

## **Genetic Influence on Pediatric Acute Myeloid Leukemia Survival**

**Short title:** Genetic Influence on Pediatric AML Survival

**Authors:** Zhaoheng Li\*<sup>1</sup>, Pablo S. Monterroso\*<sup>2</sup>, Lauren J. Mills<sup>2</sup>, Erin L. Marcotte<sup>2,3</sup>

**Affiliations:** <sup>1</sup>Department of Biostatistics, University of Washington, Seattle, WA, USA;

<sup>2</sup>Division of Epidemiology and Clinical Research, Department of Pediatrics, University of Minnesota, Minneapolis, MN, USA; <sup>3</sup>Masonic Cancer Center, University of Minnesota, Minneapolis, MN, USA

\*These authors contributed equally to this work.

**Corresponding author:** Erin L. Marcotte, PhD, Department of Pediatrics, Division of Epidemiology & Clinical Research, MMC 715, 420 Delaware St. S.E., Minneapolis, MN 55455; Phone: +1 6126263281; Email: [marcotte@umn.edu](mailto:marcotte@umn.edu)

**Data access and data sharing:** The RNA-seq and microarray data that support the findings of this study are available at <https://ocg.cancer.gov/programs/target>.

*Word count text: 2939 (max. 4000)*

*Word count abstract: 250 (max. 250)*

*Figure/Table count: 6/6*

*Number of references: 30*

## Abstract

Racial disparities in pediatric acute myeloid leukemia (AML) outcomes are well-documented, yet underlying factors, including potential genetic contributors, remain poorly understood. We conducted a retrospective cohort study using data from the Therapeutically Applicable Research to Generate Effective Treatments (TARGET) AML project to investigate genetic contributions to disparities. We analyzed data from 811 pediatric AML patients of non-Hispanic White, non-Hispanic Black, or Hispanic race/ethnicity with complete clinical and genomic data. We conducted differential gene expression and multivariate Cox regression analyses adjusted for race/ethnicity, gender, age at diagnosis, white blood cells at diagnosis, minimal residual disease, and cytogenetics to estimate hazard ratios (HR) for event-free survival (EFS) and overall survival (OS). Gene expression analysis identified 338 differentially expressed genes among racial/ethnic groups, with *ZHX2* upregulated among Hispanic patients compared to White counterparts (log<sub>2</sub> fold change: 1.074; adjusted p-value: 0.042), correlating with decreased hazard ratios for EFS (HR: 0.58, 95% CI: 0.45-0.76). Conversely, the lactase gene (*LCT*) was downregulated among Black patients (log<sub>2</sub> fold change: -2.47; adjusted p-value: 0.021), corresponding to increased hazard ratios for EFS (HR: 1.50, 95% CI: 1.21-1.87). In the microarray data, *KLRC1*, *CLUHP3*, and *PRC1* expression emerged as significant predictors for EFS among intermediate-risk patients, with high *PRC1* expression mitigating adverse survival effects among Black patients. Our findings underscore the complex interplay of genetic factors in pediatric AML survival disparities. Future research should utilize large, diverse datasets and explore non-coding genes to

further elucidate underlying mechanisms and inform risk stratification for equitable care and improved outcomes in pediatric AML.

**Key points:**

- Differential gene expression reveals distinct prognostic markers among Black and Hispanic pediatric AML patients.
- Genetic markers such as ZHX2 and LCT demonstrate significant associations with event-free survival.