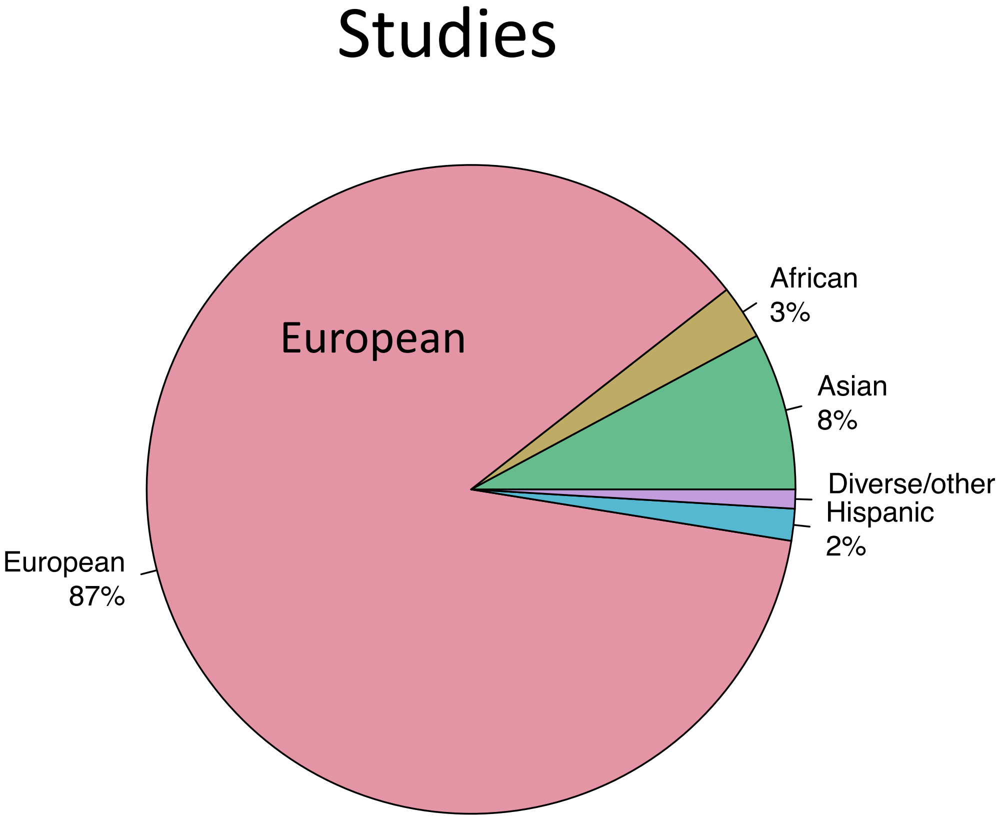
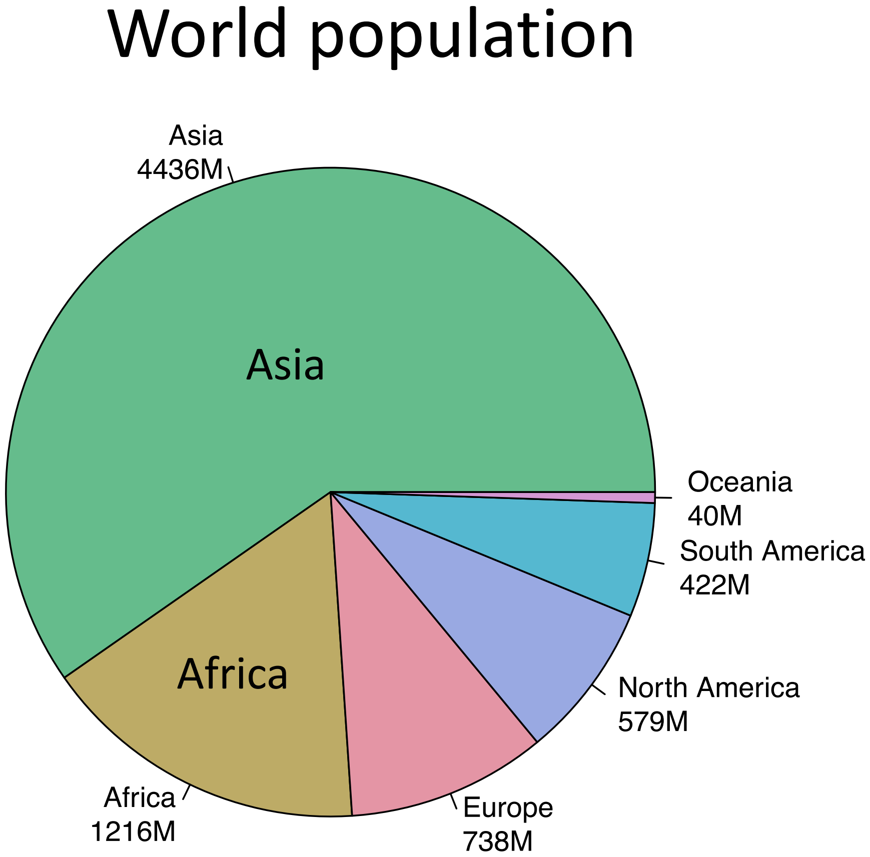
# Trans-ethnic colocalization: A novel approach to assess the transferability of trait loci across populations

Dr Karoline Kuchenbäcker

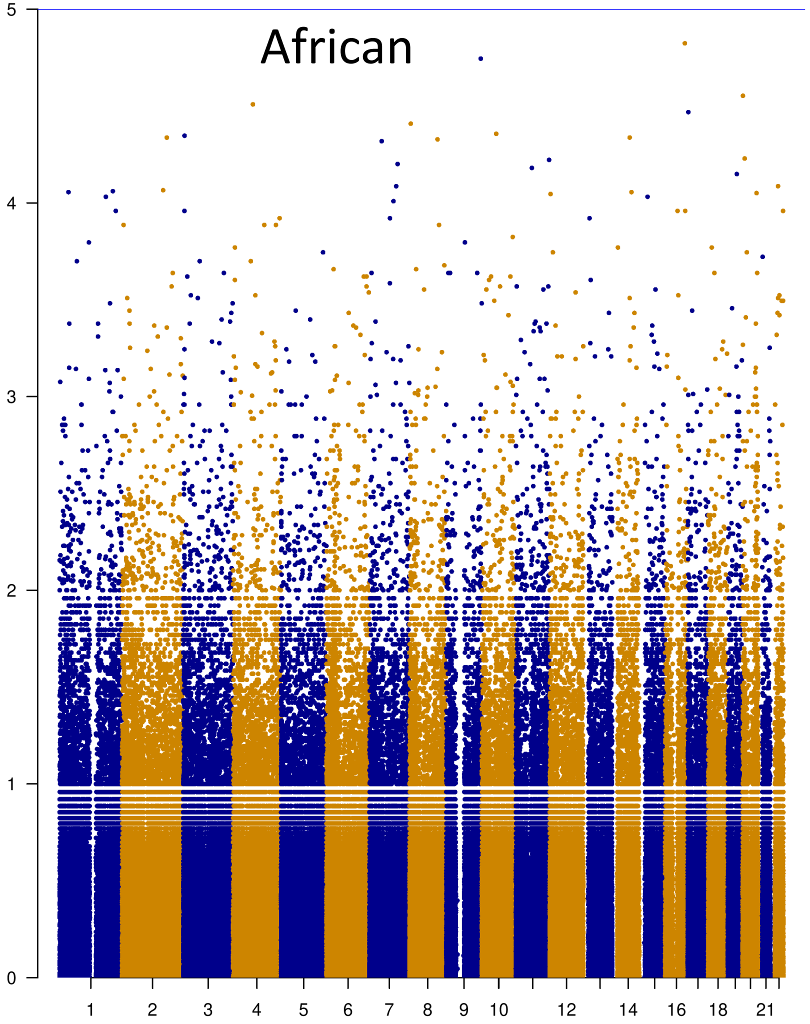
UCL Genetics Institute & Division of Psychiatry
email: k.kuchenbaecker@ucl.ac.uk
Twitter: @KKuchenbaecker

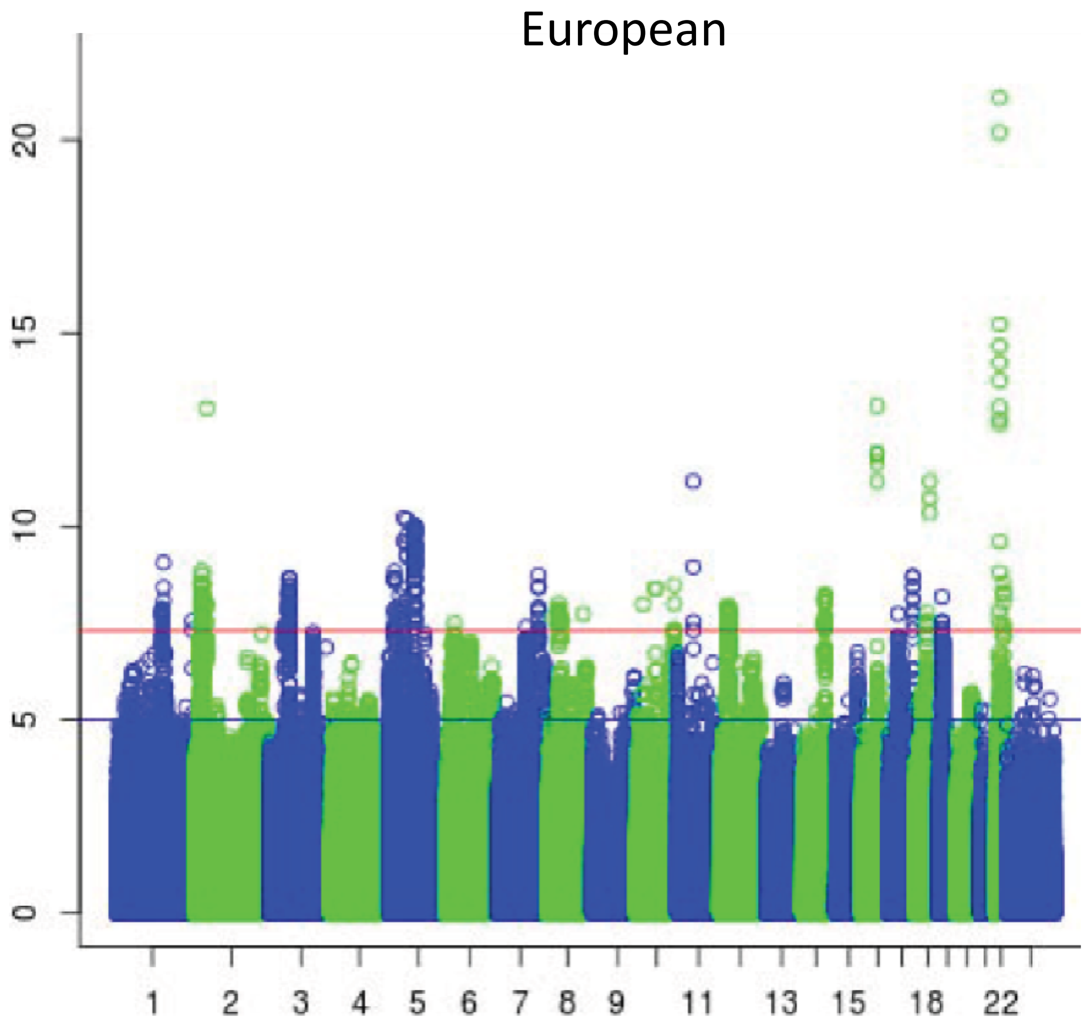
# Genome WHITE association studies



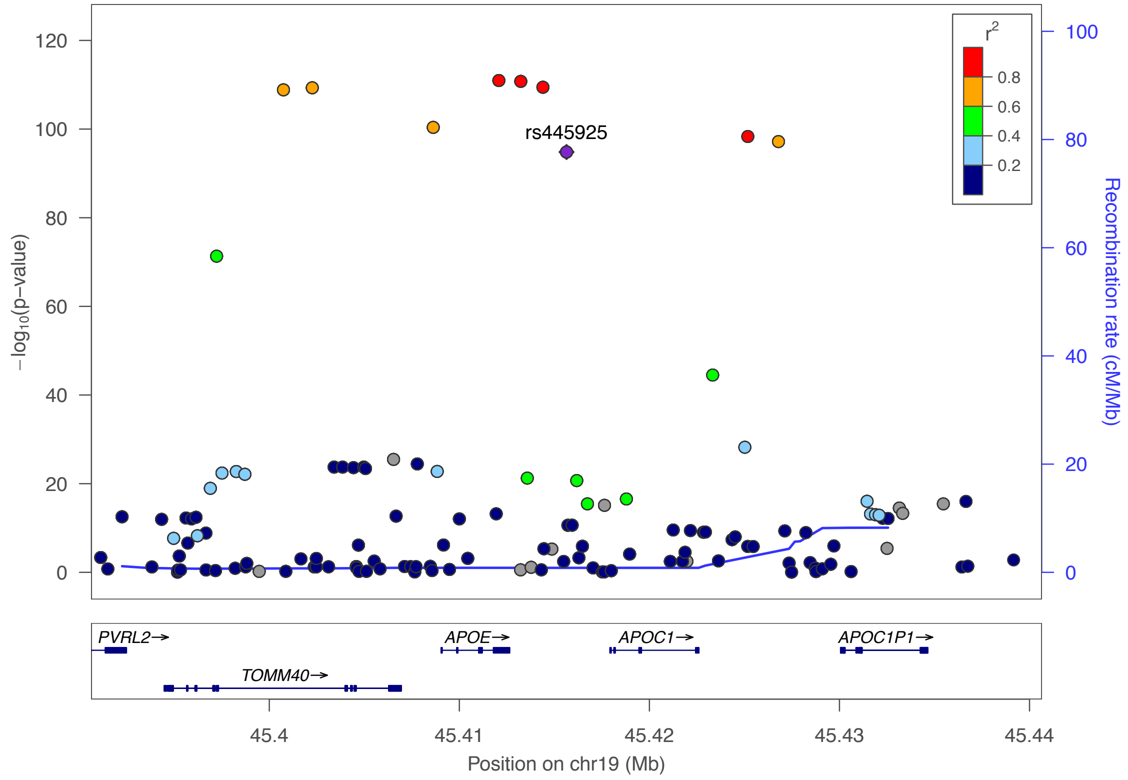


# Comparing GWAS - BC

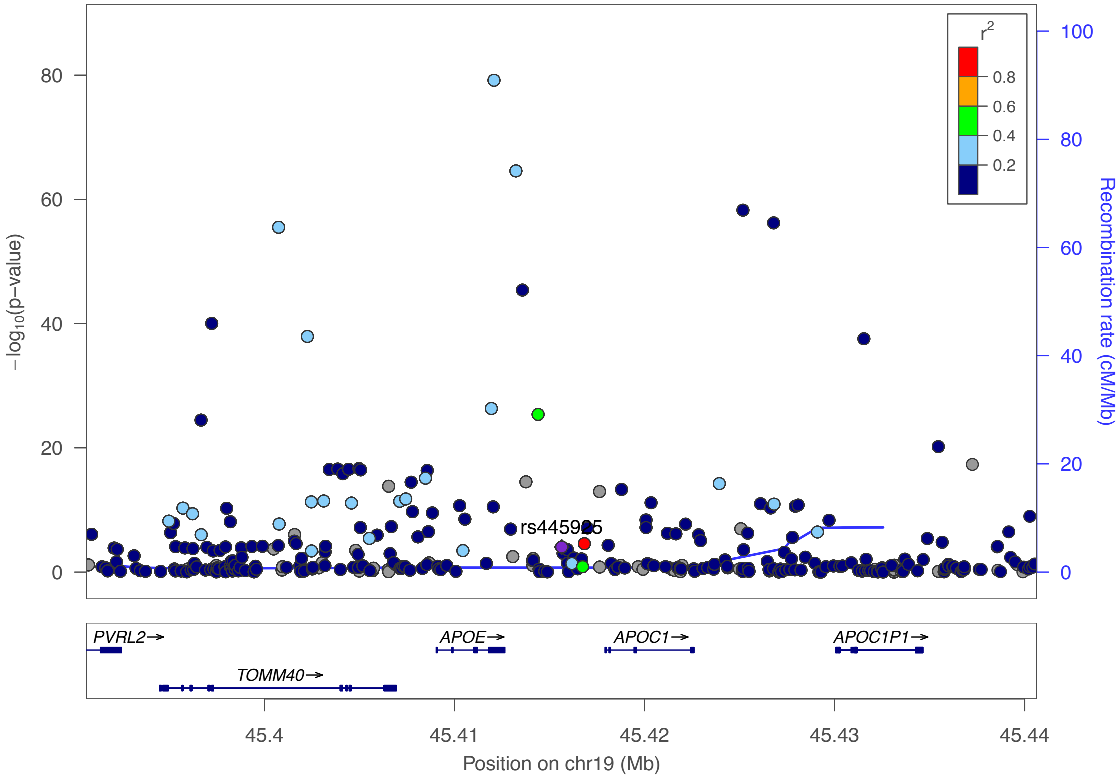




# Same causal variant?

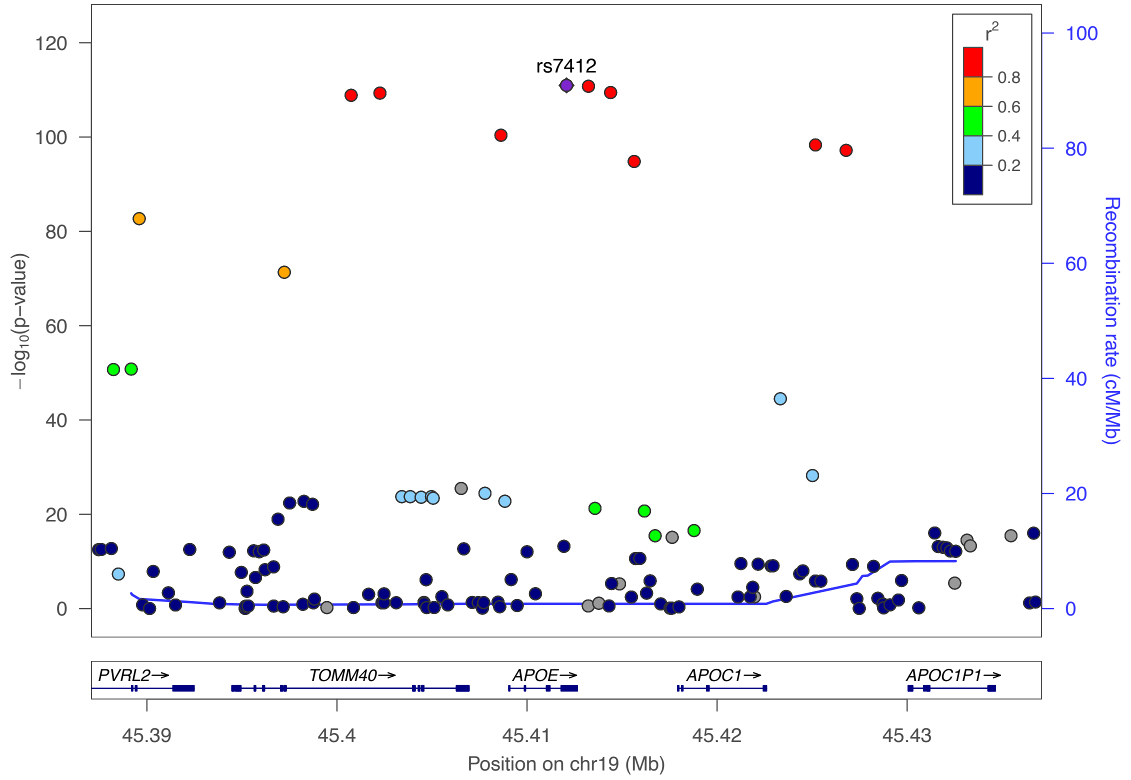


Asian samples

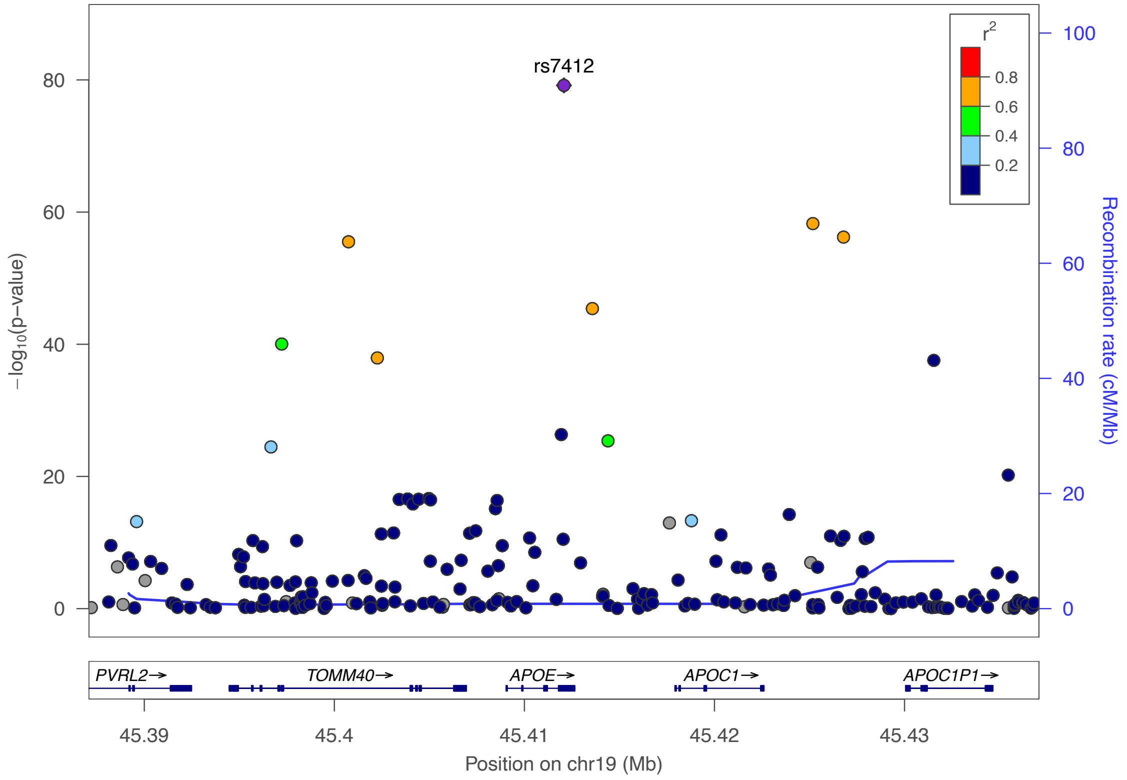


African samples

# Same causal variant?

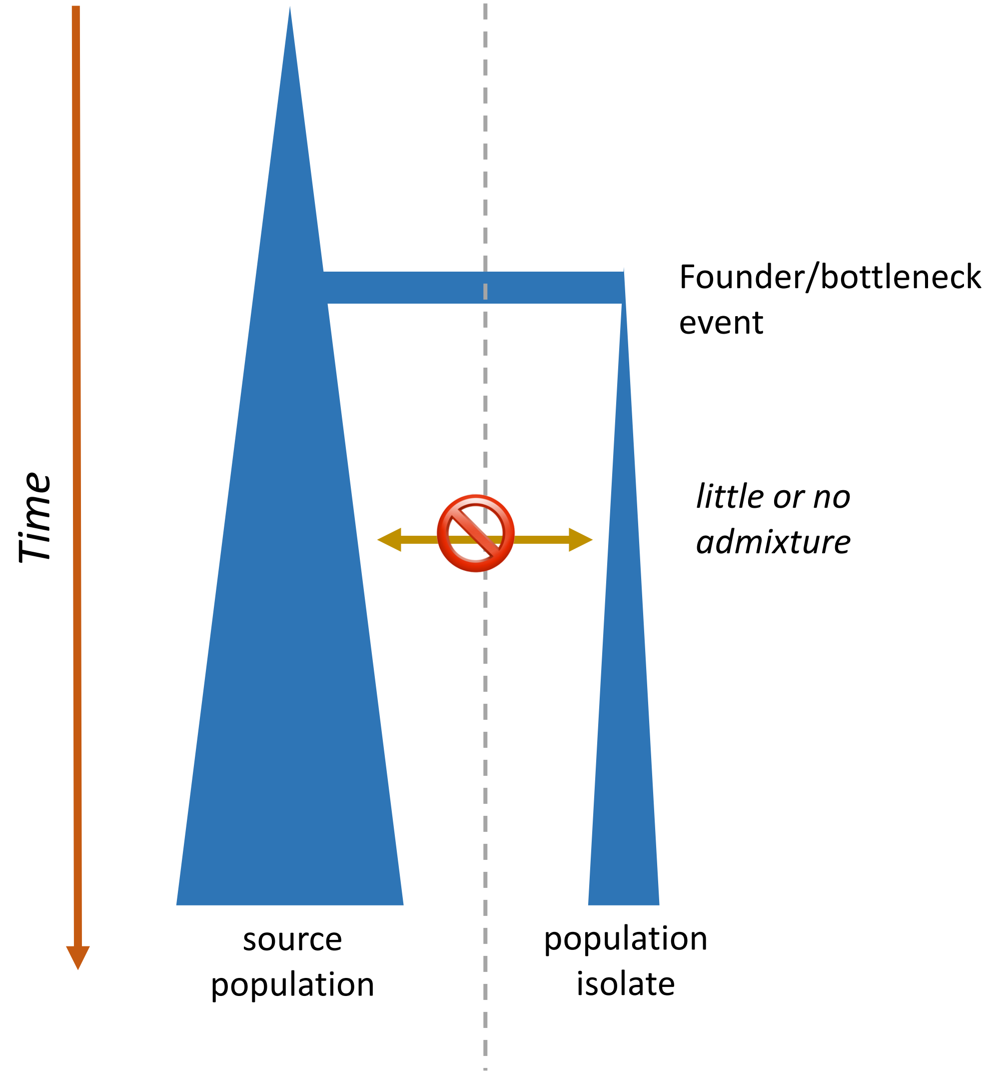


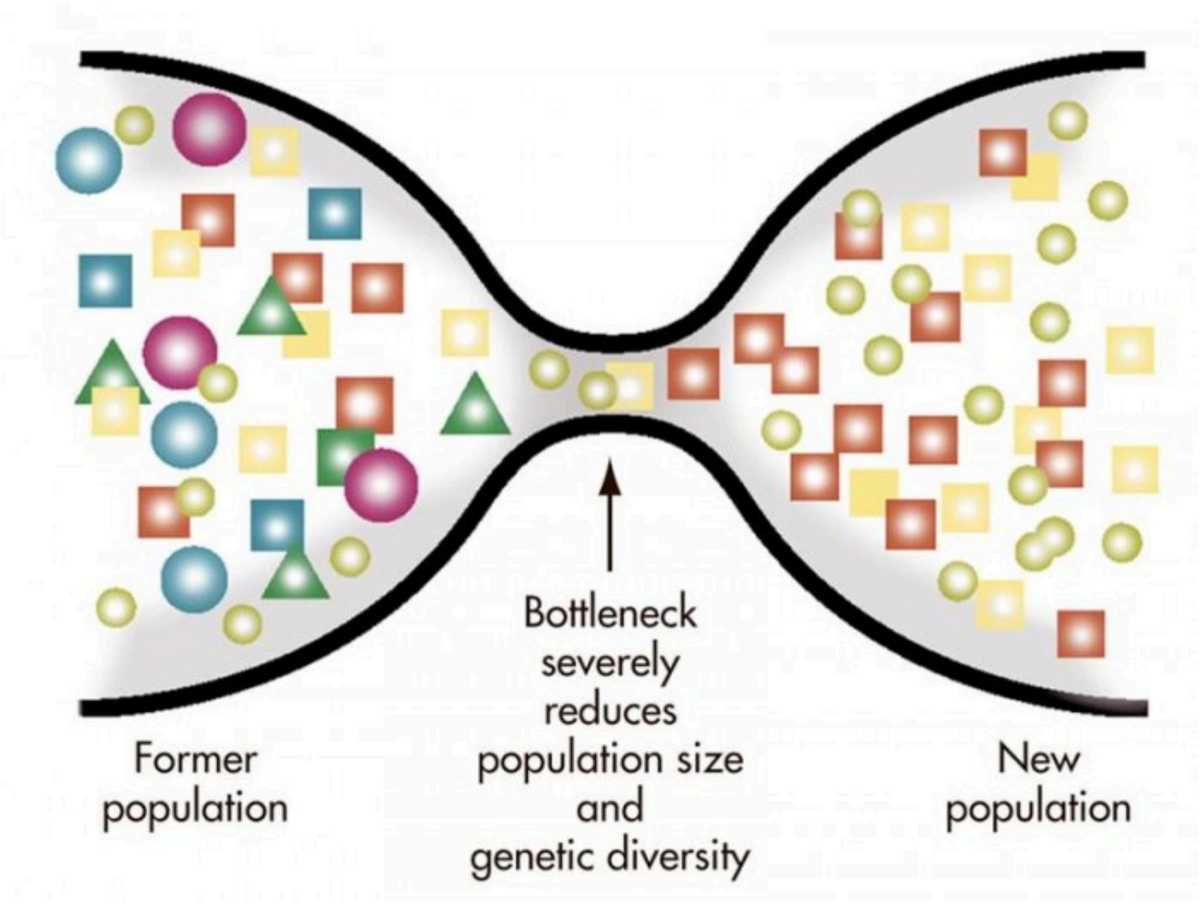
Asian samples



African samples

# Populations and genetic diversity



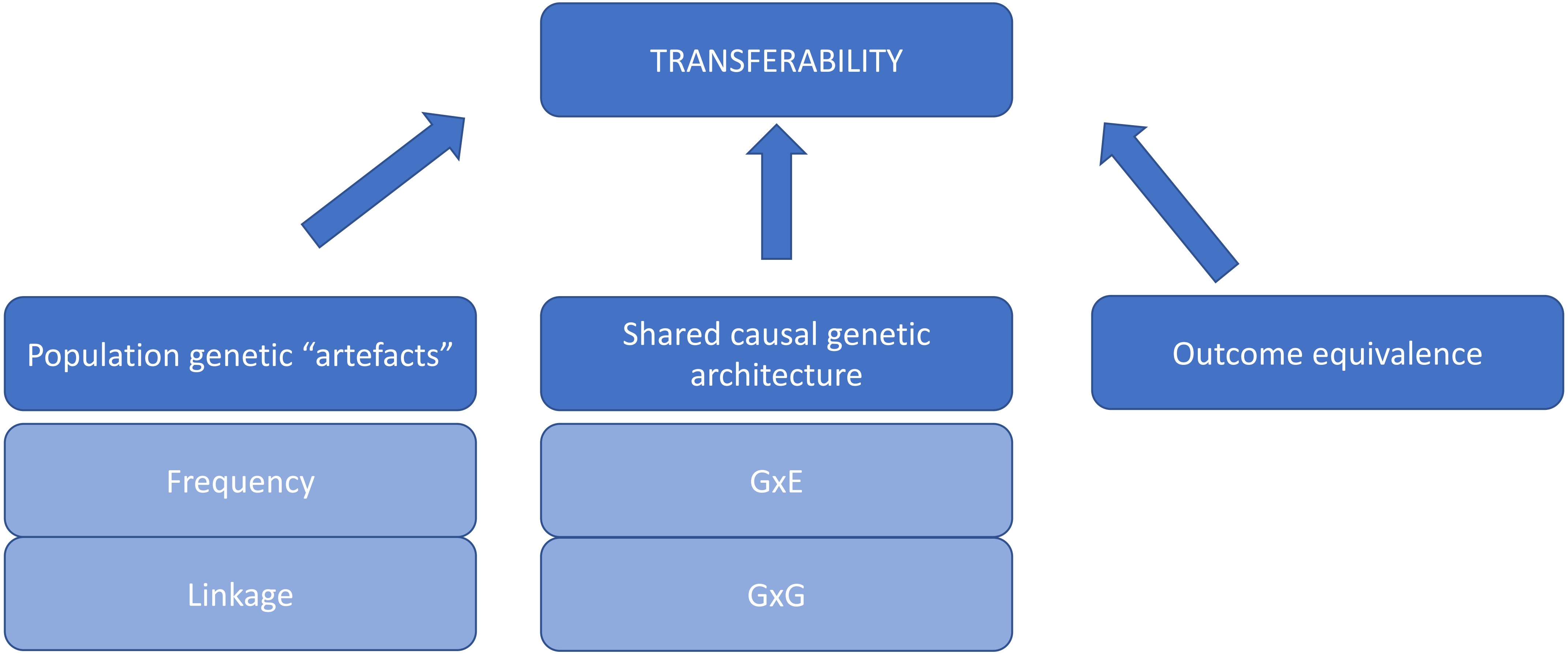


# Frequency, linkage, causal effects

$$y = b_j x + e$$

$$\hat{b}_j = \sum_{k=1}^{K} r_{jk}\, b_k^{(J)} + e_j$$

$$(1 - \beta) = f\left(r_{jk}, b_k^{(J)}, N, p, r_{info}^2\right)$$



TRANSFERABILITY

Population genetic “artefacts”

- Frequency
- Linkage

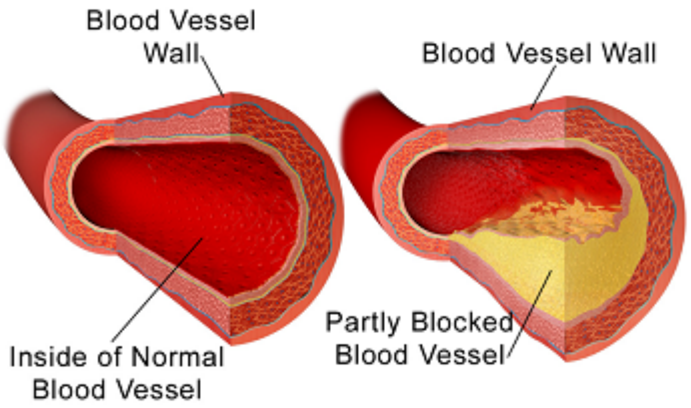
Shared causal genetic architecture

- GxE
- GxG

Outcome equivalence

# Lipid Biomarkers

- Lipid levels are a risk factor for cardiovascular disease
- Lipid levels are heritable
  - e.g., LDL: narrow sense heritability ($h^2$= 19% Europeans)
- >440 loci identified in GWAS
- can help identify potential drug target (PCSK9)



**Normal and Partly Blocked Blood Vessel**

Nordestgaard, 2010, *BG*

# Sample sets

**European ancestry (GLGC)**
N = 188,577 / 237,050
Willer et al, 2013, *Nat Gen*
Liu et al, 2017, *Nat Gen*

**UK (UKHLS)**
N = 9,961
Prins et al, 2017, *Sci Rep*

**Uganda (APCDR)**
N = 6,407

**China (CKB)**
N = 25,000

**Japan (RIKEN)**
N = 128,305
Kanai et al, 2018, *Nat Gen*

**Isolated Greek populations (HELIC)**
N = 3,586
Gilly et al, 2018, *Nat Com*

https://doi.org/10.1101/525170

# Trans-ethnic colocalization

$$\Lambda = \sum_{i \in N_\theta^1(m^*)} L_1(i) \times log \frac{L_1(i)L_2(i)}{\max\limits_{j \notin N_\theta^2(i)} L_1(i)L_2(j)}$$

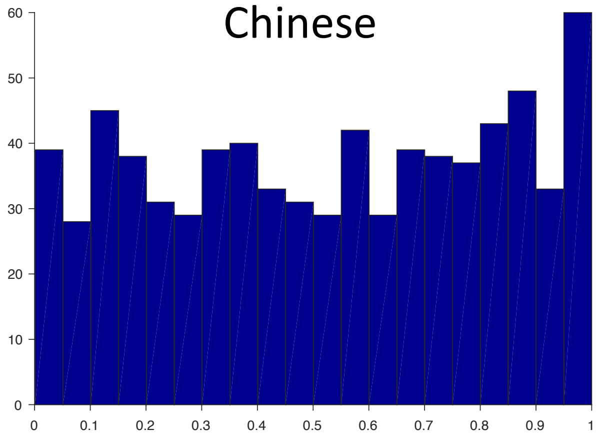
$$logL\left(Z; \lambda_i^{mle}\right) = \frac{1}{2\left(-Z^T\Sigma^{-1}Z + z_i^2\right)} - 1/2log\left((2\pi)^M|\Sigma|\right)$$

Chun et al, 2017, *Nat Gen*

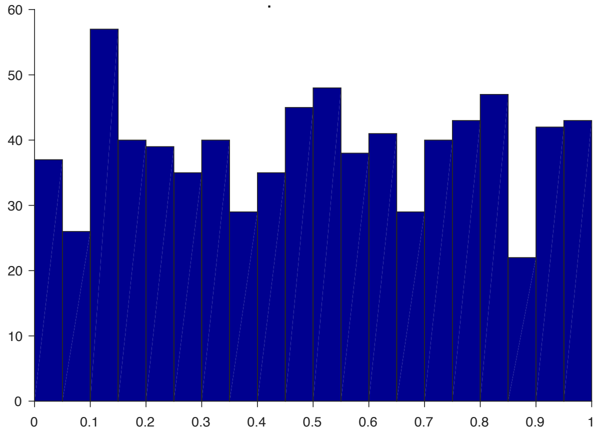
$i$ SNP
$m^*$ lead SNP
$L_1(i)$ likelihood of SNP i being causal for trait 1
$L_2(i)$ likelihood of SNP i being causal for trait 2
$N_\theta^1(i)$, $N_\theta^2(i)$ sets of SNPs in LD with i
$\theta$ LD threshold

# Simulation

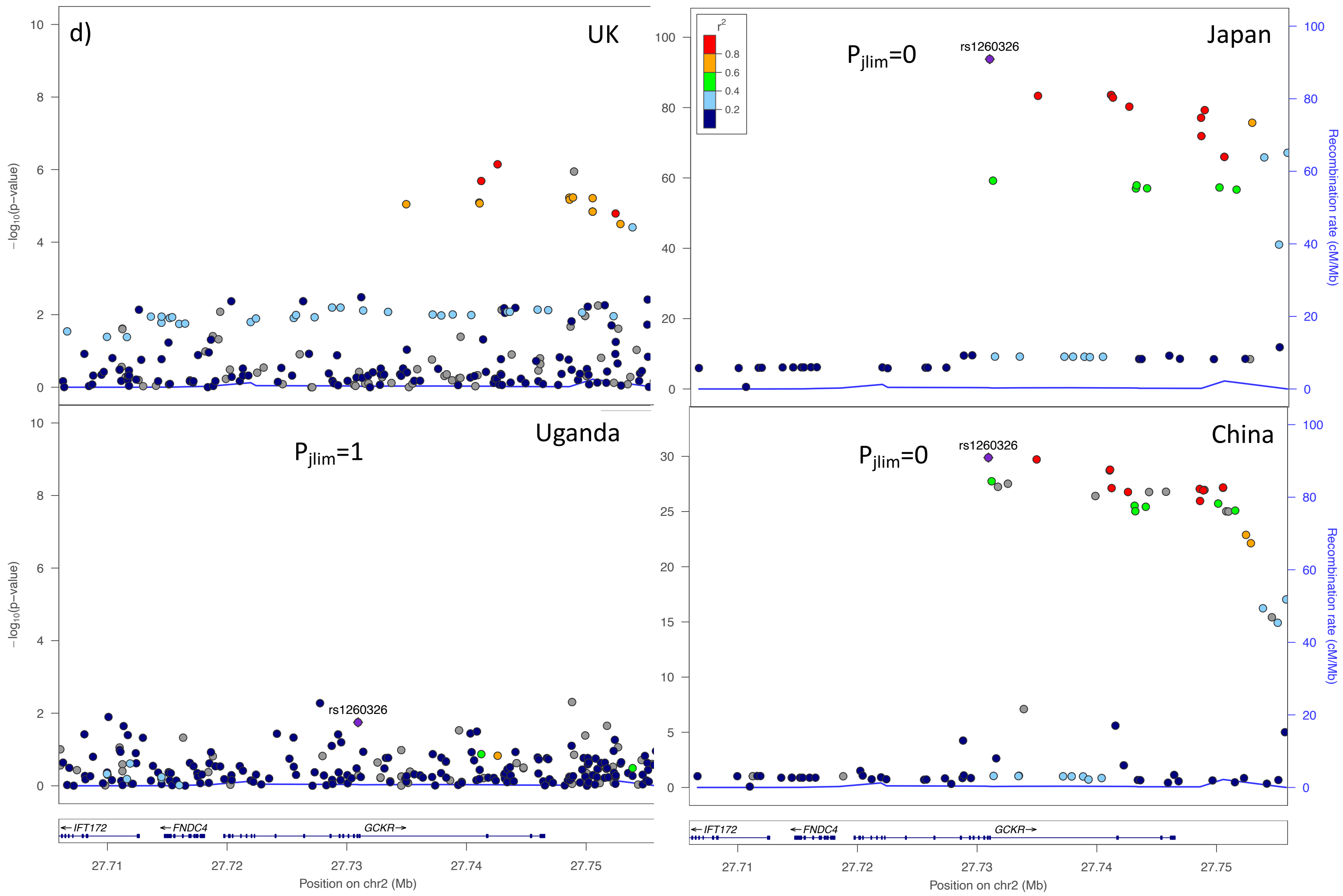
| Sample, N | Type I error rate | Power |
|---|---|---|
| CKB, 72,473 | 5.2% | 93.1% |
| APCDR-Uganda, 4,597 | 4.6% | 73.1% |
| CKB, 4,597 | 4.5% | 94.8% |


Chinese

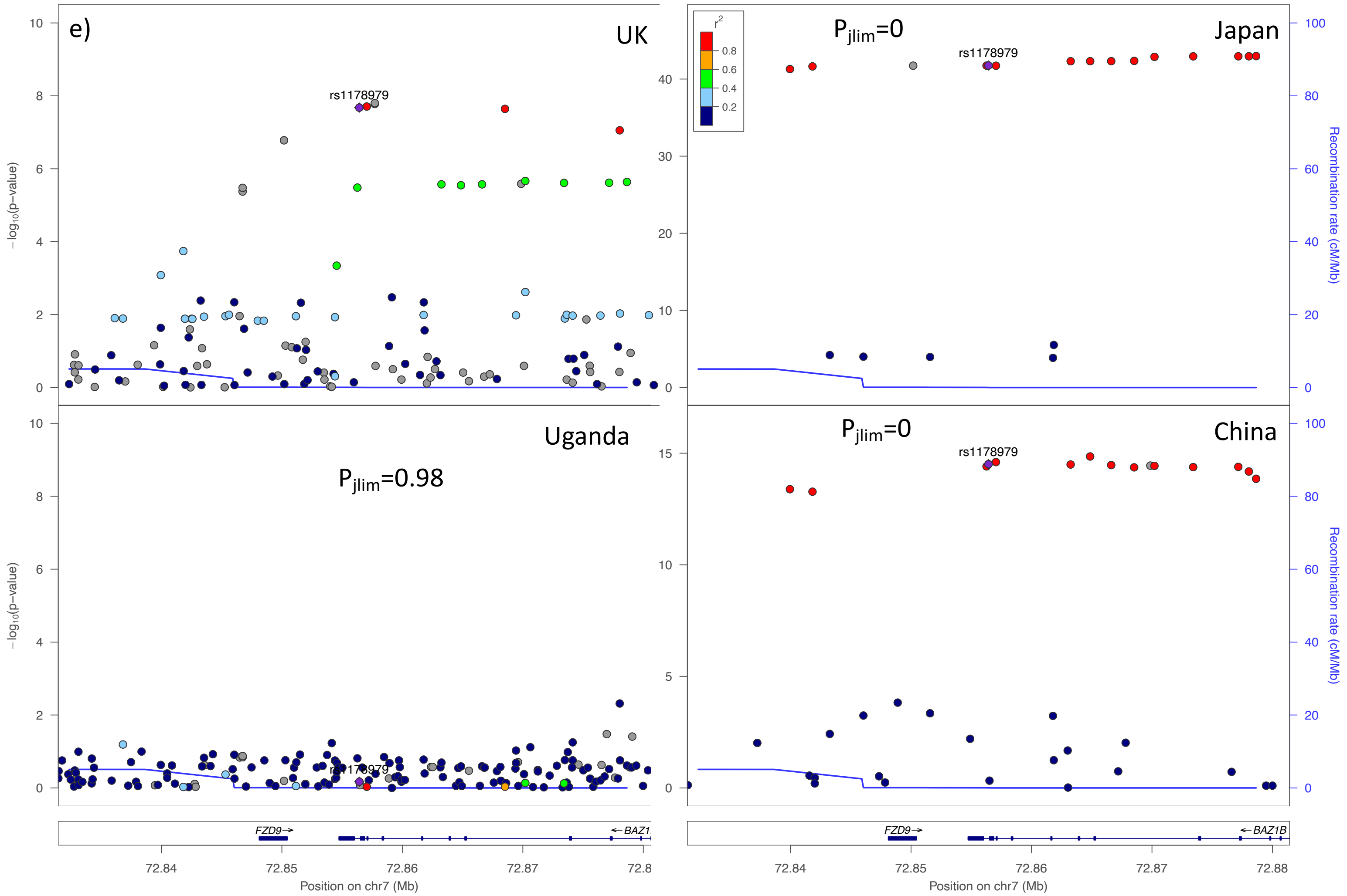
Ugandan



# 2p23 *GCKR* - TG


d)
UK
Uganda
Japan
China
$P_{jlim}=1$
$P_{jlim}=0$
$P_{jlim}=0$
rs1260326
rs1260326
rs1260326
$r^2$
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
IFT172
FNDC4
GCKR
27.71
27.72
27.73
27.74
27.75
Position on chr2 (Mb)
0
2
4
6
8
10
20
40
60
80
100
5
15
25
30

# 7q11.23 *BAZ1B* - TG



# BMI associations

by replication status in Ugandan samples

| **Repli.** | **Trait** | **rsid** | **Annotation** | **P-value** |
|---|---|---|---|---|
| no | HDL | rs11755393 | UHRF1BP1 | **9.8x10$^{-48}$** |
| | | rs1178979 | BAZ1B | **3.1x10$^{-6}$** |
| | | rs4731702 | KLF14 | **3x10$^{-7}$** |
| | | rs2954033 | NSMCE2 | **6.4x10$^{-8}$** |
| | LDL | rs4245791 | ABCG8 | 0.22 |
| | | rs3846662 | HMGCR | **1.9x10$^{-35}$** |
| | | rs2737229 | TRPS1 | **1.9x10$^{-15}$** |
| | | rs635634 | IL6R | 0.03 |
| | | rs2000999 | HPR | **8.6x10$^{-8}$** |
| | TG | rs1260326 | GCKR | **1.2x10$^{-10}$** |
| | | rs2943641 | IRS1 | **5.8x10$^{-4}$** |
| | | rs6905288 | VEGFA | **1.9x10$^{-9}$** |
| | | rs11820589 | APOA5 | 0.41 |
| | | rs58542926 | TM6SF2 | 0.33 |
| yes | HDL | rs643531 | TTC39B | 0.92 |
| | | rs1800588 | LIPC | 0.25 |
| | | rs3764261 | CETP | 0.39 |
| | | rs16942887 | PSKH1 | 0.06 |
| | LDL | rs12740374 | CELSR2 | 0.18 |
| | | rs1367117 | APOB | 0.19 |
| | | rs6511720 | LDLR | 0.03 |

# Conclusions

- TEColoc can distinguish shared from population-specific causal variants
- Lack of replication of TG associations in Ugandan samples
- Possible gene-environment interactions (diet)

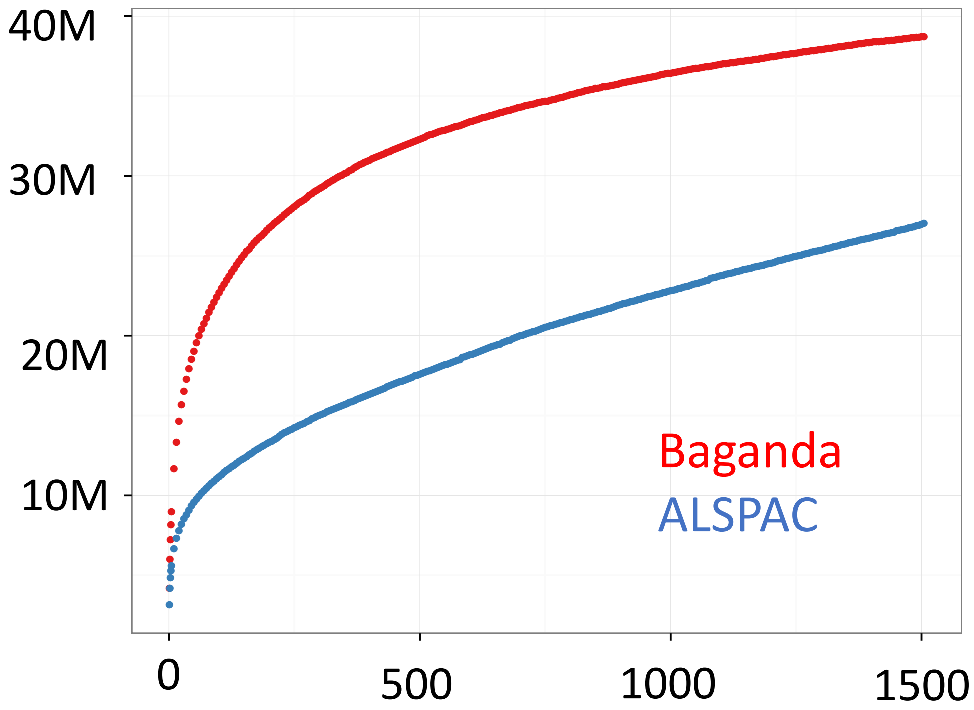
➢ cautions against generalising findings across populations

➢ locus comparison across ancestry groups can yield new insights

# Diverse populations

## European samples contain only a subset of human genetic variation

- \+ Different variants
- \+ Increased frequency
- \+ Stronger effects
- \+ Different LD empowers FM



# Acknowledgements

DOI: 10.1038/s41467-019-12026-7

The study participants

Anders Eriksson, KCL
Chris Harrison, Sanger

Oxford
Robin Walters
Zhengming Chen

Helmholtz Munich
Eleftheria Zeggini
Arthur Gilly

Cambridge
Manj Sandhu
Deepti Gurdasani