



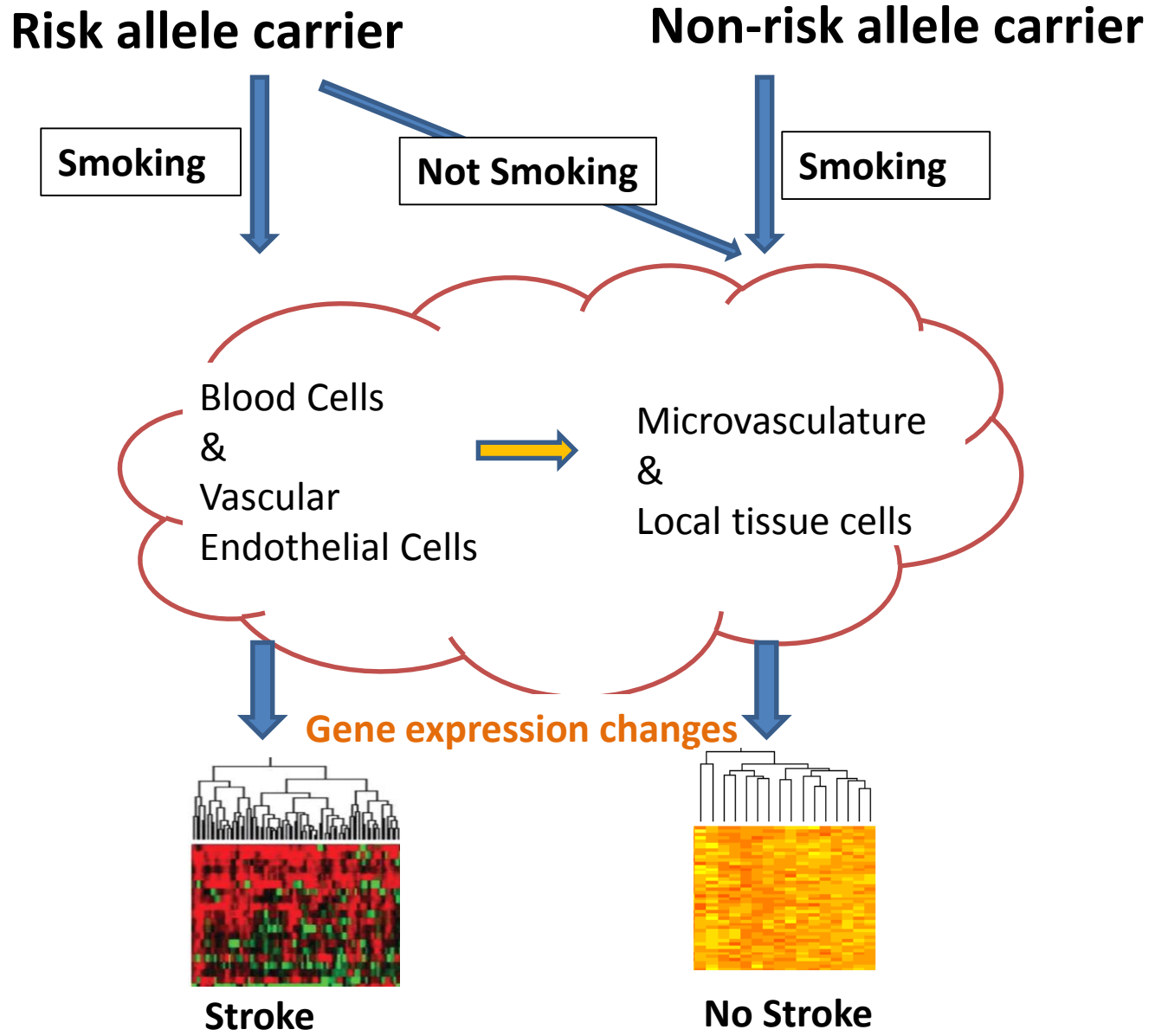
UNIVERSITY of MARYLAND
SCHOOL OF MEDICINE

Systemic Identification of Genes and Environment Interaction for the Risk of Stroke Using Functional Genomics

Huichun Xu, MD, PhD

University of Maryland School of Medicine

G*E interaction in Stroke



G*E interaction in Stroke:

Functional Candidate Gene Approach

STEP 1: Define genes whose expression altered by smoking

- tissue types: blood, alveolar macrophages, endothelial cells
- literature and publicly available microarray datasets

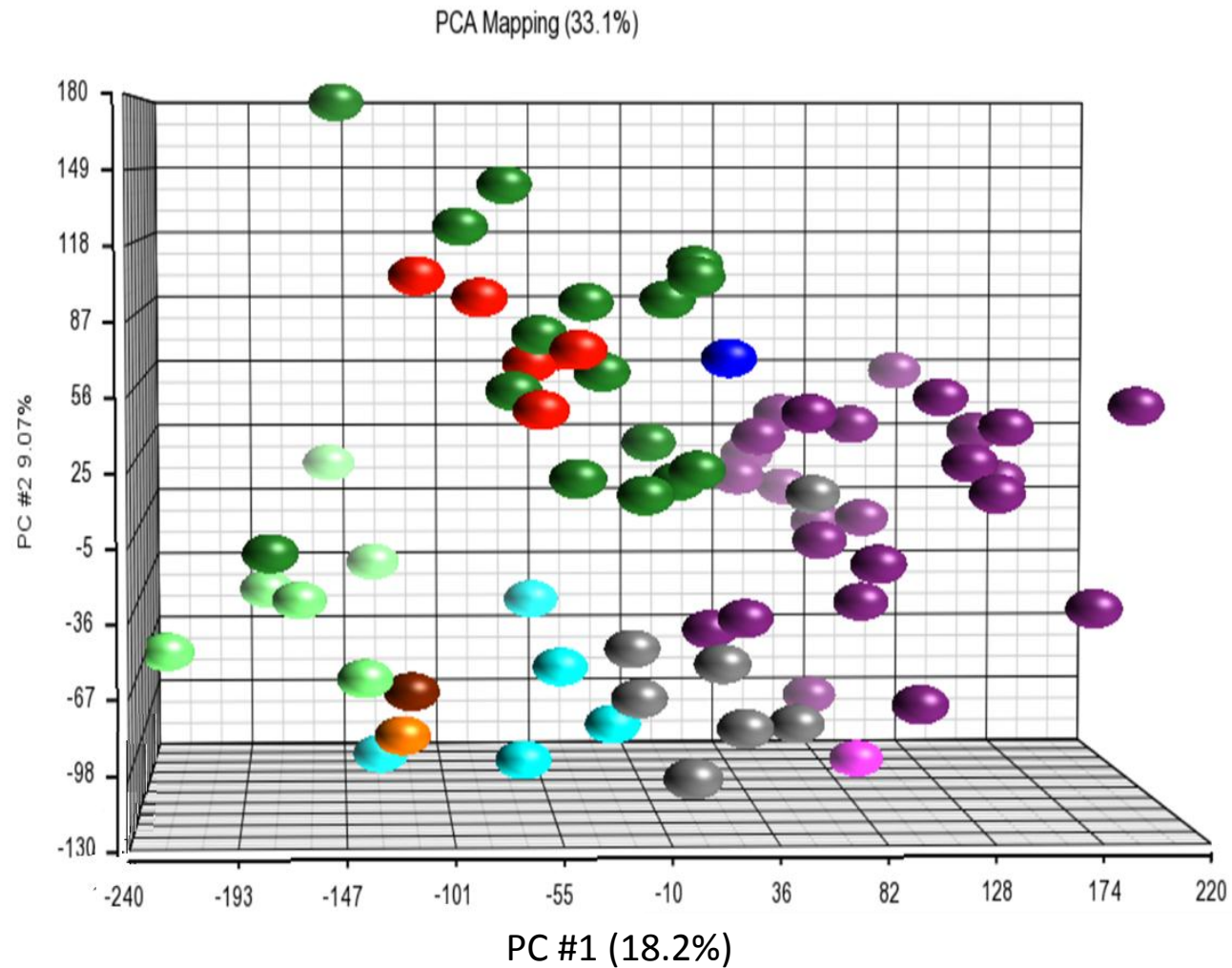
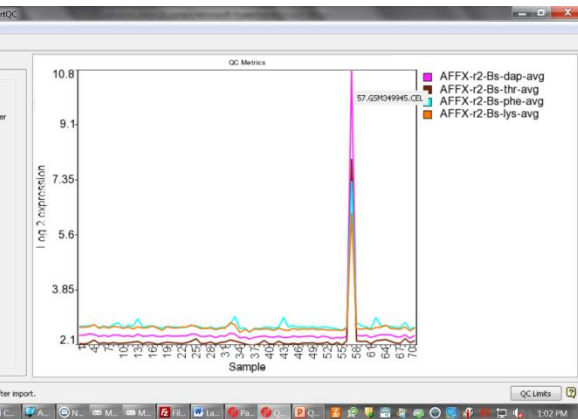
STEP 2: Define SNPs relevant to target genes

- within (+/- 10 kb) target genes
- eQTLs of target genes

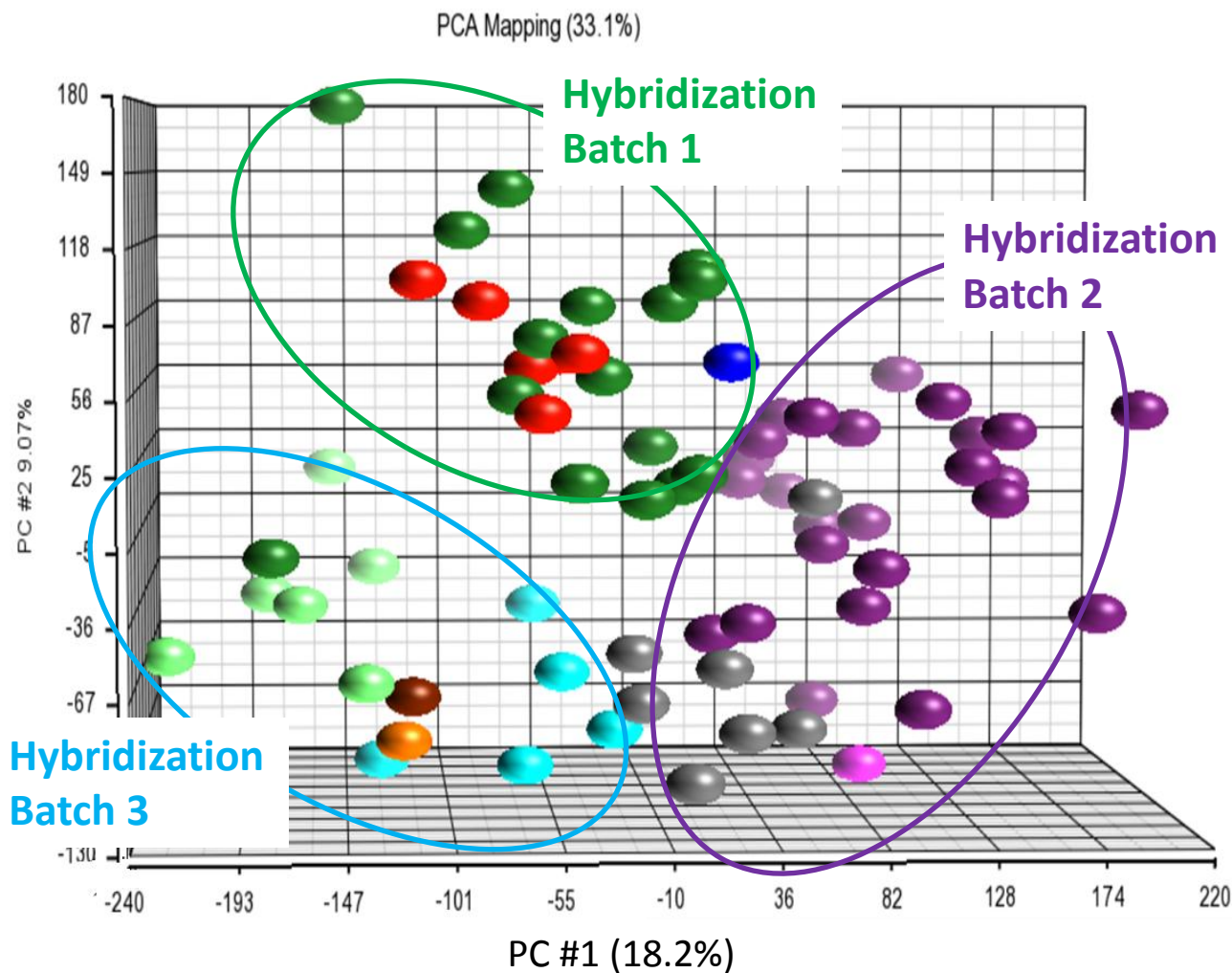
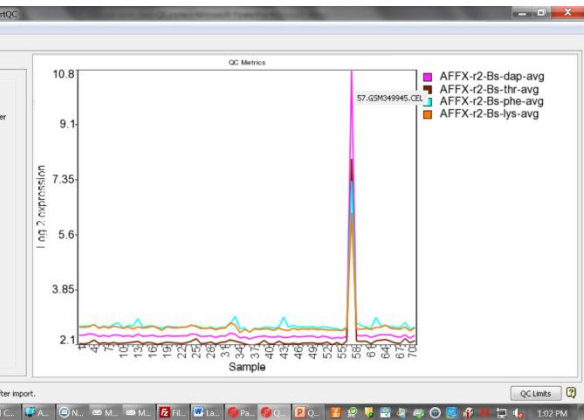
STEP 3: Run SNP x smoking interaction GWAS on filtered SNPs

STEP 4: Replicate in independent dataset

Public Microarray Data Curation: Quality Control

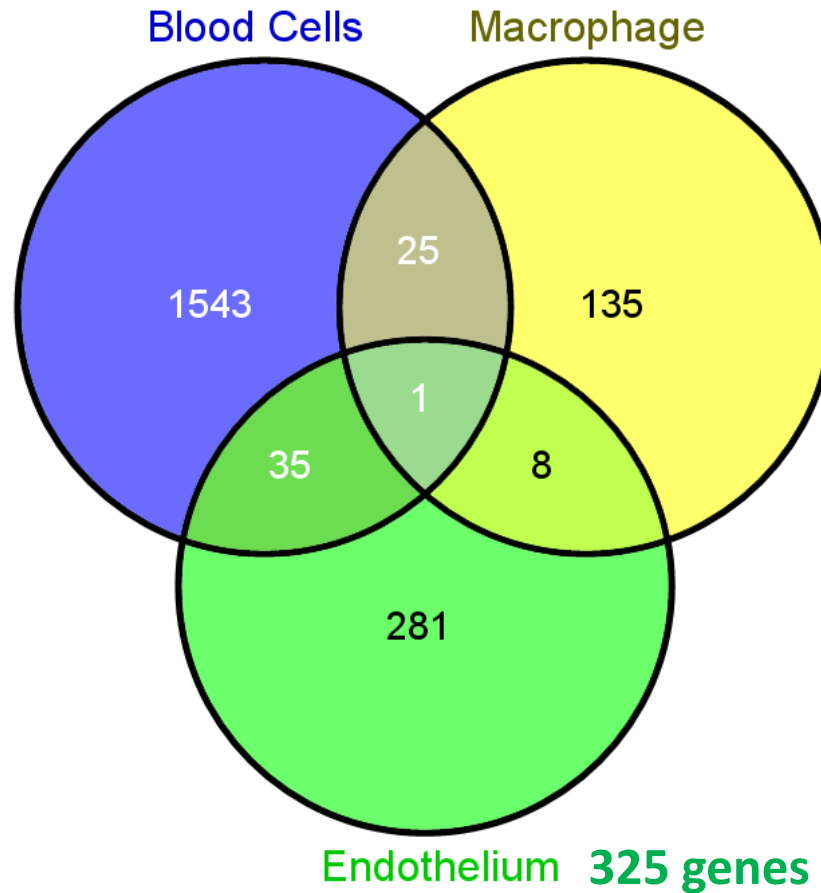


Public Microarray Data Curation: Quality Control



STEP 1: Define genes whose expression altered by smoking

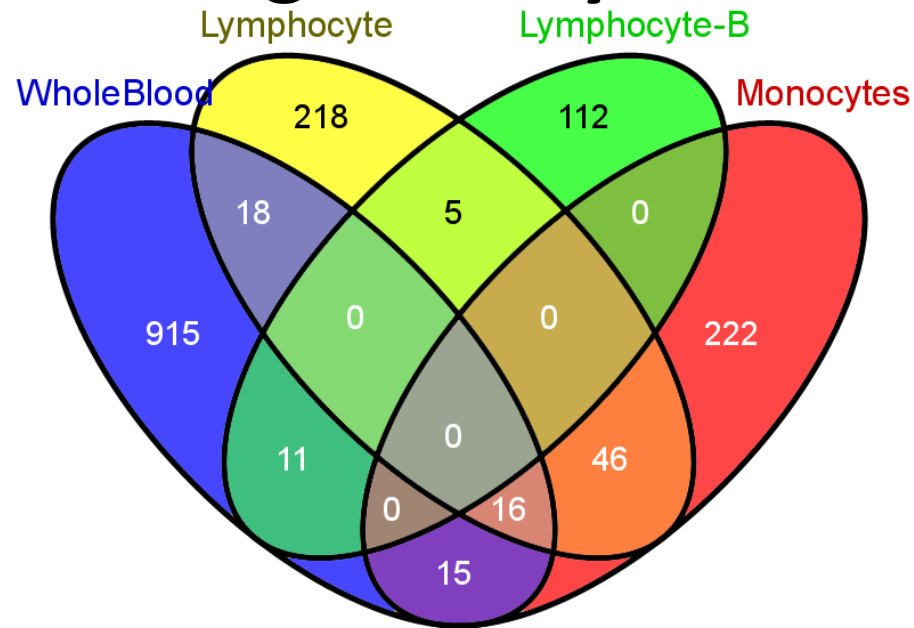
1604 genes
6 studies:
Smokers vs
Non-smokers



169 genes
3 studies:
Smokers vs
Non-smokers

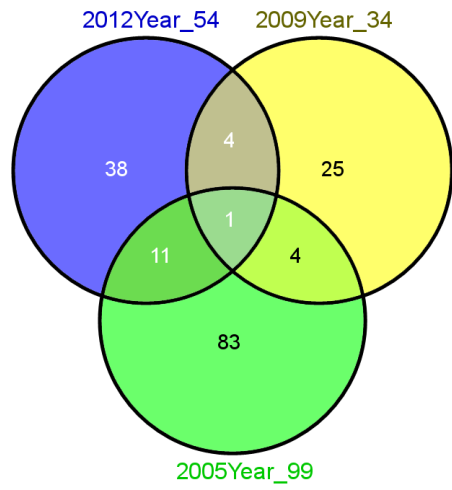
325 genes
4 studies: In Vitro Culture with
Smoking Extract or Nicotine for 1hr-24hr

Smoking-regulated genes: by individual study

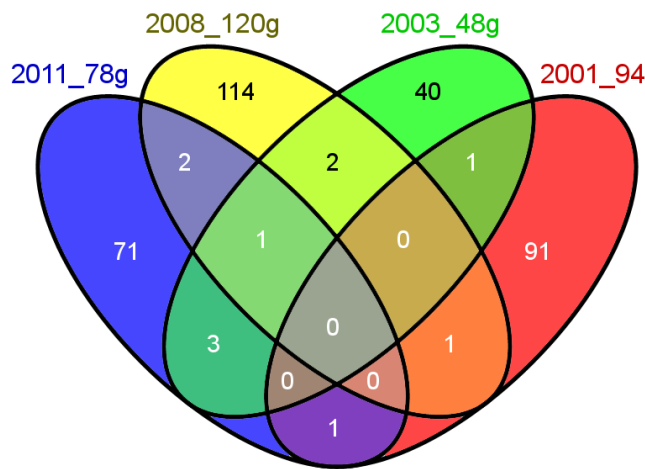


Blood Cells

(4 out of 6 studies only)



Aveolar Macrophage



Endothelial Cells

Method overview

STEP 1: Define genes whose expression altered by smoking

- tissue types: blood, alveolar macrophages, endothelial cells
- literature and publicly available microarray datasets

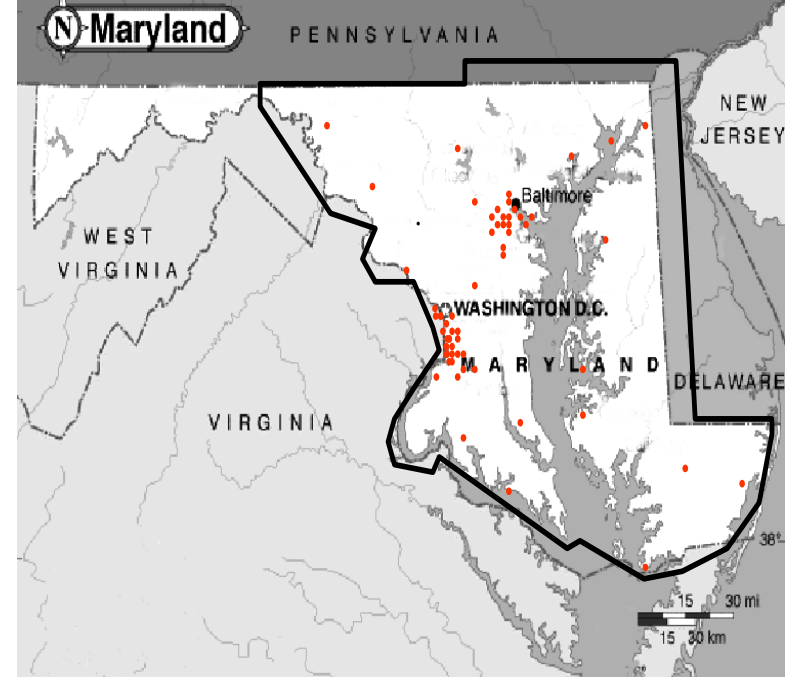
STEP 2: Define SNPs relevant to target genes

- within (+/- 10 kb) target genes on Illumina Omni1_Quad_v1-0_B
52,385 SNPs for 1668 genes (Blood+Macrophage)
8,787 SNPs from 317 genes (Endothelial Cells)
- eQTLs of target genes: (in progress) GTEx , SCANDB databases

STEP 3: Run SNP x smoking interaction GWAS on filtered SNPs

STEP 4: Replicate in independent dataset

Genetics of Early Onset Stroke



Goal: Identify genes associated with early onset stroke (stroke event 16 - 50 yrs)

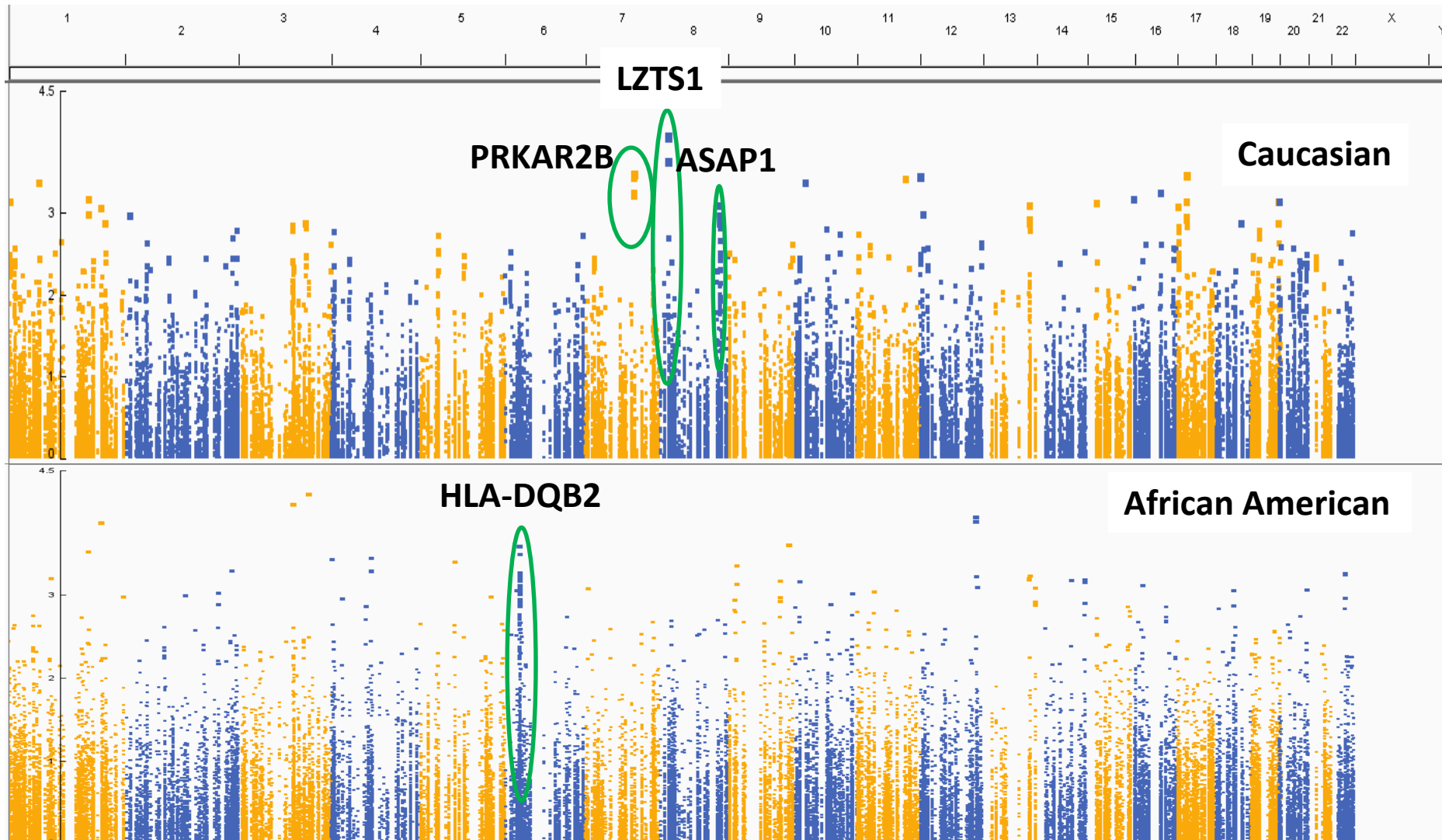
889 ischemic stroke cases from discharge surveillance from Baltimore-Washington area (1992 – 2007)

927 Controls by random-digit dialing

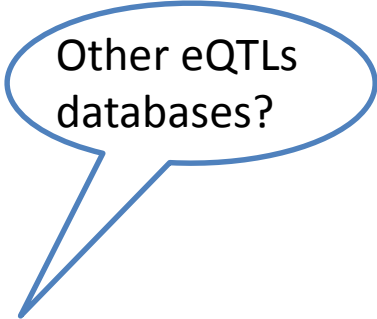
Preliminary Results: GEOS study

52,385 SNPs from 1668 genes (Blood+Macrophage)

Stroke = Age + PC1+ smoking + SNP + SNP × smoking



Future Directions



Other eQTLs
databases?

1. Expand SNPs list by including eQTLs SNPs of smoking-regulated genes: GTEx database, SCANDB database
2. Run SNP x smoking interaction statistical analysis on defined SNPs in GEOS studies
3. Replicate in independent dataset:
 - Geneva Germany Dataset
 - Pakistan Dataset



Other datasets?

Summary and Reflections

1. Candidate gene approach using functional genomics annotation derived from transcriptome microarray studies can be effectively utilized to study SNP*smoking interaction in stroke;
2. Transparency, quality, and accessibility of omics data are critical to ensure the success of such approach;
3. More tissue-specific & cell-specific studies are needed to enrich the functional annotation of genes associated with various “environment” factors including smoking.

ACKNOWLEDGEMENTS

GEOS Study:

Braxton D. Mitchell, PhD

Steven J. Kittner, MD

John Cole, MD

Yu-Ching Cheng, PhD

Patrick F. McArdle, PhD

O. Colin Stine, PhD

Jeffrey R. O'Connell, PhD

Nicole Dueker, BS

Mary J. Sparks, RN,BSN

Kathy Ryan, MS

Program in Personalized and Genomic Medicine

Alan R. Shuldiner, MD

NIH/NIA T32 Training Grant 5T32AG000219 (PI: Andrew P Goldberg,
MD)

The GENEVA Consortium, especially:

At Univ Washington:

Cathie C. Laurie, PhD

Jess Shen

At CIDR:

Kim Doheney, PhD

Elizabeth Pugh, PhD

Hua Ling, PhD

At NHGR:

Teri A. Manolio, MD, PhD