Publication Committee Review Date: [**04/09/24**]
ARIC Manuscript Proposal Number: #[**4428**]

1.a. Full Title: Association between prior infections and immune profiles with clonal hematopoiesis in ARIC.

b. Abbreviated Title (Length 26 characters): Infections and CH

2. Writing Group [please provide a middle name if available; EX: Adam Lee Williams]:

Writing group members: Ajibike Dorothy Lapite, Meng Ru, James DeCuir, Apoorva Thatavarty, Temitope Olarinde, Christie Ballantyne, Kunihiro Matsushita, Koichi Takahashi, Junichi Ishigami, Katherine Y. King, and Elizabeth A. Platz

I, the first author, confirm that all the coauthors have given their approval for this manuscript proposal. [please confirm with your initials electronically or in writing] ***AL***

First author [please provide a middle name if available; EX: Adam Lee Williams]:
Ajibike Dorothy Lapite, MD MPHTM
Address: King Lab (attn: Ajibike Lapite)
1102 Bates Street Suite 1150
Houston, Texas 77030
Phone: 832-824-4222
E-mail: adlapite@texaschildrens.org

ARIC author to be contacted if there are questions about the manuscript and the first author does not respond or cannot be located (this must be an ARIC investigator).

Name: Elizabeth A. Platz, ScD, MPH
Address: Department of Epidemiology
Johns Hopkins Bloomberg School of Public Health
615 N. Wolfe St., Room E6132
Baltimore, MD 21205
Phone: 410-614-9674
E-mail: eplatz1@jhu.edu

3. Timeline: Expect to complete this work by February 2024.

4. Rationale:

Clonal hematopoiesis (CH) is the age-related expansion of hematopoietic stem cells (HSCs) that harbor somatic mutations, which confer survival advantage. CH is associated with increased risk for hematologic

malignancies in addition to cardiovascular disease and all-cause mortality.[1] Drivers for CH aside from aging include environmental conditions and health status. Inflammatory bowel disease and smoking[3] have both been demonstrated in clinical studies to be correlated with a higher prevalence of CH. Of particular interest is the growing evidence that infection is a driver for CH.

Our group has contributed to the work that describes the impact of infection on HSCs and CH. We demonstrated in a mouse model that chronic infection depletes HSCs.[4] In addition, we demonstrated that chronic mycobacterial infection is a driver of *Dnmt3a* mutant CH in a murine model.[5] Mechanistically, mycobacterial infection depletes wildtype stem cells through excessive terminal differentiation and impaired self-renewal, whereas *Dnmt3a*-mutant HSCs are resistant to these effects and therefore expand in this condition. Others have demonstrated that LPS or microbial gut translocation drives expansion of *Tet2*-mutant HSC clones in animal models.[6] In combination, these efforts make clear that infection is a driver of two of the most common mutant clones in CH - *Dnmt3a* and *Tet2*.

5. Main Hypothesis/Study Questions:

We hypothesize that CH is more likely to develop or expand in persons who have had major or repeated bouts of prior bacterial or viral infections as compared to persons who have not had documented prior infections (or few or minor). In this study, we will evaluate whether the number and type of documented prior infections and immune profiles are associated with CH.

Question 1. Among participants without CH at Visit 2 (or the first visit before Visit 5), are documented prior infections and immune profiles associated with CH metrics at Visit 5, including:

a. detected CH (yes/no)
b. large CH (ε10%) (large, small, no CH)
c. number (0, 1, 2+) of CH mutations
d. repeat of a. and b. for *DNMT3A* and *TET2*

Question 2. Among participants with detected CH at Visit 2 (or the first visit before Visit 5), are documented prior infection and immune profiles associated with change in CH metrics at Visit 5, including, possibly:

a. change in total VAF (e.g., no change, small to large CH, large to small CH)
b. change in number of CH mutations detected (no change, increase, decrease)
c. change in type of CH mutations detected (no change, gain, loss)
d. repeat of a. for *DNMT3A* and *TET2*

6. Design and analysis (study design, inclusion/exclusion, outcome, and other variables of interest with specific reference to the time of their collection, summary of data analysis, and any anticipated methodologic limitations or challenges if present).

Study design and cohort. We will conduct a closed, prospective cohort study among participants who have information available on CH status at Visit 2 (or other visits before Visit 5, if Visit 2 was not available; for the purpose of this proposal, the first visit with CH status will be referred to as Visit 2) and Visit 5. We will exclude participants whose ancestry is not Black or White, who have a history of a hematologic malignancy at Visit 5, and who did not provide full consent. We expect to include 4,187

participants based on Uddin et al. (PMID: 37732181), 3,730 without CH at Visit 2 (among whom 735 had incident CH at Visit 5) for Question 1 and 457 with CH at Visit 2 for Question 2.

Assessment of CH. The analysis will be conducted stratified by CH status (not detected, detected) at Visit 2 The outcome is CH metrics (CH detected, VAF, large CHIP (ε10%), number and type of mutations). CH was assessed at Visits 2 (Broad Institute; HiSeq 2000) and 5 (Baylor College of Medicine; NovaSeq 6000) by whole exome sequencing of DNA extracted from peripheral blood leukocytes coupled with Mutect2 software and pipeline (BioRxiv, 861054); details have been reported including concordance between the two platforms (83% for VAFε2%, and 93% for VAFε10% (Uddin et al. PMID: 37732181). CH was defined as ε2% variant allele frequency (VAF) in any of 69 driver genes for myeloid malignancy.

Assessment of infections. Hospital discharge summaries are routinely collected for ARIC participants. ARIC investigators have previously identified participants hospitalized for infections based on ICD-9 and ICD-10 codes and categorized these as respiratory, urinary tract, digestive tract, skin, blood or circulatory system (PMID: 36622673). We will include hospitalizations for infections that occur after Visit 1 and before Visit 5. Note that Visit 5, the second time point with CH status pre-dates COVID. We will classify participants by whether they have been hospitalized for infection, the number of hospitalizations for infection, whether they have been hospitalized for the above categories of organ system of infection.

At Visit 4, an inflammation interview was performed. We will also include the following infections self-reported in response to specific questions during that interview: 1) a recent or prior urinary tract infection or pyelonephritis, 2) recent or prior pneumonia, 3) recent or prior tuberculosis infection, 4) recent or prior course of antibiotics, 5) at least one prior episode of fever blister/cold sore on the lips, and/or 6) at least one episode of shingles/herpes zoster.

Assessment of immune profiles. Immune profiles (measured at Visits 2 and 3) will include classic cytokines (e.g., IFN- β, IFN- γ, IL-1β, IL-6, TNF- α), total antibodies (IgG, IgM, IgA), and C-reactive protein from the Somalogic panel.

Assessment of covariates. Covariates will be from Visit 2 and include age, race-center, education, smoking status (never, current, former) and packyears (for current and former smokers), body mass index (calculated from height and weight measured by trained technicians), diabetes status (diagnosed diabetes defined as self-reported physician diagnosis or used diabetes drugs, undiagnosed diabetes defined as those without diagnosed diabetes having fasting glucose ε126 mg/dL or non-fasting glucose ε200 mg/dL, at risk for diabetes defined as having fasting glucose 100 to <126 mg/dL or non- fasting glucose 140 to <200 mg/dL, and not diabetic or not at risk for diabetes), and dyslipidemia (as defined in PMID: 37732181 - total cholesterol ≥240, triglycerides ≥200, LDL- cholesterol ≥160, HDL-cholesterol <40 in men or <50 in women [all in mg/d]L, and/or taking a statin).

Statistical analysis. We will perform the analysis stratified by CH status at Visit 2 and evaluate the association of infections and immune profiles in relation to CH metrics using logistic regression (binary outcomes) or polytomous logistic regression (ordinal or nominal for category outcomes).

We will evaluate associations for:

- any hospitalization for infection
- number of hospitalizations for infection
- hospitalization for each category of organ system infection separately (see *Assessment of Infections*)
- any of the self-reported infections (see *Assessment of Infections*)
- the common self-reported infections separately
- immune profiles: cytokines, antibodies, and CRP RFUs (these measurements do not have clinical cutpoints; we will used quantile cutpoints)
- joint categories of hospitalization for infection and immune profiles

For Question 1, we will adjust for age, sex, and dyslipidemia, which were found to be associated with incident CH in Uddin et al. (PMID: 37732181). For Question 2, we will first assess whether any of the covariates are associated with change in CH metrics, and adjust for those.

Strengths.

- The availability of CH status at two time points; can evaluate incident CH and change in CH over time.
- Large sample size and number of incident CH for Question 1

Limitations.

- Small sample size for change in CH in participants with CH for question 2
- Two different platforms were used for the two visits, leading to reduced confidence in change in VAF
- For hospitalizations, we do not expect to be able to study specific infectious agents; however, our hypothesis is not specific to particular bacteria or viruses
- For the Visit 4 inflammation questionnaire, histories are self reported; no confirmation was done
- For some participants, the first visit with CH status is after Visit 2 (but before Visit 5). We will use covariates from the corresponding visit, but the time from their first visit to Visit 5 will be shorter, resulting is less time to observe CH or change in CH. Thus, we will conduct a sensitivity analysis restricting to participants whose first visit is Visit 2.

7.a. Will the data be used for non-ARIC analysis or by a for-profit organization in this manuscript? No

b. If Yes, is the author aware that the current derived consent file ICTDER05 must be used to exclude persons with a value RES_OTH and/or RES_DNA = "ARIC only" and/or "Not for Profit" ? (The file ICTDER has been distributed to ARIC PIs, and contains the responses to consent updates related to stored sample use for research.) N/A

8.a. Will the DNA data be used in this manuscript? Yes

b. If yes, is the author aware that either DNA data distributed by the Coordinating Center must be used, or the current derived consent file ICTDER05 must be used to exclude those with value RES_DNA = "No use/storage DNA"? Yes

9. The lead author of this manuscript proposal has reviewed the list of existing ARIC Study manuscript proposals and has found no overlap between this proposal and previously approved manuscript proposals either published or still in active status.
ARIC Investigators have access to the publications lists under the Study Members Area of the web site at: http://www.cscc.unc.edu/aricproposals/dtSearch.html
Yes

10. What are the most related manuscript proposals in ARIC (authors are encouraged to contact lead authors of these proposals for comments on the new proposal or collaboration)?

We are collaborating with Christie Ballantyne and his team on CH.

We will investigate a small number of candidate proteins involved in the immune response in the context of infections and CH. This manuscript proposal appears to be investigating each of the Somalogic panel proteins with CH. If there is overlap, we will investigate the joint association of the candidate proteins and infections. #4166 - Studying the Proteomic Predictors of Incidence and Expansion of Clonal Hematopoiesis of Indeterminate Potential (CHIP): The Atherosclerosis Risk in Communities (ARIC) Study

This study does not specifically include infection, but covers a wide array of possible exposures that may be related to CH. #4130 Studying the Clinical Conditions, Lifestyle Factors, and Laboratory Markers Associated with Incidence and Expansion of Clonal Hematopoiesis of Indeterminate Potential (CHIP): The Atherosclerosis Risk in Communities (ARIC) Study

11.a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data? Yes

11.b. If yes, is the proposal
X A. Primary the result of an ancillary study (list number)
1995.04
2011.07
2020.102022.24

B. primary based on ARIC data with ancillary data playing a minor role (usually control variables; list number(s)*
***ancillary studies are listed by number https://sites.cscc.unc.edu/aric/approved-ancillary-studies**

12a. Manuscript preparation is expected to be completed in one to three years. If a manuscript is not submitted for ARIC review at the end of the 3-years from the date of the approval, the manuscript proposal will expire.

12b. The NIH instituted a Public Access Policy in April, 2008 which ensures that the public has access to the published results of NIH funded research. It is **your responsibility to upload manuscripts to PubMed Central** whenever the journal does not and be in compliance with this policy. Four files about the public access policy from http://publicaccess.nih.gov/ are posted in http://www.cscc.unc.edu/aric/index.php, under Publications, Policies & Forms. http://publicaccess.nih.gov/submit_process_journals.htm shows

References

1. 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. doi:10.1056/nejmoa1408617

2. Zhang CRC, Nix D, Gregory M, et al. Inflammatory cytokines promote clonal hematopoiesis with specific mutations in ulcerative colitis patients. *Exp Hematol*. 2019;80:36-41.e3. doi:10.1016/j.exphem.2019.11.008

3. Levin MG, Nakao T, Zekavat SM, et al. Genetics of smoking and risk of clonal hematopoiesis. Sci Rep. 2022;12(1). doi:10.1038/s41598-022-09604-z

4. Matatall KA, Jeong M, Chen S, et al. Chronic Infection Depletes Hematopoietic Stem Cells through Stress-Induced Terminal Differentiation. *Cell Rep*. 2016;17(10):2584-2595. doi:10.1016/j.celrep.2016.11.031

5. Hormaechea-Agulla D, Matatall KA, Le DT, et al. Chronic infection drives Dnmt3a-loss-of-function clonal hematopoiesis via IFNγ signaling. Cell Stem Cell. 2021;28(8):1428-1442.e6. doi:10.1016/j.stem.2021.03.002

6. Meisel M, Hinterleitner R, Pacis A, et al. Microbial signals drive pre-leukaemic myeloproliferation in a Tet2-deficient host. *Nature*. 2018;557(7706):580-584. doi:10.1038/s41586-018-0125-z