Capturing and defining population descriptors in research and health systems

Public Workshop on Use of Race, Ethnicity, and Ancestry as Population Descriptors in Genomics Research - Session II

Eimear Kenny, PhD
Founding Director, The Institute for Genomic Health
Professor of Medicine and Genetics
Icahn School of Medicine at Mount Sinai
Twitter: @EimearEKenny

Icahn School of Medicine at Mount Sinai

Institute for Genomic Health

Disclosures

Eimear Kenny has received personal fees from Regeneron Pharmaceuticals, 23&Me, and Illumina, and serves on the advisory boards for Encompass Biosciences and Galateo Bio.

Use of population descriptors in academic research

International HapMap Project

Polygenic RIsk MEthods in Diverse populations (PRIMED) Consortium

Population Architecture Using Genomics and Epidemiology (PAGE) Consortium

Human Genome Reference Program

1000 Genomes Project

Trans-Omics for Precision Medicine (TOPMed) Program

Project began

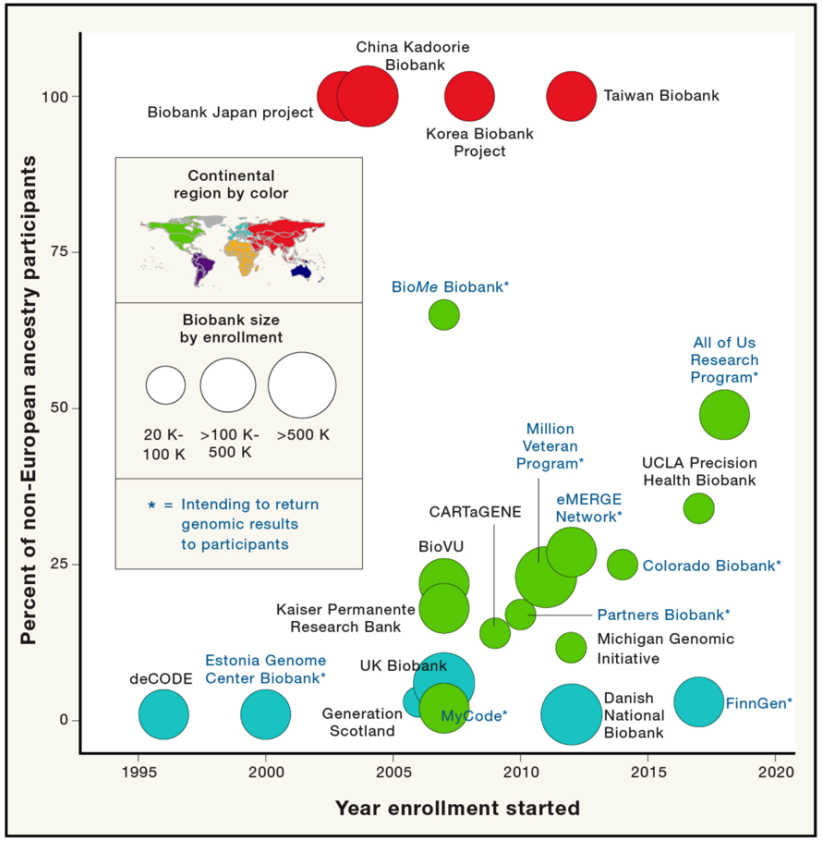
Clinical Sequencing Evidence-Generating Research (CSER)

NHLBI Grand Opportunity Exome Sequencing Project (ESP)

Use of population descriptors in academic research

- Upfront sampling and assignation of population descriptor
- Interested in looking at differences within and between populations as a means to understand both genetic and environmental factors impacting disease
- Long term goal to translate findings to understand impact on real-world populations

Population-based recruitment to biobanks as engines of translation



NS Abul-Husn, EE Kenny. *Cell* 2019; 177(1): 58-69

All of Us RESEARCH PROGRAM | The Precision Medicine Initiative

CHINA KADOORIE BIOBANK 中国慢性病前瞻性研究

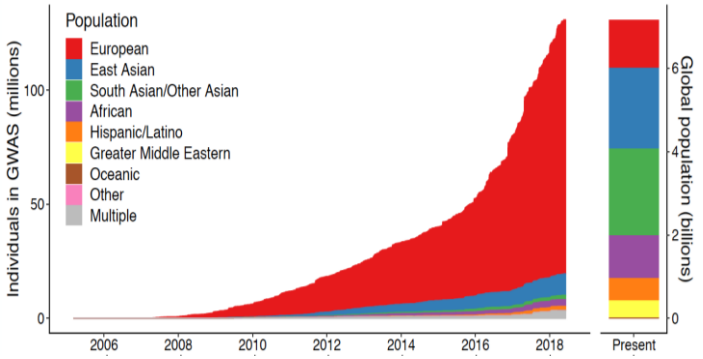
emerge network
ELECTRONIC MEDICAL RECORDS & GENOMICS

MVP

biobank uk
Improving the health of future generations

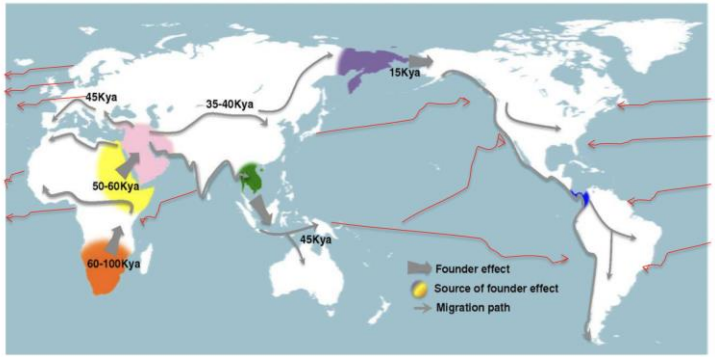
BIOBANK JAPAN

Challenges for understanding genetic risk at a population-level



Biases in representation in our genomic databases ….

…. complexity and admixture in our human demographic history…



Henn B M et al. PNAS 2012;109:17758-17764

…. and need to integrate genomic information with clinical history, family history, lifestyle factors, and social determinants of health

Impact of diversity on polygenic risk prediction

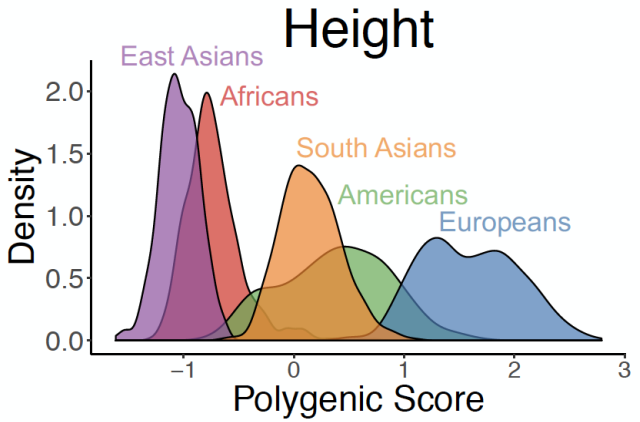
Please cite this article in press as: Martin et al., Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations, The American Journal of Human Genetics (2017), http://dx.doi.org/10.1016/j.ajhg.2017.03.004

ARTICLE

Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations

Alicia R. Martin,[1,2,3,4] Christopher R. Gignoux,[4] Raymond K. Walters,[1,2,3] Genevieve L. Wojcik,[4] Benjamin M. Neale,[1,2,3] Simon Gravel,[5,6] Mark J. Daly,[1,2,3] Carlos D. Bustamante,[4] and Eimear E. Kenny[7,8,9,10,*]

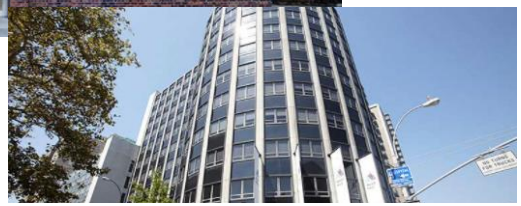
Genetic prediction accuracy decays with increasing divergence between the discovery populations and the target populations



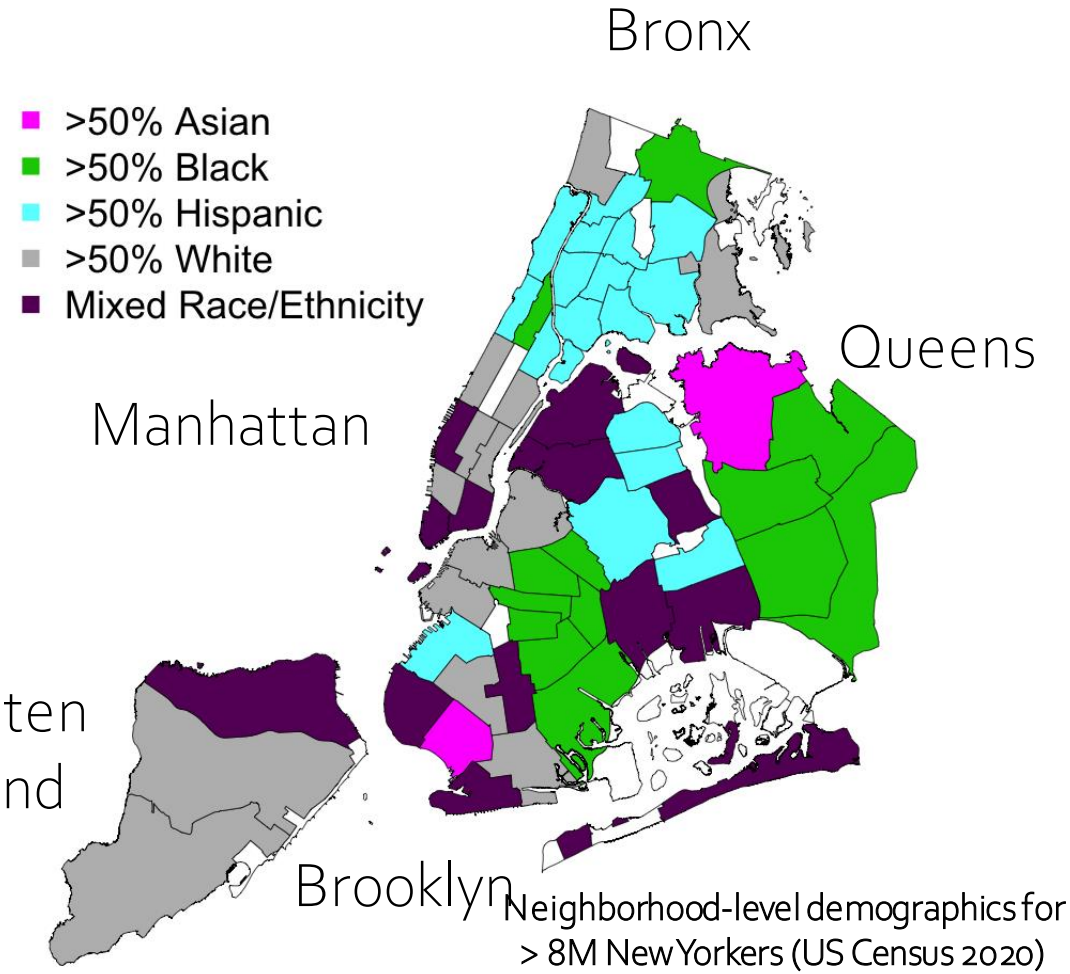
Martin AR et. al. AJHG 2017

Correction: Martin AR et. al. AJHG 2020

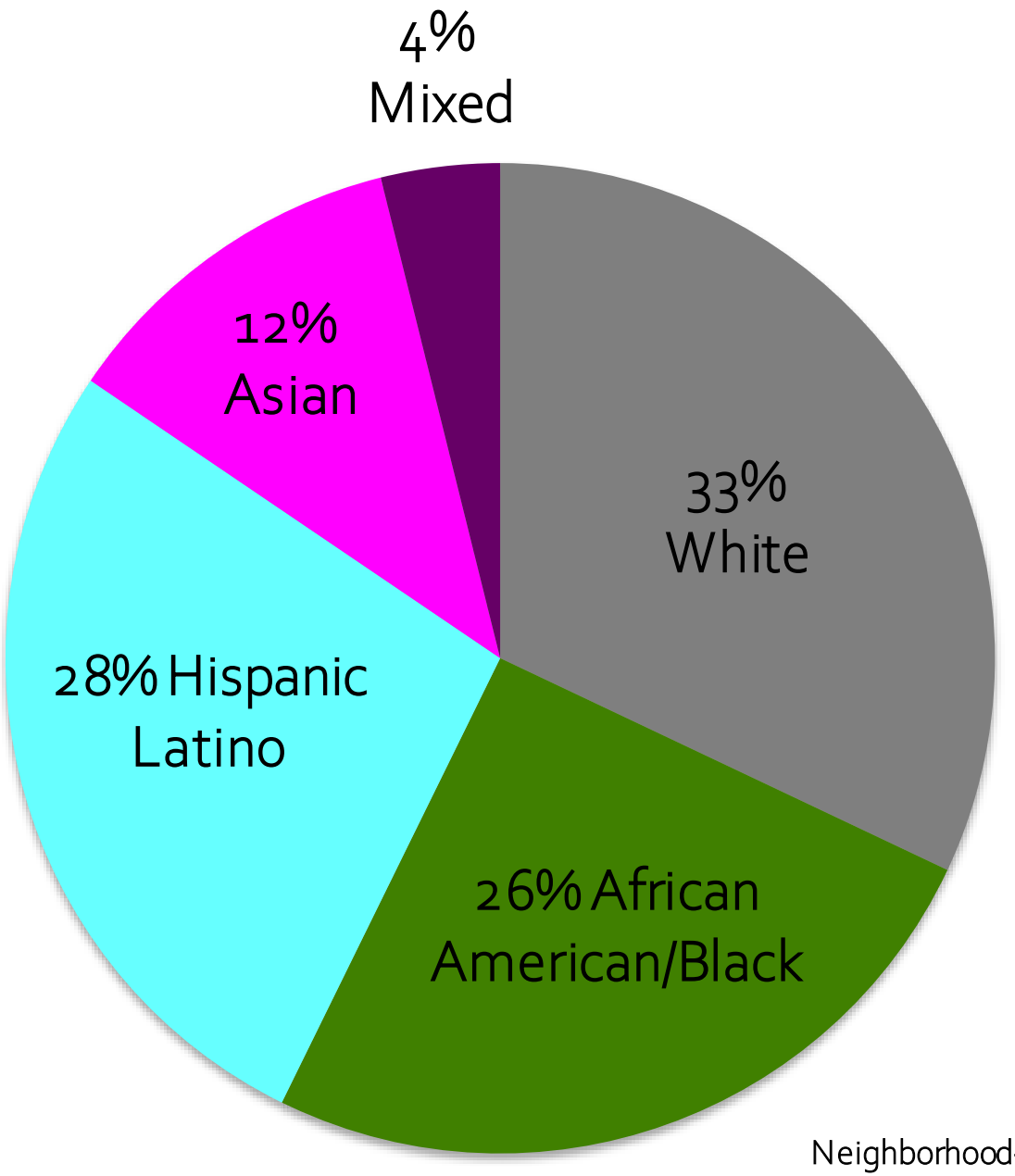
Mount Sinai Health System covers the lives of >3M New Yorkers



Mount Sinai Hospital



Neighborhood-level demographics for > 8M New Yorkers (US Census 2020)



Neighborhood-level demographics for
> 8M New Yorkers (US Census 2020)

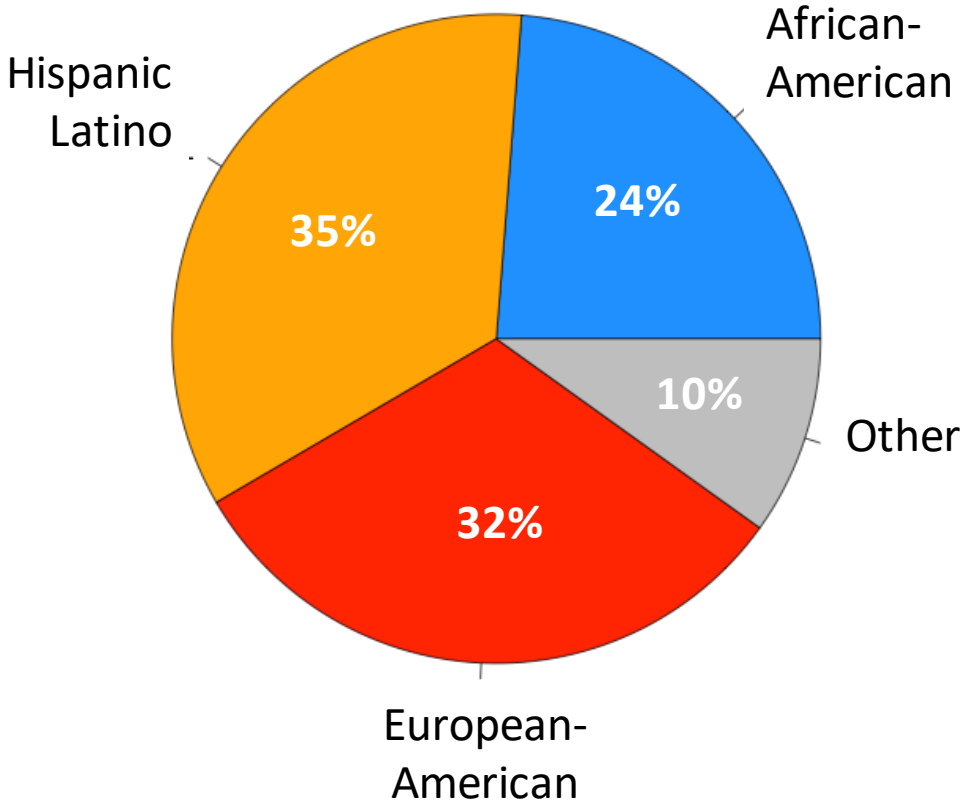
Mount Sinai The Charles Bronfman Institute for Personalized Medicine

BioMe Biobank

Your history is our future.

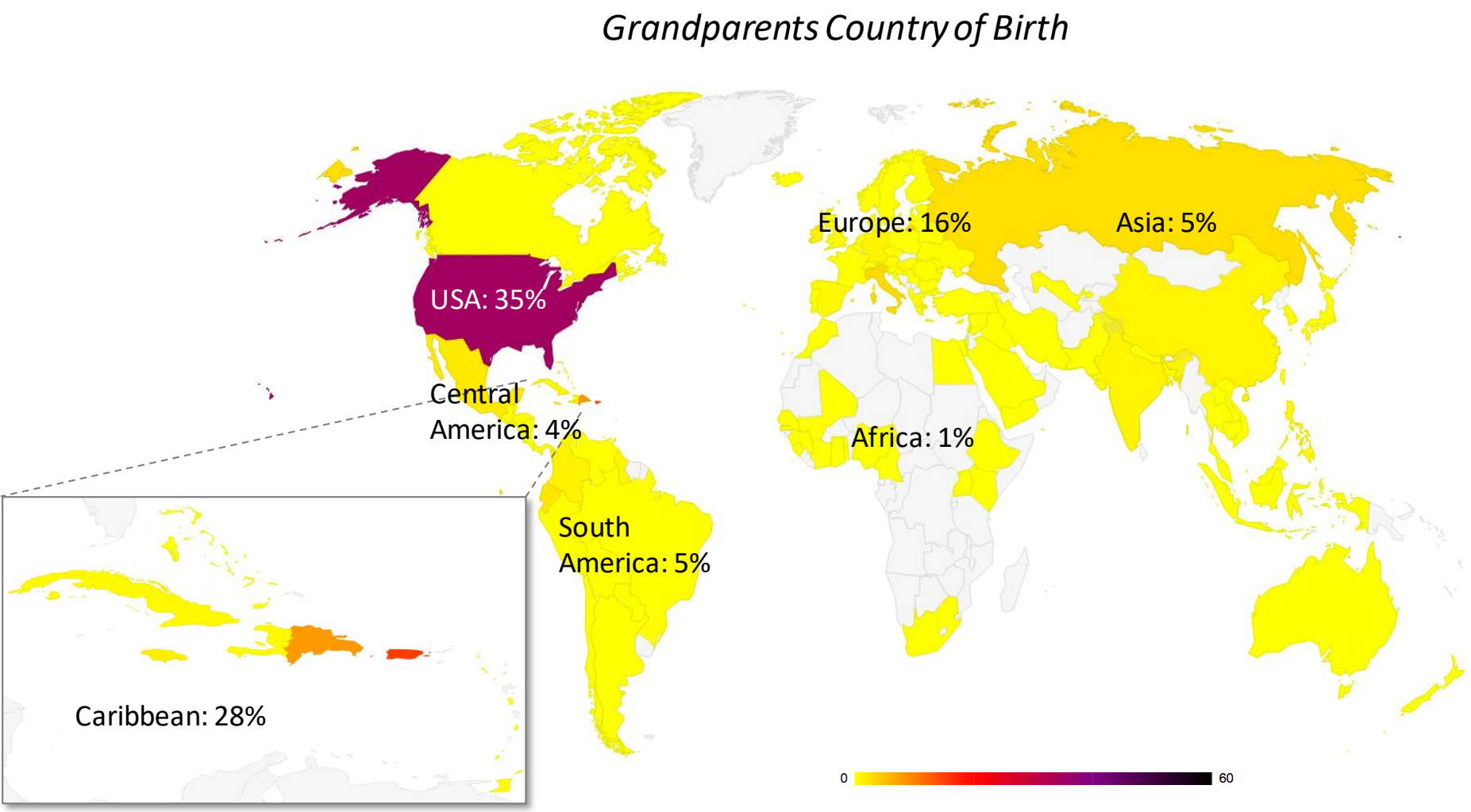
Genomic discovery and clinical genomics in a diverse patient population

- Research biobank recruiting since 2007
- >70,000 participants currently enrolled
- GS, ES, array data
- Rich survey information for those enrolled

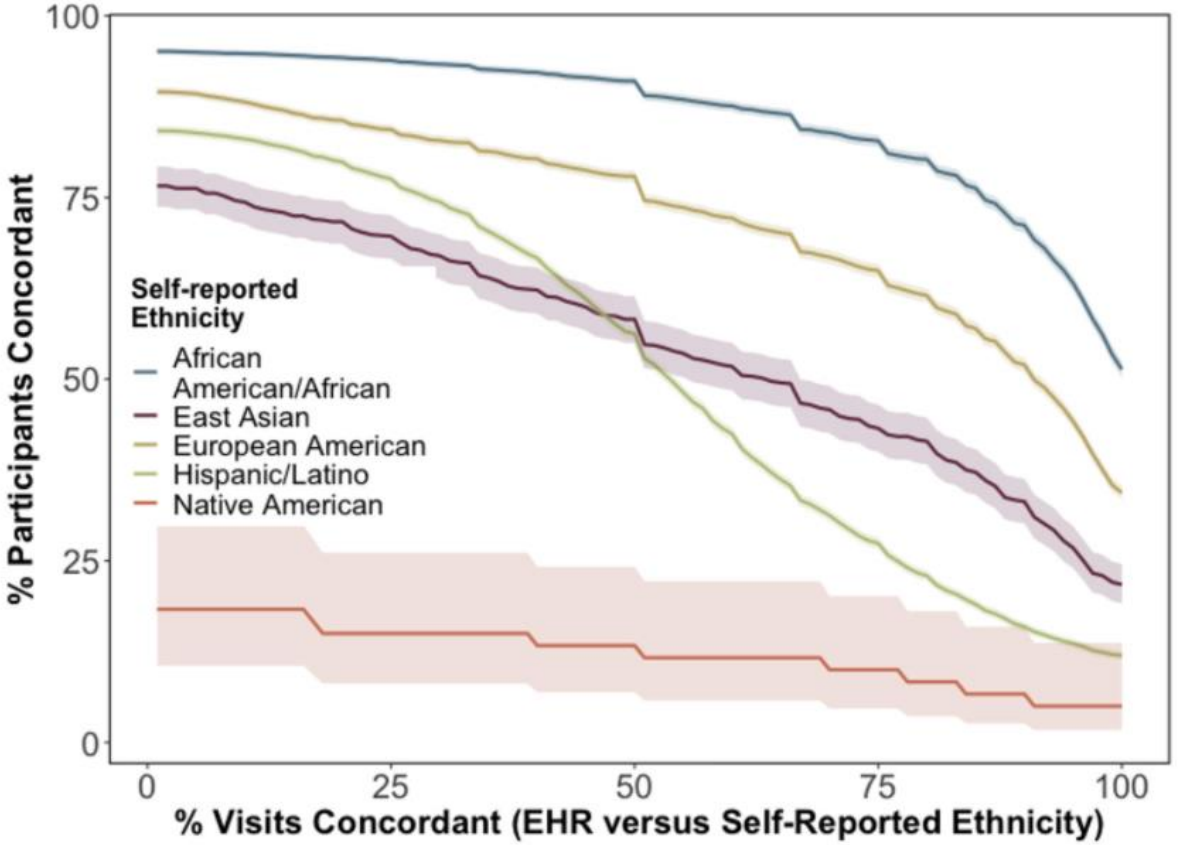


Mount Sinai Hospital

Diversity is deep and multidimensional



But diversity is poorly captured in a health system



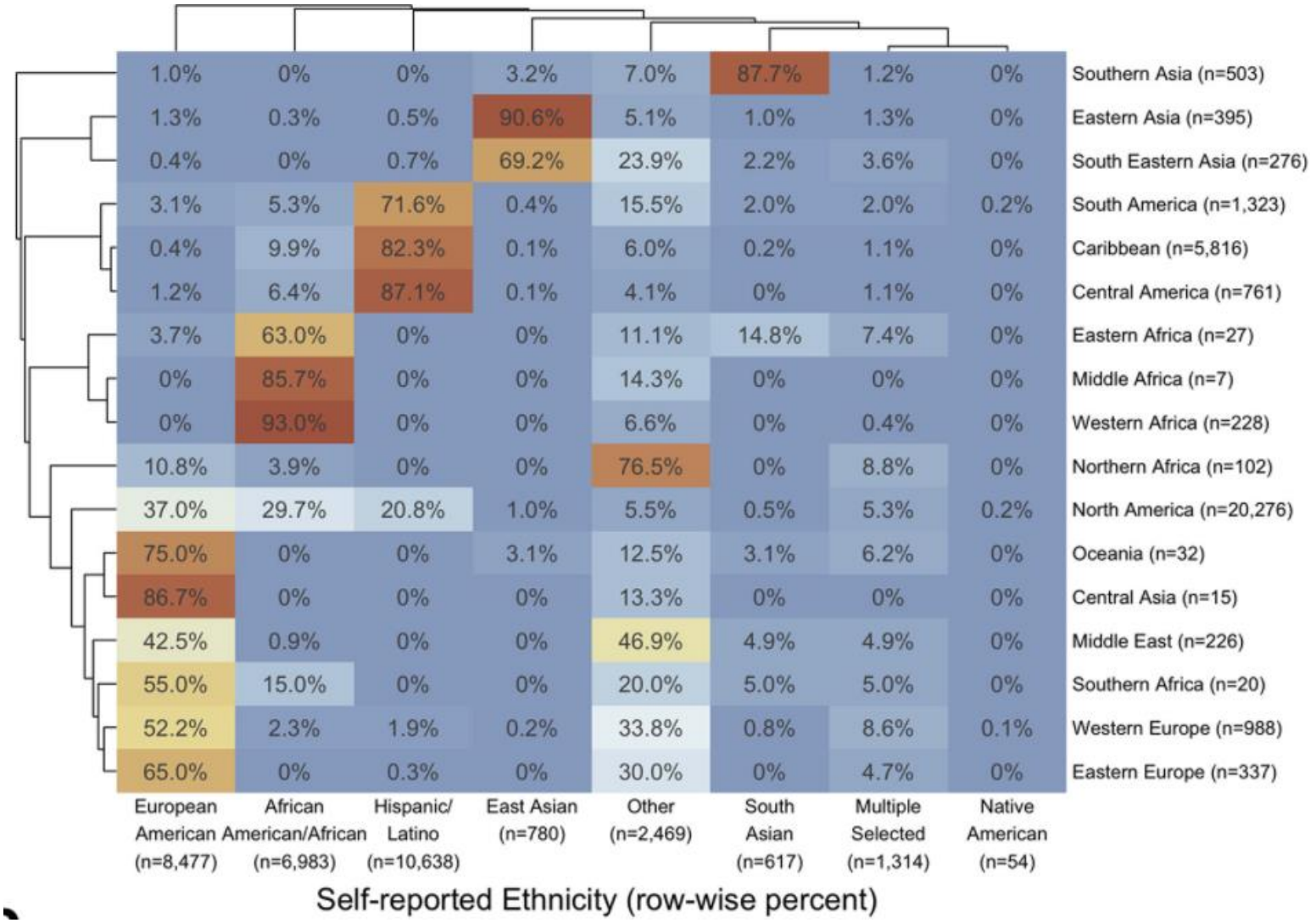
30,376 Bio*Me* participants

1,310,279 total health system visits

January 2007 and December 2014

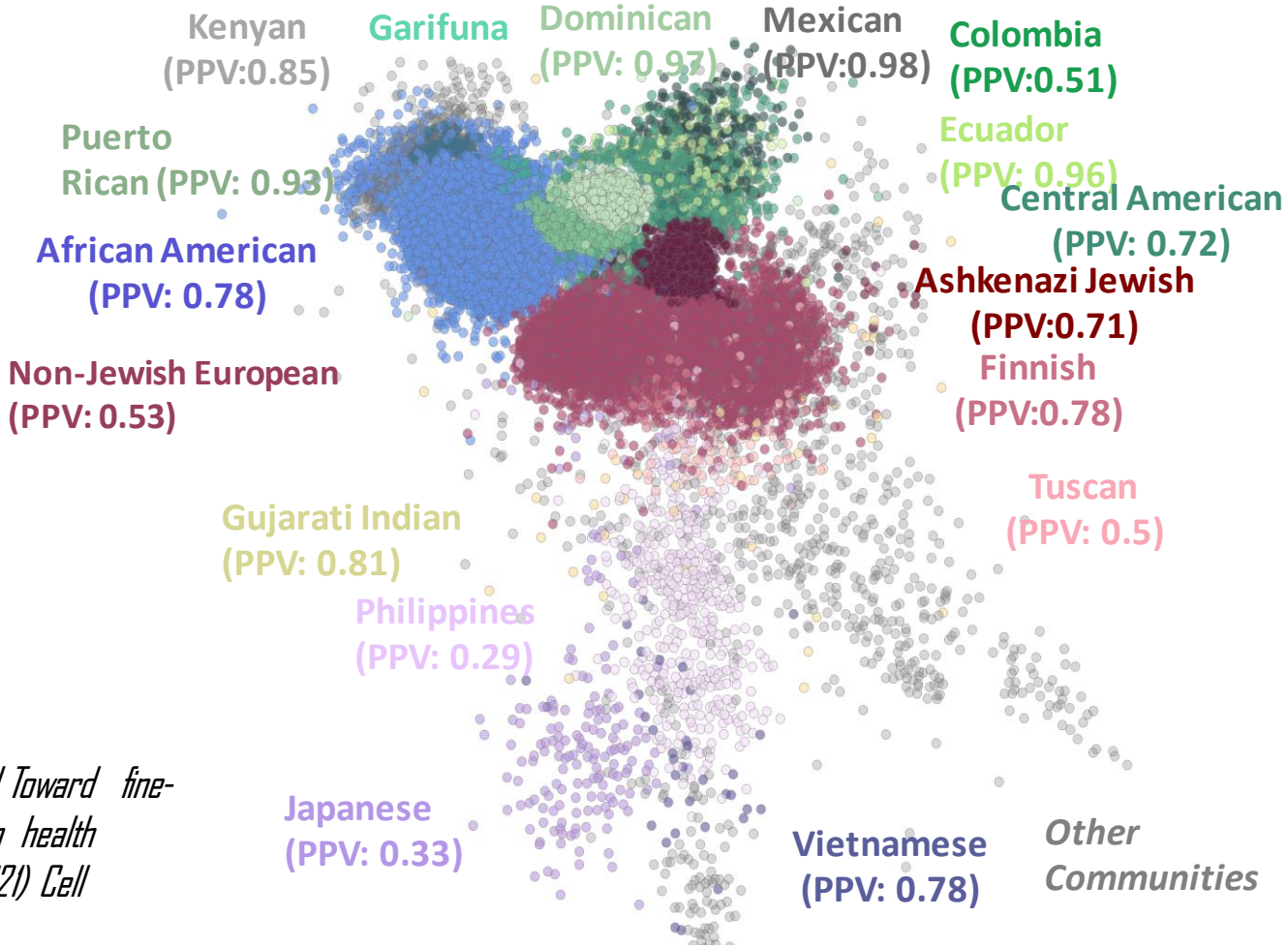
Belbin GM, et al Toward fine-scale population health monitoring Cell (2021)

Nuance and complexity in self-identity



Belbin GM, et al Toward fine-scale population health monitoring Cell (2021)

Genetic ancestry unlocks nuance and scale



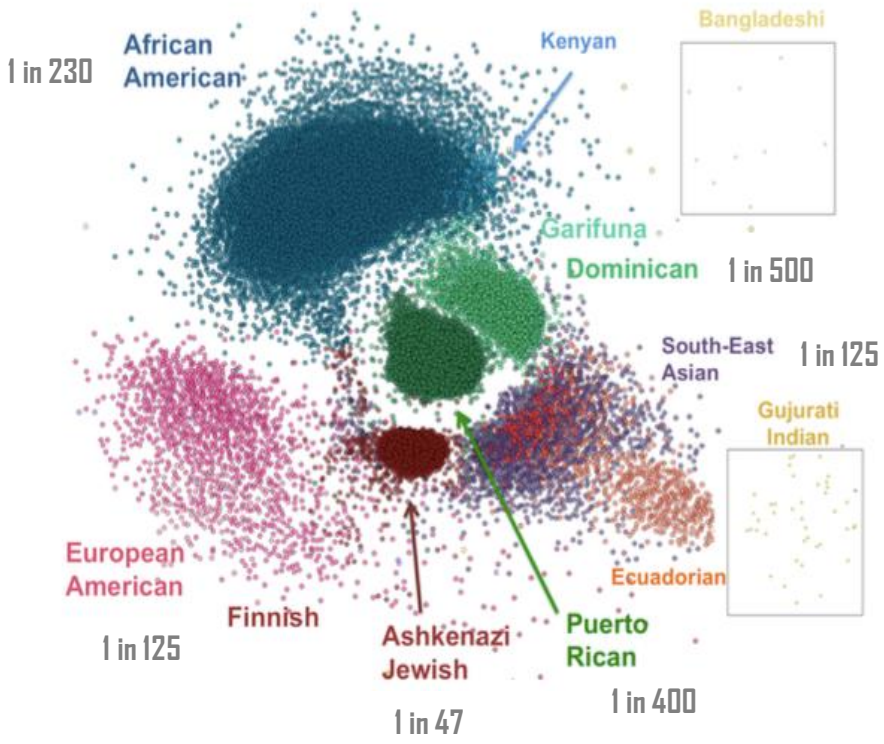
Belbin GM, et al Toward fine-scale population health monitoring (2021) Cell

$$\mathrm{PPV} = \frac{\text{number of true positives}}{\text{number of true positives} + \text{number of false positives}} = \frac{\text{number of true positives}}{\text{number of positive calls}}$$

Some Populations Are at Higher Risk of Many Types of Cancers

Cancer type	Average lifetime risk	Risk by age 70 when *BRCA1+*	Risk by age 70 when *BRCA2+*
Female breast	12%	46 – 87%	38 – 84%
Male breast	<0.1%	1 – 2%	Up to 9%
Ovarian	1 – 2%	39 – 63%	16 – 27%
Prostate	11%	Elevated	15 – 20%
Pancreatic	<1%	1 – 3%	2 – 7%
Melanoma	1.6%	?	Elevated

Prevalence of *BRCA1 and BRCA2* across populations



Exome sequencing reveals a high prevalence of *BRCA1* and *BRCA2* founder variants in a diverse population-based biobank

Noura S. Abul-Husn, Emily R. Soper, Jacqueline A. Odgis, Sinead Cullina, Dean Bobo, Arden Moscati, Jessica E. Rodriguez, CBIPM Genomics Team, Regeneron Genetics Center, Ruth J. F. Loos, Judy H. Cho, Gillian M. Belbin, Sabrina A. Suckiel & Eimear E. Kenny

Genome Medicine **12**, Article number: 2 (2019) | Cite this article

For inherited large effect risk variants, genetic ancestry may be a better tie to risk than population descriptors.....

Table 1. *Continued*

Gene	Mode of inheritance	Physical position (GRCh38)	c.DNA position	Protein modification	Self-reported ethnic group	Self-reported country of birth	IBD community	Condition	IBD community prevelance
DNAI2	AR	17: 74309345	c.1304G>A	p.Trp435Ter	AJ(5); C(9); B(4); O(1)	US(17); O(2)	AJ(19)	rimary ciliary yskinesia	1 in 216
CFAP298	AR	21: 32602299	c.735C>G	p.Tyr245Ter	AJ(5); C(6); B(1); O(2)	US(12); O(2)	AJ(15)	rimary ciliary yskinesia	1 in 293
DHDDS	AR	1: 26438228	c.124A>G	p.Lys42Glu	AJ(4); C(18); B(7); O(6)	US(31); O(4)	AJ(35)	etinitis pigmentosa, onsyndromic	1 in 117
PCDH15	AR	10: 54317414	c.733C>T	p.Arg245Ter	AJ(5); C(10); B(1); O(1)	US(15); O(2)	AJ(17)	sher syndrome ype 1	1 in 241
CLRN1	AR	3: 150972565	c.144T>G	p.Asn48Lys	AJ(8); C(29); B(5); O(3)	US(41); O(4)	AJ(44); O(1)	sher syndrome ype 3A	1 in 93
PR founder variants (N = 3)									
BRCA2	AD	13: 32338277	c.3922G>T	p.Glu1308Ter	HL(7)	US(3); PR(4)	PR(7)	ereditary breast ancer	1 in 729
COL27A1	AR	9: 114195977	c.2089G>C	p.Gly697Arg	HL(82); B(1); O(4)	US(47); PR(38); O(2)	PR(87); O(2)	teel syndrome	1 in 60
SGCG	AR	13: 23324452	c.787G>A	p.Glu263Lys	HL(49); O(6)	US(31); PR(22); O(2)	PR(54); O(2)	mb girdle dystrophy	1 in 96

Belbin GM, et al Toward fine-scale population health monitoring Cell (2021)

Population Characteristics	All *BRCA1/2*-Associated Cancers	Evidence of clinical genetic testing
	Personal History N (%)	N (%)
All Variant Positive (N = 217)	**54 (24.9)**	**57 (26.3)**
By Gender		
Female (N = 134)	48 (35.8)	52 (38.8)
Male (N = 83)	6 (7.2)	5 (6.0)
P-value (chisq test)	$3.4x10^{-7}$	$3.1x10^{-7}$
By Founder Variants		
With AJ Founder Variant (N = 80)	22 (27.5)	**31 (38.8)**
Without AJ Founder Variant (N = 137)	32 (23.4)	**26 (19.0)**
P-value (chisq test)	0.18	**$1.5x10^{-4}$**

Abul-Husn et al Gen Med *(2019)*

.... but knowledge of population risk may improve health outcomes

Thoughts for the future

- Population labels are labile, nuanced, and tied to social and political forces -> but are an important anchor to understand the gene x environment impact on disease

- Genetic ancestry can be an excellent biomarker for the environment factors and social determinants impacting disease -> an additional way to incorporate population information into genomic research

- Genomic information will be increasing centered in routine clinical care, but many questions remain about how to translate and implement genomic medicine -> need better approaches to reconciling individual and population-based risk

Acknowledgements

Institute for Genomic Health:
Noura Abul-Husn, Clinical Director
Gillian Belbin, Population Genetics Director
Giovanna Braganza
Sinead Cullina
Kate Donahue
Julia Galasso
David Wang
Ayuko Iverson
Karla Lopez
Jacqueline Odgis
Jessica Rodriguez
Joe Rothstein
Roohy Shemirani
Emily Soper
Sabrina Suckiel
Nicole Yelton

University of Colorado:
Chris Gignoux

Harvard/Broad:
Alicia Martin

Stanford University:
Carlos Bustamante

Johns Hopkins University:
Geneviève Wocjik

THANK YOU
to the thousands of research participants who make this work possible.

PAGE Collaborators:
Lucia Hindorff (NHGRI)
Kari North (UNC)
Chris Carlson (Fred Hutch)
Charles Kooperberg (Fred Hutch)
Chris Haiman (MEC)
Ruth Loos (MSSM)
Tara Matise (Rutgers)
Steve Buyske (Rutgers)
Jose-Luis Ambite (USC)
Christy Avery (UNC)
Ulrike Peters (Fred Hutch)

NIH Funding:
NHGRI/U01HG109391
NHGRI/U01HG009080
NHLBI/X01HL134588
NIDDK/R01HL104608
NHLBI/R01DK110113