Childhood Cancer Survivor Study
Analysis Concept

Title: Analysis of Genetic Variants from Candidate Anthracycline-responsive Genes in iPSC-Cardiomyocytes as Predictors of Cardiotoxicity

Working Groups: Genetics
Chronic Diseases

Investigators:

Monica E. Reyes, Ph.D.	melinan@mdanderson.org
Lindsay Morton	mortonli@mail.nih.gov
Smita Bahtia, M.D., Ph.D.	sbhatia@peds.uab.edu
Kevin Oeffinger, M.D.	kevin.oeffinger@duke.edu
Gregory T. Armstrong, M.D., M.S.	greg.armstrong@stjude.org
Wendy Leisenring, Sc.D.	wleisenr@fredhutch.org
Michelle A.T. Hildebrandt, Ph.D.	mhildebr@mdanderson.org
Yuanquing Ye, Ph.D.	yqye@mdanderson.org

Background and Rationale:

Treatment-related cardiac death is one of the leading non-cancer causes of mortality in survivors of childhood cancers[1]. Anthracyclines, one of the most commonly used chemotherapy drugs for childhood malignancies, are known to have cardio toxic effects that can increase with cumulative dose[2]. With an increase in the number of studies being performed for risk assessment of anthracycline-induced cardiotoxicity (ACT), several genetic variants associated with cardiotoxicity are surfacing[3,4]. However, our current understanding of how these mechanisms contribute to the risk of developing ACT is limited. Further knowledge of these mechanisms relies on using a suitable experimental model such as human induced pluripotent stem cell (iPSC)-derived cardiomyocytes. These cells display many characteristics of human myocardium[5]. Using this cardiomyocyte model, which is the best model for recapitulating drug response in human cardiomyocytes, we performed RNA sequencing (RNAseq) to investigate altered gene expression following treatment with doxorubicin at various doses and time points. From the RNAseq data, we identified genes significantly differentially expressed following doxorubicin exposure. These

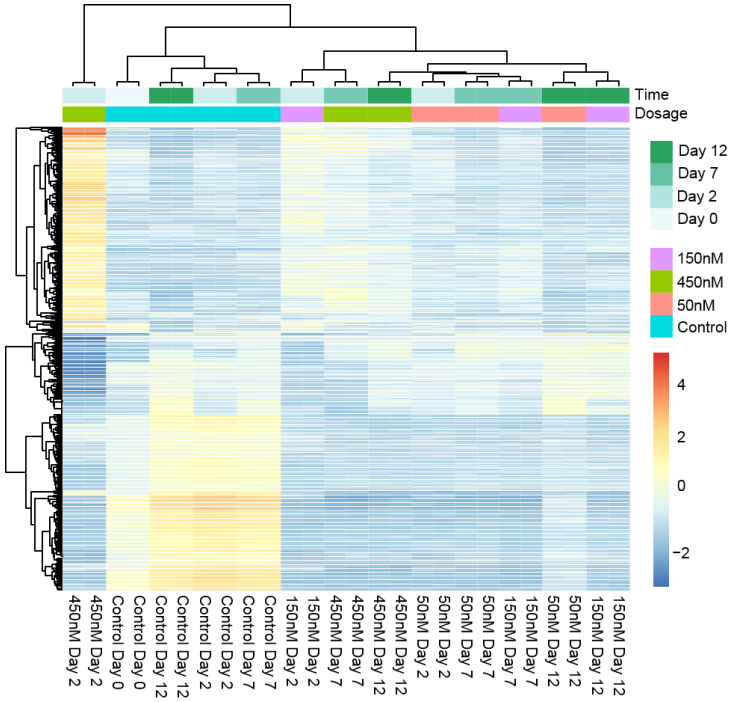

Figure 1. A heat map of the 500 transcripts with significant global variation in expression between control and doxorubicin treated samples.

genes have roles in DNA damage response and cell cycle regulatory pathways. Specifically, genes with significant changes in expression with increase in dose at the earliest time point (day 2) were found to be involved in the regulation of chromosomal replication and mitotic spindle formation[6].

Significant genes identified in RNAseq data from doxorubicin-treated iPSC-cardiomyocytes

From an RNAseq pilot study, we identified transcripts differentially expressed between untreated and doxorubicin-treated iPSC-cardiomyocytes. From 18,424 transcripts, a list of 220 transcripts (216 down-regulated and 4 up-regulated) which encodes for 212 protein-coding genes, were found to have the most significant fold-change in expression (log2 fold change >1; P<0.05) between control and doxorubicin-treated conditions (Figure 1). Gene Ontology (GO) terms in DESeq2[7], a bioinformatics package, were obtained for the complete data set (N=18,424), and transcripts were filtered for the top significant GO Term Accession ID (GO:0000278) for mitotic cell cycle. We found that 101 genes identified for this GO Term overlapped with the 220 transcripts. Following global variation analysis, we identified a set of eight genes with functions in pathways relevant to DNA damage response, chromosomal replication, and cell cycle progression that were found to be significantly differentially expressed between control and treated conditions on day 2 and day 7.

Overlapping genes and pathways between our findings and a study using patient-derived iPSC-cardiomyocytes treated with doxorubicin

In a study conducted by Burridge et al, several genes were found to be significantly downregulated, following a 24-hour exposure to doxorubicin in culture, in hESC-derived cardiomyocytes or iPSC-cardiomyocytes derived from patients with doxorubicin-induced cardiotoxicity[8]. For their study, iPSC-derived cardiomyocytes were treated with a range of doxorubicin doses (0.1-10 uM) for 24 hours, which is within the range of physiological serum concentrations for clinical doses. The exposure time and concentrations used in our study, a 48-hour exposure at 0.05-0.45 uM doxorubicin, are comparable to the conditions used in the Burridge et al. study, and in addition, we have included recovery time points days 7 and 12. In hESC-derived cardiomyocytes, they highlighted differentially expressed genes, following 1 uM or 10 uM doxorubicin exposure, have functions in p53 signaling, pro-apoptosis, cardiac hypertrophy and renal necrosis/cell death, cell cycle: G1/S checkpoint regulation, TGF-β signaling, and RAR activation pathways. Altered expression of genes encoding for transcription factors were recapitulated in doxorubicin –treated hiPSC-cardiomyocytes, including *RELA*, *NFKB1*, *RARA* and *STAT3*. Comparing untreated DOXTOX (derived from patients with cardiotoxicity) hiPSC-cardiomyocytes to 1 uM doxorubicin treated DOXTOX hiPSC-cardiomyocytes revealed increased expression of genes with functions in programmed cell death, p53 pathway, and transcription factors, such as *TP53*, *BRCA*, *PALB2*, *STAT3*, *CEBPA*, and *PPARG*. Similar to their study, in which a significant downregulation of *BRCA1* and *MYBL2* was observed following doxorubicin exposure, we found both *BRCA1* (log2 fold-change= -1.51; P=0.01529) and *MYBL2* (log2 fold-change= -3.43; P=5.73×10^{-11}) to be significantly downregulated in doxorubicin-treated iPSC-cardiomyocytes. The similarity between our study and the study by Burridge et al. supports the use of non-patient derived iPSC-cardiomyocytes as a model to investigate anthracycline response in cardiomyocytes.

We hypothesize that variants identified in the genes identified with differential expression following anthracycline exposure will be significantly associated with risk of cardiotoxicity in long term childhood cancer survivors and that there are variants predictive of risk stratified by total cumulative dose. Of the 212 protein-coding genes, 148 genes were identified with genetic variation (MAF>0.05) within the exonic and 5'UTR regions using 1000 Genomes Ensembl GRCh38 browser. Approximately 600 variants in total have been identified at a global minor allele frequency (MAF) cutoff of 0.05, this includes about 200

variants within regulatory regions that were identified using Ensembl Variant Effect Predictor[9]. In our approach, we will determine whether variants within these genes, differentially expressed following doxorubicin exposure in iPSC-cardiomyocytes, are predictors of cardiotoxicity. Furthermore, variants associated with this risk will be assessed for risk stratified by cumulative anthracycline doses to identify variants associated with those at highest risk.

Primary Aims:

1. **In 148 genes identified from the analysis of iPSC-cardiomyocyte RNAseq data, we will identify variants associated with risk of cardiotoxicity.**

2. **Identify genetic variants in genes modulated in iPSC-cardiomyocytes during doxorubicin exposure that can predict dose-dependent associations with risk of cardiotoxicity.**

Analysis Framework:

For the study analysis, we will use genotyped data from the CCSS cohort. The analysis plan for each aim is described in detail below along with a description of the study population and a list of the variables of interest. In collaboration with CCSS statisticians and members of the proposed working groups, we will develop a final plan for the methods.

Outcomes of interest: The primary outcome of interest is cardiotoxicity defined by cardiomyopathy/heart failure with CTCAE grades 3 (severe, heart failure medication), 4 (life-threatening, heart transplantation), or 5 (fatal), which is described previously in the methods used by Feijen et al.[10], in those treated with anthracyclines.

Study population: The study population will be derived from the CCSS original cohort of individuals of European ancestry who are childhood cancer survivors, diagnosed from 1970-1986 (N=5,324), and that were previously genotyped for the CCSS GWAS study, with imputation of the data based on the 1000 Genomes Project. For this study, we will only select individuals from this population who received anthracyclines. Cases will be defined as individuals who received anthracyclines and diagnosed with cardiomyopathy/CHF (i.e. CTCAE grades 3-5 heart failure).

Exploratory variables: Primary variables of interest include genotypes of individuals with/without cardiotoxicity. From the imputed GWAS dataset, we are interested in genetic variants within protein-coding, 5'UTR regions, and regulatory regions of the 148 genes identified as differentially expressed in doxorubicin-treated iPSC-cardiomyocytes from our RNAseq data; restricted to MAF>0.05.

Additional covariates to be considered in the analysis include:

- Gender
- Diagnosis year
- Age at diagnosis
- Year of last treatment
- Age at last follow-up
- Cancer type
- Co-morbidities (e.g., hypertension, diabetes, obesity)
- Date of last follow-up
- Length of follow-up (date of diagnosis to last follow-up)

- Anthracycline cumulative dose (updated anthracycline calculated dosage)
- Chest radiation therapy (Y/N)

Analytical Approach: From the CCSS individuals (N=5,324), we will only select for individuals who were previously treated with anthracyclines. Those diagnosed with a congenital heart abnormality or diagnosed with a secondary tumor will be excluded from our study population. From this population we will identify cardiotoxicity cases as those with cardiomyopathy/CHF (CTCAE grades 3-5) we will compare the characteristics of those with cardiotoxicity to those without cardiotoxicity (Table 1).

For Aim 1, our primary focus will be to investigate variants (MAF>0.05) within the 148 genes (listed in Appendix A) and cardiotoxicity to determine associations with the risk of cardiotoxicity. To do this, we will calculate hazard ratios (HR), 95% confidence interval (CI), and a p-value for each association between variant and cardiotoxicity (Y/N) using an a Cox proportional hazard regression model. For each variant, a statistical significance will be defined as less than the Bonferroni test $p<8.33\times10^{-5}$ (0.05/600)

The focus of Aim 2 will be to investigate the relationship between total cumulative anthracycline doses and variants in Aim 1 that were found to be significantly associated with an increased risk of cardiotoxicity. High total cumulative dose is a known risk factor for anthracycline-induced cardiotoxicity. However, cardiotoxicity can also develop at lower total cumulative doses (<100 mg/m^2). Therefore, investigating the associations between genetic variants and stratified cumulative doses, we can identify which genes and pathways may predict high risk of anthracycline-induced cardiotoxicity in long-term childhood cancer survivors. Briefly, to identify the genetic variants that may contribute to different levels of sensitivity to anthracycline-induced cardiotoxicity, we will stratify individuals who received anthracyclines by total cumulative dose as follows: <250 mg/m^2 and ≥250 mg/m^2.

Power: We are sufficiently powered for the proposed statistical analyses. For example, restricting to only those patients who were exposed to anthracyclines (N=1,912) and anticipating 137 cases, we have 80% power to detect an odds ratio of 2.04 for association with cardiotoxicity for variants with a MAF of 10% at a P-value of 0.05. This power increases with increasing MAF of the risk allele. As the selection of genes for inclusion in the analysis was hypothesis-driven based on anthracycline-response in the iPSC-cardiomyocyte, our ability to detect a true association is further enhanced.

Special Considerations:

Replication: From a cohort of long-term childhood cancer survivors at MD Anderson Cancer Center (N=700), we will identify survivors who received anthracyclines for inclusion in a replication analysis. We will perform replication genotyping for the genetic variants that were found to be significantly associated with cardiotoxicity risk in Aim 1 and Aim 2. These genetic variants will be genotyped on the TaqMan Genotyping Assay and ABI 7900HT Sequence Detection System following the standard protocol.

Functional Validation: We are continuing to collect phenotypic data of anthracycline-response in the iPSC-cardiomyocytes. This includes measurements for changes in cardiac function (contractility, beating rate, and amplitude), cell viability, and mitochondrial function. Observed treatment-induced alterations in gene expression at each time point and dose will be assessed for correlations with these phenotypes. We also have the ability to conduct knockout/knockdown experiments in the cell lines to further define the role of candidate genes on anthracycline-response.

Table 1. Characteristics of study population			
	Cases (with cardiotoxicity) N (%)	Controls (without cardiotoxicity) N (%)	P value
Total			
Gender			
Male			
Female			
Cancer type			
Age at diagnosis (yr)			
Age at last follow-up (yr)			
Mean time since diagnosis (yr)			
Anthracycline treatment			
Doxorubicin			
Daunorubicin			
Epirubicin			
Idarubicin			
Anthracycline cumulative dose (mg/m^2)			
<250			
≥250			
Chest RT			
Yes			
No			

References

1 Armstrong, G. T. *et al.* Comprehensive Echocardiographic Detection of Treatment-Related Cardiac Dysfunction in Adult Survivors of Childhood Cancer: Results From the St. Jude Lifetime Cohort Study. *J Am Coll Cardiol* **65**, 2511-2522, doi:10.1016/j.jacc.2015.04.013 (2015).

2 Harake, D., Franco, V. I., Henkel, J. M., Miller, T. L. & Lipshultz, S. E. Cardiotoxicity in childhood cancer survivors: strategies for prevention and management. *Future Cardiol* **8**, 647-670, doi:10.2217/fca.12.44 (2012).

3 Krajinovic, M. *et al.* Polymorphisms of ABCC5 and NOS3 genes influence doxorubicin cardiotoxicity in survivors of childhood acute lymphoblastic leukemia. *Pharmacogenomics J* **17**, 107, doi:10.1038/tpj.2016.86 (2017).

4 Wojnowski, L. *et al.* NAD(P)H oxidase and multidrug resistance protein genetic polymorphisms are associated with doxorubicin-induced cardiotoxicity. *Circulation* **112**, 3754-3762, doi:10.1161/CIRCULATIONAHA.105.576850 (2005).

5 Maillet, A., Tan, K. P. & Brunham, L. R. Use of Human Pluripotent Stem Cell Derived-Cardiomyocytes to Study Drug-Induced Cardiotoxicity. *Curr Protoc Toxicol* **73**, 22 25 21-22 25 22, doi:10.1002/cptx.30 (2017).

6 Reyes, M. E. *et al.* RNA sequence analysis of inducible pluripotent stem cell-derived cardiomyocytes reveals altered expression of DNA damage and cell cycle genes in response to doxorubicin. *Toxicol Appl Pharmacol* **356**, 44-53, doi:10.1016/j.taap.2018.07.020 (2018).

7 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550, doi:10.1186/s13059-014-0550-8 (2014).

8 Burridge, P. W. *et al.* Human induced pluripotent stem cell-derived cardiomyocytes recapitulate the predilection of breast cancer patients to doxorubicin-induced cardiotoxicity. *Nat Med* **22**, 547-556, doi:10.1038/nm.4087 (2016).

9 McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *Genome Biol* **17**, 122, doi:10.1186/s13059-016-0974-4 (2016).

10 Feijen, E. A. *et al.* Equivalence Ratio for Daunorubicin to Doxorubicin in Relation to Late Heart Failure in Survivors of Childhood Cancer. *J Clin Oncol* **33**, 3774-3780, doi:10.1200/JCO.2015.61.5187 (2015).

Appendix A

Gene	Log2 Fold Change	Adjusted P value[a]	Gene	Log2 Fold Change	Adjusted P value[a]	Gene	Log2 Fold Change	Adjusted P value[a]
ANLN	-2.86	1.73E-39	***DEPDC1***	-3.91	1.84E-26	***KIF4A***	-4.14	7.68E-32
ASF1B	-3.29	3.68E-13	***DLGAP5***	-3.83	1.55E-15	***KNTC1***	-2.03	8.09E-09
ASPM	-4.17	8.10E-29	***DSCC1***	-2.98	3.04E-12	***KPNA2***	-1.52	0.00096
ATP8B3	-2.78	9.83E-14	***DTL***	-3.18	8.71E-10	***MCM10***	-3.18	2.19E-09
AURKB	-4.23	4.04E-32	***ECT2***	-1.89	1.04E-08	***MCM2***	-1.91	0.00041
BORA	-2.31	1.78E-09	***EME1***	-2.10	5.04E-06	***MCM5***	-2.53	1.52E-07
BRCA1	-1.51	0.01529	***ESCO2***	-3.60	5.80E-21	***MCM6***	-1.90	3.59E-06
BRCA2	-3.21	1.46E-18	***EXO1***	-3.17	1.41E-18	***MCM7***	-1.88	1.71E-05
BRIP1	-3.49	2.61E-16	***FAM111B***	-2.60	0.00018	***MELK***	-3.70	1.30E-25
BUB1	-2.67	9.26E-17	***FAM72D***	-2.58	1.44E-06	***MKI67***	-4.06	7.29E-19
BUB1B	-2.62	1.74E-10	***FANCA***	-2.04	2.22E-08	***MND1***	-2.99	2.31E-12
C21orf58	-2.03	2.41E-07	***FANCD2***	-2.89	1.82E-18	***MXD3***	-2.41	3.25E-06
C11orf82	-2.38	6.46E-15	***FANCG***	-1.95	1.38E-05	***MYBL2***	-3.43	5.73E-11
C5orf34	-2.07	6.18E-05	***FANCI***	-3.12	1.87E-30	***NCAPG***	-4.18	1.74E-29
CASC5	-3.58	3.58E-32	***FBXO5***	-1.74	0.00077	***NCAPH***	-3.92	1.00E-23
CCNA2	-4.17	2.28E-25	***FOXM1***	-4.13	1.05E-27	***NDC80***	-3.70	1.05E-14
CCNB1	-3.15	7.35E-29	***GGH***	-1.81	1.42E-05	***NUF2***	-3.86	9.46E-25
CCNB2	-3.93	9.72E-29	***GINS1***	-1.87	2.72E-08	***NUSAP1***	-4.18	1.84E-26
CCNF	-1.85	5.71E-08	***GINS2***	-2.16	5.00E-07	***PARPBP***	-2.72	5.05E-18
CDC20	-3.21	2.53E-15	***GPSM2***	-2.39	8.36E-18	***PBK***	-3.13	3.58E-08
CDC25B	-1.64	0.00392	***GSG2***	-3.22	5.45E-21	***PHF19***	-2.22	3.44E-11
CDC25C	-3.71	1.06E-17	***GTSE1***	-3.86	2.28E-25	***PKN3***	-1.87	1.30E-05
CDC45	-3.05	4.07E-08	***HELLS***	-2.48	3.42E-11	***PLK1***	-4.14	2.77E-39
CDC6	-3.42	7.34E-18	***HIST1H1B***	-4.36	3.44E-29	***PLK4***	-2.59	8.72E-08
CDC7	-2.04	1.95E-08	***HIST1H1D***	-3.00	1.31E-16	***POLA2***	-1.65	5.42E-05
CDCA2	-3.93	8.58E-30	***HIST1H2AB***	-3.04	5.70E-18	***POLE***	-1.56	0.0483
CDCA5	-3.46	1.96E-23	***HIST1H2AL***	-3.69	2.72E-25	***POLE2***	-2.56	5.74E-11
CDCA8	-4.19	2.11E-29	***HIST1H2BE***	-1.80	0.00297	***POLQ***	-3.16	4.54E-27
CDK1	-3.72	1.07E-15	***HIST1H2BL***	-2.45	7.93E-12	***PRC1***	-3.35	1.74E-29
CDKN3	-3.23	2.19E-22	***HIST1H3I***	-2.13	1.00E-04	***PRIM1***	-2.10	4.98E-05
CDT1	-2.48	1.69E-09	***HIST1H3J***	-3.17	3.47E-11	***PRR11***	-3.51	9.06E-35
CENPE	-3.85	6.47E-31	***HISTH4C***	-1.90	6.05E-05	***PSRC1***	-3.44	2.56E-21
CENPF	-4.20	1.02E-28	***HJURP***	-3.85	2.58E-16	***RACGAP1***	-3.15	4.63E-30
CENPI	-2.70	1.18E-10	***HMMR***	-3.20	2.82E-28	***RAD51AP1***	-3.11	6.46E-14
CENPK	-3.27	1.66E-17	***INCENP***	-1.51	0.02973	***RBL1***	-1.57	0.00802
CENPM	-3.70	2.82E-19	***IQGAP3***	-3.30	4.96E-10	***RMI2***	-1.95	0.00014
CENPU	-2.68	1.13E-13	***KIAA1524***	-1.88	5.92E-08	***RRM2***	-3.94	9.18E-17
CENPW	-1.97	0.0003	***KIF14***	-4.18	2.00E-29	***RTKN2***	-2.25	5.72E-05
CEP128	-2.04	6.67E-14	***KIF15***	-2.66	2.59E-10	***SAPCD2***	-3.21	5.87E-15
CEP55	-3.67	2.97E-18	***KIF18A***	-2.84	1.63E-11	***SGOL1***	-3.45	1.31E-15
CHAF1A	-2.22	2.15E-09	***KIF20A***	-3.85	1.32E-18	***SGOL2***	-2.45	4.05E-13
CHEK1	-1.57	0.0074	***KIF20B***	-1.95	3.60E-06	***SKA3***	-3.53	1.33E-24
CIT	-3.97	5.03E-24	***KIF23***	-2.92	4.47E-23	***SMC4***	-1.82	3.44E-07

CKAP2L	-4.09	2.52E-19	***KIF24***	-2.09	3.84E-09	***SPC24***	-3.52	2.33E-20
CLSPN	-3.79	6.50E-21	***KIF2C***	-3.68	8.58E-30	***STIL***	-2.35	2.49E-15