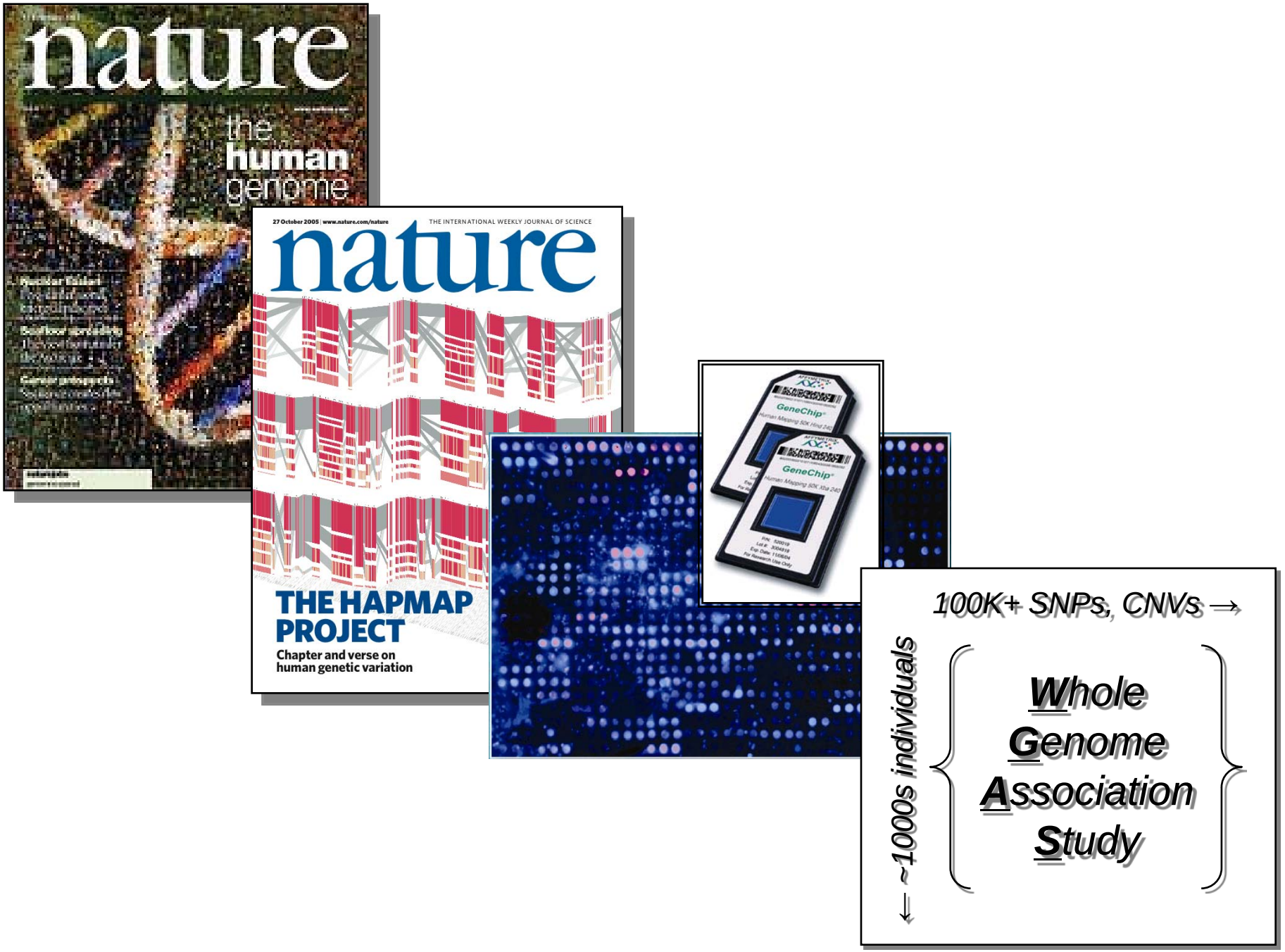
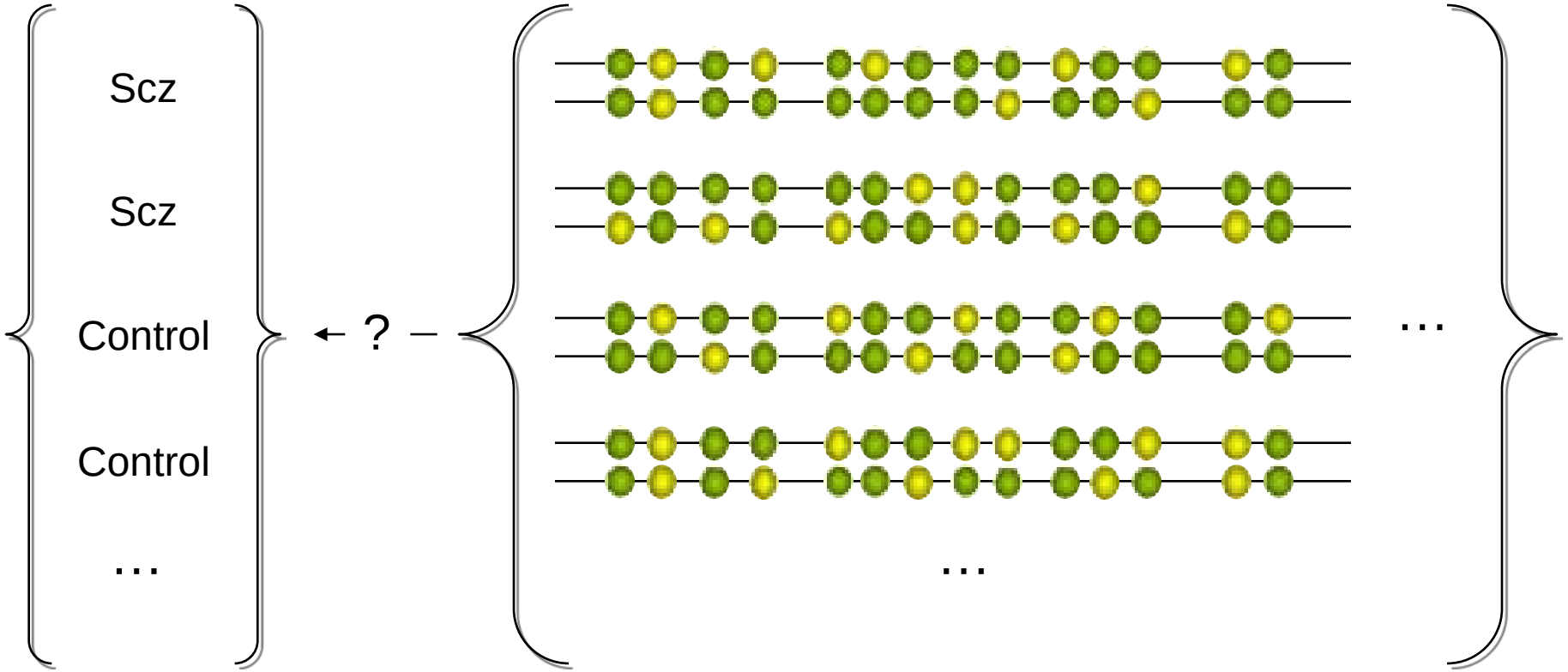
Whole genome association studies

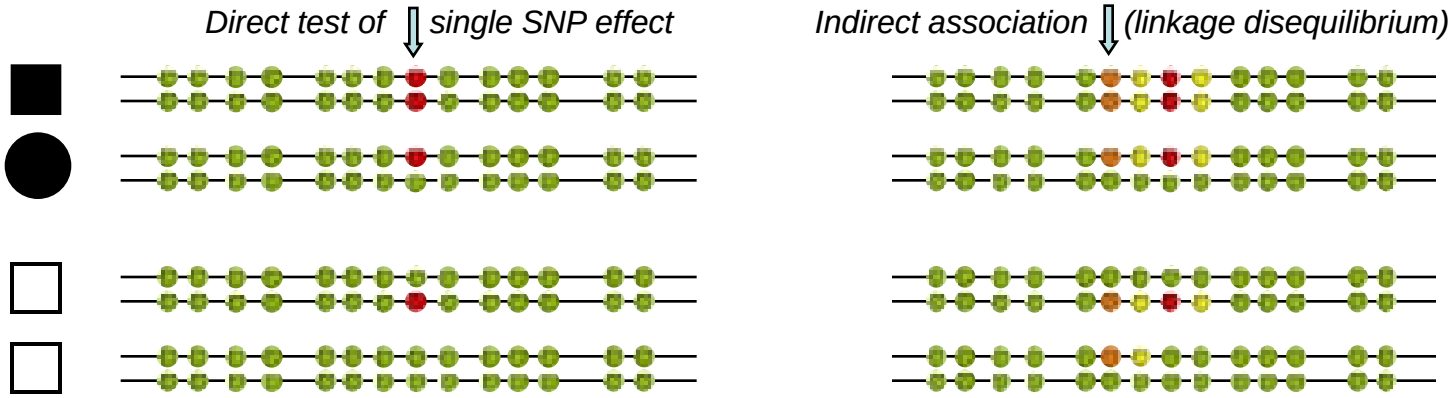
Introduction and practical

Boulder, March 2009

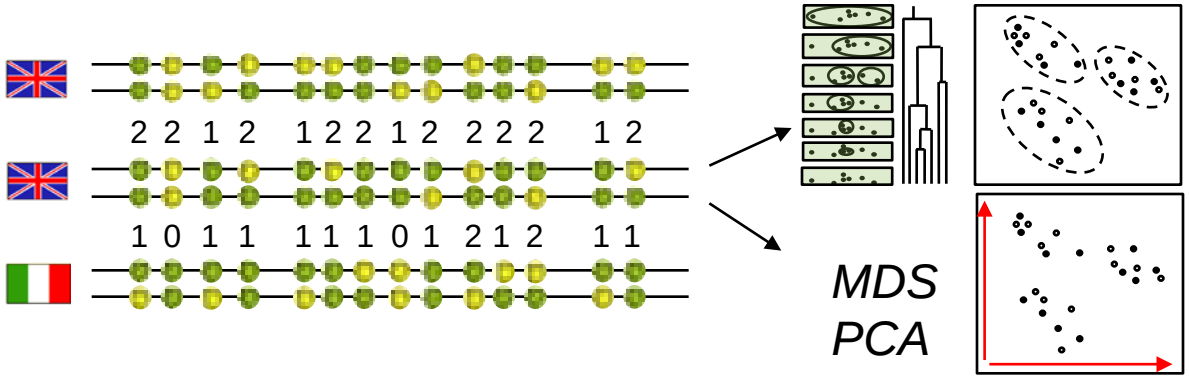

nature
the human genome
nature
27 October 2005 | www.nature.com/nature
THE INTERNATIONAL WEEKLY JOURNAL OF SCIENCE
THE HAPMAP PROJECT
Chapter and verse on human genetic variation
GeneChip
GeneChip
100K+ SNPs, CNVs →
← ~1000s individuals
Whole Genome Association Study

Associating phenotypic and genotypic variation





Analytic tools to perform, validate and enhance basic single SNP WGAS, e.g.:

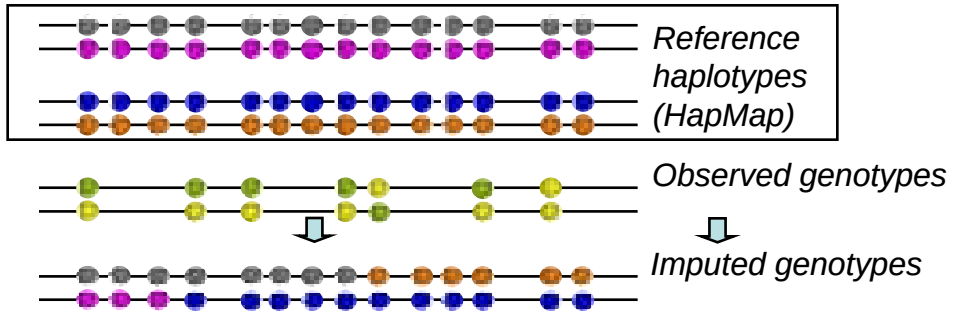


Empirical assessment of ancestry

1) Detect outliers, substructure
2) Clusters or continuous indices
3) Batch effects, relatedness, sample swaps, contamination, etc

Imputation of ungenotyped SNPs

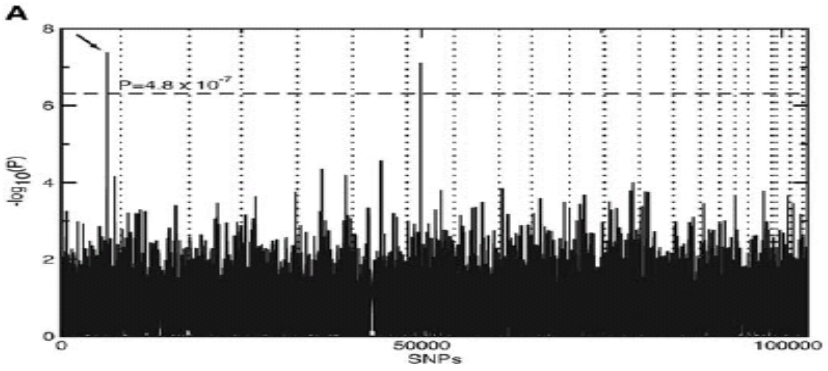
1) Increase coverage
2) Facilitate meta-analysis across platforms
3) Quality control (drop SNP/re-impute)



Age-related macular degeneration

Complement Factor H Polymorphism in Age-Related Macular Degeneration

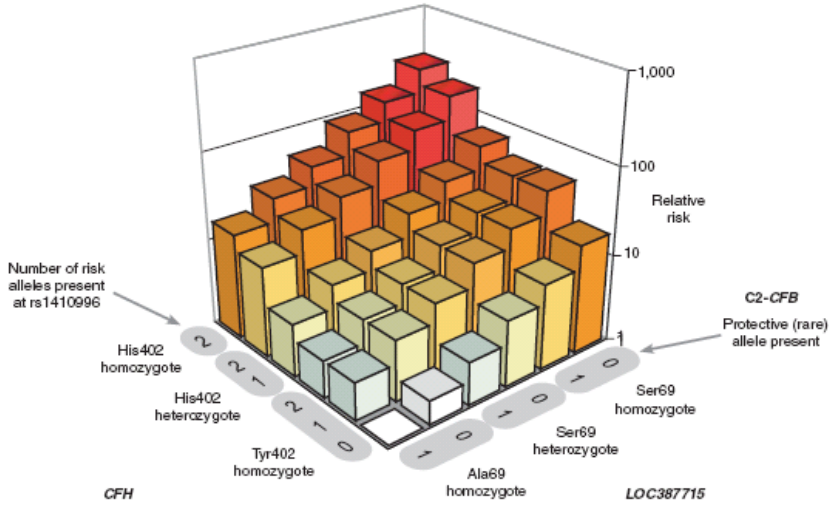
Robert J. Klein,[1] Caroline Zeiss,[2*] Emily Y. Chew,[3*]
Jen-Yue Tsai,[4*] Richard S. Sackler,[1] Chad Haynes,[1]
Alice K. Henning,[5] John Paul SanGiovanni,[3] Shrikant M. Mane,[6]
Susan T. Mayne,[7] Michael B. Bracken,[7] Frederick L. Ferris,[3]
Jurg Ott,[1] Colin Barnstable,[2] Josephine Hoh[7†]



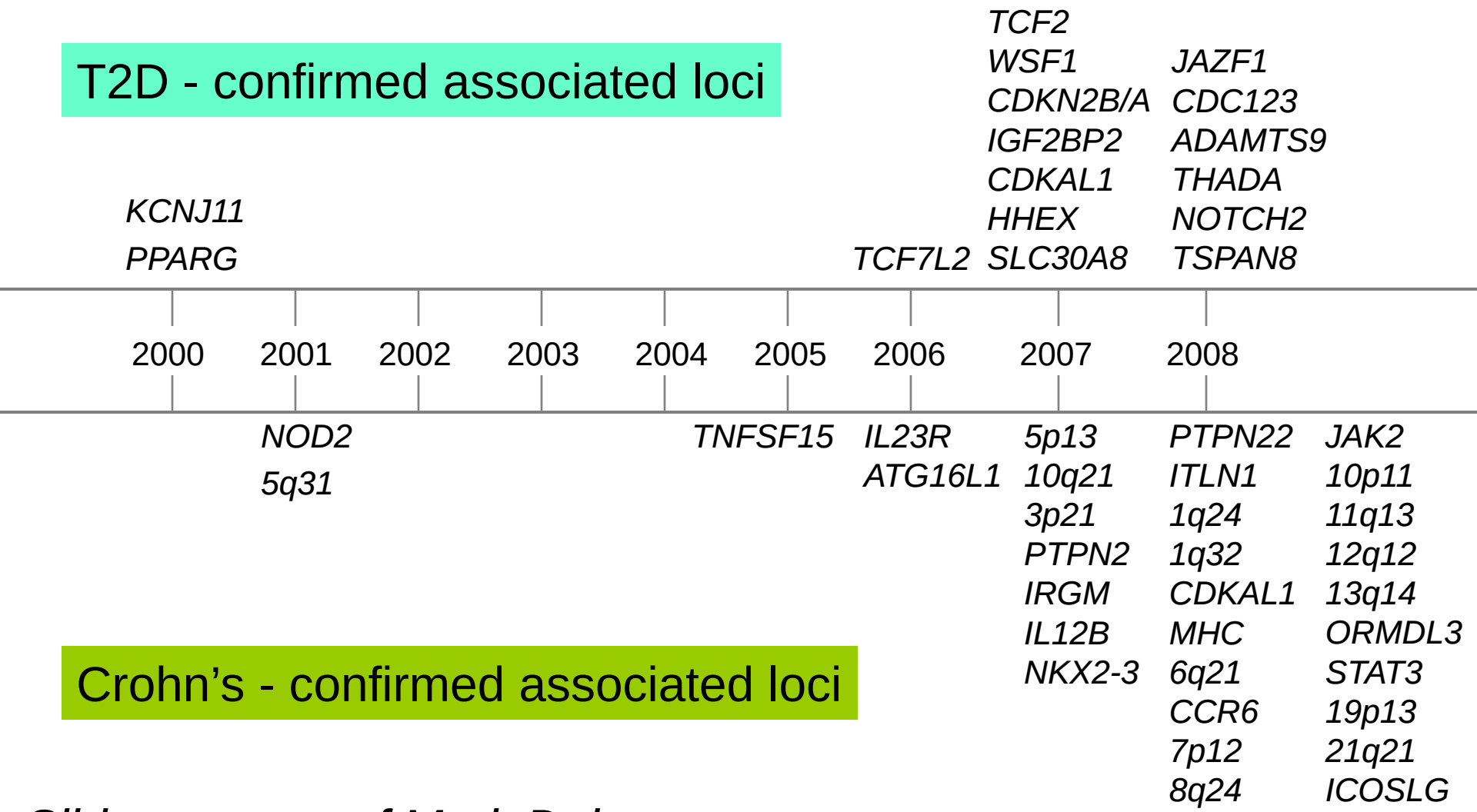
nature genetics

Common variation in three genes, including a noncoding variant in *CFH*, strongly influences risk of age-related macular degeneration

Julian Maller[1,3], Sarah George[2], Shaun Purcell[1,3], Jes Fagerness[1,3], David Altshuler[1,3,4], Mark J Daly[1,3,4] & Johanna M Seddon[2,4]



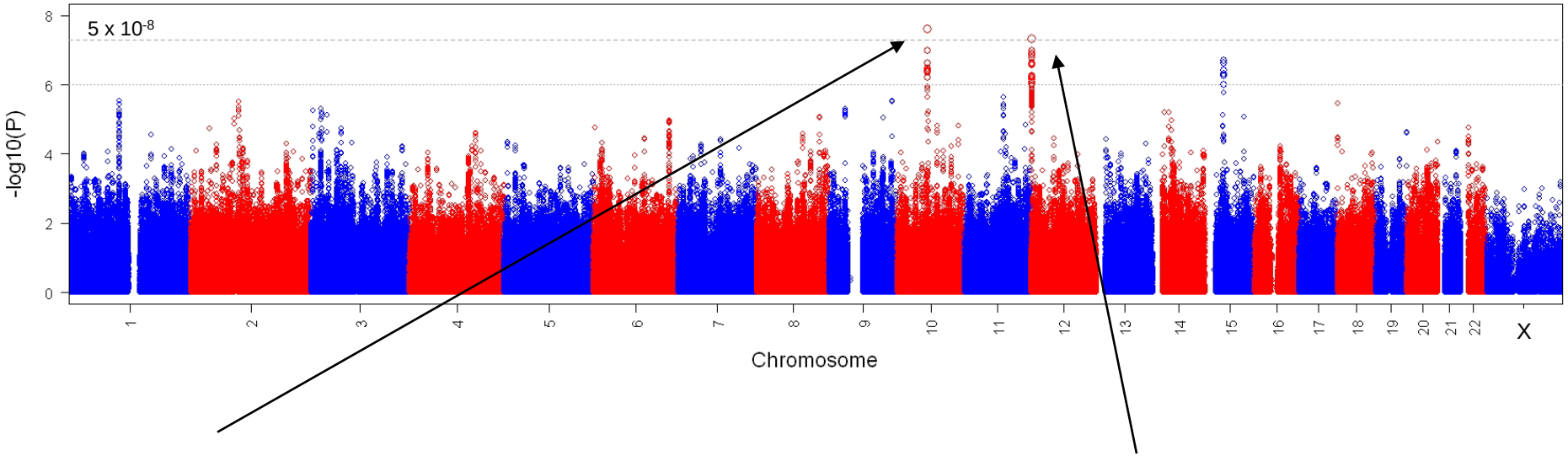
Progress in type 2 diabetes and Crohn's disease



Slide courtesy of Mark Daly

Bipolar WGAS of 10,648 samples

>1.7 million genotyped and (high confidence) imputed SNPs



Ankryin-G (ANK3)

CACNA1C

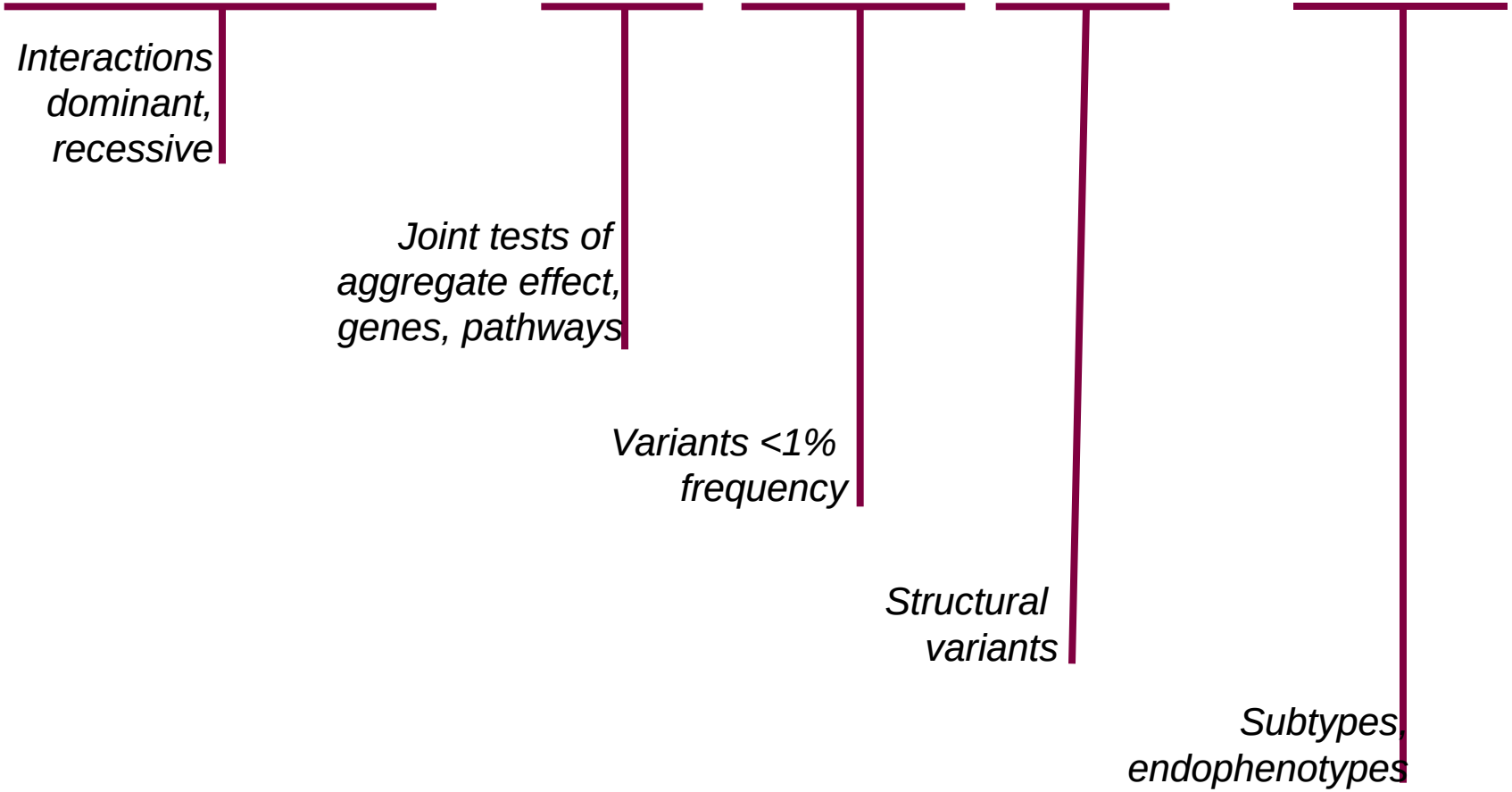
Sample	Cases	Controls	*P*-value
STEP	7.4%	5.8%	0.0013
WTCCC	7.6%	5.9%	0.0008
EXT	7.3%	4.7%	0.0002
Total	**7.5%**	**5.6%**	**9.1×10^{-9}**

Sample	Case	Controls	*P*-value
STEP	35.7%	32.4%	0.0015
WTCCC	35.7%	31.5%	0.0003
EXT	35.3%	33.7%	0.0108
Total	**35.6%**	**32.4%**	**7×10^{-8}**

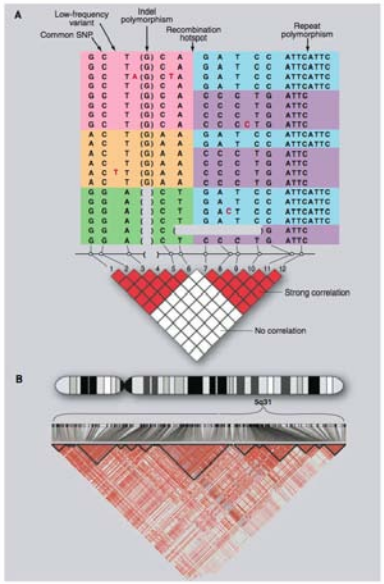
Ferreira *et al* (*Nature Genetics*, 2008)

Main focus of many association studies:

additive effects of single common SNPs on disease



Further reading on association mapping and interpretation of GWAS findings



Altshuler, Daly & Lander (2008) *Science*

Manolio, Brooks & Collins (2008) *JCI*

The case of the missing heritability

When scientists opened up the human genome, they expected to find the genetic components of common traits and diseases. But they were nowhere to be seen. **Brendan Maher** shines a light on six places where the missing loot could be stashed away.

Maher (2008) *Nature*

Practical session

- Data are in ~pshaun/prac2/
- Software required: PLINK and Haploview
- PDF with instructions is ~pshaun/instruct.pdf
- Work through until section “Empirical assessment of population stratification”
- Use PLINK website for help (http://pngu.mgh.harvard.edu/purcell/plink/)