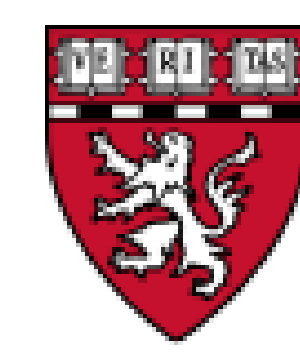# A customized genotype array to investigate the genetic architecture of tuberculosis in Peru

LIMAA

Y. Luo[1,2], L. Lecca[3], R. Calderon[3], C. Conteras[3], R. Yataco[3]; Z.B. Zhang[5], J. Galea[3]; B. Moody[8], M. Murray[3,4,5,6,7], S. Raychaudhuri[1,2]

HARVARD MEDICAL SCHOOL

BWH BRIGHAM AND WOMEN'S HOSPITAL
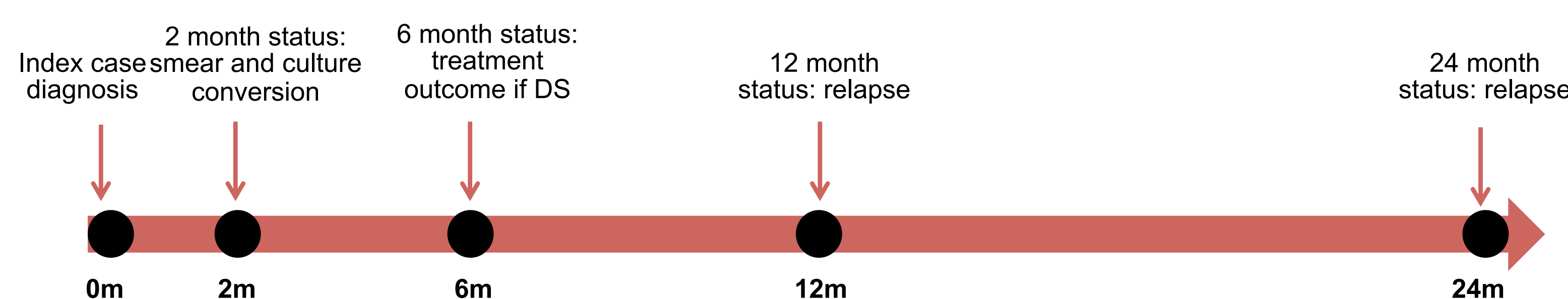
Background & Study Design

As a part of the Tuberculosis Research Unit program (TBRU) that aims to integrate basic and clinical research to study tuberculosis (TB) in countries where the disease is endemic, we are conducting

➤ the first genome-wide association study (GWAS) in a Latin American cohort in Lima, Peru

➤ a GWAS of **1500** active TB cases and **1500** *Mycobacterium tuberculosis* (*M.tb*) infected household contact controls
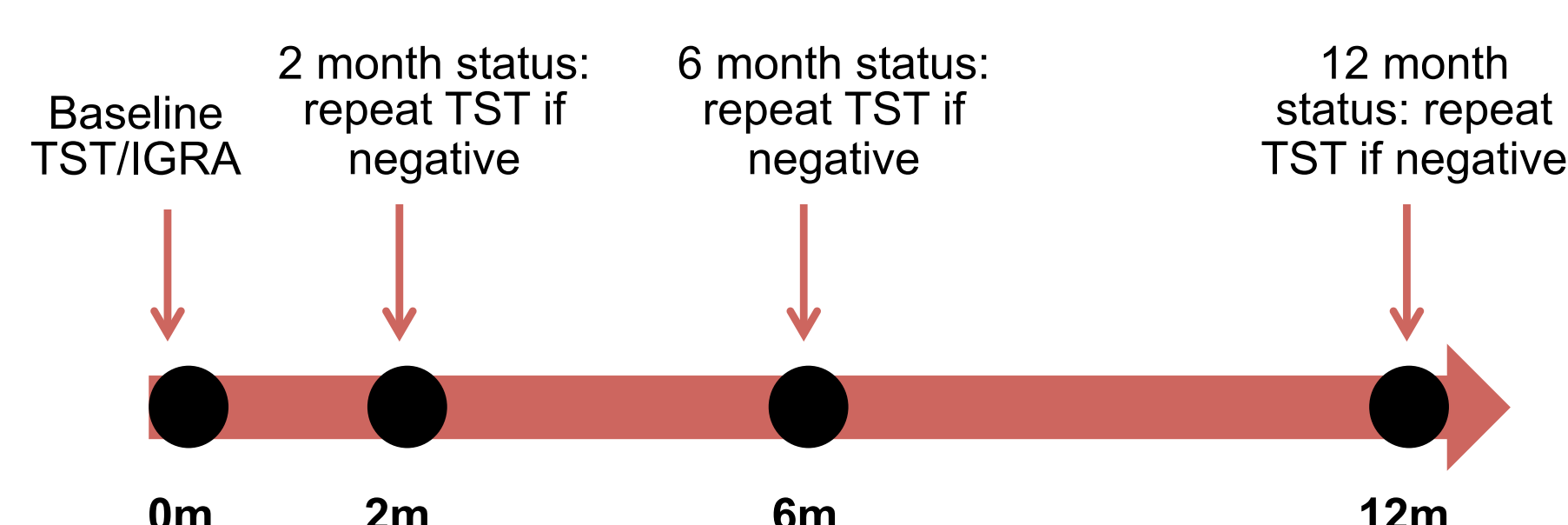
➤ first high phenotypic TB GWAS data

To further characterize variants specific to this group, especially those with tuberculosis, we initiated an exome sequencing pilot study of

➤ **116** patients with active TB

➤ cases without non-metabolic risk factors for TB (smoking, alcohol use, HIV) to enrich for possible genetic susceptibility to disease progression
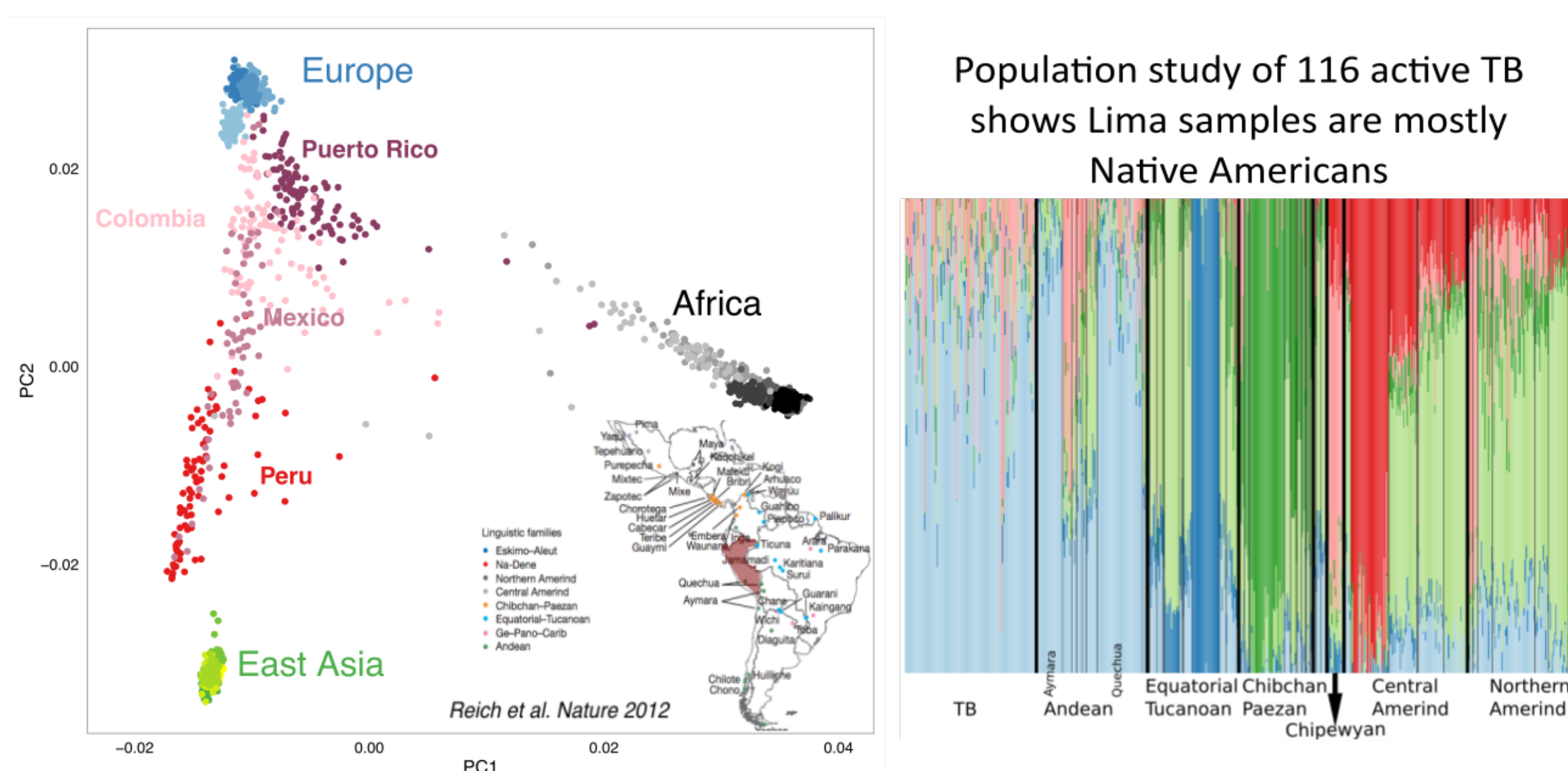
Index cases

Index case diagnosis — 0m
2 month status: smear and culture conversion — 2m
6 month status: treatment outcome if DS — 6m
12 month status: relapse — 12m
24 month status: relapse — 24m

Household contacts

Baseline TST/IGRA — 0m
2 month status: repeat TST if negative — 2m
6 month status: repeat TST if negative — 6m
12 month status: repeat TST if negative — 12m

Index case outcomes:
- drug resistance status
- strain type
- treatment outcomes
- relapse

Unique Genetic Heritage of Peruvian Population

Reich et al. Nature 2012

Population study of 116 active TB shows Lima samples are mostly Native Americans

LIMAArray Design

Tuberculosis related genes
✓ all GWAS reported TB genes (Tier 1)
✓ literature based TB/TB-related genes (Tier 2)

Whole exom sequencing (WES) of 116 active TB cases at 41x
✓ variants have MAF$_{Lima}$> MAF$_{all}$
✓ all coding variants

Whole genome sequencing (WGS) of 85 Peruvians at 7x from 1000 Genome Project
✓ all variants with MAF > 0.1% in upstream/downstream of TB genes in Tier 1
✓ all variants with MAF > 0.5% in upstream/downstream of TB genes in Tier 2
✓ all variants with MAF > 1% & MAF$_{Peru}$ > MAF$_{all}$

NHGRI GWAS catalog

Ancestry markers

HLA markers

Autoimmune/inflammatory markers

~320,000 core content

| Module	Number of Markers
GWAS / HLA / Immune-related markers	8.1K
Ancestry markers	4.5 K
WES/WGS core markers	302.4 K
Markers for genomic coverage	397.2 K
Total	712.2 K

Evaluation

To evaluate the performance of *LIMAArray*, we

➤ imputed using 1000 Genomes Project Phase 3 reference panel

➤ cross-validated of **1/5** of Peruvian samples in 1000GP (10 runs)

➤ compared against *Axiom UK BioBank* genotype array

➤ calculated minor allele frequency in Peruvian population
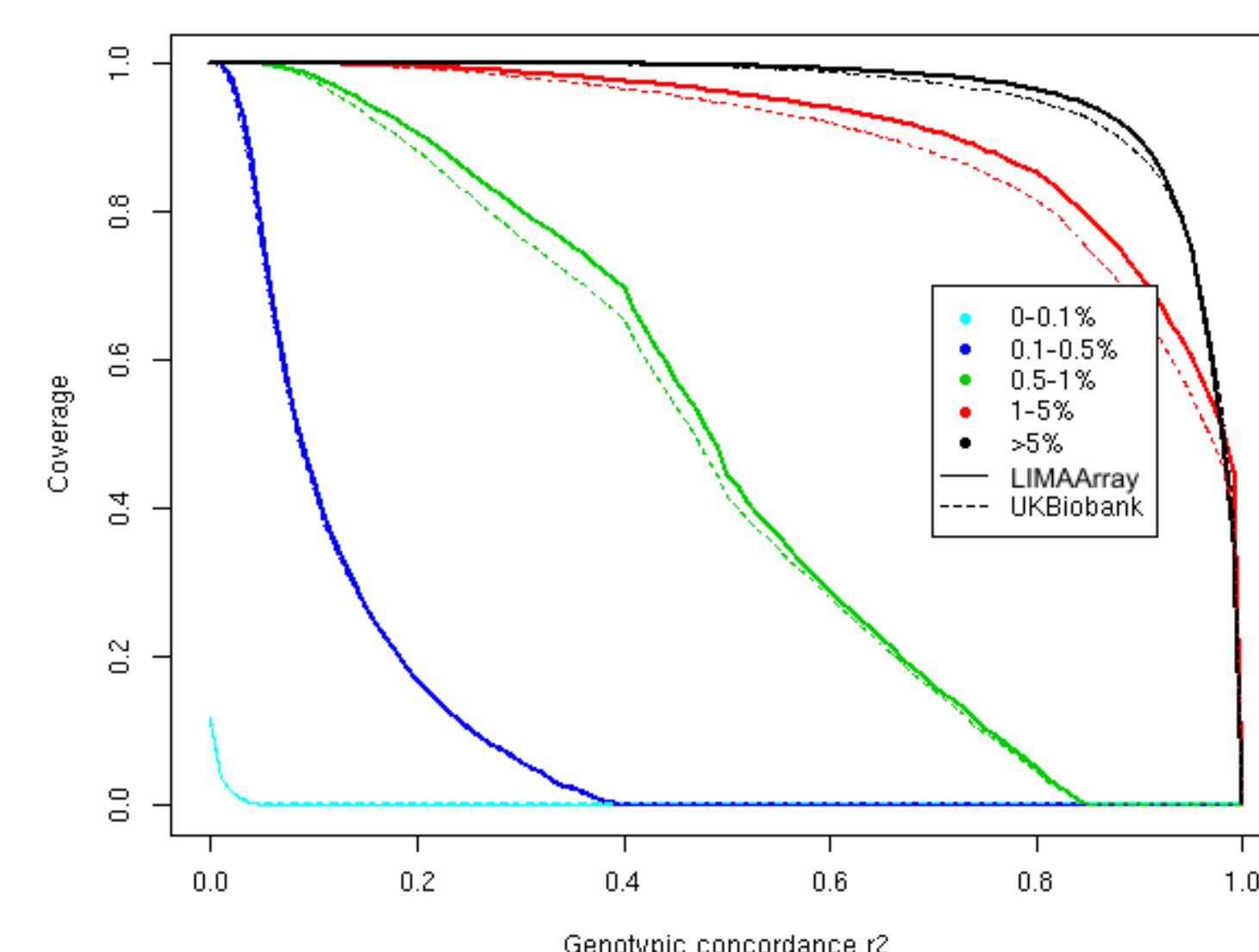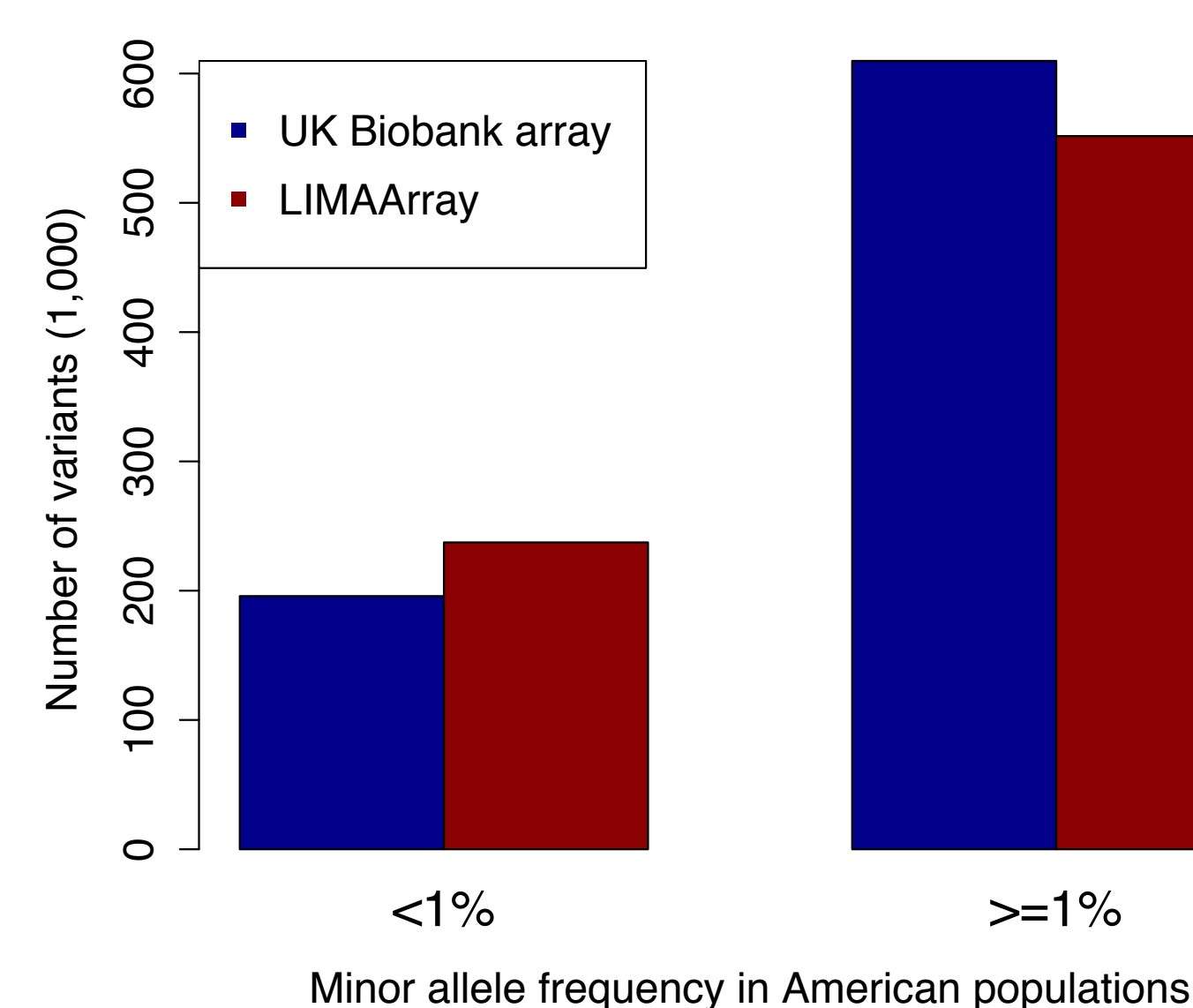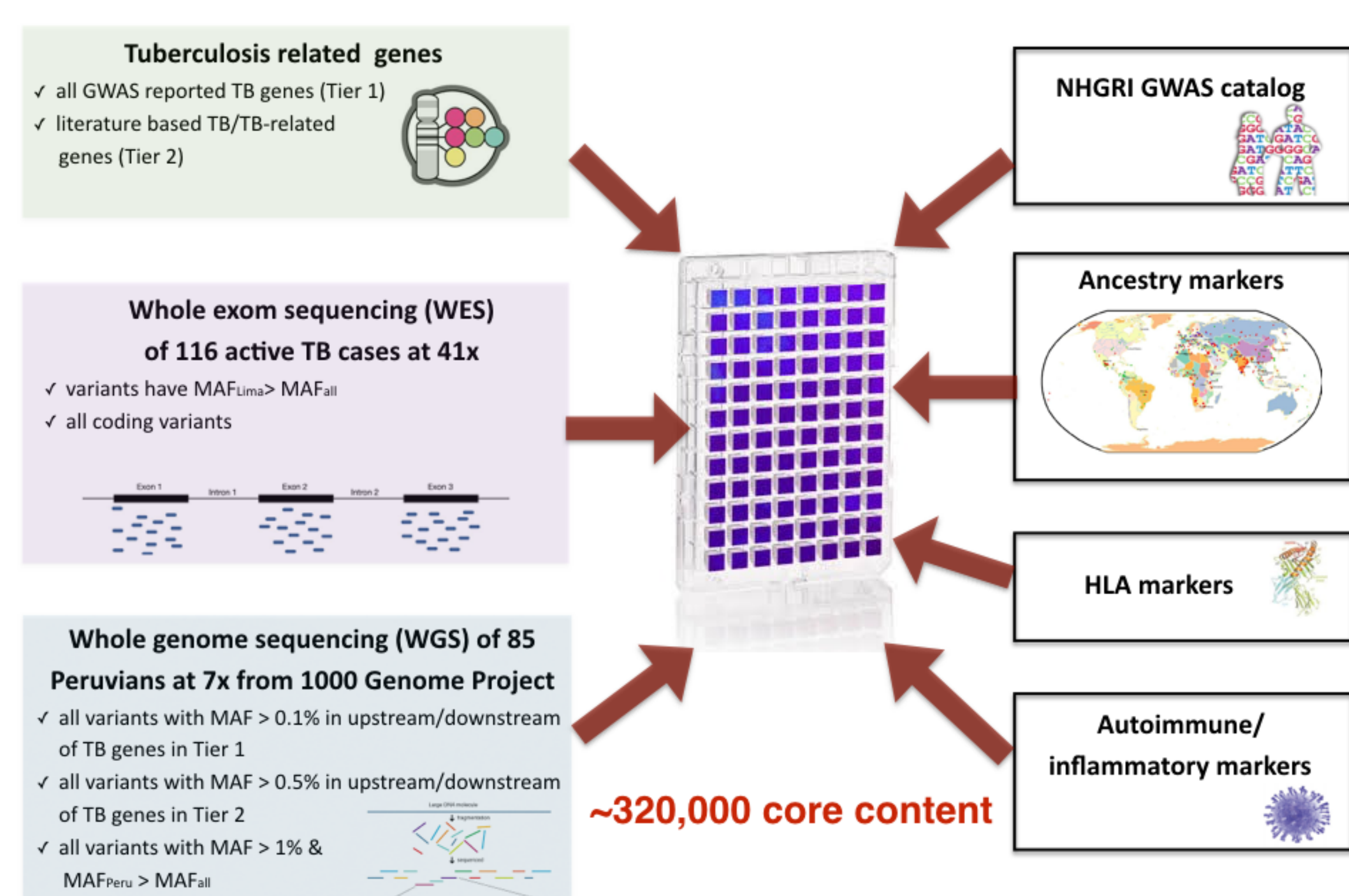
Summary and Ongoing Work

➤ Identified unique genetic heritage of Peruvian population

➤ Designed a customized genotyping array tailored for Lima population, and demonstrated higher accuracy and coverage in Peruvian population in both common and rare variants

➤ First TB GWAS in Latin America

➤ First TB GWAS cohort has the ability to adjust for exposure differences

➤ Opportunities to investigate relationships between TB and metabolic traits

Contact Information

Yang Luo, PhD
E-mail: yangluo@broadinstitute.org
Twitter: yluo86

[1]Division of Genetics and Rheumatology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA.; [2] Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA, USA; [3] Socios En Salud, Lima, Peru; [4] Partners In Health, Boston, MA, USA; [5] Division of Global Health Equity, Brigham and Women's Hospital, Boston, MA, USA; [6] Department of Global Health and Social Medicine, Harvard Medical School, Boston, MA, USA; [7] Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA; [8] Division of Rheumatology, Immunology and Allergy, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA