


SCIENTIFIC REPORTS



OPEN

Copy number variants implicate cardiac function and development pathways in earthquake-induced stress cardiomyopathy

Cameron J. Lacey¹, Kit Doudney², Paul G. Bridgman³, Peter M. George^{2,9}, Roger T. Mulder¹, Julie J. Zarifeh⁴, Bridget Kimber¹, Murray J. Cadzow⁵, Michael A. Black⁵, Tony R. Merriman⁵, Klaus Lehnert⁶, Vivienne M Bickley², John F. Pearson⁷, Vicky A. Cameron⁸ & Martin A. Kennedy^{1,9} 

The pathophysiology of stress cardiomyopathy (SCM), also known as takotsubo syndrome, is poorly understood. SCM usually occurs sporadically, often in association with a stressful event, but clusters of cases are reported after major natural disasters. There is some evidence that this is a familial condition. We have examined three possible models for an underlying genetic predisposition to SCM. Our primary study cohort consists of 28 women who suffered SCM as a result of two devastating earthquakes that struck the city of Christchurch, New Zealand, in 2010 and 2011. To seek possible underlying genetic factors we carried out exome analysis, genotyping array analysis, and array comparative genomic hybridization on these subjects. The most striking finding was the observation of a markedly elevated rate of rare, heterogeneous copy number variants (CNV) of uncertain clinical significance (in 12/28 subjects). Several of these CNVs impacted on genes of cardiac relevance including *RBFOX1*, *GPC5*, *KCNRG*, *CHODL*, and *GPBP1L1*. There is no physical overlap between the CNVs, and the genes they impact do not appear to be functionally related. The recognition that SCM predisposition may be associated with a high rate of rare CNVs offers a novel perspective on this enigmatic condition.

Stress cardiomyopathy (SCM), also known as “broken heart syndrome” or takotsubo syndrome^{1,2}, is a condition that captures widespread public interest. The cardiomyopathy is distinctive and the precipitating emotional event is typically clearly defined, however the mechanism for the cardiomyopathy and links between the psychological event and the physical illness are not understood.

Sporadic cases of SCM are estimated to account for 1–5% of acute coronary syndrome presentations^{3–5}. Predominantly the condition occurs in post-menopausal women^{6,7}, and because of this, 5–10% of female presentations with suspected acute coronary syndrome are attributed to SCM^{8–11}. Although SCM can be fatal, the symptoms are commonly transient and patients generally have a good prognosis and recover well over a period of days to weeks¹². In classic descriptions the cardiomyopathy has a typical pattern but a number of variations are now widely recognised and it is increasingly apparent that cases can be quite heterogeneous¹³.

SCM occurring in clusters around the time of major disasters such as earthquakes, floods and bushfires is also well recognised^{6,14–16}. Due to the large impact these events have upon hospital resources and medical infrastructure, it is rare for such clusters of SCM to be studied in any depth. This was made clear in reports from the Great East Japan Earthquake¹⁷. In the Canterbury (New Zealand) earthquake sequence of 2010 and 2011 the

¹Department of Psychological Medicine, University of Otago, Christchurch, New Zealand. ²Molecular Pathology Laboratory, Canterbury Health Laboratories, Canterbury District Health Board, Christchurch, New Zealand.

³Department of Cardiology, Christchurch Hospital, Christchurch, New Zealand. ⁴Psychiatric Consultation Service, Christchurch Hospital, Canterbury District Health Board, Christchurch, New Zealand. ⁵Department of Biochemistry, University of Otago, Dunedin, New Zealand. ⁶School of Biological Sciences, University of Auckland, Auckland, New Zealand. ⁷Biostatistics and Computational Biology Unit, University of Otago, Christchurch, New Zealand. ⁸Christchurch Heart Institute, Department of Medicine, University of Otago, Christchurch, New Zealand. ⁹Department of Pathology and Biomedical Science, University of Otago, Christchurch, New Zealand. Cameron J. Lacey and Kit Doudney contributed equally to this work. Correspondence and requests for materials should be addressed to C.J.L. (email: cameron.lacey@otago.ac.nz) or M.A.K. (email: martin.kennedy@otago.ac.nz)

two main events precipitated large case clusters of SCM^{17–22}. Unusually for a major natural disaster, the tertiary hospital in Christchurch continued to function, allowing the collection of a relatively large homogenous cohort of cases which have been followed over several years^{18–23}. Most research around this disorder has focused on sporadic SCM associated with heterogenous triggers^{4,24,25}. Although the presentation of earthquake-associated SCM (EqSCM) appears to be similar to that of sporadic cases, a key difference is the homogenous nature of the trigger.

Various mechanisms have been postulated for takotsubo cardiomyopathy, including that the syndrome arises from stunning of the heart muscle (myocardium) as a result of either ischemia from spasm of the coronary arteries, or from the direct effect of catecholamines (dopamine, adrenaline or noradrenaline) on cardiac myocytes^{4,24,26,27}. Despite suggestive pathophysiological observations and theories, most authors conclude that the aetiology of SCM is poorly understood, and we do not yet have satisfactory explanations for the origins of this condition^{12,24,26,28–30}. Some retrospective case series have suggested that the incidence of SCM is increased in patients with anxiety conditions, but in our studies we did not find any correlation with psychiatric or anxiety disorders^{19,23}.

Amongst the models that may be proposed for SCM aetiology, it is worth considering the possible contribution of genetic factors. Many forms of cardiomyopathy have genetic origins^{31,32}. Hypertrophic cardiomyopathy is the most common form of familial heart disease and a leading cause of sudden cardiac death. It is inherited in an autosomal dominant Mendelian manner with variable expressivity and age-related penetrance³¹. These cardiomyopathies show considerable genetic heterogeneity, with cases now attributed to some 1400 mutations in 11 genes, all of which contribute to cardiac sarcomere function. Familial dilated cardiomyopathy is also frequently attributable to an underlying genetic predisposition and at least 50 genes have now been implicated, with most eliciting disease as dominant mutations³².

Evidence for genetic contributions to SCM are not as strong as for other cardiomyopathies. However, there are several examples of familial occurrence of SCM involving siblings^{33–35} or mother-daughter pairs^{36–40}, and a large Swedish study of SCM identified three families in which several close relatives developed the condition⁴¹. The overall rarity of SCM would suggest that these familial clusters are significant, and it is quite possible that more overt familial relationships in this disorder are obscured by the simultaneous requirement for two key circumstances (in most cases): post-menopausal status and environmental exposure to a sudden major stressful event. Occasional cases of SCM occur in younger women or males, and a proportion of patients report no preceding stressor⁴², suggesting that an intrinsic pathogenic mechanism is involved. The recurrence of SCM in some patients, including one Christchurch EqSCM case²², also implies a biological vulnerability.

These observations have prompted consideration of genetic susceptibility to this condition^{37,40,43,44}. Until recently, genetic studies were restricted to candidate gene analysis in case series of sporadic SCM patients^{45–49}, but these have yielded mainly negative findings. One candidate gene study reported a significant difference in the frequency of a *GRK5* polymorphism in cases⁵⁰, but this has not been replicated and past history of single gene association studies suggests it is unlikely to be meaningful⁵¹. More recently, another candidate gene study has implicated estrogen receptor genes as potential risk factors for SCM⁵². In an effort to capture genome-wide data, exome sequencing^{53–56} was recently applied to a sample of sporadic SCM cases⁴⁰. Although this analysis did not reveal any difference in allele frequency or burden between SCM cases and population controls (28 adults with normal echocardiograms), it was noted that two thirds of the cases carried a rare deleterious variant within at least one gene of a large set of adrenergic pathway genes, and 11 genes harboured a variant in two or more cases. However, the significance of these rare variants remains unclear.

In this study, we set out to explore the role of genetic factors in predisposition to EqSCM. We specifically tested three discrete hypotheses for potential genetic contributions to risk of SCM: (i) an essentially Mendelian hypothesis that rare genetic variants in one or a few key genes cause predisposition, which was tested by whole exome sequencing (WES); (ii) that SCM was a complex disorder with genetic contributions from multiple common variants, which was tested using the Cardio-MetaboChip genotyping array; and (iii) that rare copy number variants (CNV) impacting on relevant genes contribute to risk, which was tested by array comparative genomic hybridization (aCGH).

Results

Exome Analysis. To test the potential for an essentially Mendelian predisposition to EqSCM, WES was carried out on 24 of the 28 Christchurch EqSCM cases. Several approaches to analysis of the identified variants were used, all of them hypothesising the involvement of gene variants with a low population minor allele frequency (MAF), that were over-represented in the EqSCM cohort. We carried out various iterations of filtering using variant allele frequency data derived from large population databases (1000 Genomes; NHBLI Exome Sequencing Project), followed by careful manual inspection of remaining variants. For example, excluding all variants present in these databases with an allele frequency >3%, and selecting for any present in at least 4/24 EqSCM exomes, identified variants in 131 genes, none of which proved to be convincing on closer analysis. We also carried out ranking of gene variants by predicted functional impact using various approaches^{57–59}. Once again, none of the variants identified in these analyses proved to be significantly enriched amongst our EqSCM exomes (see supplementary information for more details).

Finally, the 11 genes listed by Goodloe *et al.*⁴⁰, as well as a gene recently proposed to play a role in SCM, *BAG3*⁴⁴, were carefully examined for presence of any rare variants in the EqSCM dataset. None of the previously identified variants^{40,44}, and no other convincing rare variants in these genes, were detected.

Cardio-MetaboChip Analysis. To test the possibility of a more complex, polygenetic basis to SCM risk, involving multiple variants of small effect size, Cardio-MetaboChip analysis was carried out. The custom Illumina iSelect Cardio-MetaboChip genotyping array (see supplementary information for more details) was designed to test ~200,000 single gene polymorphisms (SNPs), identified through genome-wide meta-analyses for metabolic,

Case ID	Prediction*	Location	Del/Dup	Key genes	Size	Frequency	Location
EqSCM 01	UCS	16p13.2	DEL	RBFOX1**	139 kb	1/20000	HG19 chr16:7,054,481-7,193,526
EqSCM 03	UCS	21q21.1	DEL	CHODL**	96 kb	1/9000	HG19 chr21:19,537,178-19,633,275
	CS	22q11.22	DUP	From MIR650 to IGL	2 Mb		HG19 chr22:23,055,148-24,991,669
EqSCM 04	UCS	1p34.1	DUP	IPP** PIK3R3 (+3 others)	455 kb	1/7000	HG19 chr1:46,144,479-46,599,815
EqSCM 05	UCS	Xq22.3	DUP	IRS4** COL4A5**	113 kb	1/2000	HG19 chrX:107,896,435-108,009,609
EqSCM 06	UCS	13q31.5	DEL	GPC5**	207 kb	1/5000	HG19 chr13:92,075,673-92,283,600
	UCS	4q21.21	DUP	FRAS1 ANX3** LINC01094	329 kb	1/29000	HG19 chr4:79,281,048-79,610,796
EqSCM 07	UCS	12q13.13	DUP	17 Keratin genes	411 kb	0	HG19 chr12:52,657,396-53,069,013
EqSCM 10	UCS	19q13.42	DUP	GP6**	130 kb	1/15000	HG19 chr19:55,439,927-55,570,442
EqSCM 11	UCS	22q11.21	DUP	YPEL1** (+2 others)	180 kb	0	HG19 chr22:22,008,249-22,189,094
	UCS	Xp11.4	DUP	TSPAN7**	140 kb	1/1000	HG19 chrX:38,485,991-38,626,762
EqSCM 15	UCS	10q23.1	DUP	upstream NRG3**	78 kb	0	HG19 chr10:83,506,502-83,585,097
EqSCM 17	UCS	13q33.1	DEL	ITGBL1**	148 kb	0	HG19 chr13:102,148,514-102,296,766
EqSCM 19	?CS	13q14.2	DEL	KCNRG** (+6 others)	886 kb	0	HG19 chr13:50,585,186-51,452,033
	UCS	13q34	DUP	LINC00346, ANKRD10	146 kb	0	HG19 chr13:111,385,673-111,532,564
	UCS	13q14.11	DUP	ENOX1**	149 kb	0	HG19 chr13:89,219,432-89,359,036
EqSCM 27	UCS	7q31.1	DEL	LRRN3** IMMP2L	105 kb	1/12000	HG19 chr7:110,744,611-110,849,681

Table 1. Rare CNVs detected in a cohort of 28 EqSCM cases. *UCS = unclear clinical significance, ?CS = potential clinical significance. **Genes with clear cardiovascular association. Frequency calculations are based on larger (>1000 individuals) studies included in the DGV. A full list of all CNVs detected for each individual is presented in Supplementary Table S1.

atherosclerotic and cardiovascular diseases and traits⁶⁰. The Cardio-MetaboChip data for 27 EqSCM cases and 133 heart-healthy controls (see supplementary information for more details) were compared by logistic regression (adjusted for ethnicity and gender, additive genetic model), first performed for pairwise comparisons across groups. No SNPs reached statistical significance of <0.05 after adjusting for false discovery rate (FDR) when comparing either the EqSCM and HVOLs, or the EqSCM and CHD samples. To investigate whether the top 100 of these SNPs mapped to gene pathways that might assist in understanding potential disease mechanisms underlying SCM, pathway analysis of the leading 100 SNPs in the EqSCM versus HVOLs pairwise comparison was performed in MetaCore. Disease Biomarker Pathway analysis identified Myocardial Ischemia as the third most enriched pathway (FDR-adjusted $p = 1.3e^{-2}$), featuring 11 SNP loci on our list out of 886 pathway objects, including annexin V, ANRIL, COL4A1, dynein, HXK4, nectin-2, PPAR-gamma, prolidase, Tcf(Lef), UGT, and VEGFR-2.

aCGH Analysis. To test for potential involvement of CNVs in SCM, we applied aCGH to all cases. Of the 28 EqSCM cases examined by aCGH, twelve (42%) showed evidence of large, rare heterozygous CNVs classified as being of unclear clinical significance (Table 1), meaning that insufficient evidence is available for unequivocal determination of clinical significance⁶¹. Of all 17 CNVs detected in the 12 patients, six were deletions and 11 were duplications. All of the CNVs were different, and there was no physical overlap between the various CNVs. Each of these rare CNVs encompasses one or more genes, or their immediate upstream regulatory regions, and many of the genes included within the CNVs have functions of cardiac relevance. A full list of all CNVs detected in the cohort is presented in Supplementary Table S1.

Three cases (EqSCM 01, 06 and 19) harboured deletions very likely to impact genes of high relevance to cardiomyopathy or cardiac function. In EqSCM 01, intragenic deletion of *RBFOX1* results in a single copy loss of one exon used by the majority of transcripts predicted for the gene (Fig. 1). This exon contains the start methionine for the *RBFOX1* protein, meaning the gene is most likely rendered non-functional. *RBFOX1* is an important RNA-binding protein mediating the incorporation of microexons into many transcripts associated with neurological patterning and tissue development^{62,63}, particularly in the brain, heart and muscles. Intragenic deletions in *RBFOX1* have been observed in a range of conditions, including occasional cases with cardiac defects^{64–66}. Furthermore, *RBFOX1*-mediated RNA splicing was also recently shown to be an important regulator of cardiac hypertrophy and heart failure⁶⁷.

In the second case (EqSCM 06, Fig. 2), a heterozygous deletion encompassed exon 2 and the majority of intron 2 of the *Glypican 5* (*GPC5*) locus. *GPC5* encodes a cell surface proteoglycan, which binds to the outer surface of the plasma membrane in the cardiovascular system and displays diverse functions including blood vessel formation after ischemic injury and proliferation of smooth muscle cells during atherogenesis⁶⁸. *GPC5* was also implicated by GWAS as a protective locus for sudden cardiac arrest⁶⁹, and other glypicans (*GPC3*, 4 and 6) have been associated with cardiac dysfunction⁷⁰. This case (EqSCM 06) also harbours a duplication on 4q21.21 involving the *ANXA3* gene, which encodes a member of the annexin family, annexin A3. Members of this calcium-dependent phospholipid-binding protein family have a range of functions in the regulation of cellular growth and signal transduction pathways. Annexin A6 for example is the most abundant annexin expressed in the heart and its overexpression in mice has been shown to cause physiological alterations in contractility leading to dilated cardiomyopathy, while Annexin A6 knockout has been found to induce faster changes in Ca^{2+} transience and



Figure 1. CNV detected in EqSCM case 01. **(A)** Chromosomal location of the CNV at the RBFOX1 locus of chromosome 16. **(B)** Enlargement of the fifteen probe deletion (139 kb, delimited by vertical green lines and blue shading) illustrating loss of the fMet-containing exon (pale blue vertical bar) for three major RBFOX1 isoforms. DGV track of known CNVs shown at bottom of figure, beneath the genes and regions of interest tracks. Graphical views from Genoglyphix (PerkinElmer) software.

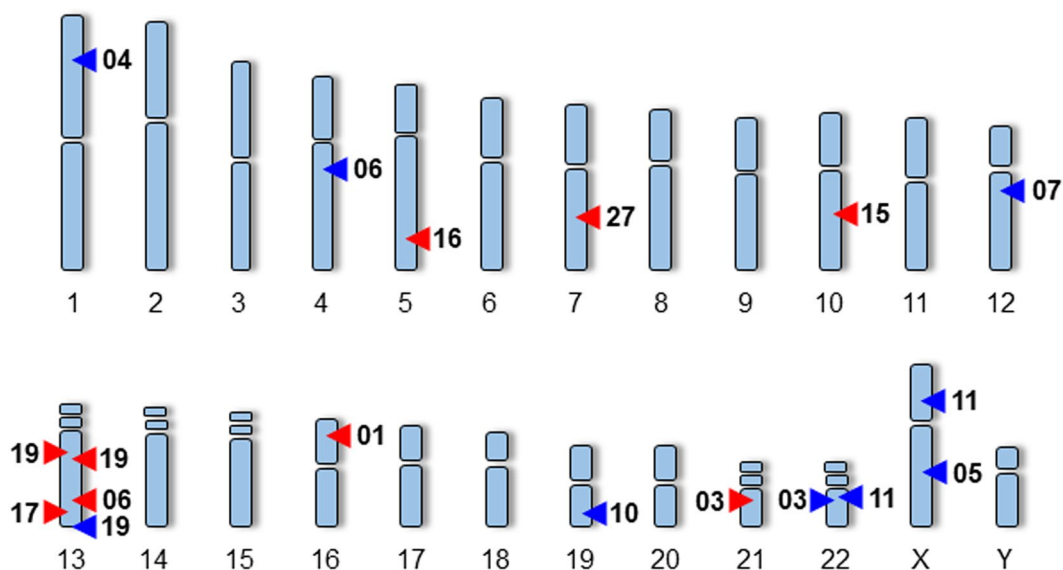


Figure 2. Genome wide distribution of CNVs. CNVs detected in 12 (of 28) EqSCM individuals by aCGH analysis. Numbers beside arrows relate to EqSCM patient number. Red arrows denote deletions, blue arrows duplications. Note that EqSCM 03, EqSCM 06, EqSCM 11 carry two rare CNVs, while EqSCM 19 contains three.

increased contractility^{71,72}. Alterations in expression and activity of annexins A5 and A7 have also been found to be associated with regulation of Ca²⁺ handling in the heart⁷³. The function of annexin A3 is not fully understood, however it has been shown to play a role in endothelial migration and vascular development⁷⁴.

The third case (EqSCM 19), contained a deletion at chr13q14.3. This region harbours at least 10 genes (*DLEU2*, *TRIM13*, *KCNRG*, *MIR16-1*, *MIR15A*, *DLEU1*, *DLEU1AS-1*, *ST13P4*, *DLEU7AS-1*, *DLEU7*, and *RNASEH2B-AS1*), including several non-coding RNAs (*DLEU* genes and micro-RNA genes) and a gene (*KCNRG*) encoding a protein involved in the regulation of voltage-gated potassium channel activity. The micro-RNA genes *mir-16-1* and *mir-15a* in this interval have been implicated in a range of cardiovascular phenotypes, including a role for *mir-15a* in postnatal mitotic arrest of cardiomyocytes^{75–77}. Two further chromosome 13 duplicated CNVs of approximately 150 kb were classified as uncertain significance - one involving *LINC00346* and *ANKRD10* and the other containing *ENOX1*, postulated to affect vascular development based on zebrafish expression patterns⁷⁸ (Table 1).

Beyond these three cases, cardiac or relevant neurological impacts appeared likely for many of the other rare CNVs identified in EqSCM cases (Table 1, Fig. 2, Supplementary Table S1), several of which are discussed below.

A 96 kb heterozygous deletion in EqSCM 03 disrupts all predicted transcripts of the chondrolectin gene (*CHODL*), a membrane bound C-type lectin involved in muscle organ development, whose protein product is detected in heart and skeletal muscle by immunohistochemistry⁷⁹. In addition to this CNV, this patient carries a 1.94 Mb duplication at 22q11.25, a locus containing 45 genes or miRNAs, associated with learning difficulties^{80,81}. This individual, who exhibited a degree of cognitive impairment, had consented for clinically relevant findings to be forwarded to their General Practice clinician, who subsequently recommended genetic counselling for this individual.

A 455 kb duplication within EqSCM 04 at 1p34.1 affects the genes *GPBP1L1*, *TMEM69*, *IPP*, *MAST2*, and *PIK3R3*. Smaller rare duplications in this region have been reported by the DGV database⁸² but none span the genes within this CNV. *GPBP1L1* is widely expressed in many tissues, including heart muscle⁸³ and predicted to be involved in transcriptional regulation. *TMEM69*, a gene of unknown function, is most strongly expressed in heart tissue⁸³. *IPP* is a transcription factor with a 50 amino acid Kelch repeat known to interact with actin, while *MAST2* contains a PDZ domain and is another gene highly expressed in heart and skeletal muscle⁸³. The protein product of *PIK3R3*, phosphoinositide-3-kinase regulatory subunit 3, acts downstream of G-protein-coupled receptors in cardiac function⁸⁴, and is also a target for isoproterenol, which can trigger SCM-like conditions in humans and rodents^{12,85–87}.

Duplication of a long non-coding RNA (LOC101928358), the 3' segment of *COL4A5* and the entire *IRS4* gene at Xq22.3 (9 probes, 113 kb) was identified within case EqSCM 05. *IRS4* is an insulin receptor molecule expressed in heart and skeletal muscle cells⁸⁸ and other tissues such as brain, kidney and liver⁸⁹. Duplication of *IRS4* may be of functional significance, although copy number increases on the X chromosome of females may be counteracted to a degree by random X inactivation. Of note is a Genoglyphix Chromosome Aberration Database (GCAD⁹⁰) case 52414 with a phenotype of low muscle tone, which has an identical duplication at this locus (as well as a 1p33 deletion). With regard to *IRS4*, Schreyer *et al.*⁸⁸ found a more restricted tissue distribution than *IRS1* and *IRS2*, in primary human skeletal muscle cells and rat cardiac muscle and isolated cardiomyocytes. Although *IRS4* protein function is still relatively unknown, the role of IRS proteins in general, acting as mediators of intracellular signalling from insulin and insulin-like growth factor 1 receptors, implicates *IRS4* in cell growth and survival⁹¹. It is interesting to note that PI 3-kinase (PI3K) signalling in HEK293T cells depends on *IRS4*, and that the IRS proteins relay signals from receptor tyrosine kinases to downstream components of signalling pathways⁸⁸, which is a connection with the *PIK3R3* gene duplicated in one of our other cases, EqSCM 04.

EqSCM 10 harboured a 130 kb duplication of *NLRP7*, *NLRP2*, *GP6* and *RDH13* at 19q13.42. One similar DGV duplication has been seen in this area (*nsv1062047*⁸²), but otherwise duplicated CNVs are generally much smaller and rare. *NLRP2* and *NLRP7* are genes that encode members of the NACHT, leucine rich repeat, and PYD containing (NLRP) protein family. These proteins are implicated in the activation of pro-inflammatory caspases. Recessive mutations in *NLRP2/7* in humans are associated with reproductive disorders⁹². Another gene in this duplicated cluster associated with disease is *GP6*, a platelet membrane glycoprotein, involved in collagen-induced platelet aggregation and thrombus formation, which is expressed at high levels in heart, kidney and whole blood⁸³. A *GP6* SNP (c.13254TC) has been implicated in recurrent cardiovascular events and mortality⁹³. Another study involving this SNP⁹⁴, found that hormone replacement therapy (HT) reduced the hazard ratio (HR) of CHD events in patients with the *GP6* 13254TT genotype by 17% but increased the HR in patients with the TC + CC genotypes by 35% (adjusted interaction $P < 0.001$). The authors found that in postmenopausal women with established CHD, the *GP6* polymorphism, and another in *GP1B*, were predictors of CHD events and significantly modified the effects of HT on CHD risk⁹⁴.

Duplication of *PPL2*, *YPEL1* and *MAPK1* on chromosome 22q11.21 (180 kb) observed in EqSCM 11, does not appear in the DGV catalogue of CNVs in healthy individuals, and the consequences of overexpression of these genes, miRNAs or regulatory sequences are unknown. One of the affected genes, *MAPK1* (previously named *ERK* or *ERK2*), may constitute a link to another kinase intracellular signalling pathway – the RAF-MEK-ERK kinase cascade, which in mice and human has an established role in the induction of cardiac tissue hypertrophy⁹⁵. Although not the kind of left ventricular enlargement seen in SCM, subtle copy number variation at *MAPK1* may influence signalling through this pathway. Another duplication in EqSCM 11 involving the 3' half of *TSPAN7*, a member of the tetraspanin protein superfamily (Xp11.4) was noted as rare, and a similar (though larger) duplication was recently observed in a patient with Rolandic epilepsy⁹⁶.

An agenic duplication 50 kb upstream from *NRG3* (10q23.1) was seen in case EqSCM 15. This CNV could conceivably disrupt upstream regulatory regions of *NRG3*, which encodes an important ligand for the transmembrane tyrosine kinase receptor ERBB4. *NRG3* has been shown to activate tyrosine phosphorylation of its cognate receptor, ERBB4, and is thought to influence neuroblast proliferation, migration and differentiation by signalling

through ERBB4. *NRG3* is a strong candidate gene for schizophrenia, and neuregulin molecules and their receptors are involved in rat cardiac development and maintenance⁹⁷.

Finally, the two rare deletions observed on chromosome 13 (13q21.33 and 13q33.1) in case EqSCM 17, fall into largely uncharacterised areas of the genome. The first is agenic, although there is a prediction of a spliced EST in the NCBI database, and the second occurs as two 50 kb blocks within the *ITGEB1* gene. *ITGEB1* is most strongly expressed in aorta^{98,99}.

Discussion

Monogenic and polygenic models of risk. The Christchurch earthquakes repeatedly exposed the entire population of the city, approximately 350,000 people, to major stress and life disruption. Almost all patients presenting with EqSCM were post-menopausal females, consistent with other reports²⁶. We set out to explore three categories of genetic contributions to SCM predisposition, using WES to explore Mendelian models of risk, Cardio-MetaboChip analysis to test for polygenic risk factors, and aCGH analysis to evaluate the role of genomic structural variants.

Extensive analysis of the WES data did not yield any apparent enrichment of rare, damaging variants within exome regions amongst the EqSCM cases. Therefore, it seems unlikely that point mutations or small insertion-deletion (indels) in a single gene underlie predisposition to earthquake SCM. A limitation of this analysis is that it would have been unable to detect regulatory mutations, or other important variants, not included or well represented within the captured exome regions. Whole genome sequencing may therefore be warranted to further test the hypothesis of Mendelian underpinnings of SCM, as this approach could identify any regulatory variants not obtained with WES, and due to the absence of a DNA capture step, would also provide more uniform coverage of exons. Of course, it is quite possible that SCM has variable or complex penetrance, and heterogeneous origins involving rare mutations, as do many other cardiomyopathies^{31,32}, and it may prove difficult to identify underlying genetic contributors in a cohort of this small size. In addition, there may be environmental factors, in addition to earthquake-induced anxiety, that contribute to onset of the condition.

In a second approach, we explored the alternative hypothesis of polygenic risk alleles of small effect size using a case-control association study, with genotypes generated by the Cardio-MetaboChip. This chip allowed genotyping of ~200,000 SNPs previously identified through genome-wide association studies (GWAS) for risk of metabolic, atherosclerotic and cardiovascular diseases and traits⁶⁰. The traits covered by the panel of genetic variants on the chip include myocardial infarction (MI) and coronary heart disease (CHD), type 2 diabetes (T2D), T2D age diagnosed, T2D early onset, mean platelet volume, platelet count, white blood cell, HDL cholesterol, LDL cholesterol, triglycerides, total cholesterol, body mass index, waist hip ratio (BMI adjusted), waist circumference (BMI adjusted), height, percent fat mass, fasting glucose, fasting insulin, 2-hour glucose, HbA1c, systolic blood pressure, diastolic blood pressure and QT interval. This analysis did not yield variants of genome-wide significance in the SCM cases compared to either healthy controls or patients with coronary disease. Exploratory pathway analysis suggested that the EqSCM cases carried a greater burden of SNPs that mapped to a myocardial ischemia pathway compared to the healthy controls, although this must be interpreted with caution as our small sample set meant very limited statistical power. Two limitations of this analysis were the relatively constrained content of the Cardio-MetaboChip, which is less able to provide a rich dataset of genome-wide SNP genotypes than the chips commonly used for GWAS, and the relatively small cohort of cases available for study. Recruitment of a much larger SCM cohort with a view to a well-powered GWAS with a more extensive genotyping chip would therefore be a worthwhile future goal to more fully explore possible polygenic underpinnings of this disorder. We note the recent publication of a preliminary GWAS on 96 SCM cases and 475 healthy controls¹⁰⁰, and believe extension of this approach to larger cohorts is an important goal.

Involvement of copy number variants. CNVs have been implicated in many diseases since the recognition a decade ago of their widespread distribution through the genome^{101–103}. Of note, rare CNVs are implicated in autism, epilepsy, schizophrenia, developmental delay and intellectual disability^{82,104–108}. Cardiac conditions which involve CNVs include congenital left-sided heart disease^{109,110}, congenital heart disease^{111,112}, some cases of long QT syndrome¹¹³, and Tetralogy of Fallot¹¹⁴. Our final analysis, therefore, was to explore the potential involvement of CNV in risk of EqSCM, using aCGH analysis of all cases. Results from this analysis were striking, with 42% of EqSCM cases having a rare CNV of unclear clinical significance. The CNV detection rate for diagnostic aCGH in childhood developmental disorders such as autism, developmental delay and intellectual disabilities, is approximately 20–30%^{115–117}. A recent report of a large New Zealand aCGH case series (5,300 pre- and post-natal tests) reported CNVs in 28.3% of these clinically-selected cases¹¹⁸. Our observation of a rate of 42% for the EqSCM case series is significantly greater ($P < 0.02$) than rates for the enriched case cohorts normally referred for clinical aCGH testing^{115–118}. The CNVs detected in EqSCM cases were all different, and there were no physical overlaps between them. This situation is similar to the pattern of CNVs seen in other conditions, including Rolandic epilepsies⁹⁶ and congenital heart disease^{109,111,112}. Many of the CNVs we observed are likely to impact genes of potential relevance to physiological processes implicated in SCM.

We have taken a relatively conservative approach to categorising CNVs, in terms of rarity and predicted functional significance. For example, we did not include two CNVs located at 9p24.3, involving individuals EqSCM 14 and 28 - a deletion and duplication, respectively. These CNVs encompassed a region including the large isoform of *DOCK8* gene. *DOCK8* encodes a protein implicated in the regulation of the actin cytoskeleton¹¹⁹, and *DOCK8* mutations cause autosomal recessive hyper-IgE syndrome¹²⁰. One reported case of a homozygous 129 kb deletion in this region was associated with Graves's disease and aortic aneurysm¹²¹. However, several deletions of *DOCK8* are recorded in the DGV for unaffected individuals, therefore the CNVs in EqSCM 14 and 28 were categorised as TBB (thought to be benign). The approximately 160kb duplication in EqSCM 28 was larger than the 44 kb deletion in EqSCM 14, and it encompassed a second gene, *KANK1*. A small number of similar duplications have been

recorded in the DGV, and therefore we did not consider this to be pathogenic. However, it is of interest that we see two relatively rare CNVs at this locus in our small EqSCM cohort.

Of the twelve EqSCM cases with rare CNVs, we consider that three (EqSCM 01, 06 and 19) contain CNVs that affect genes of high relevance to cardiomyopathy or cardiac function. The remaining CNVs are also strong candidates with potential functional relevance. In one of our most highly-ranked CNV containing cases (EqSCM 01) a large genomic deletion removes an exon of *RBFox1* which contains the start codon used by the majority of transcripts predicted for the gene. A recent report by Gao *et al.* (2016) provided strong functional data that would support our hypothesis of this gene's involvement as a susceptibility locus for SCM⁶⁷. Their work with mouse models has shown *RBFox1* deficiency in the heart promoted pressure overload-induced heart failure, and induction of *RBFox1* over-expression in these murine pressure-overload models, substantially attenuated cardiac hypertrophy and pathological manifestations⁶⁷. The haploinsufficiency seen in EqSCM 01 at the *RBFox1* locus may, in concert with other environmental or genetic factors, contribute to SCM through reduced global RNA splicing changes in the heart.

Conclusion. Beginning with a cohort of 28 SCM cases triggered by two major earthquakes that caused extensive death and damage in Christchurch (New Zealand), we carried out exploratory analyses of three models for genetic predisposition to this disorder. Using WES and Cardio-MetaboChip genotyping analyses we did not detect an obvious role for exonic mutations in a monogenic model, or SNPs in a polygenic model, for SCM risk. However, our analysis of copy number variation in SCM cases revealed a high rate of occurrence of CNV categorised as of uncertain clinical significance. Most of the CNV we detected in SCM cases were rare, or not previously seen (Table 1).

These observations lead us to propose that SCM is a copy number variant disorder, whereby haploinsufficiency of genes overlapping deletions or over-expression of duplicated genes leads to relatively subtle modification of cardiac or adrenergic physiology, such that these individuals are at increased risk of suffering SCM when exposed to specific environmental triggers. Although no obvious single pathway relationships between the genes affected by these CNVs is apparent, most of the CNVs encompass loci relevant to cardiac function or cardioneuronal development.

In order to confirm whether SCM predisposition does indeed arise from CNVs, four key areas for future work need to be pursued. First, more widespread analysis is required of CNVs in many SCM cases. This would confirm whether our observation of a high rate of CNV in EqSCM also prevails in sporadic cases, and it will broaden the catalog of affected genes, helping to discern underlying signalling networks and physiological processes. In addition, with increasing numbers of cases, physical overlaps between CNVs in different individuals should become apparent, pinpointing key genomic regions for more intensive analysis. Second, the inheritance patterns of these CNVs must be established. It is unclear what proportion are *de novo* versus inherited from either parent. Third, there is a clear need for detailed physiological and gene expression analyses on appropriate cells, including cardiomyocytes, derived from SCM cases. Given the diversity of CNV seen in our SCM cases, this goal would most effectively be achieved by generation of induced-pluripotent stem cell (iPSC) lines from many patients and appropriate controls^{122,123}. Finally, although our data implicate CNV as a significant genetic factor underlying SCM risk, it would seem wise to pursue an effective GWAS strategy to identify other genetic contributors to SCM and build on the initial study in this area¹⁰⁰. International initiatives to collate SCM cases¹³ should therefore ensure that consented DNA is available to provide appropriately large numbers of well phenotyped cases and controls to facilitate this goal.

Finally, we hope our observations implicating CNV in this unique case series of EqSCM will stimulate further studies of copy number variation in other SCM cohorts, and lead to an improved understanding of this perplexing and intriguing condition.

Material and Methods

Cases. The September 2010 earthquake of magnitude 7.1 on the Richter scale (Mw 7.1) in Christchurch (New Zealand) triggered eight cases of EqSCM, and the shallow highly destructive quake (Mw 6.3) that followed in February 2011 triggered 21 cases over four days. One woman presented after both quakes²², and one was in hospital during the initial quake. Enrolment of this latter participant was delayed, and her sample was available for aCGH analysis ($n = 28$) but not for Cardio-Metabolome analysis ($n = 27$). The steps leading to recruitment of our EqSCM cohort are detailed elsewhere^{20,21}, but briefly, our study commenced the day of the first earthquake with the creation of a register of prospectively identified earthquake stress cardiomyopathy cases. As our hospital was still functioning we could build a cohort with first-world data from complete single centre capture. After the second earthquake the study was extended^{20–22}.

Inclusion criteria: i) Meeting modified Mayo criteria for stress cardiomyopathy and admitted to Christchurch Hospital within one week of either the September 2010 or February 2011 earthquake; ii) age over 18; iii) informed consent given. Exclusion criteria: i) unable to understand English sufficiently to be able to complete questionnaires.

All participants were recruited with informed consent, including discussion of the possibility of incidental findings from genetic analyses, and return of such findings after consultation with a medical geneticist. The Southern Health and Disability Ethics Committee (New Zealand) approved this study (reference URA/11/07/033). All methods used in this research were performed in accordance with the relevant guidelines and regulations.

DNA extraction. Peripheral blood samples were obtained from consenting participants. Genomic DNA was extracted from 3 mL peripheral blood using NucleoMag extraction kits (Machery-Nagel GmbH, Düren, Germany) on a KingFisher™ Flex Magnetic liquid-handling robot (Thermo Fisher Scientific, Inc, Waltham, MA).

Stage of analysis	Number of SNPs	Comments
Assayed	196,725	
No genotype	10,927	
Mono allelic	28,367	
MAF < 0.004	44,424	SNPs with only 1 minor allele found
Missing > 0.02 of samples	36,884	
HWE fail	130	HWE $P < 1e-20$ and $1.5\times$ more heterozygotes than expected
Total Removed	65,622	Note that SNPs could fail on two or more of the criteria above
SNPs remaining in analysis	131,103	

Table 2. SNP markers removed from Cardio-MetaboChip analysis in quality control.

DNA was quantified by analysis with the Nanodrop™ (ThermoFisher), and, where appropriate, the TapeStation 4200 system (Agilent Technologies).

Exome analysis. We applied WES to a subset (24 of 28) EqSCM cases. The exome capture and sequencing was carried out in two batches of 12, during 2012–13 (New Zealand Genomics Limited, Dunedin, New Zealand). DNA was processed with Illumina TruSeq sample preparation and exome enrichment kits (which capture ~62 Mb of genomic DNA), and sequencing (100 bp paired-end reads) was carried out on an Illumina HiSeq. 2000 system. Good quality sequence was obtained across all exomes, with very few unassigned reads. Raw read data were aligned to the GRCh37 human reference genome using the Burrows-Wheeler Aligner (BWA)¹²⁴, and processed through the Broad GATK pipeline¹²⁵. The alignment process included removal of reads from duplicate fragments, realignment around known indels, and recalibration of all base quality scores. Read depth and coverage for all exome datasets were derived using Picard CollectHsMetrics function. On average, 22.7 million reads (range 16.9–30.7 M) at mean quality scores (Phred) of Q37 were mapped to the exome target intervals; 70% (range 61–77%) of targeted bases were covered at least 20-fold. Full metrics for each sample are provided in Supplementary Table S2.

Joint variant calling was performed with GATK's HaplotypeCaller. This included *de novo* assembly at each potential variant locus. Variants were annotated and analysed using Ingenuity Variant Analysis (IVA) software (QIAGEN, Redwood City, CA, USA), MutationTaster⁵⁹, SnpEff⁵⁸, SeattleSeq annotation server⁵³, and Galaxy (via usegalaxy.org)¹²⁶. Allele frequencies and additional annotations were drawn from 1000 Genomes project¹²⁷, NHLBI GO Exome Sequencing Project (ESP), Seattle, WA (<http://evs.gs.washington.edu/EVS/>), ClinVar¹²⁸, and Exome Aggregation Consortium (ExAC), Cambridge, MA (<http://exac.broadinstitute.org>). Promising gene variants were inspected by Sanger sequence analysis on the appropriate genomic DNA samples.

Cardio-MetaboChip Analysis. Three groups were genotyped using the Illumina Cardio-MetaboChip: 27 out of 28 female Christchurch EqSCM cases, plus an opportunistic control set consisting of 133 heart-healthy controls from the Canterbury Healthy Volunteers Study (HVOLs, 54 F/79 M), and 157 patients recruited for an ongoing study of premature coronary heart disease and consented for genotyping (CHD, 64 F/93 M). The heart-healthy controls were a subset of the Canterbury Healthy Volunteers Cohort (total $n > 3,400$), who were randomly selected from Canterbury electoral rolls, age- and gender-matched to existing patient cohorts and had no personal history of diagnosed cardiovascular disease at the time of recruitment¹²⁹. (See supplementary information for more details). DNA samples were run on the Cardio-MetaboChip and scanned on the Illumina® iScan platform by AgResearch Limited (Invermay, New Zealand).

Quality control with summary analysis of allele and genotype frequencies, Hardy-Weinberg equilibrium tests, and missing genotype rates were performed with PLINK version 1.07 software¹³⁰. SNPs with a minor allele frequency of < 0.05 and those that failed the Hardy-Weinberg equilibrium test ($p < 0.001$) were excluded from the analysis, leaving 141,095 SNPs in the analysis (Table 2). Three samples from the CHD Study were also removed after analysis of relatedness. Principal Component Analysis (PCA, Eigenstrat 4.2) was performed on an independent subset of almost 50,000 SNPs; the first principal component explained 6% of the variation, subsequent components all less than 0.5%, and matched self-reported ethnicity (visual inspection). Hence the first principal component was subsequently included as a factor in the logistic regression. Logistic regression was performed to evaluate differences in SNP minor allele frequencies between groups, adjusted for ethnicity and gender, using an additive genetic model (R 3.01 software¹³¹). P values were adjusted for false discovery rate (FDR) using the Benjamini Yekutieli method¹³². Pathway analysis was performed for the leading 100 SNPs in each pairwise group comparison, using MetaCore from GeneGo (Thomson Reuters).

CNV detection and analysis. Array comparative genomic hybridisation (aCGH) was undertaken on 28 EqSCM cases to examine structural variants in the cohort. For this analysis, we used either the Nimblegen 135k oligo array (CGX12) (Roche NimbleGen Inc, Madison, WI, USA), capable of genome-wide screening for CNV to a resolution of 10 kb in well-categorised pathogenic genomic regions, and 50 kb elsewhere, or the Agilent 180k HD oligo array (Sureprint G3 Human 4 × 180k) (Agilent Technologies, Santa Clara, CA, USA), which has a similar resolution (see supplementary information for more details).

Pooled reference DNA samples (catalogue numbers G147A and G152A) were purchased from Promega (Madison, WI, USA). EqSCM case and reference DNA samples (0.5–1 µg each) were labelled with Cy3 and Cy5

dyes respectively, purified, hybridized, and washed according to Nimblegen and Agilent protocols. Microarrays were scanned on a GenePix 4000B laser scanner (Axon Instruments, CA, USA) or a G2600D Agilent SureScan microarray scanner (Agilent Technologies). Data was processed using NimbleScan (Roche Nimblegen Inc) or Cytogenomics software (Agilent Technologies) with the default algorithms and analysis settings, but with a 5 probe minimum calling threshold. All arrays passed QC metrics for derivative log ratio spread (DLRS) values of <0.2. CNV data was visualised and interpreted using Genoglyphix software (Perkin Elmer) and NCBI genome browser software (genome build hg19 (GRCh37)). The EqSCM aCGH data were assessed against many CNV databases including Genoglyphix Chromosome Aberration Database (containing over 14,000 validated variants from 50,000 samples) (Perkin Elmer, Waltham, MA, USA)⁹⁰, DECIPHER¹³³, and the Database of Genomic Variation (DGV, containing CNV data from over 35,000 unaffected individuals)¹³⁴. CNVs were classified as thought to be benign (TBB), uncertain clinical significance (UCS), or clinically significant (CS) using an evidence-based approach^{61,135–137} which included database comparisons (frequency in cases/controls and relation to phenotype), gene content, gene function and dosage sensitivity. A broad summary of our CNV interpretation algorithm is depicted in Supplementary Figure S1. Rare CNVs are defined as those that occur at a frequency of ≤1%. We classified our rare CNV frequency using larger DGV studies containing >1000 individuals^{82,138–141}.

Data availability. The datasets generated during the current study are not publicly available for ethical reasons but can be made available by the corresponding author (M.A.K) on reasonable request.

References

- Dote, K., Sato, H., Tateishi, H., Uchida, T. & Ishihara, M. [Myocardial stunning due to simultaneous multivessel coronary spasms: a review of 5 cases]. *J Cardiol* **21**, 203–214 (1991).
- Tsuchihashi, K. *et al.* Transient left ventricular apical ballooning without coronary artery stenosis: a novel heart syndrome mimicking acute myocardial infarction. Angina Pectoris-Myocardial Infarction Investigations in Japan. *Journal of the American College of Cardiology* **38**, 11–18 (2001).
- Sharkey, S. W. Takotsubo cardiomyopathy: natural history. *Heart Fail Clin* **9**, 123–136 (2013).
- Sharkey, S. W. *et al.* Natural history and expansive clinical profile of stress (tako-tsubo) cardiomyopathy. *J Am Coll Cardiol* **55**, 333–341 (2010).
- Eshtehardi, P. *et al.* Transient apical ballooning syndrome—clinical characteristics, ballooning pattern, and long-term follow-up in a Swiss population. *Int J Cardiol* **135**, 370–375 (2009).
- Watanabe, H. *et al.* Impact of Earthquakes on Takotsubo Cardiomyopathy. *Jama* **294**, 305–307 (2005).
- Strunk, B. *et al.* High incidence of focal left ventricular wall motion abnormalities and normal coronary arteries in patients with myocardial infarctions presenting to a community hospital. *The Journal of invasive cardiology* **18**, 376–381 (2006).
- Sy, F. *et al.* Frequency of Takotsubo cardiomyopathy in postmenopausal women presenting with an acute coronary syndrome. *Am J Cardiol* **112**, 479–482 (2013).
- Eitel, I. *et al.* Clinical characteristics and cardiovascular magnetic resonance findings in stress (takotsubo) cardiomyopathy. *JAMA* **306**, 277–286 (2011).
- Prasad, A. *et al.* Incidence and angiographic characteristics of patients with apical ballooning syndrome (takotsubo/stress cardiomyopathy) in the HORIZONS-AMI trial: an analysis from a multicenter, international study of ST-elevation myocardial infarction. *Catheter Cardiovasc Interv* **83**, 343–348 (2014).
- Deshmukh, A. *et al.* Prevalence of Takotsubo cardiomyopathy in the United States. *American heart journal* **164**, 66–71 (2012).
- Wright, P. T., Tranter, M. H., Morley-Smith, A. C. & Lyon, A. R. Pathophysiology of takotsubo syndrome: temporal phases of cardiovascular responses to extreme stress. *Circ J* **78**, 1550–1558 (2014).
- Templin, C. *et al.* Clinical Features and Outcomes of Takotsubo (Stress) Cardiomyopathy. *N Engl J Med* **373**, 929–938 (2015).
- Stephoe, A. The impact of natural disasters on myocardial infarction. *Heart* **95**, 1972–1973 (2009).
- Butterly, S. J. *et al.* Stress-induced takotsubo cardiomyopathy in survivors of the 2011 Queensland floods. *Med J Aust* **198**, 109–110 (2013).
- Leor, J., Poole, W. K. & Kloner, R. A. Sudden cardiac death triggered by an earthquake. *N Engl J Med* **334**, 413–419 (1996).
- Aoki, T. *et al.* Effect of the Great East Japan Earthquake on cardiovascular diseases—report from the 10 hospitals in the disaster area. *Circ J* **77**, 490–493 (2013).
- Bridgman, P. G. *et al.* CTG-repeat expansions in the DMPK gene do not cause takotsubo syndrome. *Int J Cardiol* **203**, 107–108 (2016).
- Lacey, C. *et al.* Broken heart syndrome - is it a psychosomatic disorder? *J Psychosom Res* **77**, 158–160 (2014).
- Chan, C. *et al.* Acute myocardial infarction and stress cardiomyopathy following the Christchurch earthquakes. *PLoS One* **8**, e68504, <https://doi.org/10.1371/journal.pone.0068504> (2013).
- Chan, C., Troughton, R., Elliott, J., Zarifeh, J. & Bridgman, P. One-year follow-up of the 2011 Christchurch Earthquake stress cardiomyopathy cases. *N Z Med J* **127**, 15–22 (2014).
- Bridgman, P. G., Chan, C. W. & Elliott, J. M. A case of recurrent earthquake stress cardiomyopathy with a differing wall motion abnormality. *Echocardiography* **29**, E26–27, <https://doi.org/10.1111/j.1540-8175.2011.01568.x> (2012).
- Zarifeh, J. A., Mulder, R. T., Kerr, A. J., Chan, C. W. & Bridgman, P. G. Psychology of earthquake-induced stress cardiomyopathy, myocardial infarction and non-cardiac chest pain. *Intern Med J* **42**, 369–373 (2012).
- Wittstein, I. S. Stress cardiomyopathy: a syndrome of catecholamine-mediated myocardial stunning? *Cell Mol Neurobiol* **32**, 847–857 (2012).
- Coupez, E. *et al.* A single pathophysiological pathway in Takotsubo cardiomyopathy: Catecholaminergic stress. *Arch Cardiovasc Dis* **107**, 245–252 (2014).
- Wittstein, I. S. *et al.* Neurohumoral features of myocardial stunning due to sudden emotional stress. *N Engl J Med* **352**, 539–548 (2005).
- Nef, H. M. *et al.* Tako-Tsubo cardiomyopathy: intraindividual structural analysis in the acute phase and after functional recovery. *Eur Heart J* **28**, 2456–2464 (2007).
- Prasad, A., Lerman, A. & Rihal, C. S. Apical ballooning syndrome (Tako-Tsubo or stress cardiomyopathy): a mimic of acute myocardial infarction. *Am Heart J* **155**, 408–417 (2008).
- Summers, M. R., Lennon, R. J. & Prasad, A. Pre-morbid psychiatric and cardiovascular diseases in apical ballooning syndrome (tako-tsubo/stress-induced cardiomyopathy): potential pre-disposing factors? *J Am Coll Cardiol* **55**, 700–701 (2010).
- Nef, H. M., Mollmann, H., Akashi, Y. J. & Hamm, C. W. Mechanisms of stress (Takotsubo) cardiomyopathy. *Nat Rev Cardiol* **7**, 187–193 (2010).
- Maron, B. J., Maron, M. S. & Semsarian, C. Genetics of hypertrophic cardiomyopathy after 20 years: clinical perspectives. *J Am Coll Cardiol* **60**, 705–715 (2012).

32. McNally, E. M., Golbus, J. R. & Puckelwartz, M. J. Genetic mutations and mechanisms in dilated cardiomyopathy. *J Clin Invest* **123**, 19–26 (2013).
33. Musumeci, B. *et al.* Simultaneous Takotsubo syndrome in two sisters. *Int J Cardiol* **165**, e49–50 (2013).
34. Ikutomi, M. *et al.* Takotsubo cardiomyopathy in siblings. *Heart Vessels* **29**, 119–122 (2014).
35. Pison, L., De Vusser, P. & Mullens, W. Apical ballooning in relatives. *Heart* **90**, e67, <https://doi.org/10.1136/hrt.2004.046813> (2004).
36. Sharkey, S. W., Lips, D. L., Pink, V. R. & Maron, B. J. Daughter-mother tako-tsubo cardiomyopathy. *Am J Cardiol* **112**, 137–138 (2013).
37. Cherian, J., Angelis, D., Filiberti, A. & Saperia, G. Can takotsubo cardiomyopathy be familial? *Int J Cardiol* **121**, 74–75 (2007).
38. Kumar, G., Holmes, D. R. Jr. & Prasad, A. “Familial” apical ballooning syndrome (Takotsubo cardiomyopathy). *Int J Cardiol* **144**, 444–445 (2010).
39. Subban, V., Ramachandran, S., Victor, S. M., Gnanaraj, A. & Ajit, M. S. Apical ballooning syndrome in first degree relatives. *Indian Heart J* **64**, 607–609 (2012).
40. Goodloe, A. H., Evans, J. M., Middha, S., Prasad, A. & Olson, T. M. Characterizing genetic variation of adrenergic signalling pathways in Takotsubo (stress) cardiomyopathy exomes. *Eur J Heart Fail* **16**, 942–949 (2014).
41. Schultz, T. *et al.* Stress-induced cardiomyopathy in Sweden: evidence for different ethnic predisposition and altered cardio-circulatory status. *Cardiology* **122**, 180–186 (2012).
42. Pilgrim, T. M. & Wyss, T. R. Takotsubo cardiomyopathy or transient left ventricular apical ballooning syndrome: A systematic review. *Int J Cardiol* **124**, 283–292 (2008).
43. Limongelli, G. *et al.* Takotsubo cardiomyopathy: do the genetics matter? *Heart Fail Clin* **9**, 207–216 (2013).
44. d’Avenia, M. *et al.* A novel miR-371a-5p-mediated pathway, leading to BAG3 upregulation in cardiomyocytes in response to epinephrine, is lost in Takotsubo cardiomyopathy. *Cell death & disease* **6**, e1948, <https://doi.org/10.1038/cddis.2015.280> (2015).
45. Figtree, G. A. *et al.* No association of G-protein-coupled receptor kinase 5 or beta-adrenergic receptor polymorphisms with Takotsubo cardiomyopathy in a large Australian cohort. *Eur J Heart Fail* **15**, 730–733 (2013).
46. Sharkey, S. W. *et al.* Adrenergic receptor polymorphisms in patients with stress (tako-tsubo) cardiomyopathy. *J Cardiol* **53**, 53–57 (2009).
47. Handy, A. D., Prasad, A. & Olson, T. M. Investigating genetic variation of adrenergic receptors in familial stress cardiomyopathy (apical ballooning syndrome). *Journal of cardiology* **54**, 516–517 (2009).
48. Spinelli, L. *et al.* L41Q polymorphism of the G protein coupled receptor kinase 5 is associated with left ventricular apical ballooning syndrome. *European journal of heart failure* **12**, 13–16 (2010).
49. Vriza, O. *et al.* Analysis of beta1 and beta2-adrenergic receptors polymorphism in patients with apical ballooning cardiomyopathy. *Acta Cardiol* **66**, 787–790 (2011).
50. Novo, G. *et al.* G-protein-coupled receptor kinase 5 polymorphism and Takotsubo cardiomyopathy. *J Cardiovasc Med (Hagerstown)* **16**, 639–643 (2014).
51. Munafò, M. R. & Flint, J. Meta-analysis of genetic association studies. *Trends Genet* **20**, 439–444 (2004).
52. Pizzino, G. *et al.* Takotsubo syndrome and estrogen receptor genes: partners in crime? *J Cardiovasc Med (Hagerstown)* **18**, 268–276 (2017).
53. Ng, S. B. *et al.* Targeted capture and massively parallel sequencing of 12 human exomes. *Nature* **461**, 272–276 (2009).
54. Bamshad, M. J. *et al.* Exome sequencing as a tool for Mendelian disease gene discovery. *Nature reviews. Genetics* **12**, 745–755 (2011).
55. Yang, Y. *et al.* Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *N Engl J Med* **369**, 1502–1511 (2013).
56. Chugh, S. S. & Huertas-Vazquez, A. Inherited arrhythmia syndromes: exome sequencing opens a new door to diagnosis. *J Am Coll Cardiol* **63**, 267–268 (2014).
57. Kircher, M. *et al.* A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* **46**, 310–315 (2014).
58. Cingolani, P. *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)* **6**, 80–92 (2012).
59. Schwarz, J. M., Cooper, D. N., Schuelke, M. & Seelow, D. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat Methods* **11**, 361–362 (2014).
60. Voight, B. F. *et al.* The metabochip, a custom genotyping array for genetic studies of metabolic, cardiovascular, and anthropometric traits. *PLoS Genet* **8**, e1002793, <https://doi.org/10.1371/journal.pgen.1002793> (2012).
61. Kearney, H. M. *et al.* American College of Medical Genetics standards and guidelines for interpretation and reporting of postnatal constitutional copy number variants. *Genet Med* **13**, 680–685 (2011).
62. Li, Y. L., Sanchez-Pulido, L., Haerty, W. & Ponting, C. P. RBFOX and PTBP1 proteins regulate the alternative splicing of micro-exons in human brain transcripts. *Genome Res* **25**, 1–13 (2015).
63. Irimia, M. *et al.* A highly conserved program of neuronal microexons is misregulated in autistic brains. *Cell* **159**, 1511–1523 (2014).
64. Zhao, W. W. Intragenic deletion of RBFOX1 associated with neurodevelopmental/neuropsychiatric disorders and possibly other clinical presentations. *Molecular cytogenetics* **6**, 26, <https://doi.org/10.1186/1755-8166-6-26> (2013).
65. Lale, S., Yu, S. & Ahmed, A. Complex congenital heart defects in association with maternal diabetes and partial deletion of the A2BP1 gene. *Fetal Pediatr Pathol* **30**, 161–166 (2011).
66. Lal, D. *et al.* Rare exonic deletions of the RBFOX1 gene increase risk of idiopathic generalized epilepsy. *Epilepsia* **54**, 265–271 (2013).
67. Gao, C. *et al.* RBFOX1-mediated RNA splicing regulates cardiac hypertrophy and heart failure. *J Clin Invest* **126**, 195–206 (2016).
68. Earle, N. J. *et al.* Genetic markers of repolarization and arrhythmic events after acute coronary syndromes. *Am Heart J* **169**, 579–586 e573, <https://doi.org/10.1016/j.ahj.2014.11.015> (2015).
69. Arking, D. E. *et al.* Genome-wide association study identifies GPC5 as a novel genetic locus protective against sudden cardiac arrest. *PLoS One* **5**, e9879, <https://doi.org/10.1371/journal.pone.0009879> (2010).
70. Rosenberg, R. D., Shworak, N. W., Liu, J., Schwartz, J. J. & Zhang, L. Heparan sulfate proteoglycans of the cardiovascular system. Specific structures emerge but how is synthesis regulated? *J Clin Invest* **100**, S67–75 (1997).
71. Guntjeski-Hamblin, A. M. *et al.* Annexin VI overexpression targeted to heart alters cardiomyocyte function in transgenic mice. *Am J Physiol* **270**, H1091–1100 (1996).
72. Song, G. *et al.* Altered mechanical properties and intracellular calcium signaling in cardiomyocytes from annexin 6 null-mutant mice. *FASEB J* **16**, 622–624 (2002).
73. Camors, E., Monceau, V. & Charlemagne, D. Annexins and Ca²⁺ handling in the heart. *Cardiovasc Res* **65**, 793–802 (2005).
74. Meadows, S. M. & Cleaver, O. Annexin A3 Regulates Early Blood Vessel Formation. *PLoS One* **10**, e0132580, <https://doi.org/10.1371/journal.pone.0132580> (2015).
75. Adhikari, N. *et al.* Identification of a new target of miR-16, Vacuolar Protein Sorting 4a. *PLoS One* **9**, e101509, <https://doi.org/10.1371/journal.pone.0101509> (2014).
76. Porrello, E. R. *et al.* MiR-15 family regulates postnatal mitotic arrest of cardiomyocytes. *Circ Res* **109**, 670–679 (2011).
77. Hullinger, T. G. *et al.* Inhibition of miR-15 protects against cardiac ischemic injury. *Circ Res* **110**, 71–81 (2012).
78. Venkateswaran, A. *et al.* The NADH oxidase ENOX1, a critical mediator of endothelial cell radiosensitization, is crucial for vascular development. *Cancer Res* **74**, 38–43 (2014).

79. Weng, L. *et al.* Isolation and characterization of chondrolectin (Chodl), a novel C-type lectin predominantly expressed in muscle cells. *Gene* **308**, 21–29 (2003).
80. Coppinger, J. *et al.* Identification of familial and de novo microduplications of 22q11.21–q11.23 distal to the 22q11.21 microdeletion syndrome region. *Hum Mol Genet* **18**, 1377–1383 (2009).
81. Wentzel, C., Fernstrom, M., Ohrner, Y., Anneren, G. & Thuresson, A. C. Clinical variability of the 22q11.2 duplication syndrome. *Eur J Med Genet* **51**, 501–510 (2008).
82. Coe, B. P. *et al.* Refining analyses of copy number variation identifies specific genes associated with developmental delay. *Nat Genet* **46**, 1063–1071 (2014).
83. Pipes, L. *et al.* The non-human primate reference transcriptome resource (NHPRT) for comparative functional genomics. *Nucleic Acids Res* **41**, D906–914, <https://doi.org/10.1093/nar/gks1268> (2013).
84. Vadas, O. *et al.* Molecular determinants of PI3Kgamma-mediated activation downstream of G-protein-coupled receptors (GPCRs). *Proc Natl Acad Sci USA* **110**, 18862–18867 (2013).
85. Izumi, Y. Drug-induced takotsubo cardiomyopathy. *Heart Fail Clin* **9**, 225–231 (2013).
86. Ueyama, T. Emotional stress-induced Tako-tsubo cardiomyopathy: animal model and molecular mechanism. *Ann N Y Acad Sci* **1018**, 437–444 (2004).
87. Kido, K. & Guglin, M. Drug-Induced Takotsubo Cardiomyopathy. *Journal of cardiovascular pharmacology and therapeutics*, 1074248417708618, <https://doi.org/10.1177/1074248417708618> (2017).
88. Schreyer, S., Ledwig, D., Rakatzi, I., Kloting, I. & Eckel, J. Insulin receptor substrate-4 is expressed in muscle tissue without acting as a substrate for the insulin receptor. *Endocrinology* **144**, 1211–1218 (2003).
89. Fantin, V. R. *et al.* Cloning, tissue expression, and chromosomal location of the mouse insulin receptor substrate 4 gene. *Endocrinology* **140**, 1329–1337 (1999).
90. Neill, N. J., Torchia, B. S., Bejjani, B. A., Shaffer, L. G. & Ballif, B. C. Comparative analysis of copy number detection by whole-genome BAC and oligonucleotide array CGH. *Molecular cytogenetics* **3**, 11 (2010).
91. Mertens, F. *et al.* The t(X;6) in subungual exostosis results in transcriptional deregulation of the gene for insulin receptor substrate 4. *Int J Cancer* **128**, 487–491 (2011).
92. Duenez-Guzman, E. A. & Haig, D. The evolution of reproduction-related NLRP genes. *J Mol Evol* **78**, 194–201 (2014).
93. Snoep, J. D. *et al.* The minor allele of GP6 T13254C is associated with decreased platelet activation and a reduced risk of recurrent cardiovascular events and mortality: results from the SMILE-Platelets project. *J Thromb Haemost* **8**, 2377–2384 (2010).
94. Bray, P. F., Howard, T. D., Vittinghoff, E., Sane, D. C. & Herrington, D. M. Effect of genetic variations in platelet glycoproteins Ibalpha and VI on the risk for coronary heart disease events in postmenopausal women taking hormone therapy. *Blood* **109**, 1862–1869 (2007).
95. Lorenz, K., Schmitt, J. P., Schmitteckert, E. M. & Lohse, M. J. A new type of ERK1/2 autophosphorylation causes cardiac hypertrophy. *Nat Med* **15**, 75–83 (2009).
96. Dimassi, S. *et al.* A subset of genomic alterations detected in rolandic epilepsies contains candidate or known epilepsy genes including GRIN2A and PRRT2. *Epilepsia* **55**, 370–378 (2014).
97. Zhao, Y. Y. *et al.* Neuregulins promote survival and growth of cardiac myocytes. Persistence of ErbB2 and ErbB4 expression in neonatal and adult ventricular myocytes. *J Biol Chem* **273**, 10261–10269 (1998).
98. Berg, R. W. *et al.* Cloning and characterization of a novel beta integrin-related cDNA coding for the protein TIED (“ten beta integrin EGF-like repeat domains”) that maps to chromosome band 13q33: A divergent stand-alone integrin stalk structure. *Genomics* **56**, 169–178 (1999).
99. Wheeler, H. E. *et al.* Survey of the Heritability and Sparse Architecture of Gene Expression Traits across Human Tissues. *PLoS Genet* **12**, e1006423, <https://doi.org/10.1371/journal.pgen.1006423> (2016).
100. Eitel, I. *et al.* Genome-wide association study in takotsubo syndrome - Preliminary results and future directions. *Int J Cardiol* **236**, 335–339 (2017).
101. Redon, R. *et al.* Global variation in copy number in the human genome. *Nature* **444**, 444–454 (2006).
102. McCarroll, S. A. & Altshuler, D. M. Copy-number variation and association studies of human disease. *Nat Genet* **39**, S37–42 (2007).
103. Stranger, B. E. *et al.* Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science* **315**, 848–853 (2007).
104. Purcell, S. M. *et al.* A polygenic burden of rare disruptive mutations in schizophrenia. *Nature* **506**, 185–190 (2014).
105. Helbig, I. *et al.* Structural genomic variation in childhood epilepsies with complex phenotypes. *Eur J Hum Genet* **22**, 896–901 (2014).
106. Hooli, B. V. *et al.* Rare autosomal copy number variations in early-onset familial Alzheimer’s disease. *Mol Psychiatry* **19**, 676–681 (2014).
107. Gilissen, C. *et al.* Genome sequencing identifies major causes of severe intellectual disability. *Nature* **511**, 344–347 (2014).
108. Mefford, H. C. *et al.* Rare copy number variants are an important cause of epileptic encephalopathies. *Ann Neurol* **70**, 974–985 (2011).
109. Hitz, M. P. *et al.* Rare copy number variants contribute to congenital left-sided heart disease. *PLoS Genet* **8**, e1002903, <https://doi.org/10.1371/journal.pgen.1002903> (2012).
110. White, P. S. *et al.* Analysis of chromosomal structural variation in patients with congenital left-sided cardiac lesions. *Birth Defects Res A Clin Mol Teratol* **100**, 951–64 (2014).
111. Fakhro, K. A. *et al.* Rare copy number variations in congenital heart disease patients identify unique genes in left-right patterning. *Proc Natl Acad Sci USA* **108**, 2915–2920 (2011).
112. Glessner, J. T. *et al.* Increased frequency of de novo copy number variants in congenital heart disease by integrative analysis of single nucleotide polymorphism array and exome sequence data. *Circ Res* **115**, 884–896 (2014).
113. Barc, J. *et al.* Screening for copy number variation in genes associated with the long QT syndrome: clinical relevance. *J Am Coll Cardiol* **57**, 40–47 (2011).
114. Silversides, C. K. *et al.* Rare copy number variations in adults with tetralogy of Fallot implicate novel risk gene pathways. *PLoS Genet* **8**, e1002843, <https://doi.org/10.1371/journal.pgen.1002843> (2012).
115. Roberts, J. L., Hovanes, K., Dasouki, M., Manzardo, A. M. & Butler, M. G. Chromosomal microarray analysis of consecutive individuals with autism spectrum disorders or learning disability presenting for genetic services. *Gene* **535**, 70–78 (2014).
116. Battaglia, A. *et al.* Confirmation of chromosomal microarray as a first-tier clinical diagnostic test for individuals with developmental delay, intellectual disability, autism spectrum disorders and dysmorphic features. *Eur J Paediatr Neurol* **17**, 589–599 (2013).
117. Chong, W. W. *et al.* Performance of chromosomal microarray for patients with intellectual disabilities/developmental delay, autism, and multiple congenital anomalies in a Chinese cohort. *Molecular cytogenetics* **7**, 34 (2014).
118. Mc Cormack, A. *et al.* Microarray testing in clinical diagnosis: an analysis of 5,300 New Zealand patients. *Molecular cytogenetics* **9**, 29 (2016).
119. Ruusala, A. & Aspenstrom, P. Isolation and characterisation of DOCK8, a member of the DOCK180-related regulators of cell morphology. *FEBS Lett* **572**, 159–166 (2004).
120. Engelhardt, K. R. *et al.* Large deletions and point mutations involving the dedicator of cytokinesis 8 (DOCK8) in the autosomal-recessive form of hyper-IgE syndrome. *J Allergy Clin Immunol* **124**, 1289–1302 e1284 (2009).

121. Al Mutairi, M. *et al.* Grave aortic aneurysmal dilatation in DOCK8 deficiency. *Mod Rheumatol* **24**, 690–693 (2014).
122. Sternecker, J. L., Reinhardt, P. & Scholer, H. R. Investigating human disease using stem cell models. *Nat Rev Genet* **15**, 625–639 (2014).
123. Ma, D. *et al.* Generation of patient-specific induced pluripotent stem cell-derived cardiomyocytes as a cellular model of arrhythmogenic right ventricular cardiomyopathy. *Eur Heart J* **34**, 1122–1133 (2013).
124. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
125. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research* **20**, 1297–1303 (2010).
126. Goecks, J., Nekrutenko, A., Taylor, J. & Galaxy, T. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol* **11**, R86, <https://doi.org/10.1186/gb-2010-11-8-r86> (2010).
127. Durbin, R. M. *et al.* A map of human genome variation from population-scale sequencing. *Nature* **467**, 1061–1073 (2010).
128. Landrum, M. J. *et al.* ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res* **42**, D980–985 (2014).
129. Ellis, K. L. *et al.* Genomic risk variants at 1p13.3, 1q41, and 3q22.3 are associated with subsequent cardiovascular outcomes in healthy controls and in established coronary artery disease. *Circ Cardiovasc Genet* **4**, 636–646 (2011).
130. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559–575 (2007).
131. R Development Core Team, R: A Language and Environment for Statistical Computing. Vienna, Austria: the R Foundation for Statistical Computing. ISBN: 3-900051-07-0 (2011).
132. Benjamini, Y. & Yekutieli, D. The control of the false discovery rate in multiple testing under dependency. *Ann Stat* **29**, 1165–1188 (2001).
133. Firth, H. V. *et al.* DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. *Am J Hum Genet* **84**, 524–533 (2009).
134. MacDonald, J. R., Ziman, R., Yuen, R. K., Feuk, L. & Scherer, S. W. The Database of Genomic Variants: a curated collection of structural variation in the human genome. *Nucleic Acids Res* **42**, D986–992 (2014).
135. Riggs, E. R. *et al.* Towards an evidence-based process for the clinical interpretation of copy number variation. *Clin Genet* **81**, 403–412 (2012).
136. Hanemaaijer, N. M. *et al.* Practical guidelines for interpreting copy number gains detected by high-resolution array in routine diagnostics. *Eur J Hum Genet* **20**, 161–165 (2012).
137. Breckpot, J. *et al.* Challenges of interpreting copy number variation in syndromic and non-syndromic congenital heart defects. *Cytogenet Genome Res* **135**, 251–259 (2011).
138. Mills, R. E. *et al.* Mapping copy number variation by population-scale genome sequencing. *Nature* **470**, 59–65 (2011).
139. Cooper, G. M. *et al.* A copy number variation morbidity map of developmental delay. *Nat Genet* **43**, 838–846 (2011).
140. Shaikh, T. H. *et al.* High-resolution mapping and analysis of copy number variations in the human genome: a data resource for clinical and research applications. *Genome Res* **19**, 1682–1690 (2009).
141. Vogler, C. *et al.* Microarray-based maps of copy-number variant regions in European and sub-Saharan populations. *Plos One* **5**, e15246, 16, <https://doi.org/10.1371/journal.pone.0015246> (2010).

Acknowledgements

This study was supported by an Earthquake research grant jointly awarded by the Health Research Council of New Zealand and Canterbury Medical Research Foundation (12/737). This study makes use of data generated by the DECIPHER community. A full list of centres who contributed to the generation of the data is available from <http://decipher.sanger.ac.uk> and via email from decipher@sanger.ac.uk. Funding for the DECIPHER project was provided by the Wellcome Trust. We thank all participants for their agreement to be involved in this study, Alan Aitchison and Allison Miller for preparation of DNA samples and Sanger sequencing analyses, and Whitney Whitford (University of Auckland) for generating the exome read count tables.

Author Contributions

C.J.L. and M.A.K. led the project and drafted the paper. K.D. and V.M.B. carried out the array C.G.H. analysis and data analysis, and contributed to writing of the manuscript. J.J.Z. and B.K. recruited and interviewed patients, and collected samples. P.G.B. provided clinical cardiology expertise, and P.M.G. provided molecular pathology expertise. M.A.K., M.J.C., M.A.B., K.L., T.R.M. developed the bioinformatic pipelines and carried out exome analysis. J.F.P. and V.A.C. carried out Cardio-MetaboChip analysis. All authors contributed to editing of the manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-25827-5>.

Competing Interests: The authors declare no competing interests.

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018