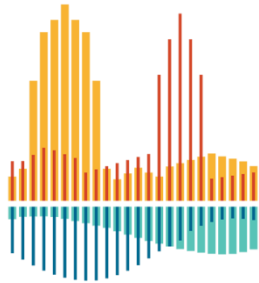


# THE ABCs OF DIVERSITY IN GENOMIC RESEARCH: ASCERTAINMENT, BIOINFORMATICS AND COMPUTATION



INSTITUTE FOR  
COMPUTATIONAL  
BIOLOGY

October 14, 2016

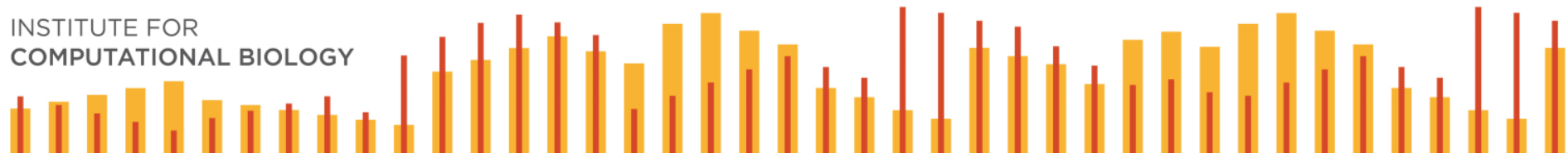
Dana C. Crawford, PhD  
Associate Professor  
Epidemiology and Biostatistics  
Institute for Computational Biology



# THE STATE OF GENOMIC DISCOVERY

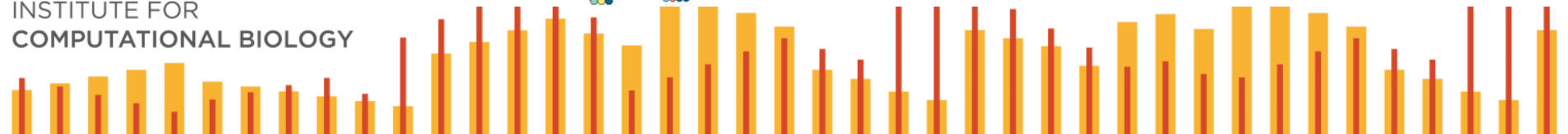


INSTITUTE FOR  
COMPUTATIONAL BIOLOGY



GWAS as of 2015:  
2,111 publications  
15,396 associated  
SNPs ( $10^{-5}$ )

<https://www.ebi.ac.uk/gwas/>

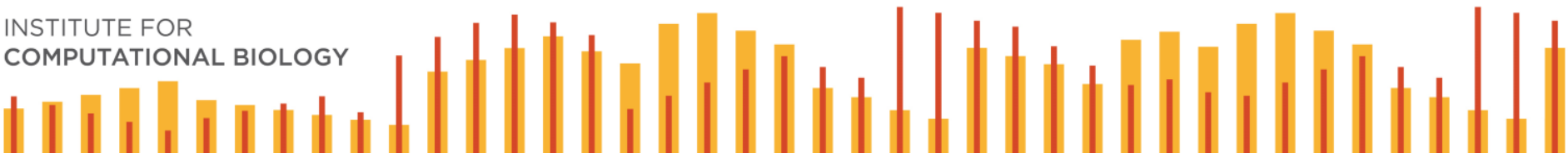
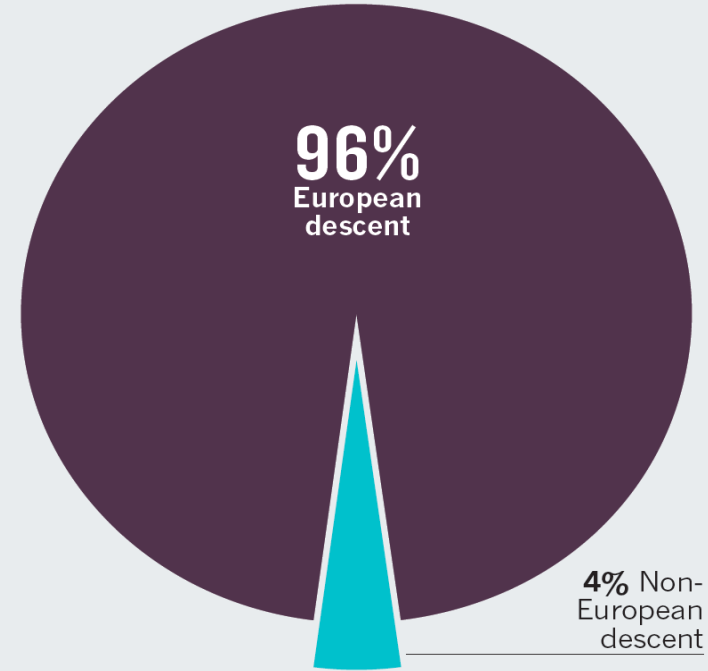


# MOSTLY EUROPEAN-DESCENT

Bustamante, Burchard, de la Vega (2011) *Nature*  
475:163-165

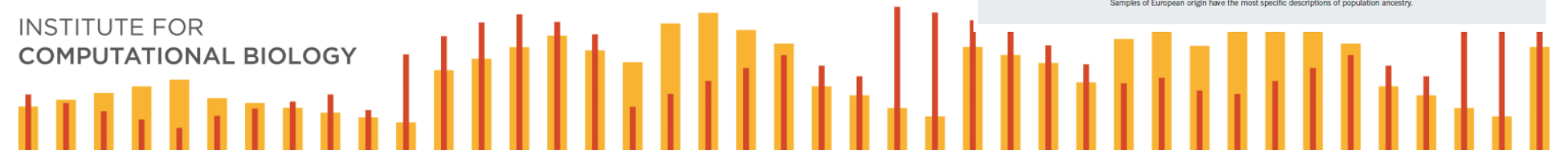
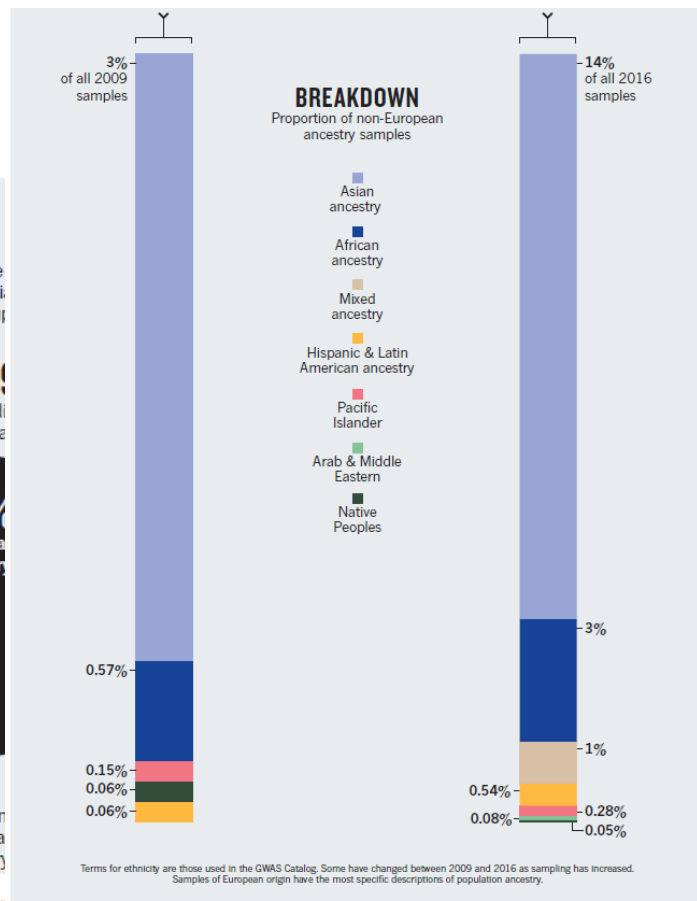
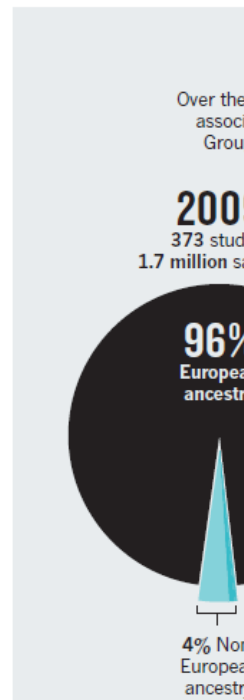
## SAMPLING BIAS

Most genome-wide association studies have been of people of European descent.



# GETTING BETTER?

Popejoy and Fullerton (2016) *Nature*  
538: 161-164



# DIVERSITY IN RESEARCH MATTERS

SPECIAL ARTICLE

## Genetic Misdiagnoses and the Potential for Health Disparities

Arjun K. Manrai, Ph.D., Birgit H. Funke, Ph.D., Heidi L. Rehm, Ph.D., Morten S. Olesen, Ph.D., Bradley A. Maron, M.D., Peter Szolovits, Ph.D., David M. Margulies, M.D., Joseph Loscalzo, M.D., Ph.D., and Isaac S. Kohane, M.D., Ph.D.

N Engl J Med 2016; 375:655-665 | [August 18, 2016](#) | DOI: 10.1056/NEJMsa1507092

REPORT

## Association of Trypanolytic ApoL1 Variants with Kidney Disease in African Americans

Giulio Genovese<sup>1,2,\*</sup>, David J. Friedman<sup>1,3,\*</sup>, Michael D. Ross<sup>4</sup>, Laurence Lecordier<sup>5</sup>, Pierrick Uzureau<sup>5</sup>, Barry I. Freedman<sup>6</sup>, Donald W. Bowden<sup>7,8</sup>, Carl D. Langefeld<sup>8,9</sup>, Taras K. Oleksyk<sup>10</sup>, Andrea L. Uscinski Knob<sup>4</sup>, Andrea J. Bernhardt<sup>1</sup>, Pamela J. Hicks<sup>7,8</sup>, George W. Nelson<sup>11</sup>, Benoit Vanhollebeke<sup>5</sup>, Cheryl A. Winkler<sup>12</sup>, Jeffrey B. Kopp<sup>11</sup>, Etienne Pays<sup>5,†</sup>, Martin R. Pollak<sup>1,13,†</sup>

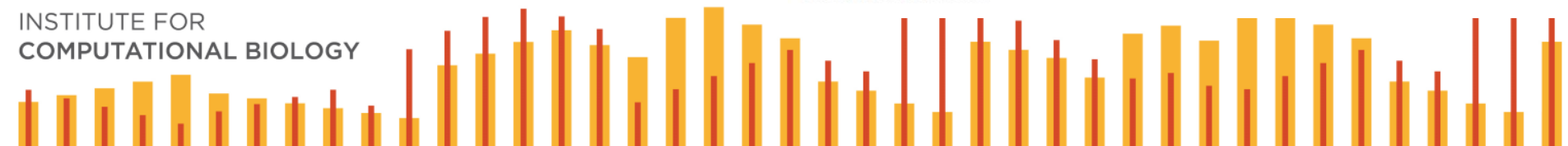
+ Author Affiliations

†To whom correspondence should be addressed. E-mail: [mpollak@bidmc.harvard.edu](mailto:mpollak@bidmc.harvard.edu) (M.R.P.); [epays@ulb.ac.be](mailto:epays@ulb.ac.be) (E.P.)

\* These authors contributed equally to this work.

*Science* 13 Aug 2010:  
Vol. 329, Issue 5993, pp. 841-845  
DOI: 10.1126/science.1193032

INSTITUTE FOR  
COMPUTATIONAL BIOLOGY



# PRECISION MEDICINE INITIATIVE COHORT PROGRAM



1 Million Cohort

Large

Linked to EHR and payer databases

Oversample subgroups

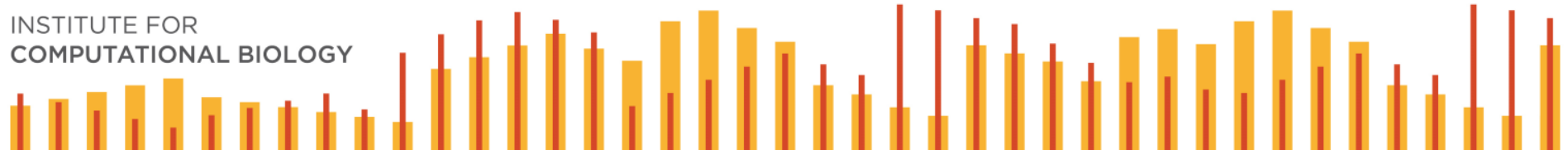
Re-contact for return of results

Wide age range

Sophisticated lifestyle data collection

Rare and common disease

Extensive follow-up





# SESSION SPEAKERS



- Study Design 101: The dos and don'ts in ascertainment for genomic studies

**Farren Briggs, PhD, ScM**

Case Western Reserve University



- Emerging resources for genomic assays that emphasize diverse populations

**Janina Jeff, PhD, MS**

Illumina, Inc.



- Crunching numbers: Bioinformatics and computational approaches for genomic studies

**William Bush, PhD, MS**

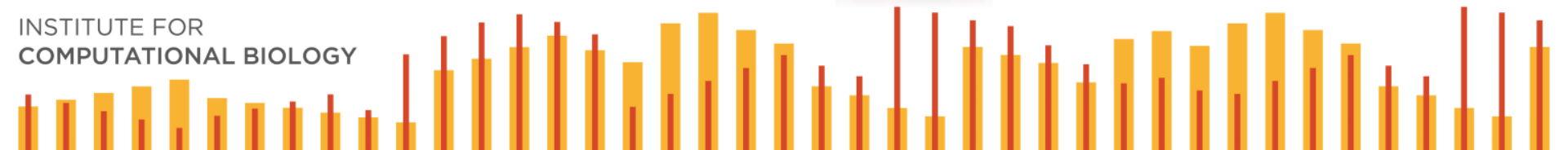
Case Western Reserve University



- Genomic research in action: phenome-wide association studies in diverse populations

**Sarah Pendergrass, PhD, MS**

Geisinger Health System





# Q&A TOPICS

Diversity in Research

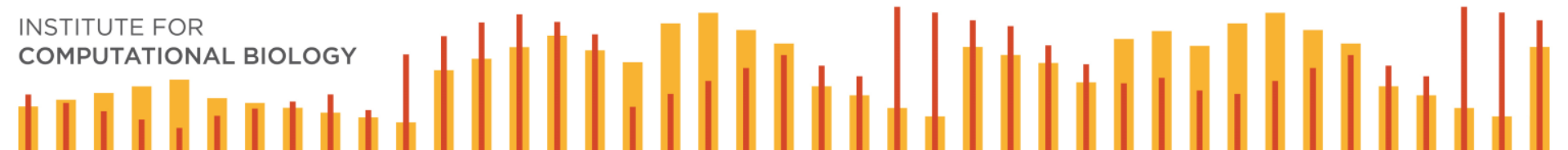
Study Design

Genotyping

Precision Medicine Initiative

Bioinformatics

PheWAS



# Q&A TOPICS

Careers

Degrees

Disciplines

Collaborations

Colleagues

Friends



# Q&A TOPICS



Big Data and Disease  
JNU, New Delhi, India



ABRCMS  
Nashville, TN



American Society for  
Human Genetics  
Baltimore, MD



Pacific Symposium on  
Biocomputing  
Big Island, Hawaii

INSTITUTE FOR  
COMPUTATIONAL BIOLOGY

