Session 1

Replication

Replication

Replication

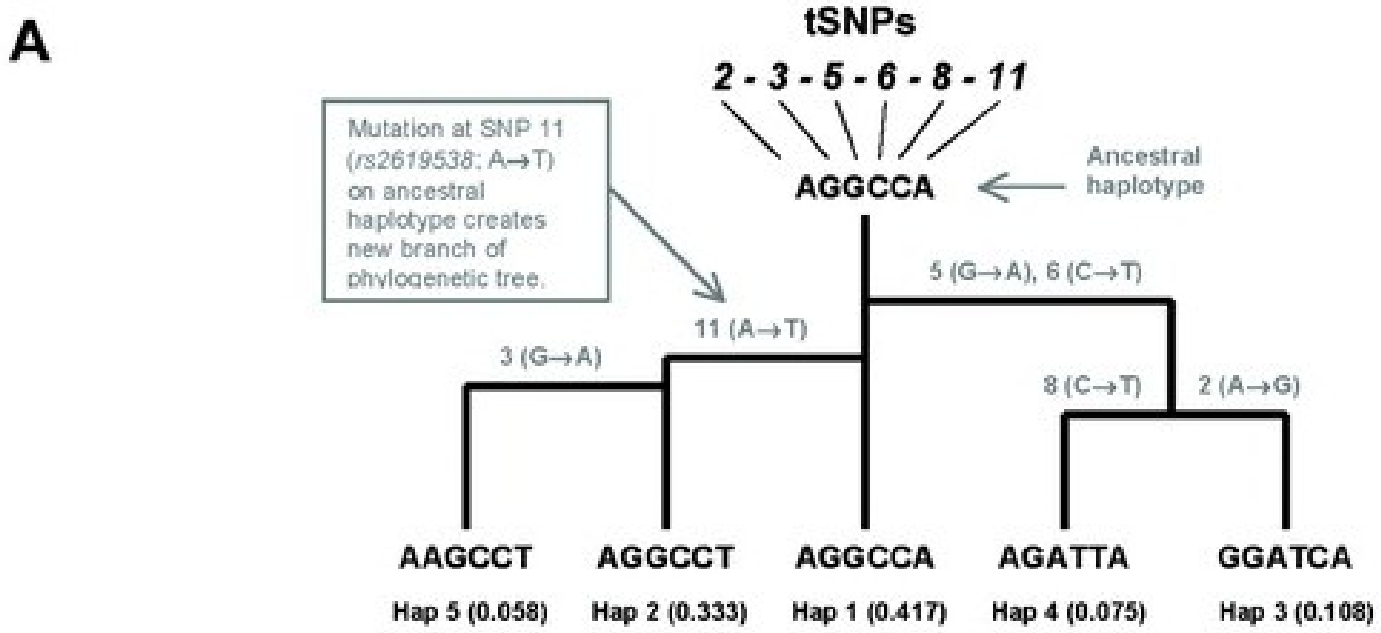
Session 1

Stephen Chanock, NCI
Cashell Jaquish, NHLBI
Teri Manolio, NHGRI
Winnifred Rossi, NIA

Need for Consensus on What Constitutes Replication

- **Avalanche of GWAS and candidate gene studies**
- **Replication: *sine qua non***
- **Likelihood of single study establishing an association is low *until* studies are sufficiently powered.**
  - **Sample sizes increase sufficiently**
  - **Analytical methods improve substantially**
- **Common problem of how to interpret confusing and spurious findings**
  - **Multiple markers & study designs**

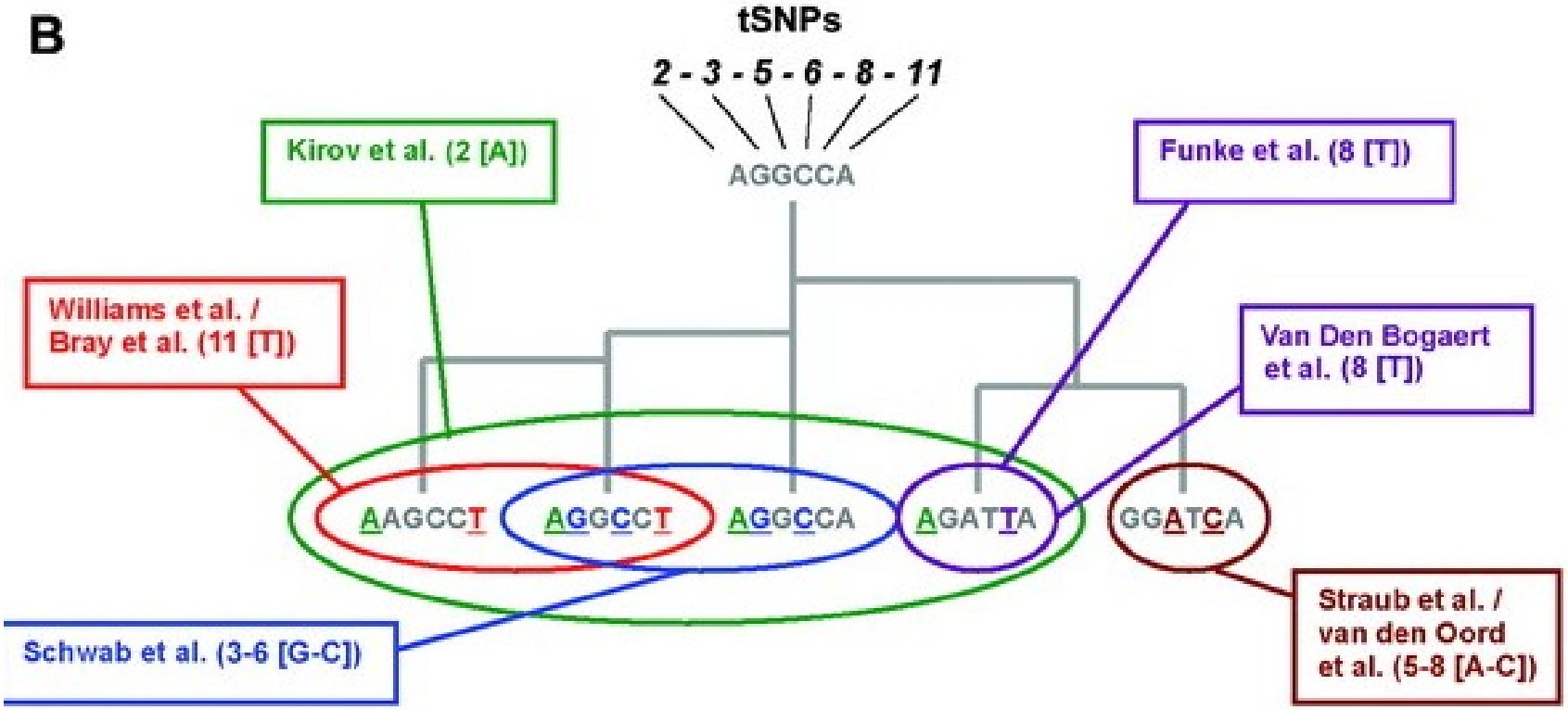
Phylogenetic Tree of Five Common Haplotypes of *DTNBP1*



Possible Association with Schizophrenia

Mutsuddi et al, *Am J Hum Genet* 2006; 79:903-909.

“Positively” Associated Haplotypes Differ in All Six Studies



Each common DTNBP1 haplotype was tagged but multiple association signals observed. Strong possibility that more than one common variant contributes to schizophrenia risk at DTNBP1 locus

Mutsuddi et al, *Am J Hum Genet* 2006; 79:903-909.

Challenges in Replication Studies: aka…how not to do a study

- **Use different phenotypes**
- **Use different markers**
- **Mix fine-mapping and replication**
- **Use different analytic methods (haplotype vs. single marker)**
- **Fail to adequately account for differences in populations**
  - **Distinct populations**
  - **High heterogeneity within a study**

Proposed Criteria for Positive Replication*

- **Sufficient sample size in soundly designed study**
- **Same gene**
- **Same SNP (or SNP in complete LD with prior SNP, $r^2 = 1$)**
- **Comparable genetic model**
- **Same direction as original finding**
- **Highly significant association reporting**
  - **Initial Report**
  - **Joint or Combined Analysis Preferred**
- **Same or very similar population**

Chanock et al *Nature* in press

NCI-NHGRI Working Group on Criteria for Genotype-Phenotype Association

How Do We Get to Replication?

SNP Health Association Resource (SHARe)

- Genotyping of all consenting subjects from 3 generations (N~10,000) of the Framingham Heart Study
- 550K SNPs (Affymetrix) in each subject
- Plan to genotype two more cohorts
- Include all available phenotypes, clinical outcomes and genotypes in dbGAP
- Individual data will be shared openly for replication and gene finding

Candidate Gene Association Resource (CARE)

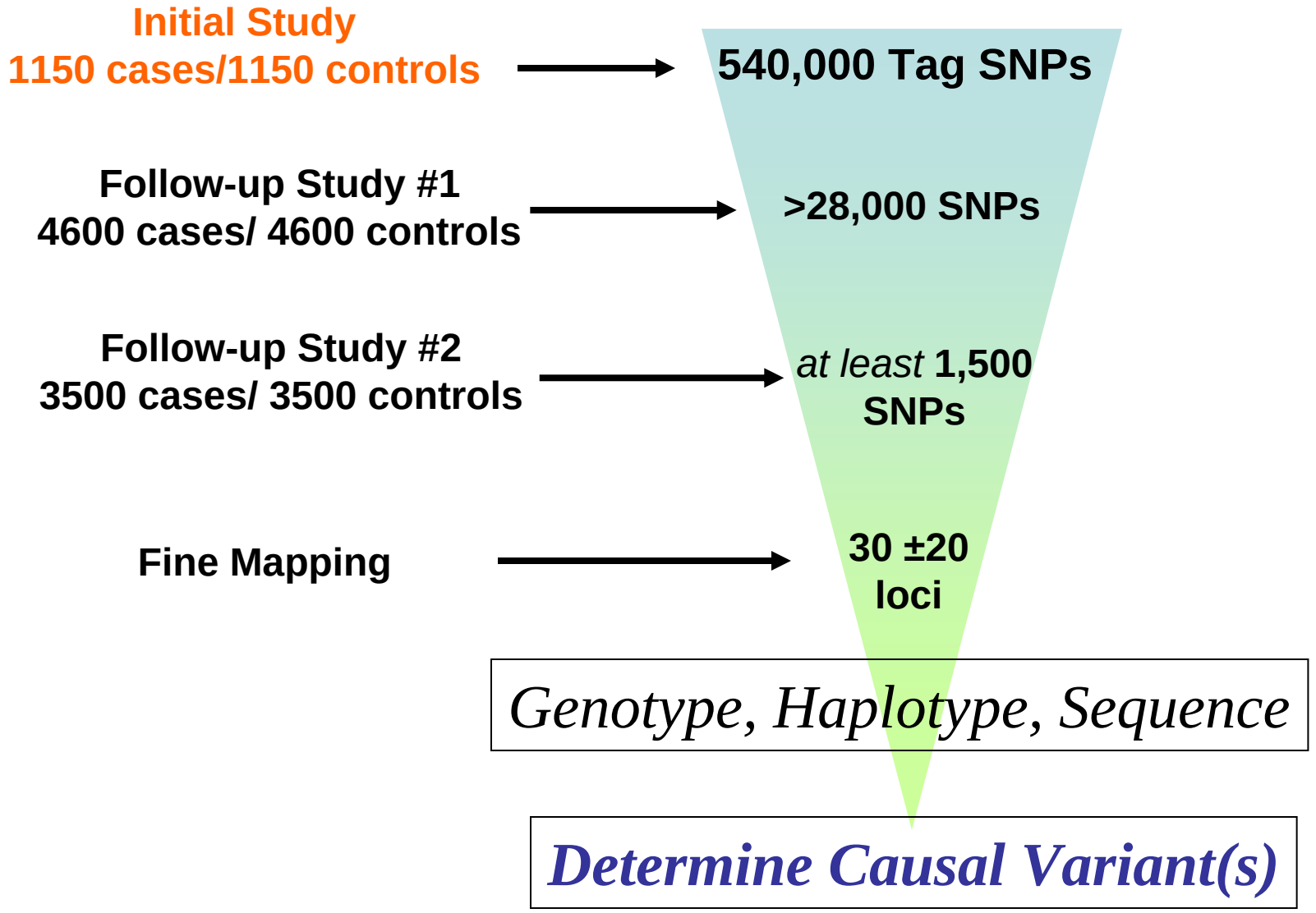
- Create 50,000 person cohort with genotyping of 1,700+ candidate genes
- Complete 500k SNP genotyping on a subset of the combined cohorts
- Merge harmonized phenotype data for multiple studies with genotype data
- Openly shared data source
- Utilizes a more standardized consortium effort for replication

SNP Typing for Association with Multiple Phenotypes from Existing Epidemiologic Data (STAMPEED)

- Supports genome wide SNP typing and analysis in existing clinical trial, cohort, case-control or family studies.
- 13 R01s investigating heart, lung and blood phenotypes
- Close to 30,000 people will be genotyped
- Studies were required to have a plan in place for replication and validation
- Investigators initiated working groups within STAMPEED to replicate findings
- Investigator initiated collaborative approach to replication
- Investigators expressed a need for “rapid response” resource for replication and validation

General Strategy for Prostate & Breast Cancer GWAS

CGEMS
Cancer Genetic Markers of Susceptibility



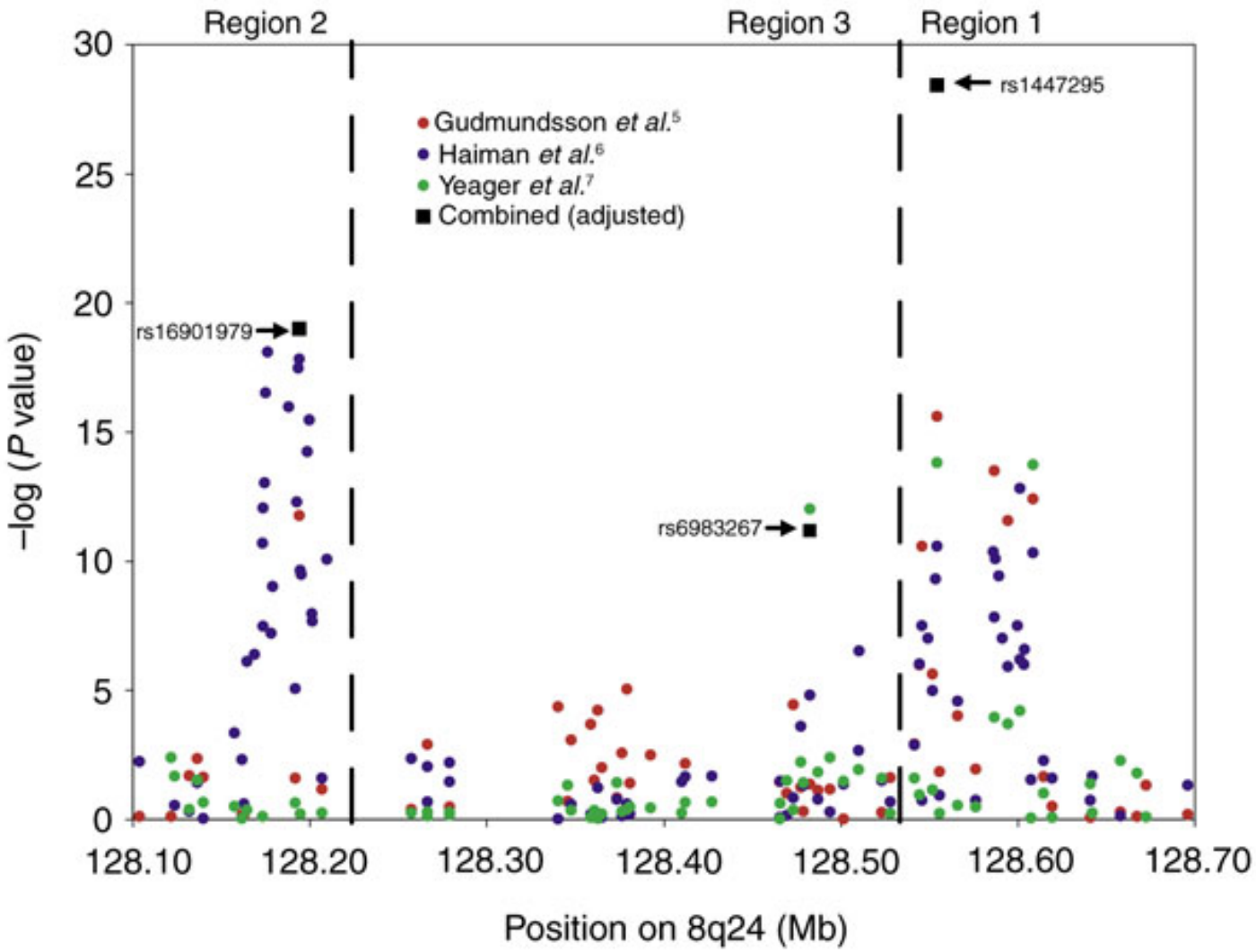
Replication Studies in CGEMS Prostate Cancer GWAS

	Subjects		rs6983267 Predisposing allele frequency		rs6983267 P-value	rs1447295 Predisposing allele frequency		rs1447295 P-value
			Cases	Cont.		Cases	Cont.	
PLCO	1157	1172	0.55	0.49	2.4×10^{-05}	0.14	0.10	9.8×10^{-05}
ACS	1151	1150	0.55	0.50	3.2×10^{-03}	0.12	0.08	2.7×10^{-05}
ATBC	896	894	0.57	0.51	1.9×10^{-03}	0.21	0.17	2.9×10^{-02}
FPCC	459	455	0.56	0.51	1.2×10^{-01}	0.12	0.07	4.4×10^{-03}
HPFS	636	625	0.57	0.51	1.0×10^{-02}	0.13	0.09	2.7×10^{-03}
ALL	4299	4296	0.56	0.50	9.4×10^{-13}	0.15	0.11	1.5×10^{-14}

Estimated Odds Ratios Overall

Heterozygotes	1.26	1.43
Homozygotes	1.58	2.23

Meta-Analysis of 8q24 papers in *Nature Genetics*: J Witte



J Witte Nature Genetics 2007

Issues in Replication

- Built-in Replication vs “1000 Flowers Blooming”
- Public Health or Clinical Significance
- Population Genetics
  - When is it Not Replication?
  - Special Populations
  - Barriers to Timely Conduct Follow-up
- Data Availability/Access
- Reporting of Negative Results
- Sufficient Detail for Inspection of Data

Issues in Developing Consortia

- Why do we need consortia?
- When are they needed?
- What aspects of GWA studies should they address?
- How should they be developed?
- Who should participate?
- What are the incentives for collaboration?
  - Publication?
  - Data access?