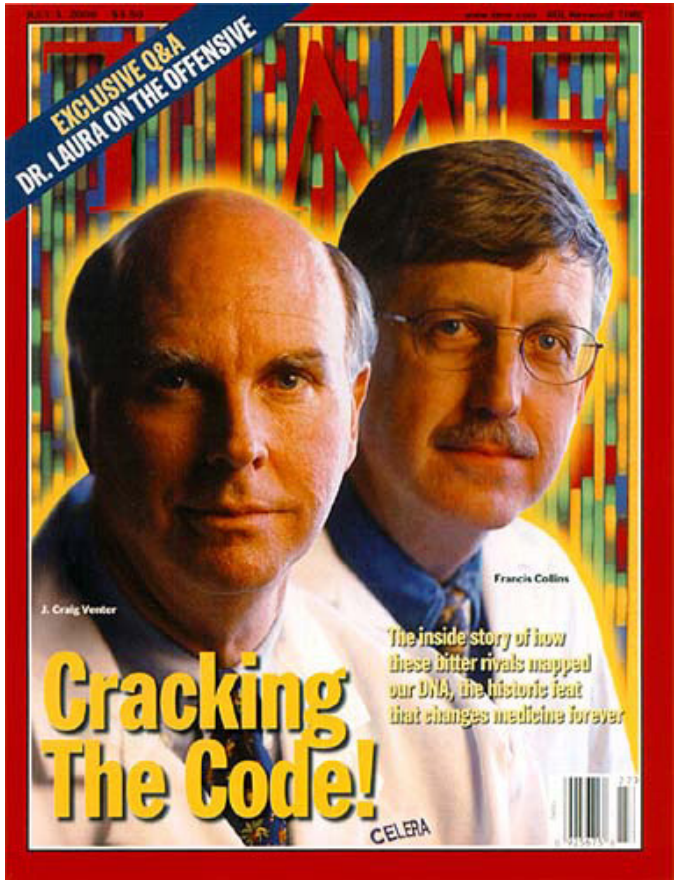
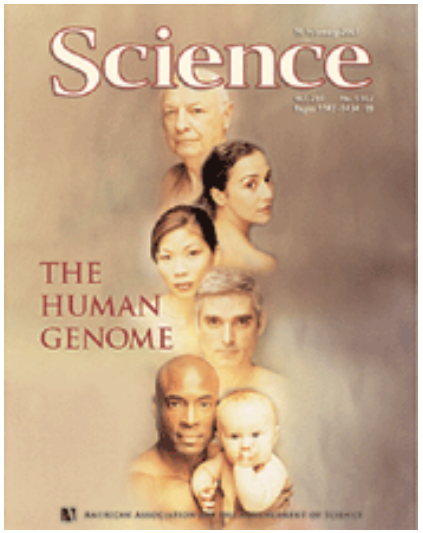
Rekenen aan het humane genoom

Paul de Bakker

Universitair Medisch Centrum
Utrecht

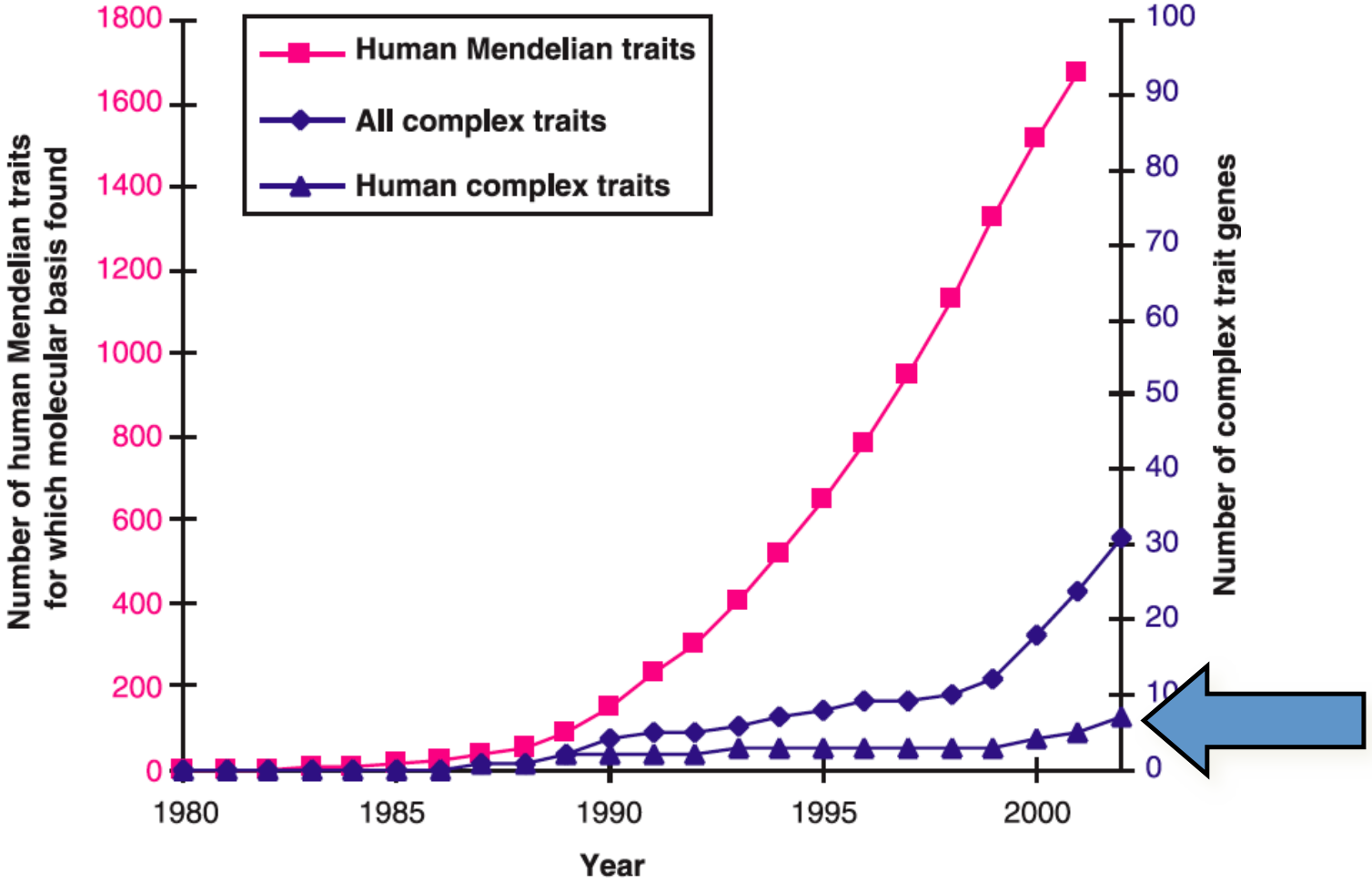
The Human Genome Project (2001)





Genotype ⇔ Phenotype

Early successes for monogenic diseases

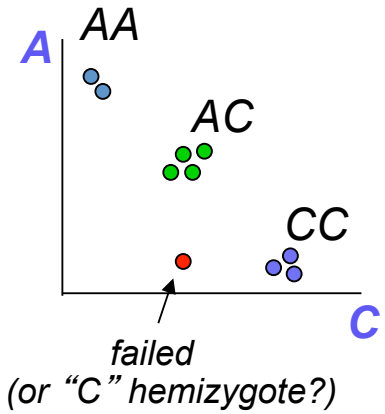


Glazier et al. Science 2002

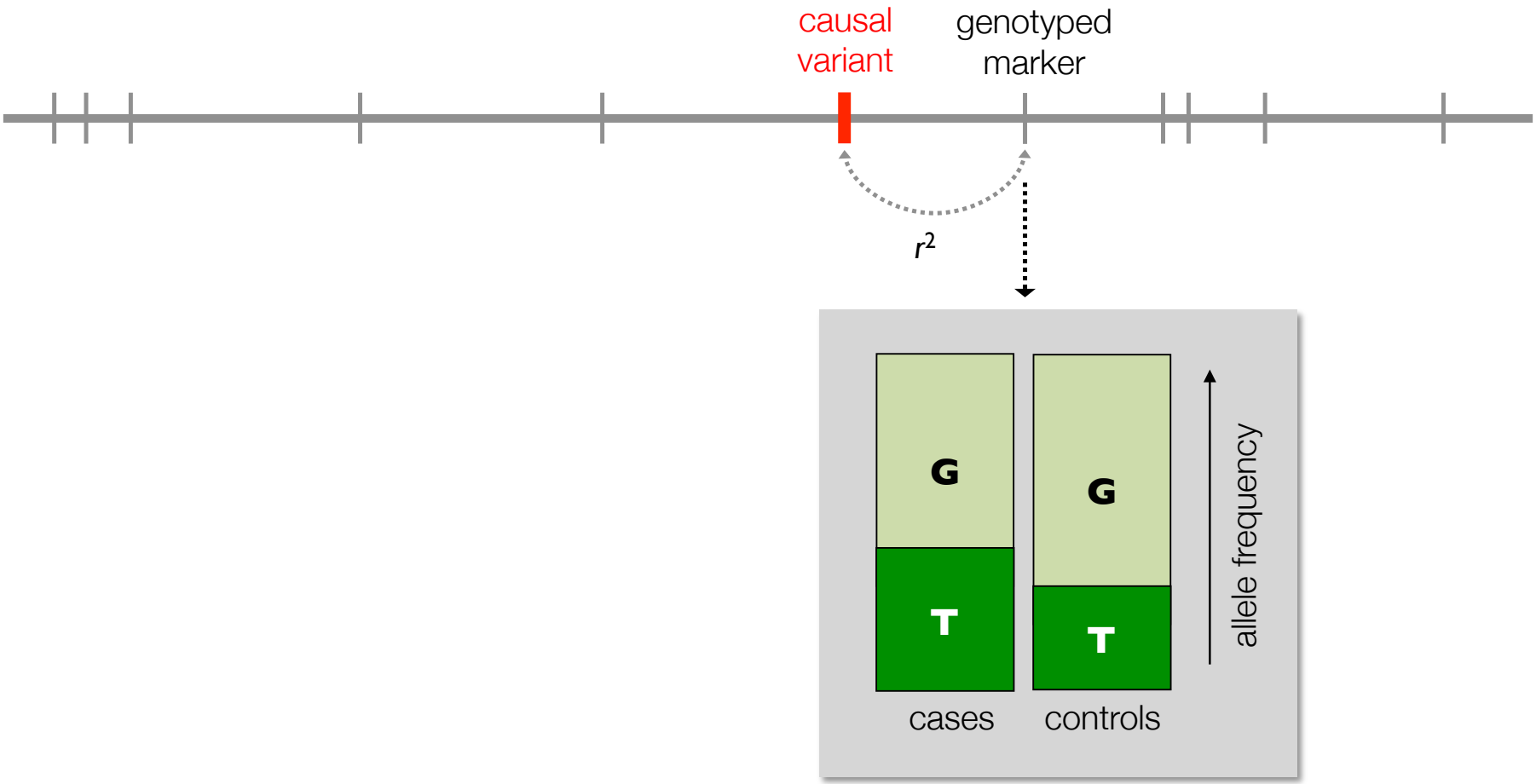
Catalogs of inherited DNA variation

SNP microarray platforms

- Allow genotyping of 100,000's of SNPs within a single experiment
- Variety of microarrays available
- Genome-wide association studies have revolutionized human genetics of complex traits



Genome-wide association



Statistical genetics

	A	a
Case		
Control		

Allelic (default) and trend test (Cochran-Armitage)

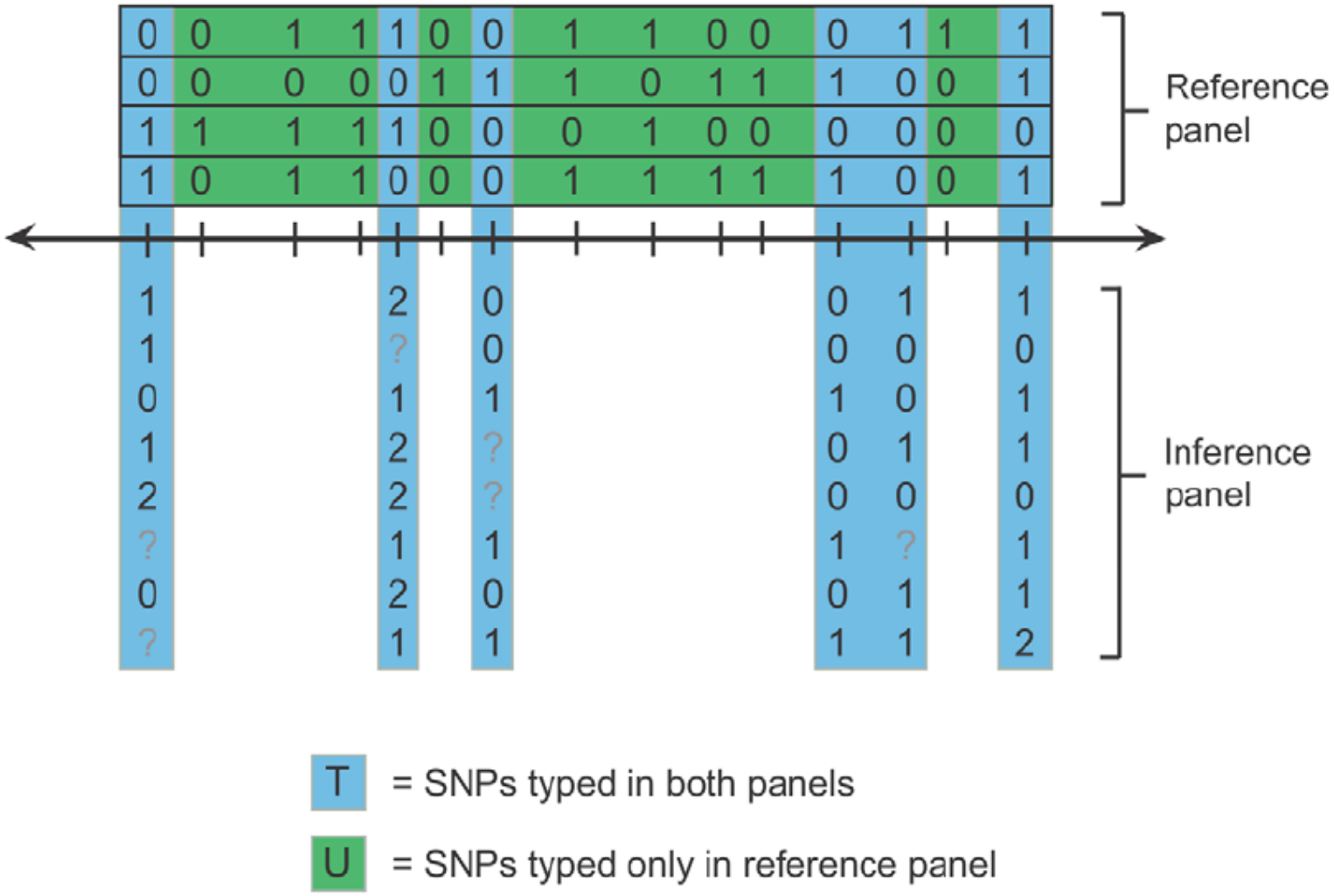
	AA	Aa	aa
Case			
Control			

Genotypic

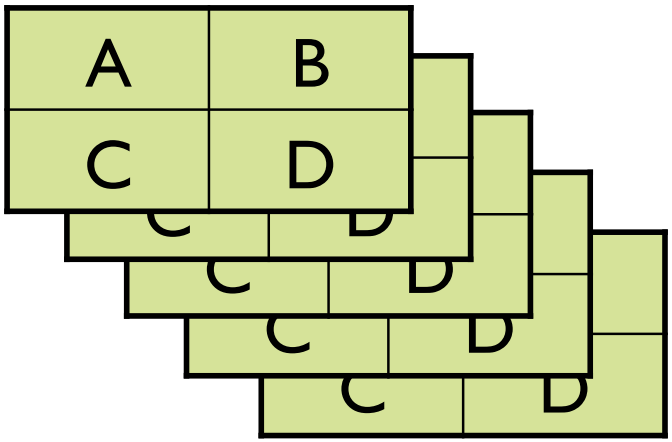
	AA/Aa	aa
Case		
Control		

Dominant
(recessive)

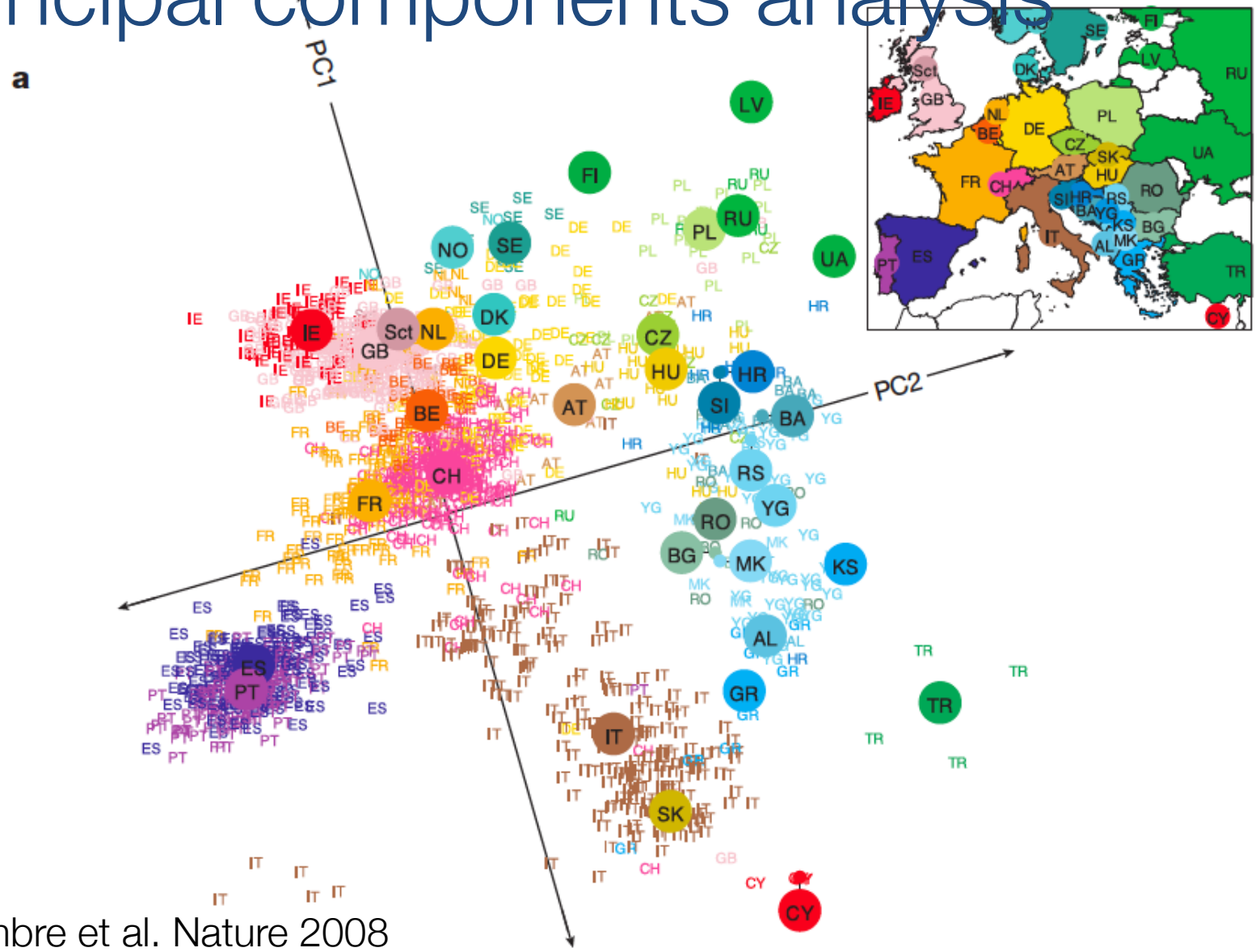
Imputation of non-genotyped SNPs



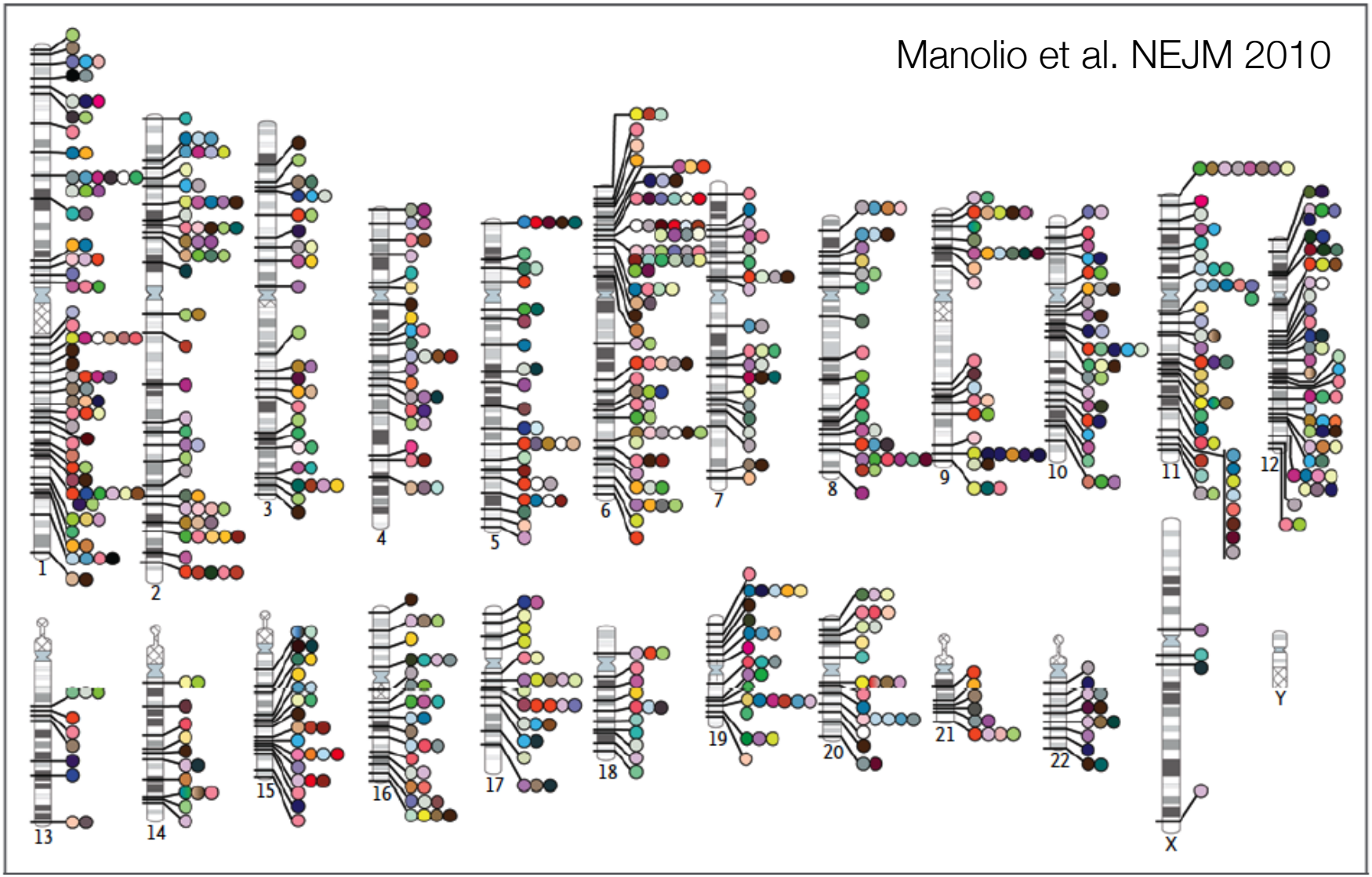
Meta-analysis of multiple data sets



Principal components analysis

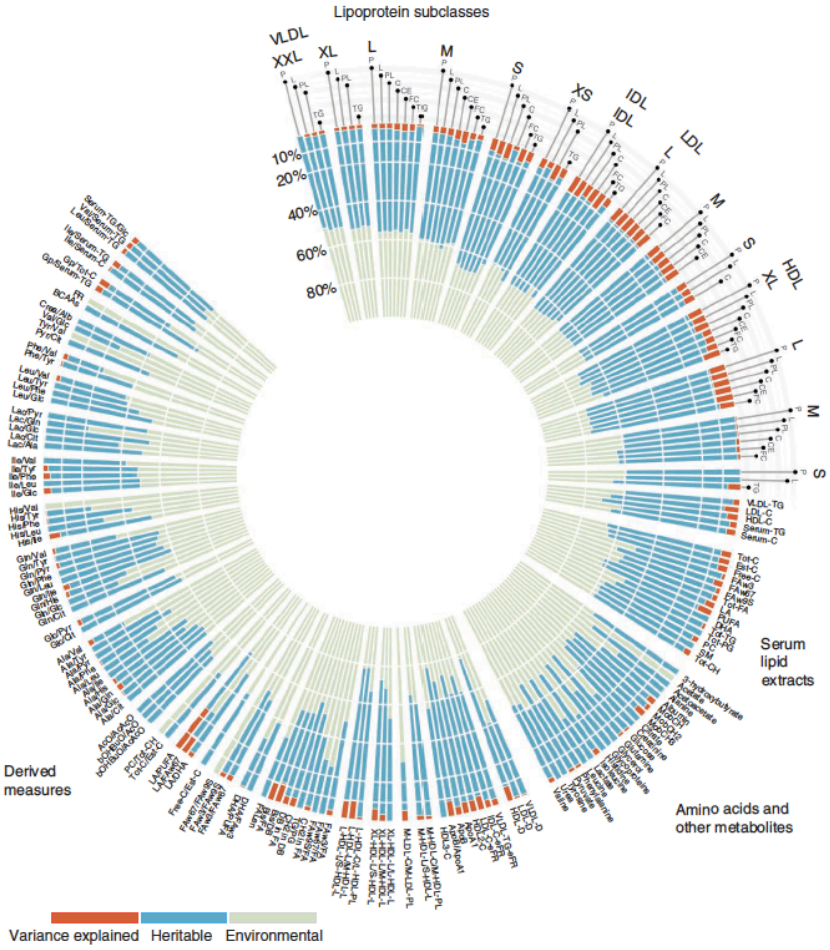


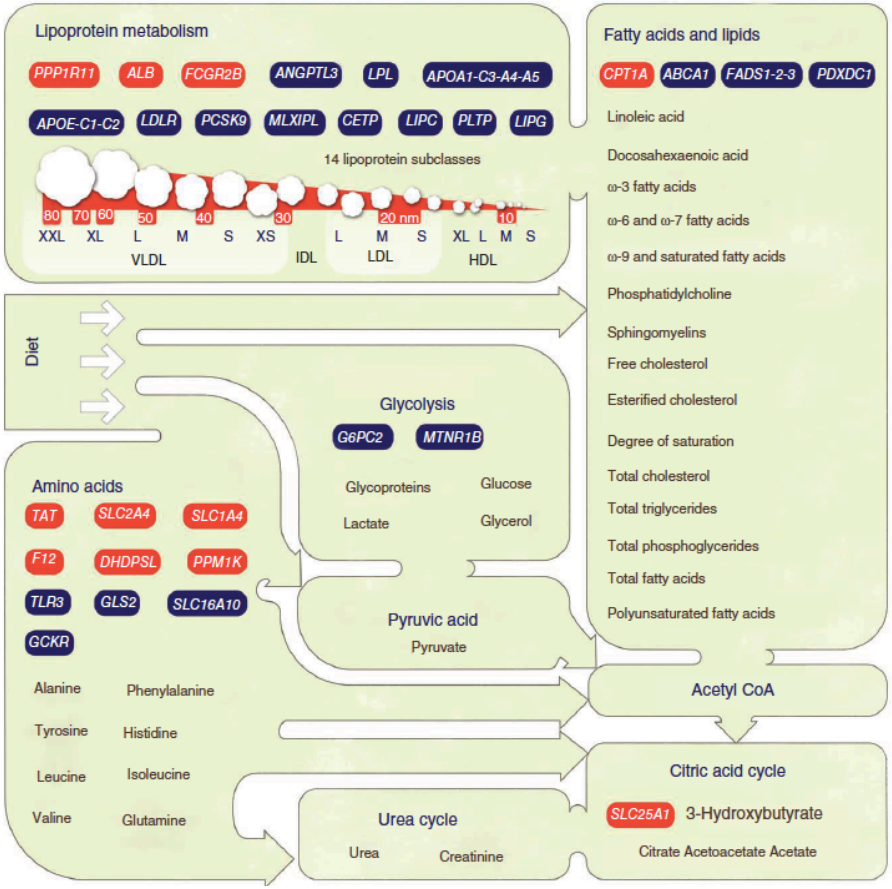
Novembre et al. Nature 2008



At each of these loci, genetic variation is involved in disease etiology

Metabolomics + genomics





Causal variants are still unknown

- Non-synonymous
- Alternative splice forms
- Expression levels (eQTLs)
- Other regulatory (miRNAs)?

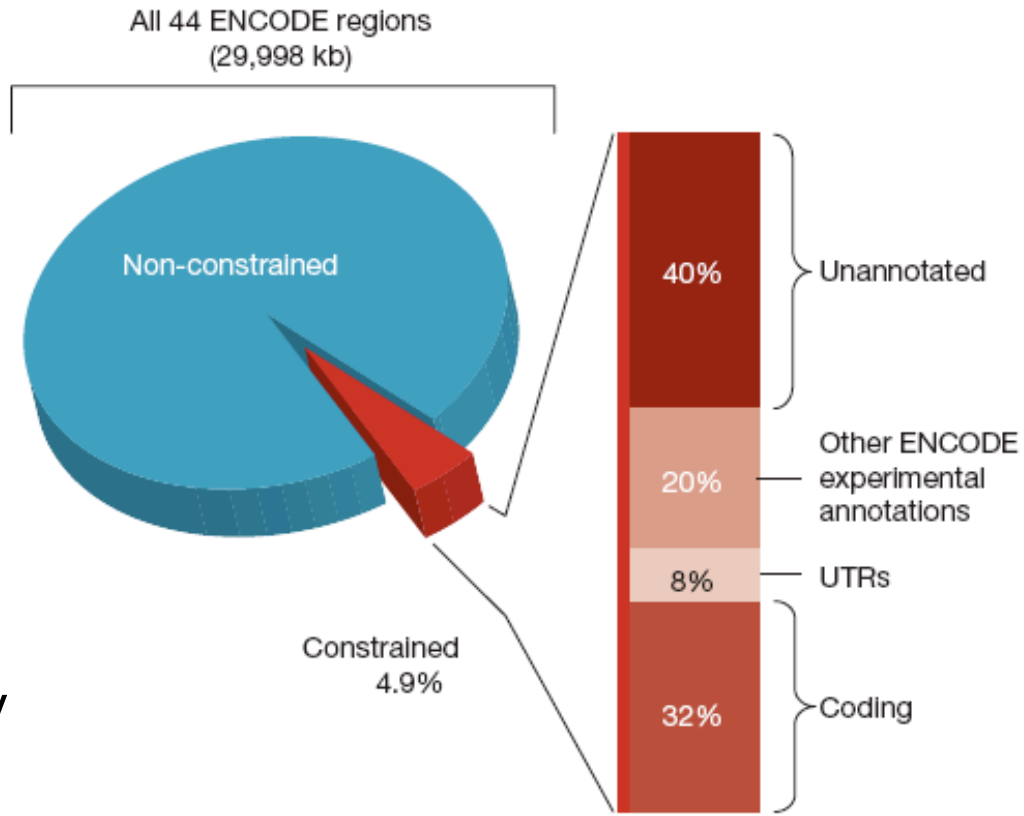
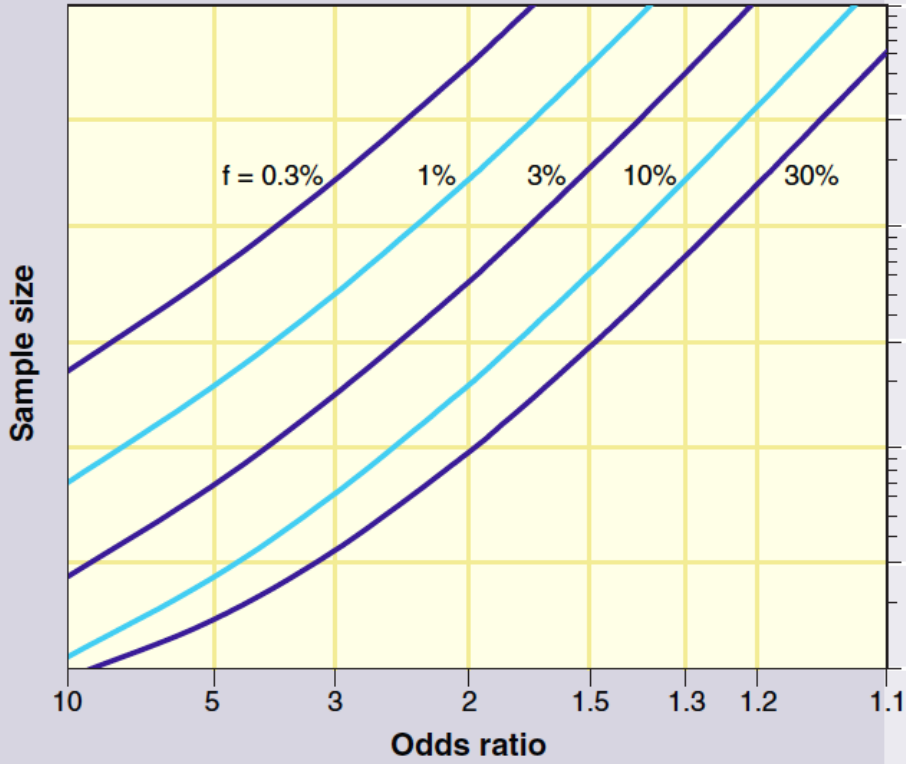


Figure from ENCODE Consortium. Nature. 2007

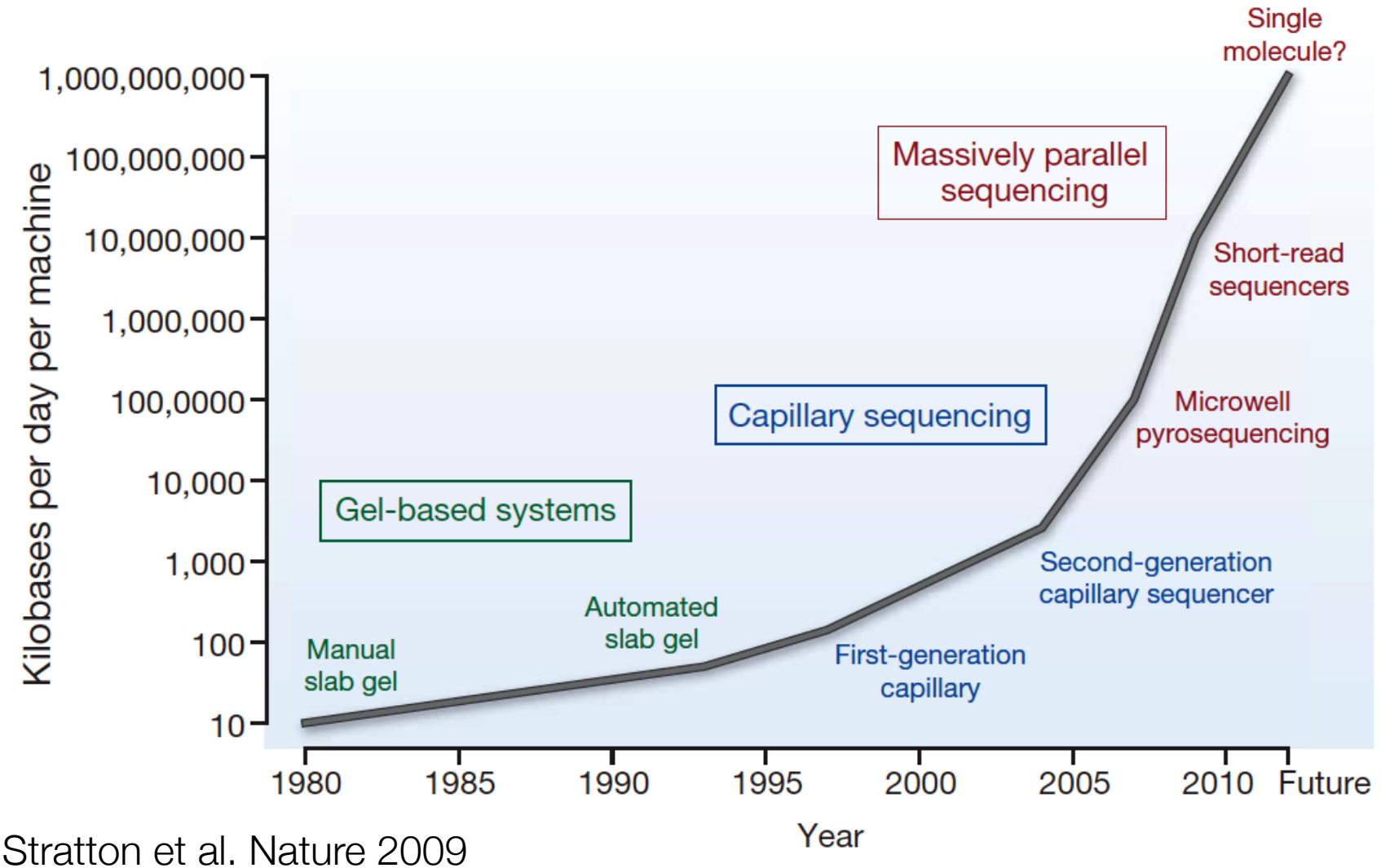
More samples = more power



	GWAS			Rare variant	Nominal
	90%	50%	10%	90%	90%
100,000	66,790	40,260	66,040	21,370	
30,000	20,037	12,078	19,812	6411	
10,000	6679	4026	6604	2137	
3000	2004	1208	1981	641	
1000	668	403	660	214	
300	200	121	198	64	
100	67	40	66	21	

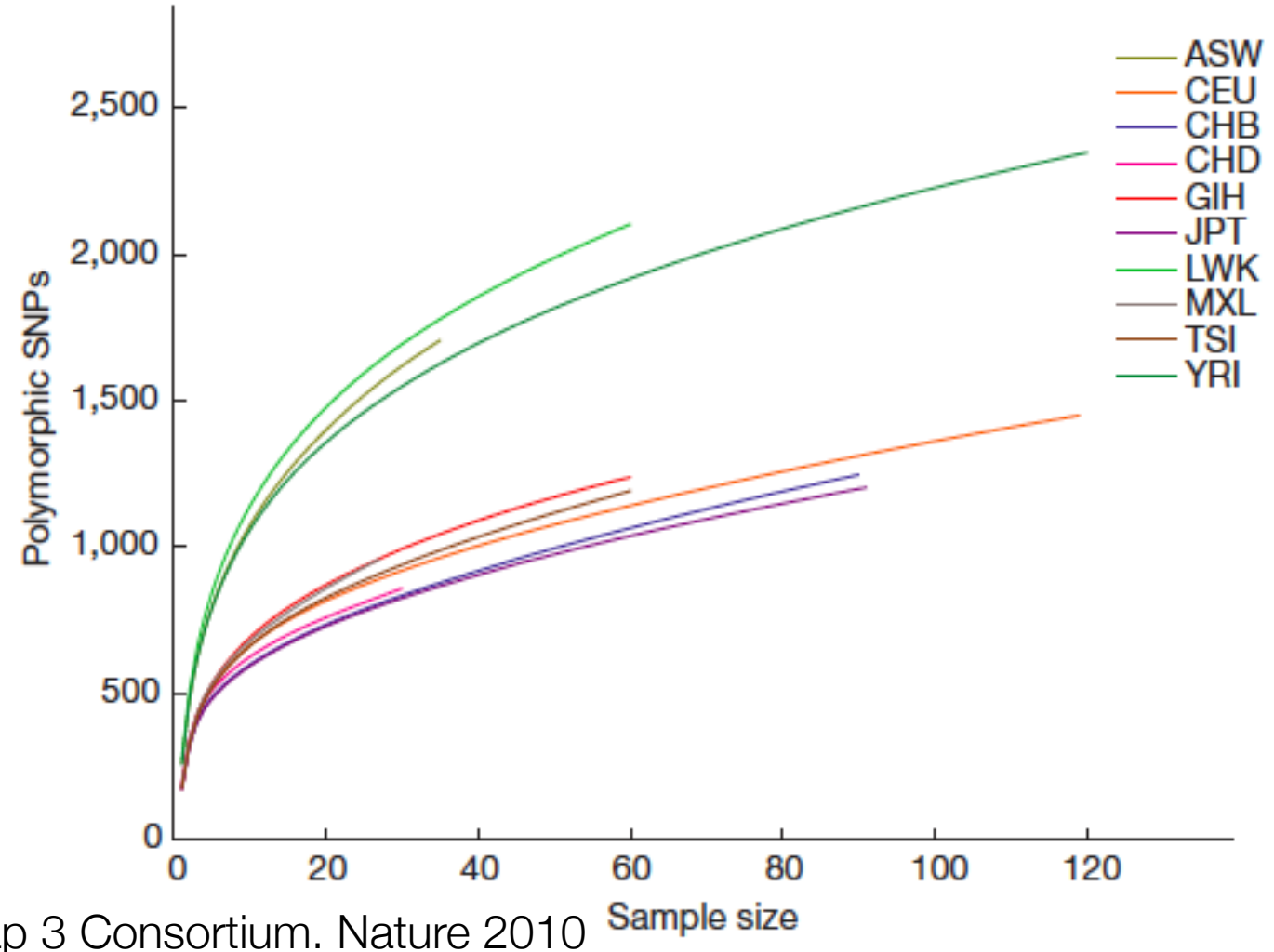
Altshuler, Daly & Lander Science 2008

Next-generation sequencing



Stratton et al. Nature 2009

Sequencing in more people will continue to find novel variants



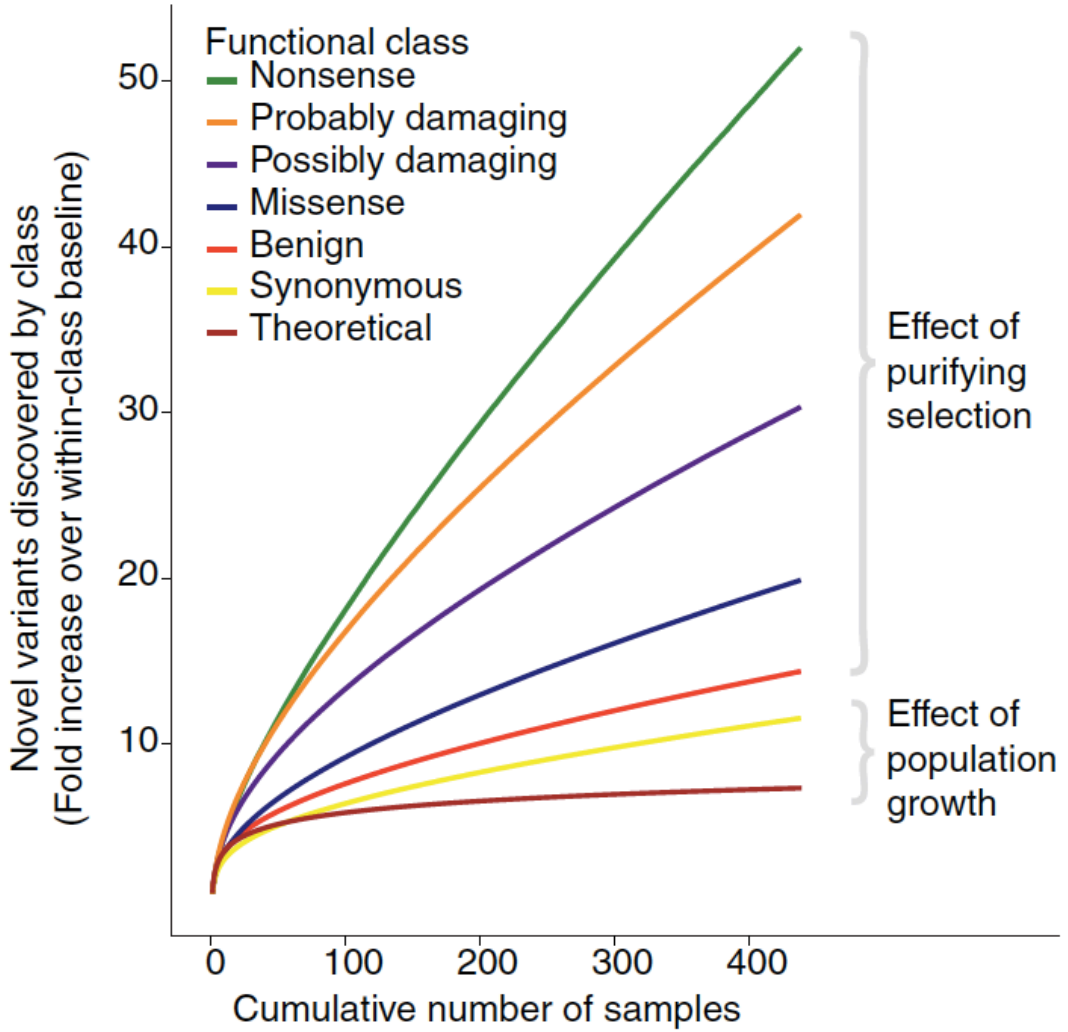
HapMap 3 Consortium. Nature 2010

“Normal” individuals carry hundreds of deleterious and big-effect mutations

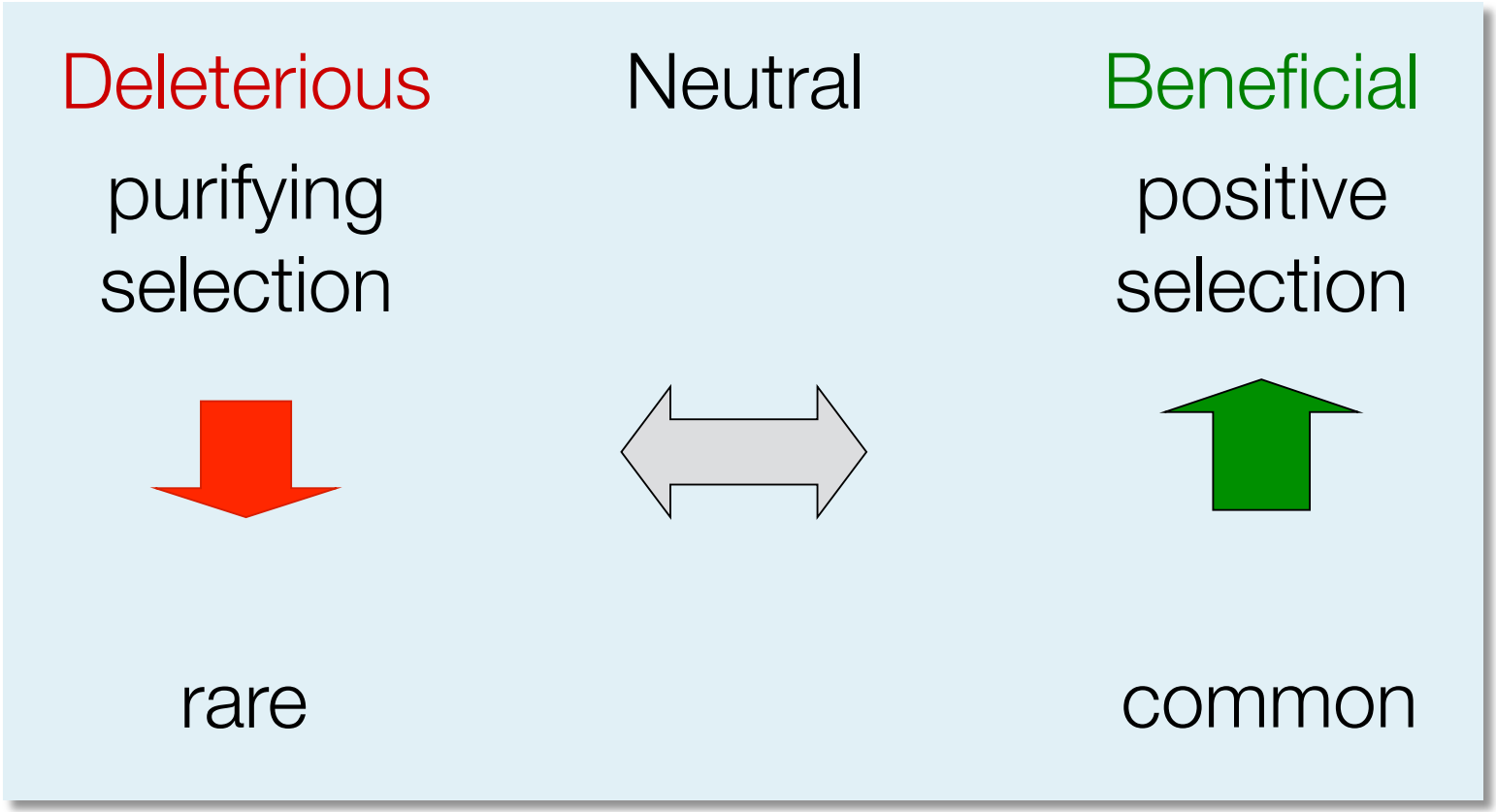
	Low Coverage	
class	**Total**	**Per sample**
synonymous SNPs	55217	10572-12126
nonsynonymous SNPs	61284	9966-10819
small in-frame indels	666	198-205
stop losses	71	9-11
stop-introducing SNPs	951	88-101
splice-site-disrupting SNPs	500	41-49
small frameshift indels	890	227-242
genes disrupted by large deletions	143	28-36
total genes containing LOF variants	1795	272-297
HGMD "damaging mutation" SNPs	578	57-80

Interpretation of genomic variation in a single patient is very challenging!!!

Abundance of rare, functional variation

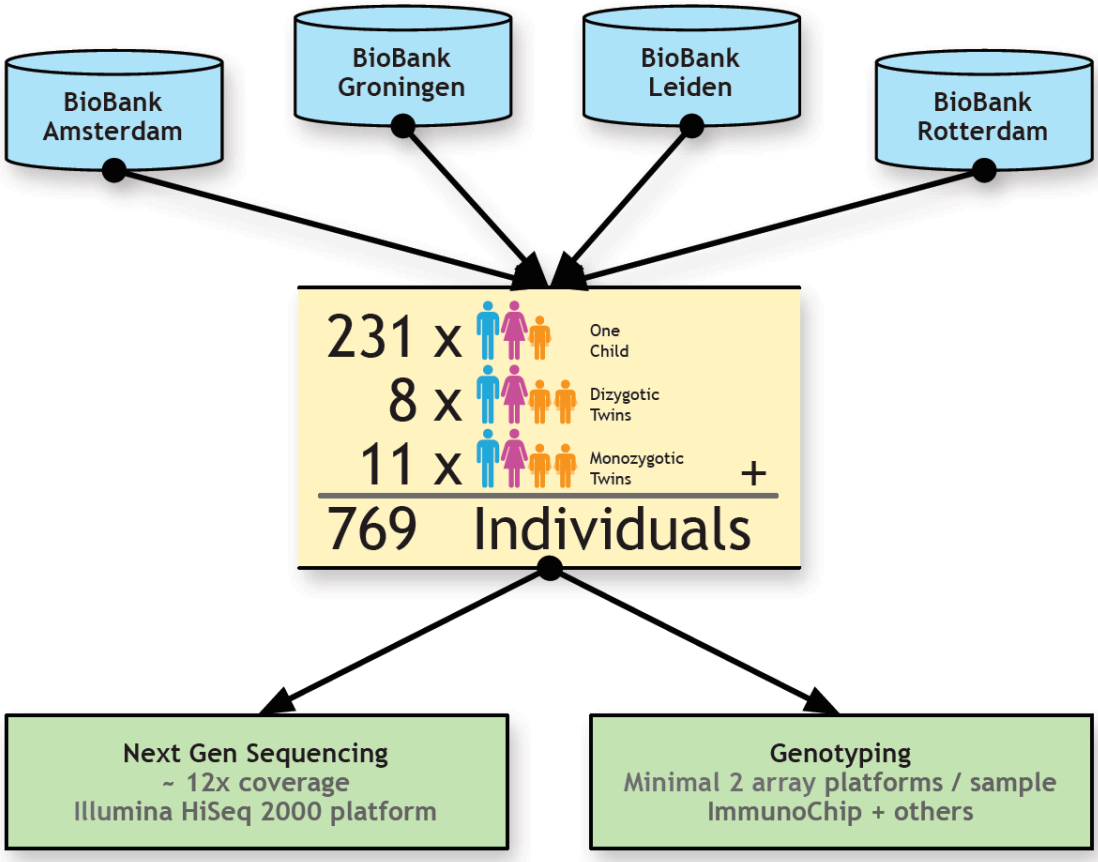


Effect of natural selection on population frequency of disease alleles



Go•NL

GENOMEoftheNETHERLANDS



Genoom van Nederland

- Catalogus van hoogfrequente en laagfrequente varianten in de Nederlandse populatie
- Internationale samenwerking
- Data gegenereerd door BGI
- Gefinancierd door BBMRI-NL

华大基因 BGI

BBMRI·NL

Opinions

- Cost-effectiveness of national HPC infrastructure
- Accessibility and transparency is key
- Nature of research
- Worldwide infrastructure impractical?
- Sharing of data and tools
- Cost model

Attention to teamwork has secured wins for the New England Patriots.


pdebakker@umcutrecht.nl
POŠTA