Original Article

Identification of Cadherin 2 (*CDH2*) Mutations in Arrhythmogenic Right Ventricular Cardiomyopathy AQ1 AQ2

Bongani M. Mayosi, MD, DPhil; Maryam Fish, PhD; Gasnat Shaboodien, PhD; Elisa Mastantuono, MD; Sarah Kraus, MB, ChB; Thomas Wieland, MSc; Maria-Christina Kotta, MSc, PhD; Ashley Chin, MPhil; Nakita Laing, MSc; Ntobeko B.A. Ntusi, DPhil; Michael Chong, MSc; Christopher Horsfall, MB, ChB; Simon N. Pimstone, PhD; Davide Gentilini, PhD; Gianfranco Parati, MD; Tim-Matthias Strom, MD; Thomas Meitinger, MD; Guillaume Pare, MD; Peter J. Schwartz, MD; Lia Crotti, MD, PhD AQ4

Background—Arrhythmogenic right ventricular cardiomyopathy (ARVC) is a genetically heterogeneous condition caused by mutations in genes encoding desmosomal proteins in up to 60% of cases. The 40% of genotype-negative cases point to the need of identifying novel genetic substrates by studying genotype-negative ARVC families.

Methods and Results—Whole exome sequencing was performed on 2 cousins with ARVC. Validation of 13 heterozygous variants that survived internal quality and frequency filters was performed by Sanger sequencing. These variants were also genotyped in all family members to establish genotype–phenotype cosegregation. High-resolution melting analysis followed by Sanger sequencing was used to screen for mutations in cadherin 2 (*CDH2*) gene in unrelated genotype-negative patients with ARVC. In a 3-generation family, we identified by whole exome sequencing a novel mutation in *CDH2* (c.686A>C, p.Gln229Pro) that cosegregated with ARVC in affected family members. The *CDH2* c.686A>C variant was not present in >200 000 chromosomes available through public databases, which changes a conserved amino acid of cadherin 2 protein and is supported as the causal mutation by parametric linkage analysis. We subsequently screened 73 genotype-negative ARVC probands tested previously for mutations in known ARVC genes and found an additional likely pathogenic variant in *CDH2* (c.1219G>A, p.Asp407Asn). *CDH2* encodes cadherin 2 (also known as N-cadherin), a protein that plays a vital role in cell adhesion, making it a biologically plausible candidate gene in ARVC pathogenesis.

Conclusions—These data implicate *CDH2* mutations as novel genetic causes of ARVC and contribute to a more complete identification of disease genes involved in cardiomyopathy. **(*Circ Cardiovasc Genet*. 2017;10:e001605. DOI: 10.1161/CIRCGENETICS.116.001605.)**

Key Words: arrhythmogenic right ventricular dysplasia ■ cadherins ■ cardiomyopathies ■ genetics ■ mutation AQ7

Arrhythmogenic right ventricular cardiomyopathy (ARVC) is a genetically transmitted disease, characterized by fibrofatty replacement of the right ventricular myocardium, favoring ventricular arrhythmia and sudden death especially in the young.[1]ARVC is primarily caused by mutations in genes encoding proteins of the desmosome, a type of intercellular junction abundant in tissues subjected to a high degree of mechanical stress, such as heart and skin.[2]These genes include plakophilin 2 (*PKP2*), desmoplakin (*DSP*), desmocollin 2 (*DSC2*), desmoglein 2 (*DSG2*), and junction plakoglobin (*JUP*), with *PKP2* being the main contributor to ARVC pathogenesis.[3,4] Other genes that have been implicated in ARVC include transmembrane protein 43 (*TMEM43*),[5] cardiac ryanodine receptor (*RYR2*),[6,7] and transforming growth factor beta 3 (*TGFB3*).[8] The putative mechanism of disease involves destabilization of the desmosomal complex by

Received August 23, 2016; accepted February 22, 2017.

From the Cardiovascular Genetics Laboratory, Hatter Institute for Cardiovascular Research in Africa, Cape Town (B.M.M., M.F., G.S., S.K., N.L.); The Cardiac Clinic, Department of Medicine, Groote Schuur Hospital (B.M.M., S.K., A.C., N.B.A.N.) and Division of Human Genetics, Department of Pathology, National Laboratory Service, South Africa (N.L.), University of Cape Town, South Africa; Technische Universität München (E.M., T.M.) and Helmholtz Zentrum München (E.M., T.W., T.-M.S., T.M.), Institute of Human Genetics, Neuherberg, Germany; Center for Cardiac Arrhythmias of Genetic Origin and Laboratory of Cardiovascular Genetics (M.-C.K., P.J.S., L.C.) and Molecular Biology Laboratory (D.G.), and Department of Cardiovascular, Neural and Metabolic Sciences, San Luca Hospital (G.P., L.C.), IRCCS Istituto Auxologico Italiano, Milan, Italy; Department of Molecular Medicine, University of Pavia, Italy (M.-C.K., L.C.); Population Health Research Institute, Hamilton Health Sciences, McMaster University, Hamilton, Ontario, Canada (M.C., G.P.); Department of Medicine, Frere Hospital and Walter Sisulu University, East London, South Africa (C.H.); Division of General Internal Medicine, University of British Columbia, Vancouver, Canada (S.N.P.); Department of Medicine and Surgery, University of Milano-Bicocca, Milan, Italy (G.P.); and DZHK (German Centre for Cardiovascular Research), partner site Munich Heart Alliance, Germany (T.M.). AQ5

The Data Supplement is available at http://circgenetics.ahajournals.org/lookup/suppl/doi:10.1161/CIRCGENETICS.116.001605/-/DC1.

Correspondence to Bongani M. Mayosi, MD, DPhil, Department of Medicine, J Floor Old Groote Schuur Hospital, Observatory 7925, Cape Town, South Africa. E-mail bongani.mayosi@uct.ac.za or Lia Crotti, MD, PhD, Department of Molecular Medicine, University of Pavia, 27100 Pavia, Italy. E-mail l.crotti@auxologico.it AQ6

mutant proteins, consequent myocyte death and replacement by fibrofatty tissue. The genes listed above only account for up to 60% of ARVC cases.[9,10]

Clinical Perspective on p xxx

Incomplete penetrance of the disease and the fact that a clear phenotype often manifests only in adulthood constitute a challenge for disease gene identification in ARVC. We therefore applied a 2-step approach. In a first step, an extended family helped to pinpoint the putative disease-causing gene. In a second step, a cohort of unrelated genotype-negative ARVC cases was screened for other pathogenic mutations in that gene, to reaffirm causality.

Methods

Study Subjects

This study was approved by the Institutional Review Board of the University of Cape Town. All patients gave written informed consent before undergoing evaluation and testing. One hundred healthy, anonymous blood donors with no history of cardiac disease from the Western Province Blood Transfusion Service provided samples for DNA isolation. Self-reported ethnicity of these individuals was white South African as were the members of the family under study.

The participants in this study were enrolled in the ARVC Registry of South Africa whose phenotyping scheme has been described previously.[11] Briefly, a scientific panel made up of cardiologists (A.C., B.M.M.) and clinical research fellows (N.B.A.N. and S.K.) determined, after achieving consensus, whether the referred cases and family members met the revised Task Force criteria for the diagnosis of ARVC. The presence of epsilon waves was reclassified based on the study of Platonov et al[12] using the 2010 Task Force criteria.[1]

Whole Exome Sequencing

DNA was extracted from blood samples using the PureGene DNA Purification Kit (Gentra Systems, Inc., Minneapolis, MN). Whole exome sequencing (WES) was performed on 2 affected cousins using the SureSelect Human All Exon 50 Mb kit (Agilent Technologies, Inc.), which covers the exonic sequences of ≈24 000 genes corresponding to 50 Mb of genomic DNA. Library preparation and sequencing on the Illumina HiSeq 2000 were performed at the Institute of Human Genetics, Helmholtz Zentrum München, Germany. The Burrows-Wheeler Alignment tool (version 0.7.5)[13] was used to align the reads to the human genome assembly hg19 (GRCh37). Variant calling was performed with SAMtools (v0.1.19)[14] and PINDEL (v0.2.4t).[15] Variant quality was determined using the SAMtools varFilter script: default parameters were applied, with the exception of the minimum *P* value for base quality bias (−2), which was set to 1e−400. Variant annotation was performed applying custom Perl scripts, including information about known transcripts (UCSC genes and RefSeq genes), known variants (dbSNP v135), type of mutation, and amino acid change (where applicable). The obtained variants were then inserted into the Helmholtz Zentrum München database, including 7916 exomes.

To identify putative pathogenic variants, we queried the Helmholtz database for variants present in both affected subjects with a minor allele frequency (MAF) cutoff of ≤0.0001, using an internal frequency filter. We selected an MAF≤0.0001 as a cutoff considering an estimated prevalence of ARVC of 1:5000.[16] Indeed, if 1:5000 subjects are affected by the autosomal dominant ARVC, a rare variant causing the disease should be present in the general population at genotype frequency <1/5000, which means an allele frequency <1/10 000. All synonymous and intronic (other than canonical splice sites) variants were excluded, and quality control filters were applied as well (variant quality ≥30 and mapping quality ≥50). Genetic variants were also interrogated in the 1000 Genomes project (www.1000genomes.org) and the Exome Aggregation Consortium (ExAC) browsers (http://exac.broadinstitute.org). Predicted functional effect of a coding variant was surveyed using Polyphen 2 (http://genetics.bwh.harvard.edu/pph/),[17] SIFT (http://sift.jcvi.org/),[18] MutationTaster[19] and Combined Annotation Dependent Depletion (CADD) (http://cadd.gs.washington.edu/).[20] Expression of the encoded proteins in the heart (www.genecards.org) was evaluated together with their biological plausibility in causing ARVC. As a further step, the resulting variants were visually inspected using the Integrative Genomics Viewer.[21] Validation of the variants that survived internal quality and frequency filters was performed by Sanger sequencing, and these variants were also genotyped in all family members to produce genotype–phenotype cosegregation data (Table I in the Data Supplement). Finally, parametric linkage analyses data (Supplemental Methods in the Data Supplement) were used as a further prioritization tool.

Targeted Genetic Screening of CDH2

CDH2 was screened in a population of unrelated ARVC probands through high-resolution melting analysis on the Rotor-Gene 6000 instrument (Corbett Life Science Mortlake, Australia). Samples displaying variations in melting patterns with high-resolution melting analysis were sequenced by Sanger sequencing with the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA) on the ABI PRISM 3130xl Genetic Analyser (Applied Biosystems) and analyzed using the BioEdit software (Ibis Biosciences, Carlsbad, CA; for primers, reagents and reaction conditions, please refer to Supplemental Methods and Table II in the Data Supplement).

Results

A kindred of white South African ancestry with familial ARVC was initially referred to our institution in 1996 for genetic investigation (hereafter, the arrhythmogenic cardiomyopathy 2 family; Figure 1). Five family members were classified as having definite ARVC (patients II:4, III:2, III:3, III:4, and III:5), whereas 1 member was classified as having borderline ARVC (patient II:7). The age of disease onset for arrhythmogenic cardiomyopathy 2 family members ranged from 13 to 57 years. In an attempt to identify the underlying genetic cause of ARVC in this family, genotyping of 25 microsatellite markers was originally performed to assess putative linkage to any of the ARVC loci described at the time (ARVC1 to ARVC6).[22] That analysis had identified a region of putative linkage on chromosome 10p12-p14.[22] However, candidate screening of several genes within that critical region and whole-genome copy number variation analysis through microarray genotyping (authors' unpublished data; Supplemental Methods in the Data Supplement) failed to identify the causative gene. F1 AQ8

The original phenotyping of this family had been performed mainly at Wentworth Hospital in Durban between 1990 and 2000 as part of the first report of ARVC in South Africa.[23] The proband (patient III:2) was a young man who presented at the age of 16 with recurrent episodes of palpitations associated with ventricular tachycardia requiring cardioversion. The diagnosis of ARVC in this individual was based on echocardiographic and angiographic findings of a dilated dysplastic RV, T-wave inversion in V_1 through V_5, delayed terminal activation duration of the QRS on ECG, and the presence of recurrent ventricular tachycardia with a left bundle branch block morphology and superior axis (Figure 2). He had a sudden cardiac death while bathing in 1992 at the age of 21 years. His sister (patient III:3) also had a witnessed sudden cardiac death in 1997 at the age of 24 years while she had previously been admitted to hospital several times for ventricular tachycardia (left bundle branch block morphology F2

Saritha	03/6/17	4 Color Fig(s): F1,3-5	18:50	Art: HCG001605

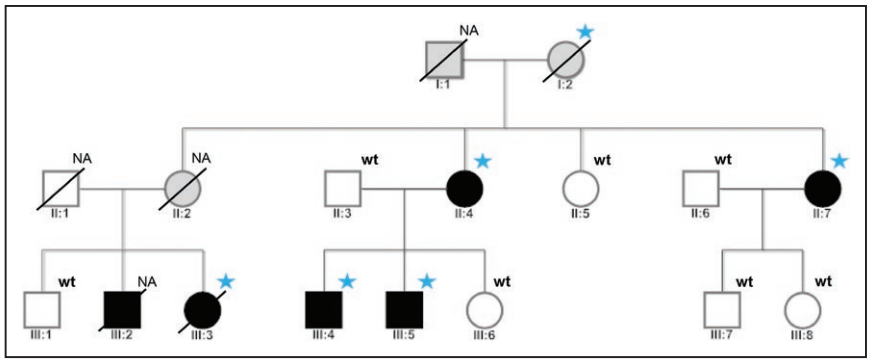

Figure 1. Arrhythmogenic cardiomyopathy 2 family pedigree depicting the cosegregation of the *CDH2* c.686A>C (p.Gln229Pro) mutation in the South African family with arrhythmogenic right ventricular cardiomyopathy. Filled symbols, clinically affected subjects; open symbols, unaffected subjects; gray-shaded symbols denote subjects with phenotype unknown or insufficient phenotypic data: *CDH2* c.686A>C (p.Gln229Pro) indicates mutation carrier; NA: *CDH2* genotype unknown; and wt: *CDH2* wild-type.

with superior axis) requiring cardioversion and amiodarone maintenance therapy. At the request of her family, an autopsy was not performed. Patient III:4 presented in 1995 with symptoms of chest pain and palpitations and the diagnosis of ARVC was confirmed on investigations outlined in Table 1. The other family members were ascertained through prospective family screening. T1

The scientific panel reassessed the clinical status and phenotypic classification of all the members of this family following the publication of the revised diagnostic criteria for ARVC in 2010 (Table 1) and found patients III:6 and III:8, who were initially considered to be affected,[22] to not meet the revised criteria for the diagnosis of ARVC.

Patient III:6 (III:7 in the 2006 report)[22] was re-examined at Groote Schuur Hospital in Cape Town in 2015 because her clinical records were missing after the closure of the Cardiac Unit at Wentworth Hospital. She was found to be asymptomatic with a normal ECG, echocardiogram, cardiac magnetic resonance imaging, and 24-hour ECG monitoring, and the patient was therefore reclassified as unaffected. The reclassification of patient III:6 to normal was substantiated by correspondence from her cardiologist at Wentworth Hospital, which indicated that she had no history of syncope or premature ventricular contractions >1000/24 h. Therefore, the phenotype allocation in the 2006 report was erroneous.[22]

On review of the clinical records of patient III:8 (III:9 in 2006 report),[22] we found that she had insufficient evidence to support the diagnosis of ARVC. Although she had presented with syncope subsequent to the 2006 report,[22] ventricular tachycardia was neither induced on 2 separate electrophysiological studies

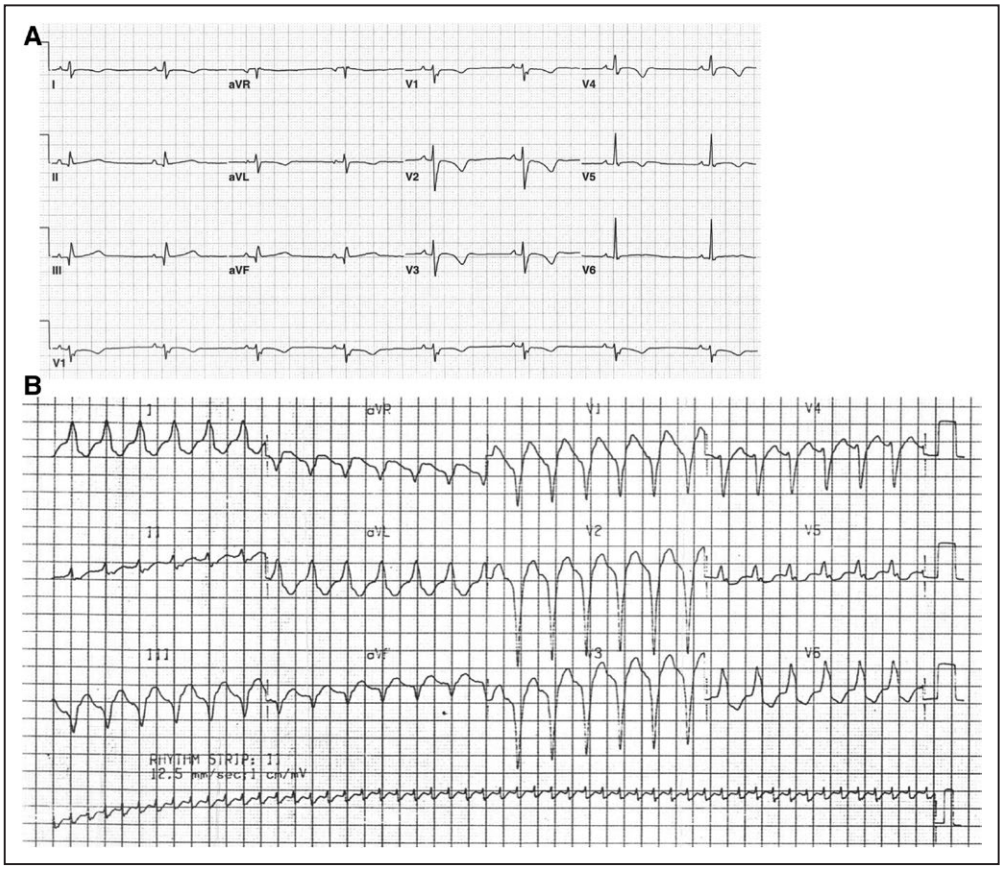

Figure 2. Electrocardiograms of the index case in family Arrhythmogenic cardiomyopathy (ACM) 2 (ie, ACM III:2). **A**, Twelve-lead resting ECG showing T-wave inversion in V_1 through V_5 and delayed terminal activation duration of the QRS in the chest leads. **B**, Sustained ventricular tachycardia with left bundle branch morphology and superior axis.

Table 1. Clinical Characteristics of Patients Affected With ARVC

Family	ID	Sex	Age at Diagnosis, y	Symptoms	Cardiac Imaging	Tissue Characteristics	ECG Repolarization Abnormalities
ACM 2	II:4	F	44	Palpitations, chest pain, presyncope	Echo: normal	Myocardial biopsy not done	TWI V_1 through V_3 (major)
					MRI: not done		
					RVA: not done		
ACM 2	II:7	F	44	Palpitations, chest pain	Echo: normal	Myocardial biopsy not done	TWI V_1 through V_2 (minor)
					MRI: normal		
					RVA: not done		
ACM 2	III:2	M	16	Palpitations	Echo: unavailable	Myocardial biopsy not done	TWI V_1 through V_5 (major)
					MRI: not done		
					RVA: dilated RV with regional akinesia of outflow tract (major)		
ACM 2	III:3	F	13	Palpitations	Echo: reported as dilated RV with hypokinesia but dimensions not available	Myocardial biopsy not done	Information not available
					MRI: not done		
					RVA: not done		
ACM 2	III:4	M	23	Palpitations, chest pain	Echo: dilated RV (42 mm),	Fatty infiltration (not quantified), scattered chronic inflammatory cells	TWI V_1 through V_5 (major)
					LVEDD 6.5 cm, LVEF 43%		
					MRI not done		
					RVA: bulging hypokinetic		
					RV inferior wall and subtricuspid dyskinesia (major)		
ACM 2	III:5	M	21	Palpitations, chest pain	Echo: RVOT 40 mm	Fibrofatty replacement (not quantified)	No TWI
					MRI: not done		
					RVA: anterior hypokinesia (major)		
ACM 11	ACM 11.2	M	15	Palpitations, presyncope	Echo: multiple aneurysms, RVEF 32%, RVOT 27 mm	Fibrofatty infiltration (not quantified)	TWI V_1 through V_6 (major)
					MRI: marked dilatation of RV, dyskinesia, multiple aneurysms,		
					RVA: Dilated with regional wall motion abnormalities (major)		

(*Continued*)

(performed in 2009 and 2014 by her attending cardiologist) nor had any episodes of ventricular tachycardia been recorded on her implantable cardioverter defibrillator, which had been inserted prophylactically in 2009 at the request of the patient. In addition, neither the ECG nor the echocardiogram performed in 2009 showed any features suggestive of ARVC. A review of the

Table 1. Continued

Depolarization/ Conduction Abnormalities	Arrhythmias	Family History	Diagnostic Criteria	Outcome
No epsilon wave	No history of VT	Son meets TFC for ARVC (major)	DEFINITE ARVC	Alive
No late potentials	24-h Holter, EST, and EPS not done		2 major (repolarization abnormalities and family history)	Last seen 2012
No epsilon wave	No history of VT	Sister meets criteria for ARVC (major)	BORDERLINE ARVC	Alive
No late potentials	EST normal		1 major (family history), 1 minor (repolarization abnormalities)	Last seen 2004
	24-h Holter ECG and EPS not done			
No epsilon wave	Sustained VT with LBBB morphology and superior axis (major)	Sister had an SCD at the age of 24 h because of ARVC (not considered as it occurred after his SCD)	DEFINITE ARVC	Died at the age of 21 y (1992)
SAECG not done	EPS: VT induced		3 major (structural abnormalities, repolarization changes, arrhythmia), 1 minor (depolarization abnormalities)	
TAD≥55 ms (minor)	24-h Holter ECG not done			
Information not available	Sustained VT with LBBB morphology and superior axis (major)	Brother met criteria for ARVC (major)	DEFINITE ARVC	Died the age of 24 y (1997)
			2 major (Arrhythmia, family history)	
No epsilon wave	EPS: NSVT LBBB morphology inferior axis (minor)	Cousin (third-degree relative) meets criteria for ARVC and had an SCD at the age of 21 y	DEFINITE ARVC with LV involvement	Alive
No late potentials	PVC 12620/24 h (minor)		2 major (structural abnormalities, repolarization changes), 1 minor (arrhythmia)	Last seen 2012
No epsilon wave	No history of VT	Brother meets criteria for ARVC (major)	DEFINITE ARVC	Alive
No late potentials	PVC 2035/24 h (minor)		2 major (structural abnormalities, family history), 1 minor (arrhythmia)	Last seen 2016
	EPS: VF induced			
No epsilon wave	Sustained VT with LBBB morphology and superior axis	No family history of ARVC or SCD	DEFINITE ARVC	Alive with ICD
Late potentials on SAECG	EPS: VT induced (multiple morphologies from RV free wall and RVOT)		3 major (structural abnormalities, repolarization abnormailies, and arrhythmia), 1 minor (depolarization abnormalities)	Last seen 2006
TAD≥55 ms (minor)	24-h Holter ECG not done			

ACM indicates arrhythmogenic cardiomyopathy; ARVC, arrhythmogenic right ventricular cardiomyopathy; Echo, echocardiogram; EPS, electrophysiological study; EST, exercise stress test; F, female; ICD, implantable cardioverter defibrillator; LBBB, left bundle branch block; MRI, magnetic resonance imaging; M, male; NSVT, nonsustained ventricular tachycardia; RV, right ventricle; PVC, premature ventricular contractions; RVA, right ventricular angiography; SAECG, signal averaged ECG; SCD, sudden cardiac death; TAD, terminal activation duration; TWI, T-wave inversion; and VT, ventricular tachycardia. AQ9 AQ10

cardiac magnetic resonance imaging study that was originally reported as showing mild RV dilation with wall motion change and aneurysmal abnormality (in the 2006 report)[22] was within normal limits. Therefore, patient III:8 was also reclassified as unaffected. Individuals III:1 and III:7 were fully investigated in 2014, and both were found to be clinically unaffected.

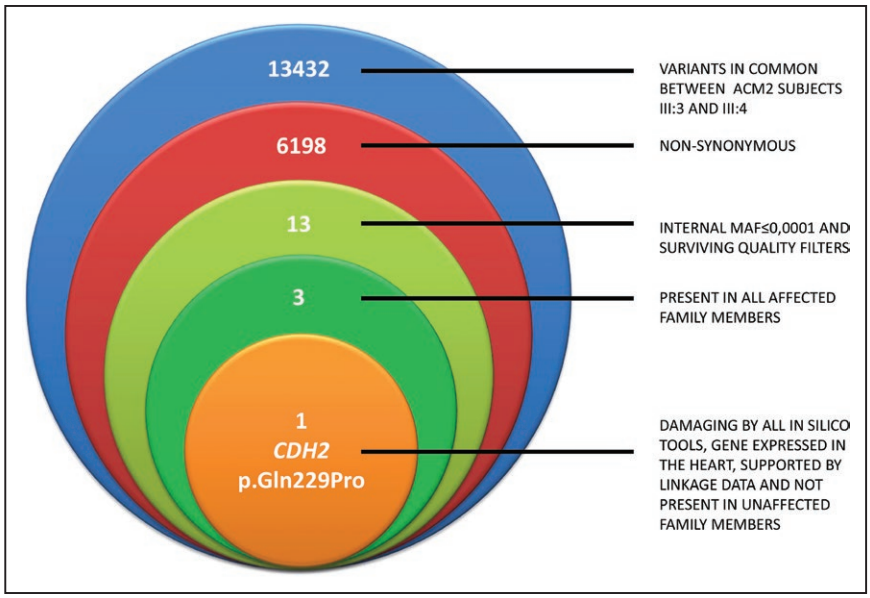

Figure 3. Whole exome sequencing variants filtering scheme.

The phenotypic reclassification of 2 family members from affected to unaffected status prompted the re-evaluation of the previously performed genetic linkage analysis. Because the latter assesses the probability that the disease and genetic loci cosegregate within a pedigree because of the linkage and not because of the chance, it was highly probable that the phenotypic misattributions of 2 family members had introduced significant bias in the results obtained. With the advent of more powerful genetic technologies available in recent times, WES was used to identify the genetic cause of ARVC in this family.

WES was performed in 2 clearly affected individuals (patients III:3 and III:4; Figure 1) and identified ≈13 000 genetic variants common to both individuals (Figure 3). Among these, approximately half were nonsynonymous (nucleotide substitutions, insertions, and deletions), but only 13 heterozygous variants survived after applying the internal quality and frequency filters (Figure 3). These 13 variants (Table 2) were all confirmed by Sanger sequencing in the 2 affected individuals, but only 3 (*CDH2*, *NEK10*, and *ADD1*) were found to be also present in all family members who were clearly affected (Figure 3; Table 2). The 3 variants were either absent (*CDH2* and *NEK10*) or present at a low frequency (*ADD1*) in the 1000 Genomes and ExAC browsers (Table 2), whereas they were absent in 200 white South African control chromosomes.

CDH2 encodes cadherin 2, previously known also as N-cadherin, a cell surface transmembrane protein with a vital

Table 2. List of Nonsynonymous Variants (n=13) Identified Through Whole Exome Sequencing in Both Affected Subjects (III:3 and III:4) Who Survived the Internal Quality and Frequency Selection Filters

Chromosome Position	Nucleotide Change	Amino Acid Change	Type	Gene	Protein Function
chr13:109365050-109365050	c.268G>A	p.Val90Ile	Missense	*MYO16*	Unconventional Myosin
chr17:39465289-39465289	c.217T>C	p.Cys73Arg	Missense	*KRTAP16-1*	Keratin-Associated Protein
chr17:39550368-39550368	c.1151C>T	p.Pro384Leu	Missense	*KRT31*	Keratin
chr17:57759757-57759757	c.3568A>G	p.Ile1190Val	Missense	*CLTC*	Clathrin
chr17:62049087-62049087	c.606G>C	p.Met202Ile	Missense	*SCN4A*	Voltage-gated Sodium Channel
chr18:12107373-12107375	c.973_975del	p.Asn325del	Indel	*ANKRD62*	Ankyrin Repeat Domain-Containing Protein
chr18:19153979-19153979	c.826C>G	p.Leu276Val	Missense	*ESCO1*	N-Acetyltransferase
chr18:25589697-25589697*	c.686A>C*	p.Gln229Pro*	Missense*	*CDH2**	Cadherin*
chr1:230401006-230401006	c.1333A>G	p.Ile445Val	Missense	*GALNT2*	Polypeptide N-Acetylgalactosaminyltransferase
chr2:10192525-10192525	c.1430T>G	p.Met477Arg	Missense	*KLF11*	Zinc finger transcription factor
chr3:27326353-27326353*	c.1889G>A*	p.Arg630Lys*	Missense*	*NEK10**	Serine/threonine kinase*
chr4:2927788-2927788*	c.1903C>G*	p.Gln635Glu*	Missense*	*ADD1**	Adducin*
chr7:127239495-127239495	c.1181G>A	p.Arg394His	Missense	*FSCN3*	Fascin

(*Continued*)

structural and functional role in the intercalated disc, and a major contributor in cell-to-cell adhesion. *NEK10* encodes a serine/threonine kinase, with low level of expression in the heart, whereas *ADD1* codes for adducin α, a cytoskeleton protein that binds with high affinity to Ca^{2+}/calmodulin and is a substrate for protein kinases A and C.[24]

Although both *CDH2* and *ADD1* are highly expressed in the heart,[24] we prioritized the *CDH2* gene as the causative gene because only the *CDH2* variant (c.686A>C; p.Gln229Pro; F4 Figure 4A) was (1) predicted to impact the protein by all in silico tools that were used (Table 2), (2) predicted to affect a highly conserved nucleotide (with maximum and positive PhastCons[25] and PhyloP[26] scores, respectively), (3) was not present in all clinically evaluated clearly unaffected family members (Figure 1), (4) was the biologically most plausible candidate, and (5) was supported by parametric linkage analysis data (see below). In contrast, the *ADD1* variant was predicted to be benign by all in silico tools, was present at a low frequency in the Helmholtz and publicly available exome databases (Table 2), was not supported by linkage analysis (see below), and was also present in clearly unaffected family members. Therefore, this variant is unlikely to be the cause of the disease in the family. The *CDH2* mutation was also present in individual I:2 who died before phenotype assessment could be done and unfortunately its presence could not be confirmed in the sudden death victim III:2 because no blood or postmortem biological material had been retained.

In addition to the biological plausibility of *CDH2* as the causal gene underlying ARVC in this family, parametric linkage analysis data provided further support in this direction. We decided to use microarray genotyping data, which were at our disposal from previously performed whole-genome copy number variation analysis in this family (see Supplemental Methods in the Data Supplement) to identify genetic regions of positive linkage that would aid the prioritization of the variants in the last WES filtering steps and thereby further strengthen the final selection of the causative gene. Genotyping data from a HumanCytoSNP-12 BeadChip array (Illumina) were available for 11 individuals of the arrhythmogenic cardiomyopathy 2 family (I:2, II:3, II:4, II:7, III:1, III:3, III:4, III:5, III:6, III:7, and III:8 in Figure 1). To perform a parametric linkage analysis, we assigned as unknown the phenotypic status of individual AQ11 I:2, whereas all the other family members were either affected or not affected. We then considered all genetic regions of positive linkage (LOD score>0; Table III in the Data Supplement) AQ12 and prioritized the regions with an LOD score>2 (Table IV in the Data Supplement). We then compared the latter with the genomic positions of all 13 genes in which variants were identified through WES in the 2 affected subjects III:3 and III:4 (Table 2). No genetic region with an LOD score>2 contained any of the genes with the 13 WES-resulting variants except the *CDH2* gene (Tables 2 and Table V in the Data Supplement). Furthermore, the chromosomal region containing *CDH2* obtained the maximum LOD score of 2.24. By contrast, the genetic regions containing the *NEK10* and *ADD1* genes had low LOD scores, 0.24 and 0.1, respectively. We do not expect the phenotypic reclassification of the family members III:6 and III:8 to have an impact on these results. Indeed, even under different settings of penetrance (from near-complete to low), the results obtained were not significantly different (data not shown) and the region containing the *CDH2* gene repeatedly obtained the highest LOD score. Based on this analysis, we were able to reinforce the prioritization of the *CDH2* variant, providing a cross-validation of the overall WES results.

We then sought to investigate whether *CDH2* mutations could be the underlying genetic cause of ARVC in an independent population of patients whose screening was negative for

Table 2. Continued

Present in All Affected	Polyphen2	SIFT	CADD	MutationTaster	IHG Helmholtz	1000 Genomes	ExAC EUR	ExAC AFR
No	0.113	Np	0.023	Polymorphism	0.0001	0	0.0002	0
No	0.879	0.02	22.3	DiseaseCausing	0.0001	0.0006	0	0.004
No	0.001	0.01	23.4	DiseaseCausing	0.00006	0.0007	0	0.0025
No	0.001	0.24	14.71	DiseaseCausing	0	0	0	0
No	0.009	1	12.07	Polymorphism	0	0	0	0
No	np	np	np	Polymorphism	0.00006	0	0	0.003
No	0.01	0.3	0.0003	Polymorphism	0	0.003	0.00005	0.008
Yes*	0.785*	0.03*	16.6*	DiseaseCausing*	0*	0*	0*	0*
No	0.001	0.76	11.97	DiseaseCausing	0.00006	0	0.0002	0.0002
No	0.483	0.09	26.2	DiseaseCausing	0	0	0	0
Yes*	0.034*	0.35*	22.4*	Polymorphism*	0*	0*	0*	0*
Yes*	0.006*	1*	0.213*	Polymorphism*	0.00006*	0*	0.0001*	0.0001*
No	0.99	0.09	28.3	DiseaseCausing	0.00006	0	0.00005	0.0001

Minor allele frequency of each variant in the different exome databases (IHG Helmholtz, 1000 Genomes, ExAC) is reported. AFR indicates the African subpopulations of ExAC; EUR, the European non-Finnish; Combined Annotation Dependent Depletion (CADD): scores >15 indicate damaging predictions; Polyphen2, values >0.5 indicate possibly or probably damaging predictions; and SIFT: values <0.05 indicate damaging predictions.

*The 3 variants present in all affected subjects of the family.

Saritha | 03/6/17 | 4 Color Fig(s): F1,3-5 | 18:50 | Art: HCG001605

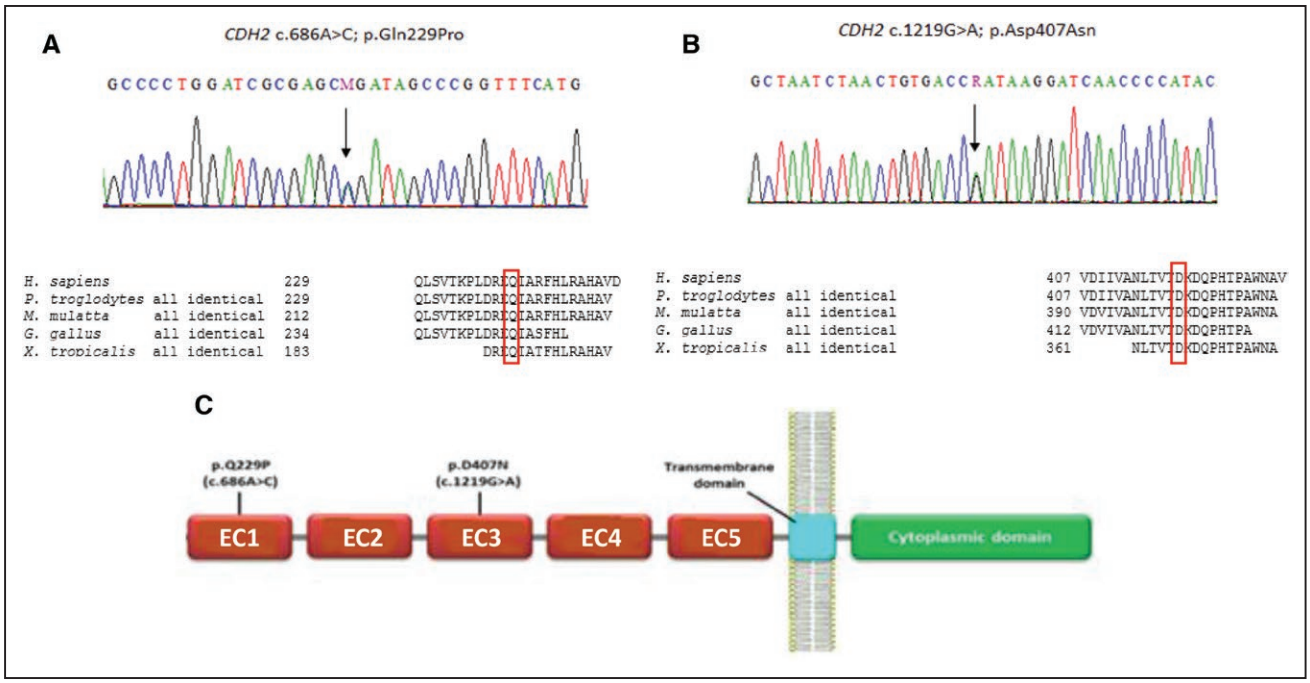

Figure 4. *CDH2* mutations in arrhythmogenic right ventricular cardiomyopathy (ARVC) families. **A**, *CDH2* c.686A>C (p.Gln229Pro) electropherogram and multiple species alignment. **B**, *CDH2* c.1219G>A (p.Asp407Asn) electropherogram and multiple species alignment. **C**, Schematic representation and cellular localization of the topological domains of the cadherin 2 protein and the relative positions of the identified *CDH2* mutations. EC indicates extracellular cadherin repeat.

mutations in the known ARVC genes. To this end, we screened 73 genotype-negative unrelated cases and identified another likely *CDH2* pathogenic variant (c.1219G>A; p.Asp407Asn; Figure 4B) in one proband of white South African ancestry with definite ARVC (arrhythmogenic cardiomyopathy 11.2; Table 1). The proband, a young male from East London, presented in 2003, in a similar manner to patient III:2, with palpitations and presyncope associated with ventricular tachycardia at the age of 15 years. He was referred to Groote Schuur Hospital in Cape Town where the diagnosis of ARVC was made based on the presence of T-wave inversion from V_1 through V_6, late potentials (signal averaged ECG), sustained ventricular tachycardia with left bundle branch block morphology (superior axis), and structural changes (marked RV dilatation and multiple aneurysms of the RV free wall) seen on echocardiography, angiography, and cardiac magnetic resonance imaging. Histological examination of the RV endomyocardial biopsy showed fibrofatty replacement of cardiomyocytes (Figure 5). He had an implantable cardioverter defibrillator inserted in 2004, and he has subsequently been followed up by a local cardiologist in East London.

The *CDH2* heterozygous mutation identified was (1) predicted to result in the alteration of a fully conserved residue, (2) predicted to impact the protein by all in silico tools, (3) absent in the unaffected mother of the proband, and (4) absent in the in-house control population and in the Helmholtz database. The variant is present once in the 1000 Genomes browser (rs568089577) with an MAF of 0.0002, and it is also present as a low-quality variant once in ExAC (MAF=0.000008). Because ExAC incorporates part of the 1000 Genomes population, it is unclear whether this finding may refer to the same individual.

In summary, we identified *CDH2* mutations that are likely to be pathogenic in 2/74 (2.7%) South African ARVC probands in which a disease-causing mutation was excluded in known ARVC genes. Such a prevalence is significantly higher than what is reported in the ExAC database. Indeed, in the 60706 unrelated individuals sequenced as part of various disease-specific and population genetic studies included in the ExAC database, 244 nonsynonymous variants with MAF<0.0001 are reported, but only 61, present in 123 subjects (0,2%), are predicted to be damaging. The difference in prevalence of rare variants with a

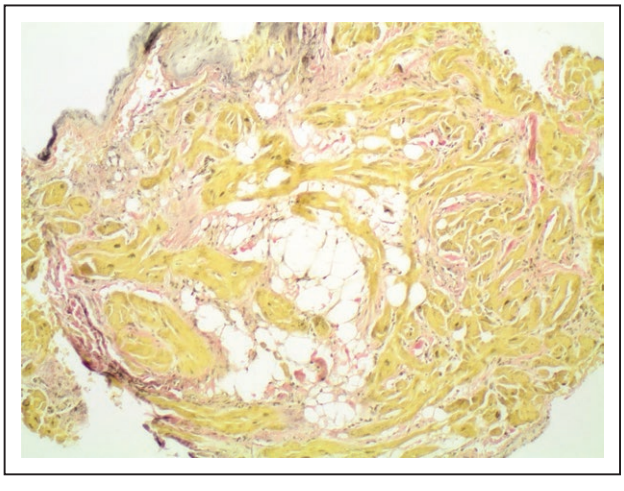

Figure 5. Endomyocardial biopsy sample from arrhythmogenic right ventricular cardiomyopathy index case ACM11.2. Histological specimen of the right ventricular myocardium showing evidence of fibrofatty infiltration (elastic von Gieson stain, magnification ×100).

predicted damaging effect between our ARVC cohort (2.7%) and the general population (0.2%) is highly significant (P=0.0005).

Discussion

The present study provides several lines of evidence implicating mutations in *CDH2* as a likely novel genetic cause of ARVC. Even though the role of cadherin 2 (*CDH2*) as a disease gene in human cardiomyopathy and arrhythmia has long been suspected, this is, to our knowledge, the first report of an association of *CDH2* mutations with ARVC.

The use of WES in a genotype-negative 3-generation ARVC family unmasked the presence, in all affected members, of a novel mutation in *CDH2* and the subsequent search in our cohort of genotype-negative patients with ARVC identified another likely pathogenic variant in the same gene.

Multiple considerations strongly support a causative role for *CDH2* in ARVC. First, the *CDH2* c.686A>C mutation cosegregates with ARVC in the South African family. Second, the *CDH2* gene is a plausible candidate gene for ARVC because it encodes the cadherin 2 protein, a cell-adhesion protein highly expressed in the heart. Third, the mutation changes a conserved amino acid of the cadherin 2 protein and was not found in >200 000 chromosomes in the general population and public genetic databases. Fourth, parametric linkage analysis supports only *CDH2* as the causal gene for ARVC in the affected family. Finally, we have validated our finding by identifying another *CDH2* mutation in an independent ARVC cohort.

WES Results and Prioritization Process

After several unsuccessful attempts, in the pre–next-generation sequencing era, to identify the disease-causing gene in a 3-generation white South African family with an autosomal dominant ARVC, we eventually used WES selecting 2 cousins (the most distant-related affected family members). As expected, the number of variants in common between the 2 was high, >13 000, forcing the application of stringent selection criteria. This is a well-known challenge of WES analysis. It is possible to get the sequence of the entire exome or even of the entire genome of a subject within a few days, but the genomic data are of limited use without rigorous postsequencing analysis. Furthermore, WES analysis for gene discovery is not always successful, and in the absence of large families in which genotype–phenotype cosegregation data can be obtained, the probability of confirming a positive result is low, as it is when the disease under investigation is a non-Mendelian trait. By contrast, when WES is used as a diagnostic tool to screen known disease-causing genes, the genetic yield is much higher and varies according to the disease under investigation.

In the present study, we were dealing with ARVC, a well-defined autosomal dominant disease, with a 3-generation family comprised an acceptable number of affected and unaffected individuals, and this provided a good a priori probability of identifying the disease-causing gene.

As a first step of the postsequencing analysis, after the bioinformatics alignment, variant calling, and annotation, we focused on variants that were expected to modify the encoded protein. This step allowed the exclusion of >7000 variants shared by the 2 affected cousins, leaving ≈6000 variants in common. By using quality and frequency filters, this number of variants was significantly reduced to 13. These 13 variants were investigated in detail, as explained in the Methods and Results sections of this article, leaving 3 variants that were present in all clearly affected subjects (*CDH2*-p.Gln229Pro, *NEK10*-p.Arg630Lys, and *ADD1*-p.Gln635Glu). Among these 3 genes, only the *CDH2* and *ADD1* genes are highly expressed in the heart, whereas the expression of *NEK10* is poor. Furthermore, only the *CDH2* variant was predicted to be damaging/disease-causing by all prediction tools. Finally, biological plausibility was strongly supporting *CDH2* as the ARVC disease-causing gene (see below).

CDH2 is not only exclusively expressed in the heart but also in other tissues, such as the neuronal one[24]; however, the clinical phenotype of all our affected individuals was purely cardiological. This is not surprising. There are many examples in inherited arrhythmogenic cardiac diseases showing that genes expressed in multiple tissues give rise to a pure cardiac phenotype. For instance, heterozygous mutations in *KCNQ1* cause long-QT syndrome with a pure cardiac phenotype, despite the fact the gene is also clearly expressed in many other tissues, including brain, smooth muscle, skin, thyroid, pancreas, and inner ear. However, when homozygous or compound heterozygous mutations are present in *KCNQ1* then the Jervell–Lange Nielsen syndrome, in which a severe cardiac phenotype is associated with neurosensorial deafness, manifests.[27] Similarly, in the field of ARVC, desmosomal genes such as *JUP* and *DSP* are highly expressed in other tissues such as brain and hair follicle, respectively,[24] without necessarily giving rise to an extracardiac phenotype. So probably, although the same gene is expressed in different tissues, the importance of the gene product in a given tissue, the compensatory mechanisms or the functional reserve, can vary significantly from tissue to tissue. In addition, the type of mutation (missense or nonmissense) and the mode of inheritance (recessive or dominant) are altogether expected to influence the phenotypic manifestations of a given mutation.

To further reinforce the prioritization process and the association between *CDH2*-p.Gln229Pro and ARVC in our family, we also used parametric linkage analysis. Linkage analysis was used to identify a chromosomal region in common between affected subjects of the family, where the disease-causing gene was more likely to reside. This method has long been used to identify the disease-causing gene in different monogenic disorders but requires large families with adequate number of clearly affected and unaffected family members.[28] We hereby took advantage of linkage analysis to complement WES data and strengthen the prioritization process of the variants identified. The fact that WES and linkage analysis can complement one another to identify novel disease genes has been demonstrated previously.[29]

Biological Plausibility of CDH2 as a New ARVC-Associated Gene

Cadherin 2, also known as N-cadherin, is a member of the cadherin superfamily of predominantly Ca^{2+}-dependent cell surface adhesion proteins.[30] In the heart, cadherin 2 is located at the intercalated disc, a complex and highly organized intercellular structure that ensures structural integrity and functional synchronization across the myocardium through the tight electromechanical coupling of cardiomyocytes.[30,31] In the

intercalated disc, intercellular communication and adhesion are achieved through 3 main junctional structures forming functional zones, the gap junctions, the fascia adherens junctions, and the desmosomes, with the latter 2 being the main contributors to cell–cell adhesion.[30–32] In desmosomes, desmosomal cadherins (desmocollin and desmoglein) are mainly anchored to the intermediate filaments of the cytoskeleton through many intracellular protein partners, whereas in fascia adherens junctions, the classical cadherin, N-cadherin, is primarily anchored to the actin microfilaments of the cytoskeleton and promotes cell–cell adhesion through extracellular associations of its cadherin repeat domains.[31–33] Interestingly, the protein components of desmosomes and fascia adherens junctions are not mutually exclusive.[33,34]In fact, mechanical junctions of the intercalated disc are an admixture of desmosomal and fascia adherens proteins that form a hybrid functional zone, nowadays known as area composita.[32–36] Therefore, even if ARVC has been traditionally considered as a disease of the desmosome, it is now reasonable to hypothesize that the mechanistic basis of ARVC may extend beyond the strict functional zone of the desmosome, to that of the area composita. In support of this concept is a recent description of another gene in the area composita, *CTNNA3*, which encodes for αT-catenin, in which mutations have been recently identified in patients with ARVC who were negative for mutations in the main desmosomal genes.[37] α-Catenins are natural partners of the cytoplasmic domain of classical cadherins, that is, N- and E-cadherins, and in the case of N-cadherin act as its go-between for anchoring to the actin cytoskeleton.[32]

The fact that cadherin 2, like its desmosomal cadherin counterparts, is indeed a major player in the intercalated disc is also supported by the *Cdh2* mouse models that have been established. Using a conditional knockout, a cardiac-specific deletion of cadherin 2 (N-cadherin) in the adult mouse heart was obtained and this caused dissolution of the intercalated disc structure including loss of both desmosomes and adherens junctions, thus demonstrating for the first time that desmosome integrity is also cadherin 2 dependent.[38] These mice also exhibited modest, albeit atypical, dilated cardiomyopathy, and spontaneous ventricular arrhythmias that resulted in sudden cardiac death.[38] This increased arrhythmic propensity (all mice suffered sudden cardiac death ≈2 months after deleting N-cadherin from the heart) was probably because of a reduced and heterogeneously distributed connexin-43, causing loss of functional gap junctions and partial cardiomyocyte uncoupling.[39] These data highlight the prominent role of cadherin 2 in all types of functional junctions in the intercalated disc, a finding also compatible with the original descriptions of abnormal connexin-43 expression and gap junction remodeling in the recessive ARVC form of Naxos disease.[40]

Gap junction decreases in number and size, with accompanying increased arrhythmia susceptibility, has subsequently been demonstrated also in the context of N-cadherin heterozygous null mice; 30% to 60% of these mice develop ventricular tachycardia, which suggests that cadherin 2 haploinsufficiency may create an important arrhythmogenic substrate.[41] Intercalated disc remodeling with concomitant reduction of localization of desmosomal proteins, connexin-43 and cadherin 2, has been demonstrated also in ventricular tissues of transplanted hearts of patients with ARVC,[42] further supporting the involvement of cadherin 2 in ARVC pathogenesis.

Cadherin 2 comprised 5 extracellular cadherin repeats (EC1–EC5), a transmembrane region and a highly conserved cytoplasmic tail.[43] The EC repeats are vital for the adhesive function of cadherin 2 and the 2 mutations we identified in our patients with ARVC were both located in these repeats. Calcium ions bind to each cadherin repeat to ensure correct protein folding and confer rigidity to the extracellular domain. Multiple cadherin domains form Ca^{2+}-dependent rod-like structures with conserved Ca^{2+}-binding pockets at the interdomain region.[43] The p.Gln229Pro mutation is located in EC1, whereas the p.Asp407Asn mutation is located in EC3 (Figure 3C). It is possible that these mutations in the EC domains may affect the adhesive function of the cadherin 2 protein. On the contrary, it is also possible that *CDH2* mutations may not have a primary effect on the adhesive properties of the protein but, as recently suggested by functional studies on mutant desmosomal proteins, may have an effect on cell responses to mechanical stress.[44]

The fact that cadherin 2 mutations may underlie a cardiomyopathy with increased arrhythmic propensity in humans has been suspected and speculated for over a decade.[38,39,41] A candidate-gene analysis screened the cadherin 2 gene in a series of genotype-negative patients with ARVC and failed to identify disease-causing mutations in a small series of 14 patients.[45] This is not surprising because our results suggest that *CDH2* mutations are not a frequent cause of ARVC (≈2.7% in our series). However, they do provide evidence that *CDH2* mutations may explain a proportion of the 40% of genotype-negative patients with ARVC.

In conclusion, our data provide novel insight into the pathogenesis of ARVC by changing the focus of field of action in ARVC from the desmosome, to the intercalated disc.

Limitations

The undeniable attraction of identifying cadherin 2 as a likely new gene for ARVC is tempered by the objective realization that our study is not without limitations. The most important one is the current lack of functional data, and such a study will be needed. The number of affected patients is still limited, and low prevalence variants of *CDH2* with a possible functional role are also present in the general population, but at a significant lower frequency. The findings need to be validated in a large cohort of patients with ARVC.

Acknowledgments

We thank the family and patients who participated in this study, and Sisters Unita September and Veronica Francis for technical assistance. We are also grateful to Serena Pisoni for assisting with linkage analysis, Associate Professor Helen Wainwright for supplying the histology slide, and to Pinuccia De Tomasi for expert editorial support. We would also like to thank the Exome Aggregation Consortium and the groups that provided exome variant data for comparison. A full list of contributing groups can be found at http://exac.broadinstitute.org/about. AQ13

Sources of Funding

This study was funded through grants from the South African Medical Research Council, the National Research Foundation of South Africa, the Lily and Ernst Hausmann Research Trust, the International Centre

for Genetic Engineering and Biotechnology, the Italian Government, the Italian Ministry of Foreign Affairs and International Cooperation grant PGR00198 entitled "Modifier genes and the risk for sudden cardiac death for channelopathies and cardiomyopathies: when molecular biology refines risk stratification", the Wellcome Trust (099313/Z/12/Z and 099313/B/12/Z), and Xenon Pharmaceuticals. Dr Kraus was a Helen and Morris Mauerberger Clinical Scholar during the conduct of this study. This work was also supported by the DZHK (German Centre for Cardiovascular Research) and by the BMBF (German Ministry of Education and Research).

Disclosures

None.

References

1. Marcus FI, McKenna WJ, Sherrill D, Basso C, Bauce B, Bluemke DA, et al. Diagnosis of arrhythmogenic right ventricular cardiomyopathy/dysplasia: proposed modification of the task force criteria. *Circulation*. 2010;121:1533–1541. doi: 10.1161/CIRCULATIONAHA.108.840827.
2. Delmar M, McKenna WJ. The cardiac desmosome and arrhythmogenic cardiomyopathies: from gene to disease. *Circ Res*. 2010;107:700–714. doi: 10.1161/CIRCRESAHA.110.223412.
3. Watkins H, Ashrafian H, Redwood C. Inherited cardiomyopathies. *N Engl J Med*. 2011;364:1643–1656. doi: 10.1056/NEJMra0902923.
4. van Tintelen JP, Entius MM, Bhuiyan ZA, Jongbloed R, Wiesfeld AC, Wilde AA, et al. Plakophilin-2 mutations are the major determinant of familial arrhythmogenic right ventricular dysplasia/cardiomyopathy. *Circulation*. 2006;113:1650–1658. doi: 10.1161/CIRCULATIONAHA.105.609719.
5. Merner ND, Hodgkinson KA, Haywood AF, Connors S, French VM, Drenckhahn JD, et al. Arrhythmogenic right ventricular cardiomyopathy type 5 is a fully penetrant, lethal arrhythmic disorder caused by a missense mutation in the TMEM43 gene. *Am J Hum Genet*. 2008;82:809–821. doi: 10.1016/j.ajhg.2008.01.010.
6. Tiso N, Stephan DA, Nava A, Bagattin A, Devaney JM, Stanchi F, et al. Identification of mutations in the cardiac ryanodine receptor gene in families affected with arrhythmogenic right ventricular cardiomyopathy type 2 (ARVD2). *Hum Mol Genet*. 2001;10:189–194.
7. Roux-Buisson N, Gandjbakhch E, Donal E, Probst V, Deharo JC, Chevalier P, et al. Prevalence and significance of rare RYR2 variants in arrhythmogenic right ventricular cardiomyopathy/dysplasia: results of a systematic screening. *Heart Rhythm*. 2014;11:1999–2009. doi: 10.1016/j.hrthm.2014.07.020.
8. Beffagna G, Occhi G, Nava A, Vitiello L, Ditadi A, Basso C, et al. Regulatory mutations in transforming growth factor-beta3 gene cause arrhythmogenic right ventricular cardiomyopathy type 1. *Cardiovasc Res*. 2005;65:366–373. doi: 10.1016/j.cardiores.2004.10.005.
9. Lazzarini E, Jongbloed JD, Pilichou K, Thiene G, Basso C, Bikker H, et al. The ARVD/C genetic variants database: 2014 update. *Hum Mutat*. 2015;36:403–410. doi: 10.1002/humu.22765.
10. Groeneweg JA, Bhonsale A, James CA, te Riele AS, Dooijes D, Tichnell C, et al. Clinical Presentation, Long-Term Follow-Up, and Outcomes of 1001 Arrhythmogenic Right Ventricular Dysplasia/Cardiomyopathy Patients and Family Members. *Circ Cardiovasc Genet*. 2015;8:437–446. doi: 10.1161/CIRCGENETICS.114.001003.
11. Watkins DA, Hendricks N, Shaboodien G, Mbele M, Parker M, Vezi BZ, et al.; ARVC Registry of the Cardiac Arrhythmia Society of Southern Africa (CASSA). Clinical features, survival experience, and profile of plakophylin-2 gene mutations in participants of the arrhythmogenic right ventricular cardiomyopathy registry of South Africa. *Heart Rhythm*. 2009;6(11 Suppl):S10–S17. doi: 10.1016/j.hrthm.2009.08.018.
12. Platonov PG, Calkins H, Hauer RN, Corrado D, Svendsen JH, Wichter T, et al. High interobserver variability in the assessment of epsilon waves: Implications for diagnosis of arrhythmogenic right ventricular cardiomyopathy/dysplasia. *Heart Rhythm*. 2016;13:208–216. doi: 10.1016/j.hrthm.2015.08.031.
13. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25:1754–1760. doi: 10.1093/bioinformatics/btp324.
14. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al; 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25:2078–2079. doi: 10.1093/bioinformatics/btp352.
15. Ye K, Schulz MH, Long Q, Apweiler R, Ning Z. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. *Bioinformatics*. 2009;25:2865–2871. doi: 10.1093/bioinformatics/btp394.
16. Pilichou K, Thiene G, Bauce B, Rigato I, Lazzarini E, Migliore F, et al. Arrhythmogenic cardiomyopathy. *Orphanet J Rare Dis*. 2016;2:11–33.
17. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010;7:248–249. doi: 10.1038/nmeth0410-248.
18. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res*. 2003;31:3812–3814.
19. Schwarz JM, Rödelsperger C, Schuelke M, Seelow D. MutationTaster evaluates disease-causing potential of sequence alterations. *Nat Methods*. 2010;7:575–576. doi: 10.1038/nmeth0810-575.
20. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet*. 2014;46:310–315. doi: 10.1038/ng.2892.
21. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. *Nat Biotechnol*. 2011;29:24–26. doi: 10.1038/nbt.1754.
22. Matolweni LO, Bardien S, Rebello G, Oppon E, Munclinger M, Ramesar R, et al. Arrhythmogenic right ventricular cardiomyopathy type 6 (ARVC6): support for the locus assignment, narrowing of the critical region and mutation screening of three candidate genes. *BMC Med Genet*. 2006;7:29. doi: 10.1186/1471-2350-7-29.
23. Munclinger MJ, Patel JJ, Mitha AS. Follow-up of patients with arrhythmogenic right ventricular cardiomyopathy dysplasia. *S Afr Med J*. 2000;90:61–68.
24. Safran M, Dalah I, Alexander J, Rosen N, Iny Stein T, Shmoish M, et al. GeneCards Version 3: the human gene integrator. *Database (Oxford)*. 2010;2010:baq020.
25. Siepel A, Bejerano G, Pedersen JS, Hinrichs AS, Hou M, Rosenbloom K, et al. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res*. 2005;15:1034–1050. doi: 10.1101/gr.3715005.
26. Pollard KS, Hubisz MJ, Rosenbloom KR, Siepel A. Detection of nonneutral substitution rates on mammalian phylogenies. *Genome Res*. 2010;20:110–121. doi: 10.1101/gr.097857.109.
27. Schwartz PJ, Spazzolini C, Crotti L, Bathen J, Amlie JP, Timothy K, et al. The Jervell and Lange-Nielsen syndrome: natural history, molecular basis, and clinical outcome. *Circulation*. 2006;113:783–790. doi: 10.1161/CIRCULATIONAHA.105.592899.
28. Barrett JH, Teare MD. Linkage analysis. *Methods Mol Biol*. 2011;760:19–33. doi: 10.1007/978-1-61779-176-5_2.
29. Yan J, Takahashi T, Ohura T, Adachi H, Takahashi I, Ogawa E, et al. Combined linkage analysis and exome sequencing identifies novel genes for familial goiter. *J Hum Genet*. 2013;58:366–377. doi: 10.1038/jhg.2013.20.
30. Knudsen KA, Soler AP, Johnson KR, Wheelock MJ. Interaction of alpha-actinin with the cadherin/catenin cell-cell adhesion complex via alpha-catenin. *J Cell Biol*. 1995;130:67–77.
31. Sheikh F, Ross RS, Chen J. Cell-cell connection to cardiac disease. *Trends Cardiovasc Med*. 2009;19:182–190. doi: 10.1016/j.tcm.2009.12.001.
32. Franke WW, Borrmann CM, Grund C, Pieperhoff S. The area composita of adhering junctions connecting heart muscle cells of vertebrates. I. Molecular definition in intercalated disks of cardiomyocytes by immunoelectron microscopy of desmosomal proteins. *Eur J Cell Biol*. 2006;85:69–82. doi: 10.1016/j.ejcb.2005.11.003.
33. Balse E, Steele DF, Abriel H, Coulombe A, Fedida D, Hatem SN. Dynamic of ion channel expression at the plasma membrane of cardiomyocytes. *Physiol Rev*. 2012;92:1317–1358. doi: 10.1152/physrev.00041.2011.
34. Knudsen KA, Wheelock MJ. Plakoglobin, or an 83-kD homologue distinct from beta-catenin, interacts with E-cadherin and N-cadherin. *J Cell Biol*. 1992;118:671–679.
35. Sacco PA, McGranahan TM, Wheelock MJ, Johnson KR. Identification of plakoglobin domains required for association with N-cadherin and alpha-catenin. *J Biol Chem*. 1995;270:20201–20206.
36. Borrmann CM, Grund C, Kuhn C, Hofmann I, Pieperhoff S, Franke WW. The area composita of adhering junctions connecting heart muscle cells of vertebrates. II. Colocalizations of desmosomal and fascia adhaerens molecules in the intercalated disk. *Eur J Cell Biol*. 2006;85:469–485. doi: 10.1016/j.ejcb.2006.02.009.
37. van Hengel J, Calore M, Bauce B, Dazzo E, Mazzotti E, De Bortoli M, et al. Mutations in the area composita protein αT-catenin are associated with arrhythmogenic right ventricular cardiomyopathy. *Eur Heart J*. 2013;34:201–210. doi: 10.1093/eurheartj/ehs373.
38. Kostetskii I, Li J, Xiong Y, Zhou R, Ferrari VA, Patel VV, et al. Induced deletion of the N-cadherin gene in the heart leads to dissolution of the

intercalated disc structure. *Circ Res*. 2005;96:346–354. doi: 10.1161/01.RES.0000156274.72390.2c.

39. Li J, Patel VV, Kostetskii I, Xiong Y, Chu AF, Jacobson JT, et al. Cardiac-specific loss of N-cadherin leads to alteration in connexins with conduction slowing and arrhythmogenesis. *Circ Res*. 2005;97:474–481. doi: 10.1161/01.RES.0000181132.11393.18.
40. Kaplan SR, Gard JJ, Protonotarios N, Tsatsopoulou A, Spiliopoulou C, Anastasakis A, et al. Remodeling of myocyte gap junctions in arrhythmogenic right ventricular cardiomyopathy due to a deletion in plakoglobin (Naxos disease). *Heart Rhythm*. 2004;1:3–11. doi: 10.1016/j.hrthm.2004.01.001.
41. Li J, Levin MD, Xiong Y, Petrenko N, Patel VV, Radice GL. N-cadherin haploinsufficiency affects cardiac gap junctions and arrhythmic susceptibility. *J Mol Cell Cardiol*. 2008;44:597–606. doi: 10.1016/j.yjmcc.2007.11.013.
42. Chen SN, Gurha P, Lombardi R, Ruggiero A, Willerson JT, Marian AJ. The hippo pathway is activated and is a causal mechanism for adipogenesis in arrhythmogenic cardiomyopathy. *Circ Res*. 2014;114:454–468. doi: 10.1161/CIRCRESAHA.114.302810.
43. Ivanov DB, Philippova MP, Tkachuk VA. Structure and functions of classical cadherins. *Biochemistry (Mosc)*. 2001;66:1174–1186.
44. Hariharan V, Asimaki A, Michaelson JE, Plovie E, MacRae CA, Saffitz JE, et al. Arrhythmogenic right ventricular cardiomyopathy mutations alter shear response without changes in cell-cell adhesion. *Cardiovasc Res*. 2014;104:280–289. doi: 10.1093/cvr/cvu212.
45. Campuzano O, Alcalde M, Berne P, Castro V, Guzzo G, Iglesias A, et al. Genetic testing of candidate genes in arrhythmogenic right ventricular cardiomyopathy/dysplasia. *Eur J Med Genet*. 2012;55:225–234. doi: 10.1016/j.ejmg.2012.02.007.

CLINICAL PERSPECTIVE

Arrhythmogenic right ventricular cardiomyopathy (ARVC) is characterized by replacement of the right ventricular myocardium by fibrofatty tissue, causing ventricular arrhythmia and sudden death in young people. This heritable condition is caused mainly by mutations in genes coding for cell adhesion proteins of the desmosome. The causative gene is, however, not known in ≈40% of cases. Here, we report the identification of a novel mutation in cadherin 2 gene (*CDH2*; c.686A>C; p.Q229P) by exome sequencing that segregates with ARVC in a South African family. *CDH2* encodes cadherin 2 (also known as N-cadherin), a protein that is located at the intercalated disc, which plays a vital role in cell adhesion. The *CDH2* c.686A>C variant was not present in >200 000 chromosomes available through public databases, changes a conserved amino acid of cadherin 2 protein and was supported as the causal mutation by parametric linkage analysis. Furthermore, a rare *CDH2* mutation (c.1219G>A; p.D407N) was identified in 1 of 73 unrelated ARVC cases that were screened. We have identified mutations in *CDH2* as novel genetic causes of ARVC in 2/74 (2,7%) South African ARVC probands in whom a known disease-causing mutation was excluded. The findings of this study have important implications for diagnostic evaluation, screening, and genetic counseling of patients with ARVC. Mutation screening of the *CDH2* gene should be considered for patients with ARVC. Our data provide a novel insight into the pathogenesis of ARVC by extending the focus of pathogenic investigation from the desmosome, to the intercalated disc.

AUTHOR QUERIES

Author Please answer all Queries

AQ1—Gene designations are italicized when referring to the gene or an allele but not when referring to the gene product. Please check and make appropriate corrections throughout the article where necessary.

AQ2—Please note only those terms that are used 5 times or more can be abbreviated, except trial names, which should be expanded at first use but then can be abbreviated throughout regardless of how many times they appear.

AQ3—Please turn to page 3 of your proof and review the running head, which will appear in the upper right-hand margins of odd-numbered pages. Running heads must be 50 or fewer characters in length, including spaces and punctuation. If your original short title was longer than 50 characters, we may have shortened it. Please modify if necessary (but observe our length guidelines).

AQ4—Please confirm that all authors are included in the correct order in the byline and that all names are spelled correctly, including special characters, accents, middle initials, and degrees, if applicable. For indexing purposes, confirm author names have been correctly identified as given names (blue) and surnames (red). Color in the byline will not appear on the final published version. Note that journal style discourages listing honorary degrees (FAHA, FRCP, etc.) in the byline; please delete such degrees from the author byline

AQ5—Please confirm that all authors' institutional affiliations (including city/state/country locations) are correct as shown in the affiliations footnote.

AQ6—Please provide department name for the required affiliations.

AQ7—Key words may have been edited to match the US National Library of Medicine's Medical Subject Headings (http://www.nlm.nih.gov/mesh/MBrowser.html). If they need modification, please limit the total number of key words to 7.

AQ8—Please provide author name with initials, degree and year for unpublished data as per style.

AQ9—Please review the typeset tables carefully against copies of the originals to verify accuracy of editing and typesetting

AQ10—Please define "BORDERLINE, DEFINITE, LVEDD, LVEF, RVEF, RVOT, TFC, VF" in Table 1 footnote

AQ11—Per style, quotes should not be used for emphasis. Hence, they have been deleted throughout the article. Please confirm whether the change made throughout is appropriate.

AQ12—Please define "LOD."

AQ13—Please carefully review any Acknowledgments, Sources of Funding, and/or Disclosures listed at the end of the manuscript (before the References), and confirm that they are accurate and complete for all authors.

AQ14—Please note that per style, "etiologic" has been changed to "pathogenic."