ARIC Manuscript Proposal Form

ARIC Publication Admin Use Only

Publication Committee Review Date: **08/13/24**
ARIC Manuscript Proposal Number: **# 4511**

1.a. Full Title: Plant-based diets, ultra-processed plant foods, and epigenetic aging: results from the Atherosclerosis Risk in Communities (ARIC) Study and Multi-Ethnic Study of Atherosclerosis (MESA)

b. Abbreviated Title (Length 26 characters):**plant-based diets and epigenetic aging**

2. Writing Group [please provide a middle initial if available; EX: Adam L Williams]:
Writing group members: Hyunju Kim, Jiantao Ma, Alexis C. Wood, Morgan E. Grams, Bing Yu, Jerome Rotter, Stephen S. Rich, Kelly Ruggles, Dan Arking, Casey M. Rebholz

Others welcome

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]**

First author [please provide a middle initial; EX: Adam L Williams]:

Hyunju		**Kim**
First name	Middle initial	Last name

Address: 3980 15th Avenue NE
Rm 886, Seattle WA 98105

Phone: 206-685-7902
E-mail: hyunjuk1@uw.edu

ARIC author to be contacted if there are questions about the manuscript and the first author does not respond or cannot be located (The ARIC author should be involved enough in ARIC to be able to point the lead author to appropriate ancillary study PIs and to be able to search ARIC manuscript proposals if the lead author doesn't have the access needed to do such a search).

Name: Dr. Casey M. Rebholz
Address: 2024 E. Monument St, Suite 2-500
Baltimore, Maryland 21287

Phone: 410-502-2539
E-mail: crebhol1@jhu.edu

3. Timeline: Analyses will begin when the proposal is approved. We anticipate that a first draft of the manuscript will be available within approximately one year of manuscript proposal approval.

4. Rationale:

Plant-based diets are rich in plant foods and low in animal products. Epidemiological studies, including the ARIC Study found that quality of plant-based diets is important to consider for the development of chronic conditions.[1–4] In the ARIC Study, greater alignment with an overall plant-based diet and a provegetarian diet (diets relatively higher in plant foods and lower in animal products) was associated with a lower risk of incident CVD.[5] Greater alignment with a healthy plant-based diet (diets relatively higher in healthy plant foods, but lower in unhealthy plant foods and animal products) was associated with a lower risk of cardiovascular risk factors (chronic kidney disease, type 2 diabetes, and hypertension), and deaths due to CVD.[3,6,7] In contrast, an unhealthy plant-based diet (diets relatively higher in unhealthy plant foods, but lower in healthy plant foods and animal products) was associated with higher risk of these conditions.[3,7]

Several plant foods that are considered 'unhealthy' in unhealthy plant-based diets are highly processed and may include industrial substances such as chemical additives, colorants, emulsifiers, and flavorings. These foods, termed 'ultra-processed foods,' have been consistently linked to several adverse health outcomes, such as incident CVD.[8] Recent evidence suggests that plant foods that are ultra-processed can be detrimental to health. In a recent study of British adults, greater consumption of plant foods that are ultra-processed (e.g., confectionery, ready-made baking mixes, soft drinks, snacks, processed soy products) was associated with a higher risk of incident CVD and mortality.[9] However, the mechanisms that explain the observed associations of plant-based diets, quality of plant foods, and CVD are incompletely known.

Epigenetic aging may help explain mechanisms underlying the observed associations. Epigenetic aging refers to the biological age expressed within an individual's DNA. Acceleration in epigenetic aging based on DNA methylation, including 1) Horvath, 2) Hannum, 3) GrimAge, and 4) PhenoAge][10–13] is thought to represent future risk of coronary heart disease, heart failure, and all-cause mortality, independent of chronological age.[14,15] Plant-based diets can modify DNA methylation patterns in several ways: by influencing the availability of methyl donors (e.g., folate) which are involved in one carbon metabolism, altering activity and function of enzymes that facilitate or reverse methylation, and modulating oxidative stress.[16–18] Plant foods that are ultra-processed contain pro-inflammatory compounds (acrylamide, per- and polyfluoroalkyl substances),[19,20] which may lead to accelerated epigenetic aging. A large population-based study of US adults found that the Dietary Approaches to Stop Hypertension (DASH) diet was associated with decelerated epigenetic aging (lower levels of GrimAge and PhenoAge), and mediated 22-45% of the association between DASH diet and all-cause mortality.[21] Nonetheless,

no previous study has examined if epigenetic aging is a pathway through which plant-based diets, and quality of plant foods, are associated with incident CVD.

Considering these gaps, our proposal aims to 1) evaluate the associations between four plant-based diet indices, previously associated with incident CVD in the ARIC Study, and epigenetic aging, 2) examine the associations between plant foods that are ultra-processed and epigenetic aging, and 3) test if epigenetic aging mediates the associations between plant-based diets, quality of plant foods, and incident CVD in two large community-based cohorts.

5. Main Hypothesis/Study Aims:

Hypothesis 1: Greater adherence to the overall plant-based diet index, provegetarian diet index, and healthy plant-based diet index is associated with slower epigenetic aging. Greater adherence to the unhealthy plant-based diet index is associated with faster epigenetic aging.

Hypothesis 2: Greater intake of plant foods that are not ultra-processed is associated with slower epigenetic aging. Greater intake of plant foods that are ultra-processed is associated with faster epigenetic aging.

Hypothesis 3: Epigenetic aging mediates the association between the overall plant-based diet index, provegetarian diet index, and healthy plant-based diet index and incident CVD.

Hypothesis 4: Epigenetic aging mediates the association between ultra-processed plant foods and incident CVD.

6. Design and analysis - please address the following aspects:

a) **Inclusion/exclusion:** Individuals with dietary data collected at ARIC visit 1 (1987-1989) and visit 3 (1993-1995), and methylation data collected at visit 2 (1990-1992) or visit 3 will be eligible for the analyses. We will exclude participants with 1) implausible dietary intake (men <700 or 4500 kcal and women <500 or >3500 kcal), 2) missing more than 10 items in the food frequency questionnaire (FFQ), 3) missing DNA methylation data, 4) prevalent CVD, 5) missing covariates (described below), and 6) missing incident CVD or follow-up time.

b) **Study design:** Cross-sectional analysis (hypothesis 1 and 2) of plant-based diet indices, plant foods that are ultra-processed, and epigenetic aging. Mediation analysis (hypothesis 3 and 4) of epigenetic aging and incident CVD through the latest follow-up period.

c) **Outcome and other variables of interest with specific reference to the time of their collection:**

Hypothesis 1 and Hypothesis 2:
Exposure: **Plant-based diet indices.** At visit 1 and 3, participants' usual dietary intake was assessed by trained interviewers using a modified version of the 66-item semiquantitative Willett FFQ.[22] Participants reported the frequency with which they consumed foods and beverages of a specific serving size in the previous year. We will

average responses from the food frequency questionnaire conducted at visit 1 and visit 3, and calculate the 4 plant-based diet indices (overall plant-based diet index, provegetarian diet index, healthy plant-based diet index, and unhealthy plant-based diet index). Details on calculation of these scores are provided in our previous study.[3]

Plant foods that are ultra-processed or minimally processed. From the FFQ, we will identify the plant foods that are ultra-processed (e.g., margarine or margarine and butter blend; chocolate bars or pieces; candy without chocolate; pie (ready-made or from mix), donuts; cake or brownie; cookies, Danish pastry, sweet roll, coffee cake and croissant; cold breakfast cereal; biscuits, cornbread, cold breakfast cereal; peanut butter; potato or corn chips; French fried potatoes; low calorie soft drinks; regular soft drinks; fruit-flavored punch or non-carbonated beverages). Intake of these foods will be summed to calculate ultra-processed plant foods. We will identify plant foods are unprocessed or minimally processed by examining the list of food items in the FFQ.

Outcome: DNA methylation was assessed using the Illumina Infinium HumanMethylation 450 BeadChip platform. We will calculate four epigenetic aging outcomes: 1) Horvath, 2) Hannum, 3) GrimAge, and 4) PhenoAge. Horvath's model calculated epigenetic age as a weighted average of the 353 CpG sites, adjusting for blood cell type proportions.[10] The Hannum model is calculated in a similar manner as Horvath's model but is weighted by blood cell type proportions.[11] GrimAge is calculated using 10 methylation biomarkers, chronological age, and sex.[12] PhenoAge is calculated based on 9 clinical biomarkers and chronological age.[13] We will calculate four measures of epigenetic age using Horvath's online calculator.[23] We will also consider GrimAge2, version 2 of GrimAge which was reported to have stronger association with morbidity outcomes than GrimAge.[24] We will use residuals from a linear regression model (epigenetic age as the response variable and chronological age as an independent variable and technical variables as a covariate). Higher residuals of epigenetic aging indicate faster biological aging.

Other variables of interest: Covariates will include age, sex, race-center, education, income (such as area deprivation index), total energy intake, smoking status, physical activity, alcohol consumption, and body mass index.

Hypothesis 3 and Hypothesis 4:
Exposure: Four epigenetic measures calculated above: 1) Horvath, 2) Hannum, 3) GrimAge or GrimAge2, and 4) PhenoAge.

Outcome: Incident CVD. Incident CVD will be defined as a composite of fatal or nonfatal coronary heart disease, definite or probable myocardial infarction, definite or probable stroke, and heart failure in the ARIC Study. Incident CVD was ascertained via annual telephone calls, active surveillance of local hospital discharge lists, health department death certificate files, and linkage to the National Death Index.

Other variables of interest: Covariates will include age, sex, race-center, education, income (such as area deprivation index), total energy intake, smoking status, physical

activity, alcohol consumption, and additional clinical factors (body mass index, estimated glomerular filtration rate (Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) 2021 race-free equation), hypertension, diabetes, and low-density lipoprotein-cholesterol)

d) Summary of data analysis:

For hypothesis 1 and 2, multivariable linear regression models will be used study the association between 1) each plant-based diet index (per 10-point higher), and 2) plant foods that are ultra-processed (per serving higher) 3) plant foods that are not ultra-processed (per serving higher) and four residuals of epigenetic aging (response variables). Models will be adjusted for covariates outlined above. Then, analyses will be repeated using quartiles of each dietary exposure. The median value of each dietary exposure from each quartile will be used to test for a linear trend. Estimates will be meta-analyzed across the ARIC Study and MESA. To take into account multiple testing, we will use the Bonferroni-adjusted *P* values.

For hypothesis 3 and 4, causal mediation analysis will be used to investigate the degree to which epigenetic aging mediates the association between each dietary exposure and incident CVD ("mediation" package in R).[25] Multivariable linear regression models (same as above) will estimate the association between each dietary exposure and epigenetic aging, and Cox regression models will estimate the association between epigenetic aging and incident CVD. The proportion of the effect attributed to mediation will be calculated as indirect effects / (direct and indirect effects).

As a sensitivity analysis, we will consider 1) stratifying by individuals with major health conditions (type 2 diabetes status or prevalent CVD), and 2) a model adjusted and not adjusted for body mass index.

e) Any anticipated methodologic limitations or challenges if present

Some DNA methylation data were assessed at visit 2. Thus, timing of DNA methylation and dietary intake is not concordant. However, in the ARIC Study, differences in diet between visit 1 and visit 3 was small and average of visit 1 and visit 3 dietary intake data will improve the precision of usual dietary estimation.[26] A sensitivity analysis restricting the analytic sample to ARIC participants with DNA methylation data and dietary data at visit 3 will be conducted.

f) Will the author need Limited data to complete the proposed manuscript?
☒ **Yes, Limited data is needed (Provide a brief (2-3 sentences) justification for requesting PHI data)** . ☐ **No, De-identified data will be sufficient.***Please note, Limited dataset access is strict and rarely provided. Limited data includes identifiable information such as dates (birthdays, visit dates, etc.). CMS, Genomic, Geocoded, Proteomics/Somalogic, and other -omic data all fall under the limited data

category. De-identified data does not include dates. All dates are date adjusted to "Days since Visit 1".

This proposal aims to investigate the associations between dietary exposures (plant-based diets and plant-based diets that are ultra-processed), and epigenetic aging. Access to DNA methylation data is needed.

7.a. Will the data be used for non-ARIC analysis or by a for-profit organization in this manuscript? (Non-ARIC analysis means that the authors are not regarded as ARIC investigators and the "ARIC author" is essentially just a facilitator rather than an integral part of the writing group.) ☐ **Yes** ☒ **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" ? ☐ **Yes** ☐ **No**
(The file ICTDER is distributed to ARIC PIs annually, and contains the responses to consent updates related to stored sample use for research.)

8.a. Will the DNA data be used in this manuscript? ☒ **Yes** ☐ **No**

8.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** ☐ **No**

9. The lead author or the "sponsoring" ARIC 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 website at: https://aric.cscc.unc.edu/aric9/proposalsearch [ARIC Website□Publications□Proposal Search]

☒ Yes ☐ No

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)?

MS#3146 DNA methylation in Relation to Diet Quality in the CHARGE Consortium
The manuscript based on the proposal #3146 has been published [Ma J, Rebholz CM et al., Whole Blood DNA Methylation Signatures of Diet Are Associated With Cardiovascular Disease Risk Factors and All-Cause Mortality. *Circulation: Genomic and Precision Medicine*. 2020;13:e002766.].
The present manuscript proposal will use different dietary exposure (plant-based diets) and different outcome (epigenetic aging).

MS#3774 Replication Request for: Omic signatures of vitamins C and E

The prior proposal was a replication where the ARIC Study contributed data, and focuses on a different dietary exposure (vitamins C and E).

MS#3266 Epigenome-wide association study of nut consumption

The prior proposal focuses on a different dietary exposure (nut consumption).

MS#4195 Methylome of plant-based diets

The prior proposal aims to identify DNA methylation signatures of plant-based diets, but did not focus on epigenetic aging.

11.a. Is this manuscript proposal associated with any ARIC ancillary studies or does it use current [or ongoing] ancillary study data (this includes ACHIEVE)? ☒ **Yes** ☐ **No → Skip to question 12**

11.b. If yes to 11.a., is the proposal

☒ **A. primarily the result of an ancillary study**

☐ **B. primarily based on ARIC data with ancillary data playing a minor role (usually control variables)**

11.c. If yes to 11.a., list number*

2022.07 Proteomic and epigenetic alterations associated with plant-based diets and CVD (PI: Kim)

*ancillary studies are listed by number
https://aric.cscc.unc.edu/aric9/researchers/ancillary_studies/approved_ancillary_studies [ARIC Website□Ancillary Studies□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 https://aric.cscc.unc.edu/aric9/publications/policies_forms_and_guidelines [ARIC Website□Publications□Publication Policies, Forms, and Guidelines]. http://publicaccess.nih.gov/submit_process_journals.htm shows you which journals automatically upload articles to PubMed central.

References:

1. Qian F, Liu G, Hu FB, Bhupathiraju SN, Sun Q. Association between plant-based dietary patterns and risk of type 2 diabetes: a systematic review and meta-analysis. *JAMA Intern Med*. 2019;179:1335–1344.

2. Gan ZH, Cheong HC, Tu Y-K, Kuo P-H. Association between plant-based dietary patterns and risk of cardiovascular disease: a systematic review and meta-analysis of prospective cohort studies. *Nutrients*. 2021;13:3952.

3. Kim H, Caulfield LE, Garcia-Larsen V, Steffen LM, Grams ME, Coresh J, Rebholz CM. Plant-based diets and incident CKD and kidney function. *Clin J Am Soc Nephrol*. 2019;14:682–91.

4. Satija A, Bhupathiraju SN, Rimm EB, Spiegelman D, Chiuve SE, Borgi L, Willett WC, Manson JE, Sun Q, Hu FB. Plant-based dietary patterns and incidence of type 2 diabetes in US men and women: results from three prospective cohort studies. *PLoS Med*. 2016;13:e1002039.

5. Kim H, Caulfield LE, Garcia-Larsen V, Steffen LM, Coresh J, Rebholz CM. Plant-based diets are associated with a lower risk of incident cardiovascular disease, cardiovascular disease mortality, and all-cause mortality in a general population of middle-aged adults. *J Am Heart Assoc*. 2019;8:e012865.

6. Sullivan VK, Kim H, Caulfield LE, Steffen LM, Selvin E, Rebholz CM. Plant-Based Dietary Patterns and Incident Diabetes in the Atherosclerosis Risk in Communities (ARIC) Study. *Diabetes Care*. 2024;47:803–809.

7. Kim H, Rebholz CM, Garcia-Larsen V, Steffen LM, Coresh J, Caulfield LE. Operational differences in plant-based diet indices affect the ability to detect associations with incident hypertension in middle-aged US adults. *J Nutr*. 2019;150:842–50.

8. Lane MM, Gamage E, Du S, Ashtree DN, McGuinness AJ, Gauci S, Baker P, Lawrence M, Rebholz CM, Srour B, Touvier M, Jacka FN, O'Neil A, Segasby T, Marx W. Ultra-processed food exposure and adverse health outcomes: umbrella review of epidemiological meta-analyses. *BMJ*. 2024;384:e077310.

9. Rauber F, Louzada ML da C, Chang K, Huybrechts I, Gunter MJ, Monteiro CA, Vamos EP, Levy RB. Implications of food ultra-processing on cardiovascular risk considering plant origin foods: an analysis of the UK Biobank cohort. *The Lancet Regional Health – Europe* [Internet]. 2024 [cited 2024 Jul 8];0. Available from: https://www.thelancet.com/journals/lanepe/article/PIIS2666-7762(24)00115-7/fulltext

10. Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol*. 2013;14:R115.

11. Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sadda S, Klotzle B, Bibikova M, Fan J-B, Gao Y, Deconde R, Chen M, Rajapakse I, Friend S, Ideker T, Zhang K. Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol Cell*. 2013;49:359–367.

12. Lu AT, Quach A, Wilson JG, Reiner AP, Aviv A, Raj K, Hou L, Baccarelli AA, Li Y, Stewart JD, Whitsel EA, Assimes TL, Ferrucci L, Horvath S. DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging (Albany NY)*. 2019;11:303–327.

13. Levine ME, Lu AT, Quach A, Chen BH, Assimes TL, Bandinelli S, Hou L, Baccarelli AA, Stewart JD, Li Y, Whitsel EA, Wilson JG, Reiner AP, Aviv A, Lohman K, Liu Y, Ferrucci L, Horvath S. An epigenetic biomarker of aging for lifespan and healthspan. *Aging (Albany NY)*. 2018;10:573–591.

14. Roetker NS, Pankow JS, Bressler J, Morrison AC, Boerwinkle E. A Prospective Study of Epigenetic Age Acceleration and Incidence of Cardiovascular Disease Outcomes in the Atherosclerosis Risk in Communities (ARIC) Study. *Circ Genom Precis Med*. 2018;11:e001937.

15. Joyce BT, Gao T, Zheng Y, Ma J, Hwang S-J, Liu L, Nannini D, Horvath S, Lu AT, Bai Allen N, Jacobs DR, Gross M, Krefman A, Ning H, Liu K, Lewis CE, Schreiner PJ, Sidney S, Shikany JM, Levy D, Greenland P, Hou L, Lloyd-Jones D. Epigenetic Age Acceleration Reflects Long-Term Cardiovascular Health. *Circ Res*. 2021;129:770–781.

16. Miles FL, Mashchak A, Filippov V, Orlich MJ, Duerksen-Hughes P, Chen X, Wang C, Siegmund K, Fraser GE. DNA Methylation Profiles of Vegans and Non-Vegetarians in the Adventist Health Study-2 Cohort. *Nutrients*. 2020;12:3697.

17. Yin R, Mao S-Q, Zhao B, Chong Z, Yang Y, Zhao C, Zhang D, Huang H, Gao J, Li Z, Jiao Y, Li C, Liu S, Wu D, Gu W, Yang Y-G, Xu G-L, Wang H. Ascorbic acid enhances Tet-mediated 5-methylcytosine oxidation and promotes DNA demethylation in mammals. *J Am Chem Soc*. 2013;135:10396–10403.

18. Beetch M, Harandi-Zadeh S, Shen K, Lubecka K, Kitts DD, O'Hagan HM, Stefanska B. Dietary antioxidants remodel DNA methylation patterns in chronic disease. *British Journal of Pharmacology*. 2020;177:1382–1408.

19. Martínez Steele E, Buckley JP, Monteiro CA. Ultra-processed food consumption and exposure to acrylamide in a nationally representative sample of the US population aged 6 years and older. *Prev Med*. 2023;174:107598.

20. Hampson HE, Costello E, Walker DI, Wang H, Baumert BO, Valvi D, Rock S, Jones DP, Goran MI, Gilliland FD, Conti DV, Alderete TL, Chen Z, Chatzi L, Goodrich JA. Associations of dietary intake and longitudinal measures of per- and polyfluoroalkyl substances (PFAS) in predominantly Hispanic young Adults: A multicohort study. *Environ Int*. 2024;185:108454.

21. Kim Y, Huan T, Joehanes R, McKeown NM, Horvath S, Levy D, Ma J. Higher diet quality relates to decelerated epigenetic aging. *Am J Clin Nutr*. 2022;115:163–170.

22. Willett WC, Sampson L, Stampfer MJ, Rosner B, Bain C, Witschi J, Hennekens CH, Speizer FE. Reproducibility and validity of a semiquantitative food frequency questionnaire. *Am J Epidemiol*. 1985;122:51–65.

23. DNA Methylation Age Calculator [Internet]. [cited 2024 Jul 11];Available from: https://dnamage.genetics.ucla.edu/new

24. Lu AT, Binder AM, Zhang J, Yan Q, Reiner AP, Cox SR, Corley J, Harris SE, Kuo P-L, Moore AZ, Bandinelli S, Stewart JD, Wang C, Hamlat EJ, Epel ES, Schwartz JD, Whitsel EA, Correa A, Ferrucci L, Marioni RE, Horvath S. DNA methylation GrimAge version 2. *Aging (Albany NY)*. 2022;14:9484–9549.

25. Tingley D, Yamamoto T, Hirose K, Keele L, Imai K. mediation: R Package for causal mediation analysis. *Journal of Statistical Software*. 2014;59:1–38.

26. Xu Z, Steffen LM, Selvin E, Rebholz CM. Diet quality, change in diet quality, and risk of incident cardiovascular disease and diabetes. *Public Health Nutr*. 2020;23:329–338.

To view publications materials, click "Log in" at the top right of the ARIC website. Click "Forgot Password" if you are experiencing issues with logging in.