ARIC Manuscript Proposal #4042

PC Reviewed: 5/17/22	**Status: _____**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: Validation study in ARIC for plasma proteomics associated with incident coronary heart disease in CHS

b. Abbreviated Title (Length 26 characters): Validating CHD-related proteins

2. Writing Group:

Matthew Huber, Josh Bis, Jen Brody, Colleen Sitlani, Tom Austin, Barbara McKnight, Dan Katz, Robert Gerszten, Russell Tracy, Jorge Kizer, Bruce Psaty, Ali Shojaie, Christie Ballantyne James Floyd, Kunihiro Matsushita, Joe Coresh

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

First author: **Matthew Huber**
Address:

Phone: 352.514.3510 Fax:
E-mail: mahuber@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 (this must be an ARIC investigator).

Name: **Kunihiro Matsushita**
Address: Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health
Division of Cardiology, Johns Hopkins School of Medicine
Welch Center for Prevention, Epidemiology, and Clinical Research
2024 E. Monument St., Suite 2-600 (Rm 2-602), Baltimore, MD 21287
Phone: (443) 287-8766 Fax: (410) 367-2384
E-mail: kuni.matsushita@jhu.edu

3. Timeline: As soon as paper proposal approved, using existing ARIC Study data

4. Rationale:

Despite the dramatic decline in morbidity and mortality associated with coronary heart disease (CHD) since the 1960s (1), the rate of decline for mortality has slowed in recent years (2), and CHD remains the leading cause of death in the US and most of the developed world (3, 4). Large-scale genetic association studies have identified over 200 loci for CHD (5-8), uncovering new biology and clarifying mechanisms of disease. However, most of these genetic loci are in

noncoding regions and have not been linked to individual genes or proteins, limiting their relevance to the discovery of new therapeutic targets.

The systematic profiling of a large number of plasma proteins in large, well-characterized populations may offer a complementary approach to genetic discovery. In a study of 938 participants with prior CHD, Ganz et al. identified 200 proteins associated with recurrent cardiovascular events (Bonferroni-corrected threshold, $p < 4.7x10\text{-}5$), many of which had not previously been identified as cardiovascular risk factors, and a risk score based on nine of these proteins improved the prediction of recurrent events beyond a conventional risk prediction model (9). In that prevalent disease population, which included a cross-sectional sample of survivors of recent and distant past events, some proteins may have been associated with differential survival rather than disease progression. As a result, it is unclear whether any of the identified proteins may be causal risk factors for future CHD events.

In the Cardiovascular Health Study (CHS), we have conducted proteomics analyses of incident CHD events using the SomaScan 5k platform and have identified several proteins (e.g., MMP-12) significantly associated with incident CHD events. In this paper proposal, we will conduct targeted replication analyses (12 proteins detailed below) using existing ARIC data on SomaScan proteomics and incident CHD events.

5. Main Hypothesis/Study Questions:

We hypothesize that several proteins associated with incident CHD in CHS are associated with incident CHD in the ARIC Study.

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).

In CHS, we have used the AMNL normalized SomaScan data, and we propose to use the same dataset for the ARIC replication analyses. The specific proteins to be evaluated are listed below.

AptName	SomaId	TargetFullName	Target	Gene
MMP12.4496.60	SL000522	Macrophage metalloelastase	MMP-12	MMP12
NPPB.7655.11	SL002785	N-terminal pro-BNP	N-terminal pro-BNP	NPPB
TNNT2.5315.22	SL000052	Troponin T, cardiac muscle	Troponin T	TNNT2
DNAJB9.11214.40	SL019363	DnaJ homolog subfamily B member 9	DNJB9	DNAJB9
ARL5B.17404.5	SL019712	ADP-ribosylation factor-like protein 5B	ARL5B	ARL5B
REG3A.15304.1	SL002648	Regenerating islet-derived protein 3-alpha	PAP1	REG3A
LEAP2.5708.1	SL007589	Liver-expressed antimicrobial peptide 2	LEAP2	LEAP2
GDF15.4374.45	SL003869	Growth/differentiation factor 15	MIC-1	GDF15
VOPP1.14618.26	SL012766	Vesicular, overexpressed in cancer, prosurvival protein 1	ECOP	VOPP1
CREBBP.13614.6	SL003777	CREB-binding protein	CREB-binding protein	CREBBP
NOTUM.8252.2	SL012580	Palmitoleoyl-protein carboxylesterase NOTUM	NOTUM	NOTUM
MXRA8.10521.10	SL017989	Matrix-remodeling-associated protein 8	MXRA8	MXRA8

ANALYSIS 1

Study population: ARIC visit 5 participants (to align with the age distribution in CHS)
Exclusion criteria: No reliable proteomics data, prevalent MI
Exposure: Each protein level, log2 transformed and normalized to mean of zero and SD of 1
Outcome: Adjudicated incident CHD (probable or definite non-fatal or fatal MI, definite fatal CHD)
Adjustment variables:

- Clinic site
- Age
- Sex
- Race
- Smoking (current vs no)
- Diabetes mellitus (yes/no)
- Treated hypertension (yes/no)
- Systolic blood pressure
- Diastolic blood pressure
- Body mass index
- LDL cholesterol
- eGFR

Report sample size, number of events, mean (SD) follow up time. Cox regression to estimate Hazard ratios and 95% Cis (or betas and SEs), P values, adjusted for all variable above. Significant threshold = 0.05/number of tests. We will conduct a sensitivity analysis after excluding those with prior history of coronary revascularization at baseline.

ANALYSIS 2 (for protein MMP12 only because this was the top protein finding in CHS)
Study populations: ARIC visit 2 participants and ARIC visit 3 participants as separate analyses
Exclusion criteria: No reliable proteomics data, prevalent MI
Exposure: MMP12 level, log2 transformed and normalized to mean of zero and SD of 1
Outcome: Incident CHD (as described above)
Adjustment variables:

- Clinic site
- Age
- Sex

Race
Smoking (current vs no)
Diabetes mellitus (yes/no)
Treated hypertension (yes/no)
Systolic blood pressure
Diastolic blood pressure
Body mass index
LDL cholesterol
eGFR

Report sample size, number of events, mean (SD) follow up time. Cox regression to estimate Hazard ratios and 95% CIs (or betas and SEs), P values, adjusted for all variable above. We will conduct a sensitivity analysis after excluding those with prior history of coronary revascularization at baseline.

Secondary analysis: Same as above, but stratify by presence of subclinical atherosclerosis from carotid U/S or ankle-brachial index at time of baseline proteomics measurement. Report P value for interaction, and for each subgroup report sample size, number of events, mean (SD) follow up time.

REFERENCES

1. Ford ES, Ajani UA, Croft JB, Critchley JA, Labarthe DR, Kottke TE, Giles WH, Capewell S. Explaining the decrease in U.S. deaths from coronary disease, 1980-2000. N Engl J Med. 2007;356(23):2388-98. Epub 2007/06/08. PMID: 17554120.
2. Sidney S, Quesenberry CP, Jr., Jaffe MG, Sorel M, Nguyen-Huynh MN, Kushi LH, Go AS, Rana JS. Recent Trends in Cardiovascular Mortality in the United States and Public Health Goals. JAMA Cardiol. 2016;1(5):594-9. Epub 2016/07/21. PMID: 27438477.
3. Moran AE, Forouzanfar MH, Roth GA, Mensah GA, Ezzati M, Flaxman A, Murray CJ, Naghavi M. The global burden of ischemic heart disease in 1990 and 2010: the Global Burden of Disease 2010 study. Circulation. 2014;129(14):1493-501. Epub 2014/02/28. PMID: 24573351; PMCID: PMC4181601.
4. Collaborators GBDCoD. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet. 2018;392(10159):1736-88. Epub 2018/11/30. PMID: 30496103; PMCID: PMC6227606.
5. Hahn J, Fu YP, Brown MR, Bis JC, de Vries PS, Feitosa MF, Yanek LR, Weiss S, Giulianini F, Smith AV, Guo X, Bartz TM, Becker DM, Becker LC, Boerwinkle E, Brody JA, Chen YI, Franco OH, Grove M, Harris TB, Hofman A, Hwang SJ, Kral BG, Launer LJ, Markus MRP, Rice KM, Rich SS, Ridker PM, Rivadeneira F, Rotter JI, Sotoodehnia N, Taylor KD, Uitterlinden AG, Volker U, Volzke H, Yao J, Chasman DI, Dorr M, Gudnason V, Mathias RA, Post W, Psaty BM, Dehghan A, O'Donnell CJ, Morrison AC. Genetic loci associated with prevalent and incident myocardial infarction and coronary heart disease in the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium. PLoS One. 2020;15(11):e0230035. Epub 2020/11/14. PMID: 33186364; PMCID: PMC7665790 manufacturer (Zoll LifeCor) and on the Steering Committee of the Yale Open Data Access Project. This does not alter our adherence to PLOS ONE policies on sharing data and materials.
6. van der Harst P, Verweij N. Identification of 64 Novel Genetic Loci Provides an Expanded View on the Genetic Architecture of Coronary Artery Disease. Circ Res. 2018;122(3):433-43. Epub 2017/12/08. PMID: 29212778; PMCID: PMC5805277.
7. Klarin D, Zhu QM, Emdin CA, Chaffin M, Horner S, McMillan BJ, Leed A, Weale ME, Spencer CCA, Aguet F, Segre AV, Ardlie KG, Khera AV, Kaushik VK, Natarajan P, Consortium CAD, Kathiresan S. Genetic analysis in UK Biobank links insulin resistance and transendothelial migration pathways to coronary artery disease. Nat Genet. 2017;49(9):1392-7. Epub 2017/07/18. PMID: 28714974; PMCID: PMC5577383.
8. Nikpay M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, Saleheen D, Kyriakou T, Nelson CP, Hopewell JC, Webb TR, Zeng L, Dehghan A, Alver M, Armasu SM, Auro K, Bjonnes A, Chasman DI, Chen S, Ford I, Franceschini N, Gieger C, Grace C, Gustafsson S, Huang J, Hwang SJ, Kim YK, Kleber ME, Lau KW, Lu X, Lu Y, Lyytikainen LP, Mihailov E, Morrison AC, Pervjakova N, Qu L, Rose LM, Salfati E, Saxena R, Scholz M, Smith AV, Tikkanen E, Uitterlinden A, Yang X, Zhang W, Zhao W, de Andrade M, de Vries PS, van Zuydam NR, Anand SS, Bertram L, Beutner F, Dedoussis G, Frossard P, Gauguier D, Goodall AH, Gottesman O, Haber M, Han

BG, Huang J, Jalilzadeh S, Kessler T, Konig IR, Lannfelt L, Lieb W, Lind L, Lindgren CM, Lokki ML, Magnusson PK, Mallick NH, Mehra N, Meitinger T, Memon FU, Morris AP, Nieminen MS, Pedersen NL, Peters A, Rallidis LS, Rasheed A, Samuel M, Shah SH, Sinisalo J, Stirrups KE, Trompet S, Wang L, Zaman KS, Ardissino D, Boerwinkle E, Borecki IB, Bottinger EP, Buring JE, Chambers JC, Collins R, Cupples LA, Danesh J, Demuth I, Elosua R, Epstein SE, Esko T, Feitosa MF, Franco OH, Franzosi MG, Granger CB, Gu D, Gudnason V, Hall AS, Hamsten A, Harris TB, Hazen SL, Hengstenberg C, Hofman A, Ingelsson E, Iribarren C, Jukema JW, Karhunen PJ, Kim BJ, Kooner JS, Kullo IJ, Lehtimaki T, Loos RJF, Melander O, Metspalu A, Marz W, Palmer CN, Perola M, Quertermous T, Rader DJ, Ridker PM, Ripatti S, Roberts R, Salomaa V, Sanghera DK, Schwartz SM, Seedorf U, Stewart AF, Stott DJ, Thiery J, Zalloua PA, O'Donnell CJ, Reilly MP, Assimes TL, Thompson JR, Erdmann J, Clarke R, Watkins H, Kathiresan S, McPherson R, Deloukas P, Schunkert H, Samani NJ, Farrall M. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. Nat Genet. 2015;47(10):1121-30. Epub 2015/09/08. PMID: 26343387; PMCID: PMC4589895.

9. Ganz P, Heidecker B, Hveem K, Jonasson C, Kato S, Segal MR, Sterling DG, Williams SA. Development and Validation of a Protein-Based Risk Score for Cardiovascular Outcomes Among Patients With Stable Coronary Heart Disease. JAMA. 2016;315(23):2532-41. Epub 2016/06/22. PMID: 27327800.

7.a. Will the data be used for non-ARIC analysis or by a for-profit organization in this manuscript? ____ Yes __X__ 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 has been distributed to ARIC PIs, 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 __X__ 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 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

___X___ 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)?
To the best of our knowledge, there are no proposals exploring these 12 proteins in the Soma platform in the context of CHD risk.

11.a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data? __X__ Yes ____ No

11.b. If yes, is the proposal

___ A. primarily the result of an ancillary study (list number* _ 2017.27____)
___ B. primarily 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 you which journals automatically upload articles to PubMed central.