Title: Frequency of pathogenic germline variation in pediatric pan-cancer survivors: A report from the Childhood Cancer Survivor Study (CCSS)

Authors: Matthew Gianferante[1,*], Jung Kim[1,*], Danielle Karyadi[1], Steven Hartley[1], Megan Frone[1], Leslie L. Robison[2], Gregory T. Armstrong[2], Smita Bhatia[3], Michael Dean[1,4], Meredith Yeager[4], Bin Zhu[1], Lei Song[1], Seth A. Brodie[4], Kelvin de Andrade[1], Fernanda Forte[5], Alisa Goldstein[1], Payal P. Khincha[1], Mitchell J. Machiela[1], Mary Lou McMaster[1], Michael L. Nickerson[1], Leatrisse Oba[1], Alexander Pemov[1], Maisa Pinheiro[1], Melissa Rotunno[1], Karina Santiago[5], Talia Wegman-Ostrosky[6], Belynda D. Hicks[4], Bin Zhu[4], Mingyi Wang[4], Kristine Jones[4], Amy A. Hutchinson4, Casey Dagnall[4], Margaret Tucker[1], Lindsay Morton[1], Stephen J. Chanock[1], Sharon A. Savage[1], Douglas R. Stewart[1,*], Lisa Mirabello[1,*]

[1] Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA;
[2] Department of Epidemiology and Cancer Control, St. Jude Children's Research Hospital, Memphis, TN;
[3] Institute for Cancer Outcomes and Survivorship, University of Alabama at Birmingham, Birmingham, AL;
[4] Cancer Genomics Research Laboratory, Frederick National Laboratory for Cancer Research, Frederick, MD 20877, USA;
[5]International Research Center, A.C. Camargo Cancer Center, São Paulo, Brazil;
[6]División de Investigación, Instituto Nacional de Cancerología, CDMX 14080, Mexico

Introduction. Pediatric cancer is the leading cause of death by disease in children despite improved survival rates. Here, we conducted the largest pediatric pan-cancer study to date using exome sequencing (ES) of 5,451 long-term survivors to quantify the frequency of germline pathogenic variation in cancer-susceptibility genes (CSG).

Methods. The CCSS is a multi-center retrospective cohort of children diagnosed <21 years of age who survived at least 5 years from a diagnosis of leukemia, lymphoma, brain tumors, neuroblastoma, Wilms tumor, bone cancers, and soft tissue sarcomas. Germline DNA underwent ES; analyses focused on rare variants in 237 previously published CSGs. Pathogenic/likely pathogenic (P/LP) variants were identified using ClinVar, HGMD with manual review, and InterVar. We compared the frequency of P/LP variants with the same pathogenicity criteria in 5,105 European-ancestry cases (CEU >0.8) vs. 51,377 gnomAD non-Finnish European cancer-free controls (ES data) using the Fisher's exact test and corrected for multiple testing using a false discovery rate (FDR; q<0.05 considered significant). An unadjusted Kaplan-Meier analysis was used to estimate the differences in overall survival in children with and without P/LP variants.

Results. In 5,105 European-ancestry pediatric cancer survivors, 11% harbored a P/LP variant in a dominant CSG (n=176) vs. 9% in controls (P<0.0001). Eight dominant genes (*NF1*, *WT1*, *TSC1*, *REST*, *KMT2D*, *EZH2*, *CDKN2A*, *MEN1*) had significantly more P/LP variants in cases vs. controls (FDR q<0.05). We identified a novel germline cancer risk association in *KMT2D* (odds ratio [OR] 16.8, 95% CI 4.5-63.5, vs. controls). *KMT2D*, underlying Kabuki Syndrome, has only previously been associated somatically with cancer. In *EZH2* (OR 4.7), *CDKN2A* (OR 8.6), and *MEN1* (OR 20.1), we found multiple novel pediatric cancer associations. Children carrying a

P/LP variant had worse survival compared to children without a P/LP variant (P=0.001). Children with one P/LP variant in *SBDS*, an autosomal recessive CSG, had a significantly increased risk of cancer (FDR q<0.05).

Conclusion. To our knowledge, this is the largest pediatric pan-cancer study of pathogenic germline variation to date. In long-term cancer survivors, we found multiple novel gene-cancer associations and worse survival in children with a P/LP variant. These findings could have implications in cancer risk stratification and genetic counseling for patients and families.