JACC: CARDIOONCOLOGY VOL. 5, NO. 3, 2023

ORIGINAL RESEARCH

Haptoglobin Gene Expression and Anthracycline-Related Cardiomyopathy in Childhood Cancer Survivors

A COG-ALTE03N1 Report

Purnima Singh, MS, PhD, MSPH,[a,b] David K. Crossman, PhD,[c] Liting Zhou, MS,[a] Xuexia Wang, PhD,[d] Noha Sharafeldin, PhD,[a] Lindsey Hageman, MPH,[a] Javier G. Blanco, PhD,[e] Paul W. Burridge, PhD,[f] Saro H. Armenian, DO, MPH,[g] Frank M. Balis, MD,[h] Douglas S. Hawkins, MD,[i] Frank G. Keller, MD,[j] Melissa M. Hudson, MD,[k] Joseph P. Neglia, MD, MPH,[l] A. Kim Ritchey, MD,[m] Jill P. Ginsberg, MD,[h] Wendy Landier, PhD,[a,b] Smita Bhatia, MD, MPH[a,b]

ABSTRACT

BACKGROUND Anthracycline-related cardiomyopathy is a leading cause of premature death in childhood cancer survivors. The high interindividual variability in risk suggests the need to understand the underlying pathogenesis.

OBJECTIVES The authors interrogated differentially expressed genes (DEGs) to identify genetic variants serving regulatory functions or genetic variants not easily identified when using genomewide array platforms. Using leads from DEGs, candidate copy number variants (CNVs) and single-nucleotide variants (SNVs) were genotyped.

METHODS Messenger RNA sequencing was performed on total RNA from peripheral blood of 40 survivors with cardiomyopathy (cases) and 64 matched survivors without cardiomyopathy (control subjects). Conditional logistic regression analysis adjusting for sex, age at cancer diagnosis, anthracycline dose, and chest radiation was used to assess the associations between gene expression and cardiomyopathy and between CNVs and SNVs and cardiomyopathy.

RESULTS Haptoglobin (*HP*) was identified as the top DEG. Participants with higher *HP* gene expression had 6-fold greater odds of developing cardiomyopathy (OR: 6.4; 95% CI: 1.4-28.6). The *HP2*-specific allele among the *HP* genotypes (HP1-1, HP1-2, and HP2-2) had higher transcript levels, as did the G allele among SNVs previously reported to be associated with *HP* gene expression (rs35283911 and rs2000999). The HP1-2 and HP2-2 genotypes combined with the G/G genotype for rs35283911 and/or rs2000999 placed the survivors at 4-fold greater risk (OR: 3.9; 95% CI: 1.0-14.5) for developing cardiomyopathy.

CONCLUSIONS These findings provide evidence of a novel association between *HP2* allele and cardiomyopathy. HP binds to free hemoglobin to form an HP-hemoglobin complex, thereby preventing oxidative damage from free heme iron, thus providing biological plausibility to the mechanistic basis of the present observation. (J Am Coll Cardiol CardioOnc 2023;5:392–401) © 2023 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/).

From the [a]Institute for Cancer Outcomes and Survivorship, University of Alabama at Birmingham, Birmingham, Alabama, USA; [b]Department of Pediatrics, University of Alabama at Birmingham, Birmingham, Alabama, USA; [c]Department of Genetics, University of Alabama at Birmingham, Birmingham, Alabama, USA; [d]Department of Mathematics, University of North Texas, Denton, Texas, USA; [e]Department of Pharmaceutical Sciences, State University of New York at Buffalo, Buffalo, New York, USA; [f]Department of Pharmacology, Northwestern University, Chicago, Illinois, USA; [g]Department of Population Sciences, City of Hope,

ISSN 2666-0873 https://doi.org/10.1016/j.jaccao.2022.09.009

Anthracycline-related cardiomyopathy is one of the leading causes of premature death after childhood cancer.[1] Previous studies have identified genomic variants associated with anthracycline-related cardiomyopathy.[2] Some studies have extended these findings to examine the functional relevance of the identified genomic variants.[3,4] Nonetheless, the functional impact of several genomic variants associated with anthracycline-related cardiomyopathy remains unknown, as many of these variants are located in noncoding (intronic or intergenic) regions and do not change the protein structure; however, these may have a regulatory function (eg, by affecting expression of target genes).[5]

Interrogating differentially expressed genes (DEGs) may allow the identification of genetic variants serving regulatory functions (expression quantitative trait loci) or genetic variants that are not easily identified when using genomewide array platforms.[6-8] However, differential constitutive gene expression in survivors with and without cardiomyopathy remains understudied and could enhance our knowledge of the pathogenesis of cardiomyopathy. We have previously used microarray-based gene expression analysis to identify the role of *GSTM1* in anthracycline-related cardiomyopathy in childhood cancer survivors.[9] Gene expression microarrays, however, have limited probe sets and are not suited to identify transcript isoforms.[10] RNA sequencing (RNA-seq) can identify known and novel transcripts and measure transcript abundance and splicing isoforms in a high-throughput and quantitative manner.[11] Furthermore, messenger RNA (mRNA) sequencing by oligo (dT) selection of poly $(A)^+$ mRNA allows interrogation of the protein-coding fractions of the transcriptome. We performed mRNA sequencing using whole-blood RNA from anthracycline-exposed childhood cancer survivors who had cardiomyopathy (cases) matched to those without cardiomyopathy (control subjects) to further our understanding of the pathogenesis of anthracycline-related cardiomyopathy.

METHODS

STUDY DESIGN. Study participants were drawn from a Children's Oncology Group study (COG-ALTE03N1 [Genetic Analysis in Identifying Late-Occurring Complications in Childhood Cancer Survivors]; NCT00082745) that used a matched case-control design to understand the pathogenesis of cardiomyopathy in childhood cancer survivors. Children's Oncology Group member institutions enrolled patients after obtaining approval from local Institutional Review Boards. Written informed consent or assent was obtained from patients, parents, or legal guardians. Cases and control subjects were identified from individuals diagnosed with cancer at ≤21 years of age. Cases consisted of childhood cancer survivors who developed cardiomyopathy after exposure to anthracyclines. For each case, patients who had no signs or symptoms of cardiomyopathy after anthracycline exposure were randomly selected as control subjects from the same Children's Oncology Group childhood cancer survivor cohort, matched on primary cancer diagnosis, year of diagnosis (±5 years), and race/ethnicity. The selected control subjects also needed to have a longer duration of cardiomyopathy-free follow-up compared with time from cancer diagnosis to cardiomyopathy for the corresponding case. Participants provided blood samples in PAXgene blood RNA tubes for RNA and blood samples in dipotassium ethylenediaminetetraacetic acid tubes or saliva in Oragene kits (DNA Genotek) for DNA.

Cases fulfilled the American Heart Association criteria for cardiac compromise by presenting with signs and/or symptoms (dyspnea, orthopnea, fatigue, edema, hepatomegaly, and/or rales). In the absence of signs or symptoms, cases had echocardiographic features of left ventricular dysfunction (ejection fraction ≤40% and/or fractional shortening ≤ 28%). Lifetime anthracycline exposure was calculated by

ABBREVIATIONS AND ACRONYMS

CNV = copy number variant

DEG = differentially expressed gene

FPKM = fragments per kilobase of transcript per million mapped reads

HB = hemoglobin

HP = haptoglobin

mRNA = messenger RNA

PCR = polymerase chain reaction

RNA-seq = RNA sequencing

SNV = single-nucleotide variant

Duarte, California; [h]Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA; [i]Seattle Children's, Seattle, Washington, USA; [j]Children's Healthcare of Atlanta, Emory University, Atlanta, Georgia, USA; [k]St. Jude Children's Research Hospital, Memphis, Tennessee; [l]University of Minnesota, Minneapolis, Minnesota; and the [m]Children's Hospital of Pittsburgh of the University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania, USA.

Yuri Kim, MD, PhD, served as Guest Editor-in-Chief for this paper.

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 June 13, 2022; revised manuscript received September 27, 2022, accepted September 30, 2022.

multiplying the cumulative dose (milligrams per square meter) of individual anthracyclines by a factor that reflects the drug's cardiotoxic potential[12] and then summing the results. Radiation to the chest with heart in the field was captured as a yes-or-no variable.

RNA ISOLATION, LIBRARY CONSTRUCTION, AND SEQUENCING. RNA was isolated using the PAXgene whole-blood RNA kit (Qiagen). RNA concentration was measured using a Nanodrop ND-1000 Spectrophotometer (Thermo Fisher Scientific). RNA quality was checked on Bioanalyzer Nanochip (Agilent Technologies) and samples with RNA integrity numbers >7 were submitted to the Genomic Services Laboratory at the HudsonAlpha Institute for Biotechnology in Huntsville, Alabama. Polyadenylated RNAs were isolated using NEBNext Magnetic Oligo d(T)25 beads. Libraries were prepared using the TruSeq RNA Sample Preparation Kit (Illumina). Each library was pair-end sequenced (100 bp) by using the TruSeq SBS Kit v4-HS (Illumina), on a HiSeq 2500 platform. Raw reads were demultiplexed using bcl2fastq Conversion Software (Illumina) with default settings.

DIFFERENTIAL GENE EXPRESSION ANALYSIS. Trim Galore! (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) was used to trim off primer adapter sequences found in the raw FASTQ files. STAR was used to align the trimmed RNA-seq FASTQ reads to the human reference genome from GENCODE (GRCh38 p7 release 25).[13] HTSeq-count was used to count the number of reads mapping to each gene from the STAR alignments.[14] Normalization and differential expression was then applied to the count files using DESeq2.[15] The *q* values (adjusted *P* values) from DESeq2 are adjusted for multiple testing using the Benjamini-Hochberg procedure, which controls false discovery rate.

TRANSCRIPT EXPRESSION ANALYSIS. Following mapping of the reads to the reference genome(GRCh38 p13 release 39), the mapped reads were assembled using Cufflinks.[16] Cuffnorm was used to estimate expression levels of individual transcripts assembled by Cufflinks. Cuffnorm output contains normalized expression levels for all gene-specific transcripts. Transcripts with fragments per kilobase of transcript per million mapped reads (FPKM) ≥1 were used to estimate allele-specific transcript abundance.

DNA ISOLATION AND GENOTYPING. DNA was isolated using the Gentra Puregene blood kit (Qiagen) or prepIT•L2P reagent (saliva). DNA concentration was measured using a Nanodrop ND-1000 Spectrophotometer. Copy number variants (CNVs) were determined using polymerase chain reaction (PCR) amplification devised by Koch et al[17] with minor modifications.[18] PCR products from DNA samples were run on 1% and 2% agarose gel (E-Gel agarose gels, Invitrogen, Thermo Fisher Scientific) using an E-Gel Power Snap Electrophoresis System. Genotyping for single-nucleotide variants (SNVs) was performed on the Juno system (Fluidigm) and 96.96 Genotyping IFCs. Endpoint fluorescence values were measured on the BioMark HD system, and Fluidigm SNP Genotyping Analysis software was used to generate genotyping calls for each sample.

STATISTICAL ANALYSIS. Statistical analyses were conducted using JMP 16 Pro (JMP Statistical Discovery) and SAS version 9.4 (SAS Institute). We summarized the characteristics of the study participants according to their case-control status and presented as mean ± SD and median with IQR for continuous variables. Categorical variables are shown as absolute numbers and percentages. Continuous and categorical variables were examined using Wilcoxon rank sum or Kruskal-Wallis tests and Fisher exact or chi-square tests, respectively. Multivariable conditional logistic regression analysis (adjusting for age at diagnosis of primary cancer [continuous variable], sex [male vs female]), anthracycline dose [<250 mg/m^2 vs ≥250 mg/m^2], and chest radiation [yes vs no]) was used to determine the association between gene-specific RNA counts and gene expression–related genetic variants with cardiomyopathy. Model results are presented as ORs and corresponding 95% CIs.

RESULTS

PATIENT CHARACTERISTICS. The median ages at primary cancer diagnosis for the 40 cases and 64 matched control subjects were 8.2 years (IQR: 3.6-13.9 years) and 9.7 years (IQR: 3.3-14.4 years), respectively (**Table 1**). Cases had received higher cumulative anthracycline exposure compared with controls (≥250 mg/m^2, 62.5% vs 35.9%; $P = 0.008$) and were more likely to have received chest radiation (47.5% vs 20.3%; $P = 0.003$). The median time between cancer diagnosis and cardiomyopathy for cases was 5.3 years (IQR: 0.8-12.8 years); control subjects were followed for a significantly longer period (median 10.1 years; IQR: 7.1-14.5 years; $P < 0.001$). The phenotypic characteristics of the cases are presented in Supplemental Table 1.

GENE EXPRESSION AND ANTHRACYCLINE-RELATED CARDIOMYOPATHY. Overall, 43,198 genes (protein coding, long noncoding RNA, processed

pseudogenes, unprocessed pseudogenes, miscellaneous RNA, small nuclear RNA, small nucleolar RNA, and transcribed pseudogenes) were expressed in total RNA from peripheral blood. We filtered 28,026 low-expressed genes (RNA counts <10 in ≥70% of the samples). The remaining 15,172 genes included protein-coding genes (82%), long noncoding RNA, pseudogenes, and immunoglobulin genes. Using DESeq2, we identified 36 DEGs with adjusted *P* values <0.05 and absolute fold changes >2 (Supplemental Table 2); 35 were up-regulated among cases, and 1 was down-regulated. Prioritization of these 36 genes with >50% of matched sets showing differential expression between the cases and their matched control subjects, as well as previously published association with heart disease, identified haptoglobin (*HP*) as the top differentially expressed gene, with 25 of 40 cases (62.5%) having higher expression compared with control subjects. In the present report we focus on *HP* gene expression and its association with anthracycline-related cardiomyopathy.

HP GENE EXPRESSION AND ANTHRACYCLINE-RELATED CARDIOMYOPATHY. Figure 1 shows the normalized aggregate read RNA counts of matched case-control sets plotted side by side for the *HP* gene. Overall, median expression of *HP* gene was higher among cases compared with control subjects (median RNA count 135.3 [IQR: 63.7-300.2] vs 70.3 [IQR: 47.5-105.8]; $P < 0.001$). Dichotomizing *HP* RNA counts at the median level for the control subjects, into low vs high gene expression (≤70 vs >70), we found that 75% of cases had high *HP* expression compared with 50% of control subjects ($P = 0.011$) Multivariable conditional logistic regression analysis, adjusted for age at diagnosis of primary cancer, sex, anthracycline exposure, and chest radiation, showed that childhood cancer survivors with *HP* RNA counts >70 (higher gene expression) had 6-fold higher odds of developing cardiomyopathy (OR: 6.4; 95% CI: 1.4-28.6; $P = 0.014$) compared with those with *HP* RNA counts ≤70 (lower gene expression).

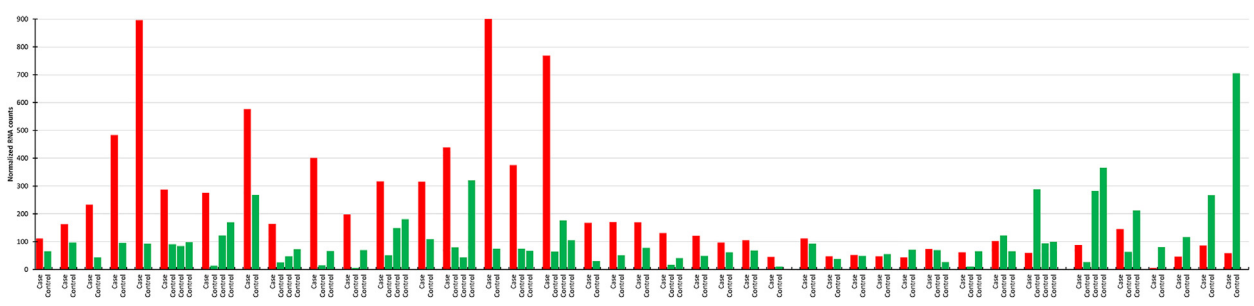
TABLE 1 Demographic and Clinical Characteristics by Case-Control Status

	Cases (n = 40)	Controls (n = 64)	*P* Value[a]
Age at primary cancer diagnosis, y	8.2 (3.6-13.9)	9.7 (3.3-14.4)	0.97
Sex			
Female	24 (60.0)	34 (53.1)	0.49
Cumulative anthracycline exposure			
≥250 mg/m^2	25 (62.5)	23 (35.9)	0.008
Chest radiation			
Yes	19 (47.5)	13 (20.3)	0.003
Race/ethnicity			
Non-Hispanic White	23 (57.5)	37 (57.8)	Matched
Hispanic	9 (22.5)	16 (25.0)	
Black/African American	5 (12.5)	7 (10.9)	
Asian	3 (7.5)	3 (4.7)	
Mixed race/ethnicity	0 (0.0)	1 (1.6)	
Primary diagnosis			
Acute lymphoblastic leukemia	9 (22.5)	16 (25.0)	Matched
Acute myeloid leukemia	2 (5.0)	3 (4.7)	
Ewing sarcoma	4 (10.0)	8 (12.5)	
Hodgkin lymphoma	7 (17.5)	10 (15.6)	
Kidney tumors	2 (5.0)	2 (3.1)	
Neuroblastoma	5 (12.5)	8 (12.5)	
Non-Hodgkin lymphoma	5 (12.5)	8 (12.5)	
Osteosarcoma	4 (10.0)	7 (10.9)	
Soft tissue sarcoma	2 (5.0)	2 (3.1)	
Time from diagnosis to cardiac event for cases or time to enrollment for control subjects, y	5.3 (0.8-12.8)	10.1 (7.1-14.5)	<0.001
HP RNA count	135.3 (63.7-300.2)	70.3 (47.5-105.8)	<0.001

Values are median (IQR) or n (%). [a]*P* values were estimated using either chi-square or Fisher exact tests for categorical variables or the Wilcoxon rank sum test for continuous variables.

HP = haptoglobin.

FIGURE 1 *HP* Gene Expression



The normalized aggregate read RNA counts of haptoglobin (*HP*) gene are shown from RNA sequencing in anthracycline-related cardiomyopathy. The matched case-control sets are plotted side by side. Matched cases are indicated in **red**, and control subjects are indicated in **green**.

JACC: CARDIOONCOLOGY, VOL. 5, NO. 3, 2023
JUNE 2023:392–401

TABLE 2 Genetic Determinants of *HP* RNA Expression in Childhood Cancer Survivors

***HP* RNA counts by *HP* genotype**				
***HP* RNA Counts**	**HP1-1 (n = 16)**	**HP1-2 (n = 77)**	**HP2-2 (n = 5)**	***P* Value[a]**
Median (IQR)	91.0 (24.0-155.4)	84.2 (55.8-166.4)	73.4 (48.7-483.5)	0.86
≤70, n (%)	7 (43.7)	31 (40.3)	2 (40.0)	0.96
>70, n (%)	9 (56.2)	46 (59.7)	3 (60.0)	
***HP* allele-specific transcript expression**				
Cumulative FPKM	***HP1* (n = 69)**	***HP2* (n = 63)**	**NA**	
Median (IQR)	0 (0-3.4)	3.6 (1.6-7.3)		
***HP* RNA counts by additional genetic determinants of *HP* gene expression**				
	rs2000999			
***HP* RNA Counts**	**A/A (n = 1)**	**G/A (n = 27)**	**G/G (n = 70)**	***P* Value[a]**
Median (IQR)	25.7 (25.7-25.7)	63.8 (36.5-108.3)	95.4 (63.2-196.6)	
≤70, n (%)	1 (100)	16 (59.3)	23 (32.9)	0.008
>70, n (%)	0	11 (40.7)	47 (67.1)	0.015
	rs35283911			
***HP* RNA counts**	**-/- (n = 1)**	**G/- (n = 27)**	**G/G (n = 70)**	***P* Value[a]**
Median (IQR)	25.7 (25.7-25.7)	68.8 (36.5-114.4)	94.5 (60.7-196.6)	
≤70, n (%)	1 (100)	15 (55.6)	24 (34.3)	0.033
>70, n (%)	0	12 (44.4)	46 (65.7)	0.049

[a]*P* values were estimated using either chi-square or Fisher exact tests for categorical variables or the Kruskal-Wallis test for continuous variables.

FPKM = fragments per kilobase of transcript per million mapped reads; *HP* = haptoglobin; NA = Not applicable.

HP GENOTYPE AND *HP* GENE EXPRESSION LEVELS.

HP gene exists as 2 CNVs (*HP1* and *HP2*), yielding 3 genotypes: HP1-1, HP1-2, and HP2-2.[19] Of the 104 participants in this study, we had genotype data for only 98 because of either unavailability of DNA (n = 4) or inconclusive PCR results (n = 2). *HP* genotypes showed the following distribution in 98 genotyped samples: 16.3% for HP1-1, 78.5% for HP1-2, and 5% for HP2-2. HP^{del} and NM_001126102.1:c.190 + 1 (splice donor mutation) were not found. The median RNA counts did not differ significantly among the 3 genotypes (HP1-1, 91.0; HP1-2, 84.2; and HP2-2, 73.4; $P = 0.86$) (Table 2). However, to gain insight into *HP1* and *HP2* allele-specific transcripts, we assessed 40 cases matched 1:1 with control subjects (n = 80) and finally analyzed (n = 74) those with genotyping data (Supplemental Results, Supplemental Table 3). FPKM represented transcript abundance in the paired-end RNA-seq data. *HP2* allele had higher transcript levels compared with *HP1* allele (median FPKM level 3.6 [IQR: 1.6-7.3] vs 0 [IQR: 0-3.4]), regardless of case-control status (Table 2). Furthermore, as shown in Table 3, cases had higher median *HP2* cumulative FPKM transcript levels compared with control subjects (median 5.1 [IQR: 2.1-11.8] vs 2.9 [IQR: 1.4-4.4]; $P = 0.027$).

TABLE 3 *HP* Allele-Specific Transcript Abundance in Cases and Control Subjects

	Cases	**Control Subjects**	***P* Value[a]**
HP2 allele (4 transcripts in HP1-2 and HP2-2)			
Cumulative FPKM	(n = 31)	(n = 32)	
Median (IQR)	5.1 (2.1-11.8)	2.9 (1.4-4.4)	0.027
HP1 allele (2 transcripts in HP1-1 and HP1-2)			
Cumulative FPKM	(n = 35)	(n = 34)	
Median (IQR)	3.9 (1.2-7.9)	2.2 (0-4.3)	0.11

[a]*P* values were estimated using the Wilcoxon rank sum test for continuous variables.

Abbreviations as in Table 2.

ADDITIONAL GENETIC DETERMINANTS OF *HP* GENE EXPRESSION LEVELS.

SNVs rs35283911 and rs2000999 are known to be associated with *HP* gene expression (independent of *HP* CNV status).[20-22] Genotyping assays for SNVs rs2000999, G>A, rs35283911 Del, G>-, and NM_001126102.1:c.190 + 1 SNV, G>C were designed using Fluidigm D3 assay design (Supplemental Figure 1). As shown in Table 2, the major allele (G) for both rs35283911 and rs2000999 was associated with higher *HP* expression. rs35283911 [G] was associated with higher *HP* gene expression (median RNA count G/G, 94.5; G/-, 68.8; and -/-, 25.7; $P = 0.033$). Similarly, rs2000999 [G] was also associated with higher *HP* expression (median RNA count G/G, 95.4; G/A, 63.8; and A/A, 25.7; $P = 0.008$).

GENETIC DETERMINANTS OF *HP* GENE EXPRESSION AND RISK FOR ANTHRACYCLINE-RELATED CARDIOMYOPATHY.

The *HP* genotype (model A) and the 2 SNVs taken together (rs35283911 and rs2000999) (model B) demonstrated 3.2-fold and 1.8-fold higher odds of cardiomyopathy, respectively, but these associations did not reach statistical significance (Table 4). We then examined the combined effect of HP genotype and the 2 SNVs, (rs35283911 and rs2000999), by creating a combined high-risk and low-risk group. The high-risk group consisted of patients with *HP2* allele (HP1-2 or HP2-2) who also had rs35283911 [G/G] or rs2000999 [G/G] genotype; the low-risk group included all others (model C). The high-risk group showed significantly higher *HP* expression (median RNA count 96.1 vs 68.8; $P = 0.006$). Multivariable conditional logistic regression analysis adjusted for age at diagnosis of primary cancer, sex, and anthracycline and chest radiation showed that patients in the high-risk group were 4 times more likely to develop cardiomyopathy (OR: 3.9; 95% CI: 1.0-14.5; $P = 0.045$) compared with those in the low-risk group (Table 4). The unadjusted conditional logistic regression analysis containing only the genetic

variables did not show a statistically significant association.

DISCUSSION

We show that childhood cancer survivors with anthracycline-related cardiomyopathy have significantly higher levels of whole blood *HP* gene expression compared with survivors without cardiomyopathy. Those with *HP* RNA counts >70 have 6-fold higher adjusted odds of developing cardiomyopathy. Cancer survivors with cardiomyopathy have higher *HP2* cumulative transcript levels compared with control subjects. Finally, patients with the HP1-2 or HP2-2 genotype and rs35283911 [G/G] and rs2000999 [G/G] genotypes are more likely to have higher *HP* gene expression and 3.9 times greater odds of developing cardiomyopathy, compared with those with HP1 allele or rs35283911 [delG].

Cell-free hemoglobin (HB) promotes accumulation of hydroxyl radicals and reactive oxygen species.[23] HP is a mammalian plasma protein that binds to free HB to form the HP-HB complex that is phagocytosed by monocyte-macrophages by binding to receptor CD163,[24-26] thereby preventing oxidative damage from free heme iron.[27] The *HP* locus is polymorphic in humans. The homozygous allelic deletion for *HP* (HP^{del}) is also known to occur, albeit at a very low frequency.[28] An intragenic duplication of exons 3 and 4 of the ancestral *HP* gene produced *HP2*,[29] yielding 3 genotypes: HP1-1, HP1-2, and HP2-2.[19] *HP1* and *HP2* differ in their ability to clear HB and differ functionally in their ability to protect from HB-induced oxidative stress.[30] *HP1* carriers have a higher HP plasma protein concentration compared with *HP2* carriers.[21,24,27,31-33] HP1-1 linear polymer has the highest affinity for cell-free HB and thereby superior antioxidant capacity compared with HP2-2-ring polymer,[33-36] by inducing a rapid clearance of HB from circulation, indicating that the potential for tissue damage is highest for individuals with the HP2-2 genotype and lowest for those with the HP1-1 genotype. Furthermore, CD163-mediated endocytosis into the macrophages is altered in patients with the HP2-2 genotype.

Preclinical models support this observation. The amount of myocardial injury after radiation was Hp genotype dependent in mice containing the Hp2 allele (generated by introducing the human *HP2* allele, as the HP2 allele does not exist in mice).[37] Although the Hp2 mice demonstrated increased production of several lipid peroxidation products in the myocardium after radiation, the Hp1 mice demonstrated increased production of the anti-inflammatory and antioxidant cytokine interleukin-10. Pharmacologic administration of an antioxidant reduced myocardial infarct size in this model. In childhood cancer survivors, the association between anthracycline exposure and cardiomyopathy is potentiated by chest radiation. In the present study, 47.5% of survivors with cardiomyopathy had received chest radiation compared with 20% of survivors without. Unfortunately, we could not evaluate the sole effect of radiation to the heart, as all patients were anthracycline exposed. HP CNVs (HP1-1, HP1-2, and HP2-2) have been identified as prognostic biomarkers of vascular disease, myocardial infarction, hypertension, heart failure, and angiogenesis in nononcology populations.[38] Free HB causes vascular injury as it intercalates into cell membranes. Once inside the cell, heme is broken down to release redox-active iron, which oxidizes low-density lipoproteins, promoting foam cell formation and vascular endothelial cell apoptosis and exacerbating plaque instability, resulting in plaque rupture. Following intraplaque hemorrhage, large amounts of cell-free HB accumulate within the atheroma because of impaired clearance by the HP-CD163 scavenging system. HP binds free HB, thus preventing this cascade of events. The potential for plaque instability is accentuated in patients with the HP2-2 genotype.[39-41] The presence of the HP2-2 genotype and low plasma levels of HP (<1.4 g/L) were associated with detrimental outcomes following acute myocardial infarction.[42] Increased intraplaque iron deposition in patients with the HP2-2 genotype may be responsible

TABLE 4 Unadjusted and Adjusted Models for Genetic Determinants of *HP* Gene Expression and Anthracycline-Related Cardiomyopathy

	Unadjusted[a]		Adjusted[b]	
	OR (95% CI)	*P* Value	OR (95% CI)	*P* Value
HP CNV status (model A)				
HP1-1	Reference	0.54	Reference	0.17
HP1-2 or HP2-2	1.5 (0.42-5.2)		3.2 (0.61-17.0)	
Combined *HP* SNV status (model B)				
rs2000999 [A/A] or [G/A] and rs35283911 [-/-] or [G/-]	Reference	0.36	Reference	0.33
rs2000999 [G/G] or rs35283911 [G/G]	1.6 (0.58-4.5)		1.8 (0.5-6.2)	
Composite *HP* CNV and SNV status (model C)				
HP1-1 and rs2000999 [A/A] or [G/A] and rs35283911 [-/-] or [G/-]	Reference	0.14	Reference	0.045
HP1-2 or HP2-2 and rs2000999 [G/G] or rs35283911 [G/G]	2.0 (0.79-5.1)		3.9 (1.0-14.5)	

[a]Unadjusted conditional logistic regression analysis for univariate genetic predictors in models A, B, and C. [b]Each conditional logistic regression model (A, B, and C) was adjusted for age at primary cancer diagnosis, sex, anthracycline dose, and chest radiation.

CNV = copy number variant; *HP* = haptoglobin; SNV = single-nucleotide variant.

for increased oxidative stress and instability of the carotid plaque.[41]

These observations in the nononcology population align with our observations that cancer survivors with cardiomyopathy have higher *HP2* cumulative transcript levels compared with control subjects. There is evidence to suggest that iron plays a role in anthracycline-related cardiomyopathy.[43] Heme iron catalyzes lipid peroxidation, which is a hallmark of ferroptosis.[44] Iron-dependent ferroptosis has been shown to have a role in anthracycline-induced cardiomyopathy.[45] Inhibition of ferroptosis by ferrostatin-1 significantly reduces cardiomyopathy risk. The cardioprotective properties of dexrazoxane, a compound that chelates iron, supports this hypothesis.[46] We propose that the alteration of iron homeostasis by anthracyclines and the associated heart damage due to iron accumulation[43,47,48] are further exacerbated by ineffective clearance of HB because of the presence of the *HP2* allele (**Central Illustration**), thus predisposing the patient to anthracycline-induced cardiomyopathy.

Determining whether dysregulated gene expression is the direct result of the outcome (cardiomyopathy) or plays a role in the pathogenesis can be difficult. SNVs such as rs35283911, rs2000999, and NM_001126102.1:c.190 + 1G>C are associated with *HP* gene expression.[20] rs35283911 is a single-base-pair deletion (delG, intron variant in *TXNL4B*) upstream of *HP*, occurs in the background of *HP2*, and influences *HP* RNA and protein levels.[21] rs2000999 (intron variant in *TXNL4B* and *HPR*) is downstream of the *HP* gene, and its effect is likely mediated via linkage disequilibrium with rs35283911.[22] Up to 45% of the variance in serum HP is explained by rs2000999.[22] In our study of childhood cancer survivors, we found that the major allele (G) for both rs35283911 and rs2000999 was associated with higher *HP* expression. A previous study showed that rs2000999 and *HP* CNVs influence HP levels independently.[49] Our findings are consistent with these observation; patients with the HP1-2 or HP2-2 genotype and with the rs35283911 [G/G] and rs2000999 [G/G] genotypes have higher *HP* gene expression and 3.9 times greater odds of developing cardiomyopathy. The association between the SNVs or CNVs and HP gene expression and cardiomyopathy provides us with evidence to state that initial exposure to anthracyclines triggers heart damage because of iron accumulation and that this is compounded by ineffective clearance of cell-free HB because of the presence of the *HP2* allele, thus predisposing the patient to anthracycline-induced cardiomyopathy.

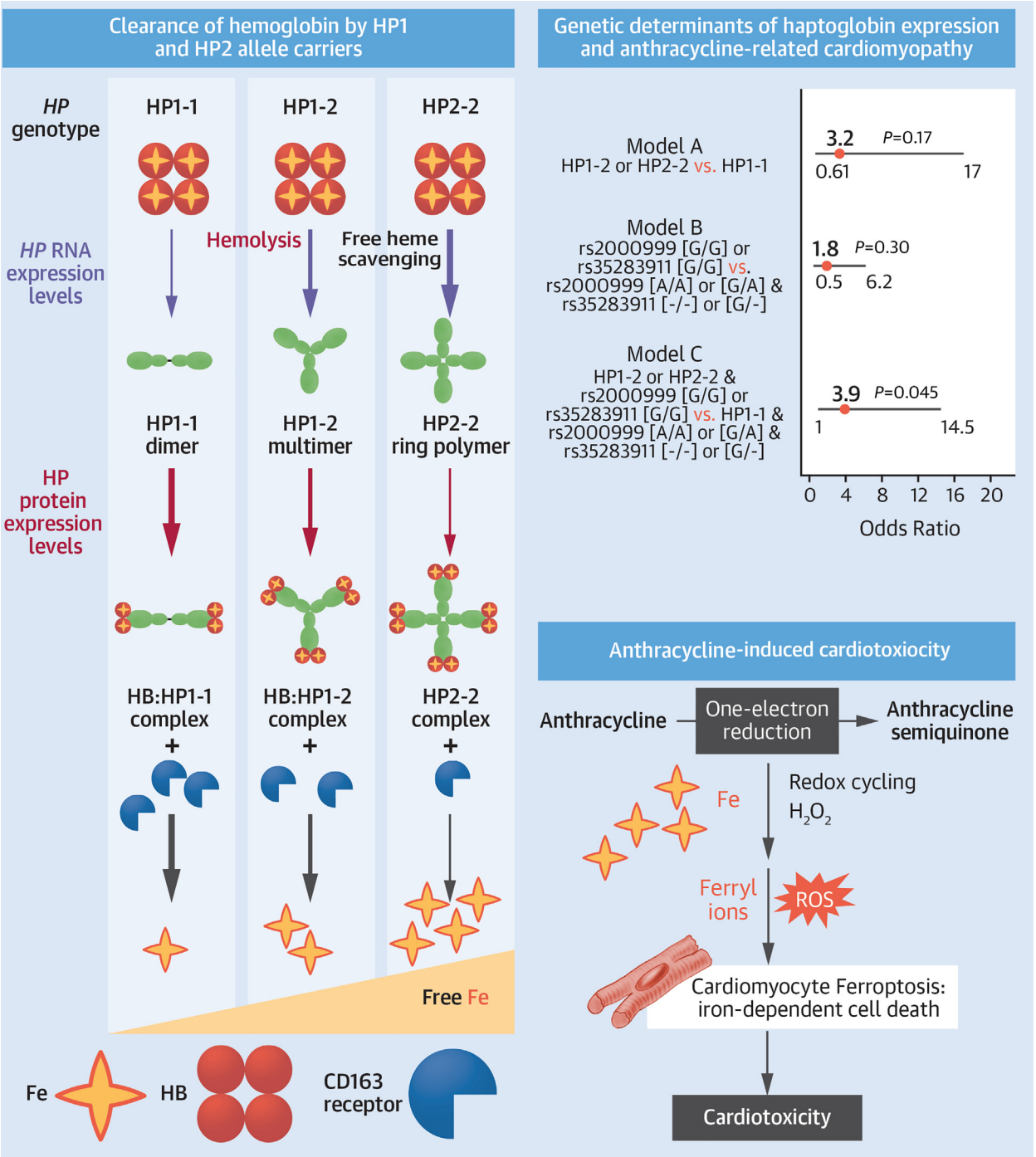
STUDY LIMITATIONS. Ideally, gene expression should be measured in the affected tissue (ie, cardiac tissue). HP binds to cell-free HB and prevents iron-mediated formation of free reactive oxygen species that cause myocardial injury and vascular damage. Thus, the ideal experiment would be to examine the extent of iron overload and ferroptosis in the myocardium from patients with high *HP* expression. However such experiments were not logistically feasible. Finally, radiation dose to the heart was not available to us; instead, we relied on receipt of radiation to the chest (with the heart in the field) as a yes-or-no variable.

The present study is the first to identify an association between *HP* RNA expression in peripheral blood in the context of genetic determinants of *HP* expression and the subsequent risk for anthracycline-related cardiomyopathy in childhood cancer survivors. Because of the structure of HP1 and HP2 CNVs, it is difficult to identify them using classic genomewide association studies, presenting the need to conduct studies such as ours to identify novel variants using differential gene expression analysis. It is important to note that the *HP* CNV was genotyped using a gold-standard PCR approach and not determined by genotype imputation. The biologic plausibility of the association between *HP* and anthracycline-related cardiomyopathy, as well as the previous demonstration of the association between HP and cardiometabolic traits that potentiate the risk for cardiomyopathy provide credence to this association in childhood cancer survivors with anthracycline-related cardiomyopathy.

CONCLUSIONS

The present study demonstrates that *HP* gene expression in blood is associated with anthracycline-related cardiomyopathy in childhood cancer survivors and that key genetic determinants contribute to *HP* expression and the risk for anthracycline-related cardiomyopathy. These findings provide evidence for the role of genetic variants associated with *HP* RNA levels in anthracycline-related cardiac dysfunction as well as a possible mechanistic explanation of the role of oxidative damage due to anthracycline exposure. Upon validation of these observations in an independent population, these findings could result in a

CENTRAL ILLUSTRATION **HP and Anthracycline-Related Cardiotoxicity**



Singh P, et al. J Am Coll Cardiol CardioOnc. 2023;5(3):392–401.

The figure shows differences in the clearance of hemoglobin (HB) by *HP1* and *HP2* carriers. On the left is the efficient clearance of HB by CD163 scavenger receptor in HP1-1 carriers, whereas HB:HP2-2 is not effectively cleared by CD163 receptor because of lower levels of CD163 and inefficient endocytosis. The inefficient clearance of HB-HP2-2 leads to increased iron and lipid peroxidation that when compounded by iron overload from anthracycline exposure renders the HP2 allele carriers more susceptible to ferroptosis and cardiomyopathy. *HP* = haptoglobin; ROS = reactive oxygen species.

blood-based screening test with limited invasiveness and high acceptance. Furthermore, patients carrying the HP2-2 genotype could derive more benefit from dexrazoxane, given that it binds free iron and/or removes iron from the doxorubicin-iron complex, thereby preventing oxygen free radical formation.

FUNDING SUPPORT AND AUTHOR DISCLOSURES

This research is supported by the National Cancer Institute (R35CA220502; principal investigator [PI], S. Bhatia), Leukemia and Lymphoma Society (6563-19; PI, S. Bhatia), and the V Foundation (DT2019-010; PI, S. Bhatia). The Children's Oncology Group study (COG-ALTE03N1 [NCT00082745]; PI, S. Bhatia) reported here is supported by the National Clinical Trials Network Operations Center Grant (U10CA180886; PI, D.S. Hawkins), the National Clinical Trials Network Statistics & Data Center Grant (U10CA180899; PI, Alonzo), the Children's Oncology Group Chair's Grant (U10CA098543; PI, Adamson), the Children's Oncology Group Statistics & Data Center Grant (U10CA098413; PI, Anderson), the National Cancer Institute Community Oncology Research Program Grant (UG1CA189955; PI, Pollock), and the Community Clinical Oncology Program Grant (U10CA095861; PI, Pollock), and the St. Baldrick's Foundation through an unrestricted grant. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors have reported that they have no relationships relevant to the contents of this paper to disclose.

ADDRESS FOR CORRESPONDENCE: Dr Smita Bhatia, University of Alabama Birmingham, 1600 7th Avenue South, Lowder 500, Birmingham, Alabama 35233, USA. E-mail: smitabhatia@uabmc.edu.

PERSPECTIVES

COMPETENCY IN PATIENT CARE AND PROCEDURAL SKILLS: *HP* gene expression in peripheral blood is associated with anthracycline-related cardiomyopathy in childhood cancer survivors, and key genetic determinants contribute to *HP* expression and the risk for cardiomyopathy.

TRANSLATIONAL OUTLOOK: These findings provide a possible mechanistic explanation of the role of oxidative damage due to anthracycline exposure and iron-dependent ferroptosis in anthracycline-related cardiomyopathy. The ability to demonstrate these findings using peripheral blood could result in a blood-based screening test with limited invasiveness and high acceptance.

REFERENCES

1. Armstrong GT, Chen Y, Yasui Y, et al. Reduction in late mortality among 5-year survivors of childhood cancer. *N Engl J Med*. 2016;374:833–842.

2. Bhatia S. Genetics of anthracycline cardiomyopathy in cancer survivors: *JACC: CardioOncology* state-of-the-art review. *J Am Coll Cardiol CardioOnc*. 2020;2:539–552.

3. 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;34:863–870.

4. Magdy T, Jouni M, Kuo HH, et al. Identification of drug transporter genomic variants and inhibitors that protect against doxorubicin-induced cardiotoxicity. *Circulation*. 2022;145:279–294.

5. Leong SL, Chaiyakunapruk N, Lee SW. Candidate gene association studies of anthracycline-induced cardiotoxicity: a systematic review and meta-analysis. *Sci Rep*. 2017;7:39.

6. Umans BD, Battle A, Gilad Y. Where are the disease-associated eQTLs? *Trends Genet*. 2021;37:109–124.

7. Yao DW, O'Connor LJ, Price AL, Gusev A. Quantifying genetic effects on disease mediated by assayed gene expression levels. *Nat Genet*. 2020;52:626–633.

8. Verlouw JAM, Clemens E, de Vries JH, et al. A comparison of genotyping arrays. *Eur J Hum Genet*. 2021;29:1611–1624.

9. Singh P, Wang X, Hageman L, et al. Association of GSTM1 null variant with anthracycline-related cardiomyopathy after childhood cancer—a Children's Oncology Group ALTE03N1 report. *Cancer*. 2020;126:4051–4058.

10. Wang Z, Gerstein M, Snyder M. RNA-seq: a revolutionary tool for transcriptomics. *Nat Rev Genet*. 2009;10:57–63.

11. Zhao W, He X, Hoadley KA, Parker JS, Hayes DN, Perou CM. Comparison of RNA-seq by poly (A) capture, ribosomal RNA depletion, and DNA microarray for expression profiling. *BMC Genomics*. 2014;15:419.

12. Feijen EAM, Leisenring WM, Stratton KL, et al. Derivation of anthracycline and anthraquinone equivalence ratios to doxorubicin for late-onset cardiotoxicity. *JAMA Oncol*. 2019;5:864–871.

13. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29:15–21.

14. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015;31:166–169.

15. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15:550.

16. Trapnell C, Roberts A, Goff L, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc*. 2012;7:562–578.

17. Koch W, Latz W, Eichinger M, et al. Genotyping of the common haptoglobin Hp 1/2 polymorphism based on PCR. *Clin Chem*. 2002;48:1377–1382.

18. Cupaioli FA, Mosca E, Magri C, et al. Assessment of haptoglobin alleles in autism spectrum disorders. *Sci Rep*. 2020;10:7758.

19. Maeda N, Yang F, Barnett DR, Bowman BH, Smithies O. Duplication within the haptoglobin Hp2 gene. *Nature*. 1984;309:131–135.

20. Kazmi N, Koda Y, Ndiaye NC, et al. Genetic determinants of circulating haptoglobin concentration. *Clin Chim Acta*. 2019;494:138–142.

21. Bjornsson E, Helgason H, Halldorsson G, et al. A rare splice donor mutation in the haptoglobin gene associates with blood lipid levels and coronary artery disease. *Hum Mol Genet*. 2017;26:2364–2376.

22. Froguel P, Ndiaye NC, Bonnefond A, et al. A genome-wide association study identifies rs2000999 as a strong genetic determinant of circulating haptoglobin levels. *PLoS ONE*. 2012;7:e32327.

23. Sadrzadeh SM, Graf E, Panter SS, Hallaway PE, Eaton JW. Hemoglobin. A biologic fenton reagent. *J Biol Chem*. 1984;259:14354–14356.

24. Andersen CBF, Stodkilde K, Saederup KL, et al. Haptoglobin. *Antioxid Redox Signal*. 2017;26:814–831.

25. Thomsen JH, Etzerodt A, Svendsen P, Moestrup SK. The haptoglobin-CD163-heme oxygenase-1 pathway for hemoglobin scavenging. *Oxid Med Cell Longev*. 2013;2013:523652.

26. Sadrzadeh SM, Bozorgmehr J. Haptoglobin phenotypes in health and disorders. *Am J Clin Pathol*. 2004;121(suppl):S97–S104.

27. Langlois MR, Delanghe JR. Biological and clinical significance of haptoglobin polymorphism in humans. *Clin Chem*. 1996;42:1589–1600.

28. Soejima M, Agusa T, Iwata H, et al. Haptoglobin genotyping of Vietnamese: global distribution of HP del, complete deletion allele of the HP gene. *Leg Med (Tokyo)*. 2015;17:14–16.

29. Boettger LM, Salem RM, Handsaker RE, et al. Recurring exon deletions in the HP (haptoglobin) gene contribute to lower blood cholesterol levels. *Nat Genet*. 2016;48:359–366.

30. Bulters D, Gaastra B, Zolnourian A, et al. Haemoglobin scavenging in intracranial bleeding: biology and clinical implications. *Nat Rev Neurol*. 2018;14:416–432.

31. Kasvosve I, Gomo ZA, Gangaidzo IT, et al. Reference range of serum haptoglobin is haptoglobin phenotype-dependent in blacks. *Clin Chim Acta*. 2000;296:163–170.

32. Carter K, Worwood M. Haptoglobin: a review of the major allele frequencies worldwide and their association with diseases. *Int J Lab Hematol*. 2007;29:92–110.

33. Naryzny SN, Legina OK. Haptoglobin as a biomarker. *Biochem Mosc Suppl B Biomed Chem*. 2021;15:184–198.

34. Melamed-Frank M, Lache O, Enav BI, et al. Structure-function analysis of the antioxidant properties of haptoglobin. *Blood*. 2001;98:3693–3698.

35. Van Vlierberghe H, Langlois M, Delanghe J. Haptoglobin polymorphisms and iron homeostasis in health and in disease. *Clin Chim Acta*. 2004;345:35–42.

36. Buehler PW, Abraham B, Vallelian F, et al. Haptoglobin preserves the CD163 hemoglobin scavenger pathway by shielding hemoglobin from peroxidative modification. *Blood*. 2009;113:2578–2586.

37. Blum S, Asaf R, Guetta J, et al. Haptoglobin genotype determines myocardial infarct size in diabetic mice. *J Am Coll Cardiol*. 2007;49:82–87.

38. Graves KL, Vigerust DJ. Hp: an inflammatory indicator in cardiovascular disease. *Future Cardiol*. 2016;12:471–481.

39. Purushothaman K-R, Purushothaman M, Levy AP, Sharma SK, Fuster V, Moreno PR. Neovascularization and intra-plaque hemorrhage: role of haptoglobin, macrophages, and heme-oxygenase-1 pathway. In: Slevin M, ed. *Therapeutic Angiogenesis for Vascular Diseases: Molecular Mechanisms and Targeted Clinical Approaches for the Treatment of Angiogenic Disease*. Dordrecht, the Netherlands: Springer Netherlands; 2011:237–256.

40. Purushothaman KR, Purushothaman M, Levy AP, et al. Increased expression of oxidation-specific epitopes and apoptosis are associated with haptoglobin genotype: possible implications for plaque progression in human atherosclerosis. *J Am Coll Cardiol*. 2012;60:112–119.

41. Lioupis C, Barbatis C, Drougou A, et al. Association of haptoglobin genotype and common cardiovascular risk factors with the amount of iron in atherosclerotic carotid plaques. *Atherosclerosis*. 2011;216:131–138.

42. Torabi A, Cleland JG, Khan NK, et al. The timing of development and subsequent clinical course of heart failure after a myocardial infarction. *Eur Heart J*. 2008;29:859–870.

43. Ichikawa Y, Ghanefar M, Bayeva M, et al. Cardiotoxicity of doxorubicin is mediated through mitochondrial iron accumulation. *J Clin Invest*. 2014;124:617–630.

44. Gutteridge JM. Lipid peroxidation and antioxidants as biomarkers of tissue damage. *Clin Chem*. 1995;41:1819–1828.

45. Fang X, Wang H, Han D, et al. Ferroptosis as a target for protection against cardiomyopathy. *Proc Natl Acad Sci U S A*. 2019;116:2672–2680.

46. Ganatra S, Nohria A, Shah S, et al. Upfront dexrazoxane for the reduction of anthracycline-induced cardiotoxicity in adults with preexisting cardiomyopathy and cancer: a consecutive case series. *Cardiooncology*. 2019;5:1.

47. Gammella E, Maccarinelli F, Buratti P, Recalcati S, Cairo G. The role of iron in anthracycline cardiotoxicity. *Front Pharmacol*. 2014;5:25.

48. Salazar-Mendiguchia J, Gonzalez-Costello J, Roca J, Ariza-Sole A, Manito N, Cequier A. Anthracycline-mediated cardiomyopathy: basic molecular knowledge for the cardiologist. *Arch Cardiol Mex*. 2014;84:218–223.

49. Soejima M, Sagata N, Komatsu N, et al. Genetic factors associated with serum haptoglobin level in a Japanese population. *Clin Chim Acta*. 2014;433:54–57.

KEY WORDS **allele-specific expression, anthracycline-related cardiomyopathy, childhood cancer survivors, ferroptosis, gene expression, haptoglobin, iron overload**

APPENDIX **For supplemental results, tables, and a figure, please see the online version of this paper.**