JACC: CARDIOONCOLOGY VOL. 2, NO. 5, 2020

ORIGINAL RESEARCH

Machine Learning Identifies Clinical and Genetic Factors Associated With Anthracycline Cardiotoxicity in Pediatric Cancer Survivors

Marie-A Chaix, MD, MSc,[a,b] Neha Parmar, BSc,[a] Caroline Kinnear, MSc,[a] Myriam Lafreniere-Roula, PhD,[c] Oyediran Akinrinade, PhD,[a] Roderick Yao, BCS,[a] Anastasia Miron, BAH,[a] Emily Lam, MSc,[a] Guoliang Meng, PhD,[d] Anne Christie, MSc,[a] Ashok Kumar Manickaraj, PhD,[1,a] Stacey Marjerrison, MD,[e] Rejane Dillenburg, MD, MSc,[e] Mylène Bassal, MDCM,[f] Jane Lougheed, MD,[f] Shayna Zelcer, MD,[g] Herschel Rosenberg, MD,[g] David Hodgson, MD,[h] Leonard Sender, MD,[i] Paul Kantor, MD,[j] Cedric Manlhiot, PhD,[k] James Ellis, PhD,[d,l] Luc Mertens, MD,[m] Paul C. Nathan, MD,[m] Seema Mital, MD[a,m]

ABSTRACT

BACKGROUND Despite known clinical risk factors, predicting anthracycline cardiotoxicity remains challenging.

OBJECTIVES This study sought to develop a clinical and genetic risk prediction model for anthracycline cardiotoxicity in childhood cancer survivors.

METHODS We performed exome sequencing in 289 childhood cancer survivors at least 3 years from anthracycline exposure. In a nested case-control design, 183 case patients with reduced left ventricular ejection fraction despite low-dose doxorubicin (≤250 mg/m^2), and 106 control patients with preserved left ventricular ejection fraction despite doxorubicin >250 mg/m^2 were selected as extreme phenotypes. Rare/low-frequency variants were collapsed to identify genes differentially enriched for variants between case patients and control patients. The expression levels of 5 top-ranked genes were evaluated in human induced pluripotent stem cell–derived cardiomyocytes, and variant enrichment was confirmed in a replication cohort. Using random forest, a risk prediction model that included genetic and clinical predictors was developed.

RESULTS Thirty-one genes were differentially enriched for variants between case patients and control patients ($p < 0.001$). Only 42.6% case patients harbored a variant in these genes compared to 89.6% control patients (odds ratio: 0.09; 95% confidence interval: 0.04 to 0.17; $p = 3.98 \times 10^{-15}$). A risk prediction model for cardiotoxicity that included clinical and genetic factors had a higher prediction accuracy and lower misclassification rate compared to the clinical-only model. In vitro inhibition of gene-associated pathways (*PI3KR2*, *ZNF827*) provided protection from cardiotoxicity in cardiomyocytes.

CONCLUSIONS Our study identified variants in cardiac injury pathway genes that protect against cardiotoxicity and informed the development of a prediction model for delayed anthracycline cardiotoxicity, and it also provided new targets in autophagy genes for the development of cardio-protective drugs. (Preventing Cardiac Sequelae in Pediatric Cancer Survivors [PCS2]; NCT01805778) (J Am Coll Cardiol CardioOnc 2020;2:690–706) © 2020 The Authors. Published by Elsevier on behalf of the American College of Cardiology Foundation. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

ISSN 2666-0873 https://doi.org/10.1016/j.jaccao.2020.11.004

Anthracycline chemotherapy, used in almost 50% of pediatric cancers, is a major cause of cardiac morbidity in cancer survivors. Anthracycline cardiotoxicity can manifest as left ventricular (LV) dysfunction and heart failure (1). Sixty percent of childhood cancer survivors develop echocardiographic cardiac dysfunction, and 10% develop symptomatic cardiomyopathy up to decades after chemotherapy (delayed toxicity) (2). Clinical factors associated with cardiotoxicity include younger age at exposure, cumulative anthracycline dose, radiation therapy involving the heart, female sex, African ethnicity, and pre-existing cardiac dysfunction (3,4). However, clinical factors have limited ability to predict patients who are at risk for cardiotoxicity (5).

Candidate single-nucleotide variant (SNV) and genome-wide association studies (6–9) have identified an association with cardiotoxicity of common exonic and intronic variants in genes involved in anthracycline transport and metabolism, oxidative stress, DNA repair, iron metabolism, nicotinamide adenine dinucleotide phosphate complex, topoisomerase 2β expression, and sarcomeric genes (10). These together explain only a small proportion of cases and do not consider the contribution of rare variants that can have larger effects on gene function (2,11). The objective of our study was to identify the contribution of rare and low-frequency SNVs that influence the susceptibility to late anthracycline cardiotoxicity, to validate the functional role of the affected genes, and to generate an integrated risk prediction model for anthracycline cardiotoxicity that combines clinical and genetic factors (**Central Illustration**).

METHODS

STUDY COHORT. The PCS2 (Preventing Cardiac Sequelae in Pediatric Cancer Survivors) study is a multicenter, prospective, longitudinal cohort study involving 6 North American sites (12). Patients followed up in survivor clinics, age <18 years at the time of anthracycline exposure, and at least 3 years from their last anthracycline dose were eligible. Patients with congenital heart disease were excluded. Clinical data were collected from medical records. The anthracycline cumulative dose was measured in doxorubicin (DOX) equivalents (mg/m^2), that is, daunorubicin total dose multiplied by 0.5, epirubicin total dose multiplied by 0.6, idarubicin total dose multiplied by 5, and mitoxantrone total dose multiplied by 4 (13). LV ejection fraction (LVEF) was measured by 2-dimensional echocardiography at enrollment and at 1 and 2 years of follow-up and analyzed at an independent core laboratory (12). All participants provided a blood sample at enrollment. Written informed consent was obtained from participants/parents/legal guardians, and the protocol was approved by institutional Research Ethics Boards. Extreme phenotypes were selected in a nested case-control design from the overall cohort. Cases included patients who received low cumulative anthracycline dose (≤250 mg/m^2) but developed either: 1) clinically defined cardiotoxicity, that is, LVEF of ≤50% or >10% LVEF decline to ≤55% from a previous echocardiogram during follow-up; or 2) low LVEF of ≤55% based on American Society of Echocardiography guidelines (14,15). Control patients were patients with preserved cardiac function (LVEF of >55%) despite high dose anthracycline (>250 mg/m^2).

WHOLE-EXOME SEQUENCING. Whole-exome sequencing was performed (average: 100× depth) by using the Illumina (San Diego, California) HiSeq X platform. High quality paired-end reads (2 × 150 bp) were

ABBREVIATIONS AND ACRONYMS

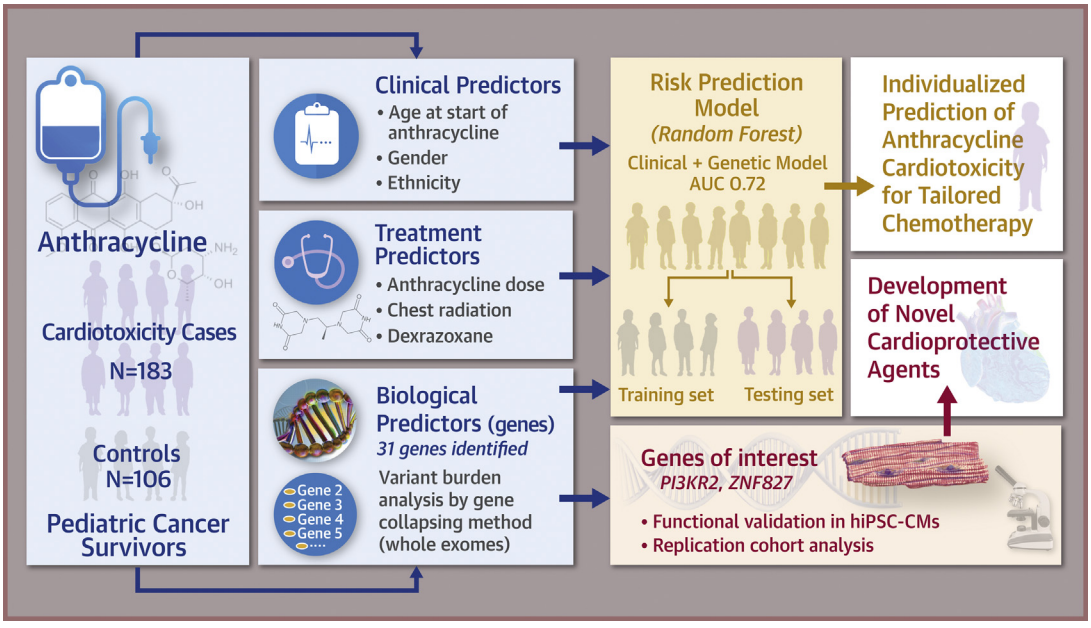
AUC = area under the curve
CI = confidence interval
DMSO = dimethyl sulfoxide
DOX = doxorubicin
GSEA = gene set enrichment analysis
H2AX = H2A family member X
hiPSC-CM = human induced pluripotent stem cell-derived cardiomyocyte
IC_{50} = half-maximal inhibitory concentration
LV = left ventricular
LVEF = left ventricular ejection fraction
MAF = minor allele frequency
mRNA = messenger RNA
OR = odds ratio
PGP = Personal Genome Project
RF = random forest
SKAT = sequence kernel association test
SNV = single-nucleotide variant

From [a]Genetics and Genome Biology, The Hospital for Sick Children, Toronto, Ontario, Canada; [b]Adult Congenital Centre, Montréal Heart Institute, Université de Montréal, Montréal, Canada; [c]Ted Rogers Computational Medicine Program, University Health Network, Toronto, Ontario, Canada; [d]Developmental and Stem Cell Biology, The Hospital for Sick Children, Toronto, Ontario, Canada; [e]Department of Pediatrics, McMaster University Children's Hospital, Hamilton, Ontario, Canada; [f]Department of Pediatrics, Children's Hospital of Eastern Ontario, University of Ottawa, Ottawa, Ontario, Canada; [g]Department of Pediatrics, Children's Hospital, London Health Sciences Centre, London, Ontario, Canada; [h]Radiation Medicine Program, Princess Margaret Cancer Centre, Department of Radiation Oncology, University of Toronto, Toronto, Ontario, Canada; [i]Department of Pediatrics, Children's Hospital of Orange County, Orange, California, USA; [j]Department of Pediatrics, Children's Hospital of Los Angeles, Los Angeles, California, USA; [k]Department of Pediatrics, Johns Hopkins Medical Center, Baltimore, Maryland, USA; [l]Department of Molecular Genetics, University of Toronto, Ontario, Canada; and the [m]Department of Pediatrics, The Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada.

The authors attest they are in compliance with human studies committees and animal welfare regulations of the authors' institutions and Food and Drug Administration guidelines, including patient consent where appropriate. For more information, visit the Author Center.

Manuscript received March 11, 2020; revised manuscript received November 3, 2020, accepted November 3, 2020.

CENTRAL ILLUSTRATION **An Integrated Precision Approach to Predict Anthracycline Cardiotoxicity**



Chaix, M.-A. et al. J Am Coll Cardiol CardioOnc. 2020;2(5):690–706.

Whole-exome sequencing data from 289 pediatric cancer survivors with extreme phenotypes identified a higher burden of rare variants in control patients compared to case patients in 31 biologically relevant genes. The top-ranked genes were functionally evaluated in human induced pluripotent stem cell-derived cardiomyocytes (hiPSC-CMs) and variant enrichment was confirmed in a replication cohort. Targeted pathway inhibitors were more effective at reducing anthracycline-induced injury than dexrazoxane. Using random forest, clinical and genetic predictors were integrated into a prediction model for anthracycline cardiotoxicity.

mapped to the human genome reference sequence (hg19, National Center for Biotechnology Information, Bethesda, Maryland) by using the bwa mem aligner, version 0.7.8 (Wellcome Trust Sanger Institute, Cambridge, United Kingdom), and variants were called using the Genome Analysis Toolkit, version 3.8.0 (16). Variants passing the default Genome Analysis Toolkit (Broad Institute, Cambridge, Massachusetts) variant caller quality metrics were annotated by using snpEff, version 4.3 (International Haplotype Map Project, National Human Genome Research Institute, Bethesda, Maryland). Quality control filters were applied to identify SNVs (17). To map ancestry, the top principal components were calculated using single-nucleotide polymorphisms common to both the HapMap (Catalogue of Somatic Mutations in Cancer, Wellcome Trust Sanger Institute, Cambridge, United Kingdom) and PCS2 cohort with minor allele frequency (MAF) of >0.01 and Hardy-Weinberg equilibrium $p > 1.00 \times 10^{-6}$. Common (MAF: >0.05) and variants associated with cancer in the COSMIC (Catalogue of Somatic Mutations in Cancer) database were excluded (18,19).

VARIANT BURDEN ANALYSIS USING GENE-COLLAPSING METHODS. All rare and low-frequency variants were collapsed within genes to generate a gene-level variant score in every sample. Three gene-based association analyses were performed: 1) burden combined multivariate and collapsing, which assumes that all functional variants in the gene have effects in the same direction (20); 2) sequence kernel association test (SKAT), a variance-component test that assumes that a variant can have either a positive or a negative effect (21); and 3) SKAT-optimized, a combination of these 2 methods (22). The association analyses were adjusted for sex, cancer diagnosis, age at the start of anthracycline, use of dexrazoxane, chest radiation, duration of follow-up from first anthracycline dose, and the first 2 principal components inferring ethnicity. Genes with an association p value of <0.001 by at least 2 gene-collapsing methods and genes in biologically relevant pathways with an association of $p < 0.001$ by at least 1 method were

included in downstream analysis and model development. Genome-wide significance was defined as a nominal $p < 2.87 \times 10^{-6}$ by using Bonferroni correction based on 17,382 genes. Results are presented as odds ratio (ORs) with 95% confidence intervals (CIs). Genes significantly associated with case/control status were further analyzed after adjustment for cumulative anthracycline dose and use of concomitant cardiotoxic drugs by using logistic regression.

PATHWAY ANALYSIS. Prioritized genes were included in a gene network analysis using GeneMania (23). Gene set enrichment analysis (GSEA) was performed by using the Java implementation of GSEA version 3.0 (Broad institute, Inc., Massachusetts Institute of Technology, Cambridge, Massachusetts) (24). The Fisher exact test p value comparing the enrichment of variants between case patients and control patients was used for gene ranking. GSEA was run using Gene Ontology pathways, version 6, with 1,000 permutations. A p value of <0.001 was considered significant. Prioritized genes were functionally evaluated in human cardiomyocytes.

FUNCTIONAL EVALUATION OF PRIORITIZED GENES. Differentiation of human induced pluripotent stem cells into cardiomyocytes. To ascertain the functional role of the prioritized genes in cardiotoxicity, human induced pluripotent stem cell-derived cardiomyocytes (hiPSC-CMs) were generated. Two previously published hiPSC cell lines (PGP17_11 and PGP14_26), reprogrammed from lymphocytes taken from 2 healthy male donors in the Personal Genome Project (PGP) were used (25). We confirmed the absence of pathogenic or likely pathogenic SNVs (based on American College of Medical Genetics criteria) in known cardiotoxicity genes and genes identified in our study in the donors through interrogation of vcf files downloaded from the PGP-Canada website (26). The STEMdiff Cardiomyocyte Differentiation Kit (STEMCELL Technologies, Vancouver, Canada) was used to differentiate hiPSCs into CMs in accordance with our previously published differentiation protocol (26). At day 16, cells were dissociated, reseeded, and maintained in STEMdiff Cardiomyocyte Maintenance medium (STEMCELL Technologies), which was replaced every 2 days, in 80% relative humidity levels and 5% CO_2 levels, to generate beating CMs.

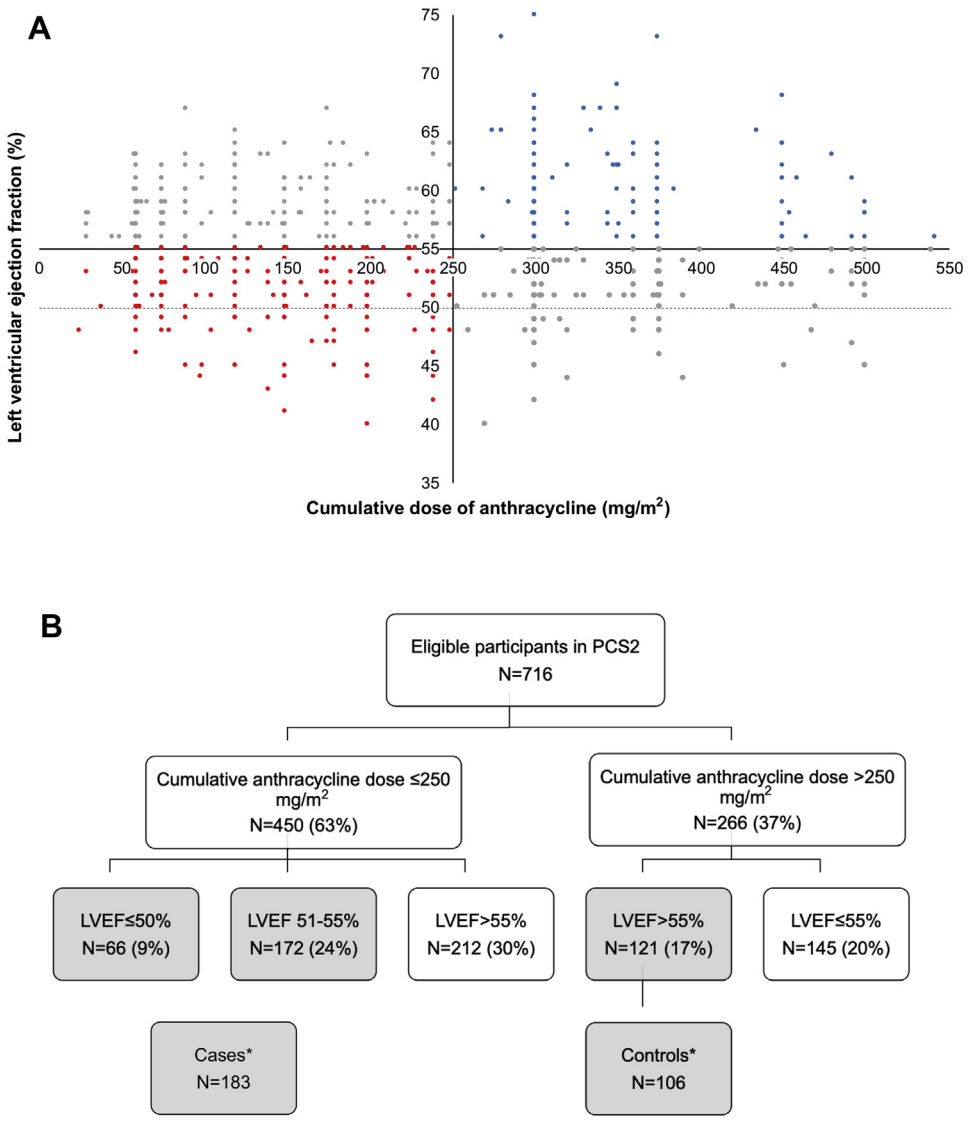
Cell contractility. hiPSC-CMs at day 16 were seeded (40,000 cells/well) in 150 μl iCell (Stony Brook, New York) maintenance medium in 48-well electronic microtiter plates (E-Plate Cardio 96, Agilent, Santa Clara, California) that contain gold microelectrode arrays fused to the bottom of each well (coated with fibronectin). Experiments were performed on day 34, when they exhibit a more mature phenotype, with regular contractility, stable electrical signals, and sarcomeric organization. CMs were treated with 0.01, 0.1, and 0.5 μmol/l of doxorubicin hydrochloride (DOX) (product no. D1515, Sigma-Aldrich, St. Louis, Missouri) on day 34 for 24 h. Cell index derived from change in impedance values, a measure of contractility and viability, was recorded and analyzed by using the xCELLigence RTCA Cardio system (Agilent, Santa Clara, California) (27).

Metabolic activity and cell viability. hiPSC-CMs at day 16 were seeded (40,000 cells/well) in 96-well plates, treated with DOX on day 34, and incubated with a 1:100 dilution of PrestoBlue Cell Viability Reagent (Life Technologies, no. A13261, Carlsbad, California) a resazurin-based assay that uses the reducing environment of cells to measure metabolic activity and cell viability, for 1.5 h at 37°C. The average fluorescence values at 560/590 nm (excitation/emission) of the empty control wells were subtracted from that of the experimental wells to provide relative fluorescence units.

Immunofluorescence for DNA damage markers. hiPSC-CMs at day 16 were seeded in black 24-well glass bottom plates for confocal microscopy and treated with DOX for 24 h on day 34. After fixation with 4% paraformaldehyde, the cells were incubated with 1:1,000 monoclonal mouse γ-H2A family member X (H2AX) (phospho S139) (Abcam, ab26350, Cambridge, United Kingdom) for 24 h to measure double-stranded DNA breaks and incubated with 1:100 Alexa Fluor 488 conjugated secondary antibody (Thermo Fisher Scientific, A21042, Waltham, Massachusetts). Images were taken on WaveFX-X1 spinning disk confocal system (Quorum Technologies Inc., Guelph, Canada) by using a Hamamatsu (Hamamatsu City, Japan) C9100-13 electron multiplying charge coupled device camera. Volocity software (Perkin Elmer, Waltham, Massachusetts) was used for the quantification of nuclear γ-H2AX foci.

Messenger RNA expression in hiPSC-CMs. Quantitative reverse-transcription polymerase chain reaction (RT-qPCR) was performed on hiPSC-CMs to measure change in messenger RNA (mRNA) expression of selected genes following 24 h of DOX exposure. Total RNA was extracted from the hiPSC-CMs by using the TRIzol reagent phenol-chloroform extraction protocol (Thermo Fisher Scientific, Waltham, Massachusetts). Complementary DNA was synthesized from the pooled RNA (200 ng RNA per sample) by using SuperScript III Reverse Transcriptase (Life

FIGURE 1 Study Cohort and CONSORT Diagram



(A) Study cohort: Scatterplot showing the distribution of left ventricular ejection fraction (LVEF) against anthracycline cumulative dose, with control patients in **blue** (n = 121), case patients in **red** (n = 238), and the rest in **gray** (n = 357). **(B)** Consolidated Standards of Reporting Trials (CONSORT) diagram: selection of case and control patients for exome sequencing in the discovery cohort. *Participants with complete data and good-quality DNA. PCS2 = Preventing Cardiac Sequelae in Pediatric Cancer Survivors.

Technologies, Carlsbad, California). RT-qPCR was performed by using different primer pairs (Supplemental Table 1). mRNA expression levels were normalized to the housekeeping gene, glyceraldehyde 3-phosphate dehydrogenase, and expressed relative to the reference group. SYBR GreenER qPCR SuperMix (Thermo Fisher Scientific, Waltham, Massachusetts) was used for transcription and amplification (ViiA7 qPCR system, Applied Biosystems, Waltham, Massachusetts) by using a total volume of 12 μl and 40 2-step cycles. Relative mRNA expression was quantified by using the $2^{-(\Delta\Delta CT)}$ method that calculates the differences in fluorescence levels at the threshold cycle between the target and reference genes (28). Genes that showed up-regulation following DOX exposure were prioritized for drug testing. Experiments were performed with 3 independent biological replicates, each containing 3 technical replicates.

Drug testing. To determine the effect of target gene inhibition on DOX-induced cardiotoxicity, CMs were

treated with 3 doses of DOX (0.01 μmol/l, 0.1 μmol/l, and 0.5 μmol/l) with or without 24 h of pre-treatment with 0.1 μmol/l of dimethyl sulfoxide (DMSO) or inhibitors of the protein products. One μmol/l DOX was associated with cell death and was not used (data not shown). The DOX doses were consistent with previously published studies (29,30). The inhibitors included: 1) dexrazoxane, an iron chelator (Sigma-Aldrich, St. Louis, Missouri, product no. D1446); 2) TGX-221, a PI3KR2 inhibitor (Selleckchem, Burlington, Ontario, Canada, product no. S1169); 3) rapamycin, an inhibitor of mammalian target of rapamycin, a protein kinase complex downstream of phosphoinositide 3-kinase, and (iv) metformin hydrochloride, an inhibitor of TR4, a nuclear receptor involved in zinc-finger protein, ZNF827, recruitment (Sigma-Aldrich, St. Louis, MO, product no. PHR1084) (31,32). Three inhibitor concentrations (0.01 μmol/l, 0.1 μmol/l, and 1 μmol/l) were studied, and the highest nontoxic dose was selected similar to published studies (33). All drugs were prepared in STEMdiff Cardiomyocyte Maintenance serum-free medium. The half-maximal inhibitory concentration (IC_{50}) values, that is, the drug dose required to inhibit a biological process by 50%, were obtained through GraphPad prism 8 (GraphPad Software Inc., San Diego, California) by using a nonlinear regression analysis function. The goodness of fit was determined by using coefficient of determination, R^2.

REPLICATION COHORT. The replication cohort comprised a subset of patients with recent enrollment in the PCS2 study. Patients meeting the clinical definition of cardiotoxicity, that is, LVEF of ≤50% or >10% LVEF drop to ≤55% from a previous echocardiogram independent of anthracycline dose, were defined as case patients (15). Control patients were defined by LVEF >55%. Case and control patients were propensity-matched by using a logistic regression model that included known clinical factors (cumulative anthracycline dose, sex, cancer type, age at first anthracycline dose, follow-up duration, use of dexrazoxane, and chest radiation). The propensity score was used to create a 1:1 propensity-matched set by using a greedy algorithm (34). The prioritized genes were sequenced by using targeted exome sequencing in the replication cohort. Exome-enriched libraries were prepared with the Agilent (Santa Clara, California) Sureselect XT Custom Kit (1 to 499 kilo base pairs, 16 reactions/package) and sequenced on an Illumina HiSeq4000 (2 × 100-base pair paired-end reads). After collapsing variants within genes, gene association with cardiotoxicity was evaluated using burden analysis (chi-square test) and a nonparametric test accounting for matched pairs (McNemar test) to compare the proportion of discordant pairs between case versus control patients (i.e., pairs where a case patient had a variant and the matched control patient did not, or vice versa). All analyses were performed by using SAS, version 9.4 (SAS Institute, Inc., Cary, North Carolina).

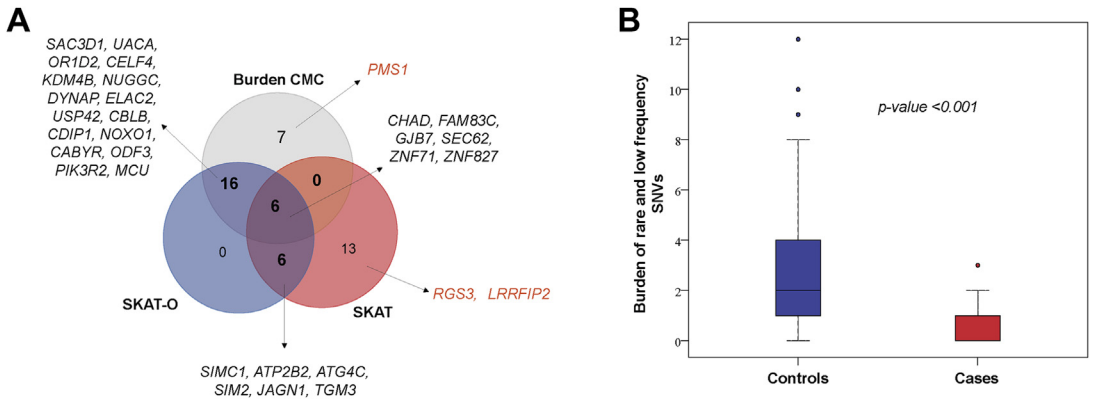
TABLE 1 Clinical Characteristics of the Study Cohort (N = 289)

	Control Patients (n = 106)	Case Patients (n = 183)	p Value
Female	57 (53.8)	92 (50.3)	0.566
Age at start of anthracycline, yrs	6.0 (2.0–10.0)	4.0 (2.0–7.0)	0.018
Cumulative anthracycline dose in doxorubicin equivalent dose, mg/m^2	371 ± 115	128 ± 59	<0.001
Use of dexrazoxane	13 (12.6)	2 (1.1)	<0.001
Radiation therapy involving the heart	44 (41.5)	64 (35.0)	0.268
Cancer diagnosis			<0.001
Leukemia (AML, ALL)	40 (37.7)	94 (51.4)	0.049
Sarcoma (osteosarcoma, Ewing, rhabdomyosarcoma)	28 (26.4)	2 (1.1)	<0.001
Neuroblastoma, hepatoblastoma	18 (17.0)	13 (7.1)	0.009
Lymphoma (NHL, HL)	10 (9.4)	35 (19.1)	0.029
Wilms tumor	2 (1.9)	23 (12.6)	0.002
LVEF at last follow-up, %	61.3 ± 6.7	51.7 ± 2.8	<0.001
Time from first anthracycline dose to last follow-up echocardiogram, yrs	8.5 (5.0–12.3)	9.0 (6.0–12.3)	0.854
Duration of treatment, days	373.3 ± 1,080.5	182.0 ± 282.5	0.026

Values are n (%), median (interquartile range), or mean ± SD.
AML = acute myeloid leukemia; ALL = acute lymphocytic leukemia; HL = Hodgkin lymphoma; LV = left ventricular; NHL = non-Hodgkin lymphoma.

RANDOM FOREST MODEL FOR RISK PREDICTION. To develop a prediction model for anthracycline cardiotoxicity, we developed 3 machine learning algorithms using clinical factors alone (sex, age at first anthracycline dose, follow-up duration, the first 2 principal components inferring ethnicity, treatment exposures [i.e., anthracycline dose, use of dexrazoxane, and chest radiation], genetic factors alone (significantly associated genes) and a combination of clinical and genetic factors. The stratified bootstrapping sampling scheme was applied to generate 1,000 replicates by using the discovery cohort. Each replicate was a pair of independent randomly selected training and testing sets with the testing set corresponding to samples not selected in the training set. A random forest (RF) classifier, based on ntree = 500, was applied to each replicate to predict cases with cardiotoxicity based on the 3 models (clinical, genetic, and combined). RF aggregates the votes from different decision trees to determine the prediction. For each replicate, we trained the RF model in a training set and evaluated the model performance in the testing set. All accuracy measures (receiver operator characteristic, area under the curve

FIGURE 2 Genes Associated With Anthracycline Cardiotoxicity



(A) A total of 28 genes showed differential enrichment between case and control patients by at least 2 methods ($p < 0.001$), and 3 additional biologically relevant genes **(orange)** were significant by at least 1 method. **(B)** Burden of rare and low-frequency single-nucleotide variants in the 31 prioritized genes was higher in control patients compared to case patients ($p < 0.001$). **(C)** Forest plot showing the estimated odds ratios (95% confidence intervals) for the 31 top genes in case versus control patients using Fisher exact test. Genes in bold were prioritized for functional studies.

Continued on the next page

[AUC], sensitivity, specificity, positive predictive value, negative predictive value (NPV), false positive rate, false negative rate] and misclassification rates were compared over 1,000 replicates, and observed versus predicted case status was calculated with 0.5 probability cutoff for each replicate. Within each replicate, accuracy measures and misclassification rates were estimated in the training set and the testing set. The overall accuracy measure was the sum of the accuracy measures and the misclassification rates from the training set and the paired testing set, weighted by 0.368 and 0.632, respectively. This was repeated for all 3 RF models.

RESULTS

COHORT CHARACTERISTICS. A total of 716 participants enrolled in the PCS2 study met eligibility criteria. The discovery cohort for exome sequencing included 289 participants, including 183 case and 106 control patients (**Figure 1**). Clinical characteristics are shown in **Table 1**. Case patients were younger at the time of anthracycline therapy ($p = 0.012$), but follow-up duration was not different between case and control patients. As expected from the selection criteria, the LVEF and the cumulative anthracycline dose were lower in case patients compared to control patients ($p < 0.001$). The duration of treatment was correlated with cumulative anthracycline dose (Spearman rho: 0.54; $p < 0.001$). Concordant with a lower cumulative anthracycline dose, case patients had a higher frequency of Wilms tumor ($p = 0.002$) and were less likely to have received dexrazoxane ($p < 0.001$). No sex-based differences were observed between case and control patients.

VARIANTS IDENTIFIED ON EXOME SEQUENCING. Average sequencing coverage was 99× (range: 69 to 234×). A total of 9,092,459 SNVs were identified. After excluding common variants (MAF: >0.05), we identified 110,558 rare/low-frequency non-synonymous SNVs in 17,382 unique genes in the cohort. Principal component analysis was performed by using 332,033 common SNVs (MAF: >0.01 in Genome Aggregation Database) mapped to HapMap to ascertain the distribution of ethnicity in the study cohort (Supplemental Figure 1).

GENE-ASSOCIATION ANALYSIS. After collapsing variants within genes, we identified 48 unique genes associated with cardiotoxicity—29 genes by burden combined multivariate and collapsing, 25 genes by SKAT, and 29 genes by SKAT-optimized ($p < 0.001$). No single gene reached exome-wide significance (Supplemental Figures 2 and 3). Of the 48 genes, 28 showed differential enrichment by at least 2 methods ($p < 0.001$), and 3 additional genes previously implicated in cardiotoxicity were identified by at least 1 method ($p < 0.001$) (**Figures 2A and 2B**, Supplemental Tables 2 and 3). Variant burden analysis showed that

FIGURE 2 Continued

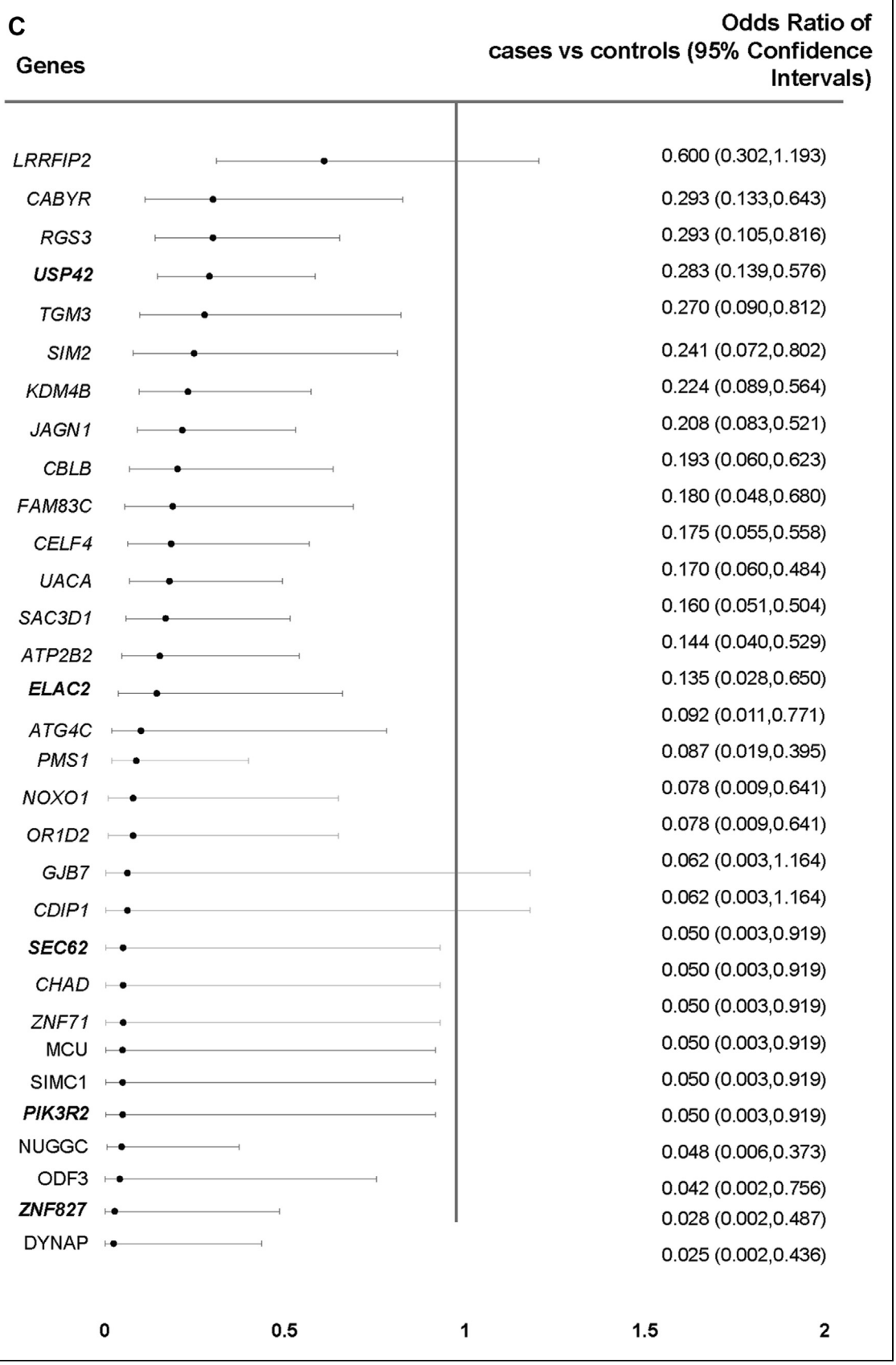
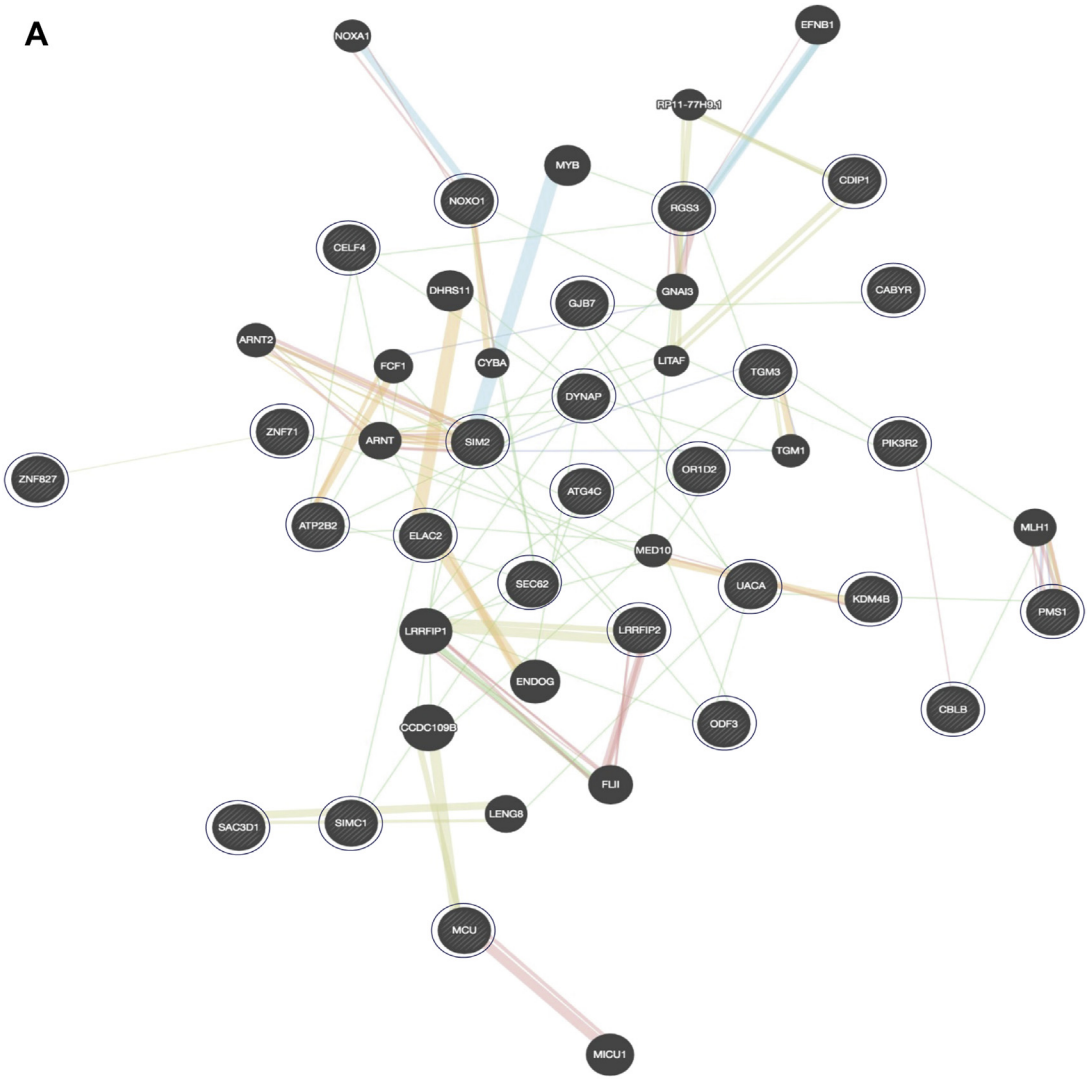


FIGURE 3 Pathways Associated With Anthracycline Cardiotoxicity

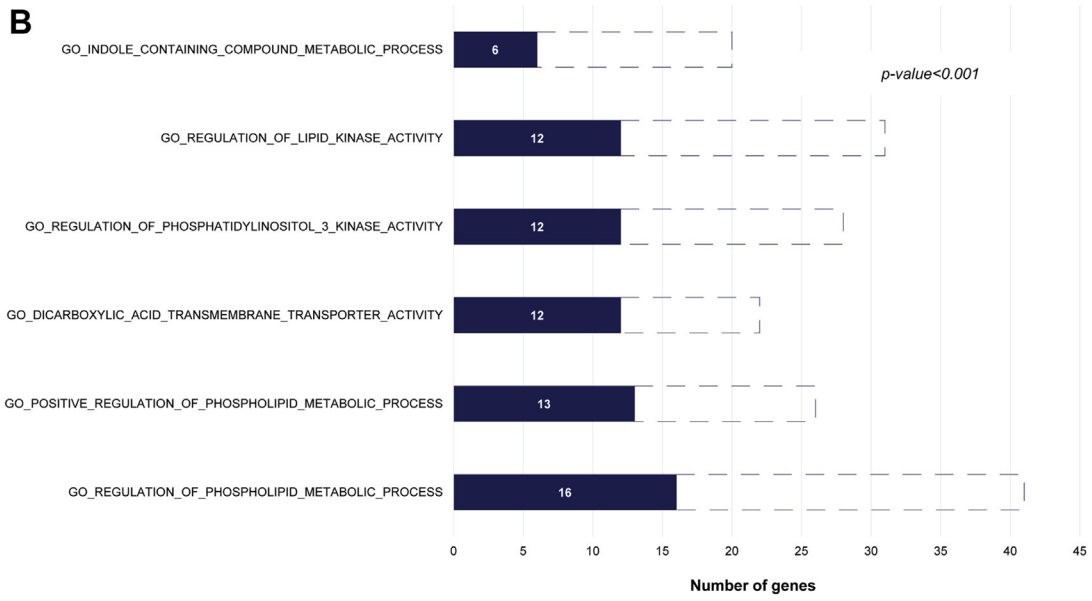


(A) GeneMania analysis identified 46 interacting genes, including 26 of the 31 top genes. Large circles represent significantly associated genes; small circles represent other interacting genes. Physical interaction **(pink lines)**, coexpression **(purple lines)**, colocalization **(blue lines)**, shared protein domains **(gray-yellow lines)**, genetic interaction **(green lines)**, and predicted **(orange lines)**. **(B)** Gene set enrichment analysis identified the top-ranked pathways to which the genes mapped ($p < 0.001$). The **solid bar** shows number of significant genes in each pathway ($p < 0.001$); the **dashed bar** represents the total genes.

Continued on the next page

fewer case patients (42.6%) compared to control patients (89.6%) harbored a variant in these 31 genes (OR: 0.09; 95% CI: 0.04 to 0.17; $p = 3.98 \times 10^{-15}$) (Figure 2C). Frequency of multiple variants was also lower in case patients (21%) compared to control patients (71%) (OR: 0.11; 95% CI: 0.06 to 0.19; $p = 4.74 \times 10^{-17}$). In a subset of 78 cases limited to LVEF of ≤50% or >10% LVEF decline to ≤55% from a previous echocardiogram, the variant burden remained lower in the case patients (42.3%) compared to the control patients (89.6%) (OR: 0.08; 95% CI: 0.04 to 0.18; $p = 5.46 \times 10^{-12}$). In subgroup analysis of 80 case patients and 63 control patients with self-reported White ethnicity, variant burden was again lower in case patients (72%) compared to control patients (89%) (OR: 0.33; 95% CI: 0.13 to 0.83; $p = 0.019$).

FIGURE 3 Continued



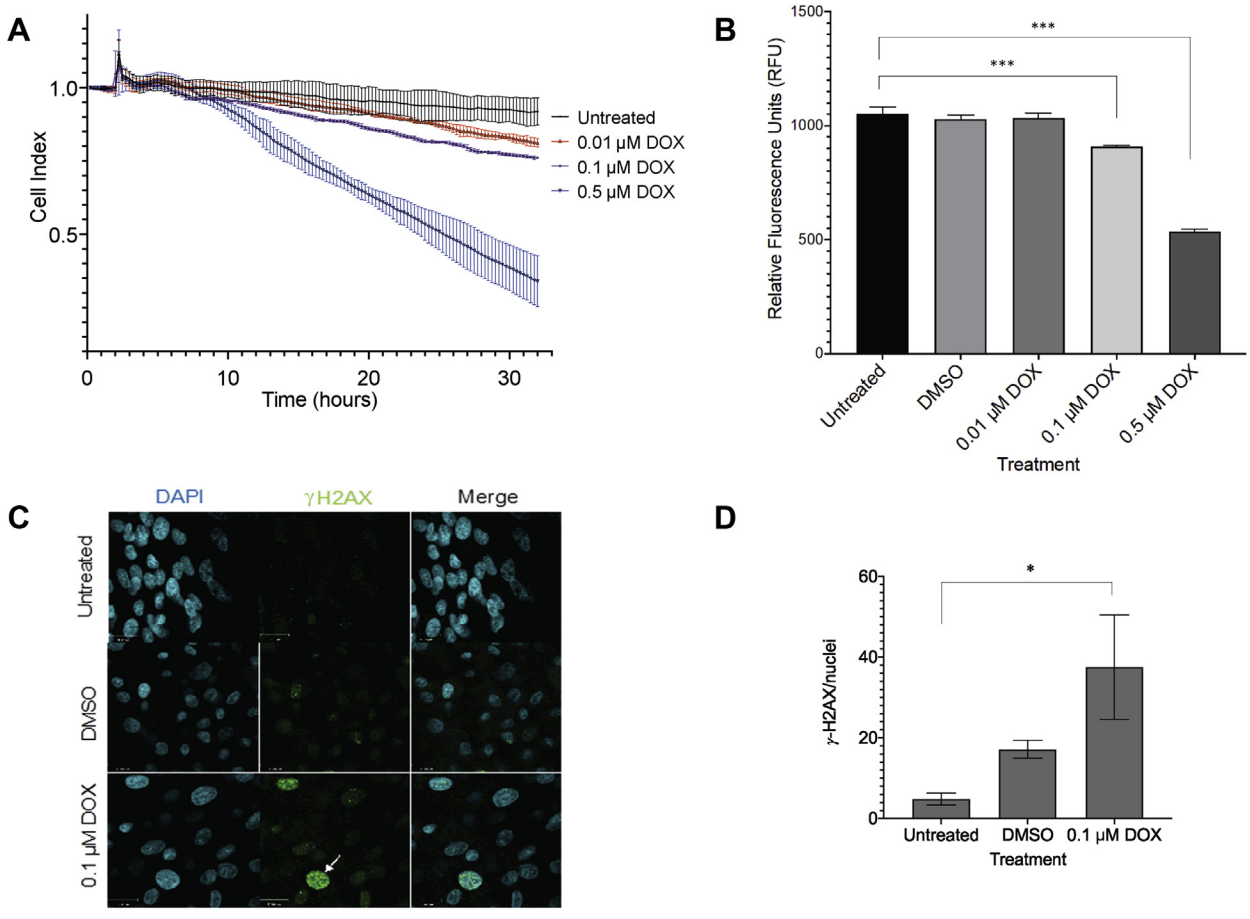
Therefore, absence of these protective variants was associated with risk of anthracycline cardiotoxicity. Logistic regression adjusted for anthracycline dose and the use of concomitant cardiotoxic agents showed that the 31 genes remained significant with similar ORs in unadjusted and adjusted analyses (Supplemental Table 4).

GENE CHARACTERIZATION AND PATHWAY ANALYSIS. The functions of the 31 prioritized genes are summarized in Supplemental Table 5. Most genes belonged to the PI3K/AKT/mTOR pathway and p53 signaling pathways. The PI3K pathway was also 1 of the 6 pathways showing differential enrichment between case and control patients using GeneMania (**Figures 3A and 3B**). *ZNF827*, *ELAC2*, *SEC62*, *USP42*, and *PIK3R2* were selected for functional analysis based on 1 or more of the following additional criteria: biological relevance, protein expression in the human heart (Supplemental Table 5), and mRNA expression in hiPSC-CMs.

FUNCTIONAL EVALUATION IN hiPSC-CMs. DOX decreased CM viability and increased *ZNF827*, *ELAC2*, and *PI3KR2* expression. A 24-hour treatment with DOX resulted in a dose-dependent decrease in cell index in hiPSC-CMs (**Figure 4A**), as well as in metabolic activity and proliferation (**Figure 4B**). Treatment with 0.1 μmol/l DOX caused an increase in γ-H2AX nuclear foci, a marker of double-stranded DNA breaks, compared to untreated cells ($p = 0.039$) (**Figures 4C and 4D**). DOX treatment up-regulated mRNA expression of *ZNF827* ($p = 0.001$), *ELAC2* ($p = 0.012$), and *PI3KR2* ($p = 0.016$) compared to DMSO but did not change *SEC62* and *USP42* mRNA expression (**Figure 5A**). Based on the availability of targeted inhibitors, we selected *PI3KR2* and *ZNF827* for further study.

Target gene inhibition prevented DOX-induced cardiotoxicity. iPSC-CMs from 2 healthy individuals were pre-treated for 24 h before DOX exposure with either a nontargeted inhibitor, dexrazoxane (iron chelator), or with targeted inhibitors: TGX-221 (PI3KR2 inhibitor), rapamycin (mTOR inhibitor downstream of PI3KR2), or metformin (ZNF827 inhibitor) (Supplemental Table 6). Dose-response curves of DOX-induced decrease in CM viability were generated, and the IC_{50} values for DOX (i.e., the dose of DOX that causes a 50% decrease in viability) were determined (**Figures 5B and 5C**). The goodness of fit R^2 values were >0.9 for all IC_{50} curves. A higher IC_{50} indicates lower drug potency or need for a higher drug dose to cause 50% target effect. The IC_{50} values (95% CI) of DOX demonstrated that targeted inhibitors, TGX-221 and metformin, were effective at blocking DOX-induced decrease in CM viability in both lines. Metformin was superior in its cardioprotective effect compared to dexrazoxane in both lines, and TGX-221 was superior in PGP17 and comparable to dexrazoxane in PGP14 (**Table 2**). Therefore,

FIGURE 4 Effect of Anthracycline in hiPSC-CMs



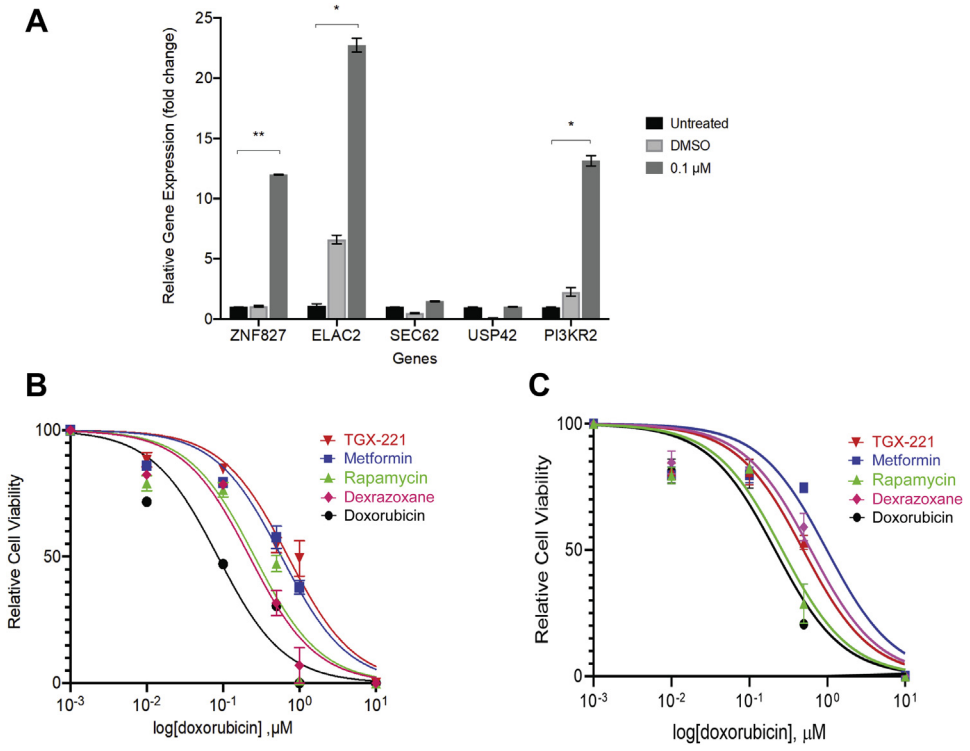
(A) The 24-h DOX treatment caused a dose-dependent decrease in cell function and viability in hiPSC-CMs measured using the cell index. **(B)** Presto blue cell viability assay demonstrated a decrease in metabolic activity and proliferation with increasing DOX doses. The values represent the average relative fluorescence from 3 independent experiments. **(C)** Representative immunofluorescence images showing increased γ-H2AX staining **(green) (white arrow)**, a DNA damage marker, in the nuclei (blue DAPI staining) of DOX-treated cells. **(D)** DOX treatment increased average γ-H2AX foci per nucleus compared to untreated cells. Error bars represent mean ± SD for 3 independent biological replicates. *$p < 0.05$; ***$p < 0.001$. CMC = combined multivariate and collapsing; DAPI = 4′,6-diamidino-2-phenylindole; DMSO = dimethyl sulfoxide; DOX = doxorubicin; H2AX = H2A family member X; hiPSC-CM = human induced pluripotent stem cell-derived cardiomyocyte; μM = μmol/l.

pharmacologic disruption of gene-associated pathways was cardioprotective, similar to the presence of putatively disruptive genetic variants in these pathways.

REPLICATION OF FUNCTIONALLY VALIDATED GENES. Variant burden in the top 2 genes, *PI3KR2* and *ZNF827*, was compared between 30 case patients and 30 propensity-matched control patients in the replication cohort (Supplemental Table 7). There was a lower variant burden in *PI3KR2* in case patients (6.7%) compared to control patients (26.7%) (OR: 0.196; 95% CI: 0.038 to 1.02; p = 0.038). McNemar test identified *PIK3R2* as the gene with the highest number of discordant pairs (n = 10). Although not statistically significant (p = 0.114), 8 of the 10 pairs involved variants present in control patients but absent in case patients. This difference was nominally significant on conditional logistic regression (p = 0.08). No test could be performed for 11 genes that did not harbor variants in the replication cohort, which reduced our power to detect discordant pairs. A comparison of variants in *ZNF827* could not be performed because only 2.4% of the cohort harbored a variant, all of which were in control patients.

RF RISK PREDICTION MODELING. We trained 3 RF prediction models in a training set and evaluated

FIGURE 5 Effect of Targeted Gene Inhibition on DOX-Induced Cardiotoxicity in hiPSC-CMs



(A) RT-qPCR analysis of PGP17 hiPSC-CMs (3 biological replicates, each containing 3 technical replicates) treated with 0.1 μmol/L DOX demonstrated a significant increase in gene expression levels of *ZNF827*, *ELAC2*, and *PI3KR2*. **(B)** PGP17 and (**C**) PGP14 hiPSC-CMs were pre-treated with DMSO or an inhibitor for 24 h and then exposed to DOX for 24 h. Dose-response curves for cell index, that is, CM viability, and IC_{50} values for DOX with and without inhibitors are shown. Data are mean ± SD for 3 independent replicates. *$p < 0.05$; **$p < 0.01$. PGP = Personal Genome Project; RT-qPCR = quantitative reverse transcription polymerase chain reaction; other abbreviations as in Figure 4.

their performance in the testing and overall set (Table 3). We chose a random resampling approach, which is superior to a training/validation split in the absence of an external validation cohort. Box-plots of the accuracy measures are shown in Figures 6A to 6C. The clinical RF model had the lowest performance across all accuracy measures and higher misclassification rate compared to the genetic and combined models. The performance metrics of the 3 models in the training set, the testing set, and overall dataset are shown in Table 3. In the testing set, the AUC for the clinical model was 0.59 (95% CI: 0.51 to 0.67), for the genetic model was 0.71 (95% CI: 0.63 to 0.80), and for the combined model was 0.72 (95% CI: 0.63 to 0.80) (Figure 6D). The combined model outperformed the clinical model with a higher AUC, higher specificity, higher positive predictive value, and a lower misclassification rate.

DISCUSSION

Despite dose-dependent anthracycline cardiotoxicity, only a proportion of patients who receive high-dose anthracycline develop cardiotoxicity (35). Our findings showed that almost 90% of patients without cardiotoxicity harbored rare/low-frequency variants in cardiac injury pathways that likely protected them from the damaging effects of anthracycline. In contrast, <50% of patients with cardiotoxicity harbored these protective variants. The resistance to cardiotoxicity with disruption of these pathways was confirmed through pharmacological inhibition of the pathways in human CMs. Using these genetic findings in combination with clinical risk predictors, we developed and internally validated a prediction model and demonstrated its superiority to the clinical model in predicting cardiotoxicity. Furthermore, we identified potentially druggable pathways of

TABLE 2 IC_{50} Values (95% CI) of DOX and Inhibitors in hiPSC-CMs Derived From PGP17 and PGP14 Individuals

	PGP17_11		PGP14_26	
	IC_{50}	p Value vs. DOX	IC_{50}	p Value vs. DOX
Doxorubicin	0.09 (0.04–0.16)		0.21 (0.14–0.31)	
TGX-221	0.71 (0.51–0.98)	0.045	0.49 (0.33–0.72)	0.043
Metformin	0.58 (0.43–0.78)	0.017	0.98 (0.56–1.80)	0.039
Rapamycin	0.26 (0.14–0.44)	0.178	0.26 (0.17–0.39)	0.164
Dexrazoxane	0.22 (0.15–0.33)	0.077	0.61 (0.43–0.88)	0.047

CI = confidence interval; DOX = doxorubicin; hiPSC-CMs = human induced pluripotent stem cell-derived cardiomyocytes; IC_{50} = half-maximal inhibitory concentration; PGP = Personal Genome Project.

cardiotoxicity that may offer better cardioprotection than currently available drugs like dexrazoxane. A strength of our study was its prospective longitudinal design with adjustment for length of follow-up, which minimized potential misclassification of case and control patients, as well as confirmation of our findings through functional validation and replication studies.

RARE VARIANT CONTRIBUTION TO ANTHRACYCLINE CARDIOTOXICITY. Previous single-nucleotide polymorphism association studies have identified common variants associated with anthracycline cardiotoxicity, often in pharmacokinetic genes, that account for a relatively small proportion of cases (10). Often these studies did not account for clinical and treatment differences, differences in anthracycline dose, or differences in follow-up between case and control patients. An important rationale for an extreme phenotype approach stratified by anthracycline dose was to enrich our cohort for rare and low-frequency variants that are expected to have a stronger effect on phenotype rather than only common variants. An extreme phenotype design is superior to a simple case-control design for detecting rare risk variants, especially protective variants (36). To minimize bias related to an imbalance in cancer types between case and control patients, we excluded cancer-causing variants and adjusted our analysis for cancer diagnosis.

TABLE 3 Accuracy Measures for Prediction Models of Anthracycline Cardiotoxicity

Accuracy Measure	Dataset	Clinical Model	Genetic Model	Combined Clinical and Genetic Model
AUC	Training	0.9990 ± 0.0026	0.8996 ± 0.0216	0.9923 ± 0.0063
	Testing	0.5896 ± 0.0431	0.7133 ± 0.042	0.7156 ± 0.0421
	Overall	0.7403 ± 0.0273	0.7819 ± 0.0262	0.8174 ± 0.0267
Sn	Training	0.9980 ± 0.0053	0.8232 ± 0.0457	0.9846 ± 0.0127
	Testing	0.4317 ± 0.0885	0.5552 ± 0.0827	0.5254 ± 0.0836
	Overall	0.6401 ± 0.056	0.6538 ± 0.054	0.6944 ± 0.0528
Sp	Training	1.0000 ± 0.0004	0.9760 ± 0.0168	0.9999 ± 0.0008
	Testing	0.7475 ± 0.0682	0.8715 ± 0.0513	0.9057 ± 0.047
	Overall	0.8404 ± 0.0431	0.9100 ± 0.035	0.9404 ± 0.0297
PPV	Training	1.0000 ± 0.0005	0.9609 ± 0.0254	0.9999 ± 0.0011
	Testing	0.4885 ± 0.0749	0.7259 ± 0.0820	0.7623 ± 0.0942
	Overall	0.6767 ± 0.0473	0.8124 ± 0.0544	0.8498 ± 0.0595
NPV	Training	0.9986 ± 0.0035	0.8893 ± 0.0253	0.9896 ± 0.0084
	Testing	0.7058 ± 0.0384	0.7684 ± 0.0376	0.777 ± 0.0365
	Overall	0.8135 ± 0.0243	0.8129 ± 0.0239	0.8552 ± 0.0231
MC	Training	0.0008 ± 0.0021	0.0864 ± 0.0182	0.0063 ± 0.0051
	Testing	0.3652 ± 0.0423	0.2465 ± 0.0385	0.2297 ± 0.0373
	Overall	0.2311 ± 0.0268	0.1876 ± 0.0241	0.1475 ± 0.0237
FPR	Training	0.0000 ± 0.0002	0.0141 ± 0.0098	0.0000 ± 0.0005
	Testing	0.1632 ± 0.0457	0.0807 ± 0.0325	0.0610 ± 0.0308
	Overall	0.1032 ± 0.0289	0.0562 ± 0.0220	0.0386 ± 0.0195
FNR	Training	0.0008 ± 0.0021	0.0723 ± 0.0185	0.0063 ± 0.0051
	Testing	0.2020 ± 0.0365	0.1658 ± 0.0333	0.1687 ± 0.0334
	Overall	0.1279 ± 0.0231	0.1314 ± 0.0215	0.1089 ± 0.0211

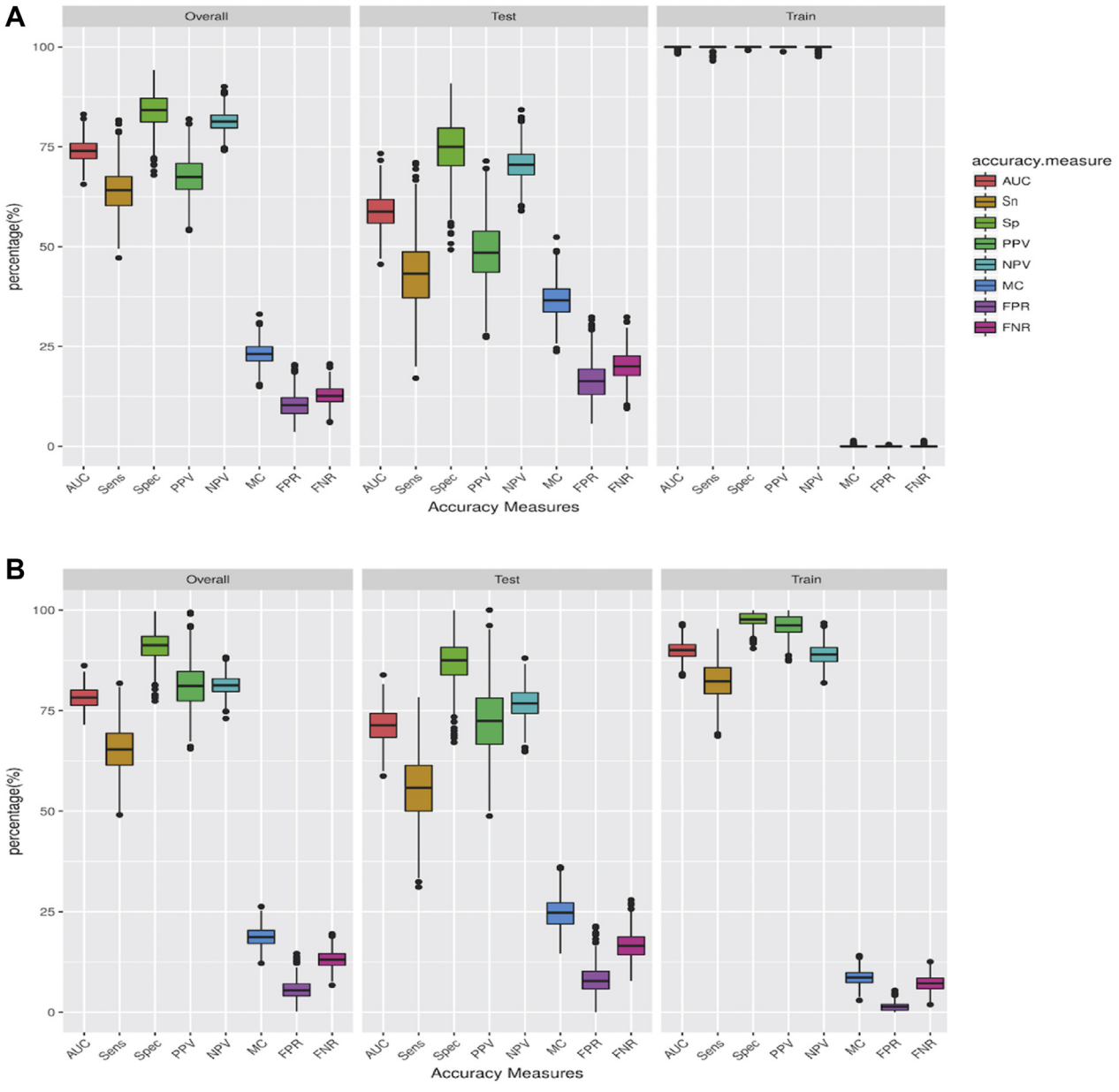
Values are mean ± SD.

AUC = area under the curve; FNR = false negative rate; FPR = false positive rate; MC = misclassification; NPV = negative predictive value; PPV = positive predictive value; Sn = sensitivity; Sp = specificity.

Of the differentially enriched genes in our study, only *CELF4* has been previously reported as a cardiotoxicity gene in human studies (8,37). *ZNF827* and *PI3KR2*, which demonstrated a higher variant burden in control patients compared to case patients, are involved in autophagy, a pathway known to be involved in anthracycline cardiotoxicity and cardiovascular disease. *PI3KR2* encodes the p85β isoform, a regulatory subunit involved in autophagy inhibition through the phosphoinositide dependent kinase-1 and protein kinase B signaling pathway (38). Although we did not measure autophagy, we did find *PI3KR2* up-regulation and accumulation of DNA damage with DOX that may be related to impaired autophagy-mediated DNA repair (39). TGX-221, a selective inhibitor of *PI3KR2*, yielded a greater cardioprotective effect against DOX compared to dexrazoxane, an iron chelator. Importantly, preclinical tumor studies have reported that PI3KR2 inhibition reduces tumorigenicity, suggesting that targeting PI3KR2 could mediate a cardioprotective effect without being tumor protective, one of the main concerns with the use of dexrazoxane (40,41).

We also noted an important role for *ZNF827* in anthracycline cardiotoxicity through enrichment of variants in control patients, increased mRNA expression on DOX exposure in CMs, and protection against DOX cardiotoxicity with metformin. Metformin, involved in 5′ adenosine monophosphate-activated protein kinase signaling, suppresses the activation of the nuclear receptor TR4, which in turn inhibits the recruitment of ZNF827 and consequently induces cardiac autophagy. Studies have demonstrated a cardioprotective effect of metformin in DOX-induced cardiotoxicity in rats through normalized autophagy processes (32). Our finding of the

FIGURE 6 **Accuracy Measures of Prediction Models Using Random Forest**



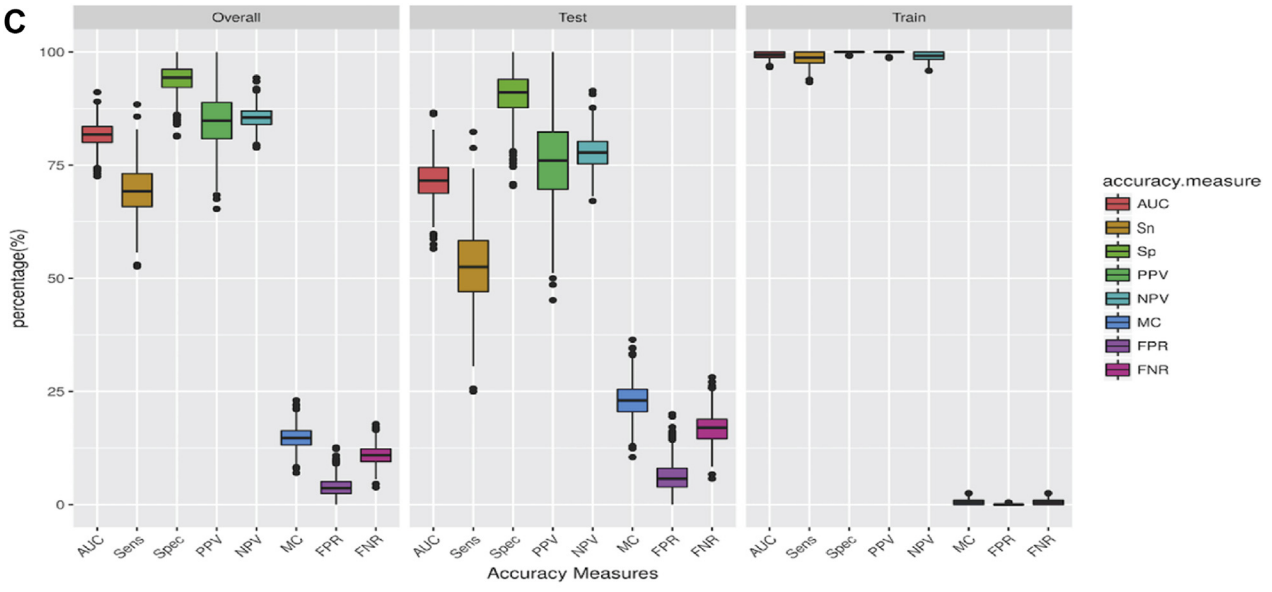
Boxplots representing the accuracy measures for the overall dataset, the testing set, and the training set for the **(A)** clinical model, **(B)** genetic model, **(C)** combined model, and **(D)** receiver operator characteristic AUC for model-derived prediction of anthracycline cardiotoxicity. Clinical model AUC: 0.59 **(black)**, genetic model AUC: 0.71 **(blue)**, combined model AUC: 0.72 **(red)**. AUC = area under the curve; FNR = false negative rate; FPR = false positive rate; MC = misclassification; NPV = negative predictive value; PPV = positive predictive value; Sens = sensitivity; Sn = sensitivity; Sp = specificity; Spec = specificity.

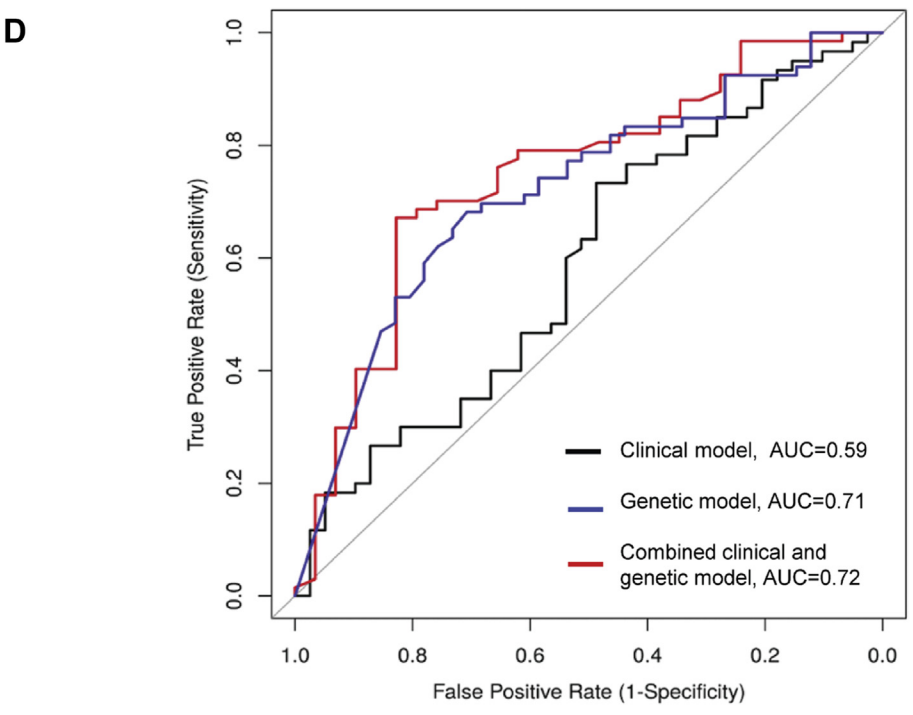
Continued on the next page

cardioprotective effect of metformin, an antidiabetic drug, in human CMs suggests a promising potential for drug repurposing of metformin for protection against anthracycline cardiotoxicity.

MACHINE LEARNING IN RISK PREDICTION MODELING. An important outcome of our study was the development of a prediction model for anthracycline cardiotoxicity using a combination of genetic

FIGURE 6 Continued





and clinical factors that outperformed the clinical model. In particular, the combined model had high specificity, good sensitivity, and a low misclassification rate. Although this model is exploratory and requires external validation, a model with high specificity is desirable for the identification of at-risk individuals with high confidence in whom the modification of chemotherapy or use of cardioprotective agents could be justified. By avoiding overestimation of the risk of anthracycline cardiotoxicity, it would prevent withholding life-saving anthracycline in patients who are not at high risk.

STUDY LIMITATIONS. The study was limited to late cancer survivors, and therefore, early cardiotoxicity was not explored. Independent validation of this prediction model is needed in an external replication cohort. Permutation tests for genes were not performed, which could provide additional support for the findings from SKAT/SKAT-optimized and burden analysis. Although we confirmed that the iPSCs did not harbor pathogenic variants in known cardiac genes, iPSC-CM studies from additional individuals with different genetic backgrounds would be helpful because all variants pre-disposing to cardiotoxicity may not have been considered. Finally, although exome sequencing enables the discovery of rare and low-frequency variants in coding regions of genes, it excludes common variants in noncoding regions that may have an association with cardiotoxicity; furthermore, studies that combine the contribution of rare and common variants is warranted.

CONCLUSIONS

Childhood cancer survivors with rare variants in genes involved in cardiac injury pathways have a lower susceptibility to anthracycline cardiotoxicity. Incorporating these genetic factors into a prediction model helped identify individuals at lower risk for anthracycline cardiotoxicity with high specificity. This knowledge may be useful in the future to individualize anthracycline-based chemotherapy tailored to a patient's risk for cardiotoxicity. The identification of promising biological targets involved in autophagy (*PI3KR2* and *ZNF827*) will inform the development and/or repurposing of cardioprotective agents in patients receiving anthracyclines. Thus, our study shows the power of a precision approach in informing strategies for tailored management of childhood cancer patients.

ACKNOWLEDGMENTS We thank Sangsoon Woo from Axio Research for statistical data analysis related to RF model development, the SickKids Centre for Applied Genomics for exome sequencing, and the SickKids Imaging Facility for technical support.

AUTHOR DISCLOSURES

The study was supported by the Canadian Cancer Society, Toronto, Canada; Canadian Institutes of Health Research, Toronto, Canada; the Pediatric Oncology Group of Ontario, Toronto, Canada; the Ontario Institute for Cancer Research, Toronto, Canada; Children's Cancer and Blood disorders, Toronto, Canada; Ted Rogers Centre for Heart Research, Toronto, Canada; and the Labatt Family Heart Center at the Hospital for Sick Children, Toronto, Canada. Dr. Sender is Senior Vice President, Medical Affairs for Pediatric, Adolescent and Young Adult Oncology at NantKwest. All other authors have reported that they have no relationships relevant to the contents of this paper to disclose.

ADDRESS FOR CORRESPONDENCE: Dr. Seema Mital, Hospital for Sick Children, 555 University Avenue, Toronto, Ontario M5G 1X8, Canada. E-mail: seema.mital@sickkids.ca. Twitter: @hcbiobank.

PERSPECTIVES

COMPETENCY IN MEDICAL KNOWLEDGE: Childhood cancer survivors harboring rare variants in pathways of cardiac injury are protected from anthracycline cardiotoxicity. A prediction model for anthracycline cardiotoxicity that combines clinical and genetic risk factors performs better than a clinical model.

TRANSLATIONAL OUTLOOK: External validation of the risk prediction model is needed to inform its use as a clinical decision support tool for tailoring anthracycline chemotherapy to individual risk. The identified genes suggest novel biological pathways for development of cardioprotective agents.

REFERENCES

1. Octavia Y, Tocchetti CG, Gabrielson KL, Janssens S, Crijns HJ, Moens AL. Doxorubicin-induced cardiomyopathy: from molecular mechanisms to therapeutic strategies. J Mol Cell Cardiol 2012;52:1213–25.

2. Nathan PC, Amir E, Abdel-Qadir H. Cardiac outcomes in survivors of pediatric and adult cancers. Can J Cardiol 2016;32:871–80.

3. Lipshultz SE, Lipsitz SR, Mone SM, et al. Female sex and higher drug dose as risk factors for late cardiotoxic effects of doxorubicin therapy for childhood cancer. N Engl J Med 1995;332:1738–43.

4. Hershman DL, McBride RB, Eisenberger A, Tsai WY, Grann VR, Jacobson JS. Doxorubicin, cardiac risk factors, and cardiac toxicity in elderly patients with diffuse B-cell non-Hodgkin's lymphoma. J Clin Oncol 2008;26:3159–65.

5. Pein F, Sakiroglu O, Dahan M, et al. Cardiac abnormalities 15 years and more after adriamycin therapy in 229 childhood survivors of a solid tumour at the Institut Gustave Roussy. Br J Cancer 2004;91:37–44.

6. Aminkeng F, Bhavsar AP, Visscher H, et al. A coding variant in RARG confers susceptibility to anthracycline-induced cardiotoxicity in childhood cancer. Nat Genet 2015;47:1079–84.

7. Schneider BP, Shen F, Gardner L, et al. Genome-wide association study for anthracycline-induced congestive heart failure. Clin Cancer Res 2017;23:43–51.

8. Wang X, Sun CL, Quinones-Lombrana A, et al. CELF4 variant and anthracycline-related cardiomyopathy: a Children's Oncology Group genome-wide association study. J Clin Oncol 2016;34:863–70.

9. Wells QS, Veatch OJ, Fessel JP, et al. Genome-wide association and pathway analysis of left ventricular function after anthracycline exposure in adults. Pharmacogenet Genomics 2017;27:247–54.

10. Linschoten M, Teske AJ, Cramer MJ, van der Wall E, Asselbergs FW. Chemotherapy-related cardiac dysfunction: a systematic review of genetic variants modulating individual risk. Circ Genom Precis Med 2018;11:e001753.

11. Visscher H, Ross CJ, Rassekh SR, et al. Pharmacogenomic prediction of anthracycline-induced

cardiotoxicity in children. J Clin Oncol 2012;30:1422–8.

12. Skitch A, Mital S, Mertens L, et al. Novel approaches to the prediction, diagnosis and treatment of cardiac late effects in survivors of childhood cancer: a multi-centre observational study. BMC Cancer 2017;17:519–28.

13. Children's Oncology Group. Long-term follow up guidelines for survivors of childhood, adolescent, and young adult cancer. Version 4.0. October 2013. Available at: http://www.survivorshipguidelines.org/pdf/ltfuguidelines_40.pdf. Accessed September 23, 2020.

14. Lang RM, Bierig M, Devereux RB, et al. Recommendations for chamber quantification: a report from the American Society of Echocardiography's Guidelines and Standards Committee and the Chamber Quantification Writing Group, developed in conjunction with the European Association of Echocardiography, a branch of the European Society of Cardiology. J Am Soc Echocardiogr 2005;18:1440–63.

15. Cardinale D, Colombo A, Bacchiani G, et al. Early detection of anthracycline cardiotoxicity and improvement with heart failure therapy. Circulation 2015;131:1981–8.

16. DePristo MA, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 2011;43:491–8.

17. International HapMap C, Frazer KA, Ballinger DG, et al. A second generation human haplotype map of over 3.1 million SNPs. Nature 2007;449:851–61.

18. Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature 2016;536:285–91.

19. Tate JG, Bamford S, Jubb HC, et al. COSMIC: the catalogue of somatic mutations in cancer. Nucleic Acids Res 2018;47:D941–7.

20. Li B, Leal SM. Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. Am J Hum Genet 2008;83:311–21.

21. Wu MC, Lee S, Cai T, Li Y, Boehnke M, Lin X. Rare-variant association testing for sequencing data with the sequence kernel association test. Am J Hum Genet 2011;89:82–93.

22. Lee S, Emond MJ, Bamshad MJ, et al. Optimal unified approach for rare-variant association testing with application to small-sample case-control whole-exome sequencing studies. Am J Hum Genet 2012;91:224–37.

23. Warde-Farley D, Donaldson SL, Comes O, et al. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. Nucleic Acids Res 2010;38:W214–20.

24. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005;102:15545–50.

25. Hildebrandt MR, Reuter MS, Wei W, et al. Precision health resource of control iPSC lines for versatile multilineage differentiation. Stem Cell Reports 2019;13:1126–41.

26. Reuter MS, Walker S, Thiruvahindrapuram B. The Personal Genome Project Canada: findings from whole genome sequences of the inaugural 56 participants. CMAJ 2018;190:E126–36.

27. Zhang X, Guo L, Zeng H, et al. Multi-parametric assessment of cardiomyocyte excitation-contraction coupling using impedance and field potential recording: a tool for cardiac safety assessment. J Pharmacol Toxicol Methods 2016;81:201–16.

28. Rao X, Huang X, Zhou Z, Lin X. An improvement of the 2-(-delta CT) method for quantitative real-time polymerase chain reaction data analysis. Biostat Bioinforma Biomath 2013;3:71–85.

29. Wang S, Konorev EA, Kotamraju S, Joseph J, Kalivendi S, Kalyanaraman B. Doxorubicin induces apoptosis in normal and tumor cells via distinctly different mechanisms. intermediacy of H_2O_2- and p53-dependent pathways. J Biol Chem 2004;279:25535–43.

30. Foglesong PD, Reckord C, Swink S. Doxorubicin inhibits human DNA topoisomerase I. Cancer Chemother Pharmacol 1992;30:123–5.

31. Conomos D, Reddel RR, Pickett HA. NuRD-ZNF827 recruitment to telomeres creates a molecular scaffold for homologous recombination. Nat Struct Mol Biol 2014;21:760–70.

32. Zilinyi R, Czompa A, Czegledi A, et al. The cardioprotective effect of metformin in doxorubicin-induced cardiotoxicity: the role of autophagy. Molecules 2018;23:1184–96.

33. Kitani T, Ong SG, Lam CK, et al. Human-induced pluripotent stem cell model of trastuzumab-induced cardiac dysfunction in patients with breast cancer. Circulation 2019;139:2451–65.

34. Division of Biomedical Statistics and Informatics, Mayo Clinic Research. Biomedical statistics and informatics software packages 2013. Available at: http://bioinformaticstools.mayo.edu/research/gmatch/. Accessed November 2020.

35. Drafts BC, Twomley KM, D'Agostino R Jr., et al. Low to moderate dose anthracycline-based chemotherapy is associated with early noninvasive imaging evidence of subclinical cardiovascular disease. J Am Coll Cardiol Img 2013;6:877–85.

36. Li D, Lewinger JP, Gauderman WJ, Murcray CE, Conti D. Using extreme phenotype sampling to identify the rare causal variants of quantitative traits in association studies. Genet Epidemiol 2011;35:790–9.

37. Gomes AV, Venkatraman G, Davis JP, et al. Cardiac troponin T isoforms affect the Ca(2+) sensitivity of force development in the presence of slow skeletal troponin I: insights into the role of troponin T isoforms in the fetal heart. J Biol Chem 2004;279:49579–87.

38. Hassan B, Akcakanat A, Holder AM, Meric-Bernstam F. Targeting the PI3-kinase/Akt/mTOR signaling pathway. Surg Oncol Clin N Am 2013;22:641–64.

39. Liu EY, Xu N, O'Prey J, et al. Loss of autophagy causes a synthetic lethal deficiency in DNA repair. Proc Natl Acad Sci U S A 2015;112:773–8.

40. Deng S, Yan T, Jendrny C, et al. Dexrazoxane may prevent doxorubicin-induced DNA damage via depleting both topoisomerase II isoforms. BMC Cancer 2014;14:842–54.

41. Chen S, Fisher RC, Signs S, et al. Inhibition of PI3K/Akt/mTOR signaling in PI3KR2-overexpressing colon cancer stem cells reduces tumor growth due to apoptosis. Oncotarget 2017;8:50476–88.

KEY WORDS anthracycline, cancer survivorship, cardiomyopathy, echocardiography, genomics, machine learning, risk prediction

APPENDIX For supplemental tables and figures, please see the online version of this paper.

Go to http://www.acc.org/jacc-journals-cme to take the CME/MOC/ECME quiz for this article.