



WHI GWAS Data



Over time, WHI has contributed samples* to numerous GWAS, including many consortium studies

*requires supplemental consent to participate in genetic studies

WHI Core GWAS

- SHARe (M5) – Cohort, African American and Hispanic
- GARNET (M13) – HT case/control: MI/Stroke/VT/T2D, majority EA
- WHIMS+ (W63) – Cohort/Subsample of HT, EA
- LLS-GWAS (W66) – LLS w/o previous GWAS

Other Key GWAS

- Hip Fracture (BAA03) – Case/control: hip fracture, majority EA
- GECCO (AS224) – Case/control: colorectal cancer, EA
- MOPMAP (AS264) – Case/control: ventricular ectopy, EA

2013 Imputation Project

- Used genotype data from **6** studies: SHARe, GARNET, WHIMS+, GECCO, HipFx and MOPMAP → ~30,000 samples
- Steps:
 - alignment ("flipping") to the same reference panel
 - imputation to the 1000 genomes
 - identification of genetically related individuals
 - computations of principal components and comparison with self-reported ethnicity

2013 Imputation Project

- Resulted in ~ 34 million SNPs on 31,807 WHI participants
- For these studies, the imputed and some genotype data have all been uploaded to dbGaP



Access to Imputed GWAS Data



From dbGaP

Apply for access through dbGaP

```
graph TD; A[Apply for access through dbGaP] --> B[Phenotype (text format)]; A --> C[Genotype for SHARE, GARNET, and WHIMS+ only (PLINK format)]; A --> D[Imputed Genotype (MACH format)];
```

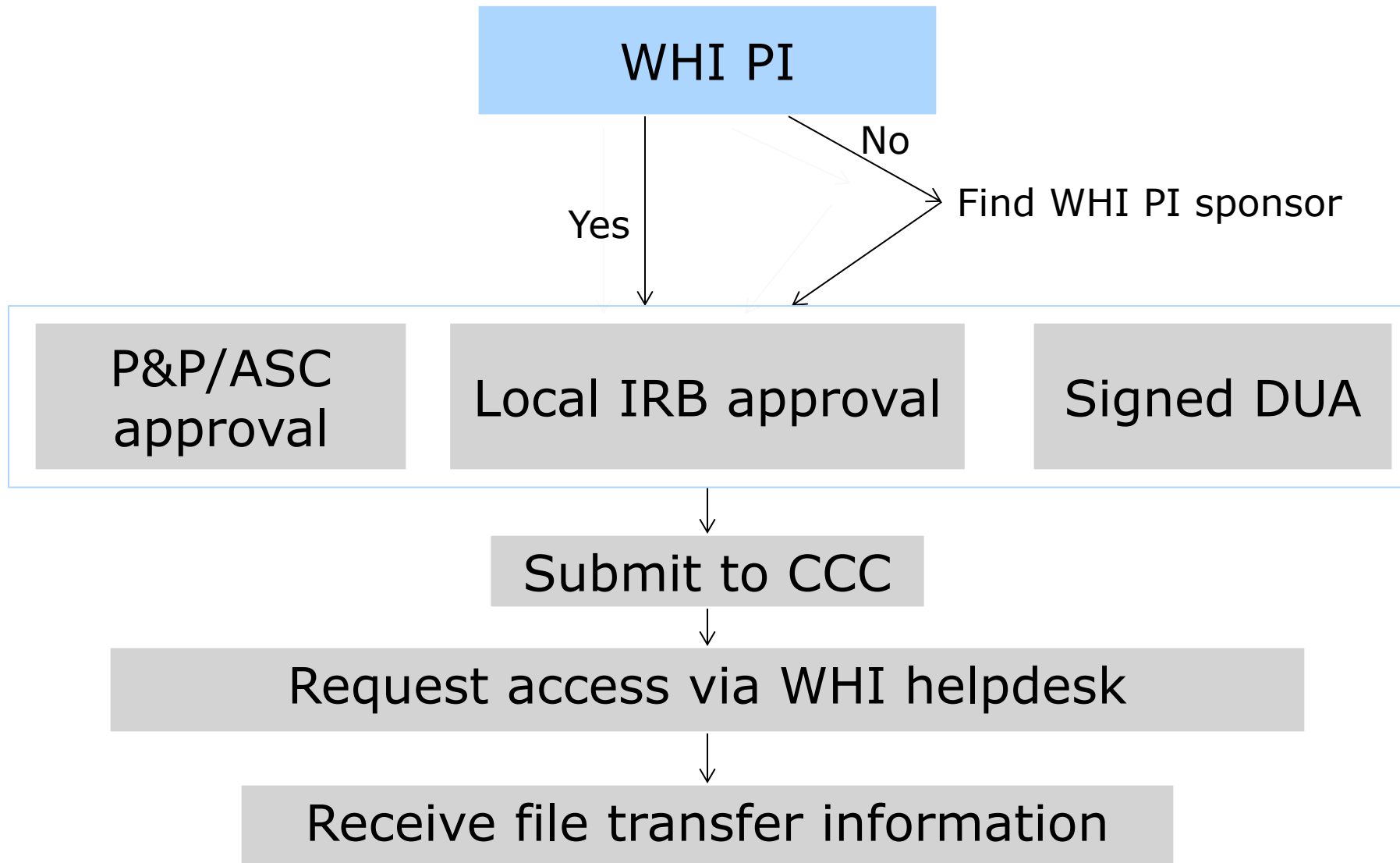
Phenotype
(text format)

Genotype for
SHARE, GARNET,
and WHIMS+ only
(PLINK format)

Imputed
Genotype
(MACH format)

* dbGaP uses it's own participant ID; not linkable with WHI ID

From WHI CCC



Other GWAS

- PAGE II (AS349): minority cohort, $N \approx 12,000$. Genotyped ~ 1.5 million SNPs, will eventually have imputed data. Currently available from PAGE investigators.
- Breast Cancer Post-GWAS (M18): EA breast cancer cases/controls ($N \approx 10,000$). 600,000 genotyped cancer related SNPs using the oncochip; imputation planned.

Other GWAS (continued)

- NCI consortium studies: pancreatic, bladder, kidney, gastric and esophageal cancer, non Hodgkin's lymphoma; used case/control design
- Stroke, 844 case only

For the most part, data from these studies has not been sent to the CCC, or uploaded on dbGaP

Other Genetic Studies

- PAGE I (M6), metabochip results, 200K SNPs
- GO-ESP WHISP (M24), SNP genotypes and whole exome sequencing data

Results have been uploaded to dbGaP

“dbGaP_availability” data file on WHI website

- Helps identify which WHI Participants have data on dbGaP from:
 - GWAS Imputation project
 - GO-ESP
 - PAGE I
- One row per participant; indicator variables identify study membership, chip type, case/control status

Online Resources

■ WHI website

- GWAS studies page:
<https://www.whi.org/researchers/data/SitePages/GWAS%20Data.aspx>
- Individual study pages (enter symbol, ex: M5, W63):
<https://www.whi.org/researchers/data/WHIStudies/StudySites/Pages/home.aspx>

■ dbGaP:

http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000200.v10.p3

■ Email: WHI helpdesk (helpdesk@whi.org)

Acronyms

- GARNET: Genome-wide Association Studies of Treatment Response in Randomized Clinical Trials
- GECCO: GWAS in Colon Cancer
- GO-ESP: "Grand Opportunity" Exome Sequencing Project
- LLS: Long Life Study
- MOPMAP: Modification of PM-mediated Arrhythmogenesis in Populations
- PAGE: Population Architecture using Genomics and Epidemiology
- SHARe: SNP Health Association Resource
- WHIMS: WHI Memory Study