17th Annual Meeting of the Organization for Human Brain Mapping
Introduction to Imaging Genetics
Quebec City, Canada – June 25-10, 2010

Structure and Analysis of Genetic Variation

Fabio Macciardi MD PhD

Dept of Psychiatry and Human Behavior

Lab of Molecular Psychiatry

University of California, Irvine

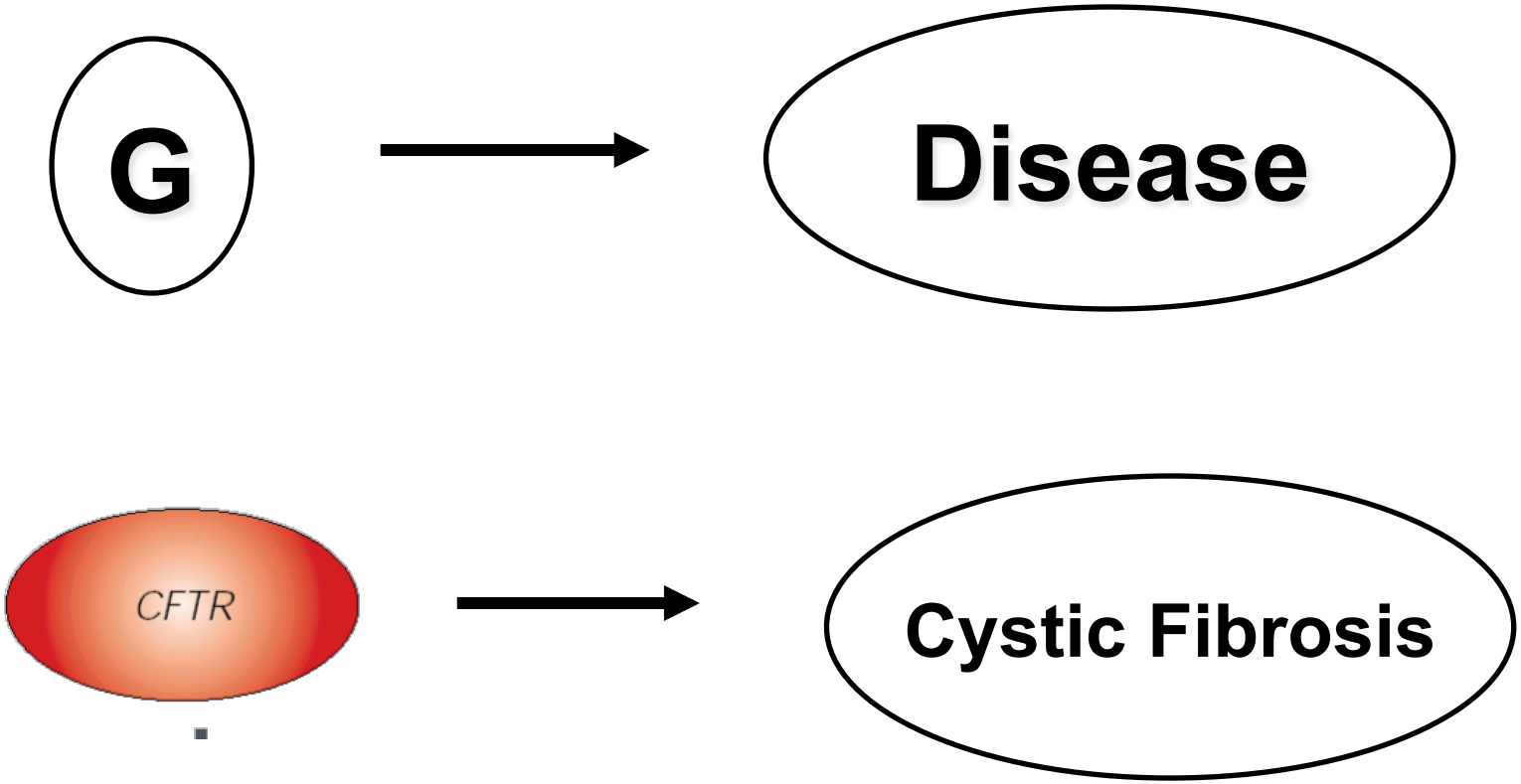
UCI
Brain Imaging Center

Structure & Analysis of Genetic Variation

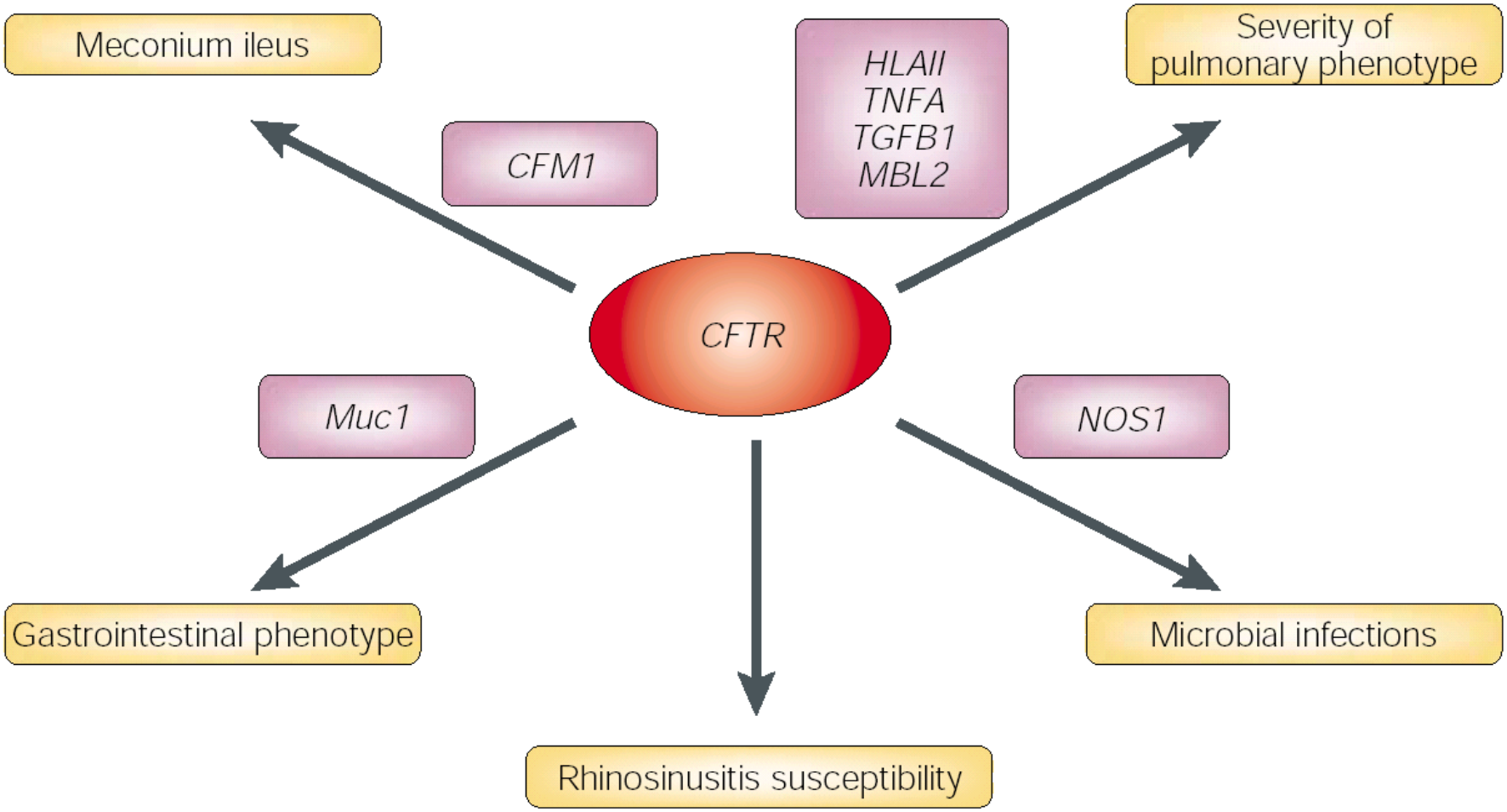
- What we will look into:
  - Genes, promoters, micro RNA
  - SNP, CNV, microsatellites
  - methylation
  - Cis - trans-acting & epistasis
  - transcriptome to Genes
  - Genes to Pathways

Let's see first how we have modified our idea of what a genetic disorder looks like ...

The "old" paradigm of genetics …

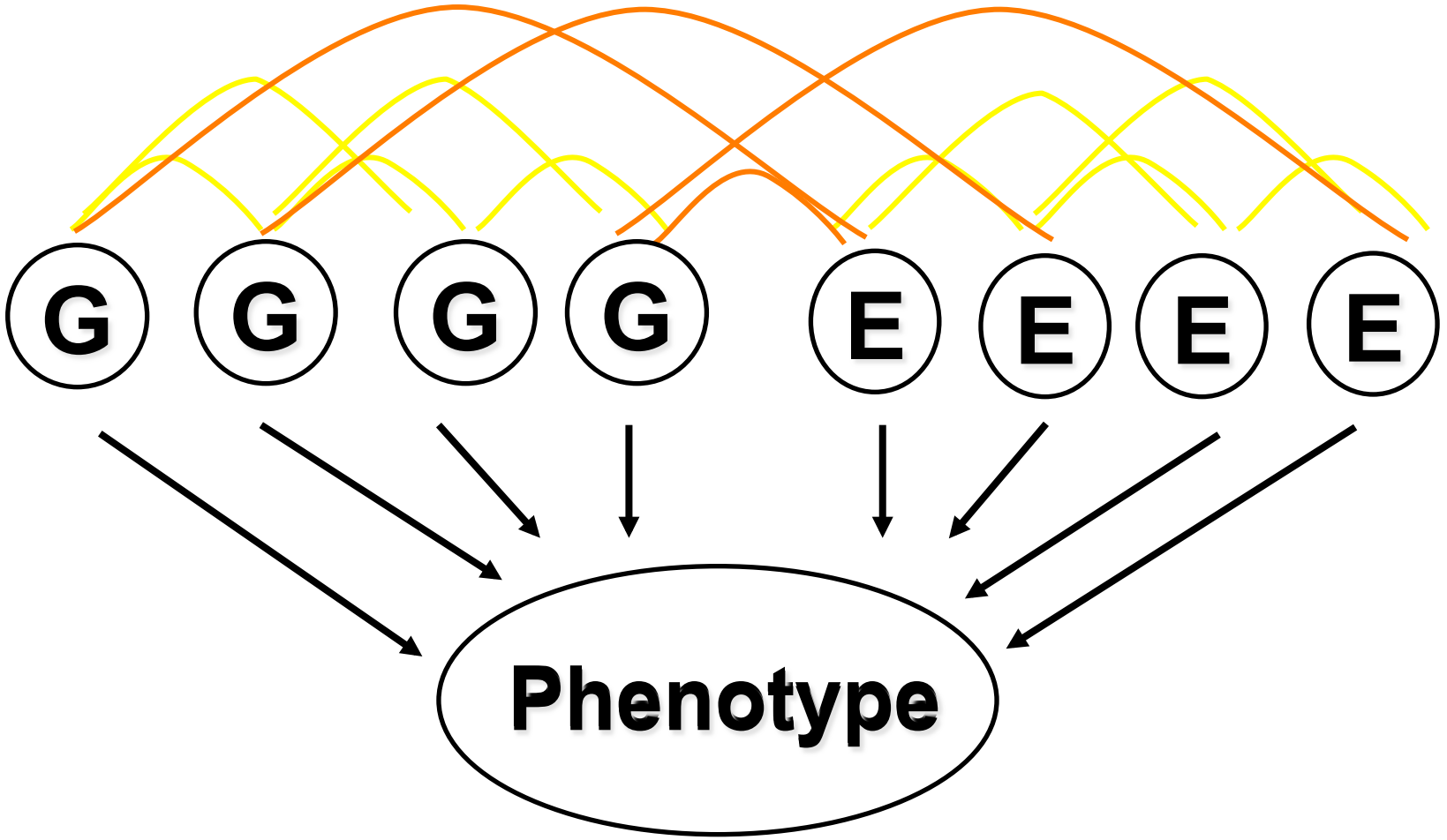


… but even simple "mendelian" disorders are not THAT simple …

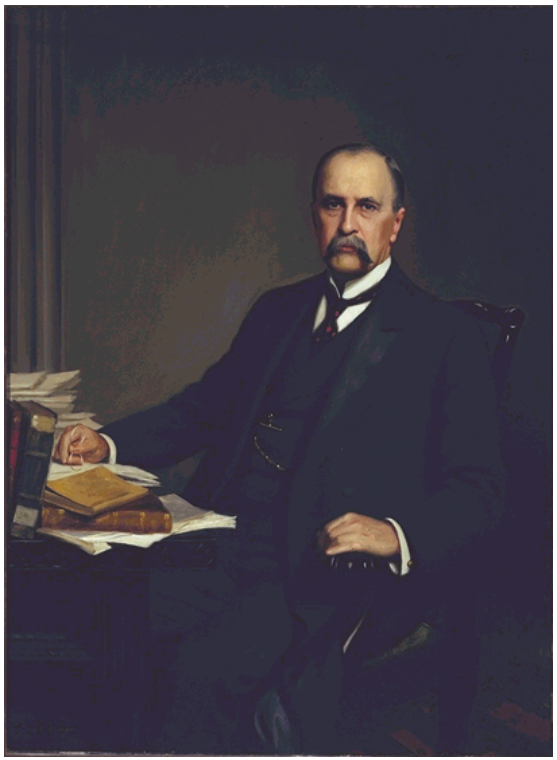


… and how now we think genetics works..

Multiple gene variants interacting with each other, and with multiple environmental factors



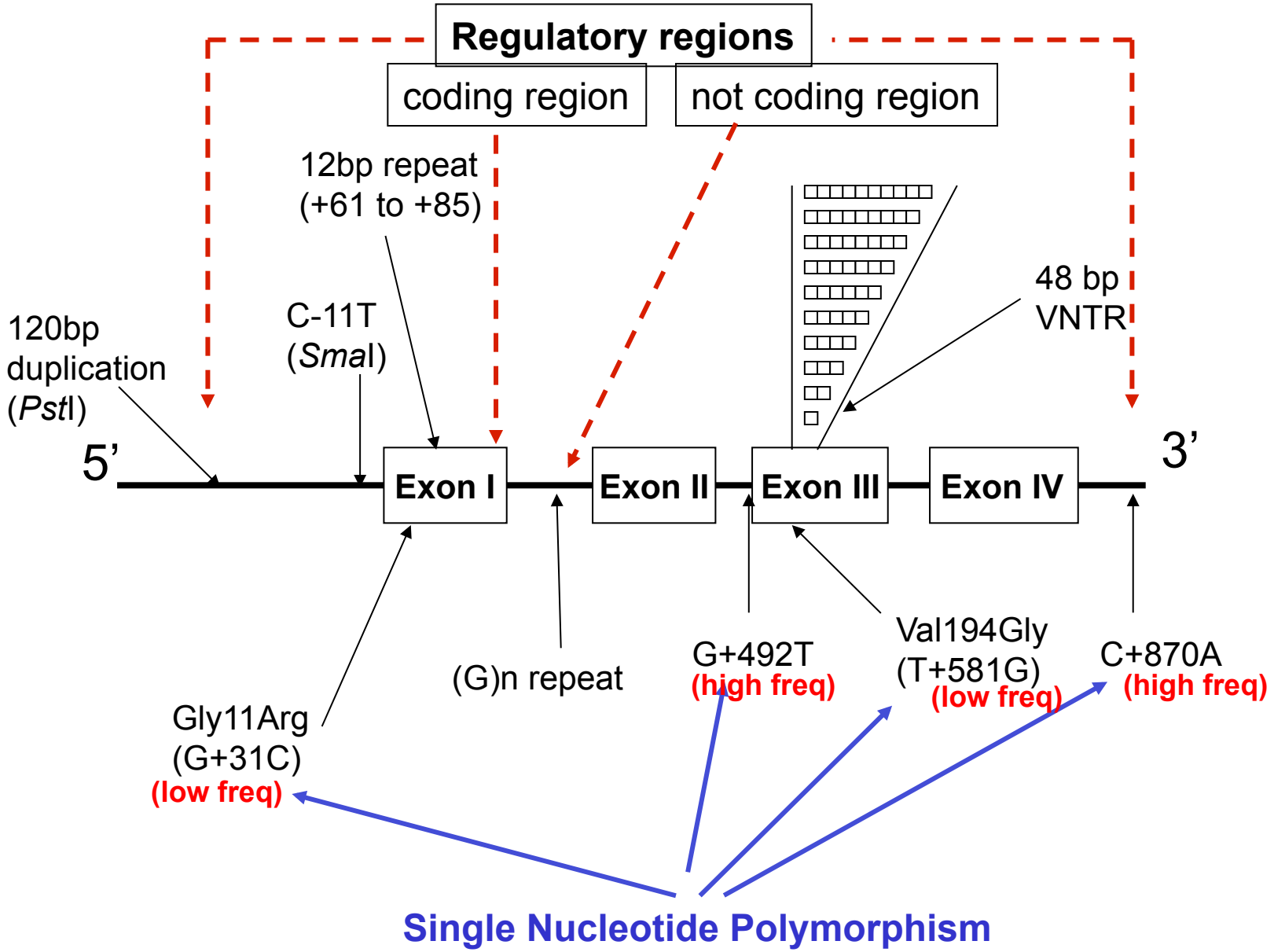
Variability in human



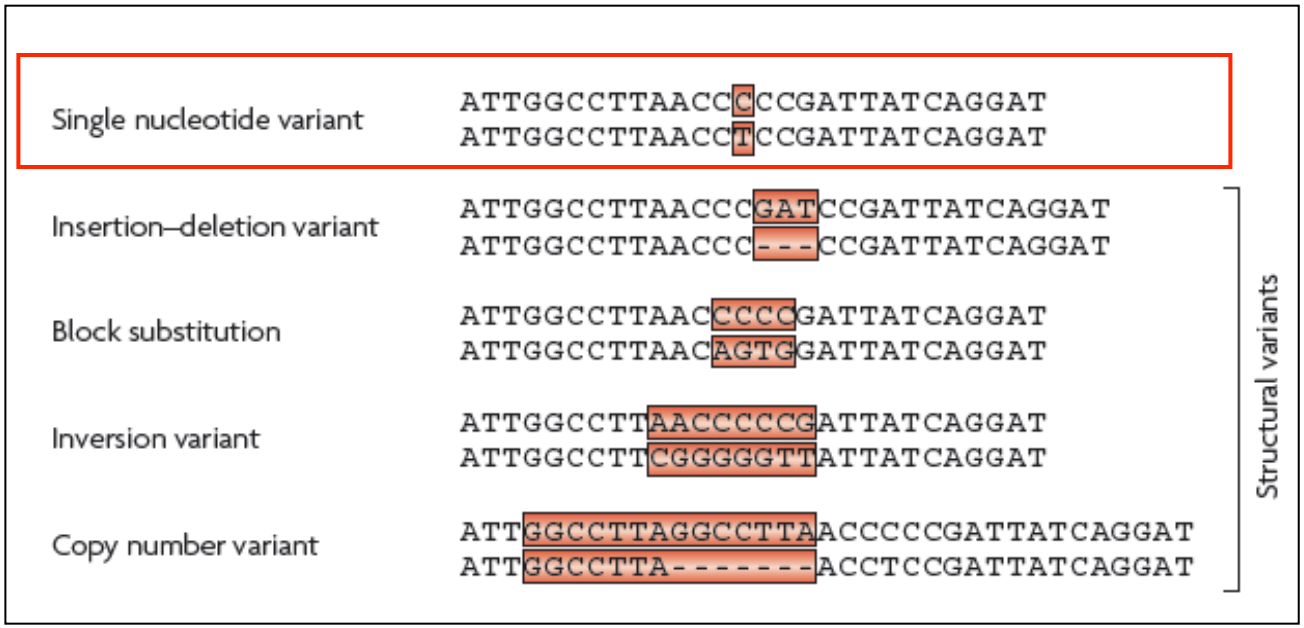
Sir William Osler
(1849-1919)

"Variability is the law of life, and as no two faces are the same, so no two bodies are alike, and no two individuals react alike, and behave alike under the abnormal conditions which we know as disease."

What is a gene? A real example, the DRD4

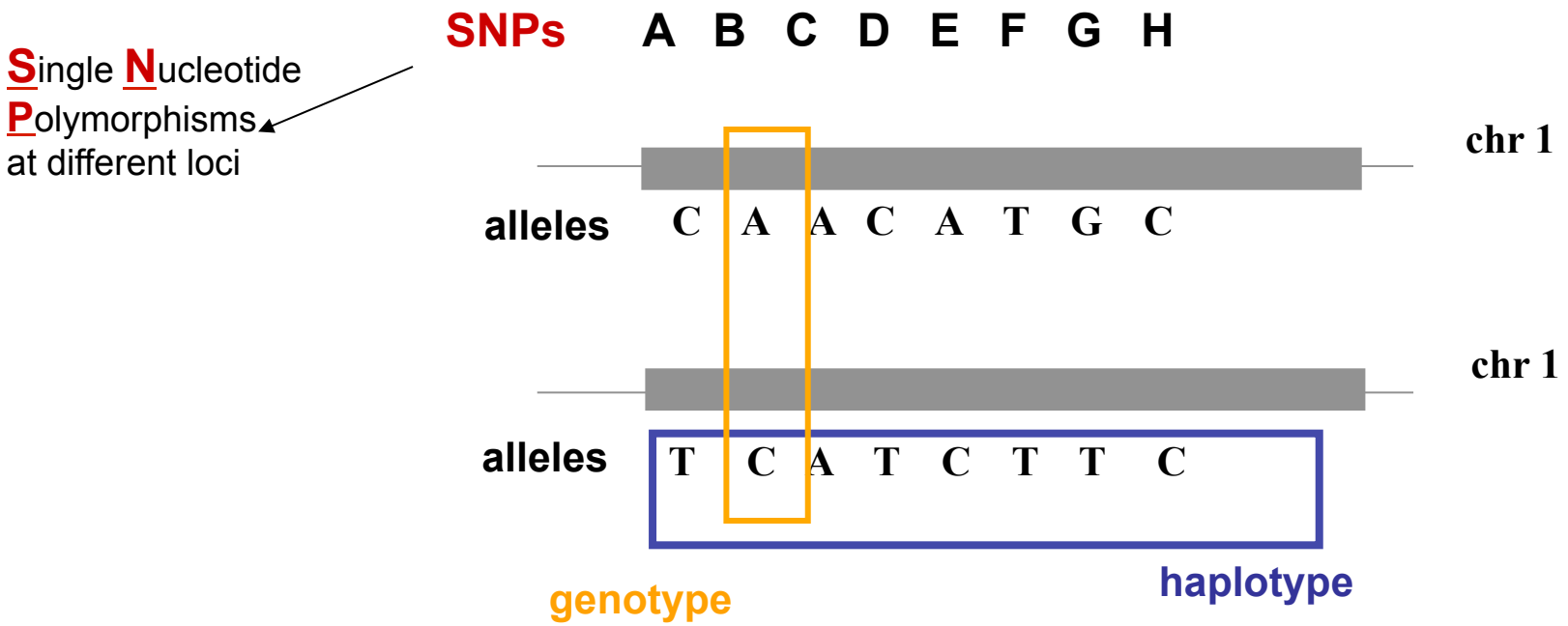


Classes of human genetic variants



Frazer et al., 2009 Nature Rev Genet. 10:241-51.

Key-words in genetics

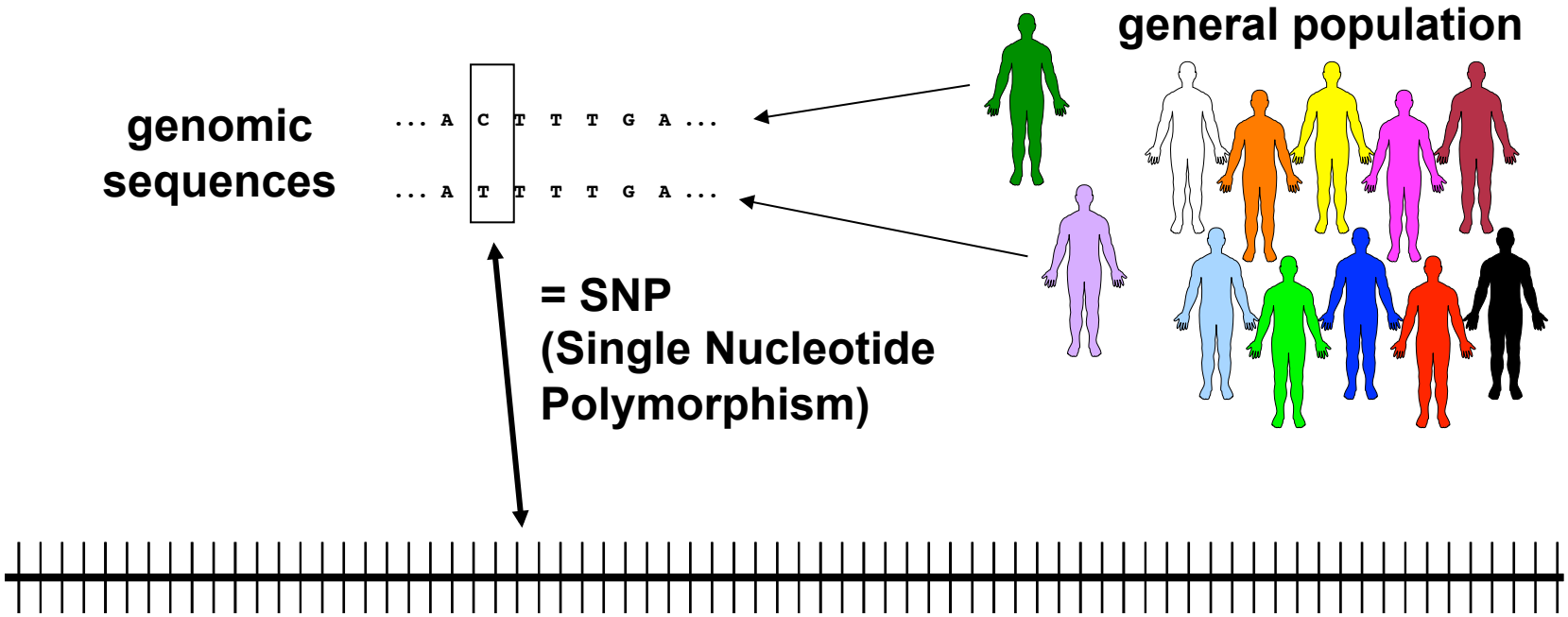


an "allele" …

Example = gene t ic – g en etic
. ..
r am

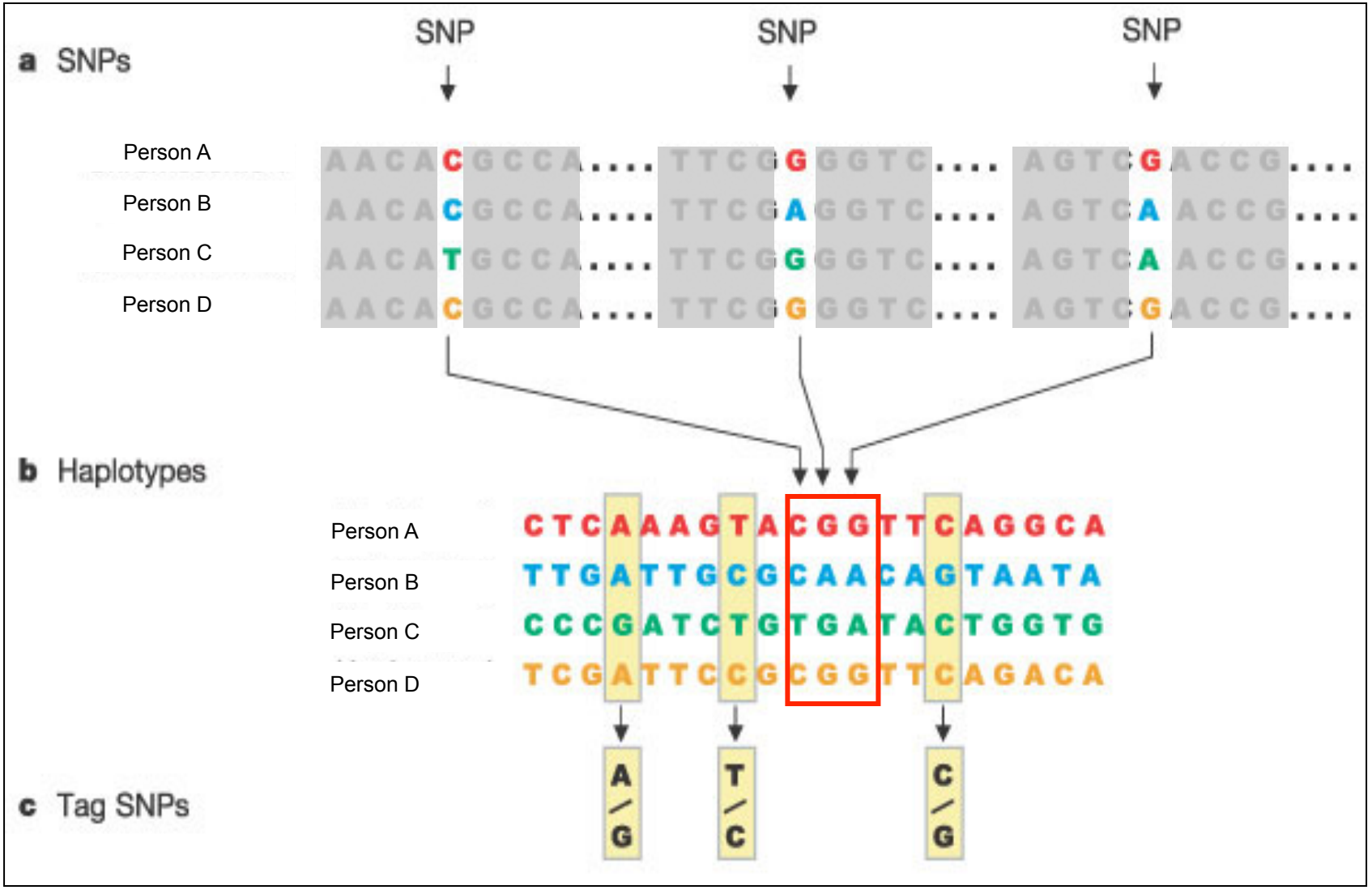
… and a haplotype

We found variable “letter” SNPs in all individuals..

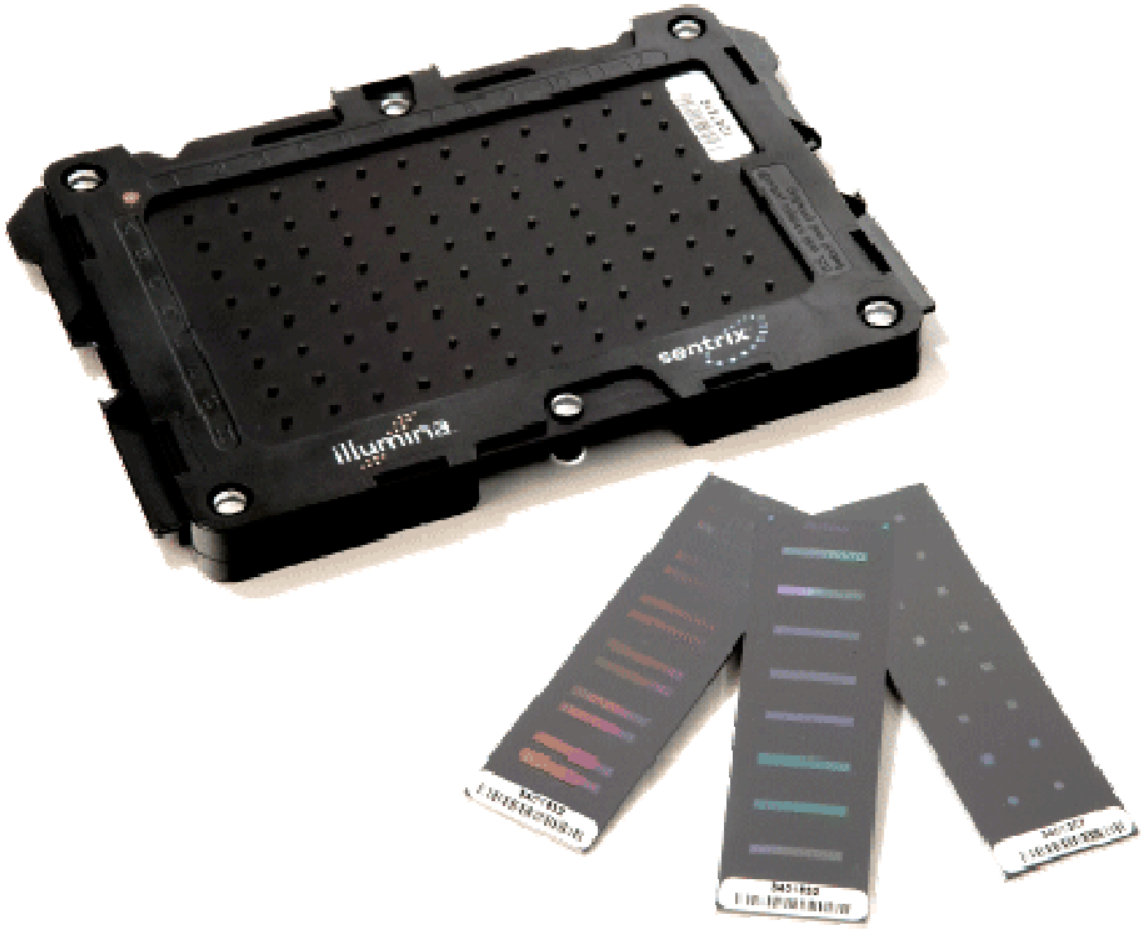


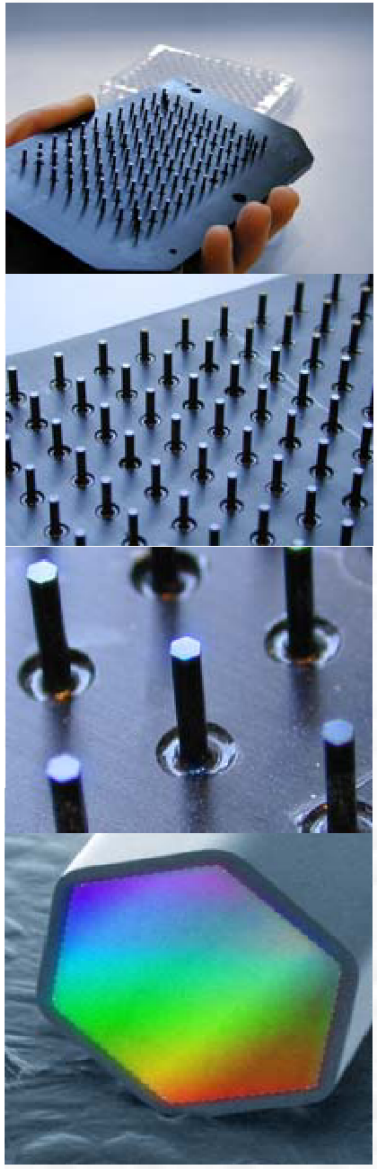
… today we know almost 15 - 18 million common SNPs (and many more not so common)

…and identify the genetic signature of each of us



A new technology: DNA MICROARRAYS Allow us to detect these SNPs ….

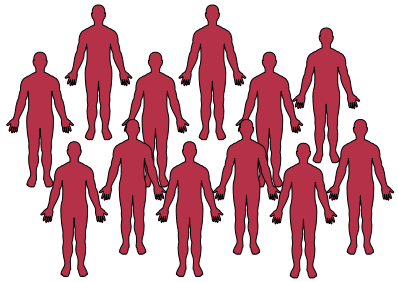




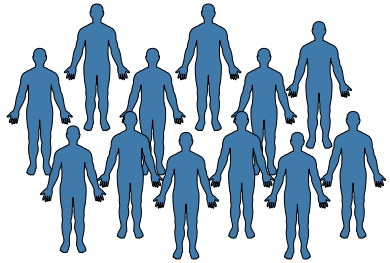
Population-based designs: what does practically means?

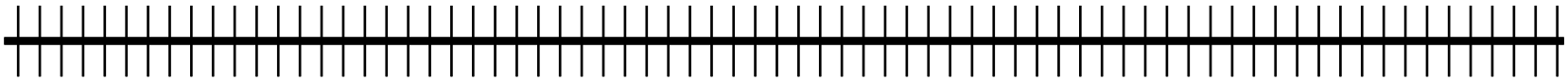
Population-based **:Cases and unrelated population controls from the same study base**

Affected Individuals (CASES)



Not affected Individuals (CONTROLS)





	A	a
CASES	656	879
CONTROLS	525	471

p-value = 0.1

	A	a
CASES	606	929
CONTROLS	555	441

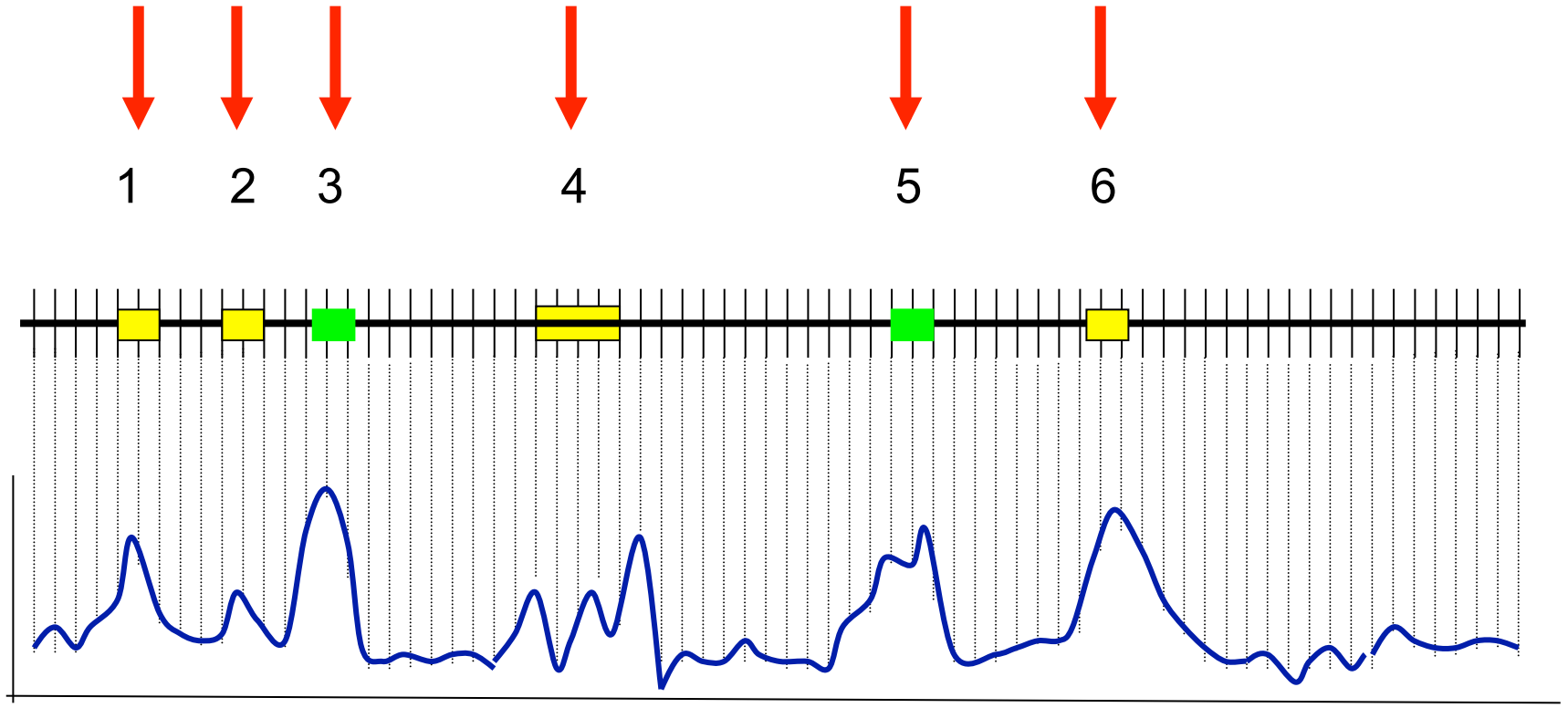
p-value = 0.3

	A	a
CASES	856	679
CONTROLS	325	671

p-value = 0.01 !!!
SIGNIFICANT !!

…and so on!

WHOLE GENOME SCAN ASSOCIATION



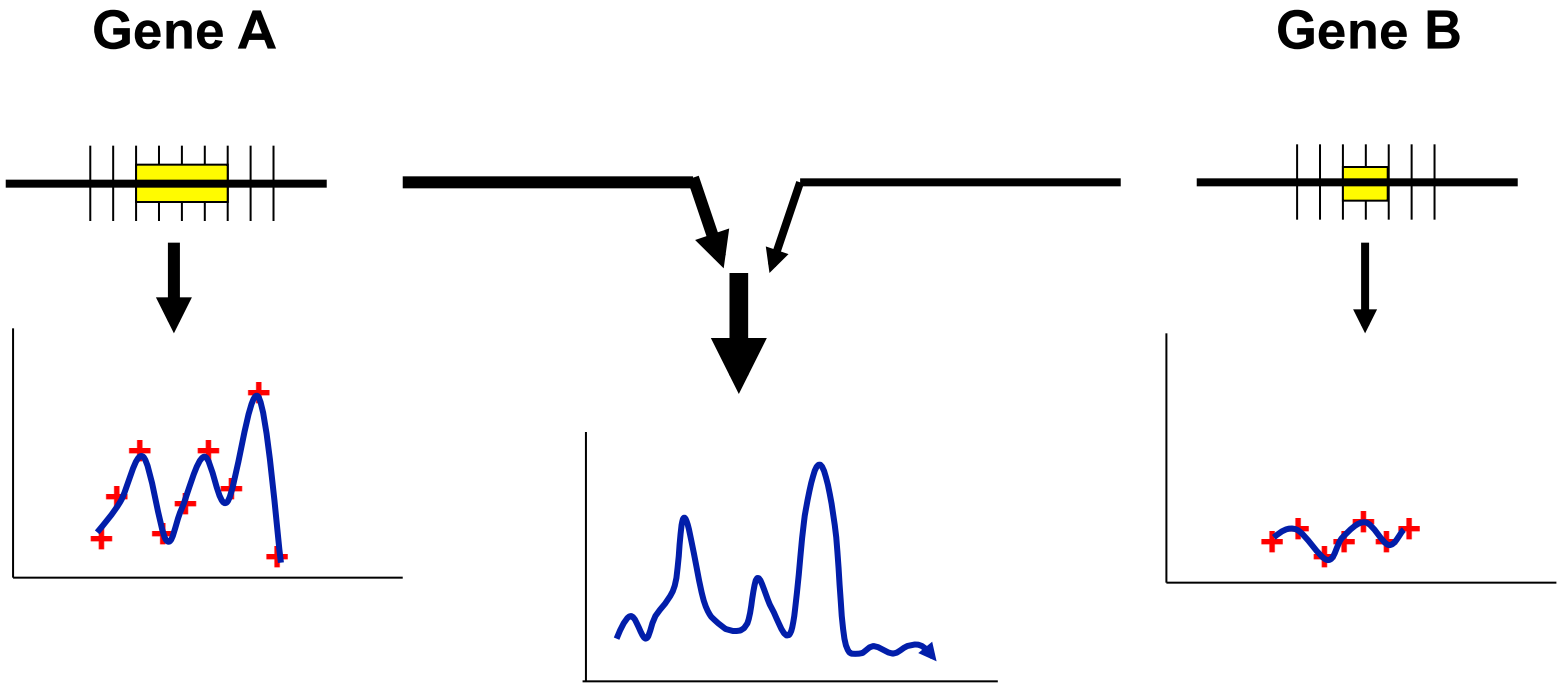
Known Gene

New Gene

How many genes?

- In complex traits, there are genes acting together and we must understand “how” if we want to understand the biology of disease:

  modeling gene^gene interactions – the Epistasis effect



Discovering interactions among BRCA1 and other candidate genes associated with sporadic breast cancer

Shaw-Hwa Lo[†‡], Herman Chernoff[‡§], Lei Cong[†], Yuejing Ding[†], and Tian Zheng[†]

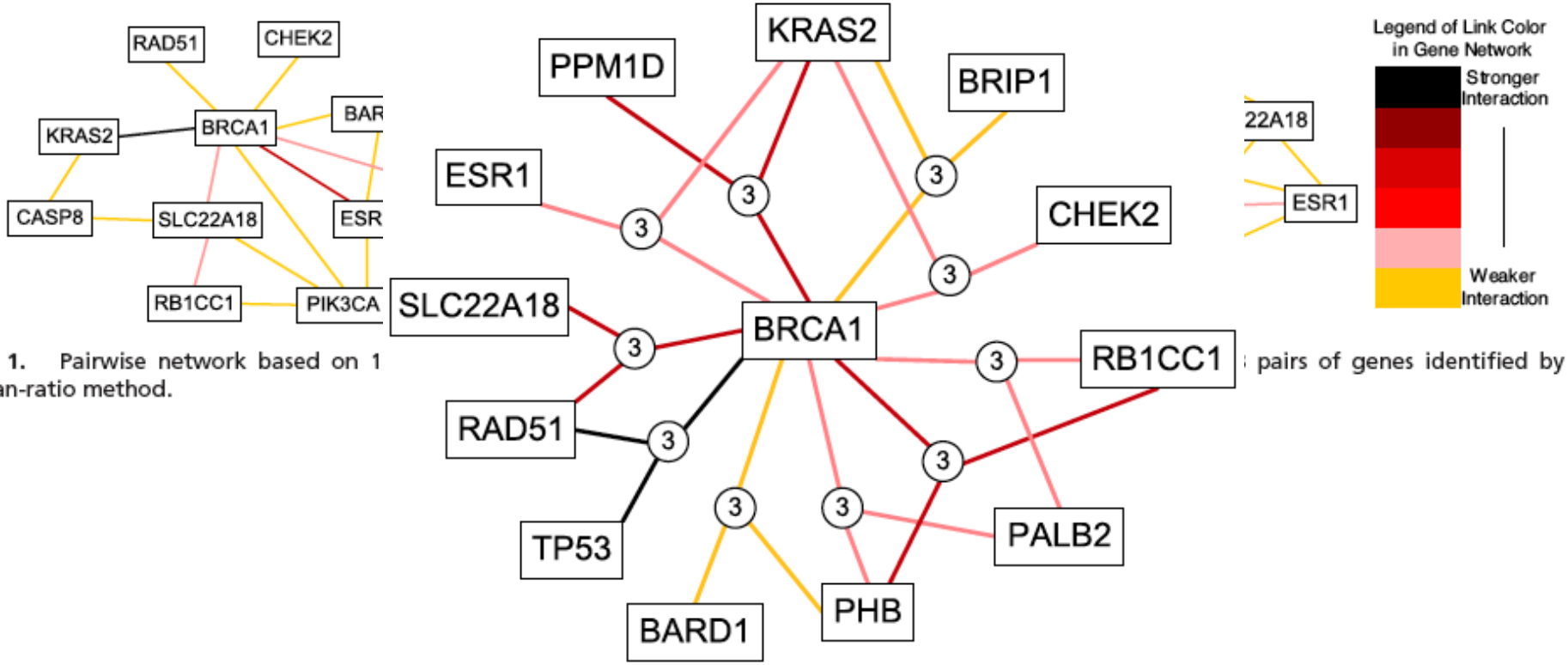
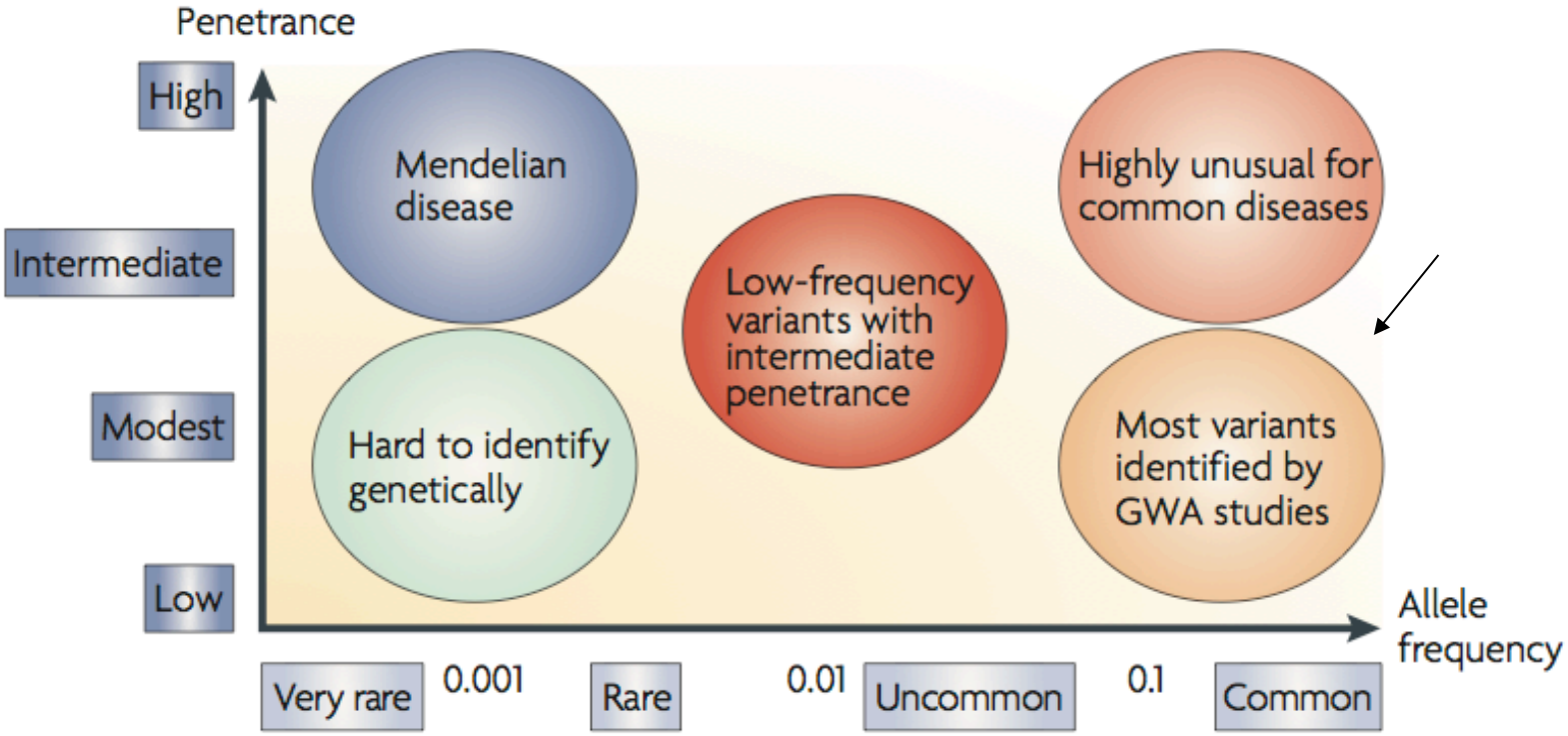


Fig. 1. Pairwise network based on 1 pairs of genes identified by the Mean-ratio method.

PNAS 105: 12387-92, 2008

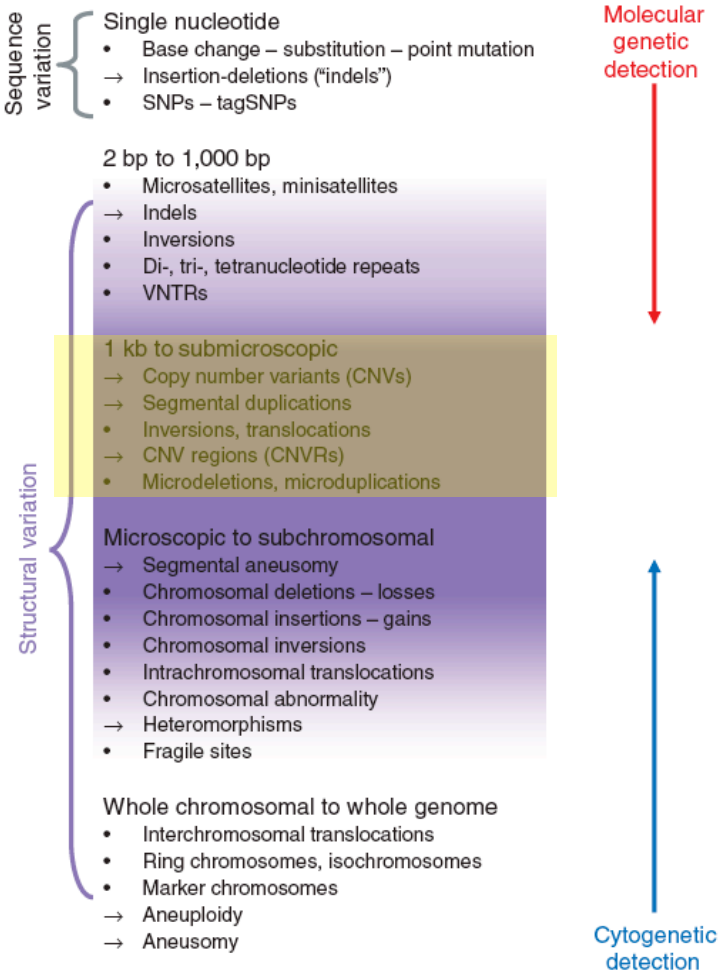
GWAS..here we are..



Outcome of a Genome wide association study

What are CNVs?

Stretches of DNA larger than 1 kb that display copy number differences in comparison to a reference genome



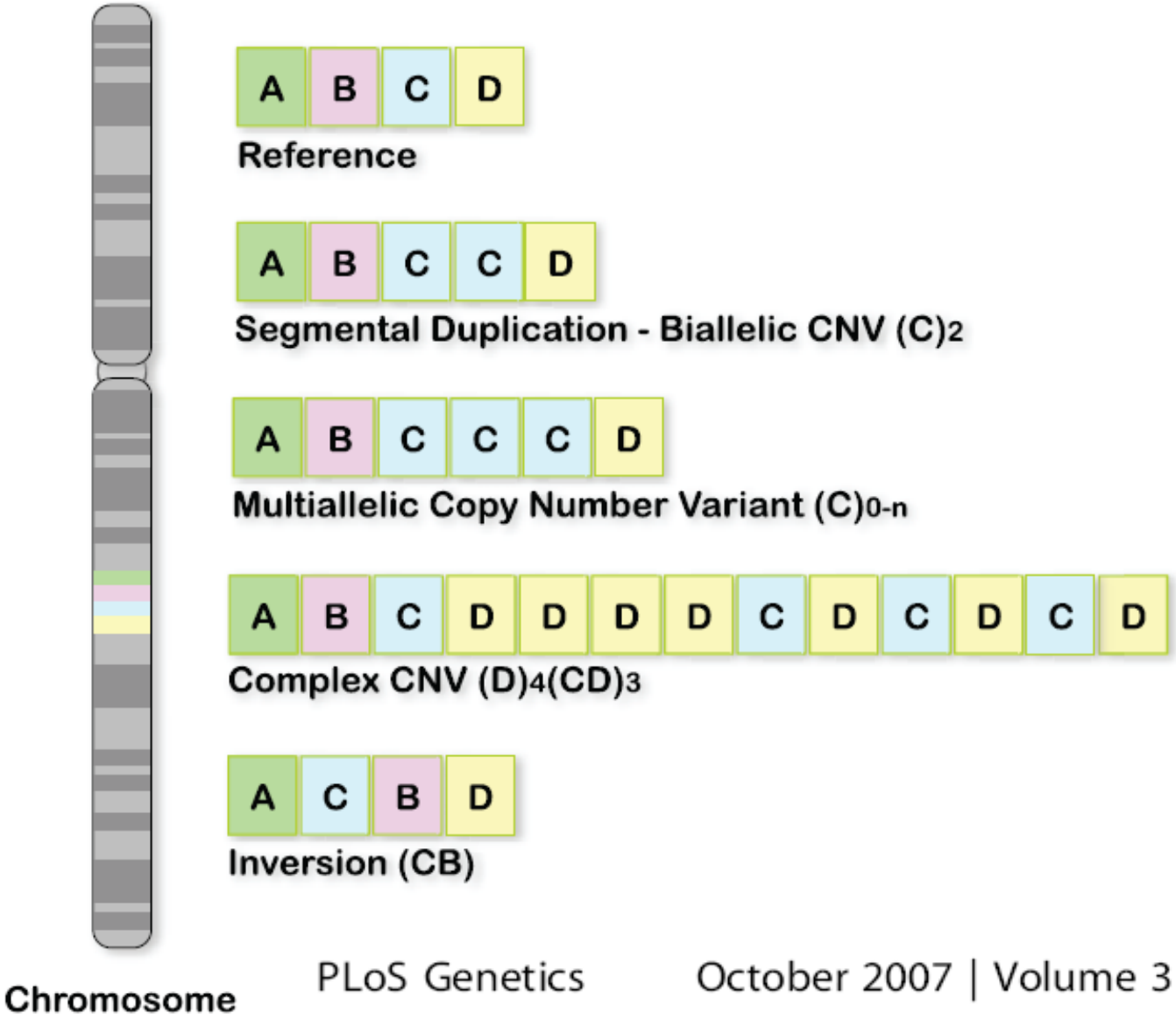
- Copy number **variation** (germline, inherited)
  - inherited: also present in parents' genome
  - *de novo:* absent in parents' genome

- Copy number **alteration** (somatic, e.g. in cancer cells)

- Copy number **polymorphism** (relatively common CNV, with a fixed starting/ending position)

- Copy number **difference** (between-species copy number differences, e.g chimpanzees and humans)

Types of Genomic Structural Changes Affecting Segments of DNA Leading to Deletions, Duplications, Inversions, and CNV Changes (biallelic, Multiallelic, and Complex



PLoS Genetics October 2007 | Volume 3 | Issue 10

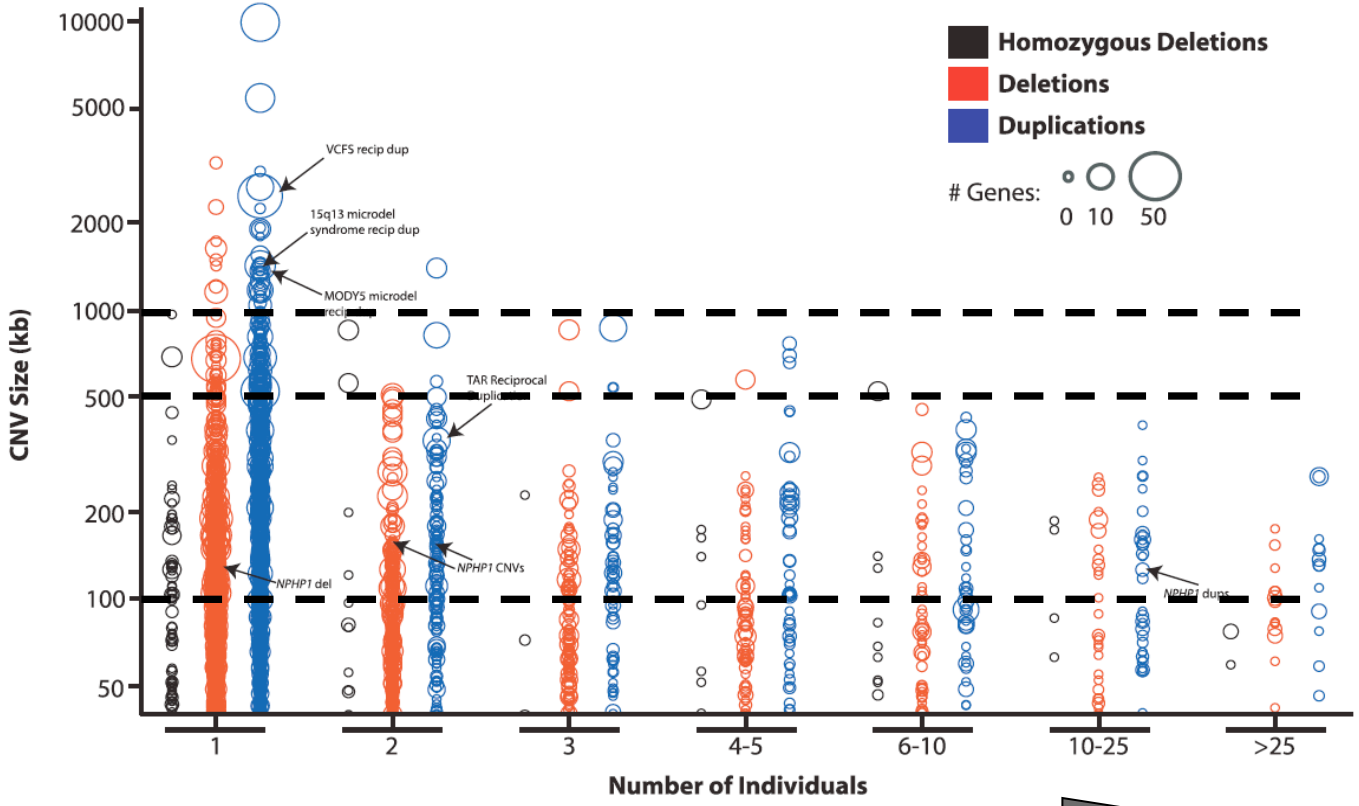
The current map of human structural variation is far from complete….

ARTICLE

Population Analysis of Large Copy Number Variants and Hotspots of Human Genetic Disease

Andy Itsara,[1,7] Gregory M. Cooper,[1,7] Carl Baker,[1,2] Santhosh Girirajan,[1,2] Jun Li,[2] Devin Absher,[3] Ronald M. Krauss,[4] Richard M. Myers,[3] Paul M. Ridker,[5] Daniel I. Chasman,[5] Heather Mefford,[1] Phyllis Ying,[1] Deborah A. Nickerson,[1] and Evan E. Eichler[1,6,*]

Am J Hum Genet. 2009;84:148-61



~16% of the genome

Low frequency

High frequency

… genomic burden of rare variants

Rare Structural Variants Disrupt Multiple Genes in Neurodevelopmental Pathways in Schizophrenia

Tom Walsh,[1*] Jon M. McClellan,[2*†] Shane E. McCarthy,[3*] Anjené M. Addington,[4*] Sarah B. Pierce,[1] Greg M. Cooper,[5] Alex S. Nord,[5] Mary Kusenda,[3,6] Dheeraj Malhotra,[3] Abhishek Bhandari,[3] Sunday M. Stray,[1] Caitlin F. Rippey,[5] Patricia Roccanova,[3] Vlad Makarov,[3] B. Lakshmi,[3] Robert L. Findling,[7] Linmarie Sikich,[8] Thomas Stromberg,[4] Barry Merriman,[9] Nitin Gogtay,[4] Philip Butler,[4] Kristen Eckstrand,[4] Laila Noory,[4] Peter Gochman,[4] Robert Long,[4] Zugen Chen,[9] Sean Davis,[10] Carl Baker,[5] Evan E. Eichler,[5] Paul S. Meltzer,[10] Stanley F. Nelson,[9] Andrew B. Singleton,[11] Ming K. Lee,[1] Judith L. Rapoport,[4] Mary-Claire King,[1,5] Jonathan Sebat[3]

Science 320: 539-543, 2008

Genes disrupted by SV breakpoints	
NBPF10	-
SLC35F3, TARBP1	-
STON1-GTF2A1L	
ERBB4	*FRRS1*
GRM7	-
PRKCD	*ROBO1*
-	*MANEA*
SKP2, SLC1A3	-
MAGI2, PHTF2	*FLJ31818*
SLC12A9, CAV1	*SND1*
PRKAG2, MLL3	*CTSB*
PTK2	*MPDZ*
SMARCA2	*SOX5, LYRM5*
-	*TMTC1*
-	*HYDIN*
HIPK3, C11orf41	*BPIL2*
DLG2	
-	
LAMA1, PTPRM	
TMC4	
LARGE	

LETTERS

nature genetics

Strong association of *de novo* copy number mutations with sporadic schizophrenia

Bin Xu[1,2], J Louw Roos[3], Shawn Levy[4], E J van Rensburg[5], Joseph A Gogos[1,6] & Maria Karayiorgou[2]

Nat Rev Genet 40: 8881-885, 2008

Genetics of complex disorders: what has been achieved so far

"Traditional" genetic methods (eg, association) in a genomic perspective point to extreme alternatives:

- GWAS → CVCD
- CNV/CNP → RVCD

Integrated detection and population-genetic analysis of SNPs and copy number variation

Steven A McCarroll[1–4,10], Finny G Kuruvilla[1–4,10], Joshua M Korn[1–6], Simon Cawley[7], James Nemesh[1], Alec Wysoker[1], Michael H Shapero[7], Paul I W de Bakker[1,4,8], Julian B Maller[3], Andrew Kirby[3], Amanda L Elliott[1], Melissa Parkin[1], Earl Hubbell[7], Teresa Webster[7], Rui Mei[7], James Veitch[7], Patrick J Collins[7], Robert Handsaker[1], Steve Lincoln[7], Marcia Nizzari[1], John Blume[7], Keith W Jones[7], Rich Rava[7], Mark J Daly[1,3,4,9], Stacey B Gabriel[1] & David Altshuler[1–4,9]

Dissecting the genetic basis of disease risk requires measuring all forms of genetic variation, including SNPs and copy number variants (CNVs) .....

..... Most common, diallelic CNPs were in strong linkage disequilibrium with SNPs, and most low-frequency CNVs segregated on specific SNP haplotypes

Nature Genetics 40: 1168-74, 2008

Looking at individual structural variants with sequencing technologies

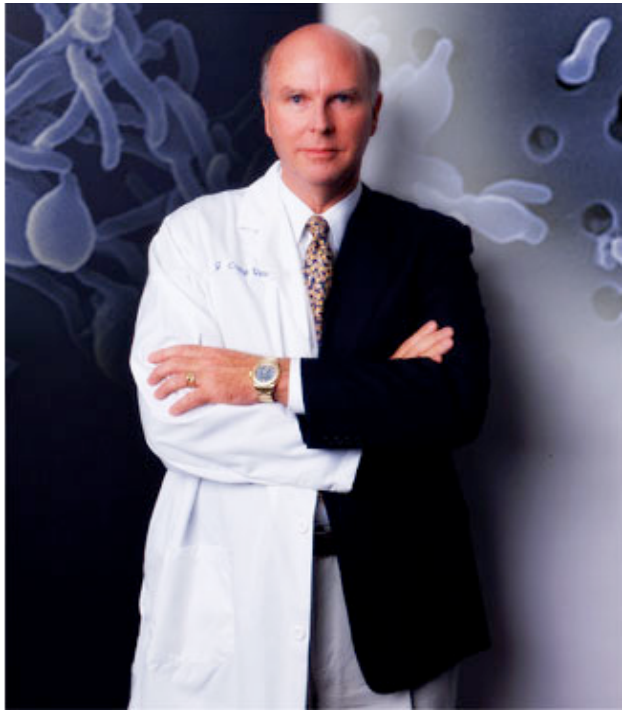
OPEN ACCESS Freely available online PLoS BIOLOGY

The Diploid Genome Sequence of an Individual Human

Samuel Levy[1*], Granger Sutton[1], Pauline C. Ng[1], Lars Feuk[2], Aaron L. Halpern[1], Brian P. Walenz[1], Nelson Axelrod[1], Jiaqi Huang[1], Ewen F. Kirkness[1], Gennady Denisov[1], Yuan Lin[1], Jeffrey R. MacDonald[2], Andy Wing Chun Pang[2], Mary Shago[2], Timothy B. Stockwell[1], Alexia Tsiamouri[1], Vineet Bafna[3], Vikas Bansal[3], Saul A. Kravitz[1], Dana A. Busam[1], Karen Y. Beeson[1], Tina C. McIntosh[1], Karin A. Remington[1], Josep F. Abril[4], John Gill[1], Jon Borman[1], Yu-Hui Rogers[1], Marvin E. Frazier[1], Stephen W. Scherer[2], Robert L. Strausberg[1], J. Craig Venter[1]

1 J. Craig Venter Institute, Rockville, Maryland, United States of America, 2 Program in Genetics and Genomic Biology, The Hospital for Sick Children, and Molecular and Medical Genetics, University of Toronto, Toronto, Ontario, Canada, 3 Department of Computer Science and Engineering, University of California San Diego, La Jolla, California, United States of America, 4 Genetics Department, Facultat de Biologia, Universitat de Barcelona, Barcelona, Catalonia, Spain

PLoS Biol. 2007 4;5:e254.



Looking at individual structural variants with sequencing technologies

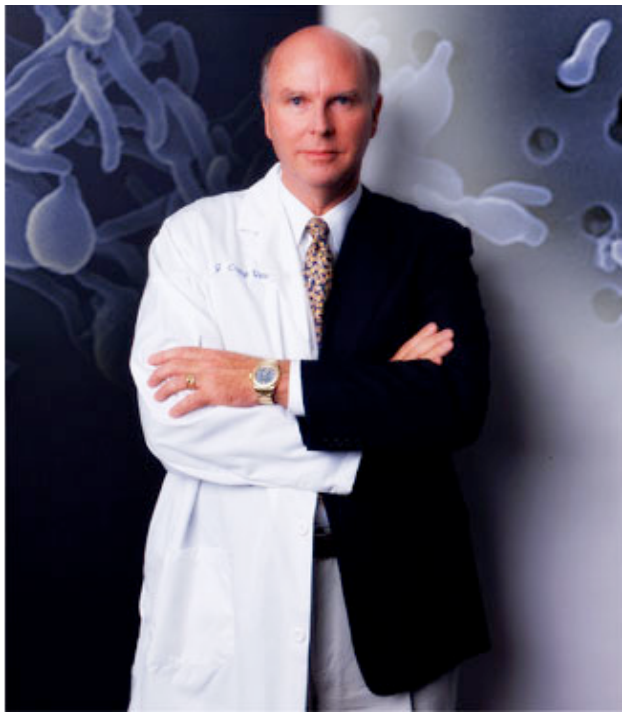
OPEN ACCESS Freely available online PLOS BIOLOGY

The Diploid Genome Sequence of an Individual Human

Samuel Levy[1*], Granger Sutton[1], Pauline C. Ng[1], Lars Feuk[2], Aaron L. Halpern[1], Brian P. Walenz[1], Nelson Axelrod[1], Jiaqi Huang[1], Ewen F. Kirkness[1], Gennady Denisov[1], Yuan Lin[1], Jeffrey R. MacDonald[2], Andy Wing Chun Pang[2], Mary Shago[2], Timothy B. Stockwell[1], Alexia Tsiamouri[1], Vineet Bafna[3], Vikas Bansal[3], Saul A. Kravitz[1], Dana A. Busam[1], Karen Y. Beeson[1], Tina C. McIntosh[1], Karin A. Remington[1], Josep F. Abril[4], John Gill[1], Jon Borman[1], Yu-Hui Rogers[1], Marvin E. Frazier[1], Stephen W. Scherer[2], Robert L. Strausberg[1], J. Craig Venter[1]

1 J. Craig Venter Institute, Rockville, Maryland, United States of America, 2 Program in Genetics and Genomic Biology, The Hospital for Sick Children, and Molecular and Medical Genetics, University of Toronto, Toronto, Ontario, Canada, 3 Department of Computer Science and Engineering, University of California San Diego, La Jolla, California, United States of America, 4 Genetics Department, Facultat de Biologia, Universitat de Barcelona, Barcelona, Catalonia, Spain

PLoS Biol. 2007 4;5:e254.



4.1 million DNA variants (~12.3 Mb)

1.288.319 (~30%) novel!

- 3,213,401 single nucleotide polymorphisms (SNPs)
- 53,823 block substitutions (2–206 bp)
- 559,473 homozygous indels (1–82,711 bp)
- 62 CNVs
- 292,102 heterozygous insertion/deletion events (indels)(1–571 bp)
- 90 inversions
- **Non-SNP DNA variation accounts for 22% of all events identified in the donor, however they involve 74% of all variant bases**

We have uncovered only the tip of the iceberg…

…each genomic region contributes a modest effect, and collectively all associated region for a given trait explain only a small fraction (5-10%) of the observed phenotypic variation attributed to genetic elements…

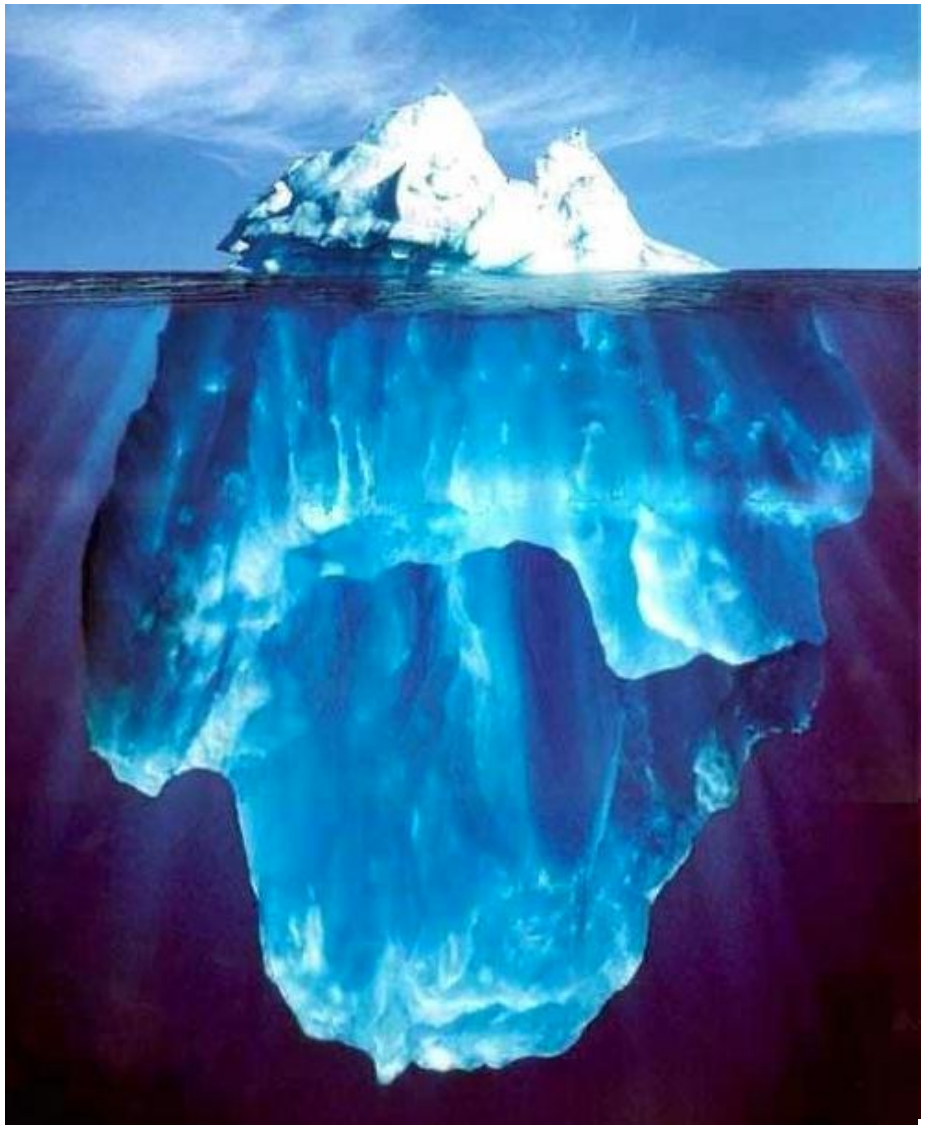
Outcome of a Genome wide association study

We have uncovered only the tip of the iceberg…

…each genomic region contributes a modest effect, and collectively all associated region for a given trait explain only a small fraction (5-10%) of the observed phenotypic variation attributed to genetic elements…

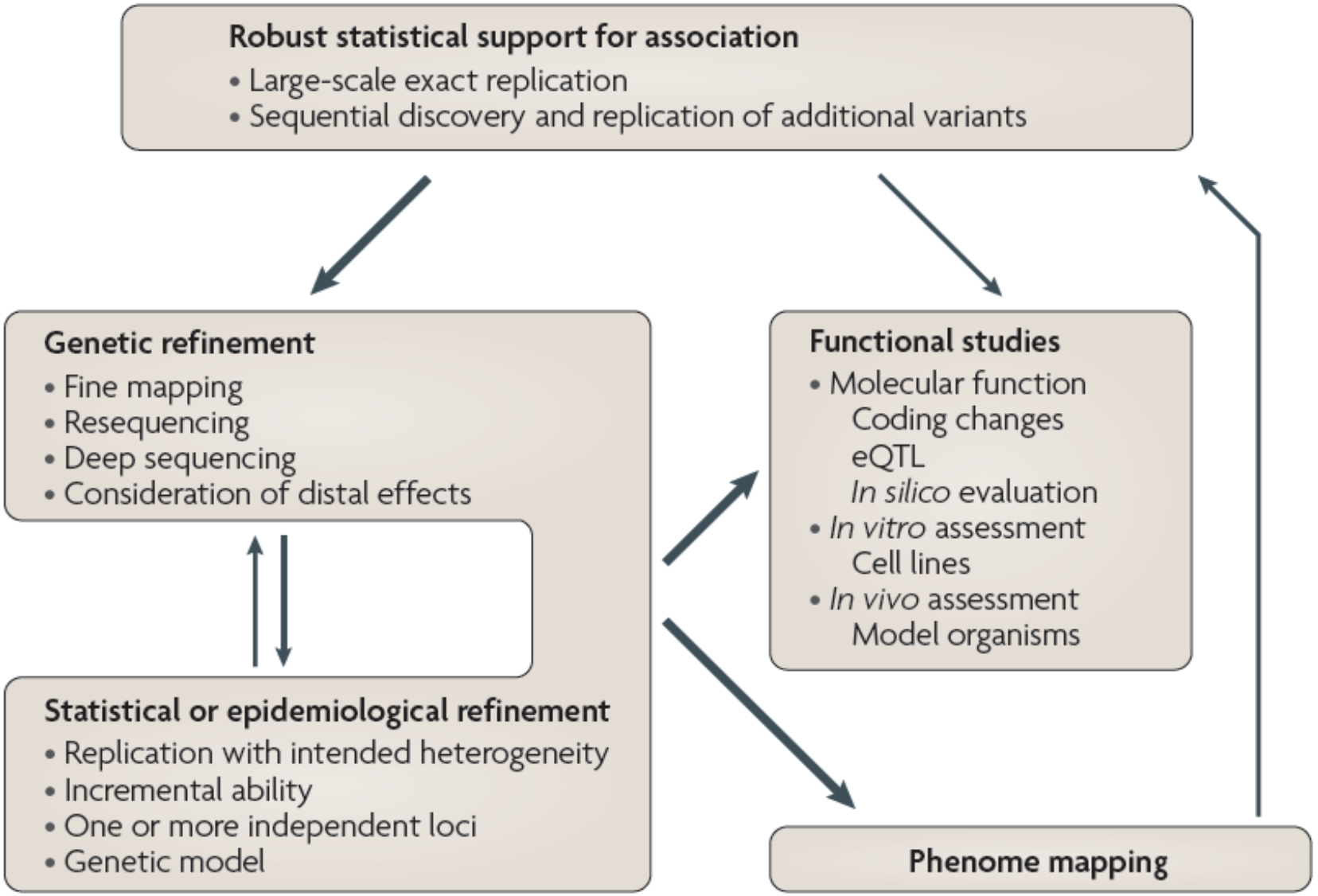
Where is the rest of the missing heritability?

- Incomplete marker coverage
- Allelic heterogeneity at a given locus
- Contribution of rare variants, including structural (CNVs and smaller) variants
- Epistatic interactions
- GxE interactions
- Epigenetic modifications
- Overestimation of heritability



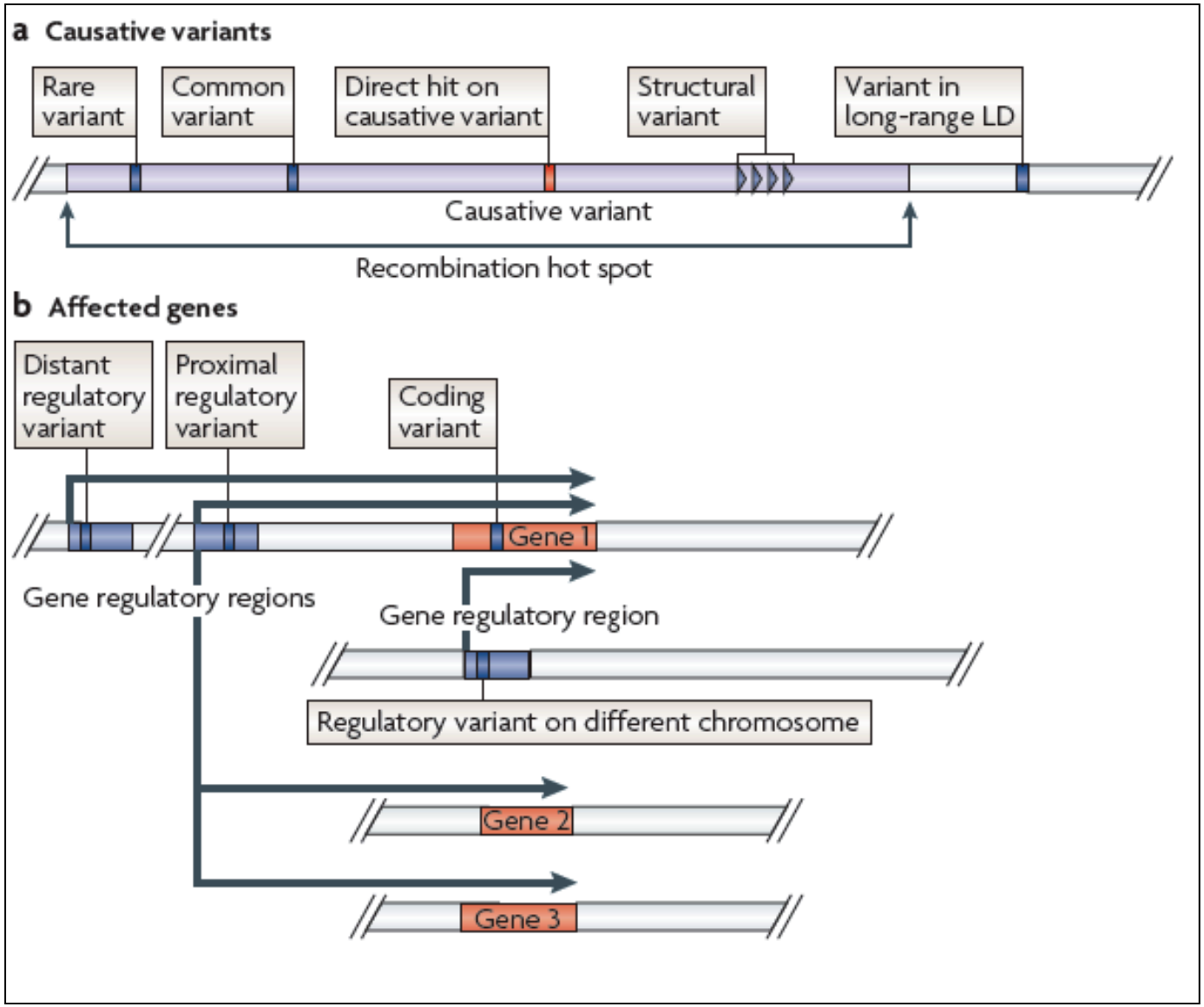
Outcome of a Genome wide association study

Once a gene has been mapped: understanding the function …



Outcome of a Genome wide association study

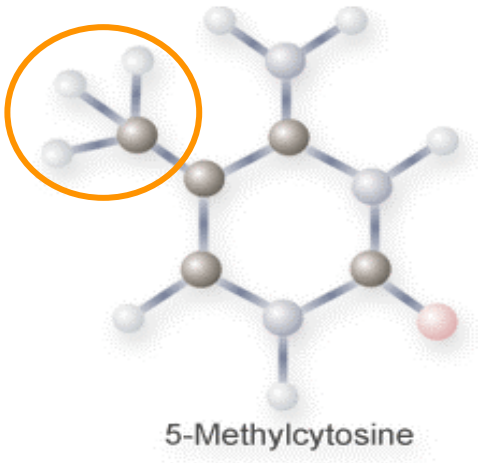
Where is the culprit?
Functional variants and affected genes



Ioannidis et al., 2009 Nature Rev Genet. 10:318-29.

DNA

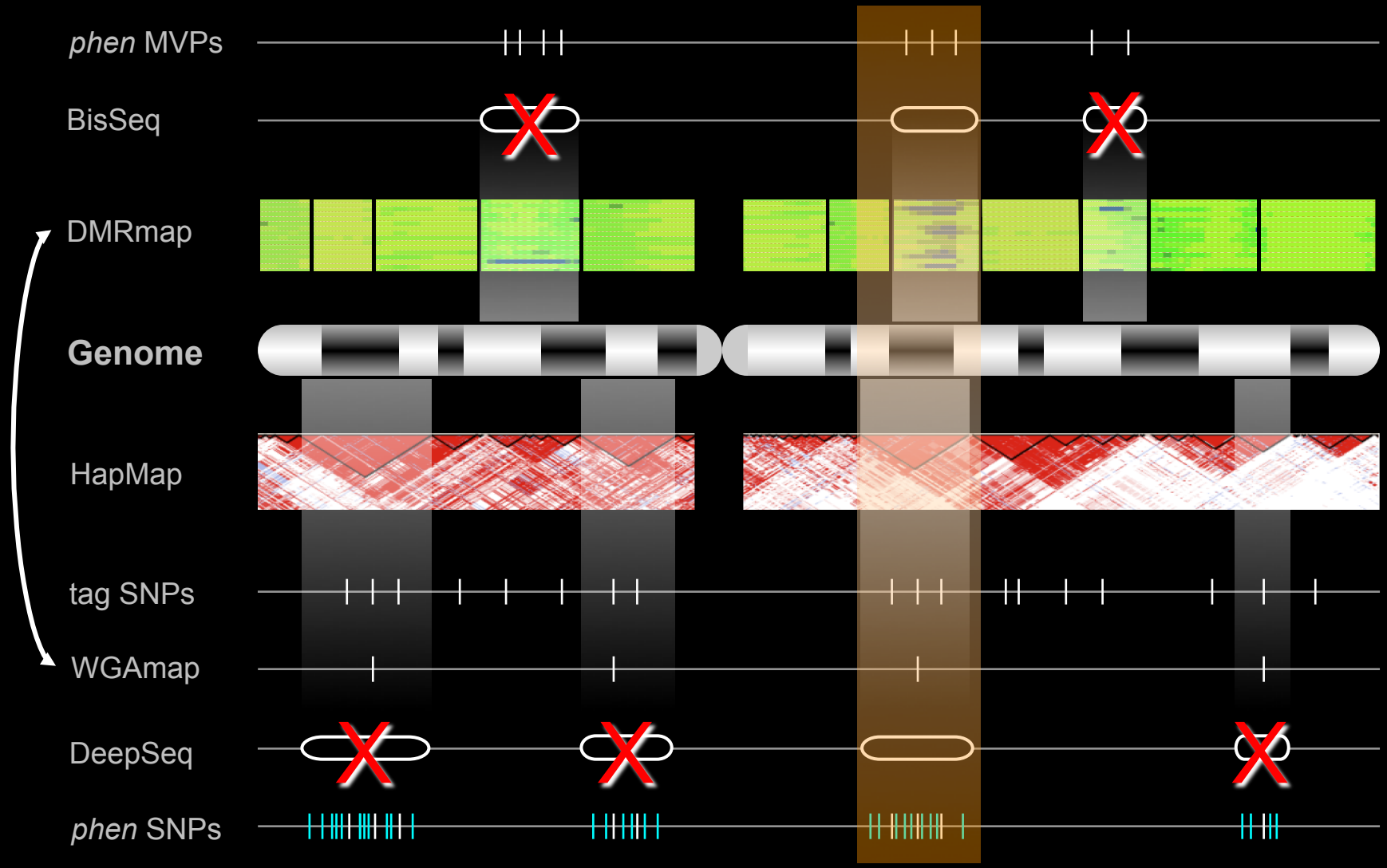
chromat



ATTCGGTCTTACCGATAT

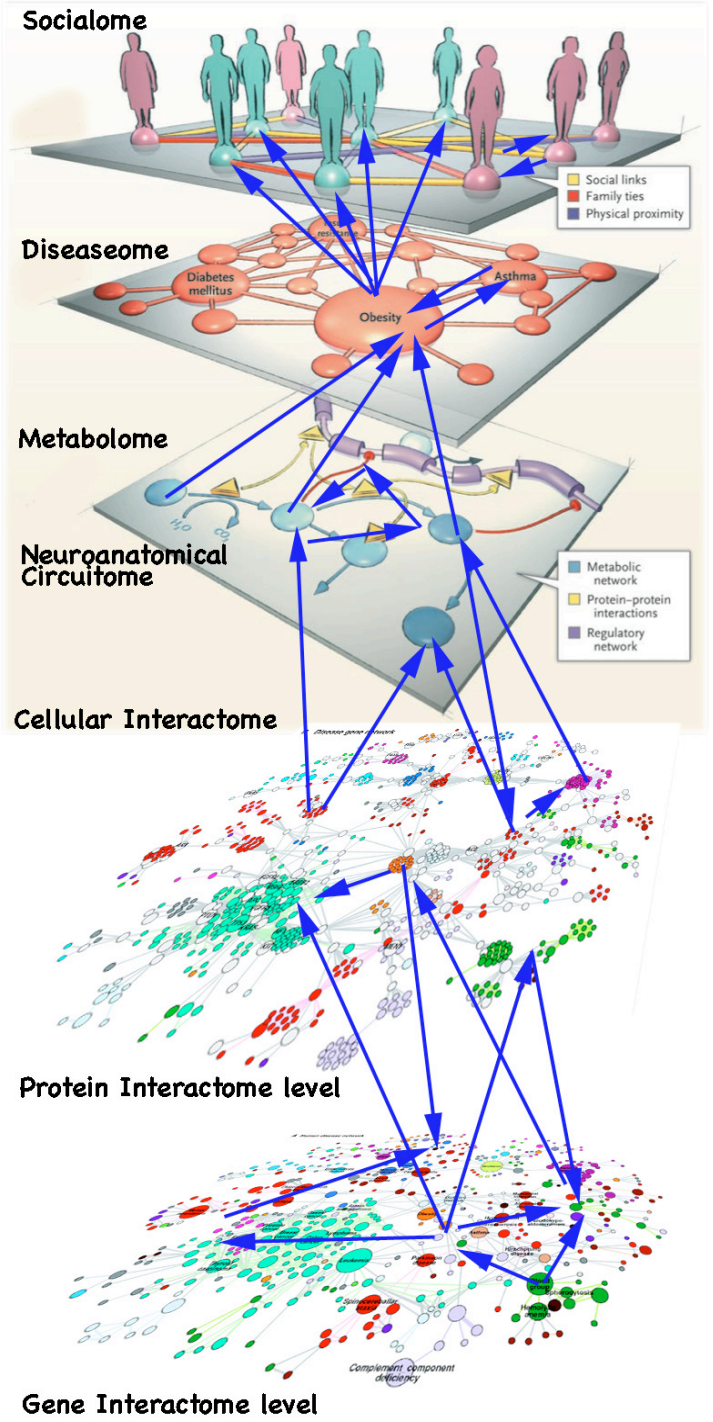
From S. Beck 2008

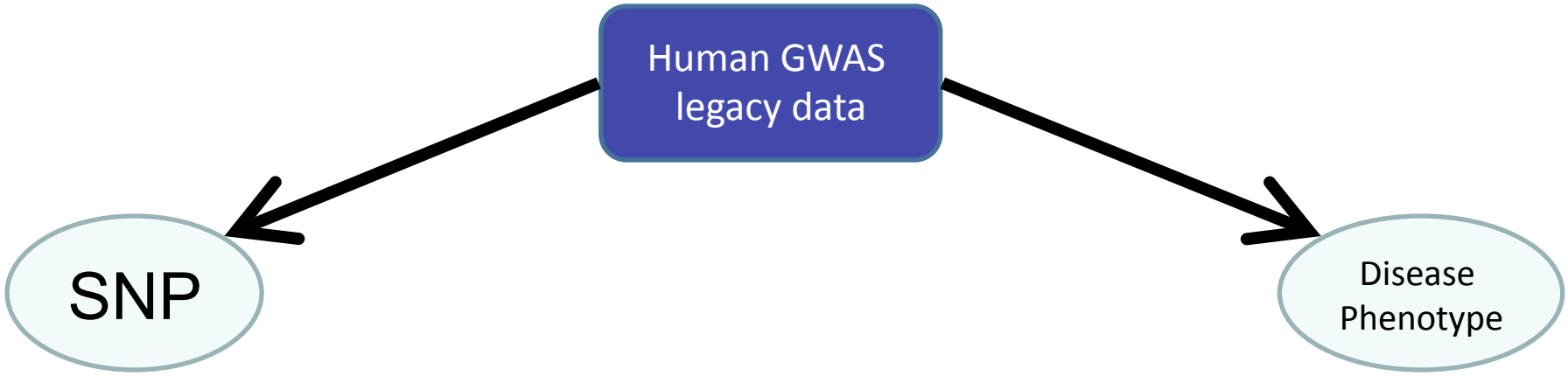
Integrated genomic approach



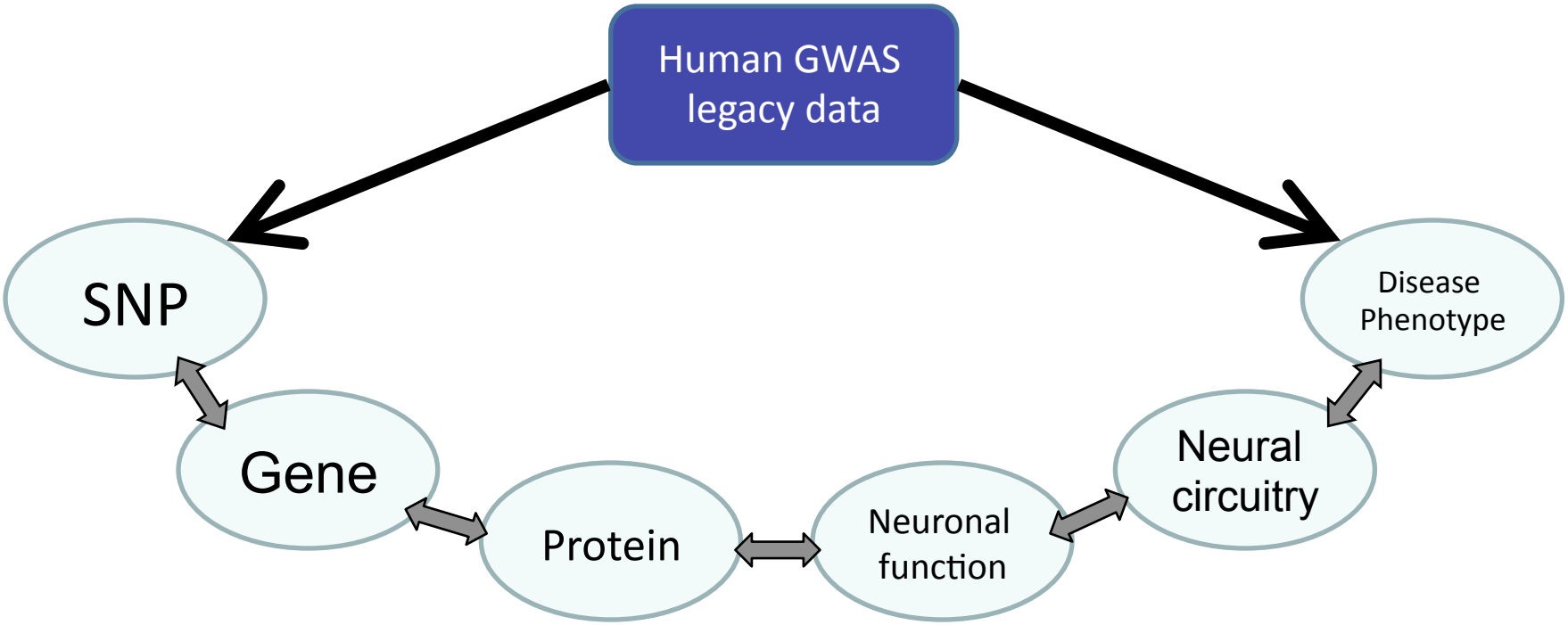
From S. Beck 2008

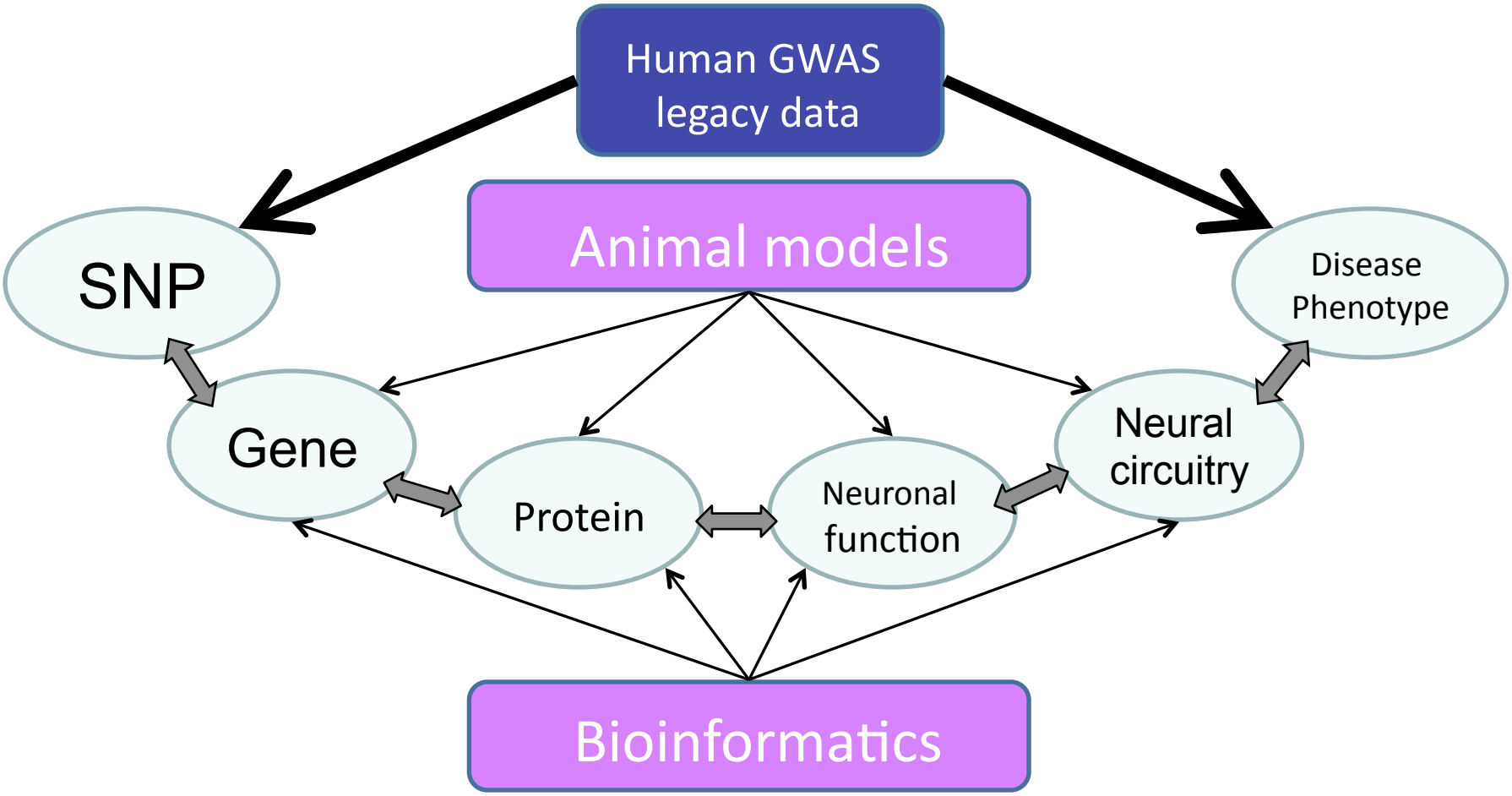
From Genes to Pathways: toward a systemic understanding of disease



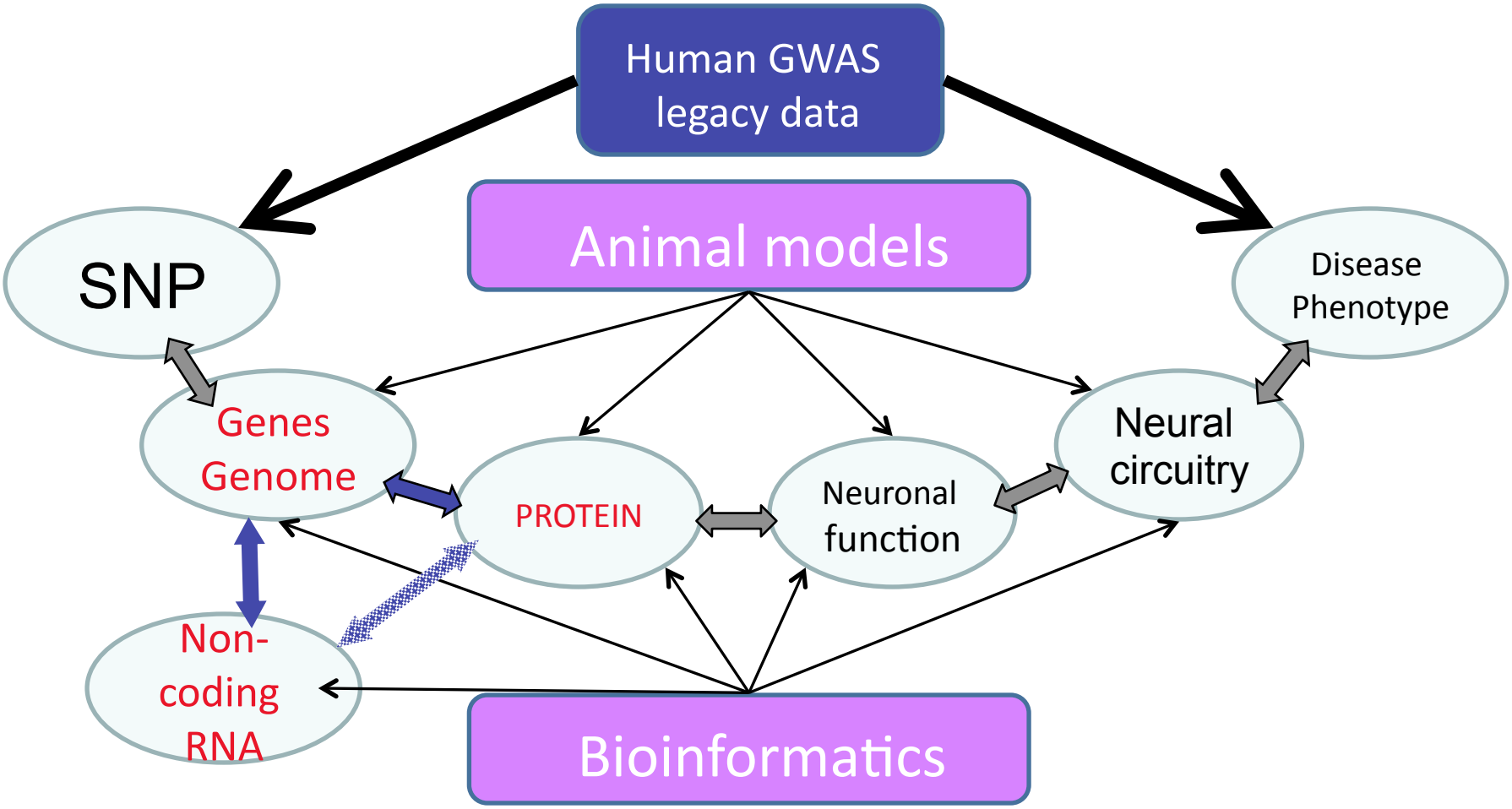


In Genome-Wide Association Studies (GWAS) our goal is to find out the relationship between a Single Nucleotide Polymorphism (SNP, as a proxy for a gene) and the Disease Phenotype of interest


Human GWAS
legacy data
SNP
Gene
Protein
Neuronal
function
Neural
circuitry
Disease
Phenotype

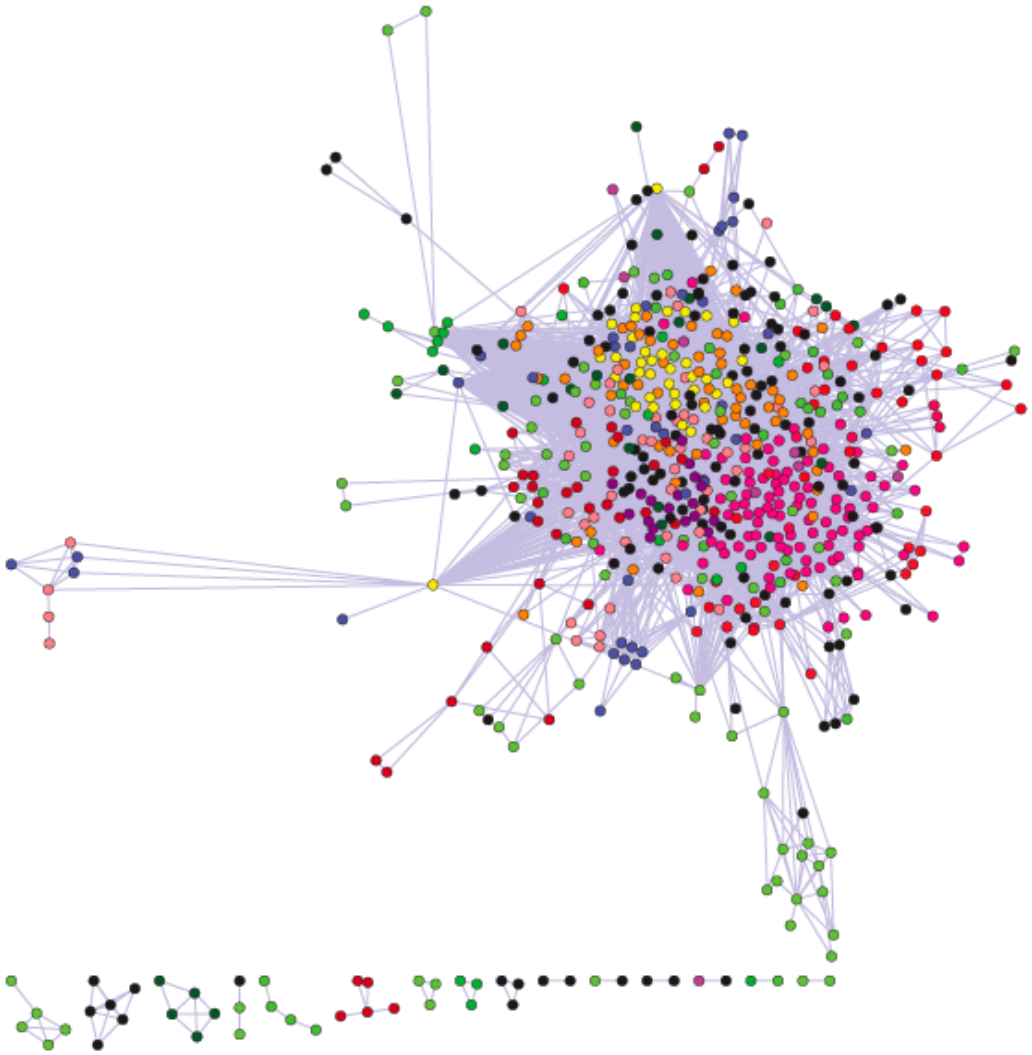


Systems Biology addresses links between SNPs and human phenotype originally identified by GWAS .....

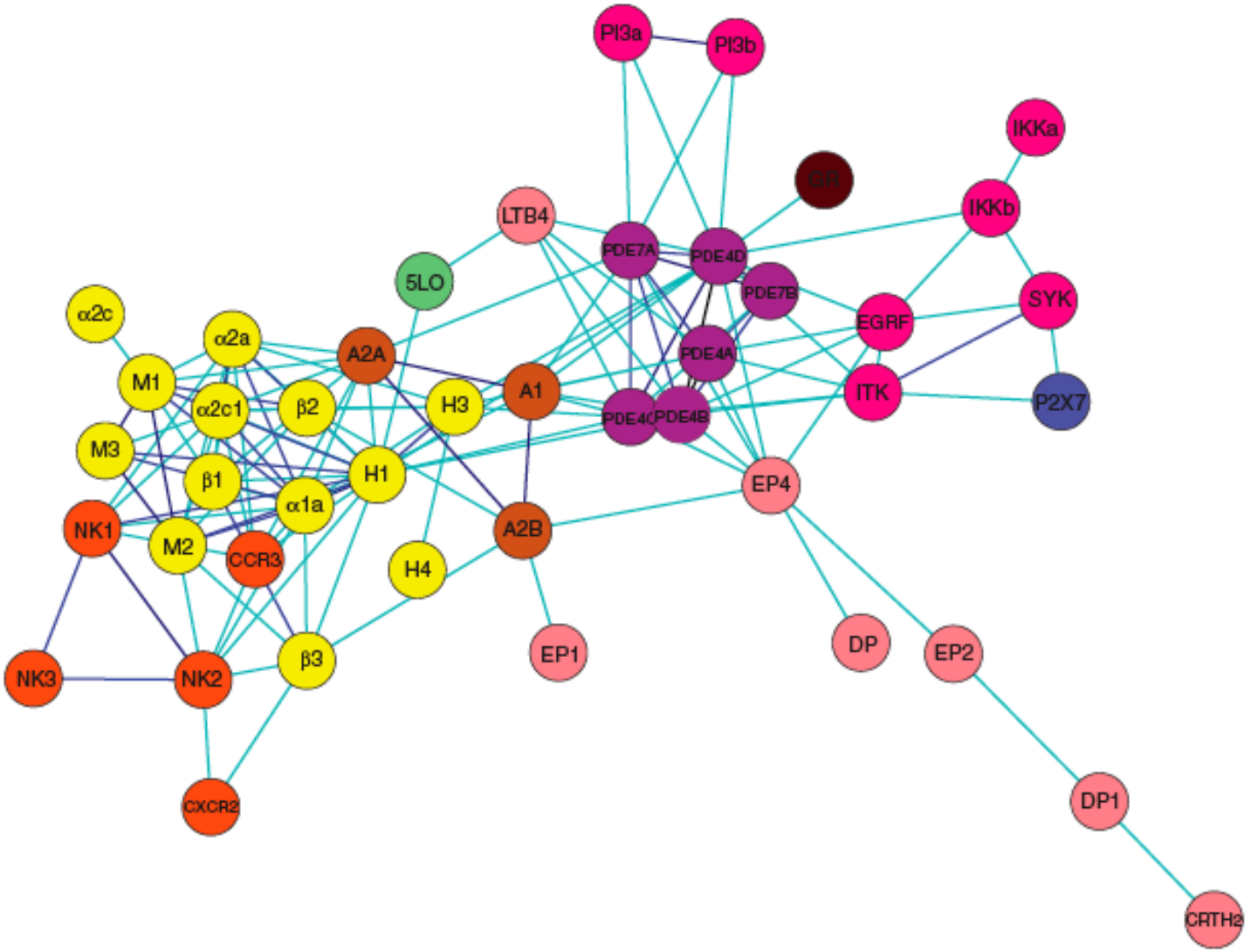


….. including information related to *WHOLE* genomic complexity

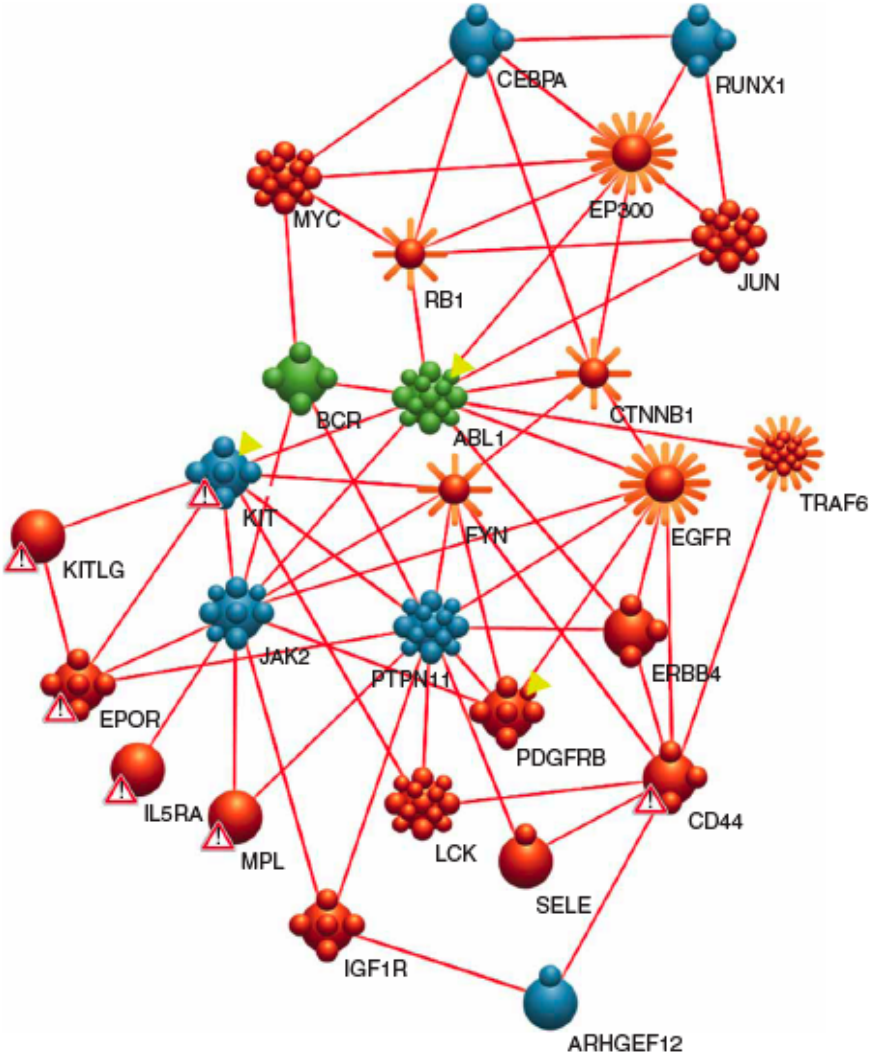
From single genes to networks



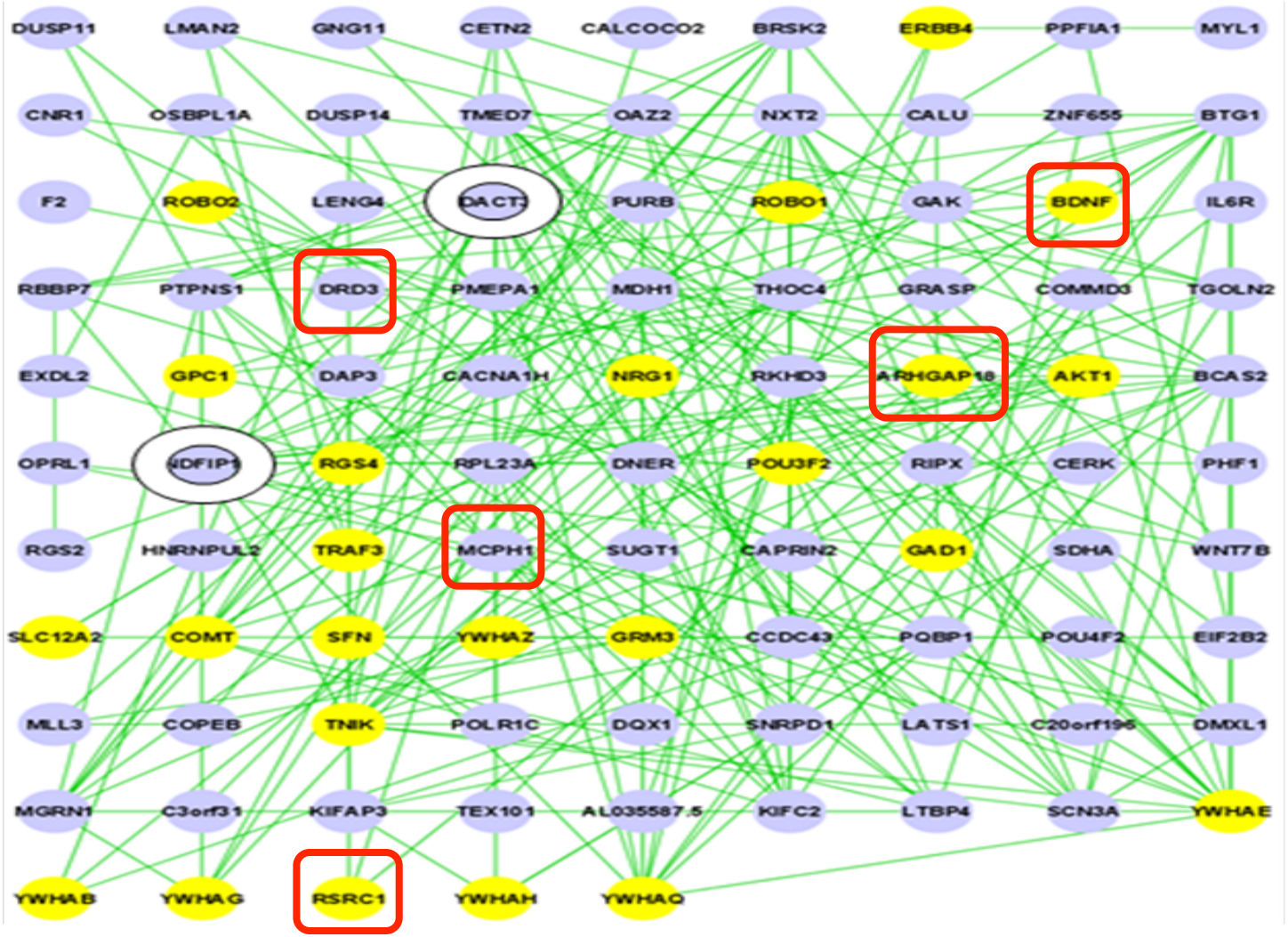
Genes associated with asthma



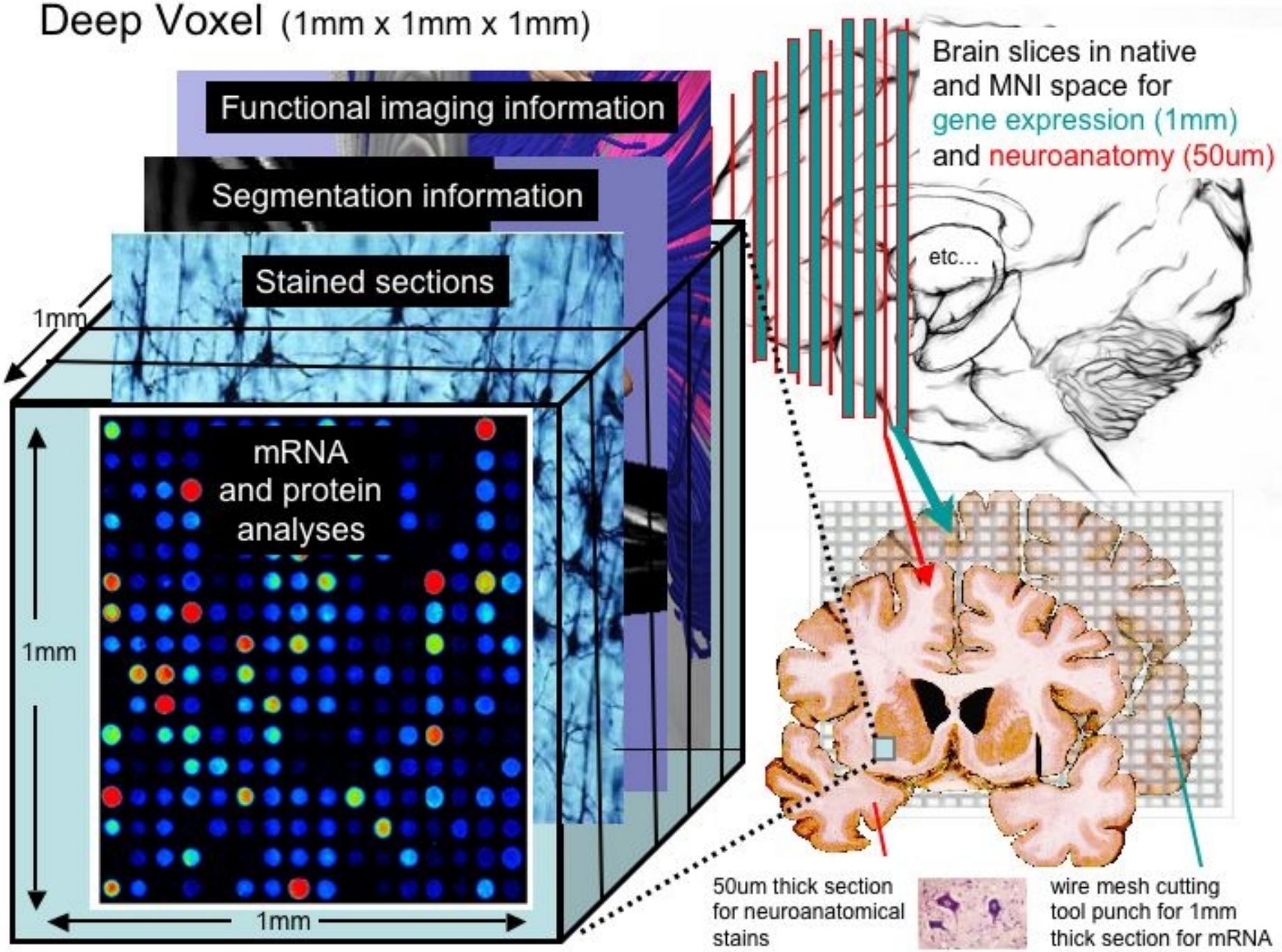
Leukemia disease network



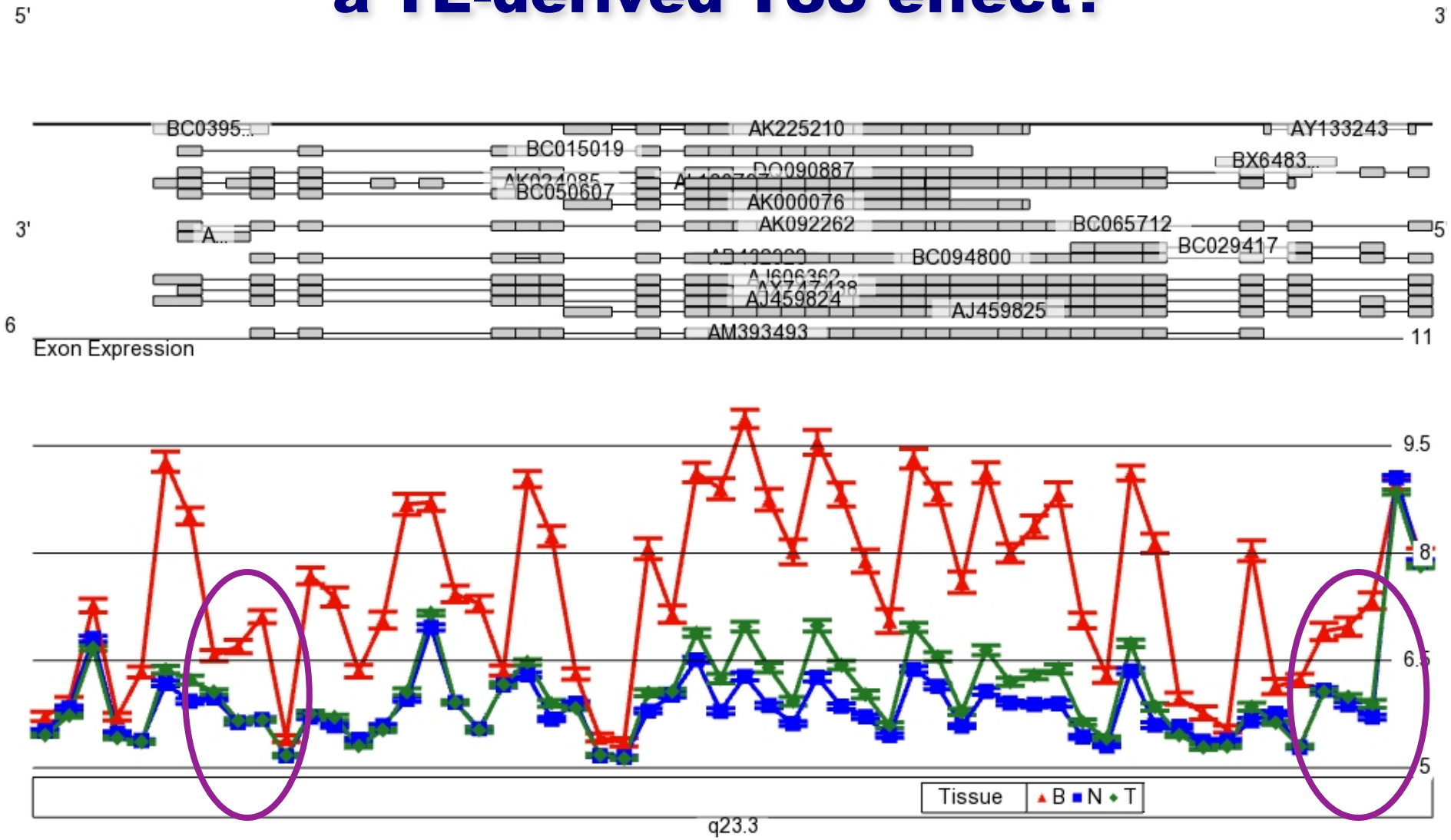
Inferred Networks based on mouse PFC expression data



Gene interaction network inferred from prefrontal cortex gene expression in 42 different inbred mouse strains. Schizophrenia candidate genes from GWAS are in yellow. Some unexpected connections: DACT3 (circled), encodes regulator of Wnt signaling that has been linking to schizophrenia[41-43].


Deep Voxel (1mm x 1mm x 1mm)
Functional imaging information
Segmentation information
Stained sections
mRNA
and protein
analyses
1mm
1mm
1mm
Brain slices in native
and MNI space for
gene expression (1mm)
and neuroanatomy (50um)
etc...
50um thick section
for neuroanatomical
stains
wire mesh cutting
tool punch for 1mm
thick section for mRNA

AHI1 exon expression in brain, LCL and immortalized cell-lines (187 subjects): a TE-derived TSS effect?



Sequencing the genome of schizophrenic patients