

TO THE EDITOR:

Accurate and reliable detection of clonal hematopoiesis in plasma cell-free DNA

Alyssa C. Parker,^{1,2} Joseph C. Van Amburg,^{1,2} Yash Pershad,^{1,2} David A. Ong,² Ketan J. Hoey,² Christopher J. Farady,³ Philipp J. Rauch,³ Michael M. Mendelson,³ J. Brett Heimlich,⁴ P. Brent Ferrell,^{1,5} and Alexander G. Bick^{1,2}

¹Program in Human Genetics, Vanderbilt Genetics Institute, Vanderbilt University School of Medicine, Nashville, TN; ²Division of Genetic Medicine and Clinical Pharmacology, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN; ³Novartis Biomedical Research, Basel, Switzerland; and ⁴Division of Cardiovascular Medicine and ⁵Division of Hematology and Oncology, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN

In clonal hematopoiesis of indeterminate potential (CHIP), hematopoietic stem cells with variants in leukemia-associated genes gain a proliferative advantage.^{1,2} Clone size is quantified as the fraction of mutant reads compared with total reads (variant allele fraction [VAF]).³ Larger VAF increases risk of hematologic and inflammatory diseases,^{4,5} making longitudinal studies of clonal growth an area of interest. Currently, VAF is measured from DNA extracted from whole blood. However, most clinical trial protocols, epidemiologic studies, and research biorepositories do not store serial whole blood samples. Conversely, serial plasma samples are frequently available.

Because plasma cell-free DNA (cfDNA) largely originates from blood cells,^{6,7} we hypothesized that sequencing cfDNA could similarly capture CHIP. Researchers investigating the use of cfDNA to monitor solid tumor recurrence have reported identification of CHIP variants, suggesting that CHIP is detectable in cfDNA.⁸ Moreover, sequencing of whole blood DNA and cfDNA from patients with hematologic disorders with somatic mutations such as aplastic anemia and myelodysplastic syndrome demonstrated a strong correlation of VAF measurements between whole blood DNA and cfDNA.^{9,10} Researchers also found that CHIP detection was not well correlated between sample types, but their study design combined various driver genes, which we anticipate would introduce variability because differentiation biases differ by driver gene and blood cell types are not equally represented in cfDNA. We tested the reliability of CHIP detection in cfDNA as a proxy for whole blood DNA, stratified by driver gene. We hypothesized that origin–cell type proportions may explain differences in VAF between sample types.

To test this, we collected blood from 67 patients suspected to have CHIP, with serial samples from 15 patients. Participants were recruited through the Clonal Hematopoiesis and Inflammation in the Vasculature cohort at Vanderbilt University Medical Center.¹¹ Participants consented to study procedures under Vanderbilt University Medical Center's institutional review board (protocols: 210022 and 201583) in accordance with the Declaration of Helsinki. Blood was drawn into K2 EDTA tubes. On the day of sample collection, 1 mL of whole blood was reserved for genomic DNA extractions.¹² Remaining blood was centrifuged at 1900g for 10 minutes at 4°C to isolate plasma. The plasma layer was transferred, centrifuged at 16 000g for 10 minutes at 4°C, then stored at –80°C for up to 14 months. cfDNA was isolated from 4 mL aliquots of thawed plasma, processed with the nRichDX Revolution cfDNA Max 20 kit (4 mL protocol); elutes were returned to –80°C. All cfDNA preparations yielded ≥ 1.76 ng/ μ L. Targeted DNA sequencing with unique molecular identifiers was performed on whole blood DNA and cfDNA using the CHIP assay at a median sequencing depth of 1249 $\times$.³ Variants were reported at VAF ≥ 0.02 .

Submitted 30 May 2025; accepted 7 October 2025; prepublished online on *Blood Advances* First Edition 1 November 2025; final version published online 16 January 2026. <https://doi.org/10.1182/bloodadvances.2025017293>.

Data is deposited in a controlled access repository (<https://zenodo.org/records/17604891>). Code supporting analysis is available at https://github.com/bicklab/CHIP_cfDNA_methylation.

The full-text version of this article contains a data supplement.

© 2026 American Society of Hematology. Published by Elsevier Inc. Licensed under Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0), permitting only noncommercial, nonderivative use with attribution. All other rights reserved.

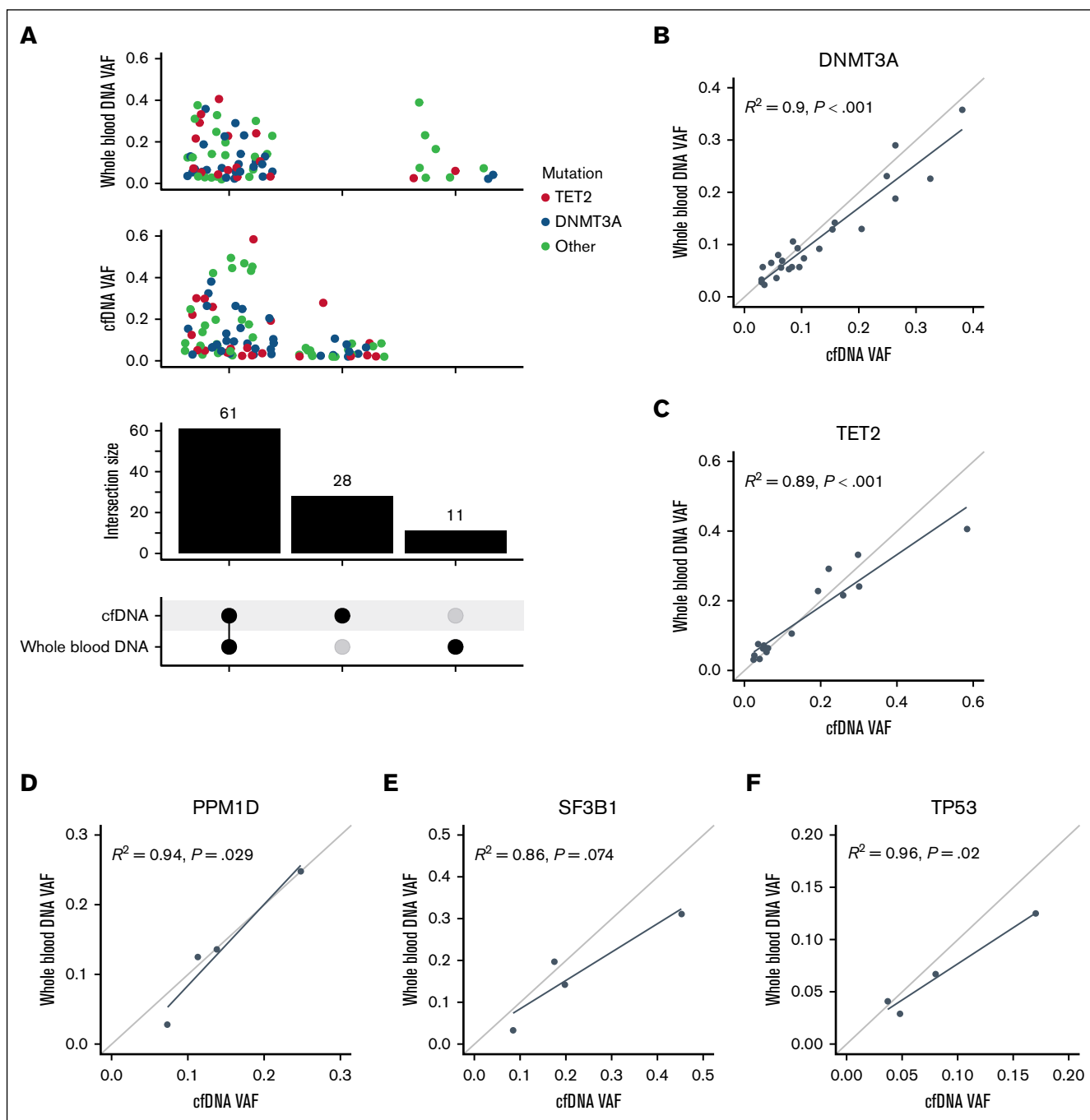


Figure 1. Variant detection and VAF measurements are similar in whole blood DNA and cfDNA. (A) Scatterplots and upset plot showing overlap of variants detected in whole blood DNA and cfDNA. (B-F) Scatterplots showing correlation of VAF measured in whole blood DNA compared with VAF measured in cfDNA, stratified by driver gene. Statistical significance determined with a Pearson correlation test.

We detected 100 unique variants among patients' initial time point samples (Figure 1A; supplemental Table 1). Of these, 61 were detected in both whole blood DNA and cfDNA, 28 were detected only in cfDNA, and 11 were detected only in whole blood DNA. If considering measurements from whole blood DNA as ground truth, this represents a positive predictive value for cfDNA of 68.5%. For low VAF (<0.1) variants in cfDNA, this represents a positive predictive value of 54.4%, and for high VAF 93.8%. Of the 28 variants detected only in

cfDNA, 26 were low VAF and 2 were high VAF. Four of the variants detected only in cfDNA were detected in whole blood DNA at later time points (supplemental Figure 1). We performed Pearson correlation tests in R (version 4.3.2)¹³ to assess the relationship between VAF in whole blood DNA and cfDNA, which showed high correlation for both *DNMT3A* (Figure 1B; $R^2 = 0.9$; $P < .001$) and *TET2* (Figure 1C; $R^2 = 0.89$; $P < .001$). Correlation was also high for *PPM1D*, *SF3B1*, and *TP53* but sample sizes were significantly smaller (Figure 1D-F).

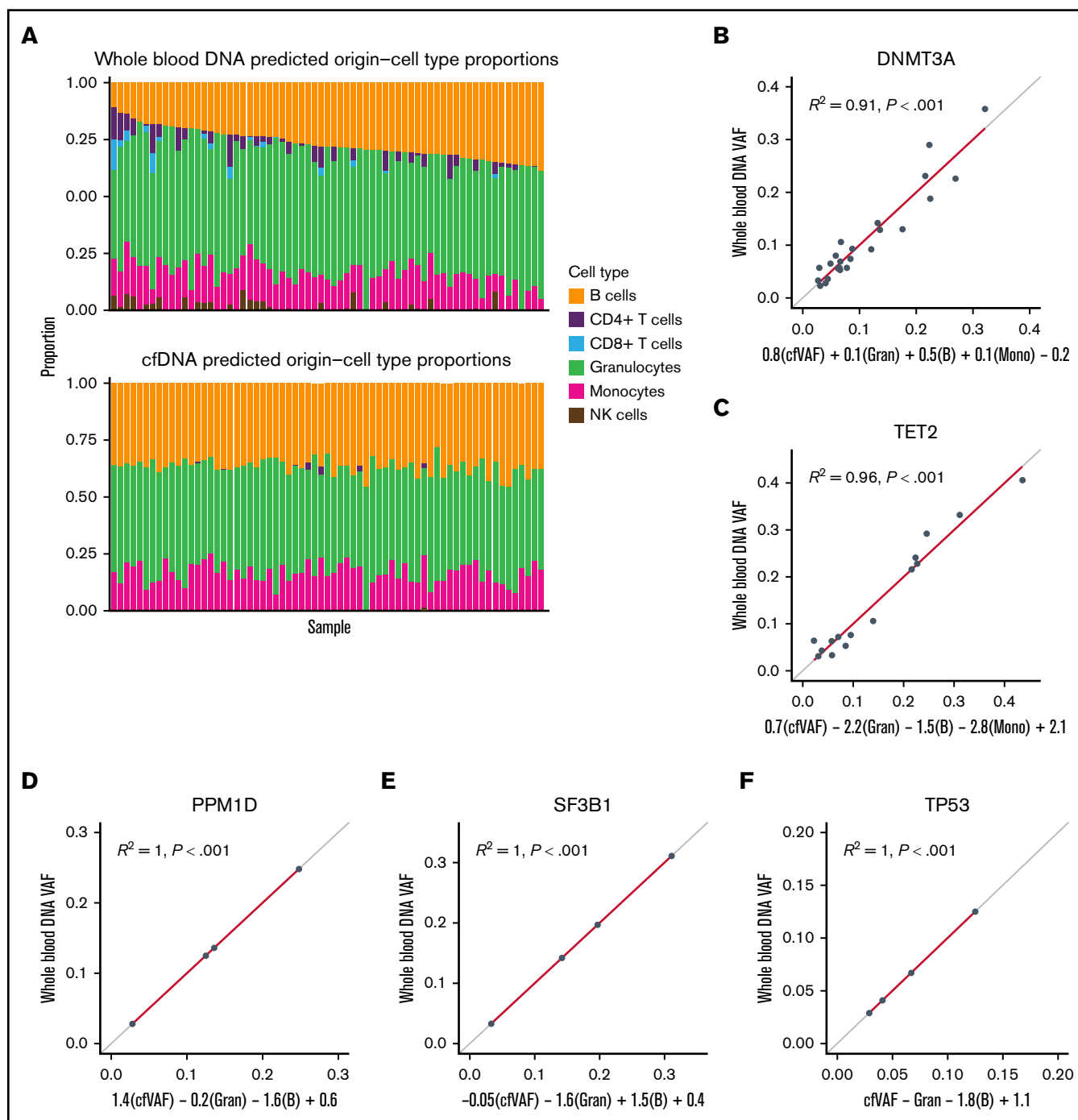


Figure 2. Linear modeling with cfDNA VAF and predicted origin-cell type proportions estimate whole blood DNA VAF with high accuracy. (A) Bar plots showing the predicted proportion of material originating from each major blood cell type as determined by the frequency of cell type-specific DNA methylation signatures observed in whole blood DNA and cfDNA. (B-F) Scatterplots showing predicted VAF compared to measured whole blood DNA VAF, stratified by driver gene. Predictions are based on measured cfDNA VAF and cfDNA origin-cell type proportions. NK, natural killer.

Because mutations in these genes are known to induce a myeloid bias, we hypothesized that the proportion of DNA originating from a particular cell type may influence the relationship between whole blood DNA VAF and cfDNA VAF. Hematopoietic cell types have unique DNA methylation signatures, enabling prediction of origin-cell type proportions from DNA methylation sequencing data. For

this reason, we performed New England Biolabs Enzymatic Methyl sequencing¹⁴ paired with a Twist Bioscience panel targeting 3.98 million CpGs for all samples. Sequencing was performed on an Illumina NovaSeq X platform at an average read depth of 30×, and DNA methylation was quantified with Bismark.¹⁵ We predicted origin-cell type proportions for whole blood DNA and cfDNA

based on the frequency of DNA methylation signatures using the R package methylCC.¹⁶ We compared origin–cell type proportions for initial time point samples (Figure 2A) and found that monocyte proportions were highly consistent between whole blood DNA and cfDNA (difference in means = 0.01, 2-sample *t* test *P* = .09). T cells and natural killer cells had very low representation in cfDNA, compared with whole blood DNA.

We sought to identify mathematical relationships between VAF in whole blood DNA, VAF in cfDNA, and cfDNA origin–cell type proportions. We performed linear modeling with the R function `lm()` with whole blood DNA VAF as the dependent variable, cfDNA VAF as the independent variable, and origin–cell type proportions of granulocytes, B cells, and monocytes as covariates. Resulting linear equations for each driver gene are listed in Figure 2B–F. We used these equations to predict whole blood DNA VAF from cfDNA values and found that essentially all differences in VAF between sample types were explainable by variation in cfDNA origin–cell type proportions for both *DNMT3A* (Figure 2B; $R^2 = 0.91$; $P < .001$) and *TET2* (Figure 2C; $R^2 = 0.96$; $P < .001$). Predictions for *PPM1D*, *SF3B1*, and *TP53* were perfectly correlated with whole blood DNA VAF measurements (Figure 2D–F; $R^2 = 1$; $P < .001$), likely because of overfitting as sample sizes were small. We performed Jackknife validation and found that the models performed consistently across iterations (supplemental Figure 2). Median R^2 values were 0.91 for *DNMT3A*, 0.96 for *TET2*, 1 for *PPM1D*, 1 for *SF3B1*, and 1 for *TP53*.

Our findings highlight the potential of cfDNA for longitudinal studies of CHIP and permit several conclusions. First, high VAF CHIP can be reliably detected in cfDNA. Despite the difference in DNA source material, detection of high VAF CHIP variants was comparable between whole blood DNA and cfDNA with most variants being detected in both sample types. We note that cfDNA reported several low VAF clones that were not detected in whole blood DNA. For this reason, we caution that researchers aiming to use cfDNA as a proxy for whole blood DNA should anticipate identification of additional low VAF clones that are not detectable in whole blood DNA.

Second, VAF can be reliably measured from cfDNA, potentially because of consistent representation of DNA from major blood cell types in cfDNA. Notably, the driver genes evaluated in this study bias hematopoiesis toward the myeloid lineage.^{17,18} For this reason, most cells carrying these variants are likely myeloid cells and the consistent representation of DNA from monocytes between sample types may be a driving factor for high correlation of VAF. Although linear modeling including origin–cell type proportions marginally improved prediction of whole blood VAF from cfDNA, we recommend that these corrections are unnecessary for most applications. Cell-type proportions may be a more important consideration if using cfDNA to measure CHIP driven by variants that bias differentiation toward other lineages.

Acknowledgments: The authors thank Brian H. Shirts for critical reading of the manuscript.

This work was supported by the National Institutes of Health (NIH) Early Independence Award grant DP5 OD029586, a Burroughs Wellcome Fund Career Award for Medical Scientists, a Pew-Stewart Scholar for Cancer Research award (supported by the Pew Charitable Trusts and the Alexander and Margaret Stewart Trust), and a Hevolution/American Federation for Aging Research

(AFAR) New Investigator Award in Aging Biology and Geroscience Research (A.G.B.). This work is also supported by funding from the Mark Foundation for Cancer Research Endeavor Award, Novartis Global Scholar Award, R56DK138826, and I01BX005991 (P.B.F.). A.C.P. and J.C.V.A. are supported by the NIH training grant T32GM145734-01, and Y.P. is supported by the NIH training grant T32GM007347.

Contribution: A.C.P. performed research, analyzed data, and wrote the paper; J.C.V.A. performed research and edited the paper; Y.P. performed research, analyzed data, and edited the paper; D.A.O. and K.J.H. performed research; and C.J.F., P.J.R., M.M.M., J.B.H., P.B.F., and A.G.B. designed the research and edited the paper.

Conflict-of-interest: C.J.F., P.J.R., and M.M.M. are employees of Novartis Biomedical Research. P.B.F. receives research funding from Novartis Biomedical Research. The remaining authors declare no competing financial interests.

ORCID profiles: A.C.P., 0000-0002-6632-7458; J.C.V.A., 0000-0002-1410-6995; Y.P., 0000-0002-2282-1403; D.A.O., 0009-0000-4260-6762; C.J.F., 0000-0002-3607-9721; P.J.R., 0000-0001-8677-0986; M.M.M., 0000-0001-7590-3958; J.B.H., 0000-0003-2812-5326; P.B.F., 0000-0003-1140-9154; A.G.B., 0000-0001-5824-9595.

Correspondence: Alexander G. Bick, Vanderbilt University Medical Center, 550 Robinson Research Building, 2200 Pierce Ave, Nashville, TN 37232; email: alexander.bick@vumc.org.

References

1. Steensma DP, Bejar R, Jaiswal S, et al. Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. *Blood*. 2015;126(1):9-16.
2. Jaiswal S, Fontanillas P, Flannick J, et al. Age-related clonal hematopoiesis associated with adverse outcomes. *N Engl J Med*. 2014;371(26):2488-2498.
3. Mack T, Vlasschaert C, von Beck K, et al. Cost-effective and scalable clonal hematopoiesis assay provides insight into clonal dynamics. *J Mol Diagn*. 2024;26(7):563-573.
4. Vlasschaert C, Mack T, Heimlich JB, et al. A practical approach to curate clonal hematopoiesis of indeterminate potential in human genetic data sets. *Blood*. 2023;141(18):2214-2223.
5. Yu Z, Fidler TP, Ruan Y, et al. Genetic modification of inflammation and clonal hematopoiesis-associated cardiovascular risk. *J Clin Invest*. 2023;133(18):e168597.
6. Sender R, Noor E, Milo R, Dor Y. What fraction of cellular DNA turnover becomes cfDNA? *eLife*. 2024;12:RP89321.
7. Moss J, Magenheimer J, Neiman D, et al. Comprehensive human cell-type methylation atlas reveals origins of circulating cell-free DNA in health and disease. *Nat Commun*. 2018;9(1):5068.
8. Fairchild L, Whalen J, D'Aco K, et al. Clonal hematopoiesis detection in patients with cancer using cell-free DNA sequencing. *Sci Transl Med*. 2023;15(689):eabm8729.
9. Gutierrez-Rodriguez F, Beerman I, Groarke EM, et al. Utility of plasma cell-free DNA for *de novo* detection and quantification

- of clonal hematopoiesis. *Haematologica*. 2022;107(8): 1815-1826.
10. Liu J, Chen X, Wang J, et al. Biological background of the genomic variations of cf-DNA in healthy individuals. *Ann Oncol*. 2019;30(3): 464-470.
 11. Shannon ML, Heimlich JB, Olson S, et al. Clonal hematopoiesis and inflammation in the vasculature: CHIVE, a prospective, longitudinal clonal hematopoiesis cohort and biorepository. *Blood Adv*. 2024; 8(13):3453-3463.
 12. Heimlich J, Shannon M, Olson S, et al. Clonal hematopoiesis and inflammation in the vasculature (CHIVE), a prospective, longitudinal clonal hematopoiesis cohort and biorepository. *Blood*. 2023; 142(suppl 1):5594.
 13. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing; 2023. Accessed 24 February 2025. <https://www.r-project.org>
 14. Vaisvila R, Ponnaluri VKC, Sun Z, et al. Enzymatic methyl sequencing detects DNA methylation at single-base resolution from picograms of DNA. *Genome Res*. 2021;31(7):1280-1289.
 15. Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics*. 2011;27(11): 1571-1572.
 16. Hicks SC, Irizarry RA. methylCC: technology-independent estimation of cell type composition using differentially methylated regions. *Genome Biol*. 2019;20(1):261.
 17. Izzo F, Lee SC, Poran A, et al. DNA methylation disruption reshapes the hematopoietic differentiation landscape. *Nat Genet*. 2020;52(4): 378-387.
 18. Heimlich JB, Bhat P, Parker AC, et al. Multiomic profiling of human clonal hematopoiesis reveals genotype and cell-specific inflammatory pathway activation. *Blood Adv*. 2024;8(14): 3665-3678.