing on patient iPSC line based on specificity score, cutting efficiency, and off-target profile. A single-stranded oligodeoxynucleotide (ssODN) was designed to be used as a repair template and a silent mutation in gRNA binding site was introduced to the ssODN to prevent re-cutting (**Supplemental Table 1**). The *KCNH2* patient-specific iPSC line was transfected with gRNA construct and ssODN using a Neon system and transfected iPSCs were subjected to puromycin selection. Single-cell colonies were picked for genotyping, and two clones with mutation correction, creating the patient's isogenic control iPSCs, were expanded for further studies.

Differentiation of iPSCs into iPSC-CMs

iPSCs were differentiated into cardiomyocytes as described previously[21].

Real-Time Quantitative Polymerase Chain Reaction

Following physical enrichment of iPSC-CMs from beating clusters, total RNA was isolated from mutant iPSC-CM clones and isogenic control iPSC-CMs using the mirVana™ miRNA isolation kit (Invitrogen, AM1560) according to the manufacturer's instruction. The cDNA was prepared from 60 ng of RNA using the QuantiTect Reverse Transcription Kit (Qiagen, 205311) according to the manufacturer's instruction. Standard PCR reaction was performed using a forward primer located in KCNH2 exon 13 and a reverse primer located in KCNH2 exon 15. The PCR products were size separated using a 2% agarose gel. The gel bands were excised, and the DNA purified using the QIAquick Gel Extraction Kit (Qiagen, 28704). The DNA derived from each band was Sanger sequenced.

Western Blot

iPSC-CM protein expression was measured through western blot using previously described methods. All primary and secondary antibodies used can be found in the **Supplemental Table 2**.

Immunocytochemistry

Immunocytochemistry was performed in iPSC-CMs using previously published methods[19]. All primary and secondary antibodies used for IF are listed in **Supplemental Table 2**.

Optical Action Potentials Measurement Using Voltage Dye on iPSC-CMs

iPSC-CMs were cultured on 35mm glass bottom dishes (MatTek, P35G-1.5-10-C) that were pre-coated with Matrigel at 37°C, 5% CO_2. On the day of imaging, cells were incubated at 37°C, 5% CO_2 for 20 minutes in Tyrode's solution containing a fluorescent voltage sensitive dye, 0.125µL FluoVolt dye and 1.25µL PowerLoad (FluoVolt Membrane Potential kit, ThermoFisher, cat#: F10488). iPSC-CMs were then washed three times with pre-warmed Tyrode's solution, and a final 2 ml of Tyrode's solution was added to these dishes. During imaging, the dishes were kept

in heated 37°C stage-top environment chamber supplied with 5% CO_2. Imaging of voltage-indicated cellular action potential duration (APD) was taken under a 40X-water objective using a Nikon Eclipse Ti light microscope. Time-lapse videos of multiple, individual beating iPSC-CMs, paced at 1Hz were recorded at a speed of 20 milliseconds per frame for 20 seconds at 15% LED power. Single regions of interest were selected for every beating iPSC-CM captured in the recordings. The raw data was exported to Excel software (Microsoft, Redmond, WA) and then analyzed with an "in-lab" developed Excel-based program[22].

Microelectrode Array (MEA) Measurement

iPSC-CMs were dissociated and seeded at 50,000 cells per well on 48-well Biocircuit MEA plate (Axion BioSystems, Inc. M768-BIO-48) pre-coated with Matrigel. Cells were cultured in a humidified incubator at 37°C and 5% CO_2 for 7-10 days after dissociation, and media was changed every 2 days. Media was changed 4-24 hours before starting experiment and the MEA plate was placed in Maestro MEA device (Axion BioSystems, Inc.) with automatically adjusted and controlled environment (37°C and 5% CO_2), and equilibrated for 2-5 mins. Field potential duration (FPD) was recorded first. MEA-based APD was recorded after local extracellular action potential (LEAP) induction using the AxIS Navigator software (Axion BioSystems, Inc). Data analysis was done using Cardiac Analysis Tool (Axion BioSystems, Inc. Version 3.1.4).

Statistical Analysis

All data points are shown as the mean value and bars represent the standard error of the mean. A Student's t-test was performed using GraphPad Prism 9 to determine statistical significance between two groups. A $P<0.05$ was considered significant.

RESULTS

Clinical Description

The proband with LQTS was a highly symptomatic 40-year-old female with a history of sudden syncope and seizures, and multiple ventricular tachycardia (VT)/ventricular fibrillation (VF)-terminating ICD therapies and a QTc of 492 ms (**Figure 1A and 1B**). Her first concerning episode occurred at age 29 years with a sudden faint without any obvious stressor. She was not formally diagnosed at this time, but placed on the beta blocker, atenolol, elsewhere. Subsequently, she presented to the emergency department on multiple occasions because of sudden falls and seizures. On one of her last emergency department visits, she displayed self-terminating VT/VF while hooked up to a cardiac monitor. LQTS was suspected and she received an ICD immediately and continued on atenolol (patient was treated elsewhere). The patient had undergone a bilateral cardiac sympathetic denervation and continued to experience appropriate VT/VF-terminating ICD shocks. Her precipitating events were often associated with emotional stress. Genetic testing for LQTS was pursued and ultimately returned a negative result, leaving the patient's LQTS phenotype genetically unexplained. A pedigree analysis was performed that identified 7 additional affected family members, including a maternal uncle who died suddenly at age 40 years (**Figure 1A**). Following the negative genetic test result, GS was pursued on 6 phenotype-positive family members (**Figure 2A**).

Genome Sequencing Discovery of a Deep Intronic Variant in *KCNH2*

No ultra-rare, heterozygous nonsynonymous variants in any protein-coding gene of the human genome co-segregated among the 6 LQTS phenotype-positive individuals. However, 2 deep intronic *KCNH2* (ENST00000262186.10, NM_000238.4) variants (c.76+199T>G, hg38 Chr7:150977639-A-C and c.3331-316G>T, hg38 Chr7:150945830-C-A) were present in all

affected individuals (**Figure 2A**). Both deep intronic variants were absent in gnomAD genomes dataset. Although *in silico* splice-site prediction tools did not predict abnormal splicing associated with c.76+199T>G, these tools predicted the c.3331-316G>T variant to cause improper RNA splicing of *KCNH2* through activation of a cryptic donor site (**Supplemental Table 3**). The c.3331-316G>T variant resides within the 1362 base pair (bp) intronic region between exon 14 and exon 15 of the *KCNH2* gene with the nucleotide substitution occurring 316 nucleotides prior to exon 15 (**Figure 2B**). According to the splice-site prediction tools, the nucleotide substitution of a G for T at position c.3331-316 leads to a new cryptic donor splice-site occurring 6 nucleotides upstream of the substitution at position c.3331-321 (**Supplemental Table 3**, **Figure 2C**). Notably, our splice-site prediction analysis of the normal 1362 bp intronic sequence between exon 14 and 15 also revealed a predicted cryptic acceptor site occurring at nucleotide position c.3331-410 with highly favorable prediction scores using 4 different splicing prediction tools (**Supplemental Table 3**, **Figure 2C**). Additionally, SpliceAI also predicted these precise boundaries of a novel cryptic exon with strong scores (scale 0-1). Specifically, SpliceAI predicted that c.3331-316G>T would result in the creation of a new cryptic donor site (score 0.46) 6 nucleotides upstream of position c.3331-316 as well as a new cryptic acceptor site (score 0.52) occurring 94 nucleotides upstream of this position, thus predicting the inclusion of an 89-nucleotide cryptic exon.

Generation of Patient-Specific iPSC-CMs

Patient-specific iPSC-CM lines/clones were generated from PBMCs. We preferentially used multiple clones of patient origin iPSCs to reduce the error from the heterogeneous nature between clones from the same patient sample, and overall did not see any significant differences in sibling clones. All clones and lines were characterized with karyotyping (**Figure 3A**),

pluripotent marker staining while in undifferentiated state (**Figure 3B**), and 3 germ layer immunofluorescence imaging (**Figure 3C**). The *KCNH2* intronic variant was confirmed by Sanger sequencing.

Characterization of Novel, Intronic Variant-Produced *KCNH2* Transcript

To determine if the c.3331-316G>T, located between exon 14 and exon 15, impacted splicing of *KCNH2*, reverse transcription polymerase chain reaction (RT PCR) was performed using a forward PCR primer located in exon 13 and a reverse PCR primer in exon 15 (**Figure 4A**). The RT-PCR reaction on RNA isolated from the patient's mutant iPSC-CMs created two bands, the predicted wild-type 255 bp band and a novel band of ~350 bp in size which was not observed in the isogenic control (**Figure 4A**). Sanger sequencing of the excised gel bands revealed that c.3331-316G>T creates a novel 89 base pair-containing exon inserted between canonical exons 14 and 15 of *KCNH2* (**Figure 4B and 4C**) resulting in a frameshift variant annotated as p.S1112Pfs*171, precisely as predicted by the *in silico* splice-site prediction tools.

Western Blot and Immunofluorescence Detecting Marked Decrease in hERG Expression

To evaluate the impact of the patient's variant on hERG expression and localization, we performed a Western blot and immunostaining using an N-terminal hERG antibody. We observed a single band at ~100 kDa (**Figure 5A**) in both the patient and isogenic control iPSC-CM-derived lysate, and hERG expression (relative to Vinculin) was reduced significantly in the patient's iPSC-CMs compared to the isogenic control ($p=0.025$, **Figure 5B**). Confocal study showed hERG protein in the patient's iPSC-CMs was retained in ER significantly compared with isogenic control ($p<0.0001$ value, **Figure 5C**).

Cardiac Action Potential Duration (APD) and Field Potential Duration (FPD) Measurements

Overall, the patient's iPSC-CMs had a significantly longer FluoVolt-derived APD_{90} values compared to the isogenic control iPSC-CMs (602.4 ± 12.2 ms, n=70 vs. 425.7 ± 9.3 ms, n=61; p<0.0001; **Figure 6A** and **6B**) thereby further strengthening this deep intronic variant's association with disease. Similarly, the FPD (358.9 ± 7.7 ms, n=65 vs. 282.2 ± 10.8 ms, n=51; p<0.0001; **Figure 6C** and **6D**) and the LEAP-based APD_{90} (390.9 ± 14.77 ms, n=14 vs. 319.6 ± 17.25 ms, n=14; p=0.0042; **Figure 6E and 6F**) were both significantly longer in the patient-specific iPSC-CMs compared to the variant-corrected, isogenic control lines.

DISCUSSION

Here, we present a GS-derived and patient-specific, re-engineered heart cell model-based functional characterization of a deep intronic *KCNH2* variant and provide evidence that this is the LQT2 disease-causing variant in the multi-generational pedigree with heretofore genetically elusive LQTS. The variant of interest, c.3331-316G>T, resides between canonical exons 14 and 15 of *KCNH2*, and impacts RNA splicing through the activation of a cryptic donor splice site. This improper splicing creates a novel 89 bp exon, resulting in a frameshift which extends the *KCNH2*-encoded Kv11.1/hERG protein sequence by 124 amino acids (from 1159 amino acids to 1283 amino acids). By itself, the identification of c.3331-316G>T would amount to only a variant of uncertain significance designation by American College of Medical Genetics (ACMG) guidelines through satisfying just the PM2 (absence in controls) criteria[23]. However, given the proven effect on splicing, the PVS1 (null variant) criteria can be applied in addition to PS3 (well established *in vitro* functional studies supporting a damaging effect), PM2, and PP1 (cosegregation with disease in multiple affected family members). Thus the c.3331-316G>T

variant satisfies the ACMG guideline criteria for designation as a disease-causative pathogenic variant[23].

Cascade genetic testing in LQTS is crucial in the diagnosis and evaluation of families with congenital LQTS. Unveiling the underlying genetic cause of a clear LQTS phenotype and designing a phenotype-based/genotype-guided treatment plan is essential in the management of LQTS and the prevention of LQTS-triggered cardiac events especially SCD. Currently, commercial genetic testing for LQTS only evaluates the exons (i.e. protein-coding regions), 10-20 bp of adjacent intronic sequences, and selected non-coding variants of known LQTS genes. Unfortunately, approximately 10% of LQTS remains unexplained following such testing. Thus, recently, GS has moved from being an important scientific tool to a powerful diagnostic tool in identifying a monogenetic cause for families with phenotype-positive but genotype-negative genetic heart disease[6, 9].

In silico splice site prediction tools aid in the analysis of variants of unknown significance (VUS). Prior to the Human Genome Project, splice site prediction tools were used to identify intron-exon boundaries, however, these tools have evolved from the development of consensus sequences to determining the transcriptional impact of VUS that reside within known splice site regions[24]. Over the years, these tools have proven their utility in the identification of deleterious genetic heart disease-causative variants, including LQTS and hypertrophic cardiomyopathy (HCM)[6, 25].

Previously, Crotti *et al*. described a *KCNH2* branch point mutation which caused aberrant splicing and resulted in LQT2[25]. The proband had exhibited several LQT2-like symptoms such as breakthrough cardiac events occurring during sleep or after an auditory stimulus (like an alarm clock). There was a family history of SCD, and the proband and the proband's mother had a

prolonged QTc and notched T waves in the precordial leads. Following a negative genetic test of the five major LQTS genes, linkage analysis for *KCNH2* was pursued because the phenotype closely resembled LQT2. An intronic variant, *KCNH2* c.2399-28A>G, was discovered and was predicted to cause abnormal RNA splicing by disrupting the intronic acceptor site, leading to the retention of an upstream intronic sequence. Ultimately, Crotti et al. revealed that intronic variants within known LQTS-causative genes can be the underlying genetic cause for a family's LQTS, and further bolstered the notion that GS can potentially provide powerful results for a "genotype-negative" case of LQTS.

More recently, deeper intronic variants, beyond the so-called branch points, have been identified as the pathogenic cause of other cardiomyopathies, namely, HCM. Bagnall *et al*. performed GS on 58 unrelated patients with HCM, and discovered pathogenic variants in 9% of genetically elusive HCM when the genetic analysis was extended to deeper intronic regions of the known, sarcomeric HCM genes[6]. Three deep intronic *MYBPC3* variants were found in 4 families, and these variants created novel pseudoexons. Nearly every proband in this study had a clear HCM phenotype (either symptomatic or exhibited septal hypertrophy), and most had a family history of cardiomyopathy and/or sudden cardiac death (39/58; 67%). Additionally, Janin et al. identified pathogenic deep intronic *MYBPC3* variants in 5% of their cohort of 93 unrelated HCM patients in which no pathogenic/likely pathogenic variants had been detected in 48 cardiomyopathy-causing genes[26]. In 2020, Lopes et al. identified four deep intronic *MYBPC3* variants (c.906-36G>A, c.1224-21A>G, c.1224-52G>A. and c.1224-80G>A) in 24/1644 (1.5%) unrelated patients with HCM overall and 2.2% (24/546) of their otherwise mutation-negative cohort[27]. These studies demonstrate that when GS is pursued and is successful in unveiling the

disease-causing variant, the proband and affected family members must be well clinically evaluated, and a phenotype must be established.

In our case, given a negative genetic test result, a search for deep intronic variants in the known LQTS-causative genes was compelled because the patient's phenotype so strongly suggested LQT2 resulting in the identification of novel deep intronic disease-causing variant in *KCNH2*. In addition, iPSC-CMs have proven to be extremely valuable for modeling several types of genetic heart diseases and helping to prove a variant's irrefutable pathogenicity[28]. iPSC-CMs provide a patient-specific model that is free from environmental factors, allowing for investigation of the genetic variant of interest as the monogenetic cause of a patient's disease phenotype. Here, iPSC-CMs derived from our patient along with CRISPR/Cas9 gene-edited variant corrected isogenic control line established that the deep intronic variant was indeed sufficient to cause APD prolongation and the LQTS phenotype in the pedigree. The iPSC-CM model also allowed us to determine and confirm the precise splicing-error that was predicted by the *in silico* tools.

CONCLUSIONS

Herein, a novel deep intronic *KCNH2* variant was identified in a multi-generational, previously genetically elusive LQTS pedigree. The patient-derived and isogenic control iPSC-CMs establish that the deeply intronic variant created the frameshift variant, p.S1112Pfs*171, in the *KCNH2*-encoded Kv11.1/hERG channel as the family's monogenetic basis for their LQT2. When evaluating "genotype-negative" LQTS, deep intronic variants within the two most common LQTS-causative genes should be considered especially when the phenotype so strongly points towards either LQT1 or LQT2.

ACKNOWLEDGMENTS

The authors would like to thank all the study subjects for participating in this study and enabling this discovery to be made.

FUNDING SOURCES

This work was supported by the Mayo Clinic Windland Smith Rice Comprehensive Sudden Cardiac Death Program.

REFERENCES

1. Schwartz PJ, Ackerman MJ, Antzelevitch C, et al. Inherited cardiac arrhythmias. Nat Rev Dis Primers 2020;6:58.
2. Schwartz PJ, Stramba-Badiale M, Crotti L, et al. Prevalence of the congenital long-QT syndrome. Circulation 2009;120:1761-1767.
3. Ackerman MJ. Cardiac channelopathies: it's in the genes. Nat Med 2004;10:463-464.
4. Giudicessi JR, Wilde AAM, Ackerman MJ. The genetic architecture of long QT syndrome: a critical reappraisal. Trends Cardiovasc Med 2018;28:453-464.
5. Adler A, Novelli V, Amin AS, et al. An international, multicentered, evidence-based reappraisal of genes reported to cause congenital long QT syndrome. Circulation 2020;141:418-428.
6. Bagnall RD, Ingles J, Dinger ME, et al. Whole genome sequencing improves outcomes of genetic testing in patients with hypertrophic cardiomyopathy. J Am Coll Cardiol 2018;72:419-429.
7. Crotti L, Johnson CN, Graf E, et al. Calmodulin mutations associated with recurrent cardiac arrest in infants. Circulation 2013;127:1009-1017.
8. Higgins EM, Bos JM, Mason-Suares H, et al. Elucidation of MRAS-mediated Noonan syndrome with cardiac hypertrophy. JCI insight 2017;2:e91225-e91225.
9. Clemens DJ, Tester DJ, Marty I, Ackerman MJ. Phenotype-guided whole genome analysis in a patient with genetically elusive long QT syndrome yields a novel TRDN-encoded triadin pathogenetic substrate for triadin knockout syndrome and reveals a novel primate-specific cardiac TRDN transcript. Heart Rhythm 2020;17:1017-1024.

10. Shehab O, Tester DJ, Ackerman NC, Cowchock FS, Ackerman MJ. Whole genome sequencing identifies etiology of recurrent male intrauterine fetal death. Prenatal Diagnosis 2017;37:1040-1045.

11. Altmann HM, Tester DJ, Will ML, et al. Homozygous/compound heterozygous triadin mutations associated with autosomal-recessive long-QT syndrome and pediatric sudden cardiac arrest. Circulation 2015;131:2051-2060.

12. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 2020;581:434-443.

13. Zhang MQ. Statistical features of human exons and their flanking regions. Hum Mol Genet May 1998;7:919-932.

14. Yeo G, Burge CB. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. J Comput Biol 2004;11:377-394.

15. Pertea M, Lin X, Salzberg SL. GeneSplicer: a new computational method for splice site prediction. Nucleic Acids Res 2001;29:1185-1190.

16. Reese MG, Eeckman FH, Kulp D, Haussler D. Improved splice site detection in Genie. J Comput Biol 1997;4:311-323.

17. Jaganathan K, Kyriazopoulou Panagiotopoulou S, et al. Predicting splicing from primary sequence with deep learning. Cell 2019;176:535-548.e524.

18. Whiffin N, Minikel E, Walsh R,et al. Using high-resolution variant frequencies to empower clinical genome interpretation. Genet Med 2017;19:1151-1158.

19. O'Hare BJ, John Kim CS, Hamrick SK, Ye D, Tester DJ, Ackerman MJ. Promise and potential peril with lumacaftor for the trafficking defective type 2 long-QT syndrome-

causative variants, p.G604S, p.N633S, and p.R685P, using patient-specific re-engineered cardiomyocytes. Circ Genom Precis Med 2020;13:466-475.

20. Dotzler SM, Kim CSJ, Gendron WAC,et al. Suppression-replacement KCNQ1 gene therapy for type 1 long QT syndrome. Circulation 2021;143:1411-1425.

21. Burridge PW, Matsa E, Shukla P, et al. Chemically defined generation of human cardiomyocytes. Nat Methods 2014;11:855-860.

22. Estes SI, Ye D, Zhou W, et al. Characterization of the CACNA1C-R518C missense mutation in the pathobiology of long-QT syndrome using human induced pluripotent stem cell cardiomyocytes shows action potential prolongation and L-type calcium channel perturbation. Circ Genom Precis Med 2019;12:e002534.

23. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 2015;17:405-424.

24. Jian X, Boerwinkle E, Liu X. In silico tools for splicing defect prediction: a survey from the viewpoint of end users. Genet Med 2014;16:497-503.

25. Crotti L, Lewandowska MA, Schwartz PJ,et al.. A KCNH2 branch point mutation causing aberrant splicing contributes to an explanation of genotype-negative long QT syndrome. Heart Rhythm 2009;6:212-218.

26. Janin A, Chanavat V, Rollat-Farnier PA, et al. Whole MYBPC3 NGS sequencing as a molecular strategy to improve the efficiency of molecular diagnosis of patients with hypertrophic cardiomyopathy. Hum Mutat 2020;41:465-475.

27. Lopes LR, Barbosa P, Torrado M,et al. Cryptic splice-altering variants in MYBPC3 are a prevalent cause of hypertrophic cardiomyopathy. Circ Genom Precis Med 2020;13:e002905.

28. Musunuru K, Sheikh F, Gupta RM, et al. Induced pluripotent stem cells for cardiovascular disease modeling and precision medicine. A scientific statement from the American Heart Association. Circ Genom Precis Med 2018;11:e000043.

FIGURES

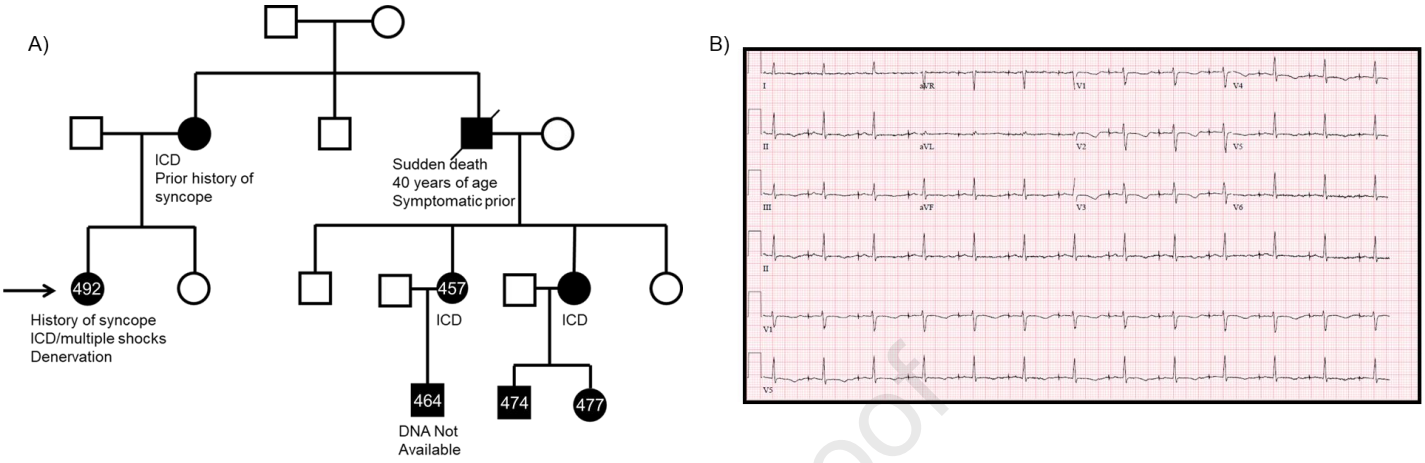


Figure 1. Autosomal dominant genetically elusive LQTS pedigree: Shown in Panel A is a multi-generational, genetically elusive LQTS pedigree. Open symbols (circles=female, squares=males) represent unaffected individuals. Black symbols represent affected family members. Numbers within the circles or squares represent the individual's QTc in milliseconds. Panel B shows a representative ECG from the index case (QTc=492 ms). No ECGs without atrial pacing were available.

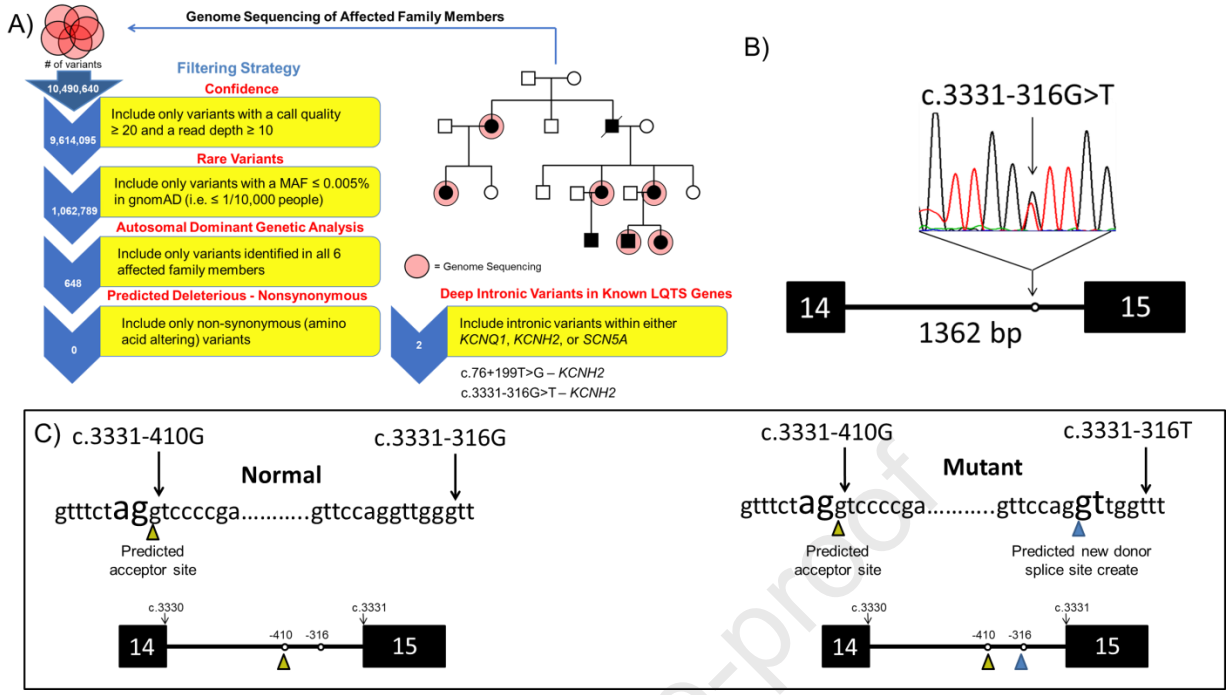


Figure 2. Genome sequencing and familial triangulation for the discovery of a novel putative disease-causing deep intronic *KCNH2* variant. Shown in panel A is a schematic representation of the variant filtering strategy used to identify potential candidate variants that segregate with the disease phenotype in the pedigree. Sanger sequencing chromatogram validating the heterozygous c.3331-316G>T variant is shown in panel B. Also shown is a schematic representation of the variant localized within the intronic region of *KCNH2* between exon 14 and exon 15. Panel C shows the schematic representation of predicted cryptic acceptor (green triangle) and donor (blue triangle) splice sites located within the intronic sequence occurring in both the normal (left) and mutant (right) allele.

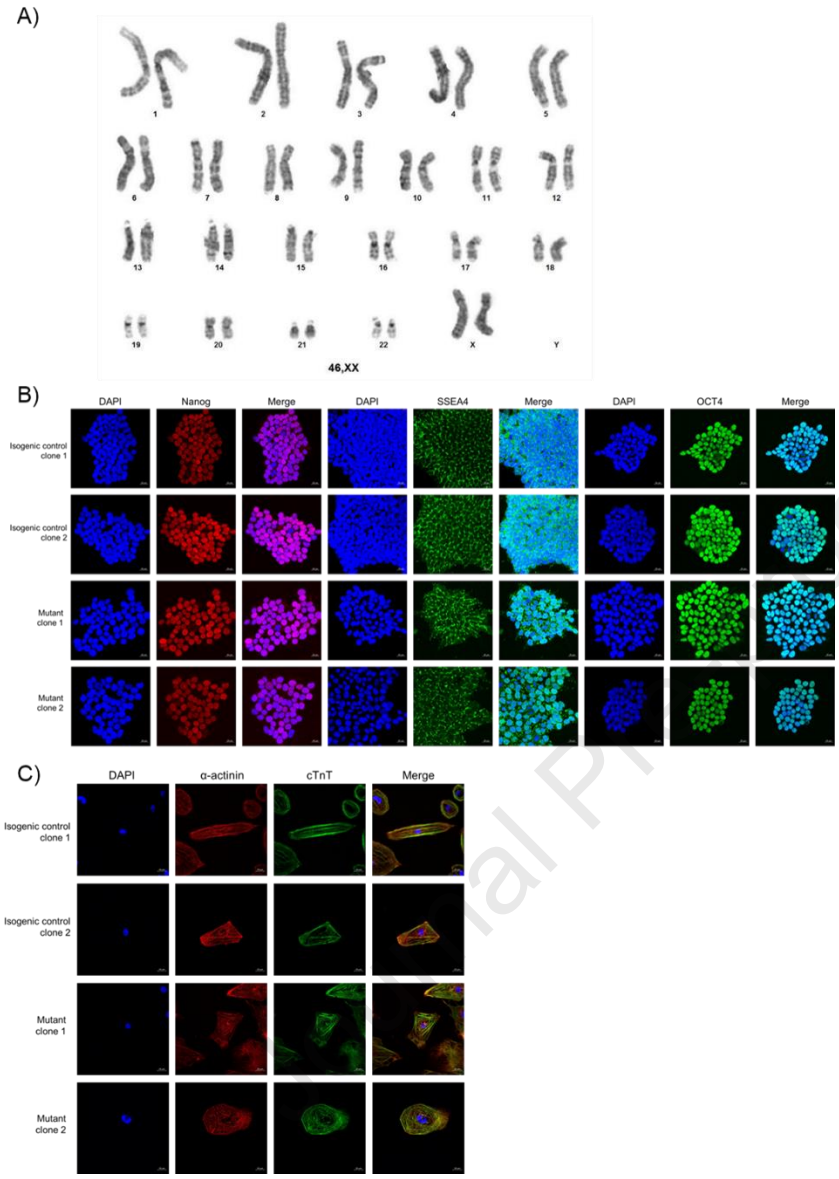
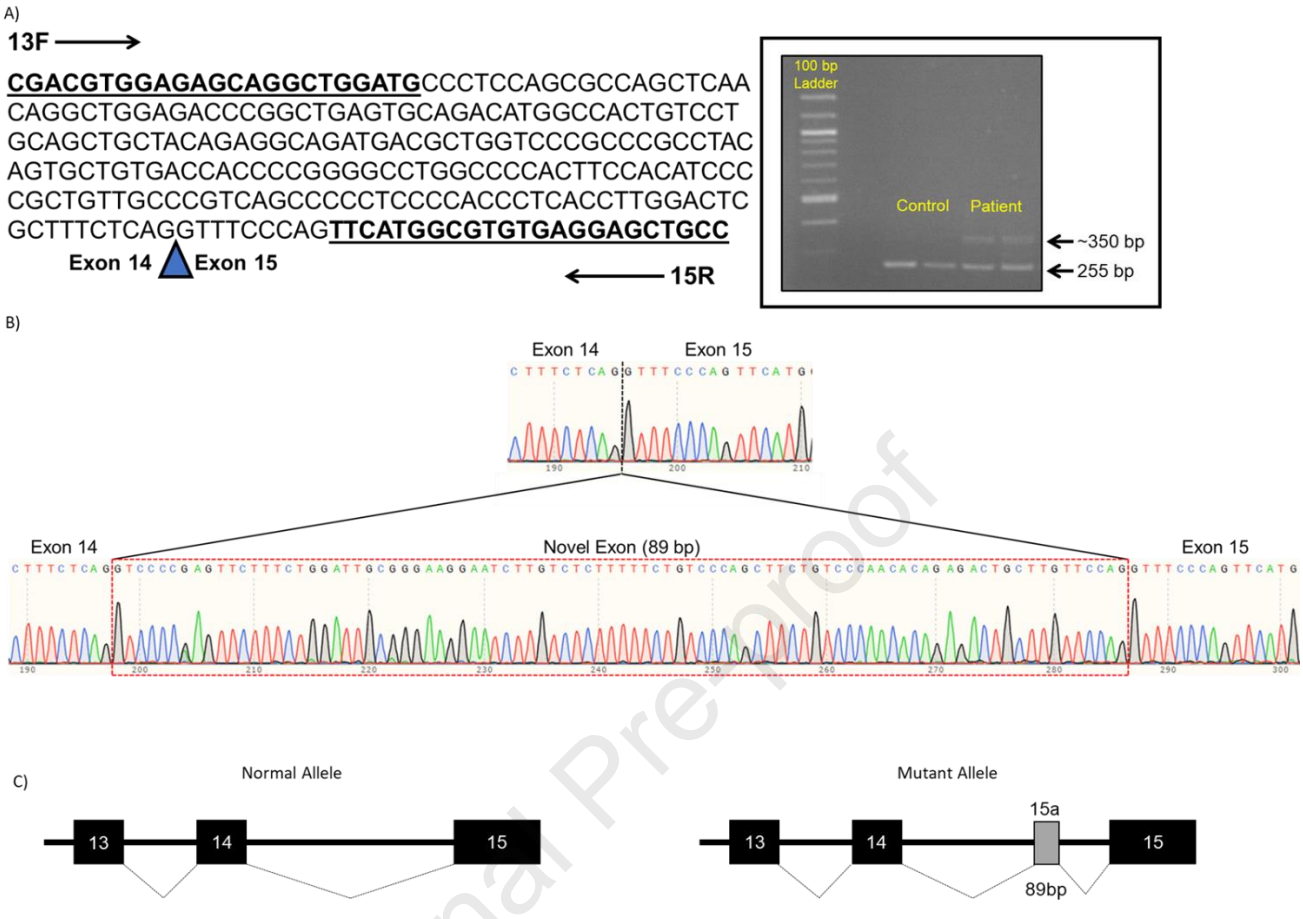


Figure 3. Generation of patient-specific iPSC-CMs. Shown in panel A is a normal female karyotype of the patient's iPSCs. Shown in panel B are representative confocal images of undifferentiated patient-specific mutant and isogenic control iPSCs demonstrating three pluripotent markers (NANOG, SSEA4, and OCT4). Scale bars equal 20μm. Following cardiac differentiation, iPSC-derived cardiomyocyte maturity was confirmed with cardiac specific markers (α-actinin and cTnT) as shown in panel C.



473

474 **Figure 4. Characterization of a novel intronic variant-produced *KCNH2* transcript.** Reverse
475 transcription polymerase chain reaction (RT-PCR) was performed using a forward PCR primer
476 located in exon 13 and a reverse PCR primer in exon 15. RT-PCR analysis of patient-specific
477 and isogenic control iPSC-CM derived RNA revealed that this deep intronic variant creates a
478 novel 89 bp exon that resides between the canonical exons 14 and 15 of *KCNH2;* resulting in a
479 frame-shift variant annotated as p.S1112Pfs*171. Panel A shows the location of the primers used
480 and an image of the agarose gel demonstrating the addition of a ~350 base pair (bp) band present
481 in the mutant iPSC-CMs that is absent in the isogenic control line. Panel C shows the Sanger
482 sequencing chromatogram of the isogenic control and mutant excised ~350 bp band which
483 reveals the inclusion of 89 nucleotides inserted between exon 14 and exon 15. Panel C is a

schematic representation illustrating the inclusion of a novel 89 bp alternative exon 15 (15a) in the mutant allele.

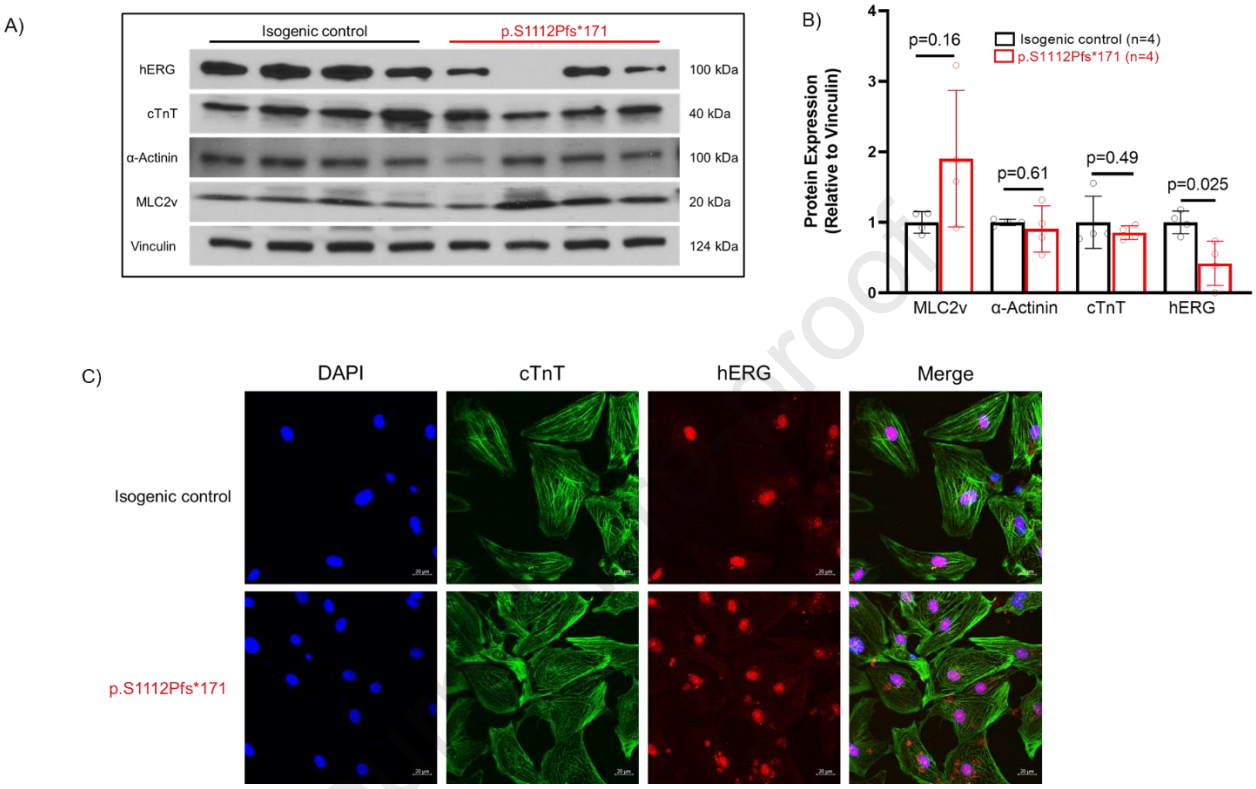


Figure 5. Western blot and immunostaining to evaluate hERG expression. Panel A is a Western blot showing expression of hERG, cTnT, α-actinin, MLC2v, and vinculin. Panel B shows protein expression levels of MLC2v, α-actinin, cTnT, and hERG, relative to vinculin, in the isogenic control and the p.S1112Pfs*171 iPSC-CMs. Compared to the isogenic control iPSC-CMs, there was a significant decrease in hERG expression in the mutant iPSC-CMs. Panel C shows immunofluorescence staining of the isogenic control and p.S1112Pfs*171 iPSC-CMs.

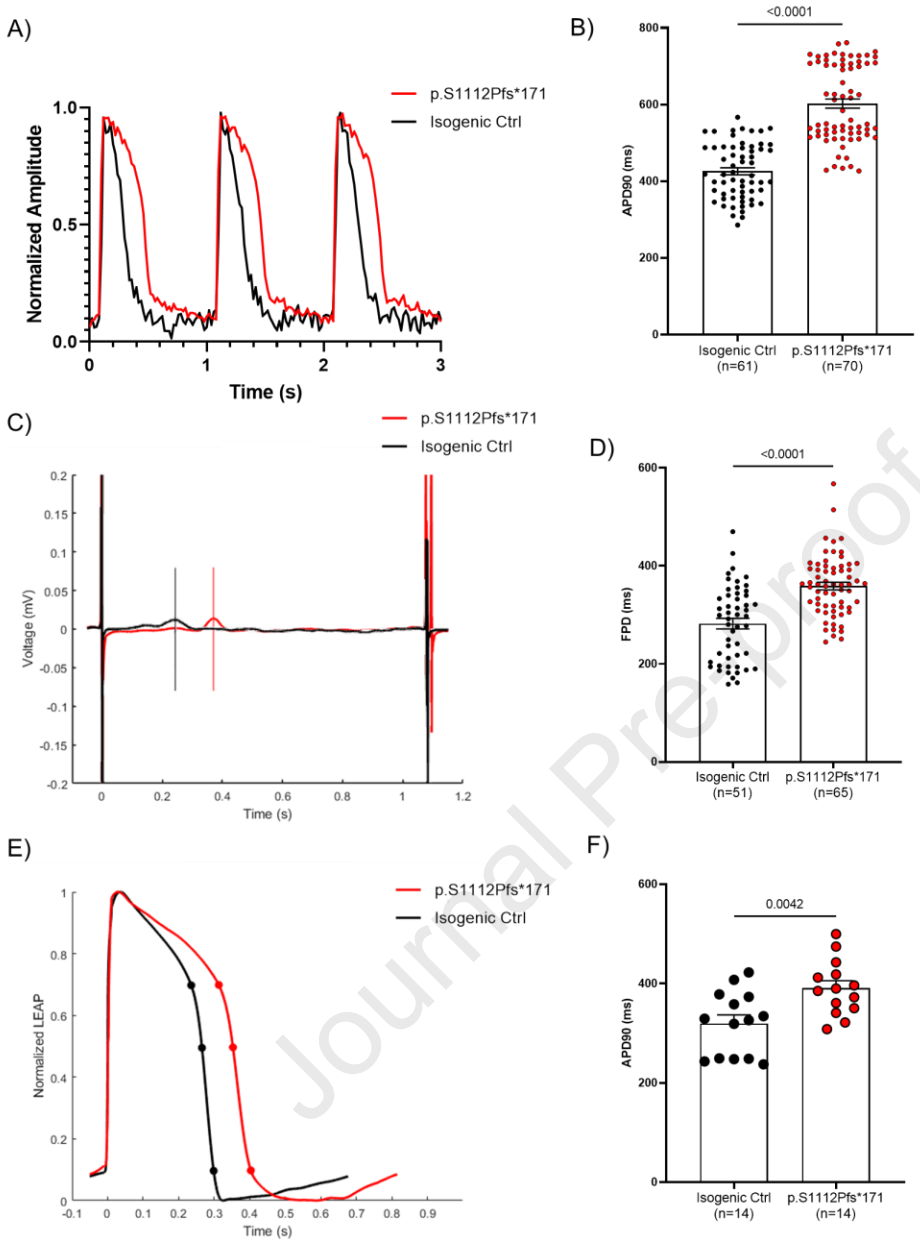


Figure 6. Action potential duration and field potential duration measurements. Shown is the action potential duration (APD) and field potential duration (FPD) measurements for the patient's mutant (p.S1112Pfs*171) iPSC-CMs (red) and the CRISPR/Cas9 gene-corrected isogenic control (black) iPSC-CMs. Panel A are the representative tracings of FluoVolt-based action potentials and in panel B are the average FluoVolt-based APD_{90} measurements. Panel C shows the representative microelectrode array (MEA)-based FPD tracings and in panel D are the

502 average MEA-based FPD measurements. Panel E are representative tracings of the local
503 extracellular action potential (LEAP)-based APD tracings and in panel F are the average LEAP-
504 based APD_{90} measurements.
505
506
507

Figure 1

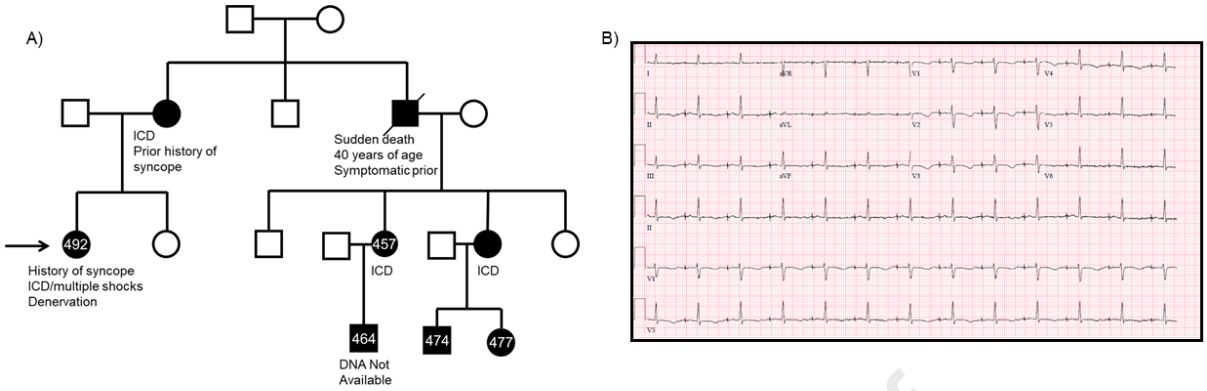


Figure 2

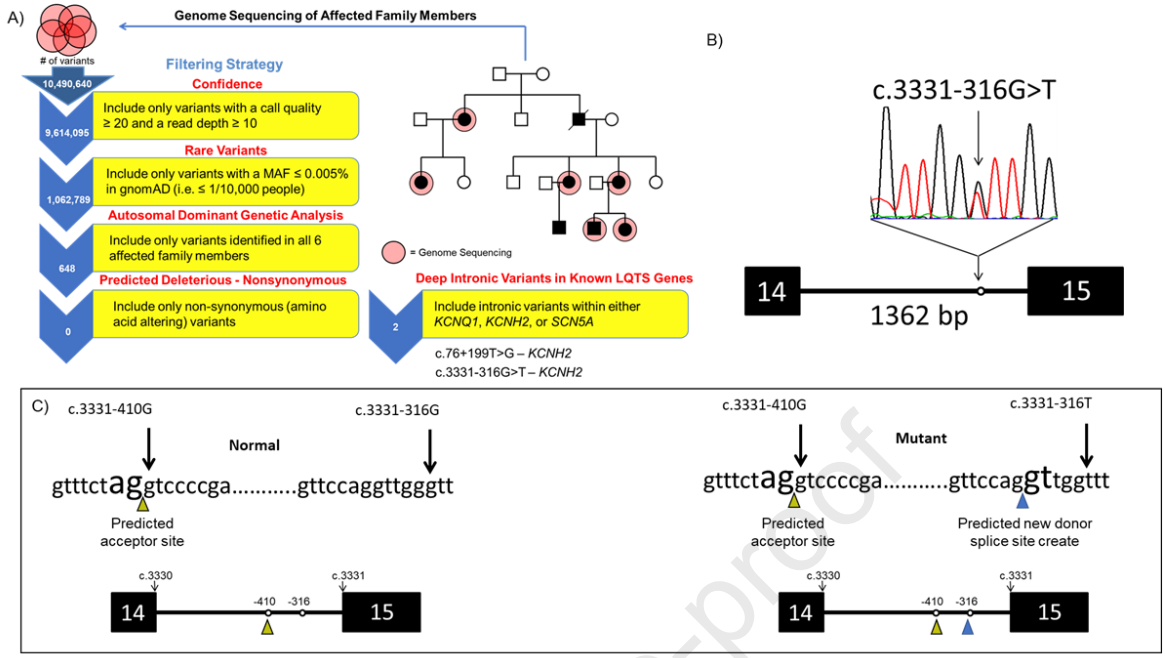


Figure 3

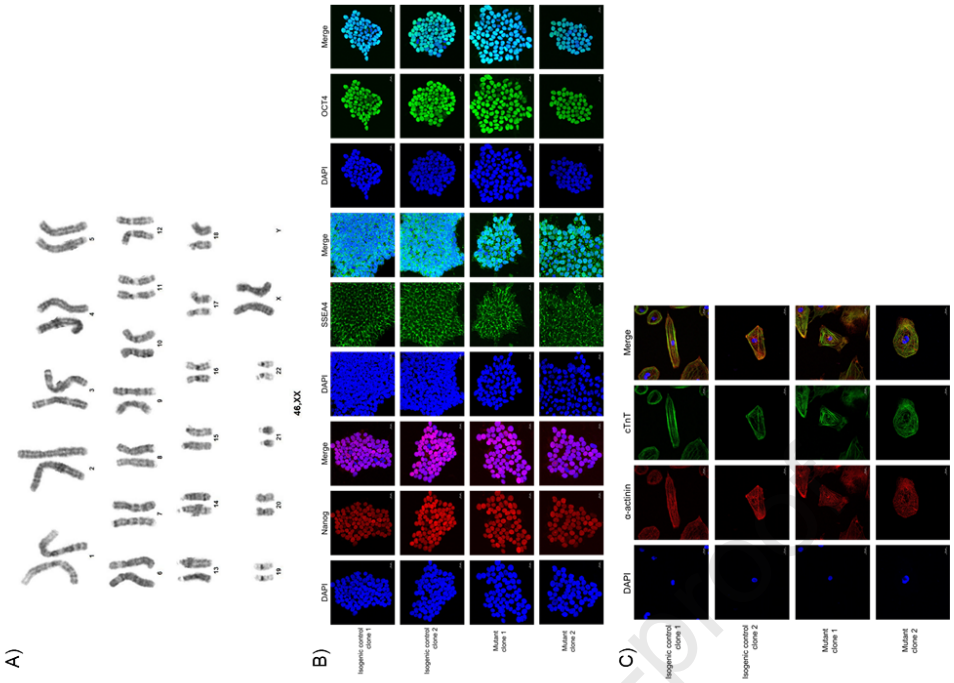


Figure 4

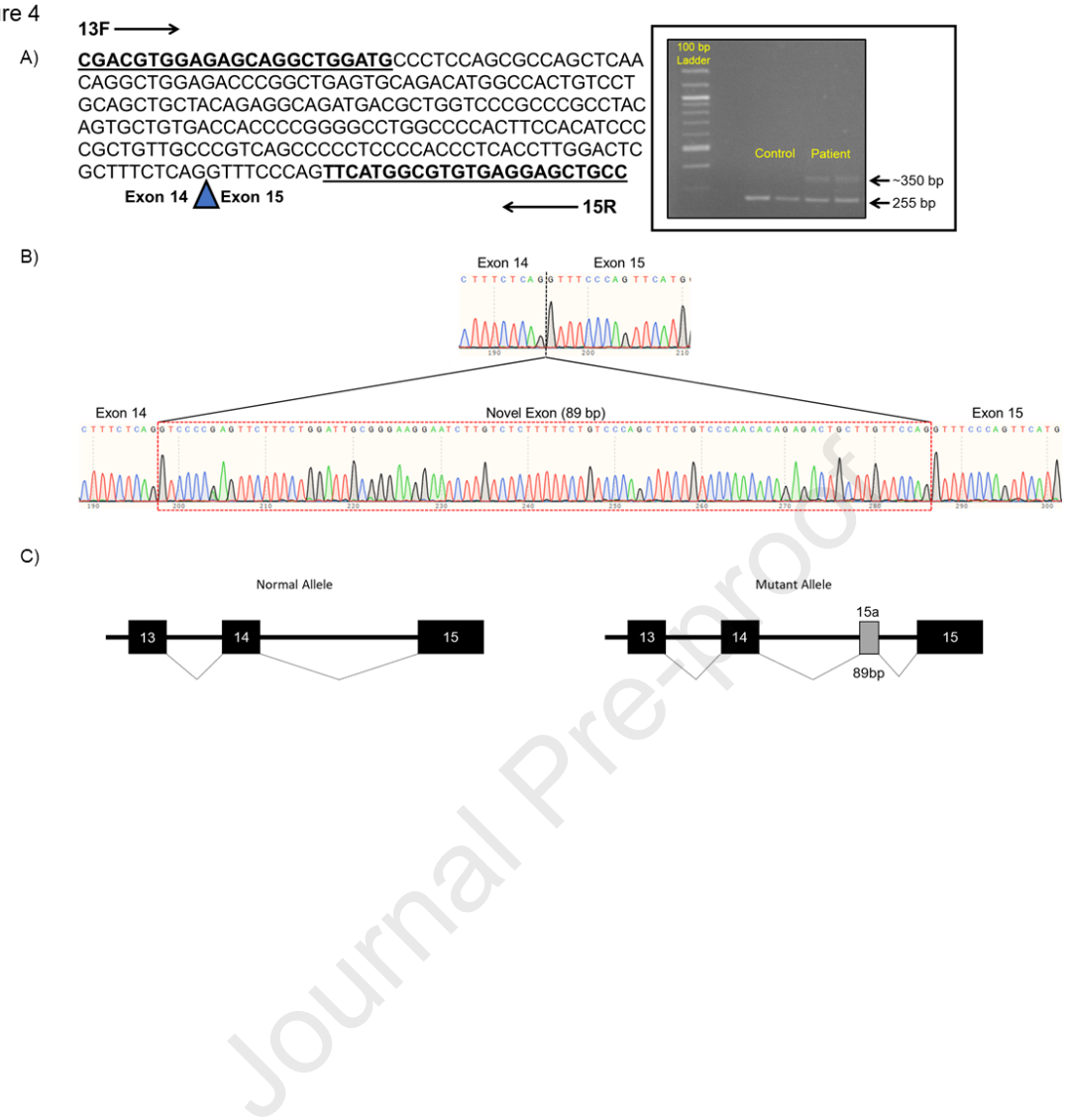


Figure 5

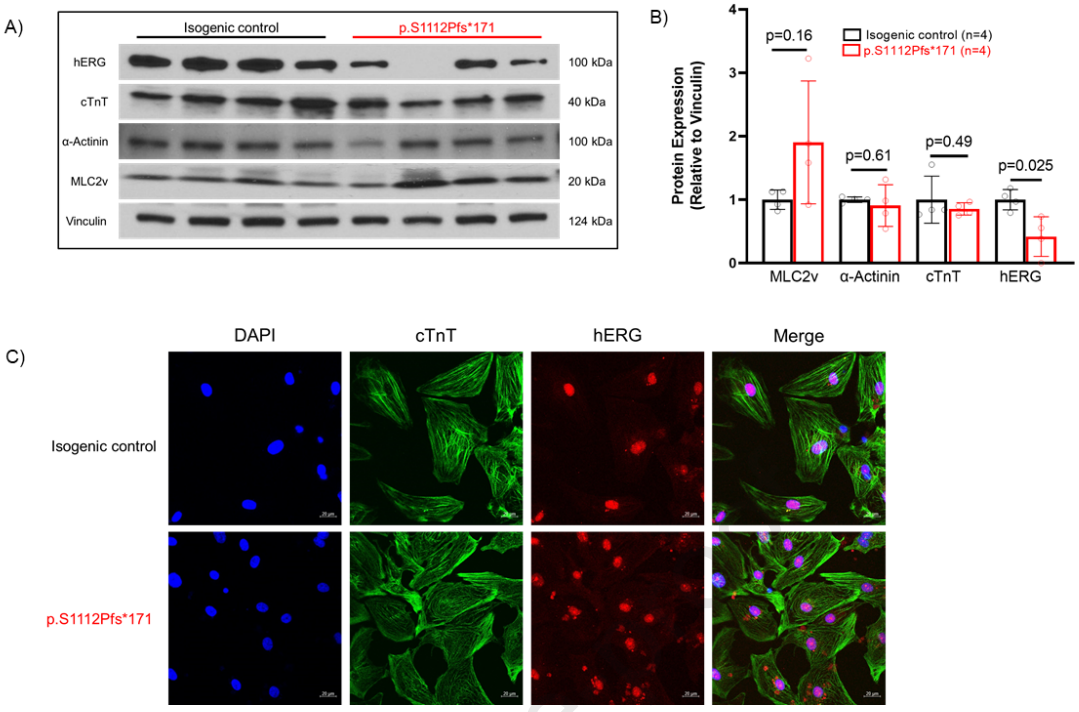


Figure 6