CHARGE Investigator Meeting
Houston 2020

Working Group Abstracts

Poster Blitz and Reception

Thursday, January 30th
5:15 – 7:00PM

Vote for the best poster:

http://www.chargeconsortium.com/houston

Poster 'blitz' participants are denoted with the symbol: ⚡
Travel award recipients are denoted with the symbol: 🚀

Poster Session Participants – by Poster Number

	Author	Abstract Title
1.	Han Chen	FiMap: A Fast Identity by Descent Mapping Test for Biobank-Scale Cohorts
2.	Michael Levin	Genetics of Smoking and Risk of Atherosclerotic Cardiovascular Diseases: A Mendelian Randomization Study
3.	Yue Shan	Phenome-wide association of functional coding variants and cardiometabolic phenotypes in American Indians.
4.	Shih-Jen Hwang	Epigenome-wide association study of coronary artery calcification and its progression
5.	Noah L. Tsao	The in Silico Analysis of ANGPTL3 Rare Exonic Missense Variants
6.	Jordi Merino, Ph.D.	Obesity mediates the adverse effects of reducing low-density lipoprotein cholesterol on type 2 diabetes risk
7.	Jin Choul Chai, Ph.D.	Metabolomic profiling and diabetes risk in US Hispanics/Latinos: Hispanic Community Health Study / Study of Latinos
8.	Magdalena Sevilla Ph.D.	Effect of fasting insulin derived SNP clusters on cardiometabolic outcomes
9.	Merle Behr, Ph.D.	Detecting epistasis with iterative random forest
10.	William Young, MBBS	Trans-ancestry GWAS of ~230,000 Individuals Identifies Over 150 Novel Loci Associated With QT, JT intervals and/or QRS duration
11.	Dan Liu	The association between cardiometabolic risk factors and DNA methylation aging across adult lifespan in the Rhineland Study
12.	Roby Joehanes	Molecular mechanisms of Vitamins C and E in Aging
13.	Kenneth Westerman, Ph.D.	Cloud-based workflows for large-scale gene-environment interaction analysis
14.	Solomon K. Musani, Ph.D.	Multi-ancestry genome-wide meta-analysis incorporating gene-depression interaction in 123,835 individuals identifies five novel lipid loci
15.	Paul S. de Vries	Whole genome sequencing and associations with coagulation factors VII and VIII and von Willebrand factor
16.	Jennifer E Huffman, Ph.D.	Multi-ethnic whole genome sequence analysis of fibrinogen, fibrin D-dimer, tissue plasminogen activator & plasminogen activator inhibitor 1 within the TOPMed program
17.	Mindy Szeto, BS	Epigenome-Wide Analysis of DNA Methylation Reveals Novel Hematologic Trait Associations for African Americans in the Jackson Heart Study
18.	Bridget Lin	Discovery of rare genetic variants from whole genome sequencing analyses of kidney function (eGFR) in 23,732 participants from multi-ethnic populations: the Trans-Omics for Precision Medicine (TOPMed) program
19.	Erica L. Kleinbrink B.S., M.S.	Putative Metabolic Sense: Anti-Sense Loci in the Global Lipids Dataset and Exome Chip Array
20.	Dipender Gill MD	Mendelian randomization: current and future perspectives
21.	Venexia Walker	Investigating unintended drug effects: what can Mendelian randomization add?
22.	Amand Floriaan Schmidt, Ph.D.	Genetic drug target validation using Mendelian randomization
23.	Marios Georgakis, MD	Identification and validation of genetic variants as instruments for studying drug effects

24. Kristin Young, Ph.D.	Distinct metabolomic signatures of central obesity in the Atherosclerosis Risk in Communities (ARIC) Study
25. Jun Liu	A multi-omics study of circulating phospholipid markers of blood pressure
26. Lidia Ximena Orozco Ruiz Ph.D.(c)	Association of Tryptophan, Tyrosine and Branched Chain Amino acids metabolites with abdominal obesity and cardio-metabolic risk factors
27. Jee-Young Moon	Milk Intake, Host LCT Genotype and Gut Bifidobacteria in relation to Obesity: Results from the Hispanic Community Health Study/Study of Latinos (HCHS/SOL)
28. Katie Meyer, MPH, ScD	Urinary hippurate, a microbiota-generated metabolite of dietary components, is negatively associated with hypertension in CARDIA
29. Jessica Williams-Nguyen	Lower Gut Bacterial Diversity in Non-alcoholic Fatty Liver Disease: Results from the Hispanic Community Health Study/Study of Latinos (HCHS/SOL)
30. Xiaoyu Zhang, Ph.D.	Metabolomics insights into osteoporosis through association with bone mineral density
31. Christine Lary, Ph.D.	Pharmacogenetic and MicroRNA Effects of Beta Blocker Association with Increased Bone Mineral Density in Humans
32. Yunju Yang, MPH	Epigenome-wide meta-analysis of cerebral white matter hyperintensities on MRI
33. Aniket Mishra, Ph.D.	The comprehensive gene-mapping study on MRI-derived extremes of cerebral small vessel disease reports role of TRIM47.
34. Ruiqi Wang, MS	Circulating metabolites associated with brain MRI markers of Alzheimer's Disease
35. Laura Ibanez, Ph.D.	Genetic influences on early neurological instability after acute ischemic stroke: GENISIS Results
36. Einat Granot-Hershkovitz Ph.D.	Admixture-mapping identifies genomic regions associated with neurocognitive function
37. Eeva Sliz, Ph.D.	Serum triacylglycerol profiles of cortical thickness and surface area of the human brain
38. Mi Kyeong Lee, Ph.D.	Genome-wide blood DNA methylation in relation to recent use of opioid medications
39. Bonnie Patchen	Genetically determined omega-3 polyunsaturated fatty acid levels and lung function: a Mendelian randomization analysis.
40. Thanh Hoang, Ph.D.	Epigenome-wide Association Study of Adult Asthma in the Agricultural Lung Health Study
41. Zhi Yu, BM, MS	Kidney Function and Blood Pressure: A Mendelian Randomization Study
42. Alexander Teumer, Ph.D.	Assessment of kidney function traits on DNA methylation by developing and applying an EWAS workflow
43. Matthias Wuttke, MD	Heritability enrichment analyses in kidney function GWAS identifies trait-specific kidney cell types

(1)

FiMap: A Fast Identity by Descent Mapping Test for Biobank-Scale Cohorts [*On behalf of the Analysis Working Group*]

Han Chen
The University of Texas Health Science Center at Houston

Han Chen(1,2),
Ardalan Naseri(2),
Degui Zhi(1,2)

1. Human Genetics Center, Department of Epidemiology, Human Genetics and Environmental Sciences, School of Public Health, The University of Texas Health Science Center at Houston, Houston, Texas, USA; 2. Center for Precision Health, School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Houston, Texas, USA

In complex disease genomic epidemiology, genome-wide association studies (GWAS) have identified thousands of genetic loci associated with complex human diseases and quantitative traits. Most GWAS have focused on testing the associations with genotypes (e.g., the number of minor alleles for single nucleotide polymorphisms or copy number variations) from genotyping arrays or DNA sequencing, including common and rare genetic variations but ignoring the phased haplotype information. However, little is known about the roles of mid-range and long-range haplotypes on the genetic architecture of complex traits. Here we leverage the Identity by Descent (IBD) segments inferred from a random projection-based IBD detection algorithm to represent shared haplotypes between individuals, in the mapping of genetic associations with complex traits, and propose a Fast IBD Mapping test (FiMap) for biobank-scale cohorts. Simulation results show that FiMap appropriately controls the type I error under the null hypothesis of no genetic association in large biobank-scale samples, and outperforms traditional GWAS approaches, especially when the causal variants are untyped and rare. We also apply FiMap to IBD mapping of multiple anthropometric quantitative traits using real data from the UK Biobank.

(2)

Genetics of Smoking and Risk of Atherosclerotic Cardiovascular Diseases: A Mendelian Randomization Study
[*On behalf of the Atherosclerosis Working Group*]

Michael Levin, MD
Division of Cardiovascular Medicine, Department of Medicine, University of Pennsylvania Perelman School of Medicine

Michael G. Levin, MD (1)
Derek Klarin, MD (2)
Tim L. Assimes, MD, PhD (3)
Matthew S. Freiberg, MD (4)
Erik Ingelsson, MD, PhD (3)
Julie Lynch, PhD (5)
Pradeep Natarajan, MD (6)
Christopher O'Donnell, MD (6)
Daniel J. Rader, MD (1)
Philip S. Tsao, PhD (3)
Kyong-Mi Chang, MD (1)
Benjamin F. Voight, PhD (1)
Scott M. Damrauer, MD (1)
on behalf of the VA Million Veteran Program

1. University of Pennsylvania Perelman School of Medicine and Corporal Michael J. Crescenz VA Medical Center, Philadelphia, PA; 2. Malcolm Randall VA Medical Center and University of Florida, Gainesville, FL; 3. Palo Alto VA Healthcare System, and Stanford University School of Medicine, Palo Alto, CA; 4. Veterans Affairs Tennessee Valley Healthcare System and Vanderbilt University Medical Center, Nashville, TN; 5. VA Informatics and Computing Infrastructure, Salt Lake City, UT; 6. VA Boston Healthcare System and Harvard Medical School, Boston, MA;

Background: Smoking is associated with atherosclerotic cardiovascular disease, but the relative contribution to each subtype (CAD, PAD, ischemic stroke) remains less well understood.
Objective: To determine the effect of smoking on risk of coronary artery disease, peripheral artery disease, and ischemic stroke.
Methods: We designed a two-sample Mendelian randomization study using summary statistics from genome-wide associations of smoking (GSCAN; up to 311,629 cases, 320,501 controls), coronary artery disease (CARDIoGRAMplusC4D 1000 Genomes + UK Biobank; up to 71,602 cases, 261,418 controls), peripheral artery disease (MVP; up to 24,009 cases, 150,983 controls), and ischemic stroke (MEGASTROKE; up to 34,217 cases, 406,111 controls). We quantified the effect of smoking initiation (ever vs. never) on risk of CAD, PAD, and ischemic stroke using inverse-variance weighted and weighted-median Mendelian randomization.
Results: Genetically-determined smoking was associated with increased risk of PAD (OR 2.57; 95% CI 1.95-3.39; P = 1.8 x 10-11), CAD (OR 1.71; 95% CI 1.41-2.08; P = 6.8 x 10-8), and ischemic stroke (OR 1.37; 95% CI 1.1-1.72; P = 0.0054). Risk of PAD in smokers was greater than risk of ischemic stroke (p for difference = 0.00051) or CAD (p for difference = 0.018).
Conclusions: Smoking is a strong, causal risk factor for CAD, PAD, and stroke, although the effect of smoking is strongest for PAD.

(3)

Phenome-wide association of functional coding variants and cardiometabolic phenotypes in American Indians. [*On behalf of the Atherosclerosis Working Group*]

Yue Shan
Biostatistics, University of North Carolina at Chapel Hill

Yue Shan(1),
Shelley A Cole(2),
Karin Haack(2),
Phillip E Melton(3, 4),
Serena Sanna(5),
Yun Li(1, 6),
Nora Franceschini (7)

1. Department of Biostatistics, University of North Carolina, Chapel Hill, NC; 2. Texas Biomedical Research Institute, San Antonio, TX; 3. The Curtin UWA Centre for Genetic Origins of Health and Disease, Faculty of Health Sciences, Curtin University and Faculty of Health and Medical Sciences, The University of Western Australia, Crawley, Western Australia; 4. School of Pharmacy and Biomedical Sciences, Faculty of Health Sciences, Curtin University, Bentley, Western Australia; 5. Department of Genetics, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands; 6. Departments of Genetics and Computer Science, University of North Carolina, Chapel Hill, NC; 7. Department of Epidemiology, University of North Carolina, Chapel Hill, NC

There has been a great interest in phenotyping individuals with predicted protein-altering functional variants (rare loss of function and missense variants) to understand the health effects in populations. In our ongoing genetic investigation in American Indians, we recently genotyped 1,127 participants from the Strong Heart Family Study, in a case-control study of chronic kidney disease, for single nucleotide variants (SNV) and small indels identified by whole exome sequencing in a subset of individuals and that had predicted protein-altering function. Among 1,206 SNVs and indels that passed quality control (average minor allele count [MAC] = 20, range of 1 to 1,064), ~43% were not present in publicly available repositories, and may be specific to American Indians. Most of the novel coding variants were missense SNVs/indels (n=228), frameshift substitutions (n=73), stop-gain or stop-loss variants (n=8), and splice acceptor/donor (n=11). We performed association analyses of SNVs/indels with a MAC>10 for 32 cardiometabolic biomarkers, using variance component models to account for relatedness, while adjusting for age, sex, center and case-control status. Using a multiple-testing adjusted significance threshold of $p< 5.5x10-6$ to account for 9,122 effective independent tests including 32 traits tested, we identified 11 trait-variant associations. For example, we detected a missense SNV at ABCA10 (p.G1369W, MAC=17, p=8x10-9) associated with increased fasting triglycerides and explaining 2.5% of trait variability. ABCA10 protein is involved in macrophage lipid homeostasis suggesting that it is a cholesterol-responsive gene. A missense SNV located at TRPM3 (p.V1249M, MAC=185, p=5x10-8) was associated with lower fasting insulin concentration and accounted for 1.7% of serum insulin variance. TRPM3 protein is a non-selective cation channel expressed in pancreatic β-cells, which has been shown to regulate insulin secretion. Additional findings include a novel missense SNV at EXTL2 (MAC=23, p=9x10-9) associated with serum creatinine, and a missense SNV at PNPLA5 (MAC=13, p=3x10-7) associated with increased HbA1c. In conclusion, our study focused on predicted damaging coding variants in American Indians identified new gene associations with cardiometabolic phenotypes, demonstrating the advantages of strategies that leverage whole exome sequencing findings to select predicted functional variants for association screenings in less genetic characterized populations.

(4)

Epigenome-wide association study of coronary artery calcification and its progression [*On behalf of the Atherosclerosis Working Group*]

Shih-Jen Hwang, Ph.D.
Population Studies Branch, National Heart, Lung, and Blood Institute, NIH

Shih-Jen Hwang (1), Yinan Zheng (2), Jiantao Ma (3), Penglong Wang (1), Chen Yao (1), Christopher I. O'Donnell (4), Daniel Levy (1)

1. Framingham Heart Study, Framingham, MA and the Population Studies Branch, National Heart, Lung, and Blood Institute, Bethesda, MD

Background
Greater coronary arterial calcification (CAC) and a longitudinal increase in CAC lead to elevated risk of atherosclerotic cardiovascular disease (ASCVD). Characterizing epigenetic signatures of subclinical ASCVD may provide insight into mechanisms of disease and highlight genes or pathways that can serve as targets for CVD prevention and treatment. To this end, we explored DNA methylation as a biomarker of CAC and its progression over time.
Methods
The discovery cohort consisted of 1188 Framingham Heart Study (FHS) Offspring and Third Generation cohort participants who had multidetector computed tomography (MDCT) assessment of CAC and measurement of DNA methylation using the Illumina Human Methylation 450K array. The replication cohort included 773 Coronary Artery Risk Development in Young Adults Study (CARDIA) participants who attended the 15-year and 20-year follow-up examinations with measurements of CAC by MDCT CAC and DNA methylation. Progression of CAC was quantified as annual change of log-transformed CAC. Cohort-specific association analyses were conducted using the linear mixed models for FHS samples and linear regression models for CARDIA. We characterized association of DNA methylation with log-transformed CAC (log(CAC + 1) and adjusted for age, sex, hypertension treatment, diabetes, lipid treatment, cigarette smoking, alcohol consumption, and body-mass-index. To test for association of DNA methylation with CAC progression from the baseline to the follow-up CT, we adjusted additionally for (log(CAC + 1) at the baseline CT scan. We performed

meta-analysis using fixed effect models and corrected for multiple testing using the false discovery rate (FDR).
RESULTS
In the discovery set, the numbers of significant probes at FDR p<0.05 for log(CAC + 1) and annual change in CAC 2 and 28, respectively. Results of meta-analysis revealed two CpG probes associated with log(CAC + 1) : cg03636183 (F2RL3; FDR P=0.04) and cg13434239 (UPF3A; FDR P=0.04). Three CpGs were associated with longitudinal change in CAC at FDR P<1E-4): cg00947324 (AK7: FDR P=1.01E-5), cg10168709 (CAMK2G; FDR P=1.43E-50), and cg12112556 (TNNC2: FDR P=3.67E-5) (Table 1). A total of 27 CpG were associated with by annual changes in CAC at FDR p<0.05.
Conclusions
We identified epigenetic signatures of subclinical ASCVD and its progression by coronary CT. Our findings for longitudinal changes in CAC warrant further investigation.

(5)

The in Silico Analysis of ANGPTL3 Rare Exonic Missense Variants [*On behalf of the Atherosclerosis Working Group*]

Noah L. Tsao., B.S.
Department of Surgery, Perelman School of Medicine, University of Pennsylvania

Noah L. Tsao (1), Nosheen Reza (2), Xiao Wang (3), Kiran Musunuru (3), Scott M. Damrauer (1)

1. Department of Surgery, Perlman School of Medicine, University of Pennsylvania, Philadelphia, PA; 2. Center for Inherited Cardiovascular Disease, Division of Cardiovascular Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA; 3. Division of Cardiology and Cardiovascular Institute, Department of Medicine, Perlman School of Medicine, University of Pennsylvania, Philadelphia, PA

Angiopoietin-like protein 3 (ANGPTL3) and its role in triglyceride metabolism has been of pertinent interest as a potential target for lipid lowering therapies. To assess the effects of missense variants of ANGPTL3, Angptl3 knockout mice were generated with CRISPR-Cas9 and treated with adenoviral vectors (AdV), expressing 82 different mutant missense ANGPTL3 alleles. When the murine functional data was annotated with various in silico structural and amino acid specific properties of the ANGPTL3 C terminal domain, we found a strong correlation between the residue level solvent accessible surface area and the serum ANGPTL3 concentration in the mice (effect estimate = 1.16 % WT ANGPTL3 Plasma Concentration / Å2, Q = 4x10-5). Moreover, some pathogenic variants resulted in the predicted decrease of the overall stability of the ANGPTL3 C-terminal domain as measured by ΔΔG. As such, our study provides substantiating evidence that specific variants in the C terminal domain of ANGPTL3 impact the proper folding and release of the protein into circulation. Furthermore, we sought to evaluate the use of molecular modeling and molecular dynamics (MD) simulations for assessing the potential structural impact of known missense variants in the N terminal domain of ANGPTL3. We documented the structural differences revealed by these models using independent 2 ns MD simulations. Our simulations revealed that dynamic fluctuation introduced by these variants at both the altered position (effect estimate = -177.18 % WT Triglyceride Concentration / Å, Q = 0.037) and the H55 residue (effect estimate = -153.6 % WT Triglyceride Concentration / Å, Q = 0.037), lead to significant negative functional consequences implicating the rigidity of the structure in LPL binding. As such, the investigation of ANGPTL3-LPL binding through docking studies could build a more comprehensive understanding of ANGPTL3 structure and function that could serve as a guide in future drug design.

(6)

Obesity mediates the adverse effects of reducing low-density lipoprotein cholesterol on type 2 diabetes risk [*On behalf of the Diabetes Working Group*]

Jordii Merino, Ph.D.
Diabetes Unit, Center for Genomic Medicine, Massachusetts General Hospital

Jordi Merino(1), Jee-Young Moon(2), Iyas Daghlas(3), Jerome I. Rotter(4), Jose C. Florez(1), James B. Meigs(5)

1. Diabetes Unit, Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA; 2. Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY; 3. Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA; 4. Institute for Translational Genomics and Population Sciences, Harbor-UCLA Medical Center, Torrance, CA; 5. Division of General Internal Medicine, Massachusetts General Hospital, Boston, MA.

Objective: To test the hypothesis that obesity modifies the association between lowering low-density lipoprotein cholesterol (LDLc) and increased type 2 diabetes (T2D) risk.
Design: Meta-analysis of prospective cohort studies and univariate and multivariate Mendelian Randomization analyses.
Setting: Cohort studies from the CHARGE Consortium (Cohorts for Heart and Aging Research in Genomic Epidemiology) and large-scale biorepositories.
Participants: A combined total of 338,186 individuals from prospective datasets, and over 1,500,000 individuals from different genome-wide association study (GWAS) consortia.
Exposures: LDLc genetic profile was characterized by a 50-variant polygenic score weighted by published effect sizes. Body mass index (BMI) was defined by clinical determinations or the use of BMI summary statistics.
Outcome measures: T2D prevalence and incidence for a predicted 1 mmol/L reduction in LDLc, and the total and direct effect (i.e., the effect not mediated via BMI) of genetically reduced LDLc on T2D odds.
Results: In a meta-analysis of cohort studies, we observed that a 1 mmol/L reduction in LDLc was associated with higher T2D prevalence independently of potential confounders (OR=1.34, 95%CI 1.08 to 1.66, I2=0.0%, τ2=0.0). These findings replicated in UK Biobank where each mmol/L decrease in LDLc was associated with higher T2D prevalence (OR=1.27, 95%CI 1.14 to 1.41). We observed no association between LDLc and T2D incidence (OR=1.12, 95%CI 0.85 to 1.46, I2=0.0%, τ2=0.0), but BMI modified the effect of LDLc

on T2D risk (Pint=0.04). We noted that the effect of lowering LDLc on T2D risk was amplified among nonobese individuals compared to obese individuals (OR=1.51, 95%CI 1.22 to 1.83 vs. OR=0.97, 95%CI 0.76 to 1.18, Q-value= 0.005). These findings were corroborated in a multivariate MR analysis, showing that BMI mediated up to 34.5% of the effect of lowering LDLc on increased T2D odds.
Conclusions: Our findings suggest that obesity mediates a substantial proportion of the increased T2D risk attributed to lowering LDLc. Findings from this study may provide supportive evidence when considering lipid lowering strategies, specifically among nonobese individuals.

(7)

Metabolomic profiling and diabetes risk in US Hispanics/Latinos: Hispanic Community Health Study / Study of Latinos [*On behalf of the Diabetes Working Group*]

Jin Choul Chai, Ph.D.
Department of Epidemiology & Population Health, Albert Einstein College of Medicine

Jin Choul Chai(1), Jun Li(2), Robert Kaplan(1,3), Eric Boerwinkle(4,5), Qibin Qi(1,2)

1. Albert Einstein College of Medicine, Bronx, NY; 2. Harvard T.H. Chan School of Public Health, Boston, MA; 3. Fred Hutchinson Cancer Research Center, Seattle, WA; 4. The University of Texas Health Science Center at Houston, Houston, TX; 5. Baylor College of Medicine, Houston, TX;

Metabolomic profiling offers the potential to reveal metabolic pathways relevant to the development of diabetes and improve diabetes risk prediction. We examined serum metabolite profiles and their relationships with incident diabetes in Hispanic Community Health Study/Study of Latinos (HCHS/SOL).
This analysis included 2010 participants (224 incident diabetes cases, 1786 control) without prevalent diabetes at baseline, aged 18-76 years with 60% female, from six Hispanic/Latino background groups in the HCHS/SOL. We profiled 624 metabolites from 8 metabolomic super-pathways. Survey Poisson regression was used to examine the associations of baseline metabolites with incident diabetes.
After correcting for multiple testing (false-positive discovery rate), 134 metabolites were significantly associated with incident diabetes. For the 134 significant metabolites, network analysis identified multiple topological modules comprised of amino acids and lipids, such as BCAA metabolites, glycine, serine and threonine metabolites, sphingolipids and ceramides, monoacylglycerol and, diacylglycerol, associated with diabetes risk.
This study identified multiple metabolite modules related to diabetes risk in a large cohort of US Hispanics/Latinos and provide new insights to metabolic pathways in the development of diabetes.

(8)

Effect of fasting insulin derived SNP clusters on cardiometabolic outcomes [*On behalf of the Diabetes Working Group*]

Magdalena Sevilla PhD
Clinical Translational and Epidemiology Unit

Magdalena Sevilla (1) Miriam Udler (1), Alisa Manning (1)

Massachussetts General Hospital

Fasting insulin concentrations reflect the state of glucose metabolism. Increased fasting insulin level or insulin resistance is associated with subsequent risk of several cardiometabolic outcomes, such as type 2 diabetes, hypertension, and coronary artery disease. We aim to cluster variants associated with fasting insulin into subsets showing similar patterns of association across a broad set of traits and to identify any broad clinical consequences in cardiometabolic outcomes (coronary artery disease [CAD], systolic blood pressure [SBP], diastolic blood pressure [DBP], body mass index [BMI] and lipid traits [fasting triglycerides, HDL, and LDL]) associated with polygenic scores in the Partners Biobank. We performed a soft clustering analysis to categorize genetic loci into subsets that represent likely mechanistic pathways of fasting insulin. We generated a weighted genetic risk scores (GRS) for each cluster for all individuals with genetic data in the Partners Biobank. We tested the association of each GRS with our outcomes (CAD, SBP, DBP, BMI, triglycerides, HDL, and LDL) using logistic regression for binary outcomes and linear regression for continuous variables. We tested the association of each GRS as a continuous variable and also dichotomized at the 90th percentile. Using 92 SNPs associated with fasting insulin, we obtained five novel groups of genetic variants. Each cluster is represented by a set of traits and a set of loci. For example, one represents insulin resistance traits with loci near SNTG1, TRAF3IP2, and BPTF genes. Another cluster represents liver traits with loci near the GCKR and SLC5A6 genes. Association tests with the Partners Biobank are ongoing.

(9)

Detecting epistasis with iterative random forest [*On behalf of the Diabetes Working Group*]

Merle Behr, Ph.D.
Department of Statistics, UC Berkeley

Merle Behr (1),
Karl Kumbier (2),
Matthew Aguirre (3),
Rima Arnaout (2),
Euan Ashley (3),
Ben Brown (1),
James Priest (3),
Bin Yu (1)

1. University of California at Berkeley, Berkeley, CA; 2. University of California at San Fransisco, San Fransisco, CA; 3. Stanford University, Stanford, CA

Background:
Epistatic interactions play a crucial component in many genetic traits. We present a computationally efficient, accurate, and stable approach to identify epistatic interactions associated with a phenotype.
We explore the power of our procedure using data from the UK Biobank.
Methods:
Most methods for detecting epistasis test a single pair of genes in a logistic regression model,
making them unreliable for traits influenced by many genes. Moreover, with approx. 10^6 data points, it is common sense that we cannot hope to differentiate among approx. 10^{12} hypotheses that arise from testing pairwise SNP interactions, and the computation is also challenging, especially
for higher order interactions.
Here we describe an approach to extract epistatic interactions based on signed iterative Random Forests (siRF) (Basu et al., 2018; Kumbier et al. 2018). Our approach models a phenotype using all genes simultaneously to search for epistatic interactions of arbitrary order.
Following the PCS (predictability, computability, and stability) framework (Yu and Kumbier, 2019), we evaluate candidate interactions based on their predictive power and stability.
Results:
We apply our pipeline to UK Biobank data, using the well-studied redhead phenotype as a positive control and recover both previously reported interactions and novel candidates.
Our full model demonstrates high prediction accuracy on held-out data (AUROC 0.99), and individual interactions show a precision of over 90%.
In addition, extending traditional regression testing to saturated models with complex interaction terms reveals potential epistatic effects beyond multiplicative interactions.
Conclusions:
Our pipeline extracts high-order, epistatic interactions that are predictive and stable in the context of a complete genetic background.
The predictive power of recovered interactions on held-out data suggests that they capture meaningful relationships, while stable relative to bootstrap replicates points to reproducibility of our data results. Together, predictability and stability suggest that our findings are strong candidates for biological validation studies.

(10)

Trans-ancestry GWAS of ~230,000 Individuals Identifies Over 150 Novel Loci Associated With QT, JT intervals and/or QRS duration [*On behalf of the EKG Working Group*]

William Young, MBBS
William Harvey Research Institute, Queen Mary University of London

William J Young(1),
Najim Lahrouchi(2,4),
Nona Sotoodehnia(3),
Patricia Munroe(1),
Christopher Newton-Cheh(4),
on behalf of the CHARGE EKG working group.

1. Queen Mary University of London, London, UK; 2. Amsterdam UMC, Amsterdam, The Netherlands; 3. University of Washington, Seattle, WA; 4. Massachusetts General Hospital & Broad Institute, Boston, MA.

Background
On the electrocardiogram, ventricular depolarisation and repolarisation is represented by QRS duration and JT interval respectively and together, make up the QT interval. These intervals are independent predictors of sudden cardiac death. They are heritable; however a large proportion of the heritability still remains unexplained. To identify novel candidate genes and genetic regions, we tested common and rare variants for association with these ECG traits.
Methods
We performed a trans-ancestry meta-analysis of GWAS summary statistics from 34 studies imputed with 1000G and/or HRC reference panels, comprising 230,192 individuals (including 189,958 European, 16,822 African-American and 19,508 Hispanic). Bivariate LD score regression was used in European samples to evaluate polygenicity and estimate genetic correlation between traits. Candidate gene prioritisation and gene-set enrichment analyses were performed using DEPICT. Gene-based meta-analysis of rare protein-coding variants was performed using SKAT.
Results
We identified 152, 147, and 104 independent loci (of which 93, 89 and 60 were novel) associated with QT, JT and QRS, respectively. A significant positive genetic correlation for QT v JT (rg=0.92, p<0.001) and negative correlation for JT vs QRS (rg=-0.25, p=0.003) were observed. There was no significant genetic correlation for QT v QRS. Gene-set enrichment highlighted cardiovascular (atria, ventricular, appendage) tissues for all three traits; however additionally for QRS duration, membranous and connective tissue cell types reached

significance. Gene-based meta-analysis indicated Mendelian Long QT syndrome genes (KCNQ1, KCNH2 and SCN5A) were associated with QT and JT, MYH7 (implicated in hypertrophic cardiomyopathy) with JT, and SCN5A with QRS. Gene-based associations were driven by more than 1 rare protein-coding variant.
Conclusions
Increased sample size and imputation to dense reference panels have improved power to detect novel associations with ECG correlates of myocardial depolarisation and repolarisation. These interim analyses highlight differences in the genetic contributions to these risk factors for sudden cardiac death.

(11)

The association between cardiometabolic risk factors and DNA methylation aging across adult lifespan in the Rhineland Study [*On behalf of the Epigenetics Working Group*]

Dan Liu
Population Health Sciences, German Centre for Neurodegenerative Diseases (DZNE)

Dan Liu (1),
N. Ahmad Aziz (1,2),
Loes C. Rutten-Jacobs (1),
Monique M.B. Breteler (1,3)

1. Population Health Sciences, German Centre for Neurodegenerative Diseases (DZNE), Bonn, Germany; 2. Department of Neurology, University of Bonn, Bonn, Germany; 3. Institute for Medical Biometry, Informatics and Epidemiology (IMBIE), Faculty of Medicine, University of Bonn, Bonn, Germany

Background: DNA methylation-based biomarkers of aging (DNAm aging) are associated with cardiometabolic diseases. Although this may be due to their relation with cardiometabolic risk factors, limited data are available on the association between cardiometabolic risk factors and DNAm aging in the general population.
Methods: We performed a cross-sectional analysis of data from the ongoing Rhineland Study, a population-based prospective cohort study in Bonn, Germany, that recruits individuals aged ≥ 30 years. Four types of DNAm aging acceleration estimates (including AgeAccelHorvath, AgeAccelHannum, AgeAccelPheno, and AgeAccelGrim), defined as within individual differences between the predicted and chronological age, were calculated using published algorithms. The effects of cardiometabolic risk factors on DNAm aging accelerations were evaluated using multiple linear regression models.
Results: Among 1869 participants (median age 54.0 years (range 31 – 91), 56.0% women), we found that each standard deviation (SD) increase in the following risk factors was significantly associated with DNAm aging acceleration (i.e. difference between estimated and chronological age in years): diastolic blood pressure (DBP) 0.41 (95% CI: 0.19, 0.64), body mass index (BMI) 0.22 (95% CI: 0.01, 0.44), hemoglobin 0.32 (95% CI: 0.04, 0.60) when using AgeAccelHorvath as the outcome, and additionally C-reactive protein (CRP) 0.45 (95% CI: 0.26, 0.63), glomerular filtration rate (GFR) -0.22 (95% CI: -0.41, -0.03) when using AgeAccelHannum as the outcome. Seven out of 13 risk factors were associated with AgeAccelPheno, including HDL -0.55 (95% CI: -0.91, -0.18), triglycerides 0.39 (95% CI: 0.06, 0.72), DBP 0.35 (95%CI: 0.01, 0.68), CRP 0.83 (95% CI: 0.51, 1.15), Cystatin C 0.65 (95% CI: 0.31, 0.98), GFR -0.49 (95% CI: -0.82, -0.15), BMI 0.99 (95% CI: 0.68, 1.31). Similar results were obtained for AgeAccelGrim.
Discussion: Multiple cardiometabolic risk factors were consistently associated with various DNAm aging acceleration estimates, which may underlie the association between DNAm aging and cardiometabolic diseases.

(12)

Molecular mechanisms of Vitamins C and E in Aging [*On behalf of the Epigenetics Working Group*]

Roby Joehanes
Framingham Heart Study

Gha Young Lee (1), Roby Joehanes (1), Caren E. Smith (2), Douglas P. Kiel (3), Daniel Levy (1)

1. Framingham Heart Study, National Heart Blood and Lung Institute, National Institutes of Health, Framingham, MA, USA; 2. Nutrition and Genomics Laboratory, Tufts University, Boston, MA; 3. Institutes for Aging Research, Harvard Medical School, Boston, MA

Introduction: Epigenetic age is a composite age prediction score calculated from levels of DNA methylation that has been shown to predict mortality. It offers the possibility of testing the effects on healthy aging of two common antioxidant vitamins, vitamins C and E, especially as they have been implicated in multiple aging pathways. Vitamin C was reported to modulate DNA demethylation in vitro, which in turn affects downstream gene expression. However, it is currently unknown how the intake of vitamins C and E affect DNA demethylation, and how this process might be involved in pathways underlying healthy aging. This study aims to examine DNA methylation and gene expression associated with vitamins C and E intake, which may also identify potential biomarkers and therapeutic targets for healthy aging.
Method: Whole blood samples of Framingham Heart Study Offspring and Third Generation cohorts were used in the analysis. Vitamin C and E intakes were calculated from Food Frequency Questionnaires obtained from the same examination cycles as the DNA methylation (n=3,866) and gene expression assays (n=5,205). Linear mixed models were used on categorized vitamin intakes, adjusting for sex, age, total caloric intake, blood counts, BMI, technical covariates and family structure. Additional models adjusted for smoking, alcohol intake, physical activity index, and dietary scores.
Results: Several CpGs passing false discovery rate of 0.05 were identified as differentially methylated with respect to vitamin C and E intake. These CpGs were enriched in inflammation pathways. Several common sets of gene transcripts were also associated with vitamin C and E intake including LRRN3, a gene previously associated with chronological aging.
Conclusion: Vitamin C and E work in tandem to alter gene expression through epigenetic effects on critical inflammation-mediating and aging-related CpGs.

(13)

Cloud-based workflows for large-scale gene-environment interaction analysis [*On behalf of the Gene-Lifestyle Working Group*]

Kenneth Westerman, Ph.D.
Clinical and Translational Epidemiology Unit, Massachusetts General Hospital

Kenneth Westerman(1,2), Yun Ju Sung(3), Yan V. Sun(4,5), D.C. Rao(3), Alisa K. Manning(1,2), Han Chen(6,7)

1. Clinical and Translational Epidemiology Unit, Massachusetts General Hospital, Boston, MA; 2. Programs in Metabolism and Medical and Population Genetics, Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, MA; 3. Division of Biostatistics, Washington University School of Medicine, St. Louis, MO; 4. Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA; 5. Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, GA; 6. Human Genetics Center, Department of Epidemiology, Human Genetics and Environmental Sciences, School of Public Health, The University of Texas Health Science Center at Houston, Houston, TX; 7. Center for Precision Health, School of Public Health and School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Houston, TX

Gene-environment interaction (GEI) studies are critical to increasing our understanding of the genetic architecture of complex diseases and explaining health disparities. However, their low inherent statistical power requires very large samples sizes, and for optimal statistical inference they require modeling capabilities compared to standard genome-wide association studies (GWAS). Furthermore, computational and practical challenges with current GEI testing tools (e.g. runtime, memory requirements, installation, etc.) and practical usage of GEI testing tools have thus far prevented the adoption of consistent workflows across the GEI research community. We developed a computational tool, GEM, for performing efficient genome-wide GEI tests in up to millions of individuals. We sought to (1) implement workflows for GEM and alternative GEI testing tools in a user-friendly cloud-computing interface and (2) compare the computational efficiency of GEM to these alternatives. A computational workspace was set up on the Broad Institute's cloud-based Terra platform, with access to imputed genotype datasets from the 1000 Genomes Project (for testing of workflows with simulated GEI exposures and phenotypes) and the UK Biobank (for benchmarking). Workflows were constructed using Workflow Description Language (WDL) and Docker environments to implement both genotype file format conversion (necessary for creating appropriate inputs for each tool) and GEI tests using GEM, QUICKTEST, and ProbABEL. GEM was benchmarked against these alternative tools using a test for gene-sex interaction influencing waist-to-hip ratio for a single chromosome in unrelated individuals of European ancestry from the UK Biobank (N=275,565). In this analysis, GEM outperformed both QUICKTEST and ProbABEL, using less CPU time and a lower memory footprint. This work introduces an accessible interface to efficient tools for GEI testing, provides the foundation for the standardization of GEI analysis workflows using a cloud-based environment, and facilitates large-scale collaborative GEI studies across many individual cohorts.

(14)

Multi-ancestry genome-wide meta-analysis incorporating gene-depression interaction in 123,835 individuals identifies five novel lipid loci [*On behalf of the Gene-Lifestyle Working Group*]

Solomon K. Musani, Ph.D
Jackson Heart Study, University of Mississippi Medical Center

Solomon K. Musani1, Ervin R. Fox1, Melissa A. Richards2, Amy R. Bentley3, Karen L. Schwander4; Myriam Fornage2, Mario Sims1, on behalf of the CHARGE Gene-Lifestyle Interactions Working Group*

1Department of Medicine, University of Mississippi Medical Center, Jackson, MS; 2Human Genetics Center, Department of Epidemiology, Human Genetics, and Environmental Sciences, School of Public Health, University of Texas, Houston, TX; 3Center for Research on Genomics and Global Health, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD; 4Division of Biostatistics, Washington University School of Medicine, St Louis, MO.

Introduction: Serum lipids (high-density lipoprotein, low-density lipoprotein cholesterol and triglycerides) are modifiable risk factors for cardiovascular disease, and are known to be influenced by genetic and non-genetic factors. Although the relationship between lipid levels and depressive symptomatology is contradictory, a number of studies have reported association between major depressive symptoms and adverse lipid profile. Motivated by this relationship or lack thereof, we hypothesized that incorporating gene-depression interaction in genome-wide association studies (GWAS) will enhance our ability to identify new loci influencing lipid levels. Methods: We performed a multi-ancestry gene-depression interaction GWAS involving 123,835 individuals from multiple ancestries with data on ~14.5 million imputed SNPs and depressive symptomatology (Yes/No). In Stage 1 (N=67,671), genome-wide analyses were conducted at all imputed variants that passed quality control and all suggestive variants (P<1x10-6) were followed up in Stage 2 (N=56,164). We used a one degree of freedom (DF) interaction test and a joint 2DF test of genetic main and interaction effects. Results: Combined meta-analyses identified five novel loci associated with lipid levels (P<5x10-8), all from the joint 2DF test, although two were driven by significant main or 1 DF interaction effects. We identified 3 loci in individuals of African ancestry and 2 in trans-ancestry. CREB3L2 identified in Africans, may be responsible for synthesis of fatty acids and cholesterol through its transcriptional activity on the endoplasmic reticulum. MACROD2 plays an important role in multiple biological processes but its role in transcriptional adipogenesis and neuropsychiatric disorders is the probable link to lipid metabolism and depression. RRP1B identified in trans-ancestry, is a breast cancer metastasis suppressor; regulates gene expression through heterochromatinization and transcriptional repression. The DEFB136 is responsible for the production of antimicrobial peptides found in white blood cells such as macrophages and NK-cells. The RNU4-73P loci is a pseudo gene whose function is unknown. Conclusions: Identified lipid loci may elucidate the role of an underlying mechanism for interaction of psychosocial factors

captured by depressive symptomatology in the genetic regulation of lipid levels and depression. We also show evidence for a possible link between lipid levels and cancer.

(15)

Whole genome sequencing and associations with coagulation factors VII and VIII and von Willebrand factor [*On behalf of the Hemostasis Working Group*]

Paul S. de Vries
University of Texas Health Science Center at Houston

Paul S. de Vries(1) and Michael R. Brown(1) on behalf of NHLBI's Trans Omics for Precision Medicine (TOPMed) Hematology and Hemostasis Working Group and the CHARGE Hemostasis Working Group

(1)Human Genetics Center, Department of Epidemiology, Human Genetics, and Environmental Sciences, School of Public Health, The University of Texas Health Science Center at Houston, Houston, TX

Coagulation factor VII (FVII), factor VIII (FVIII), and its carrier protein von Willebrand factor (vWF) are implicated in modulating the risk of arterial and venous thrombosis. We leveraged whole genome sequencing (WGS) data from NHLBI's TOPMed program to identify genetic associations with plasma levels of FVII (n=16,335), FVIII (n=19,766), and vWF (n=14,020). Phenotypes were harmonized across 9 studies that included European, African, Asian, and Hispanic ancestry participants. Association analyses were conducted across all individuals using inverse normalized and rescaled residuals adjusting for age, sex, ancestry, principal components, and a kinship matrix. Analyses were conducted on the Analysis Commons using SMMAT. Single-variant analyses assessed variants with minor allele count ≥40. Aggregate analyses grouped variants with minor allele frequency (MAF) <0.05 by gene, using 3 strategies for selecting variants within genes: 1) loss of function (LOF) variants; 2) LOF and deleterious missense variants; and 3) coding, enhancer and promoter variants.
Single-variant analyses identified associations (P<5E-8) at 4 known loci for FVII, 8 for FVIII, and 8 for vWF. New FVIII associations included rs538727675 located between FNDC3B/GHSR (MAF=0.0015; P=2.2E-8) and rs114894279 downstream of TBL1XR1 (MAF=0.012; P=3.4E-8). A new vWF association was identified with rs147142418 in DPF3 (MAF=0.017; P=1.1E-8). Conditional analyses revealed multiple independent signals at F7, VWF, STAB2, and ABO. Gene-based analyses identified associations at known locus F7 for FVII, and 3 known loci each for vWF and FVIII (VWF, STAB2, and ABO). FVIII was associated with LOF variants in CD36, a novel locus. The driving variant, rs3211938, causes CD36 deficiency and is associated with other hematological phenotypes.
Single variant analyses in CHARGE studies using data imputed to a TOPMed reference panel are currently ongoing, and we will meta-analyze the results from TOPMed with the results from CHARGE to increase the power for new discovery.

(16)

Multi-ethnic whole genome sequence analysis of fibrinogen, fibrin D-dimer, tissue plasminogen activator & plasminogen activator inhibitor 1 within the TOPMed program [*On behalf of the Hemostasis Working Group*]

Jennifer E Huffman, PhD
Center for Population Genomics, VA Boston Healthcare System

Jennifer E. Huffman(1,2), Benjamin Rodriguez(2,3), Laura M. Raffield(4), on behalf of the TOPMed Hematology and Hemostasis and CHARGE Hemostasis Working Groups

1. VA Boston Healthcare System, Jamaica Plain MA; 2. The NHLBI's Framingham Heart Study, Framingham MA; 3. NHLBI's Population Science Branch, Framingham MA; 4. Department of Genetics, University of North Carolina, Chapel Hill NC

Compared with array-based, imputation-based, and exome-focused analyses, whole genome sequencing (WGS) data provides better coverage of the genome and better representation of non-European variants.
To better understand the genetics underlying several hemostasis traits, we leverage Freeze 6 deep WGSs from NHLBI's Trans-Omics for Precision Medicine (TOPMed) program to investigate plasma levels of 4 hemostasis measures: fibrinogen (n= 32,572), fibrin D-dimer (n=19,049), tissue plasminogen activator (tPA; n=4,393), and plasminogen activator inhibitor 1 (PAI-1; n=7,857). Phenotypes were centrally harmonized across up to 12 studies of European, African, Asian, or Hispanic ancestry. Association analyses were conducted using inverse normalized and rescaled residuals adjusting for age, sex, study, TOPMed phase, study-specific parameters, self-reported ancestry, 11 ancestry informative principal components, and a kinship matrix. All analyses were conducted on the Analysis Commons using the SMMAT function implemented in GENESIS. Single-variant analyses included all variants with a minor allele count ≥40. Gene-based aggregate analyses used 3 strategies for variant selection: 1) loss of function (LOF), 2) LOF and deleterious missense (LDM), and 3) coding, enhancer and promoter variants. The latter two aggregation tests were restricted to variants with a minor allele frequency (MAF)<0.05.
Significantly associated single variant regions were found for fibrinogen (n=7) and D-dimer (n=3). All were in loci previously associated with these phenotypes, and the majority were common variants in high linkage disequilibrium with previously reported variants. The most significant association for fibrinogen was a rare missense mutation (rs148685782, p=6.8x10-48, MAF=0.003, FGG) located within the fibrinogen structural genes region. LOF and LDM aggregation tests demonstrated associations with these genes only. No significant genes with >5 alternate alleles were identified for D-dimer. No associations were detected for tPA or PAI-1.
CHARGE cohorts provided single variant results from data imputed to the TOPMed reference panel and we are currently meta-analyzing the sequence and GWAS data to increase power for new discovery.

(17)

Epigenome-Wide Analysis of DNA Methylation Reveals Novel Hematologic Trait Associations for African Americans in the Jackson Heart Study [*On behalf of the Hemostasis Working Group*]

Mindy Szeto, BS
Division of Biomedical Informatics and Personalized Medicine, University of Colorado Anschutz Medical Campus

Mindy Szeto (1),
Laura Raffield (2),
Deborah Nickerson (3),
Neil Zakai (4),
Dwight Klemm (1),
Matthew Taylor (1),
Ivana Yang (1),
Adolfo Correa (5),
Alex Reiner (3),
Ethan Lange (1),
Leslie Lange (1).

1. University of Colorado Anschutz Medical Campus, Aurora, CO; 2. University of North Carolina, Chapel Hill, NC; 3. University of Washington, Seattle, WA; 4. University of Vermont, Burlington, VT; 5. University of Mississippi Medical Center, Jackson, MS.

Background: Cardiovascular disease (CVD) is the leading cause of death in the U.S. and disproportionately affects African Americans. Routinely measured circulating red blood cell traits, which are highly heritable and differ by ethnicity, are independent predictors for CVD-related traits including hypertension, stroke, coronary heart disease, and CVD mortality. Many genetic loci associated with red blood cell count (RBC), hemoglobin (HGB), hematocrit (HCT), mean corpuscular volume (MCV), mean corpuscular hemoglobin (MCH), mean corpuscular hemoglobin concentration (MCHC), and red cell distribution width (RDW) have been identified, but do not fully explain the heritability of these traits. Epigenetic alterations in DNA methylation likely also explain a portion of red cell trait variance, and detecting methylation quantitative trait loci can provide critical insight into the development of CVD and CVD health disparities. Methods: DNA methylation at ~850,000 CpG sites was measured by the MethylationEPIC array in peripheral blood mononuclear cells from 1753 Jackson Heart Study baseline exam participants. We performed epigenome-wide association analyses with red blood cell traits using linear mixed models adjusted for age, sex, cell proportions, genetic ancestry, and experimental batch effects. A Bonferroni-corrected p-value of 9e-8 assessed statistical significance. Results: Analysis revealed a novel highly significant CpG association annotated to a non-coding RNA (cg11703701; pRBC=5.19e-22, pHGB=1.19e-11, pMCV=4.69e-59, pMCH=2.68e-67, pMCHC=2.32e-31), as well as a strong signal for a reprogramming-specific differentially methylated region (cg04321267; pRBC=1.67e-16, pHGB=3.58e-12, pMCV=2.12e-49, pMCH=1.74e-57, pMCHC=5.91e-30). Multiple CpGs annotated to genes HBA1 and HBA2 were associated with RBC, MCV, MCH, and MCHC, while both HGB and HCT were significantly associated with ITPKB (cg23740281; pHGB=4.88e-10, pHCT=6.20e-11) and ALDH2 (cg17969951; pHGB=3.03e-11, pHCT=8.31e-9). Additional CpGs were associated with HCT (PLXND1 cg22902177; p=7.54e-11 and ARL1 cg23903357; p=9.86e-11) and RDW (CPNE2 cg09018739; p=3.03e-22). Conclusions: We identified many significant differentially methylated CpG sites associated with red blood cell traits. These findings shed light on potential hematologic and CVD mechanisms in understudied populations. Future work will explore the role of neighboring SNPs in mediating observed methylation-trait associations, and replicate results in an additional multi-ethnic cohort.

(18)

Discovery of rare genetic variants from whole genome sequencing analyses of kidney function (eGFR) in 23,732 participants from multi-ethnic populations: the Trans-Omics for Precision Medicine (TOPMed) program [*On behalf of the Kidney Working Group*]

Bridget Lin
Biostatistics, University of North Carolina

Bridget Lin (1), Kelsey E Grinde (2), Jennifer Brody (2), Laura Raffield (1), Timothy A Thornton (2), Josyf C Mychaleckyj (3), James A Perry (4), Stephen Rich (3), Sharon Browning (2), Danyu Lin (1), Nora Franceschini (1)

1. University of North Carolina School of Public Health, Chapel Hill, NC; 2. University of Washington School of Public Health, Seattle, WA; 3. University of Virginia School of Medicine, Charlottesville, VA; 4. University of Maryland School of Medicine, Baltimore, MD

Estimated glomerular filtration rate (eGFR) is a heritable measure of kidney function and when low defines presence of chronic kidney disease (CKD). There is substantial evidence for ancestry-specific genetic variants contributing to eGFR and CKD, but the role of low frequency and rare variants on eGFR variation among diverse populations is understudied. We conducted whole genome sequencing analysis of 23,732 participants of 10 multi-ethnic studies (European, African, East Asian, and Hispanic) cohort studies within the NHLBI TOPMed Project. Calibrated serum creatinine was used to estimate eGFR. We applied linear mixed models adjusted for age, sex, study, and ethnicity; the variance components were modeled by the genetic relationship matrix estimated from the whole genome sequencing data. In single-variant tests for variants with minor allele count 10, we identified five loci associated with eGFR at genome-wide significance (P<5.0 x 10-9), including three novel loci driven by low frequency intronic variants (PRKAA2, rs180996919, minor allele frequency [MAF] 0.04%, P=6.1 x 10-11; METTL8, rs116951054, MAF 0.09%, P=4.5 x 10-9; and MATK, rs539182790, MAF 0.05%, P=3.4 x 10-9). The PRKAA2 and METTL8 variants were present primarily in East Asians. Ancestry-specific allele frequencies estimation (ASAFE), using inferred local ancestry, suggested that the MATK variant

is an Amerindian variant. Two additional loci identified at genome-wide significance were driven by common variants at known loci (GATM and CDK12). Gene-based SKAT tests of variants selected through functional annotation identified an additional locus at MAF driven by multiple low frequency variants. In conclusion, analyses of deep sequenced TOPMed whole genome sequencing in multi-ethnic studies have identified three novel genome-wide significant loci for eGFR, including rare variants that are ancestry specific. We demonstrated that local ancestry can help to identify ancestry-specific replication samples. Our findings underscore the challenges to study low frequency variants in multi-ethnic studies and admixed populations when identified variants are specific to an ancestral group.

(19)

Putative Metabolic Sense: Anti-Sense Loci in the Global Lipids Dataset and Exome Chip Array [*On behalf of the Lipids Working Group*]

Erica L. Kleinbrink B.S., M.S.
Center for Molecular Medicine & Genetics, Wayne State University

Erica L. Kleinbrink(1),
Jerome I. Rotter(2),
Leonard Lipovich(1,3),
L. Adrienne Cupples(4),
Gina M. Peloso(4).

1. Center for Molecular Medicine & Genetics, Wayne State University, Detroit, MI; 2. Institute for Translational Genomics and Population Sciences, Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center, Torrance, CA, Department of Pediatrics, Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center, Torrance, CA; 3. Department of Neurology, School of Medicine, Wayne State University, Detroit, MI; 4. Department of Biostatistics, Boston University, School of Public Health, Boston, MA

INTRODUCTION: Sense:anti-sense gene pairs (genes with overlapping exons on opposite strands of genomic DNA) are a common class of complex loci that maintain active capacity for co-and cross-regulation at both a transcriptional and epigenetic as well as a post-transcriptional level. Anti-sense long noncoding RNA (lncRNAs) are often overlooked as the functional unit for genetic variants associated with metabolic traits.
METHODS: We annotated 2,236 variants on the exome chip array as being sense: anti-sense gene pairs. We then analyzed these variants for 4 lipids traits (TC, HDL-C, LDL-C, and triglycerides) obtained from ~300,000 individuals contributing to the Global Lipid Genetic Consortium exome array meta-analysis. Variants associated with lipids at $\alpha < 5x10\text{-}8$ were analyzed with transcriptomics to determine potential contribution of the lncRNA or the protein coding transcript to determine functionality at each complex locus. Identified SNPs, when applicable, were analyzed with metaserver pathogenicity predictions to estimate the coding transcript and overlapping noncoding transcript's contribution.

RESULTS: Loci: We identified seven loci annotated as sense:anti-sense gene pairs that reached genome-wide significance: rs3749147, rs2167079, rs1135999, rs1136001, rs1105879, rs2070959 and rs735396.
Two lipids-associated complex loci (rs3749147 with TC and TG, and rs2167079 with TG and HDL) were characterized as coding-coding sense: anti-sense loci (CCDC121:GPN1 and ACP2:NR1H3); both loci were devoid of known lncRNA genes. Two loci were characterized as containing sense anti-sense mRNAs; devoid of lncRNA genes (rs1136001 with TG and HDL, and rs1135999 with TG and HDL). Three loci identified (all TC and LDL associated) consisted of mRNA and lncRNA that are potentially co-expressed and anti-sense to each other (rs1105879, rs2070959, rs735396). Six of 7 significant SNPs qualified for further analysis with multiple pathogenicity predictions (rs3749147, rs2167079, rs1135999, rs1136001, rs1105879, rs2070959). We present an analysis by transcriptomics, review of chromatin features (Dnase I sites, TFBS, overlap of histone peaks H3K4me1, H3K4me3 and H3K27ac), and HeliScope CAGE single-molecule promoterome quantifications for all significantly associated lipid anti-exome loci suggesting a mechanism for these loci.

(20)

Mendelian randomization: current and future perspectives
[*On behalf of the Mendelian randomization Working Group*]

Dipender Gill MD
Department of Epidemiology and Biostatistics, Imperial College London

Dipender Gill (1) and Neil M Davies (2,3)

(1) Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, UK (2) Medical Research Council University of Bristol Integrative Epidemiology Unit, Bristol, UK (3) Bristol Medical School: Population Health Sciences, University of Bristol, Bristol, UK

Background: There has been an exponential increase in the popularity of Mendelian randomization in recent years, with a corresponding growth in the array of available methods and scenarios where they may be applied. Focusing on the remit of cardiovascular disease, this works aims to review current practice and summarise the directions of further development.
Methods: Searching PubMed up to 19 May 2019, we systematically searched and reviewed all Mendelian randomization studies of cardiovascular disease, considering the methods and exposures. A review of the literature and expert opinion was undertaken to consider directions of on-going development in Mendelian randomization methodology and application.

Results: A systematic review of the literature identified 258 Mendelian randomization studies performed up to 19 May 2019, with the majority focusing on circulating chemicals (142; 55%) and the minority considering social traits (19; 7%), cellular traits (14; 5%) and drug targets (13; 5%). Approaches have developed over time to increasingly explore possible bias related to pleiotropy. On-going developments in Mendelian randomization focus on considering timing of effects, population specificity, including for specific cells and tissues, mediating mechanisms, non-linear effects, integration of information from distinct sources including proteomic and gene expression data, statistical optimisation of instrument selection and modelling pleiotropy.
Conclusions: Mendelian randomization has evolved tremendously since its inception and continues to develop. Recent advances and increasing availability of large-scale data have the potential to vastly enhance the power and applicability of the technique to dissect complex biological mechanisms.

(21)

Investigating unintended drug effects: what can Mendelian randomization add? [*On behalf of the Mendelian randomization Working Group*]

Venexia Walker
University of Bristol

Venexia Walker (1,2), Neil Davies (1,2)

1. Medical Research Council Integrative Epidemiology Unit, University of Bristol, Bristol, UK; 2. Bristol Medical School: Population Health Sciences, University of Bristol, Bristol, UK

Background:
Little is known about the risk-benefit profile of a drug when it is licensed due to the comparatively small number of patients exposed during pre-approval clinical trials. Mendelian randomization is a causal inference method that has been used to assess the relationship between a wide range of exposures and outcomes. One advantage of Mendelian randomization is that study participants do not need to be exposed to a drug and so it offers an alternative form of evidence for drug effects. This can be particularly beneficial for triangulation of evidence – where results using different data and subject to different biases are considered together to strengthen causal inference.
Methods and results:
As an exemplar, we will consider recent work to proxy antihypertensive drug classes by Walker et al and Gill et al. In both cases, instruments for the analysis were chosen to mimic the protein targets of drugs, however the authors used different instrument selection procedures. Walker et al then applied two-sample Mendelian randomization to assess whether antihypertensive drug classes could be repurposed for the prevention of Alzheimer's disease, while Gill et al applied MR-PheWAS to perform a hypothesis-free scan for unintended drug effects in UK Biobank. As part of this exemplar, we will also consider how Mendelian randomization analyses, such as these, can then be combined with other evidence in a triangulation framework to increase the robustness of results.
Conclusions:
Mendelian randomization can be used to evaluate unintended drug effects and improve evidence about risk-benefit profiles of drugs. This method has several relevant applications in this setting – such as pre-approval screening to inform safety measures in randomized controlled trials using MR-PheWAS, or hypothesis-driven investigations to evaluate drug repurposing opportunities. Ultimately, triangulation of Mendelian randomization evidence with that from other sources could provide a relatively time- and cost-effective approach to reliably evaluate unintended drug effects.

(22)

Genetic drug target validation using Mendelian randomization [*On behalf of the Mendelian randomization Working Group*]

Amand Floriaan Schmidt, PhD
ICS, UCL

A F Schmidt$ [a,b,c], C Finan$ [a,b], M Gordillo-Marañón [a], F W Asselbergs [a,b,c], D F Freitag [d], R S Patel [a,b], B Tyl [e], S Chopade [a,b], R Faraway [ab,], M Zwierzyna [a,b,f], A D Hingorani [a,b]*

a. UCL, UK; b. UCL's BHF Research Accelerator centre; c. UMC Utrecht, the Netherlands; d. Bayer, Germany; e. Servier, Suresnes, France; f. Benevolent AI, UK; $ joint first authors

Abstract
Mendelian randomisation analysis has emerged as an important tool to elucidate the causal relevance of a range of environmental and biological risk factors for human disease. However, inference on cause is undermined if the genetic variants used to instrument a risk factor of interest also associate with other traits that open alternative pathways to the disease (horizontal pleiotropy). We show how the 'no horizontal pleiotropy assumption' in MR analysis is strengthened when proteins are the risk factors of interest. Proteins are the proximal effectors of biological processes encoded in the genome, and are becoming assayable on an -omics scale. Moreover, proteins are the targets of most medicines, so Mendelian randomization (MR) studies of drug targets are becoming a fundamental tool in drug development. To enable such studies we introduce a formal mathematical framework that contrasts MR analysis of proteins with that of risk factors located more distally in the causal chain from gene to disease. Finally, we illustrate key model decisions and introduce an analytical framework for maximizing power and elucidating the robustness of drug target MR analyses.

(23)

Identification and validation of genetic variants as instruments for studying drug effects [*On behalf of the Mendelian randomization Working Group*]

Marios Georgakis, MD
Institute for Stroke and Dementia Research, LMU Munich

Marios K. Georgakis(1)

1. Institute for Stroke and Dementia Research, LMU Munich, Munich, Germany

Background: In the era of genome-wide association studies, Mendelian randomization (MR) has emerged as a powerful tool for investigating causal relationships between risk factors and disease outcomes. MR has been further proposed to provide a framework for studying the effects of pharmacological interventions by using data from observational studies. Using this approach, MR studies exploring specific drug targets have shown results comparable to those derived from clinical trials. Yet, there is no consensus on the methodological pipeline for studying drug effects with MR.
Methods and Results: This talk will provide an overview of the current literature on the methodological and conceptual challenges related to the selection of genetic variants to be used as instruments (proxies) in MR analyses focused on drug effects. There will be a focus on the selection of variants at the loci of genes encoding protein drug targets and the phenotype that can be used to proxy the effects of the drugs (e.g. expression and protein levels versus a phenotype corresponding to a downstream effects of the drug). We will discuss the options for validating the selected instruments by exploring their effects on alternative downstream effectors of the drug and by comparing the MR results with findings from clinical trials. Finally, we will complement these concepts with empirical data regarding the use of genetic proxies for studying the effects of anti-inflammatory and anti-hypertensive medications on vascular endpoints.
Conclusions: Although several lines of evidence speak for the utility of MR for studying drug effects, there is currently no systematic methodology for identifying valid genetic proxies. Given that drug targets with genetic support are more likely to show clinical efficacy in randomized trials, developing the framework for sufficiently predicting drug effects using genetic data is of substantial importance.

(24)

Distinct metabolomic signatures of central obesity in the Atherosclerosis Risk in Communities (ARIC) Study [*On behalf of the Metabolomics Working Group*]

Kristin oung, PhD
Epidemiology, UNC Chapel Hill

Kristin L. Young (1), Heather M. Highland (1), Bing Yu (2), Megan Grove-Gaona (3), Eric Boerwinkle (2), Kari E. North (1)

1. University of North Carolina at Chapel Hill, Department of Epidemiology, Chapel Hill, NC; 2. University of Texas Health Science Center at Houston, Department of Epidemiology, Human Genetics and Environmental Sciences, Houston, TX; 3. University of Texas Health Science Center at Houston, Human Genetics Center, Houston, TX

While the prevalence of obesity has more than doubled in the US, not all obese individuals have the same risk for adverse health outcomes. As central obesity is more metabolically active, metabolites may provide insights into mechanisms leading to disease. We hypothesize that central obesity (measured as waist-to-hip ratio adjusted for BMI) has an identifiable metabolic signature distinct from metabolites associated with BMI. GWAS of obesity associated metabolites (mGWAS) can identify associations linked to metabolic transporters, regulators, or enzyme coding genes, allowing for discovery of the best biological candidates in known GWAS regions, as well as identifying novel central obesity genes. Our study leverages data in the Atherosclerosis Risk in Communities (ARIC) Study (N=4,032), a prospective, longitudinal cohort study of cardiovascular disease in European (EA) and African Americans (AA), to identify metabolomic signatures of central obesity.
We estimated associations between baseline natural log (ln) transformed WHRadjBMI and 245 baseline standardized plasma metabolites, using ancestry and sex-stratified additive linear models, followed by meta-analysis in SAS 9.3. We identified 7 metabolites (all lipids) significantly associated with increased lnWHRadjBMI (p=2.04E-4 (0.05/245*2 traits), as well as 5 metabolites (4 amino acids/peptides and 1 lipid) associated with decreased lnWHRadjBMI. None of these metabolites were associated with BMI.
Metabolites significantly associated with lnWHRadjBMI were carried forward for mGWAS (n=12), followed by meta-analysis in METAL. An intronic variant in ALDH1A2 is associated with increased 1-stearoylglycerophosphoethanolamine levels in all strata (beta=0.24, SE=0.03, p=8.18E-22, EAF=0.39); this variant was previously associated with plasma lipid levels, including HDL and triglycerides. Two variants near GGT1 are associated with -glutamylleucine (beta=0.33, SE=0.05, p=6.23E-11, EAF=0.82) and -glutamylvaline (beta=0.50, SE=0.08, p=6.55E-10, EAF=0.93) levels in AA only; GGT1 codes for an enzyme involved in protein metabolism. These results highlight the potential for metabolomics to identify new central obesity signals and refine obesity phenotypes.

(25)

A multi-omics study of circulating phospholipid markers of blood pressure [*On behalf of the Metabolomics Working Group*]

Jun Liu
Nuffield Department of Population Health

Jun Liu(1,2), Paul S. de Vries(1,3)*, Fabiola Del Greco(4), Åsa Johansson(5), Katharina E. Schraut(6,7), Caroline Hayward(8), Oscar. H. Franco(1,9), Andrew A. Hicks(4), Veronique Vitart(8), Igor Rudan(7), Harry Campbell(7), Ozren Polašek(10), Peter P. Pramstaller(4), James F. Wilson(7,8), Ulf Gyllensten(5), Cornelia M. van Duijn(1,2), Abbas Dehghan(1,11)§, and Ayşe Demirkan(1,12)§.*

*1. Department of Epidemiology, Erasmus MC, University Medical Center, Rotterdam, the Netherlands; 2. Nuffield Department of Population Health, Oxford University, Oxford, United Kingdom; 3. Human Genetics Center, Department of Epidemiology, Human Genetics, and Environmental Sciences, School of Public Health, University of Texas Health Science Center at Houston, Houston, United States of America; 4. Institute for Biomedicine, Eurac Research, Affiliated Institute of the University of Lübeck, Bolzano, Italy; 5. Department of Immunology, Genetics and Pathology, Science for Life Laboratory, Uppsala University, Uppsala, Sweden; 6. Centre for Cardiovascular Sciences, Queen's Medical Research Institute, University of Edinburgh, Edinburgh, Scotland, UK; 7. Centre for Global Health Research, Usher Institute for Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, Scotland, UK; 8. Medical Research Council Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Western General Hospital, Edinburgh, Scotland, UK; 9. Institute of Social and Preventive Medicine (ISPM), University of Bern, Bern, Switzerland; 10. Department of Public Health, University of Split, School of Medicine, Split, Croatia; 11. Department of Epidemiology and Biostatistics, MRC-PHE Centre for Environment & Health, School of Public Health, Imperial College London, UK; 12. Section of Statistical Multi-omics, Department of Clinical and Experimental research, University of Surrey, Guildford, Surrey, UK.*These authors contributed equally to this work. § These authors jointly supervised this work.*

Background High-throughput techniques allow us to measure wide-range of phospholipids in high resolution with detailed characterization. Such characterized circulating phospholipids may provide insight into mechanisms of increased blood pressure. Therefore, we aimed to conduct an in-depth multi-omics study of the role of phospholipids with systolic (SBP) and diastolic blood pressure (DBP).
Methods Using 3,971 participants from five genetically isolated European populations, we conducted an electrospray ionization tandem mass spectrometry to quantify 151 plasma lipids in plasma. Linear regression was used to associate blood pressure with each of the lipids in both discovery analysis (n = 2,786) and replication analysis (n = 1,185). Furthermore, we performed a combined analysis and adjusted for a series of extra confounders. The predictive effect of selected lipids on incident hypertension was explored by Cox regression. The association analysis of these lipids with 28 available cardiometabolic traits was performed in the cross-sectional study. A two-sample-based bidirectional Mendelian Randomization (MR) approach was performed for the significant associations. Phenome-wide association study (pheWAS) was also performed.
Results We discovered and replicated six phosphatidylethanolamines (PEs, PE 38:3, PE 38:4, PE 38:6, PE 40:4, PE 40:5 and PE 40:6) and two phosphatidylcholines (PCs, PC 32:1 and PC 40:5) associated with systolic or diastolic blood pressure. The results were robust after adjusting for further covariates. The selected phospholipid profile predicted incident hypertension with area under the receiver operator characteristics (ROC) curve (AUC) of 0.61 in a 14 year follow-up. A very strong association of these lipids to triglycerides and other cardiometabolic associations were detected. The MR approach shows that the genetically increased SBP level is borderline significantly associated with the increase of circulating PE 40:5 (P-value = 0.06) after excluding the pleiotropic effect from triglycerides. These lipids are also sharing multiple genetic determinants with triglycerides, HDL subfractions, very small VLDL subfractions, some fatty acids, blood cell counts (white, red and platelet) and pulse rate, among which triglycerides, very small VLDL subfractions, fatty acids, small HDL subfractions and Apolipoprotein A1 were associated with blood pressure and related phospholipids.
Conclusion We identified eight circulating phospholipid molecules which are associated with blood pressure and also have a predictive value for incident hypertension. These eight phospholipids are likely to share genetic pathways which link them to lipoproteins, blood cell counts and pulse rate. Further causation studies on the effect of blood pressure over phospholipid metabolism are suggested.

(26)

Association of Tryptophan, Tyrosine and Branched Chain Amino acids metabolites with abdominal obesity and cardio-metabolic risk factors [*On behalf of the Metabolomics Working Group*]

Lidia Ximena Orozco Ruiz Ph.D.(c)
Population Health Sciences, German Center for Neurodegenerative Diseases

Ximena Orozco-Ruiz(1), Estrada-Leon Satiago(1), Monique M.B. Breteler(1,2)

1. German Center for Neurodegenerative Diseases (DZNE), Department of Population Health Sciences, Bonn, Germany; 2.University of Bonn, Institute for Medical Biometry, Informatics and Epidemiology (IMBIE), Bonn, Germany

Background: Obesity induces a cascade of metabolic disturbances associated with the development of obesity-related diseases. These processes mainly operate in adipose tissue and can regulate tryptophan, branched-chain amino acids (BCAAs) and tyrosine metabolism. It has been suggested that visceral and abdominal adipose tissue (VAT and SAT) are differently associated with cardio-metabolic risk factors of obesity-related diseases. To what extent this is reflected in

different metabolomic profiles of VAT and SAT and in their relation with cardio-metabolic risk factors is not fully elucidated.
Objective: To identify metabolites from tryptophan, BCAAs, and tyrosine that are influenced by VAT and SAT, and to evaluate their association with cardio-metabolic risk factors.
Methods: Cross-sectional analysis of data from Rhineland Study participants (n=893, 56% women, mean age= 53 y). Targeted LC-MS/MS based metabolomic profile was conducted on 15 tryptophan, 3 BCAAs and 5 tyrosine metabolites. VAT and SAT volumes were obtained with abdominal MR images. Cardio-metabolic risk factors considered were: systolic and diastolic blood pressure, triglycerides, HDL cholesterol (HDL-C), C-reactive protein (CRP), glucose and insulin levels. We first quantified the relation of VAT and SAT with metabolite levels. Subsequently, we evaluated how metabolites which levels depended on volumes of VAT or SAT were related with cardio-metabolic risk factors. We used multivariable linear regression analyses, adjusting for age and sex. Metabolites were logarithmically transformed and standardized to a mean of 0 and SD of 1.
Results: Larger volumes of both VAT and SAT were associated with higher levels of BCAAs (leucine, isoleucine and valine), tryptophan (3-hydroxyanthranilic acid, kynurenine and kynurenic acid) and tyrosine (phenylalanine and tyrosine) metabolites and with lower levels of indole-3-propionic acid. Higher levels of BCAAs and tryptophan metabolites were strongly associated with higher levels of triglycerides, insulin, glucose, and CRP, and lower levels of HDL-C. Tyrosine metabolites were only positively related to insulin levels.
Conclusions: VAT and SAT were significantly associated with Tryptophan, BCAAs and tyrosine metabolites, which were significantly associated with cardio-metabolic risk factors. Therefore, these metabolites may provide a better insight into the biologic mechanisms that underlie the relation of obesity with cardio-metabolic disease.

(27)

Milk Intake, Host LCT Genotype and Gut Bifidobacteria in relation to Obesity: Results from the Hispanic Community Health Study/Study of Latinos (HCHS/SOL) [*On behalf of the Microbiome Working Group*]

Jee-Young Moon
Epidemiology and Population Health, Albert Einstein College of Medicine

Jee-Young Moon(1), Zheng Wang(1), Mykhaylo Usyk(1), Yoshiki Vázquez(2), Carmen Isasi(1), Yasmin Mossavar-Rahmani(1), Daniel McDonald(2), Daniela Sotres-Alvarez(3), Martha Daviglus(4), Jianwen Cai(3), Rob D Knight(2), Robert Kaplan(1,5), Qibin Qi(1), Robert Burk(1)

1. Albert Einstein College of Medicine, Bronx, NY; 2. University of California-San Diego, La Jolla, CA; 3. University of North Carolina at Chapel Hill, Chapel Hill, NC; 4. University of Illinois at Chicago, Chicago, IL; 5. Fred Hutchinson Cancer Research Center, Seattle, WA

Background
Previous studies in Europeans found that gut Bifidobacterium genus was elevated in lactose intolerant individuals. However, the interrelationship between milk intake, host lactase (LCT) genotypes and different Bifidobacteria species, and their implications for human diseases have not been well-studied.
Methods
This study included 2141 participants aged 23-83 years with microbiome, genetics and 24 hour dietary recall data from the HCHS/SOL, a population-based cohort of US Hispanics/Latinos. Stool samples were analyzed by shotgun metagenomics sequencing.
Results
A GWAS on Bifidobacterium genus confirmed the genome-wide significant LCT locus (rs4988235, P=3.7×10−21), and 3 major Bifidobacteria species were also significantly associated with LCT (B. longum P=4.4×10−17, B. adolescentis P=2.9×10−13, B. bifidum P=3.1×10−5). 779 adults with lactose tolerant genotype (GG) had higher milk intake but lower Bifidobacteria compared to 1372 lactose intolerant adults (AA/AG). There was a significant interaction between milk intake and LCT variant on Bifidobacterium (Pint=0.0003), with a positive association between milk intake and Bifidobacterium only in lactose intolerant but not in tolerant people. The association of Bifidobacterium with obesity was also modified by LCT genotype. In lactose intolerant people, higher Bifidobacterium was associated with lower BMI, waist circumference and fat mass index, whereas in tolerant people higher Bifidobacterium was associated with higher obesity measures. Species-level analyses indicated that B. longum may account for the positive association with milk intake and inverse associations with obesity measures in lactose intolerant people, while other major species, B. adolescentis and B. bifidum, may contribute to the positive association with obesity measures in lactose tolerant people.
Conclusion
Our data indicate that diet and LCT genotype may alter the Bifidobacterium composition in gut, with different associations with obesity.

(28)

Urinary hippurate, a microbiota-generated metabolite of dietary components, is negatively associated with hypertension in CARDIA [*On behalf of the Microbiome Working Group*]

Katie Meyer, MPH, ScD
Department of Nutrition & Nutrition Research Institute, Univ. of No. Carolina at Chapel Hill

Katie A. Meyer (1), Torben Kimhofer (2), Anju Lulla (1), James M. Shikany (3), Joseph F. Petrosino (4), Luke Whiley (2), Lenore J. Launer (5), Elaine Holmes (2)

1. Univ of NC-Chapel Hill, Kannapolis, NC; 2. Murdoch Univ, Perth, Western Australia; 3. Univ of AL-Birmingham, Birmingham, AL; 4. Baylor College of Medicine, Houston, TX; National Institute of Aging, Baltimore, MD.

Background: An inverse association between urinary hippurate and blood pressure was found in the INTERMAP Study, but this finding has not been replicated. Hippurate is produced, in part, through gut microbial metabolism of dietary polyphenols. We test the association between hippurate and hypertension in an independent sample, CARDIA, and quantify associations among postulated pathway components, including gut microbial diversity and diet quality.
Methods: This is a cross-sectional analysis of data from n=480 CARDIA participants who participated in a microbiome study at the Yr. 30 exam (2015-16). We conducted 1H NMR on fasting urine samples, and quantitated hippurate by integrating spectral intensities at known locations. Hippurate was urine creatinine-normalized and ln-transformed for analysis. Kraken2 was used for taxonomic classification of reads from whole-metagenomics sequencing (stool DNA). Species-level diversity was measured using the Shannon index. Hypertension was defined as SBP≥140, DBP≥90, or antihypertensive use. Diet was assessed using a 26-item frequency-based survey. A diet quality score was created by summing over consumption quartiles of foods scored positively or negatively for health; higher scores indicate higher quality. Regression models were adjusted for age, sex, race, education, physical activity, smoking, diet quality, and species diversity.
Results: In multivariable-adjusted regression analysis, hippurate (SD-unit increase) was negatively associated with hypertension [RR: -0.28 (-0.47, -0.10)]. Diet quality was positively associated with microbial diversity [standardized beta coefficient (β): 0.54 (0.010-1.07)] and urinary hippurate [β: 0.20 (0.060, 0.34)]; microbial diversity was positively associated with urinary hippurate [β: 0.23 (0.14, 0.31)].
Conclusions: Our results replicate INTERMAP's finding that hippurate may be negatively associated with hypertension. These results were robust to covariate adjustment, including dietary quality and microbial diversity. Associations among diet, microbial diversity, and hippurate are consistent with a diet-microbiota-metabolite pathway to hypertension. Our findings contribute to a growing literature linking gut microbial metabolites of dietary components to cardiovascular disease risk factors.

(29)

Lower Gut Bacterial Diversity in Non-alcoholic Fatty Liver Disease: Results from the Hispanic Community Health Study/Study of Latinos (HCHS/SOL) [*On behalf of the Microbiome Working Group*]

Jessica Williams-Nguyen
Public Health Sciences, Fred Hutchinson Cancer Research Center

Jessica Williams-Nguyen(1),
Mark H Kuniholm(2),
Zheng Wang(3),
Jin Choul Chai(3),
Jee-Young Moon(3),
Tao Wang(3),
Carmen Isasi(3),
Jorge Kizer(4),
Scott Cotler(5),
Martha Daviglus(6),
Rob Knight(7),
Robert Burk(3),
Qibin Qi(3),
Robert Kaplan(1,3)

1 Fred Hutchinson Cancer Research Center, Seattle, WA, USA; 2 University at Albany, Rensselaer, NY, USA; 3 Albert Einstein College of Medicine, Bronx, NY, USA; 4 University of California San Francisco, San Francisco, CA, USA; 5 Loyola University Medical Center, Maywood, IL, USA; 6 University of Illinois at Chicago, Chicago, IL, USA; 7 University of California, San Diego, La Jolla, CA, USA

Nonalcoholic fatty liver disease (NAFLD) is associated with obesity and a leading cause of chronic liver disease with higher prevalence in Hispanics/Latinos compared to non-Hispanic whites. Gut microbiome (GMB) composition is associated with NAFLD and obesity in animal and human studies, but data in Hispanic/Latino populations are limited. This analysis included 2587 Hispanic/Latino adults in SOL with fecal samples. The bacterial GMB was characterized via shotgun metagenomic sequencing and taxonomic classification using the SHOGUN pipeline with RefSeq 82 prokaryotic genome database. NAFLD was defined by gender-specific liver function test thresholds. Alpha and beta diversity were compared between NAFLD/obesity groups using multinomial logistic regression and PERMANOVA, respectively. The Wilcoxon rank-sum test was used to compare the difference in abundance for the top 15 species. There were 395 NAFLD cases, and prevalence of obesity (BMI ≥ 30) was 58% and 42% among those with and without NAFLD, respectively. Higher alpha diversity (Shannon index) was associated with lower odds of NAFLD with (OR = 0.58, p = 0.0005) or without obesity (OR = 0.70, p = 0.04)

compared to participants with neither condition after adjusting for potential confounders. Higher Shannon index was also associated with lower odds of obesity in the absence of NAFLD (OR = 0.78, p = 0.01). Beta diversity (Bray-Curtis) did not differ by NAFLD/obesity groups (p = 0.78) or by NAFLD alone (p = 0.30). Among the 15 most abundant species, 4 (Bacteroides uniformis, Odoribacter splanchnicus, Oscillibacter sp. ER4, and Alistipes shahii) had lower abundance in those with NAFLD compared to those without NAFLD irrespective of obesity. Bacterial alpha diversity but not beta diversity was independently associated with the related conditions, NAFLD and obesity, in US Hispanic/Latino adults. Future work will explore associations between NAFLD and GMB functional capacity considering host genetic variants.

(30)

Metabolomics insights into osteoporosis through association with bone mineral density [*On behalf of the Musculoskeletal Working Group*]

Xiaoyu Zhang, Ph.D.
Boston University

Xiaoyu Zhang(1),
Hanfei Xu(1),
Michelle T. Long(2),
Ramachandran S. Vasan(3),
Douglas P. Kiel(4),
Ching-Ti Liu(1)

1. Department of Biostatistics, Boston University School of Public Health, Boston, MA, 02118, USA; 2. Section of Gastroenterology, Boston Medical Center, Boston University School of Medicine, Boston, MA, 02118, USA; National Heart, Lung, and Blood Institute's and Boston University's Framingham Heart Study, Framingham, MA 01702, USA; 3. National Heart, Lung, and Blood Institute and Boston University's Framingham Heart Study, Framingham MA 01702, USA; Section of Preventive Medicine and Epidemiology, Evans Department of Medicine, Boston University School of Medicine, Boston, MA 02118, USA; Whitaker Cardiovascular Institute and Cardiology Section, Evans Department of Medicine, Boston University School of Medicine, Boston, MA 02118, USA; 4. Hinda and Arthur Marcus Institute for Aging Research, Hebrew SeniorLife, Boston, MA 02131, USA; Harvard Medical School, Boston, MA 02115, USA; Broad Institute of MIT and Harvard, Boston, MA 02142, USA; Beth Israel Deaconess Medical Center, Boston, MA 02215, USA

Abstract
Background. Osteoporosis, the most common metabolic bone disease, is characterized by low bone mineral density (BMD), which is the strongest risk factor for fracture. Identifying a metabolite profile associated with BMD may provide insights into osteoporosis. The aim of this study is to identify predictive metabolomic markers associated with BMD in a general population.
Methods. We assessed 209 plasma metabolites by LC-MS/MS in 1,552 participants, aged 30-82 years, in the Framingham Study Offspring Cohort. BMD was measured at the femoral neck (FN) and lumbar spine (LS) using dual energy x-ray absorptiometry, and major osteoporotic fractures were assessed based on 22-year follow-up after metabolites measured. We implemented a least absolute shrinkage and selection operator (LASSO) to select BMD-associated metabolite concentrations with internal validation. We then assessed their capability to predict osteoporotic fracture by logistic regression and identified the pathways enriched by the BMD-associated metabolites. We further performed two-sample Mendelian randomization to explore the causal relationship between identified metabolites and BMD.
Results. We identified 27 metabolites that were associated with FN BMD or LS BMD. Incorporating them improved the prediction of osteoporotic fracture risk beyond conventional risk factors which include sex, age, body mass index, menopausal status, and smoking status (AUC=0.74 for the model with identified metabolites and risk factors vs AUC=0.7 for the model with risk factors alone, p=0.001). The glycine, serine, and threonine metabolism pathway including four identified metabolites (serine, glycine, creatine, dimethylglycine) was significantly enriched with an FDR adjusted p-value=0.028. Furthermore, three metabolites [glycine, Phosphatidylcholine (PC), and Triacylglycerol (TAG)] were found to be causally negative associated with FN-BMD while two metabolites, PC and TAG, were also causally negative associated with LS-BMD.
Conclusion. We identified 27 metabolites associated with BMD in a general population. Three of them were causally associated with BMD, which requires further experimental validation and may be clinically valuable for early detection of low BMD and osteoporotic fracture. Our findings may provide a novel insight into the pathogenesis of osteoporosis and further underlying the importance of metabolomics profiling in predicting osteoporotic fracture.

(31)

Pharmacogenetic and MicroRNA Effects of Beta Blocker Association with Increased Bone Mineral Density in Humans [*On behalf of the Musculoskeletal Working Group*]

Christine Lary, Ph.D.
Center for Outcomes Research and Evaluation, Maine Medical Center Research Institute

Kathleen T. Nevola(1)
Douglas P. Kiel(2,3)
Andrew R. Zullo(4,5)
Ines Foessl(6)
Barbara Obermayer-Pietsch(6)
Katherine J. Motyl(7)
Christine W. Lary(7)

1.Tufts University, Boston, MA; 2. Harvard Medical School, Boston, MA; 3. Hebrew SeniorLife, Boston, MA; 4. Brown University, Providence, RI; 5. Rhode Island Hospital, Providence, RI; 6. Medical University of Graz, Graz, Austria; 7. Maine Medical Center Research Institute, Portland, ME

Osteoporosis is a debilitating and costly disease. Recent studies have shown that beta blocker (BB) users have higher bone mineral density (BMD) and decreased risk of fracture as compared to non-users. This association is thought to be due to suppression of adrenergic signaling in osteoblasts, which leads to increased BMD in rodent models; however the mechanism in humans is unknown. We used the Framingham Heart Study/Osteoporosis Substudy to investigate pharmacogenetic effects from variants in the β1AR and β2AR genes and to discover potential microRNA (miRNA) mechanisms of the effect of BB use on BMD. Framingham Offspring cohort participants had clinical data, dual-energy X-ray absorptiometry (DXA) scans, miRNA and mRNA profiling of whole blood, and high density genotype data imputed to the Haplotype Reference Consortium (HRC). We found nine miRNAs associated with BB use and increased BMD as well as a miRNA network associated with BMD and BB use containing two of these nine miRNAs, miR-19a-3p and miR-186-5p, the first of which has been previously shown to directly target the β1AR mRNA transcript. To validate these miRNA associations, we show that these two miRNAs have significantly higher expression in individuals without incident fracture compared to those with incident fracture in an external data set. We have found rs1801252, a serine->glycine missense variant at positon 49 in ADRB1 to show a significant interaction with BB use on femoral neck BMD in women (p=0.048). Several variants in ADRB2 have also shown significant interaction effects with BB use in women and men. These pharmacogenetic and miRNA associations for BB use on BMD provide a starting point for understanding molecular mechanisms involved and for clinical biomarker discovery.

(32)

Epigenome-wide meta-analysis of cerebral white matter hyperintensities on MRI [*On behalf of the Neurology Working Group*]

Yunju Yang, MPH
Department of Epidemiology, University of Texas Health Science Center at Houston

Yunju Yang(1), Qiong Yang(2), Maria Knol(3), Hieab Adams(3), Sudha Seshadri(2,4), Myriam Fornage(1)

1. The University of Texas Health Science Center at Houston, Houston, TX, USA;2. Boston University, Boston, MA 02118, USA;3. Erasmus MC University Medical Center, Rotterdam, The Netherlands;4. University of Texas Health Sciences Center, San Antonio, TX, USA

Cerebral white matter hyperintensities (WMH) on MRI are an indicator of cerebral small vessel disease, a major risk factor for vascular dementia and stroke. DNA methylation may contribute to the molecular underpinnings of WMH, which are highly heritable. We performed a meta-analysis of 11 epigenome-wide association studies in 6,019 middle-aged to elderly subjects, who were free of dementia and stroke and were of African (AA) or European (EA) descent. In each cohort, association between WMH volume and each CpG was tested within ancestry using a linear mixed model, adjusted for age, sex, total intracranial volume, white blood cell count, technical covariates, BMI, smoking and blood pressure (BP). To detect differentially methylated regions (DMRs), we also calculated region-based p-values accounting for spatial correlations among CpGs.
No individual CpG reached epigenome-wide significance, but suggestive novel associations were identified with cg17577122 (CLDN5, P=2.39E-7), cg24202936 (LOC441601, P=3.78E-7), cg03116124 (TRIM67, P=6.55E-7), cg04245766 (BMP4, P=3.78E-7) and cg06809326 (CCDC144NL, P=6.14E-7). Gene enrichment analyses implicated pathways involved in regulation of cell development and differentiation, especially of endothelial cells. We identified 11 DMRs (SidakP<0.05) including two mapping to BP-related genes (HIVEP3, TCEA2). The most significant DMRs mapped to PRMT1, a protein arginine methyltransferase involved in glioblastomagenesis (P=7.9E-12), and to CCDC144NL, a coiled-coil domain-containing protein P=1.6E-11). Genes mapping to DMRs were enriched in biological processes related to lipoprotein metabolism and transport. Bi-directional Mendelian randomization analysis showed that methylation at cg06809326 influenced WMH burden (OR[95% CI]=1.7[1.2-2.5], p-value=0.001), while WMH burden did not influence methylation at this CpG (OR[95% CI]=1.0[0.95, 1.04], p=0.89). Transcriptome association analyses showed that increased methylation at cg06809326 is associated with lower expression of the CHKB gene, which encodes a key enzyme of the phospholipid biosynthesis (P=7.1E-7).
In conclusion, we identified novel epigenetic loci associated with WMH which may provide new clues about its pathology.

(33)

The comprehensive gene-mapping study on MRI-derived extremes of cerebral small vessel disease reports role of TRIM47. [*On behalf of the Neurology Working Group*]

Aniket Mishra, PhD
Bordeaux Population Health Research Center, Bordeaux, France

Aniket Mishra(1)
W T Longstreth Jr(2)
Myriam Fornage(3)
Sudha Seshadri(4,5)
Stephanie Debette(1,5,6)

1.Univ. Bordeaux, Inserm, Bordeaux Population Health Research Center, team VINTAGE, UMR 1219, F-33000 Bordeaux, France; 2.Department of Neurology and Department of Epidemiology, University of Washington, Seattle, WA, USA; 3. The University of Texas Health Science Center at Houston, Houston, TX, United States; 4. Glenn Biggs Institute for Alzheimer's and Neurodegenerative Diseases, University of Texas Health Sciences Center, San Antonio, Texas, USA; 5. Department of Neurology, Boston University School of Medicine, Boston, MA, USA; 6. CHU de Bordeaux, Department of Neurology, F-33000 Bordeaux, France

Background: Cerebral Small Vessel Disease (SVD) is a leading cause of stroke and the main vascular contributor to dementia risk. We conducted multi-cohort genome-wide association (GWAS) and whole exome sequencing (WES) studies on extremes of SVD (extreme-SVD), based on the distribution of white matter hyperintensities (WMH) and brain infarcts on cranial MRI. We aimed to identify causal genes underlying extreme-SVD risk loci and to facilitate identification of biotargets.
Methods: Fifteen population-based cohorts yielded 39,664 participants for GWAS, 15,930 for WES, and 4,783 for exome-chip. We performed inverse-variance weighted meta-analysis of GWAS using METAL and of WES study using SeqMeta. We used: VEGAS2 and SKAT-O approaches for gene-based association testing; the summary-based Mendelian randomization approach to test association between brain eQTL and extreme-SVD risk variants; and in-house scripts to screen human exomes for loss of function, stop-gain-allele carriers. Moreover, we used mice models of aging and chronic brain hypoperfusion for in-vivo characterization of putative biotargets.
Results: GWAS identified 11 risk loci associated with extreme-SVD: seven WMH risk loci (2p16.1, 2p21, 2q33.2, 6q25.1, 10q24.33, 17q21.31, 17q25.1) and four novel loci (2q32.1, 12q24.11, 16q12.1, 16q24.1). WES study identified significant association of functional variants in EFEMP1 (2p16.1), TRIM47 and FBF1 (17q25.1) genes. Corroborating evidences from brain eQTL data and profiling of human stop-gain carriers identified DCAKD at 17q21.31 and TRIM47 at 17q25.1 as the most plausible putative biotargets. Through in-vivo experiments in mice, we demonstrate that Trim47 expresses predominantly in blood-brain barrier compared to other brain tissues and is upregulated in ageing mice and in a chronic brain hypoperfusion model mimicking SVD.
Conclusion: We report a multi-cohort gene-mapping study on extreme-SVD, identifying novel genetic risk loci and functional variant and providing in-silico and experimental evidence for a putative causal role of TRIM47.

(34)

Circulating metabolites associated with brain MRI markers of Alzheimer's Disease [*On behalf of the Neurology Working Group*]

Ruiqi Wang, MS
Department of Biostatistics, Boston University

Ruiqi Wang(1), Claudia L. Satizabal(1,2), Qiong Yang(1), Sudha Seshadri(1,2)

1. Boston University, Boston, MA; 2. University of Texas Health Sciences Center, San Antonio, TX

Background
Prior research has identified metabolites associated with general cognitive function and Alzheimer's Disease (AD). However, more research is needed to identify additional metabolites related to AD endophenotypes to better understand the earliest pathological processes. We aimed to identify metabolites associated with brain MRI markers of AD, including total brain volume, lateral ventricular volume, and hippocampal volume in the CHARGE consortium.
Method
The Framingham Heart Study Offspring and Third-Generation cohorts were used to pilot this project. The sample size varied between 160 and 2,145 participants based on the availability of metabolites. We used linear mixed effect models to assess the association between metabolites and MRI phenotypes. Brain MRI measures were introduced as the percentage of total intracranial volume to correct for head size. Our primary model was adjusted for age, sex, BMI, and use of lipid-lowering medications. In secondary analyses, we ran the same model stratifying for APOE4 status subgroups (e4 carriers vs. non-carriers). After correction for multiple testing, significance was set at $P < 2.09e\text{-}4$.
Results
We found that higher levels of cotinine, fumarate-malate-valine and lactate were associated with lower total brain volumes ($P < 8.13e\text{-}5$). In subgroup analyses, we observed that higher levels of anthranilate were related to lower brain volumes in APOE-e4 carriers. Further, only in APOE-e4 non-carriers, higher phosphatidylcholine (32:1) levels were related to lower brain volumes ($P < 1.52e\text{-}4$) and higher levels of urate were related to lower hippocampal volume ($P < 1.11e\text{-}4$).

Conclusion
Several metabolites were associated with total brain and hippocampal volumes. Remarkably, some of these have been previously related to cognitive function and AD, adding to the consistency of our findings and highlighting the power of endophenotype research. Expansion to include additional brain MRI markers and replication in other CHARGE cohorts are currently underway.

(35)

Genetic influences on early neurological instability after acute ischemic stroke: GENISIS Results [*On behalf of the Neurology Working Group*]

Laura Ibanez, PhD
Department of Psychiatry, Washington University in Saint Louis

Laura Ibanez (1)
Laura Heitch(1)
Caty Carrera(2)
Daniel Strbian(3)
Agnieszka Slowik(4)
Antonio Arauz(5)
Israel Fernandez(2)
Joan Montaner(2)
Jin-Moo Lee(1)
Carlos Cruchaga(1)

1. Washington University School of Medicine, Saint Louis, MO, USA; 2. Vall d'Hebron Institute of Research, Barcelona, Spain; 3. Helsinki University Hostpital, Helsinki, Finland; 4. Jagiellonian University

Medical College, Krakow, Poland; 5. Instituto Nacional de Neurologia y Neurocirurgia, Tlalpan, Mexico

1.Objective: To identify genetic loci related to early neurological changes after ischemic stroke using a genome wide association study (GWAS).
2.Background: Following acute ischemic stroke (AIS) onset, neurological deficits can be highly unstable within the first 24 hours. We hypothesized that this early change, defined here as ΔNIHSS24h (NIHSS baseline – NIHSS24hours), could serve as a quantitative phenotype to capture the genetic architecture of ischemic brain injury.
3.Design/Methods: AIS patients were prospectively enrolled between 2008-2019 at 7 sites (St Louis, Barcelona, Helsinki, Krakow, Korea, Costa Rica and Mexico). NIHSS scores were obtained within 6 hours and again at 24 hours after stroke onset. Genome-wide genotyping was generated for rare and common variants, imputing up to 6 million single nucleotide polymorphisms (SNPs) for all subjects (N=5,487). ΔNIHSS24h was used as a quantitative trait in an association model, with covariates that included baseline NIHSS, age, sex, PCA1, PCA2 and genotyping round. All samples were analyzed together using Plink1.9. They are also analyzed by site and then, we used Mantra to perform a trans-ethnic meta-analysis.
4.Results: We found one GWAS significant locus in the joint-analysis (rs16838295 – p=2.84×10-08) and two suggestive loci (rs114248865 and rs10807797). After deep functional annotation and Mendelian randomization anlayses, the top association in chromosome 2 was driven by ADAM23.
5.Conclusions: Early neurological outcome after AIS is strongly influenced by genetics. ADAM23 is a neuronal transmembrane protein involved in the bridging of pre- and post-synaptic membranes. It appears to play a role in modulating trans-synaptic excitability, suggesting a role in excitotoxicity in the setting of AIS.

(36)

Admixture-mapping identifies genomic regions associated with neurocognitive function [*On behalf of the Neurology Working Group*]

Einat Granot-Hershkovitz Ph. D.
Sleep medicine, Brigham and Women's Hospital

Einat Granot-Hershkovitz (1), Hector Gonzalez (2), Myriam Fornage(3), Wassim Tarraf(4), Tamar Sofer(1)
(1) Brigham and Women's Hospital, Harvard Medical School, Boston, MA
(2) Department of Neurosciences and Shiley-Marcos Alzheimer's Disease Research Center, University of California, San Diego, La Jolla, CA
(3) Brown Foundation Institute of Molecular Medicine, McGovern Medical School, The University of Texas Health Science Center at Houston, Houston, TX
(4) Institute of Gerontology and Department of Health Care Sciences, Wayne State University, Detroit, MI

Harvard Medical School

Late-onset Alzheimer disease (LOAD) is a growing worldwide epidemic and one of the leading causes of death in the elderly population. It is a multifactorial disease, influenced by genetic and environmental exposures. APOEε4 is the strongest known genetic risk factor for LOAD. Several population-based studies have shown ancestry heterogeneity of the APOEε4 - LOAD association, suggesting weaker associations in Hispanics. It is hypothesized that this lower risk effect is due to ancestry-specific protective genetic factors. Admixture mapping can be used to detect genetic association regions in admixed populations, such as Hispanics, by estimating associations between local ancestry allele counts and the trait of interest. We performed admixture mapping in the Hispanic Community Health Study/Study of Latinos (HCHS/SOL) for cognitive impairment and decline traits, serving as predictors for Dementia and Alzheimer, to detect association regions potentially related to LOAD risk. We tested local ancestry intervals across the genome for each of the three ancestral populations of Hispanics/Latinos: European, African, and Amerindian. We identified six genomic regions that were significantly associated with a cognitive trait at the genome-wide admixture mapping level. African ancestry in two genomic regions on chr11 (p15.5 and p15.4), and on chr21q22.13 was associated with cognitive decline. Amerindian ancestry in a region on chr13q31.1 was associated with cognitive impairment, and European ancestry in two regions, chr11p15.1 and chr6q25.1, were associated with cognitive decline and cognitive impairment, respectively. The chr6q25.1 overlaps with the MTHFD1L gene, previously associated with LOAD. MTHFD1L may influence levels of homocysteine which is a known risk factor for LOAD. Further analysis included fine-mapping and replication of these association signals. The results may lead to population-specific risk predictions and novel therapeutics, and, in turn, help reduce health disparities in the general population.

(37)

Serum triacylglycerol profiles of cortical thickness and surface area of the human brain [*On behalf of the Neurology Working Group*]

Eeva Sliz, Ph.D.
The Hospital for Sick Children, and Departments of Physiology and Nutritional Sciences, University of Toronto

Eeva Sliz(1)
Jean Shin(1)
Catriona Syme(1)
Sandra Black(1,2)
Sudha Seshadri(3)
Tomas Paus(1,4)
Zdenka Pausova(1)

1. University of Toronto, Toronto, Canada; 2. Sunnybrook Research Institute, Toronto, Canada; 3. University of Texas Health San Antonio, San Antonio, USA; 4. Holland Bloorview Kids Rehabilitation Hospital, Toronto, Canada

Background
Prior studies have demonstrated that the circulating pool of lipids is an important determinant of brain health; it is a prominent source of fatty acids (FAs) used by the brain as structural components of cell membranes. High proportion of brain lipids are long-chain polyunsaturated FAs, such as docosahexaenoic acid (DHA). Most circulating FAs are carried by triacylglycerols (TAGs). Here we examined associations of brain cortical thickness and surface area with a comprehensive set of serum TAGs.
Methods
We studied 441 middle-aged adults from the Saguenay Youth Study in whom fasting serum concentrations of 518 TAG species were quantified with mass spectrometry (Metabolon), and cortical thickness and surface area were determined from T1-weighted magnetic resonance images of the brain (1.5T, Siemens). We used linear regression to model brain phenotypes as a function of each TAG concentration adjusted for age and sex.
Results
Serum concentrations of 6 TAGs associated ($p<0.006$) with mean cortical thickness and none with mean cortical surface area. The effect estimates of the associations between the 518 TAGs and mean cortical thickness were positively correlated with the number of acyl-chain carbons ($p=0.002$) and double bonds ($p=4.4x10-28$), indicating that TAGs with longer-chain and more unsaturated FAs are more closely associated with higher cortical thickness. Four of the 6 TAGs associated with cortical thickness included FA18:3, which is a precursor of DHA. Secondary analysis of the 6 TAGs showed variation in their association effects across 34 cortical regions delineated with FreeSurfer. Virtual histology and functional enrichment analyses suggested an important role of astrocytes and CA1 pyramidal neurons, and of long-chain acyl-CoA synthetases in particular.
Conclusions
Circulating levels of TAGs carrying longer-chain polyunsaturated FAs are associated with thicker cerebral cortex. No circulating TAGs are associated with cortical surface area.

(38)

Genome-wide blood DNA methylation in relation to recent use of opioid medications [*On behalf of the Pulmonary Working Group*]

Mi Kyeong Lee, Ph.D.
NIEHS/NIH

Mi Kyeong Lee (1), Annah B. Wyss (1), Christine Parks (1), Stephanie J. London (1) for the Agricultural Lung Health Study

1 Epidemiology Branch, National Institute of Environmental Health Sciences, National Institutes of Health, Dept. of Health and Human Services, Research Triangle Park, NC

Introduction: Opioid use is a growing concern worldwide with many adverse health impacts that include the respiratory system. Candidate gene studies have found differential DNA methylation related to opioid addiction. Few data exist on associations of genome-wide blood DNA methylation with use of opioid medications in the general population.
Methods: We evaluated associations of genome-wide DNA methylation with recent use of opioid medications in the Agricultural Lung Health Study, a case-control study of asthma nested within a farming cohort. Among 2288 adults, we analyzed DNA methylation, measured with the Illumina HumanMethylationEPIC BeadChip, in relation to reported use of a full opiate agonist in the past 7 days (131 exposed). Using robust linear regression we modeled opioid use as the predictor of methylation and adjusted for age, sex, body mass index, smoking status, pack-years, asthma (selection factor), state of residence (study center), ancestry principal components, and cell-type proportions. Given that studies suggest sex differences in prevalence of opioid use and its effects, we additionally conducted the analysis stratified by sex. Benjamini-Hochburg False Discovery Rate was used to correct for multiple-testing.
Results: Many CpGs were differentially methylated in relation to opioid use (BH FDR<0.05) in all study participants. In sex-specific analyses, many additional CpGs were differentially methylated with significance of interaction. Follow-up analyses will include confirmation in an additional population based study, correlation with gene expression and functional annotation.
Conclusions: To our knowledge, this is the first large-scale study of genome-wide DNA methylation in relation to opioid use in a population-based study. This study can help elucidate how use of opioid medications impacts human health, including the respiratory system, and lead to the development of methylation-based biomarkers of use.

(39)

Genetically determined omega-3 polyunsaturated fatty acid levels and lung function: a Mendelian randomization analysis. [*On behalf of the Pulmonary Working Group*]

Bonnie Patchen
Cornell University

Bonnie Patchen (1)
Jiayi Xu (1)
Dana Hancock (2)
Patricia Cassano (1)

1. Cornell University; 2. RTI International

We previously found positive associations of plasma omega-3 polyunsaturated fatty acids (N-3 PUFAs) with lung function cross-sectionally. To improve causal inference, we performed two sample MR using data from a GWAS meta-analysis of N-3 PUFAs in the CHARGE consortium and genetic associations with lung function, including FEV1 and FVC, in the UK biobank. We instrumented the N-3 PUFAs alpha-linolenic acid (ALA), eicosapentanoic acid (EPA), docosahexaenoic acid (DHA) and docosapentaenoic acid (DPA) with variants in the fatty acid desaturase (FADS1/FADS2) and fatty acid elongase (ELOVL2) genes. To address correlation of N-3 PUFAs we performed sensitivity analyses using GWAS estimates of N-3 PUFAs adjusted for the preceding or subsequent N-3 PUFA in the metabolic pathway. Given the pleiotropic effects of FADS1/FADS2 variants on omega-6 (N-6) PUFAs, we also performed multivariable MR (MVMR) including linoleic acid (LA), the main dietary N-6 PUFA. We used the Wald's ratio or inverse variance weighted method in all analyses. In univariable MR, ALA was negatively associated with FEV1 (-0.27±0.13 mL/% total FA, p=0.02), while EPA was positively associated with FEV1 (0.05±0.02 mL/% total FA, p=0.02). The DPA—FEV1 association was similar to EPA (p=0.05). These results align with the opposing effects of FADS1/2 variants on ALA vs EPA and DPA. DHA was not associated with FEV1 and there were no statistically significant N-3 PUFA—FVC associations. Using GWAS estimates adjusted for correlated N3-PUFAs did not alter these results. In MVMR including LA, the ALA—FEV1 associations were strengthened (p=0.007), while the EPA—and DPA—FEV1 associations were no longer statistically significant. Our analyses suggest that higher ALA has a direct negative effect on lung function, while the positive effects of EPA and DPA may be through the balance of N-3 and N-6 PUFA metabolism. However, interpretation of MVRM findings when modeling metabolic pathways needs further consideration.

(40)

Epigenome-wide Association Study of Adult Asthma in the Agricultural Lung Health Study [*On behalf of the Pulmonary Working Group*]

Thanh Hoang, Ph.D.
Epidemiology Branch, NIEHS

Thanh T. Hoang(1), Mi Kyeong Lee(1), Sinjini Sikdar(1), Stephanie J. London(1) for the Agricultural Lung Health Study

1. Epidemiology Branch, National Institute of Environmental Health Sciences, National Institutes of Health, Department of Health and Human Services, Research Triangle Park, North Carolina, USA.

BACKGROUND: The genetic variation of asthma is incompletely understood. Epigenetic mechanisms, such as DNA methylation, may contribute to the variability of asthma. Most epigenome-wide association studies of DNA methylation and asthma have been conducted in children. We conducted an epigenome-wide study to identify methylation patterns in relation to adult asthma. Because asthma is a heterogeneous disease, asthma status was stratified by atopy.
METHODS: We analyzed data from the Agricultural Lung Health Study, a case-control study of adult asthma nested within a larger, agricultural cohort in the U.S. Current asthma was based on self-reported medical diagnosis of asthma plus current symptoms or use of asthma medications. Controls were a random sample of cohort members without current asthma. Participants were atopic if at least 1 of 10 specific immunoglobulin E was ≥0.70 IU/mL. Cases and controls were stratified by atopy status, resulting in 1,157 non-cases, 185 with atopy without asthma, 673 with non-atopic asthma, and 271 with atopic asthma. DNA methylation was assessed in whole blood using the Infinium Methylation EPIC BeadChip. Analyses were conducted using logistic regression.
RESULTS: Compared to non-cases, no differentially methylated C-phosphate-G sites (CpGs) were observed in atopy without asthma. Using a false discovery p-value<0.05, 524 CpGs were differentially methylated in non-atopic asthma, and 1,086 CpGs were differentially methylated in atopic asthma. 104 CpGs overlapped between non-atopic and atopic asthma. Differentially methylated CpGs in non-atopic asthma were enriched in pathways related to the nervous system and those in atopic asthma were enriched in pathways related to inflammation. Many differentially methylated findings replicated in studies with methylation in whole blood, respiratory epithelium, and eosinophils.
CONCLUSIONS: We identified several CpGs in relation to non-atopic and atopic asthma. These results have the potential to identify novel biomarkers of disease and shed light on asthma pathogenesis.

(41)

Kidney Function and Blood Pressure: A Mendelian Randomization Study [*On behalf of the Renal Working Group*]

Zhi Yu, BM, MS
Johns Hopkins University

Zhi Yu(1),
Josef Coresh(1),
Guanghao Qi(1),
Morgan Grams(1,2),
Eric Boerwinkle(3,4),
Harold Snieder(5),
Alexander Teumer(6),
Cristian Pattaro(7),
Anna Köttgen(1,8),
Nilanjan Chatterjee(1,2),
Adrienne Tin(1)

1. Johns Hopkins Bloomberg School of Public Health, Baltimore, MD; 2. Johns Hopkins School of Medicine, Baltimore, MD; 3. University of Texas Health Science Center at Houston, Houston, TX; 4. Baylor College of Medicine, Houston, TX; 5. University Medical Centre Groningen, Groningen, Netherlands; 6. University Medicine Greifswald, Greifswald, Germany; 7. Eurac Research, Bolzano, Italy; 8. Faculty of Medicine and Medical Centre - University of Freiburg, Freiburg, Germany

BACKGROUND: Lower kidney function has been associated with hypertension and vice versa. It remains unclear whether these associations are causal. We conducted two-sample bidirectional Mendelian randomization (MR) analyses to examine the causal relations between kidney function and blood pressure.

METHODS: The primary kidney function trait was glomerular filtration rate estimated from serum creatinine (eGFRcr) with blood urea nitrogen (BUN) as an alternative GFR marker. Summary results of European-ancestry genome-wide association studies were from the CKDGen Consortium (eGFRcr, n=567,460; BUN, n=243,031) and the UKB-ICBP Consortium (SBP and DBP, n=757,601). To reduce bias due to marker-specific determinants of the kidney function traits, we used association results from an alternative GFR marker to select genetic instruments that were likely to reflect kidney function. We used weighted mode, a pleiotropy-robust MR method, for primary analyses.
RESULTS. Of the 256 eGFRcr and 75 BUN index SNPs, 33 for eGFRcr and 24 for BUN were retained as genetic instruments. For blood pressure, 240 SBP and 243 DBP index SNPs were retained. Significant evidence supported the causal effects of lower kidney function for higher blood pressure (1 SD lower in ln(eGFRcr): 0.17 SD higher in SBP, p=9.92x10-5 and 0.15 SD higher in DBP, p=5.02x10-4). Similar results were observed for BUN to SBP and DBP. The causal effects of higher blood pressure for lower kidney function were not statistically significance.
CONCLUSIONS. MR analyses support causal effects of lower kidney function to higher blood pressure. These results suggest preventing CKD can reduce the public health burden of hypertension.

(42)

Assessment of kidney function traits on DNA methylation by developing and applying an EWAS workflow [*On behalf of the Renal Working Group*]

Alexander Teumer, Ph.D.
Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany

Pascal Schlosser(1), Franziska Grundner-Culemann(1), Pamela Matias Garcia(2), Antoine Weihs(3,4), Anna Köttgen(1,5), Adrienne Tin(5), Alexander Teumer(3,6)

1. Institute of Genetic Epidemiology, University of Freiburg, Germany; 2. Helmholtz Zentrum München, Munich, Germany; 3. Institute for Community Medicine, University Medicine Greifswald, Germany; 4. Department of Psychiatry and Psychotherapy, University Medicine Greifswald, Germany; 5. Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA; 6. DZHK (German Center for Cardiovascular Research), partner site Greifswald, Germany

Background: Results of recent GWAS on kidney function implicated gene regulatory mechanisms. To get insight into the link between DNA methylation and traits related to kidney function, we conducted a meta-analyses of epigenome-wide association study (EWAS) based on Illumina arrays with DNA methylation levels assessed from whole blood using samples of up to 32 cohorts of diverse ancestries.
Methods: The discovery stage of this project included up to 22,347 individuals for the EWAS on estimated glomerular filtration rate (eGFR), 11,458 individuals for the urinary albumin-to-creatinine ratio (UACR), and 12,479 individuals for serum urate levels. For each trait, associations that reached a 5% alpha level after Bonferroni correction for the number of sites tested (p<1.1E-7) where taken forward for replication in up to 11,258 independent samples. Sites with preplication<0.05 (same effect direction) and pcombined<1.1E-7 passed replication. Analyses were conducted using linear regression of the proportion of methylated DNA (beta values) on the trait adjusting for sex, age, white blood cell composition, technical factors, and known trait-specific correlates.
Results: We developed and established a workflow for standardization and quality control of EWAS including automated harmonized generation of variables for pre-analysis data checks, an intra- and inter-study quality control pipeline of the EWAS results that systematically checks both methylation data and association results for consistency. The BACON method was used to correct for inflation of test statistics in each cohort prior to meta-analysis. In total, 69, seven, and 100 CpG sites for eGFR, UACR and serum urate, respectively, fulfilled the criteria for replication. For eGFR, 60 of the 69 associations were novel, and 10 sites were located within known eGFR

GWAS loci. The replicated sites showed a clear trend for hypomethylation (n_CpG-=60, p_binom=2.2E-10). CpGs at SLC1A5 were significantly associated with both UACR and serum urate. Initial results indicated a causal effect of DNA methylation at some CpGs on eGFR. Bi-directional Mendelian randomization analyses are currently conducted for all traits.

Conclusions: These results will improve our understanding of the role of DNA methylation in whole blood in relation to kidney function traits and will provide insights into mechanisms of gene regulation.

(43)

Heritability enrichment analyses in kidney function GWAS identifies trait-specific kidney cell types [*On behalf of the Renal Working Group*]

Matthias Wuttke, MD
Institute of Genetic Epidemiology, University of Freiburg, Germany

Matthias Wuttke(1)
Yong Li(1)
Anna Köttgen(1)
on behalf of the CKDGen Consortium

1. Institute of Genetic Epidemiology, University of Freiburg, Germany

Background: Identifying relevant tissues and cell types underlying kidney function and disease informs experimental follow-up studies to understand disease biology. Novel statistical methods allow for unbiased identification of trait-relevant cell types by incorporating RNA-seq data with GWAS summary statistics.
Methods: We used LD score regression for specifically expressed genes (LDSC-seg) to partition heritability in GWAS summary statistics of European ancestry participants from the CKDGen Consortium of estimated glomerular filtration rate (eGFR, n=567,460), urinary albumin-to-creatinine ratio (UACR, n=547,361), blood urea nitrogen (BUN, n=243,031) and serum urate (n=288,666). Additionally, GWAS of blood pressure (SBP, n=745,820; DBP, n=757,601; UKB-ICBP), asthma (UK Biobank, n=452,264), and schizophrenia (CLOZUK+PGC Consortia, n=105,318) were analysed. Publicly available kidney single-cell RNA-seq datasets (human, 24 cell types; mouse, 16 cell types) were used to construct the top 10% specifically expressed genes per cell type followed by testing heritability enrichment in each trait. For examination at tissue level, the same procedure was applied using GTEx V7 data.
Results: Across tissues, we found significant enrichment of heritability in trait-associated loci containing genes that are highly expressed in kidney (eGFR: 2.2-fold enrichment, p=9.1e-8; urate: 2.1-fold enrichment, p=1.2e-5); liver was also enriched. Within the kidney, enrichment was observed in regions containing genes specifically expressed in proximal tubule cells in human (eGFR 2.3-fold, p=8.5e-5; BUN 1.7-fold, p=0.005; urate 2.3-fold, p=7.8e-6) and mice (eGFR 2.3 fold, p=0.0003; BUN 1.8-fold, p=0.02; urate 2.3-fold, p=0.0002), as well as in human podocytes (UACR 1.7-fold, p=0.009). Suggestive heritability enrichment (1.7-fold, p=2.1e-3) was found in human mesangium for SBP. Both asthma and schizophrenia did not show significant enrichment of heritability in regions with genes that are highly expressed in kidney cell types, but instead in brain tissues (schizophrenia, smallest p=9.8e-16). At the gene level, we found high and specific expression of UACR-associated genes such as PRKCI and PTH1R in podocytes and of CUBN in proximal tubules.
Conclusion: GWAS signals of kidney function traits are enriched for genes that are highly expressed in relevant tissues and cell types such as proximal tubular cells for eGFR, BUN and urate, and glomerular cells for UACR. These results allow for identifying relevant cell types for experimental research to translate GWAS loci into a mechanistic understanding.

###