An introduction to population based data for studies of DNA methylation

Stanford Center for Population Health Sciences
Seminar Series
March 22, 2019

David Rehkopf
Associate Professor
Stanford University School of Medicine
Six Questions addressed in this talk

1) What is DNA methylation?
2) What are the suspected causes and consequences of DNA methylation?
3) Why might population health scientists be interested in studying DNA methylation?
4) What are the standard approaches to analysis of DNA methylation data?
5) What can epidemiology, demography and social sciences contribute to understanding the role of DNA methylation?
6) What population based datasets are currently (2019) available for analysis of DNA methylation?
Q1. What is DNA methylation?
methyltransferase
Methylation and the human genome

3,234,000,000 base pairs

CpG 28,000,000

[NB: 42% of genome is CG, .21*.21 is 4.4%]
Genes that can be expressed

Promoter region

Genes inactivated by DNA methylation

M M M

Methylated

Unmethylated
Typical mammalian DNA methylation landscape

- CpG Island
- Transposable element
- CpG Island
- Gene

- methylated CpG
- unmethylated CpG
Illumina EPIC array

850,000 CpGs
Q2. What are the suspected causes and consequences of DNA methylation?
### Annual Review of Public Health

Environmental Influences on the Epigenome: Exposure-Associated DNA Methylation in Human Populations

Elizabeth M. Martin and Rebecca C. Fry

Department of Environmental Sciences and Engineering, and Curriculum in Toxicology, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, North Carolina 27599, USA; email: cmsebas@live.unc.edu, rfry@unc.edu

---

### Table 1  Summary of exposures assessed within this review

<table>
<thead>
<tr>
<th>Exposures</th>
<th>Global methylation</th>
<th>Gene-specific methylation</th>
<th>Exposure-associated health impact</th>
<th>Relevant citations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aflatoxin B1</td>
<td>Hypomethylation associated with exposure</td>
<td>71 CpG sites associated with prenatal exposure</td>
<td>Hepatocellular carcinomas, reduced growth, immune deficiencies</td>
<td>73, 75, 185, 197</td>
</tr>
<tr>
<td>Air pollution</td>
<td>Hypomethylation typically associated with exposure in adults, prenatal exposure is associated with both hypomethylation</td>
<td>MAPK pathway members, ACE, iNOS, ICAM-1, TLR2, IL-6, TET1</td>
<td>Accelerated lung aging, loss of lung capacity, asthma, bronchitis, emphysema, and cancer</td>
<td>19, 20, 28, 31, 32, 36, 39, 44, 45, 61, 69, 79, 87, 88, 98, 112, 120, 165, 168, 183</td>
</tr>
<tr>
<td>Arsenic</td>
<td>Hypomethylation associated with exposure with sex-specific directionality shown as well</td>
<td>KCNQ1, SQSTM1, sex-specific profiles</td>
<td>Cancer lung conditions and diabetes in adults; prenatal exposure is associated with increased incidence of infection, neurocognitive effects, and increased neonatal mortality</td>
<td>2, 6, 9, 13, 15, 29, 33, 34, 47, 50, 54, 65, 76, 77, 84, 97, 105, 110, 121, 136, 137, 151–153, 155, 159, 177, 184, 199</td>
</tr>
<tr>
<td>Bisphenol A</td>
<td>Hypomethylation associated with exposure in females, potential nonmonotonic dose responses</td>
<td>SNORD, SULT2A1, COMT</td>
<td>Neurocognitive effects, increased incidence of cancer, and heart conditions from prenatal exposure</td>
<td>52, 53, 70, 99, 133, 134</td>
</tr>
<tr>
<td>Cadmium</td>
<td>Hypomethylation associated with exposure</td>
<td>DNMT1</td>
<td>Cancer, lung, bone, and kidney disease, developmental toxicity</td>
<td>51, 70, 78, 103, 129, 169, 170, 187, 188</td>
</tr>
<tr>
<td>Chromium</td>
<td>Hypermethylation associated with exposure</td>
<td>Not assessed at present</td>
<td>Cancer</td>
<td>3, 192</td>
</tr>
<tr>
<td>Lead</td>
<td>Not assessed at present</td>
<td>Alterations in imprinted genes, sex-specific response</td>
<td>Neurotoxicity, developmental toxicity</td>
<td>64, 70, 114, 138, 172–174</td>
</tr>
<tr>
<td>Mercury</td>
<td>Not assessed at present</td>
<td>EMID2, sex-specific profiles</td>
<td>Neurotoxicity</td>
<td>14, 34, 35, 62, 70, 119</td>
</tr>
<tr>
<td>Polycyclic aromatic hydrocarbons</td>
<td>Hypomethylation associated with exposure</td>
<td>HIN1, ESR1, TWIST1</td>
<td>Cancer</td>
<td>48, 72, 74, 101, 112, 146, 149, 194, 195, 198</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Exposures</th>
<th>Global methylation</th>
<th>Gene-specific methylation</th>
<th>Exposure-associated health impact</th>
<th>Relevant citations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Persistent organic pollutants</td>
<td>Nonmonomeric association with exposure</td>
<td>IGF2, TNF-α, NRR1C1</td>
<td>Various health effects</td>
<td>40, 71, 82, 100, 116, 126, 141, 162, 201</td>
</tr>
<tr>
<td>Tobacco smoke</td>
<td>Hypomethylation associated with exposure</td>
<td>AHR2, CNTNAP2, MYO1G</td>
<td>Cancer, developmental toxicity, cardiovascular disease, chronic respiratory conditions</td>
<td>76, 77, 57, 55, 67, 86, 89, 94, 118, 148, 157, 175, 181, 182</td>
</tr>
<tr>
<td>Nutritional factors</td>
<td>Hypermethylation associated with exposure</td>
<td>IGF2, RXR-α, PPARγ1</td>
<td>Proper development</td>
<td>1, 17, 22, 23, 71, 93, 120, 131, 142, 154</td>
</tr>
</tbody>
</table>
Cigarette smoking behaviors and time since quitting are associated with differential DNA methylation across the human genome

Emily S. Wan1,* , Weiliang Qiu1, Andrea Baccarelli2, Vincent J. Carey1, Helene Bacherman1, Stephen I. Rennard3, Alvar Agusti4, Wayne Anderson5, David A. Lomas6 and Dawn L. DeMeo1

1Channing Laboratory and the Division of Pulmonary and Critical Care Medicine, Brigham and Women's Hospital, Boston, MA, USA, 2Exposure Epidemiology and Risk Program, Harvard School of Public Health, Boston, MA, USA, 3Section of Pulmonary and Critical Care, University of Nebraska Medical Center, Omaha, NE, USA, 4Thorax Institute, Hospital Clinic, IDIBAPS, Barcelona and CIBERES, Fundacio Caubet-Cilmer, Palma de Mallorca, Spain, 5GlaxoSmithKline Research and Development, Research Triangle Park, NC, USA and 6Cambridge Institute for Medical Research, University of Cambridge, Cambridge, UK

Received November 16, 2011; Revised March 22, 2012; Accepted April 2, 2012
Epigenome-wide meta-analysis of PTSD across 10 military and civilian cohorts identifies novel methylation loci

Alicia K Smith\textsuperscript{1,2}, Andrew Ratanatharathorn\textsuperscript{3}, Adam X Maihofer\textsuperscript{4}, Robert K Naviaux\textsuperscript{5}, Allison E Aiello\textsuperscript{6}, Ananda B Amstadter\textsuperscript{7}, Allison E Ashley-Koch\textsuperscript{8}, Dewleen G Baker\textsuperscript{4,9,10}, Jean C

Sensitive periods for the effect of childhood adversity on DNA methylation: Results from a prospective, longitudinal study

Erin C. Dunn, ScD, MPH\textsuperscript{b,c,1}, Thomas W. Soare, PhD\textsuperscript{a,b,c}, Andrew J. Simpkin, PhD\textsuperscript{d}, Matthew J. Suderman, PhD\textsuperscript{d}, Yiwen Zhu, MS\textsuperscript{a}, Torsten Klengel, MD, PhD\textsuperscript{b,e}, Andrew D.A.C. Smith, PhD\textsuperscript{d}, Kerry Ressler MD, PhD\textsuperscript{b,e}, Caroline L. Relton, PhD\textsuperscript{d,g}

\textsuperscript{a}Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA 02114
\textsuperscript{b}Department of Psychiatry, Harvard Medical School, Boston, MA 02115
\textsuperscript{c}Stanley Center for Psychiatric Research, The Broad Institute of Harvard and MIT, Cambridge, MA 02142
Systematic Mendelian randomization framework elucidates hundreds of genetic loci which may influence disease through changes in DNA methylation levels

Tom G. Richardson1, Philip C. Haycock3, Jie Zheng1, Nicholas J. Timpson1, Tom R. Gaunt1, George Davey Smith1, Caroline I. Relton4, Gibran Hemani2

1 MRC Integrative Epidemiology Unit (IEU), Bristol Medical School (Population Health Sciences), University of Bristol, Oakfield House, Oakfield Grove, Bristol, BS8 2BN, United Kingdom

*Corresponding author: Dr. Tom G. Richardson, MRC Integrative Epidemiology Unit, Bristol Medical School (Population Health Sciences), University of Bristol, Oakfield House, Oakfield Grove, Bristol BS8 2BN, UK. Tel: +44 (0)117 331 3370; E-mail: Tom.G.Richardson@bristol.ac.uk

Identification of 55,000 Replicated DNA Methylation QTL

Allan F. McRae1,2, Riccardo E. Mariam1,4, Sonia Shah1,2, Jia Yang1,2, Joseph E. Powell1,2, Sarah E. Harris1, Jude Gibson5, Anjali K. Henderson1, Lisa Bowdler6, Josie N. Painter6, Lee Murphy1,2, Nicholas G. Martin1,2, John M. Starr4,7, Naomi R. Wray1,2, Ian J. Deary1,2, Peter M. Visscher1,3,8 & Grant W. Montgomery1,3

DNA methylation plays an important role in the regulation of transcription. Genetic control of DNA methylation is a potential candidate for explaining the many identified SNP associations with disease that are not found in coding regions. We replicated 51,516 cis and 2,025 trans DNA methylation quantitative traitloci (mQTL) using methylation from whole blood measured on Illumina HumanMethylation450 arrays in the Brisbane Systems Genetics Study (n = 614 from 177 families) and the Lethal Birth Cohorts of 1921 and 1931 (combined n = 1366). The trans mQTL SNPs were found to be over-represented in 1 Mb subtelomeric regions, and on chromosomes 16 and 19. There was a significant increase in trans mQTL DNA methylation sites in upstream and 5' UTR regions. The genetic
Frailty is associated with the epigenetic clock but not with telomere length in a German cohort

Lutz Philipp Breitling, Kai-Uwe Saum, Laura Perna, Ben Schöttker, Bernd Holleczek and Hermann Brenner

DNA methylation age of blood predicts all-cause mortality in later life

Mental Health

The epigenetic clock is correlated with physical and cognitive fitness in the Lothian Birth Cohort 1936

Riccardo E Marioni, Sonia Shah, Allan F McRae, Stuart J Ritchie, Graciela Muniz-Terrera, Sarah E Harris

<table>
<thead>
<tr>
<th>Epigenetic Clock</th>
<th>Reference</th>
<th>Population</th>
<th>Age Range</th>
<th>Follow-Up (yr)</th>
<th>Sample Size</th>
<th>HR Mort</th>
<th>5% Cl. Mort</th>
<th>Risk Disease or Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Perna et al. 2016), Hannum</td>
<td>German case-cohort, ESTHER</td>
<td>50-75</td>
<td>Per 5 yr</td>
<td>1548</td>
<td>1.21</td>
<td>1.14, 1.29</td>
<td>Cancer, CVD</td>
<td></td>
</tr>
<tr>
<td>(Perna et al. 2016), Horvath</td>
<td>1.11</td>
<td>1.05, 1.19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Chen et al. 2016), EEEA</td>
<td>Meta-Analysis, 13 cohorts</td>
<td>13089</td>
<td>1.04</td>
<td>1.03, 1.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Zhang et al. 2017), DNA methylation score (10 CpGs)</td>
<td>German case-cohort, ESTHER</td>
<td>50-75</td>
<td>14</td>
<td>954</td>
<td>2.16</td>
<td>1.1, 4.24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Quach et al. 2017), EEEA</td>
<td>WHI, InCHIANTI, 71 ± 16</td>
<td>--</td>
<td>402</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Zhang et al. 2018), PhenoAge</td>
<td>Subset of population-based cohort</td>
<td>62 ± 8.5</td>
<td>14</td>
<td>858</td>
<td>1.37</td>
<td>1.25, 1.51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Zhang et al. 2018), Methylation risk (MR) score</td>
<td>Subset of population-based cohort</td>
<td>62 ± 8.7</td>
<td>14</td>
<td>903</td>
<td>1.91</td>
<td>1.63, 2.22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Levine et al. 2018), PharmaAge</td>
<td>5 studies: women’s health (WHI), Framingham Heart (FHHS), Normative Aging (NAS), Jackson Heart (JHS)</td>
<td>10</td>
<td>WNI: 2691 FHS: 2593 NAS: 667 JHS: 1747</td>
<td>Meta: 1.045</td>
<td>Healthspan, CVD, cancer, AD, T2D, respiratory</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Starnavskaya et al. 2017), DNA methylation score</td>
<td>Middle-aged twins</td>
<td>10</td>
<td>486</td>
<td>NOT Cognitive function</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Marioni et al. 2015), Hannum</td>
<td>5 cohorts: Lothian Birth cohorts (1921, 1936), Framingham Heart (FHHS), Normative Aging (NAS)</td>
<td>50-80</td>
<td>Per 5 yr</td>
<td>1821: 448 1936: 520 FHS: 2635 NAS: 857</td>
<td>1.21</td>
<td>1.14, 1.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Marioni et al. 2015), Horvath</td>
<td>1.11</td>
<td>1.05, 1.19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

DNA methylation age of blood predicts future onset of lung cancer in the women’s health initiative

Morgan E. Levine, H. Dean Hosgood, Brian Chen, Devin Absher, Themistocles Assimes, and Steve Horvath

1. Human Genetics, David Geffen School of Medicine, University of California LA, Los Angeles, CA 90095, USA;
2. Center for Neurobehavioral Genetics, University of California Los Angeles, Los Angeles, California 90095, USA;
3. Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY 10461, USA;
4. Longitudinal Study Section, Translational Gerontology Branch, Intramural Research Program, National Institute on Aging, National Institutes of Health, Bethesda, MD 20892, USA;
Q3. Why might population health scientists be interested in studying DNA methylation?
Reasons to study DNA methylation

1. Because it’s there
2. Prediction of future disease
3. Biological pathway from the environment to disease
4. As an intermediate outcome to disease and mortality
Q4. What are the standard approaches to analysis of DNA methylation data?
Collaborators

Lisa McEwen

Michael Kobor

Andres Cardenas

Simone Ecker

Nicole Gadish
NOTES: Each point is a voting location. Points are proportional to population size. Canton limits shown.
1. CpG site specific regression
Fig. 5 Pyrosequencing of significantly differentially methylated single CpGs. Left Bland–Altman plot of concordance between 450k array and pyrosequencing result for each CpG. Text labels represent sample IDs. Middle scatter plot displaying correlation between 450k array and pyrosequencing for each CpG. Spearman correlation coefficients shown. Right box plots of significant difference between Nicoyans and non-Nicoyans at each CpG site, measured using pyrosequencing. Significant value from regression model of CpG methylation on group status, while controlling for sex, age, and cell-type proportions.
2. Variability

with age. We calculated the interquantile range (IQR) at each CpG site (90th–10th percentile) represented on the 450k array to account for outliers and found a significant variability difference between Nicoyans and non-Nicoyans (Wilcoxon signed-rank test; $p < 2.2 \times 10^{-16}$), with a lower level of total mean DNAm variation in Nicoyans (Fig. 3a). Furthermore, we assessed the level of DNAm
3. Summary methylation score
Figure 1 Chronological age (y-axis) versus DNA methylation age (x-axis) in the training data. Each point corresponds to a DNA methylation sample (human subject). Points are colored and labeled according to the underlying data set as described in Additional file 1. (A) Across all training data, the correlation between DNA methylation age (x-axis) and chronological age (y-axis) is 0.97 and the error (median absolute difference) is 2.9 years. Results for (B) peripheral blood mononuclear cells (cor = 0.97, error <1 year), (C) whole blood (cor = 0.98, error = 2.7 years), (D) cerebellum (cor = 0.92, error = 4.5), (E) pons (cor = 0.96, error = 3.3), (F) pre-frontal cortex (cor = 0.98, 1.4), (G) temporal cortex (cor = 0.99, error = 2.2), (H) brain samples, composed of 58 glial cell, 58 neuron cell, 20 bulk, and 9 mixed samples (cor = 0.94, error = 3.1), (I) normal breast tissue (cor = 0.73, error = 8.9), (J) buccal cells (cor = 0.95, error <1 year), (K) cartilage (cor = 0.79, error = 4), (L) colon (cor = 0.98, error = 3.7), (M) dermal fibroblasts (cor = 0.92, error = 12), (N) epidemis (cor = 0.96, error = 3.1), (O) gastric tissue (cor = 0.83, error = 5.3), (P) normal adjacent tissue from head/neck cancers (cor = 0.73, error = 5.8), (Q) heart (cor = 0.82, error = 9.2), (R) kidney (cor = 0.88, error = 3.8), (S) liver (cor = 0.90, error = 4.5), (T) lung (cor = 0.80, error = 3.1), (U) mesenchymal stromal cells (cor = 0.95, error = 5.2), (V) prostate (cor = 0.55, error = 4.2), (W) saliva (cor = 0.89, error = 2.9), (X) stomach (cor = 0.84, error = 3.7), (Y) thyroid (cor = 0.96, error = 4.1).
Q5. What can epidemiology, demography and social sciences contribute to understanding the role of DNA methylation?
Contributions to move the literature forward

1) Generalizability, representative samples

2) Replication in multiple samples

3) Attention to identification strategies

4) Appropriate statistical mediation analysis
Q6. What population based datasets are currently (2019) available for analysis of DNA methylation?
Costa Rica (CRELES)

National Health and Nutrition Examination Survey (NHANES)

Health and Retirement Study (HRS)

Danish National Birth Cohort (DBC)

Women’s Health Initiative (WHI)
Costa Rica (CRELES study)
n=500 (now) + 500 (late 2019)
Costa Rica (CRELES study)

Costa Rica Longevity and Health Aging Study
Probabilistic sample of adults age 60 and over selected from the 2000 Census database
Different sampling fractions by age: 1941-1945: 1.1%, 1900 or earlier 100%
2 waves of data from 2005 and 2007
N=2827 total, 85% response rate, of those 95% blood
90 minute in person interview (24% proxy) collecting individual and household data on social, economic, functional status and health outcomes.
Blood draw (fasting next morning), urine and biomarker assays.
Linked to national mortality database.
The PAGE Study: How Genetic Diversity Improves Our Understanding of the Architecture of Complex Traits

Genevieve L Wojcik* (1), Mariaelisa Graff* (2), Katherine K Nishimura* (3), Ran Tao* (4), Jeffrey Haessler* (3), Christopher R Gignoux* (1), Heather M Highland* (2), Yesha M Patel* (5), Elena P Sorokin (1), Christy L Avery (2), Gillian M Belbin (6), Stephanie A Bien (3), Iona Cheng (7), Sinead Cullina (6), Chani J Hodonsky (2), Yao Hu (3), Laura M Huckins (6), Janina Jeff (6), Anne E Justice (2), Jonathan M Kocarnik (3), Unhee Lim (8), Bridget M Lin (2), Yingchang Lu (6), Sarah C Nelson (9), Sung-Shim L Park (5), Hannah Poisner (6), Michael H Preuss (6), Melissa A Richard (10), Claudia Schurmann (6), Veronica W Setiawan (5), Alexandra Sockell (1), Karan Vahi (11), Abhishek Vishnu (6), Marie Verbanck (6), Ryan Walker (6), Kristin L Young (2), Niha Zubair (3), Victor Acuna-Alonso (20), Jose Luis Ambite (11), Kathleen C Barnes (21), Eric Boerwinkle (12), Erwin Bottinger (6), Carlos D Bustamante (1), Christian Caberto (13), Samuel Canizales-Quintero (22), Matthew P Conomos (9), Ewa Deelman (11), Ron Do (6), Kimberly Doheny (14), Lindsay Fernandez-Rhodes (23), Myriam Fornage (10), Gerardo Heiss (2), Brenna Henn (24), Lucia A Hindorff (15), Rebecca D Jackson (16), Benyam Hailu (17), Cecelia A Laurie (9), Cathy C Laurie (9), Yuqing Li (7), Dan-Yu Lin (2), Andres Moreno-Estrada (25), Girish Nadkarni (6), Paul Norman (21), Loreall C Pooler (5), Alexander P Reiner (9), Jane Romm (14), Chiara Sabati (1), Karla Sandoval (25), Xin Sheng (5), Eli A Stahl (6), Daniel O Stram (5), Timothy A Thornton (9), Christina L Wassel (18), Lynne R Wilkens (13), Cheryl A Winkler (26), Sachi Yoneyama (2), Steven Buyske† (19), Chris Haiman† (5), Charles Kooperberg† (3), Loic Le Marchand‡ (13), Ruth JF Loos‡ (6), Tara C Matise‡ (19), Kari E North‡ (2), Ulrike Peters‡ (3), Eimear E Kenny‡ (6), Christopher S Carlson‡ (3)

* Shared first authorship
† Shared senior authorship
* Corresponding authorship
Life expectancy at age 80

Remaining years of life expectancy at age 80

Sources: Human Mortality Data Base (HMD); CCP: http://ccp.ucr.ac.cr/observa/CRIindicadores/TVcompletas.
Fig. 1. Life expectancy by per capita GDP. World’s countries 2003–7. Data from ref. 1.
NHANES 1999-2002 (n=2641) (late 2019)

MPI with Dr. Needham at U. Michigan
Large, nationally representative sample with socioeconomic and racial/ethnic diversity
DNA methylation from 566 African-American, 898 Hispanic, 1,071 white, and 79 other race individuals aged 50+ from NHANES 1999-2002.
Mortality follow-up
Links to Medicare claims data
Exact location data (RDC)
Figure 1. Conceptual Model of Racial/Ethnic Disparities in DNAm Patterns and Subsequent CV Mortality Risk

(Aim 2)
- Health Behaviors
- Environmental Toxicants
- Socioeconomic Factors
- Neighborhood Disadvantage

(Aim 3)
Race/Ethnicity -> DNAm -> CV Mortality

(Aim 1)
Age
<table>
<thead>
<tr>
<th>Data File Name</th>
<th>Doc File</th>
<th>Data File</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acculturation</td>
<td>ACQ Doc</td>
<td>ACQ Data [XPT - 629.2 KB]</td>
</tr>
<tr>
<td>Alcohol Use</td>
<td>ALQ Doc</td>
<td>ALQ Data [XPT - 314.5 KB]</td>
</tr>
<tr>
<td>Analgesic Pain Relievers</td>
<td>RXQ_ANA Doc</td>
<td>RXQ_ANA Data [XPT - 279.2 KB]</td>
</tr>
<tr>
<td>Audiometry</td>
<td>AUQ Doc</td>
<td>AUQ Data [XPT - 966.6 KB]</td>
</tr>
<tr>
<td>Balance</td>
<td>BAQ Doc</td>
<td>BAQ Data [XPT - 601.2 KB]</td>
</tr>
<tr>
<td>Blood Pressure &amp; Cholesterol</td>
<td>BPQ Doc</td>
<td>BPQ Data [XPT - 1.7 MB]</td>
</tr>
<tr>
<td>Cardiovascular Health</td>
<td>CDQ Doc</td>
<td>CDQ Data [XPT - 250.9 KB]</td>
</tr>
<tr>
<td>Cognitive Functioning</td>
<td>CFQ Doc</td>
<td>CFQ Data [XPT - 116.5 KB]</td>
</tr>
<tr>
<td>Current Health Status</td>
<td>HSQ Doc</td>
<td>HSQ Data [XPT - 553.8 KB]</td>
</tr>
<tr>
<td>Dermatology</td>
<td>DEQ Doc</td>
<td>DEQ Data [XPT - 3.5 MB]</td>
</tr>
<tr>
<td>Diabetes</td>
<td>DIQ Doc</td>
<td>DIQ Data [XPT - 1.2 MB]</td>
</tr>
<tr>
<td>Diet Behavior &amp; Nutrition</td>
<td>DBQ Doc</td>
<td>DBQ Data [XPT - 3.5 MB]</td>
</tr>
<tr>
<td>Drug Use</td>
<td>DUQ Doc</td>
<td>DUQ Data [XPT - 112.3 KB]</td>
</tr>
<tr>
<td>Early Childhood</td>
<td>ECQ Doc</td>
<td>ECQ Data [XPT - 462.3 KB]</td>
</tr>
<tr>
<td>Data File Name</td>
<td>Doc File</td>
<td></td>
</tr>
<tr>
<td>-------------------------------------------------------------------------------</td>
<td>-------------------</td>
<td></td>
</tr>
<tr>
<td>Aflatoxin B1-lysine - Serum (Surplus)</td>
<td>SSAFB_A Doc</td>
<td></td>
</tr>
<tr>
<td>Albumin &amp; Creatinine - Urine</td>
<td>LAB16 Doc</td>
<td></td>
</tr>
<tr>
<td>Anti-Mullerian Hormone (AMH) &amp; Inhibin-B (Surplus)</td>
<td>SSAMH_A Doc</td>
<td></td>
</tr>
<tr>
<td>Autoantibodies - Immunofluorescence &amp; Immunoprecipitation Analyses (Surplus)</td>
<td>SSANA_A Doc</td>
<td></td>
</tr>
<tr>
<td>Cadmium, Lead, Mercury, Cotinine &amp; Nutritional Biochemistries</td>
<td>LAB06 Doc</td>
<td></td>
</tr>
<tr>
<td>Chlamydia &amp; Gonorrhea - Urine</td>
<td>LAB05 Doc</td>
<td></td>
</tr>
<tr>
<td>Cholesterol - LDL &amp; Triglycerides</td>
<td>LAB13AM Doc</td>
<td></td>
</tr>
<tr>
<td>Cholesterol - Total &amp; HDL</td>
<td>LAB13 Doc</td>
<td></td>
</tr>
<tr>
<td>Complete Blood Count with 5-part Differential - Whole Blood</td>
<td>LAB25 Doc</td>
<td></td>
</tr>
<tr>
<td>C-Reactive Protein (CRP)</td>
<td>LAB11 Doc</td>
<td></td>
</tr>
<tr>
<td>Cryptosporidium &amp; Toxoplasma</td>
<td>LAB17 Doc</td>
<td></td>
</tr>
<tr>
<td>Cystatin C - Serum (Surplus)</td>
<td>SSCYST_A Doc</td>
<td></td>
</tr>
</tbody>
</table>
Health and Retirement Study (n=4100) (mid 2019)

Health and Retirement Study (1992 to present)
Longitudinal Panel data, nationally representative
Baseline survey 1992, age 50+ and spouses
Biological samples taken in 2006 and 2008
Over representation of non-whites in methylation sample
Information on early life location and history of residence since age 50.
Linked to mortality and medical claims data.
Danish National Birth Cohort (n=78 x 2)

78 Danish women
DNAm (EPIC) from these women at two time points
Meta data of moms and children during and after pregnancy
– anthropometric measurements
– physical health
– mental health
– social economic status
Red = DNAm data
Blue = meta data
DNA methylation age of blood predicts future onset of lung cancer in the women’s health initiative

Morgan E. Levine\textsuperscript{1,2}, H. Dean Hosgood\textsuperscript{3}, Brian Chen\textsuperscript{4}, Devin Absher\textsuperscript{5,*}, Themistocles Assimes\textsuperscript{6,*}, and Steve Horvath\textsuperscript{1,7,*}

\textsuperscript{1} Human Genetics, David Geffen School of Medicine, University of California LA, Los Angeles, CA 90095, USA; \textsuperscript{2} Center for Neurobehavioral Genetics, University of California Los Angeles, Los Angeles, California 90095, USA; \textsuperscript{3} Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY 10461, USA; \textsuperscript{4} Longitudinal Study Section, Translational Gerontology Branch, Intramural Research Program, National Institute on Aging, National Institutes of Health, Bethesda, MD 20892, USA;
Other studies

~100ng of DNA  ~$350
Costa Rica (CRELES)

National Health and Nutrition Examination Survey (NHANES)

Health and Retirement Study (HRS)

Danish National Birth Cohort (DBC)

Women’s Health Initiative (WHI)
Thank you.

Funding:
NIA (K01 AG047280, R21 MD013296, R01 HD091262, R01 MD011721)
CEDA, UC Berkeley (P30AG012839)
Canadian Institutes of Health Research

dredekopf@stanford.edu
@dredekopf