Nutrigenetics- Concepts of genetic studies in Nutrition

Ecole Polytechnique de Lausanne

23rd September 2025

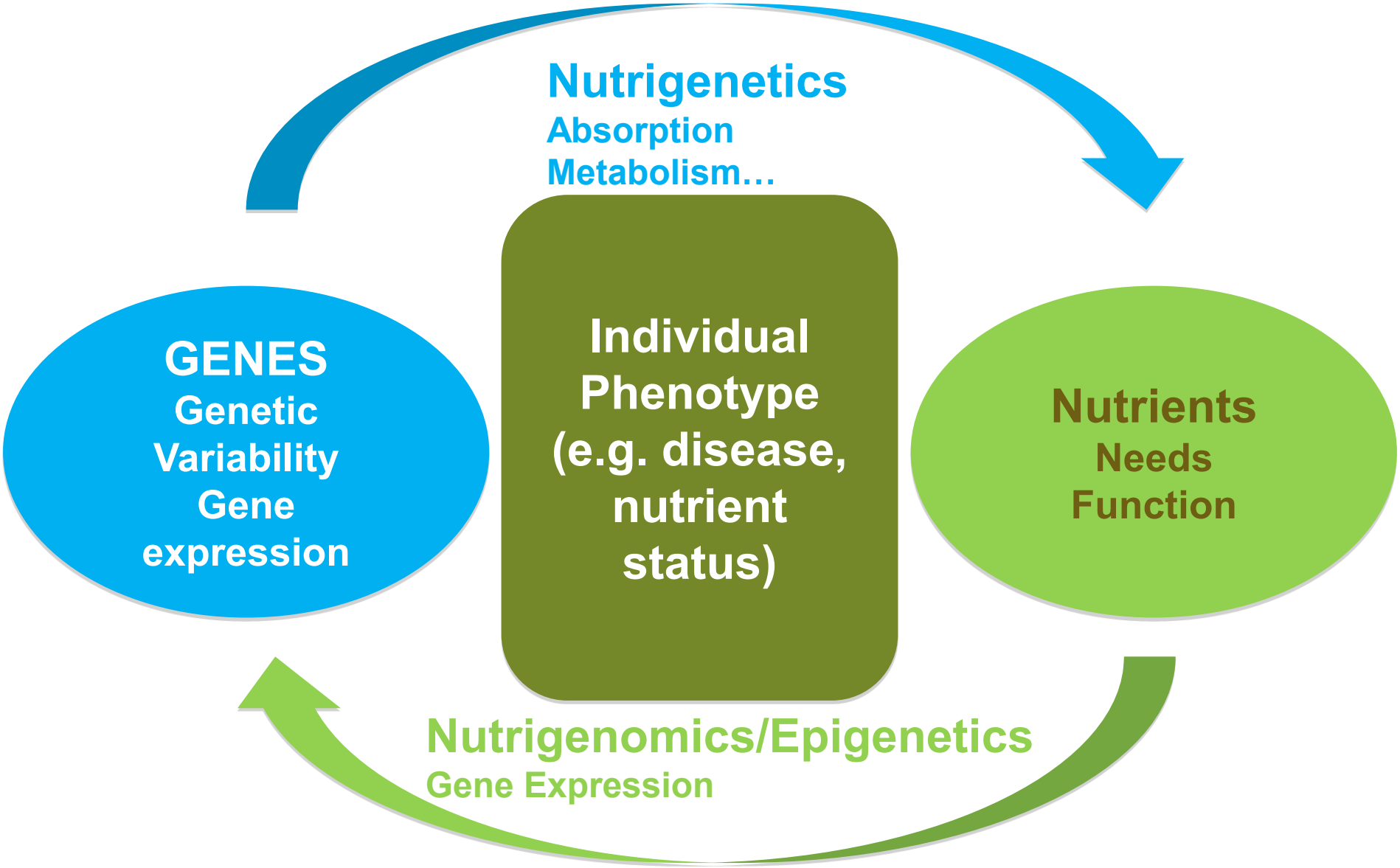
Jörg Hager PhD
Jorg.Hager@RD.Nestle.com

NutriGenetics

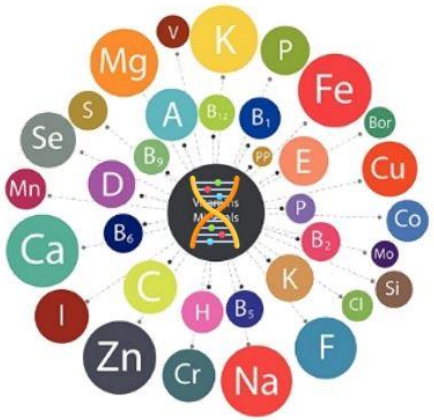
- Nutrigenetics identifies how genetic variability (i.e. differences in our DNA sequence) affect the response to and the bio-availability or activity of dietary nutrients.

NutriGenomics: A Definition

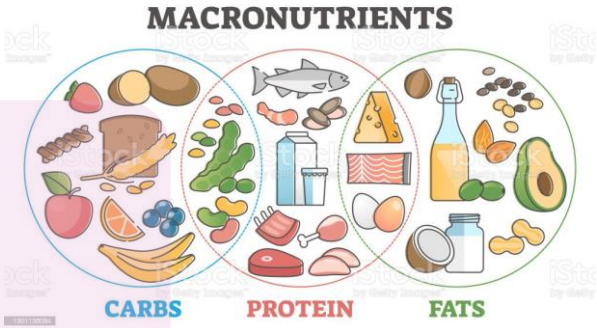
- **Nutrigenomics**
  - The study that investigates
  - How your nutrition affects the expression of your genes and vice versa
  - How subsequently your genes control your health


Nutrigenetics
Absorption
Metabolism…
GENES
Genetic
Variability
Gene
expression
Individual
Phenotype
(e.g. disease,
nutrient
status)
Nutrients
Needs
Function
Nutrigenomics/Epigenetics
Gene Expression

Genetic factors may influence nutrients on multiple levels



Micronutrients:
- Absorption
- Activity
- Clearance



Macronutrients:
- Digestion
- Transport
- Clearance

Bioactives:
- Absorption
- (In-) Activation
- Clearance

Nogueira et al. Polymers, 2020

Factors contributing to variable (micro-) nutrient levels at similar intake levels.

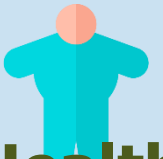
Genetic

Food intake/matrix

Micro-biome

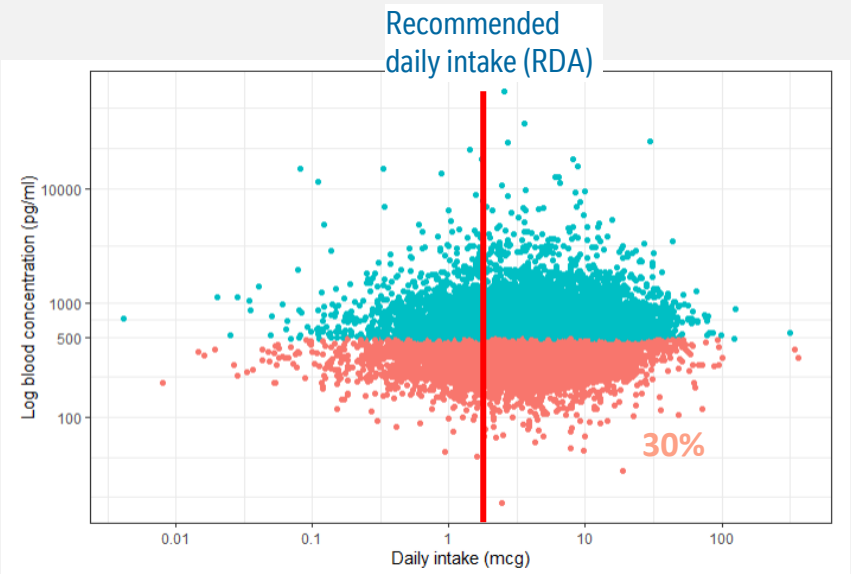
Life style

Health e.g.Obesity



Genetic factors can influence nutrient levels (at identical intake levels).

VITAMIN B12 BLOOD LEVELS CORRELATE POORLY WITH INTAKE



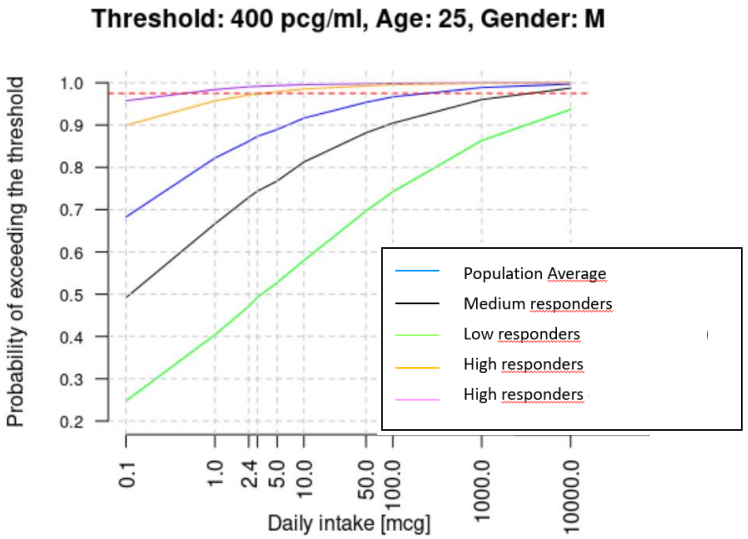
**NHANES nutrition survey 2012-018*

BLOOD LEVELS OF B12 ARE INFLUENCED BY MANY DIFFERENT GENETIC VARIANTS (SNPS) SYNERGISTICALLY

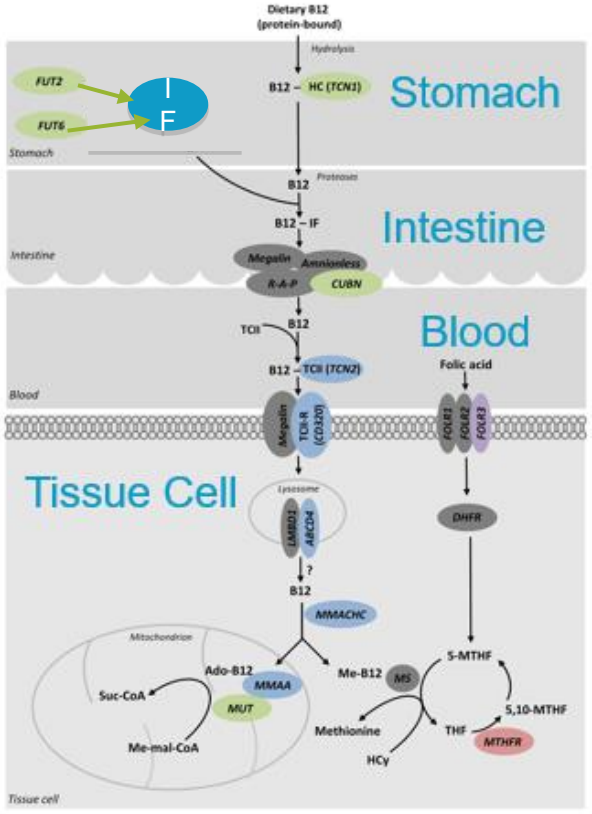
One phenotype (e.g. type 2 diabetes, height, vitamin B12 levels), many SNPs

Vitamin B12 1 Variant 4 pg/ml decrease
26 Variants **104** pg/ml decrease

Gene variants affect key proteins involved in vitamin B12 transport and absorption

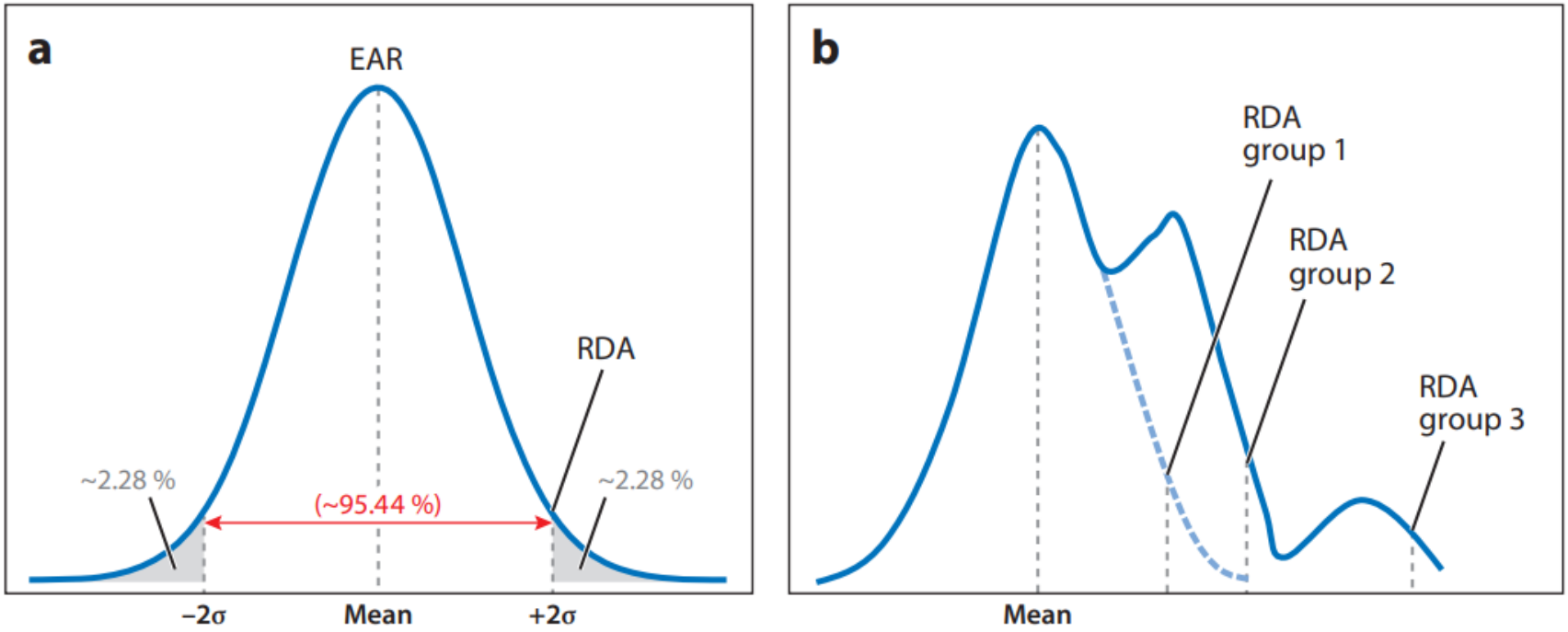


- Low responders do not reach sufficiency, even at extreme high (single) doses of B12;



Source: Grarup et al. Genetic architecture of vitamin B12 and folate levels. PLOS Genetics, 2013

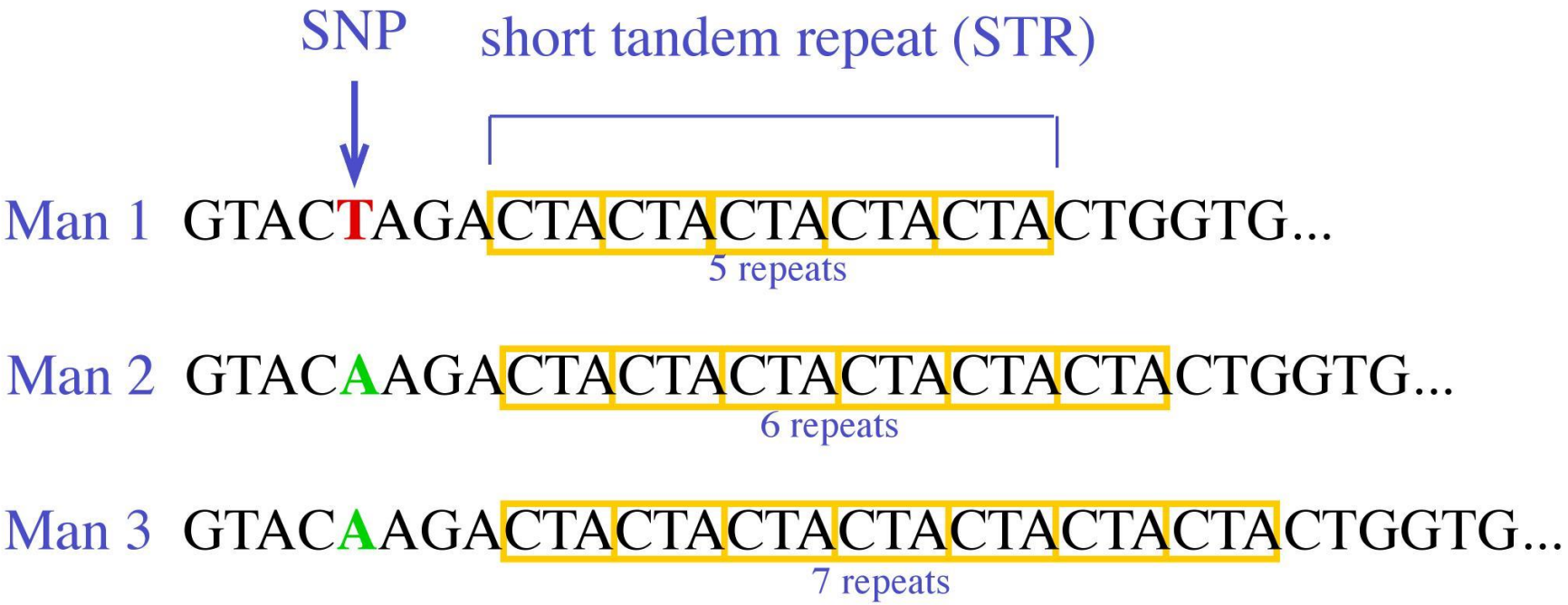
Genetic factors contribute to define our individual nutrient intake needs to reach sufficient (blood-) levels.



Borel & Desmarchelier. Annual Review of Nutrition. 2018

Genetic markers to identify genetic variation influencing phenotypic traits

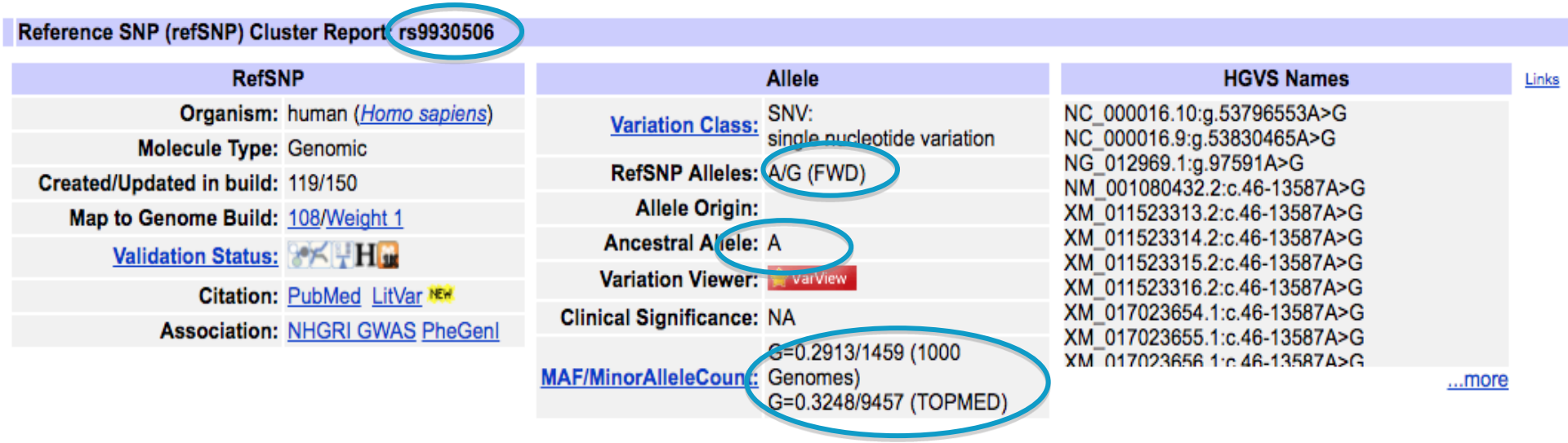
A genetic marker is a variable site in the genome characterized by it's nucleotide composition, position and frequency



- Very abundant
- Easy to analyze
- Low informativity (max. 50% Hz)

- Very informative (high heterozygosity)
- Difficult to analyze
- Sparsely distributed

Characteristics of Single Nucleotide Polymorphisms



Reference SNP (refSNP) Cluster Report: rs9930506

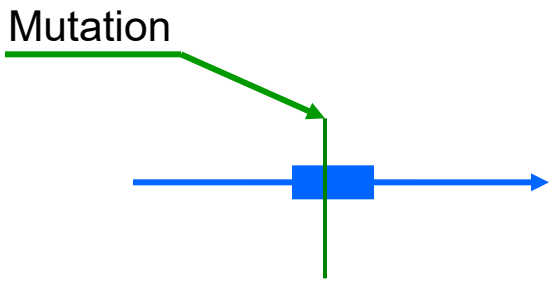
RefSNP	
Organism:	human (*Homo sapiens*)
Molecule Type:	Genomic
Created/Updated in build:	119/150
Map to Genome Build:	108/Weight 1
Validation Status:	
Citation:	PubMed LitVar NEW
Association:	NHGRI GWAS PheGenI

Allele	
Variation Class:	SNV: single nucleotide variation
RefSNP Alleles:	A/G (FWD)
Allele Origin:	
Ancestral Allele:	A
Variation Viewer:	varView
Clinical Significance:	NA
MAF/MinorAlleleCount:	G=0.2913/1459 (1000 Genomes) G=0.3248/9457 (TOPMED)

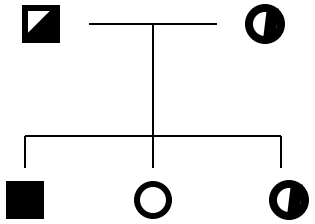
HGVS Names	Links
NC_000016.10:g.53796553A>G	
NC_000016.9:g.53830465A>G	
NG_012969.1:g.97591A>G	
NM_001080432.2:c.46-13587A>G	
XM_011523313.2:c.46-13587A>G	
XM_011523314.2:c.46-13587A>G	
XM_011523315.2:c.46-13587A>G	
XM_011523316.2:c.46-13587A>G	
XM_017023654.1:c.46-13587A>G	
XM_017023655.1:c.46-13587A>G	
XM_017023656.1:c.46-13587A>G	
...more	

https://www.ncbi.nlm.nih.gov/snp/

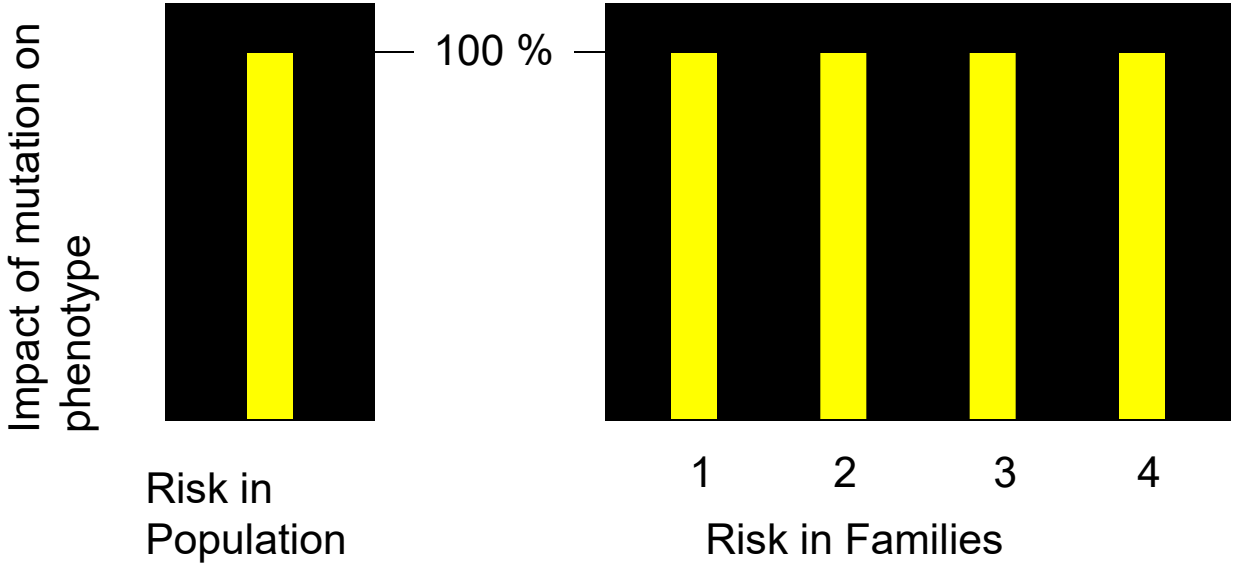
Monogenic Trait (phenotype)



Loss of function mutation
- base pair change

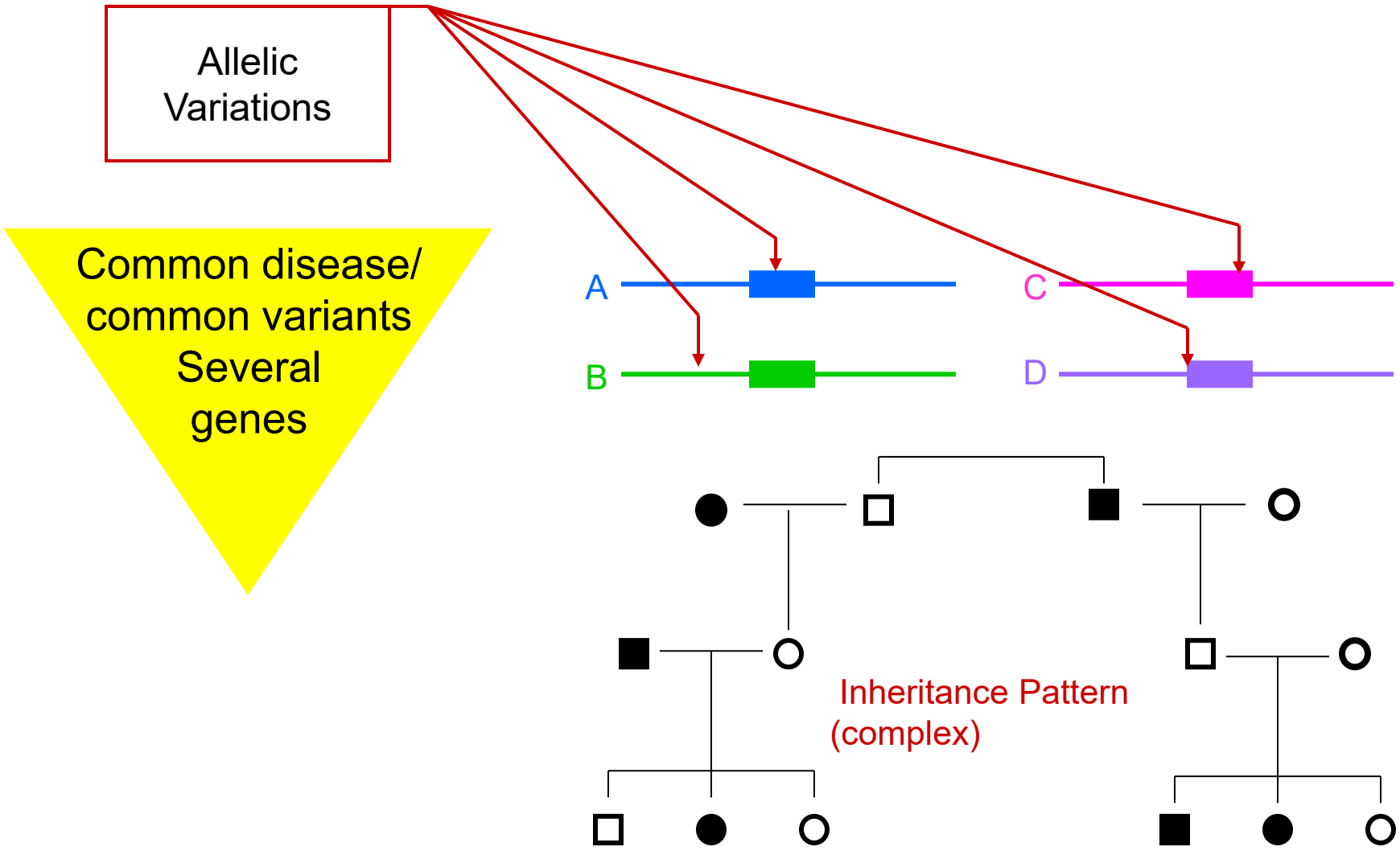


Linkage studies ID'd genes



Peltonen and McKusick, Dissecting Human Disease in the Postgenomic Era
Science Feb 16 2001: 1224-1229

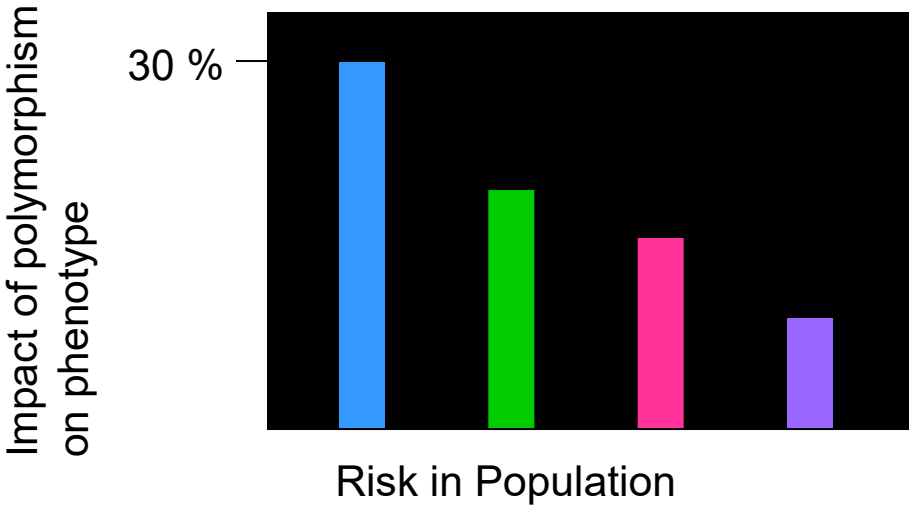
Polygenic Trait

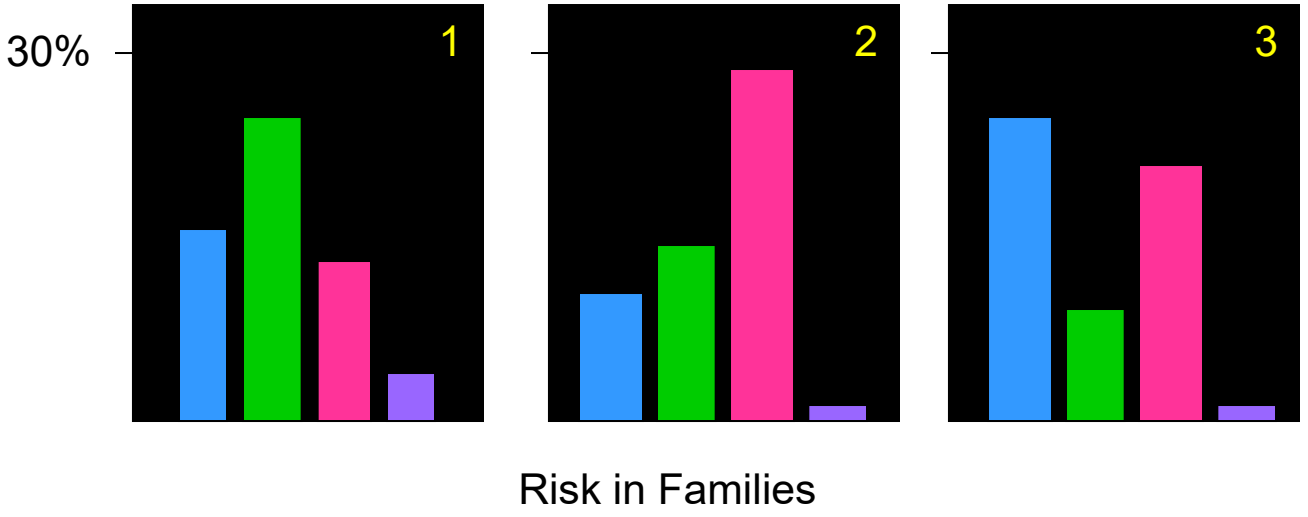


Peltonen and McKusick, Dissecting Human Disease in the Postgenomic Era *Science* Feb 16 2001: 1224-1229

Polygenic trait

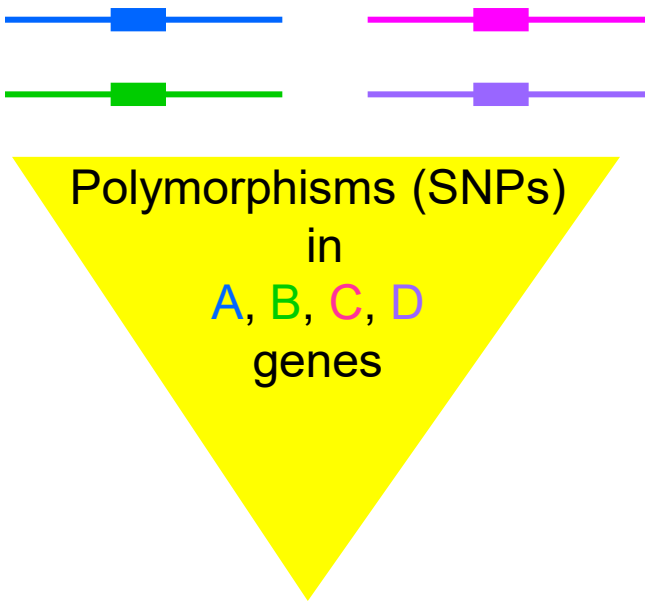
Variants in
A, B, C, D
genes

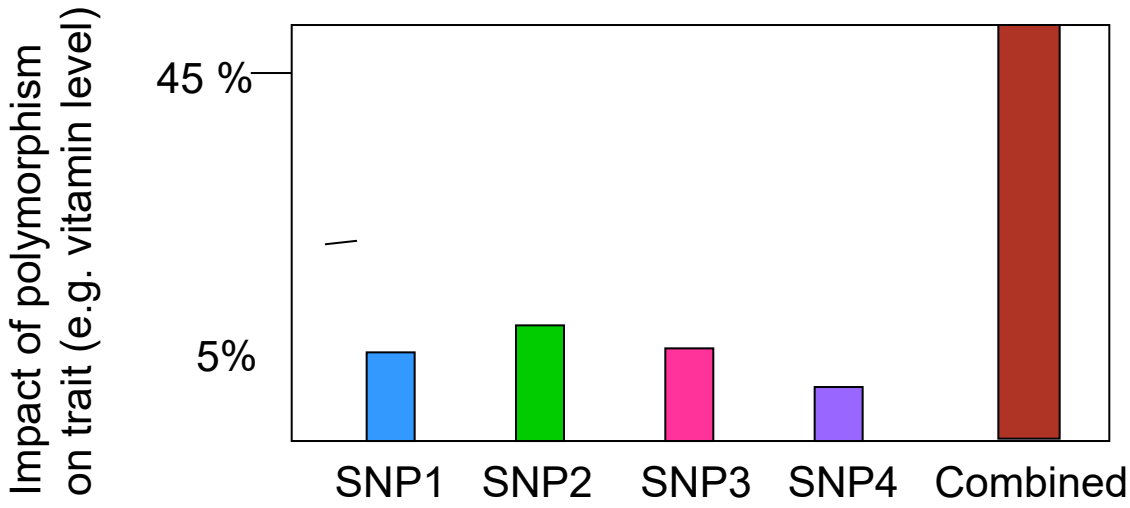




Peltonen and McKusick,
Science Feb 16 2001: 1224-1229

Complex Disorder

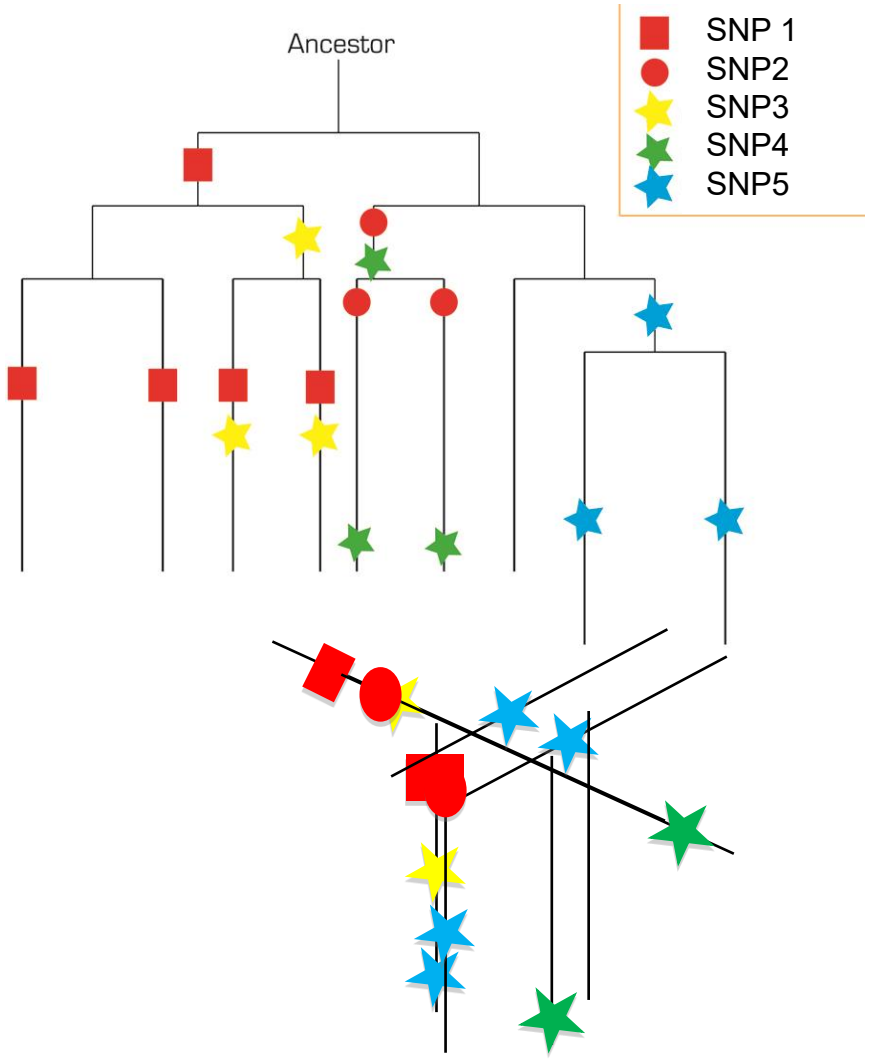


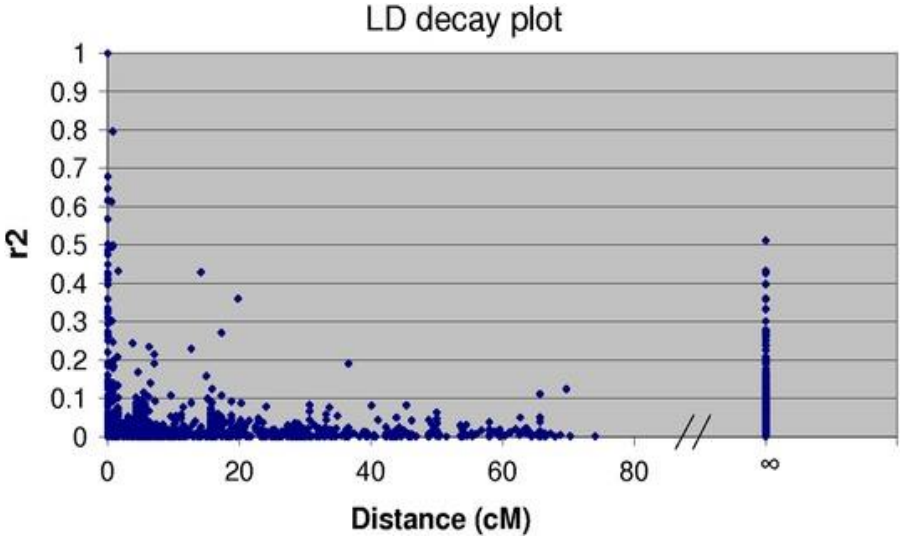


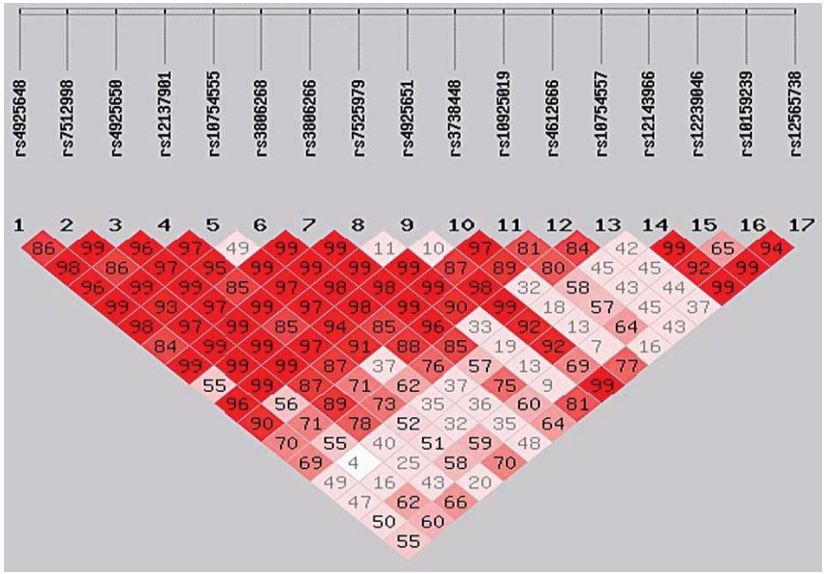
To summarize:

Most nutritional phenotypes are complex, due to polymorphisms in many different genes (loci).
The impact of each polymorphism on the phenotype is usually small but the combined impact can be large.

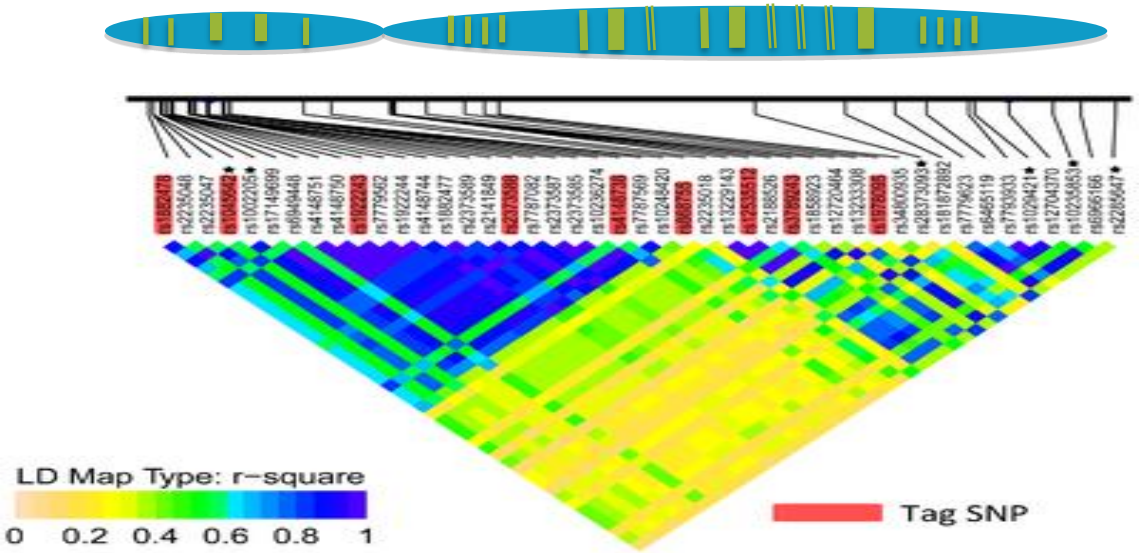
Linkage Disequilibrium







Genotyping arrays for whole genome association studies – tag SNP selection & imputation



LD maps allow to:

a) genotype only a limited number of tag SNP's to capture most information;

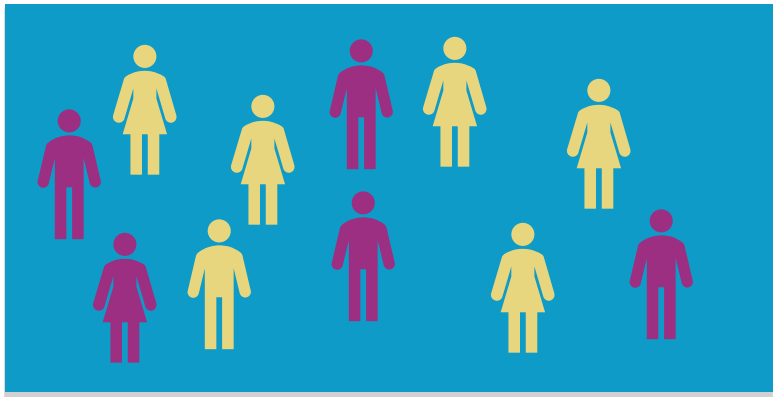
b) To impute genotypes of unobserved SNPs

SNP 1: C – T (65 – 35)
SNP 2: A – G (62 – 38)
R^2= 0.87

P (A| C or T)
P (G| C or T)

Dichotomous (qualitative) trait association studies

People **with** Disease (cases)



People **without** Disease (controls)

Genotype	With Disease	Without disease	Total
Allele- C	6	2	8
Allele- T	5	9	14
	11	11	

$X^2 = 5.38$

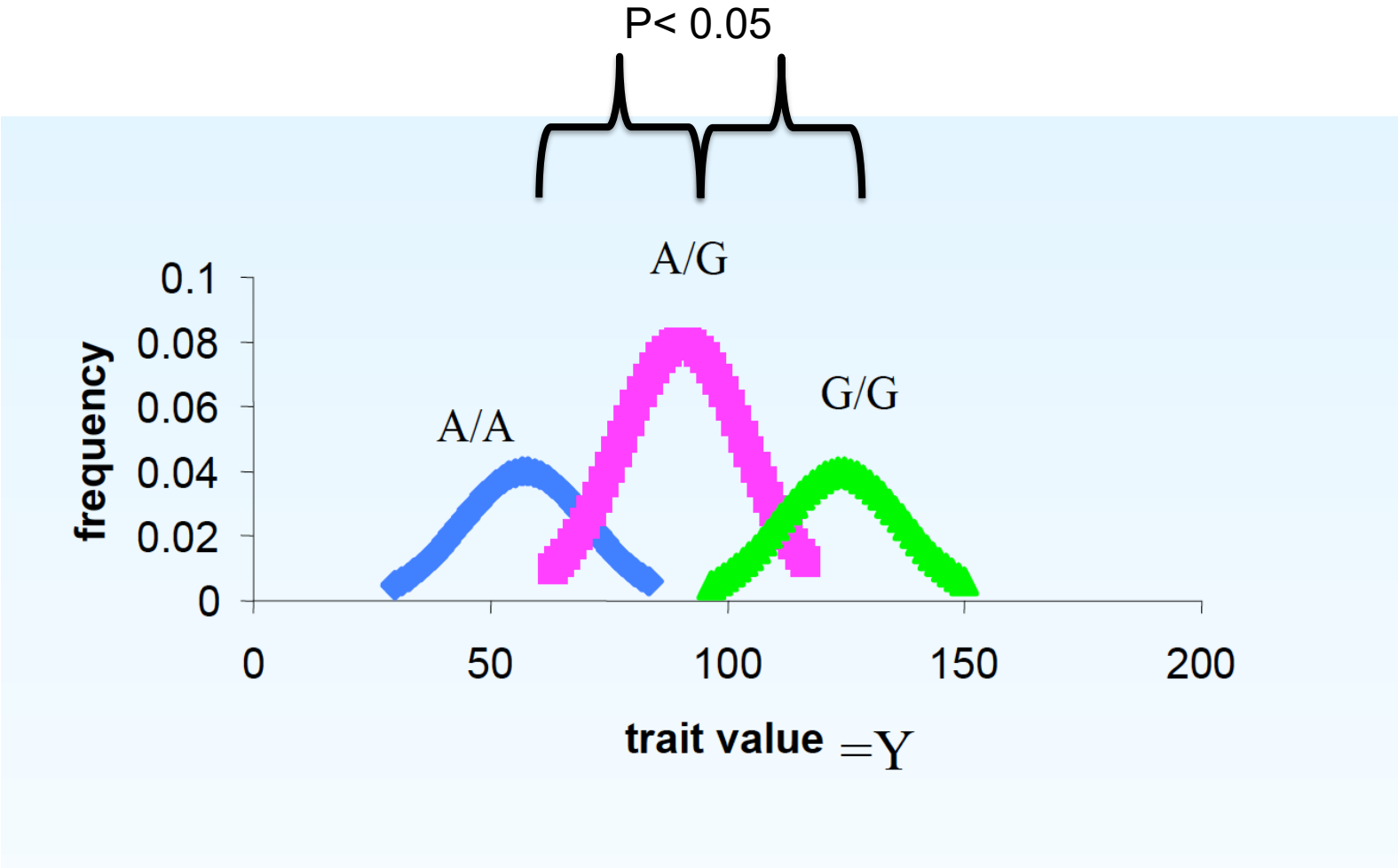
$p < 0.025$

…in reality we will use methods like logistic regression that allow to adjust for co-variates (e.g. sex, age, genetic background)

Quantitative trait association analyses: What is a quantitative trait?

- A quantitative trait has numerical values that can be ordered from lowest to highest;
- The values can be discrete or (more often) can be continuous. Most genetic association studies use continuous variables;
- Examples: cholesterol levels, blood pressure, weight, height, glucose levels, vitamin levels…

The basics: Quantitative trait values differ with genotype



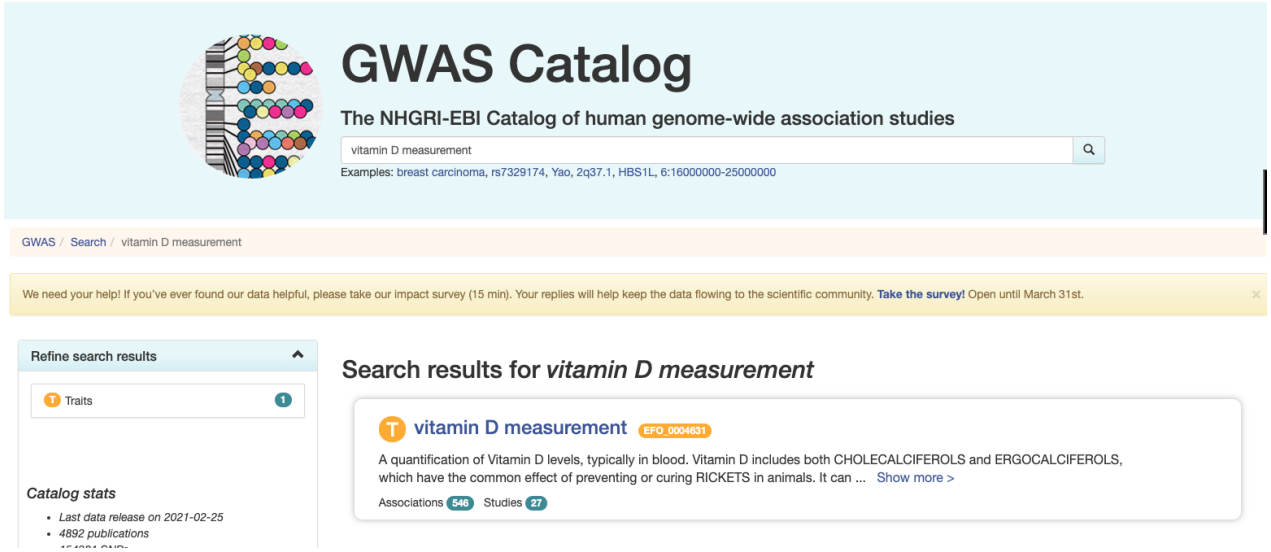
From: Janet Sinsheimer; Human Genetics, Biomathematics and Statistics

Quantitative trait association analyses: Why use them?

- More power than dichotomized traits i.e. fewer individuals are needed for the analysis;
- Genotype – phenotype relationship may be more direct (higher effect sizes) i.e. a disease may be the product of multiple affections like gene – environment interactions that may differ between people resulting in heterogeneity;
- **In nutrition we're mostly interested in quantitative outcomes on nutrients (e.g. vitamin levels).**

Genome Wide Association Studies (GWAS)

- In reality we will not test just one or two genetic variants(SNPs) for an asociation but will test hundreds of thousands of SNPs in a large number of subjects (thousands) in parallel.
- The most comprehensive data base of huamn GWAS studies is the GWAS catalog (www.ebi.ac.uk/gwas/)



Looking at GWAS data

First author	Study accession	Publication date	Journal	Title	Reported trait	Trait(s)	Discovery sample number and ancestry	Replication sample number and ancestry	Ass
Engelman CD	GCST000711	2010-06-26	J Steroid Biochem Mol Biol	Genome-wide association study of vitamin D concentrations in Hispanic Americans: the IRAS family study.	Vitamin D levels	vitamin D measurement	• 229 Hispanic or Latin American	• 961 Hispanic or Latin American	0
Traglia M	GCST010171	2020-02-11	Genetics	Genetic Contributions to Maternal and Neonatal Vitamin D Levels.	Midgestational total 25-hydroxyvitamin D levels (maternal genetic effect)	vitamin D measurement	• 323 Hispanic or Latin American • 269 European • 115 East Asian • 23 African American or Afro-Caribbean • 23 South Asian • 23 NR	-	3
Traglia M	GCST010170	2020-02-11	Genetics	Genetic Contributions	Neonatal total 25-hydroxyvitamin D	vitamin D measurement.	• 115 East Asian • 269 European	-	7

Available data: Associations 546 | Studies 27 | LocusZoom

Download Catalog data

Include child trait data

Associations 546

Search

Variant and risk allele	P-value	P-value annotation	RAF	OR	Beta	CI	Mapped gene	Reported trait	Trait(s)	Study accession	Location
rs184291421-C	1×10^{-28}		0.994155	-	0.170103 unit increase	[0.14-0.2]	AC068721.1, NPFFR2	Serum 25-Hydroxyvitamin D levels	vitamin D measurement	GCST010144	4:71887129
rs188838036-A	3×10^{-24}		0.995469	-	0.17944 unit increase	[0.14-0.21]	AC068721.1, NPFFR2	Serum 25-Hydroxyvitamin D levels	vitamin D measurement	GCST010144	4:71917668
rs186881826-A	4×10^{-77}		0.223069	-	0.045907 unit increase	[0.041-0.051]	NPFFR2, AC068721.1	Serum 25-Hydroxyvitamin D levels	vitamin D measurement	GCST010144	4:71920026
rs186441690-G	2×10^{-18}		0.997189	-	0.26711 unit decrease	[0.21-0.33]	NPFFR2, AC068721.1	Serum 25-Hydroxyvitamin D levels	vitamin D measurement	GCST010144	4:71955252

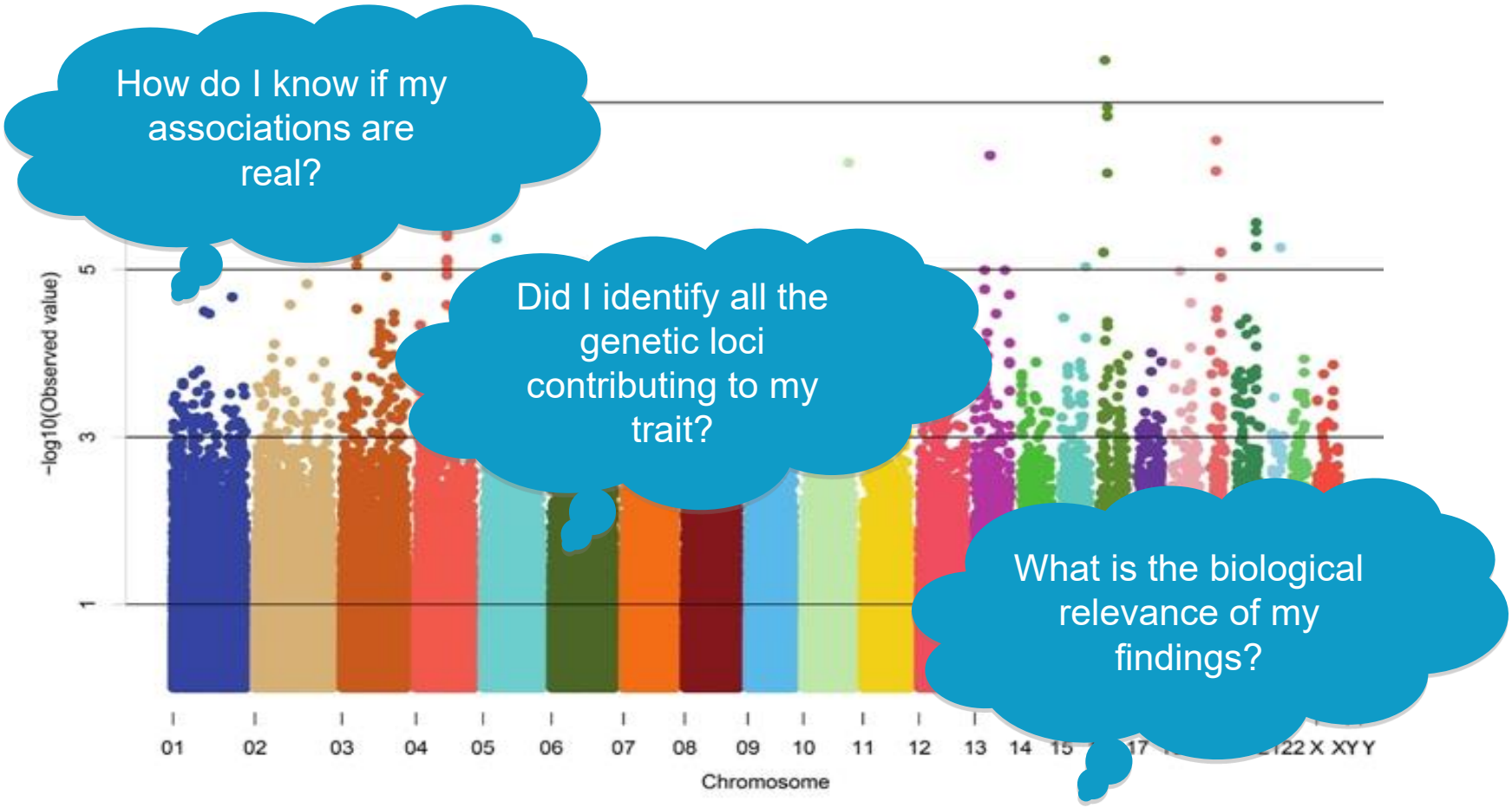
Beyond the p-value: Estimating effect sizes

Variant and risk allele	P-value	P-value annotation	RAF	OR	Beta	CI
rs3755967-T	5×10^{-343}		NR	-	0.089 unit decrease	[0.084-0.094]
rs186441690-G	2×10^{-18}		0.997189	-	0.26711 unit decrease	[0.21-0.33]

P-value= determines the likelyhood of the null hypothesis (i.e. there is no association) being true. The smaller the p-value the less likely is the null hypothesis or in oher owrds the more likely your association is true.

OR or beta (for quantitative traits)= estimates the effect of the allele. E.g. an OR of 2 for a disease trait means that the risk to get the disease given the genetic risk variant is increased 2-fold.
For quantitative traits the beta-coefficient corresponds to the average change in trait value per allelle

Caveats of genome-wide association studies



Caveats of association studies: Population stratification

- Genetic marker not associated with trait of interest but highly significant association is observed;
- Population stratification happens when allele frequencies differ between groups with and without the trait (example case – control groups) without being associated with the trait.

January 2000, Volume 5, Number 1, Pages 11–13

News & Views

Beware the chopsticks gene

D Hamer and L Sirota

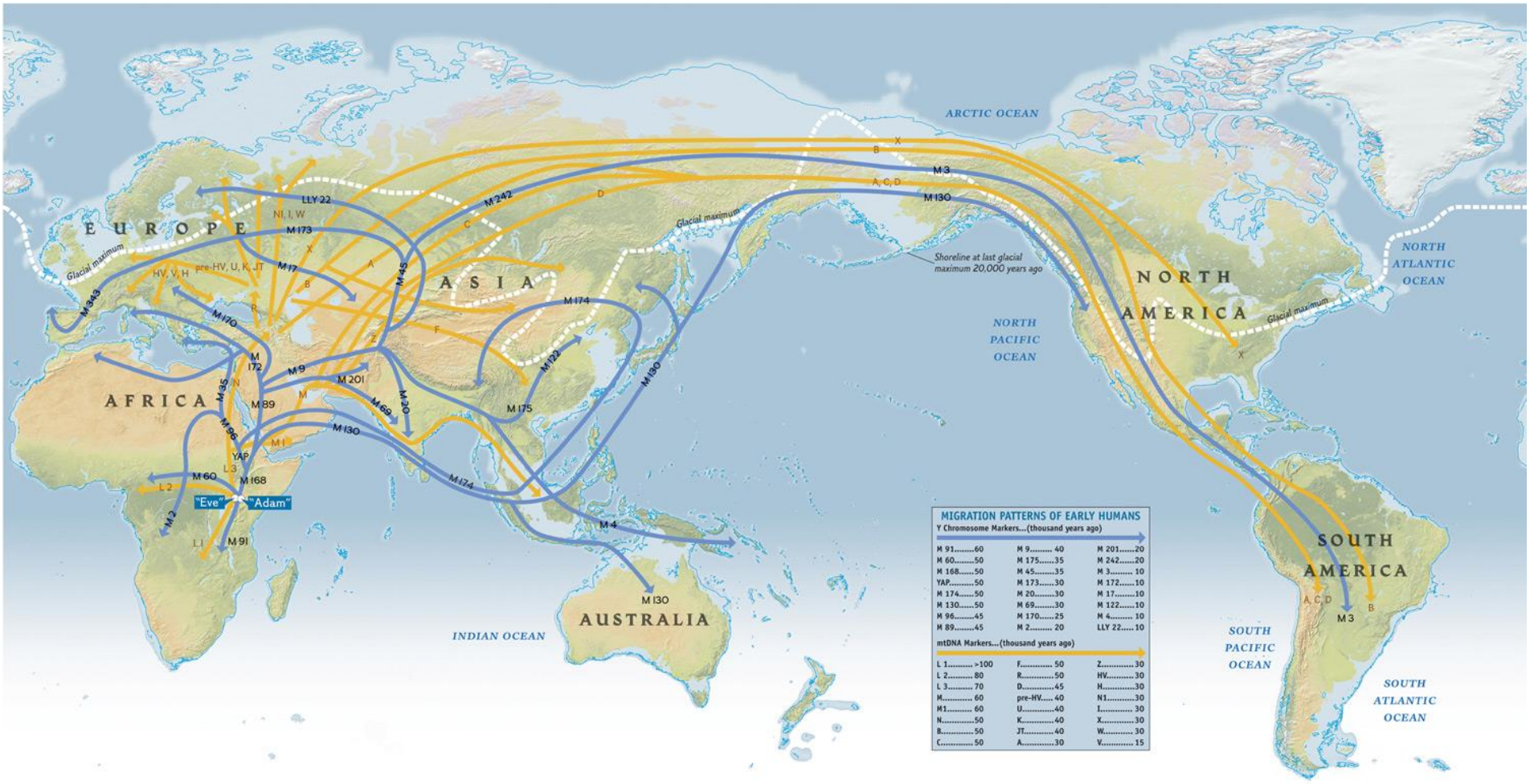
Laboratory of Biochemistry, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA

Correspondence to: Dr D H Hamer, NIH Building 37, Rm YA13, Bethesda, MD 20892, USA. deanh@helix.nih.gov

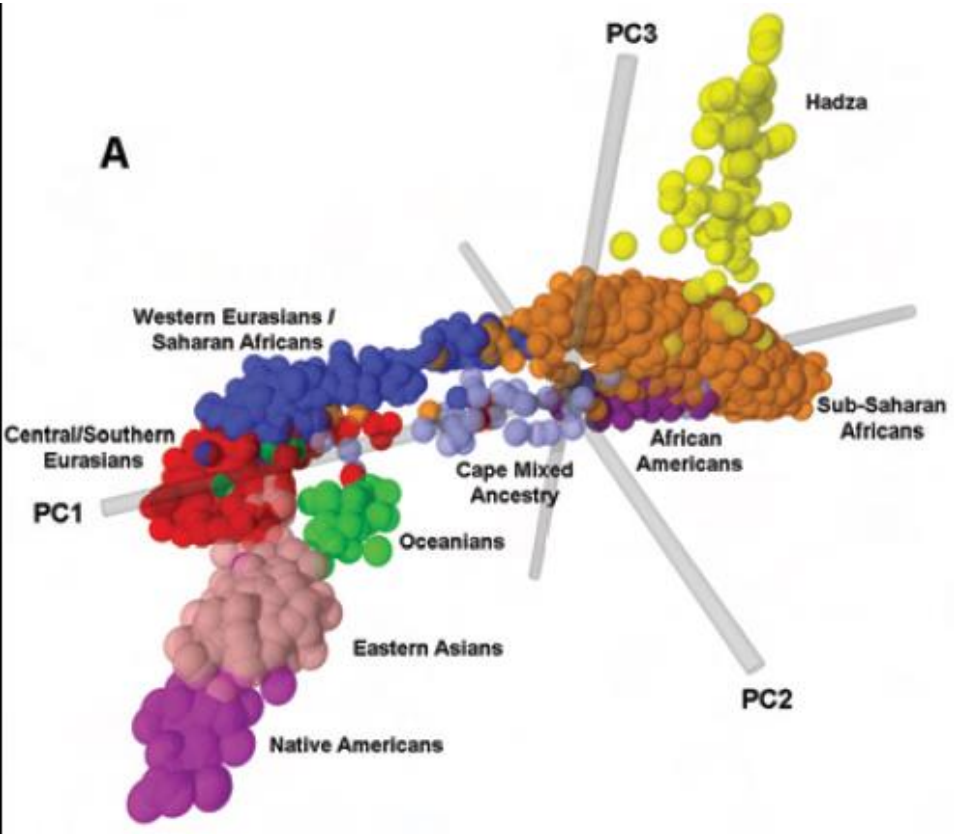
Once upon a time, an ethnogeneticist decided to figure out why some people eat with chopsticks and others do not. His experiment was simple. He rounded up several hundred students from a local university, asked them how often they used chopsticks, then collected buccal DNA samples and mapped them for a series of anonymous and candidate genes.

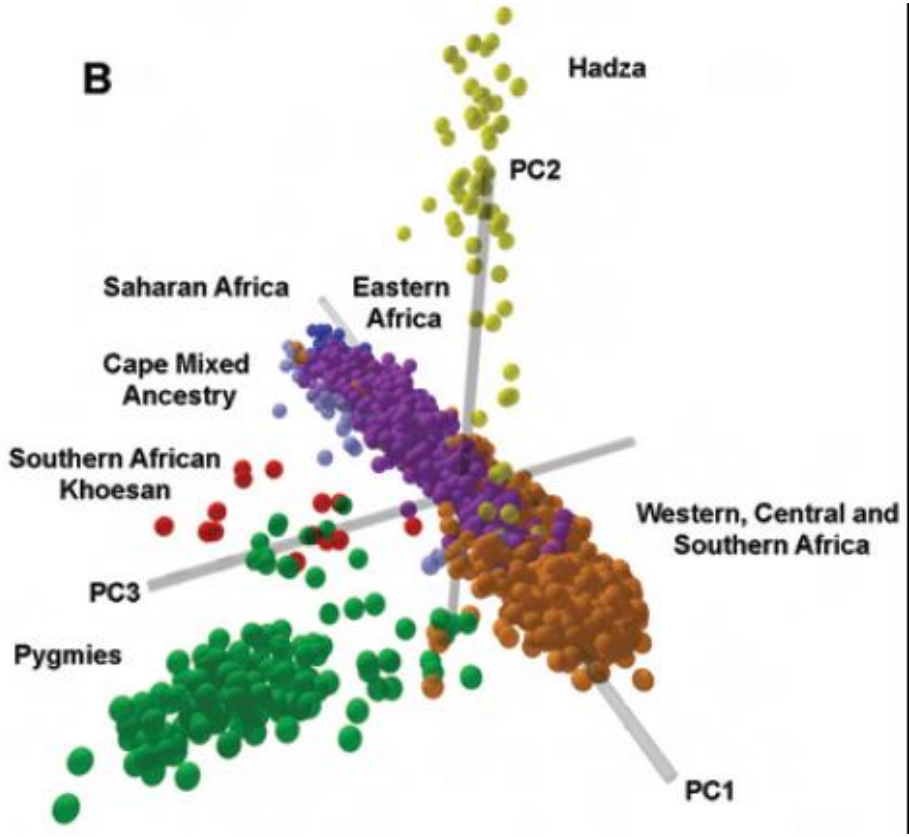
The results were astounding. One of the markers, located right in the middle of a region previously linked to several behavioral traits, showed a huge correlation to chopstick use, enough to account for nearly half of the observed variance. When the experiment was repeated with students

Facts & Challenges Human migrations



Geographical Genetics



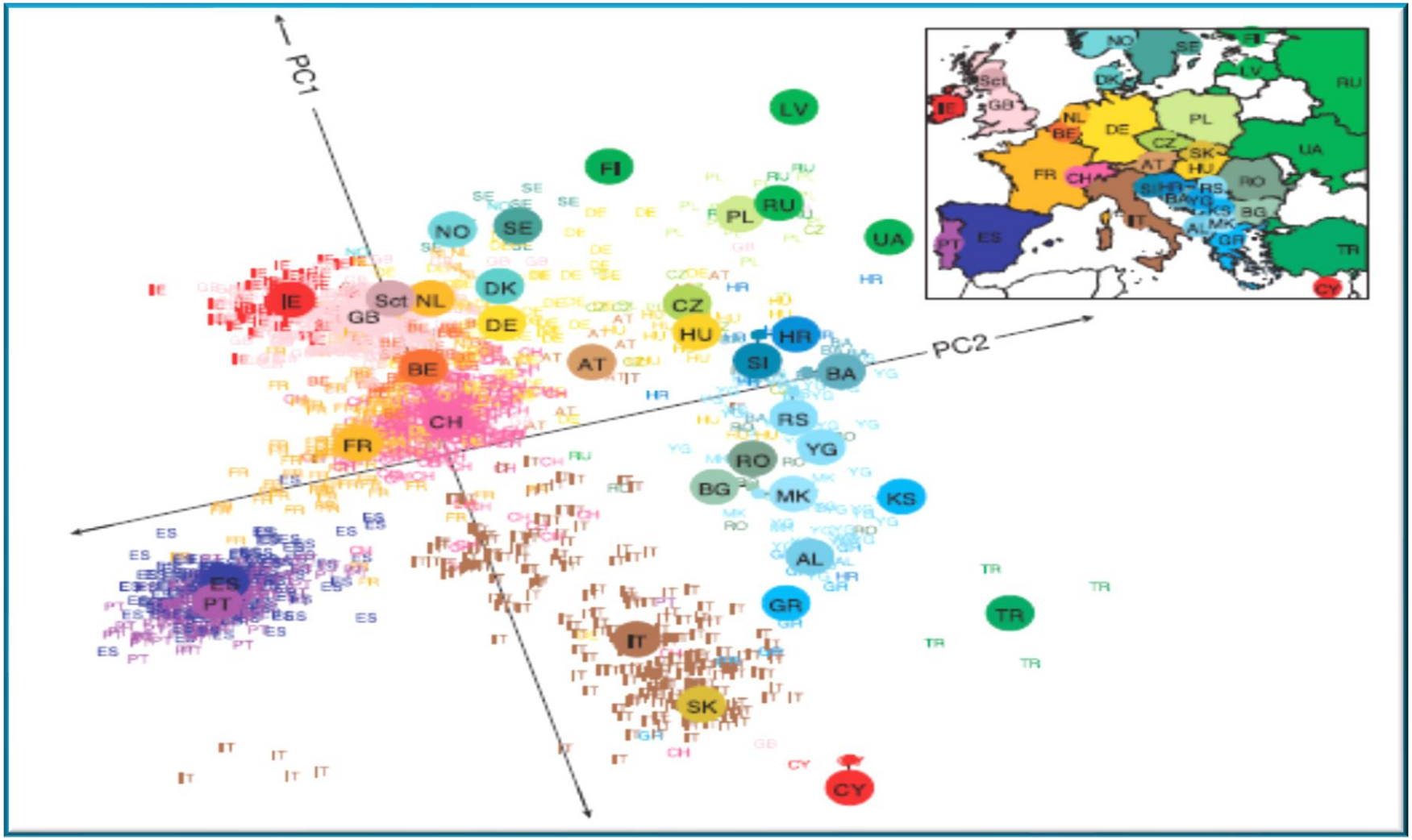


Human Populations

14 Major ancestral groups

Tishkoff et al, Science on line. 30 April 2009

Facts & Challenges Genetic Diversity



Novembre et al *Nature* 456, 98 (2008)

Caveats of association studies: Multiple testing problem

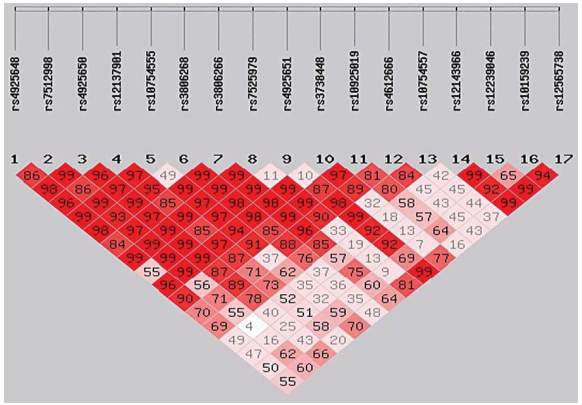
People **with** Disease (cases)

People **without** Disease (controls)

Genotype	With Disease	Without disease	Total
Allele- C	6	2	8
Allele- T	5	9	14
	11	11	

$X^2 = 5.38$

$p < 0.025$



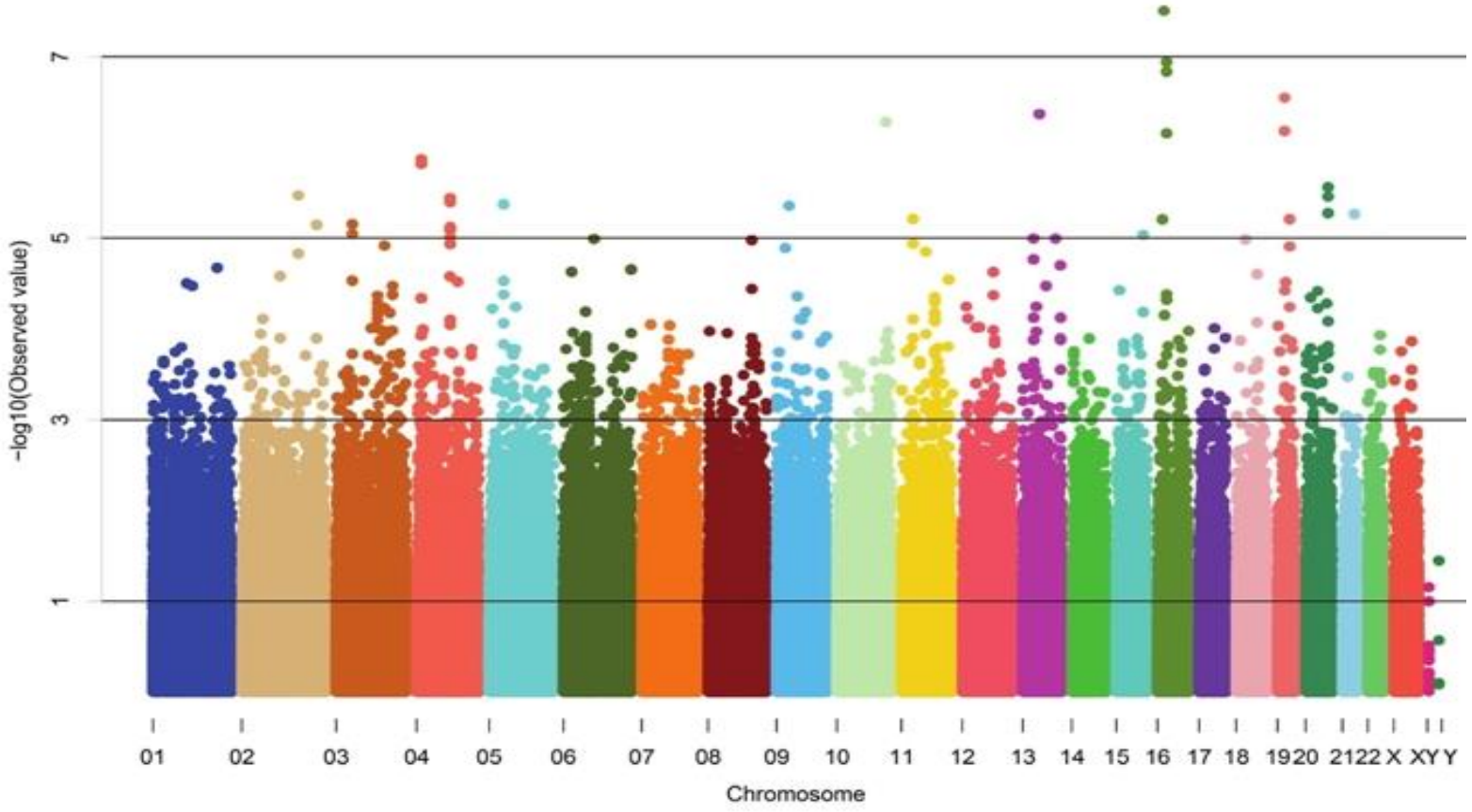
BUT…

- The more markers we test, the greater are the odds to observe a significant p-value by chance
- Hence we need to adjust our results for the number of tests we have performed;
- The simplest adjustment is the Bonferroni correction= p_{value}/number of tests

- Example: 1000 markers tested: New significance threshold p= 0.05/1000= 0.00005

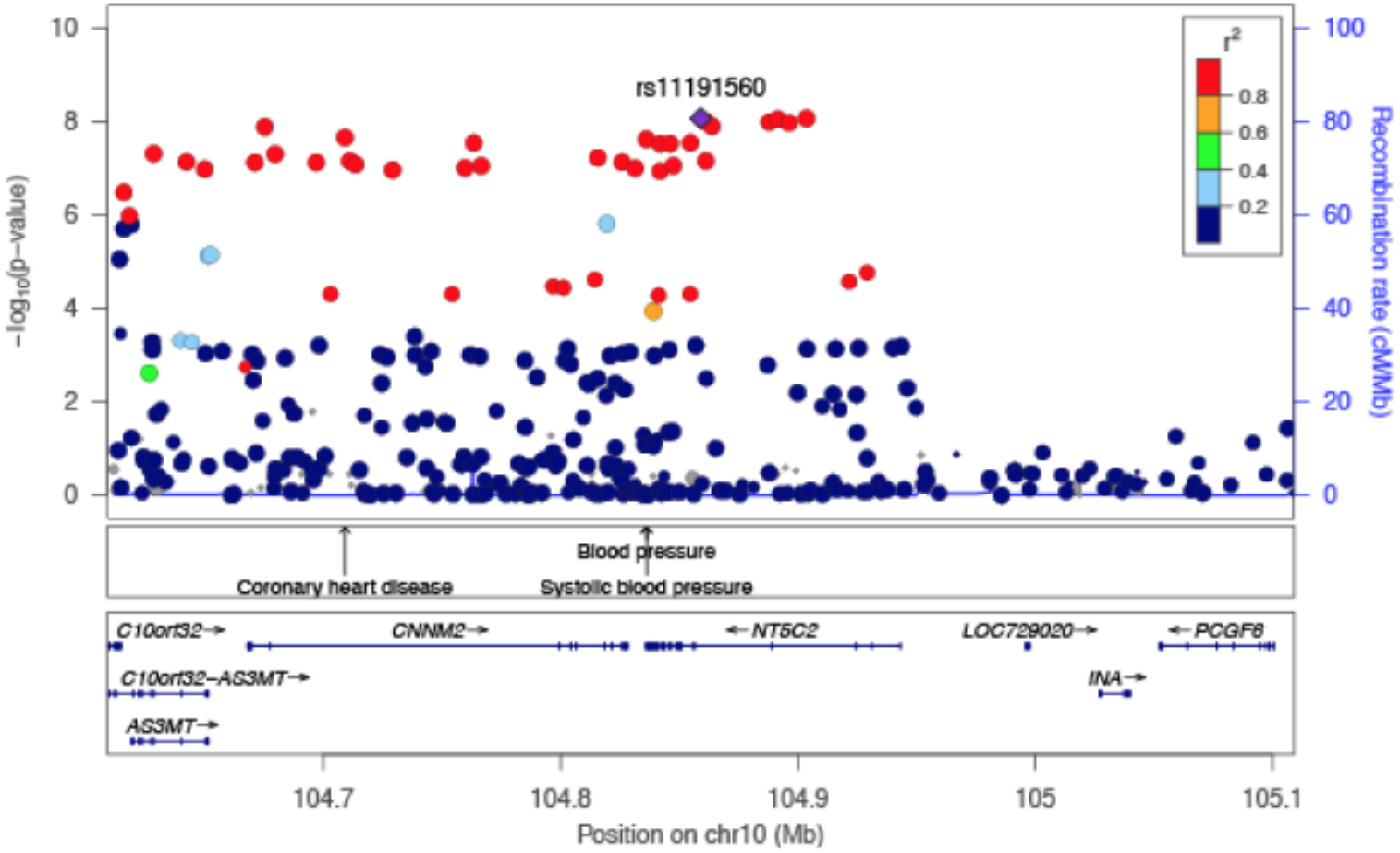
Caveats of association studies: Multiple testing problem

If we test hundreds of thousands of markers, the threshold becomes an important obstacle for trait locus identification



Limitations of GWAS studies

GWAS studies identify **genetic loci** NOT **genes**



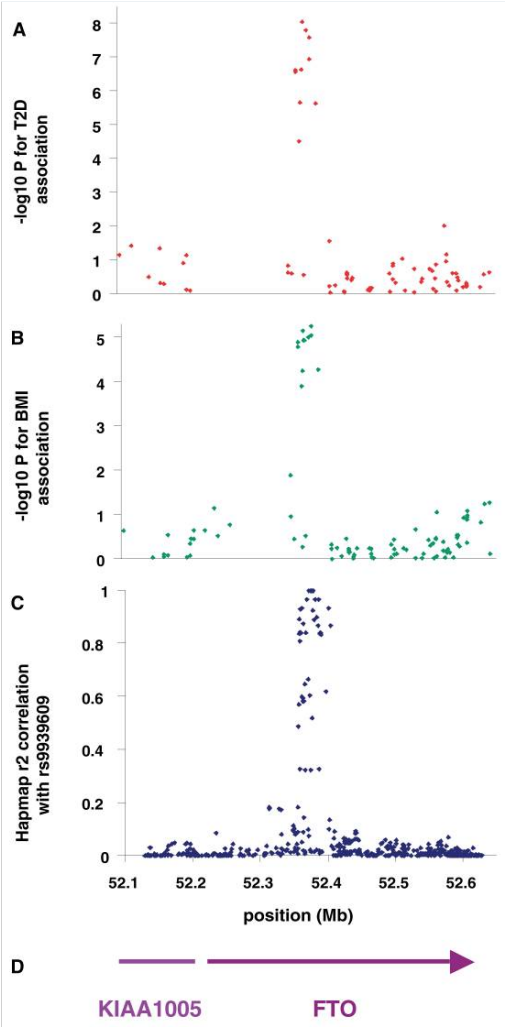
Be careful when citing gene names in association studies….

Science. Author manuscript; available in PMC 2009 Feb 22.
Published in final edited form as:
Science. 2007 May 11; 316(5826): 889–894.
Published online 2007 Apr 12. doi: 10.1126/science.1141634

PMCID: PMC2646098
NIHMSID: NIHMS45574

A Common Variant in the *FTO* Gene Is Associated with Body Mass Index and Predisposes to Childhood and Adult Obesity

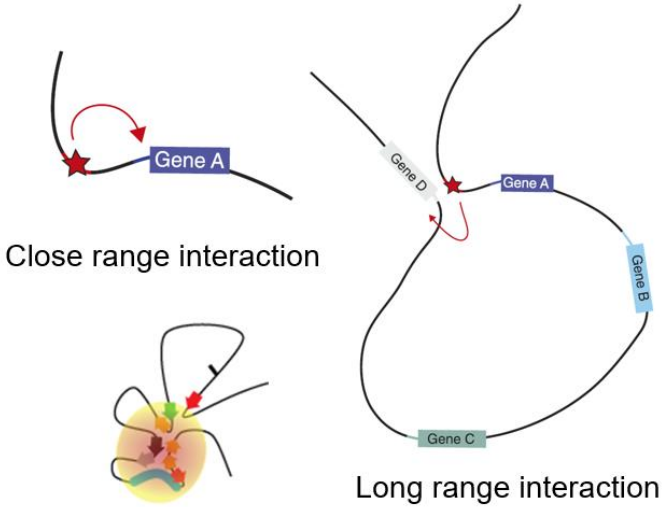
Timothy M. Frayling,[1,2,*] Nicholas J. Timpson,[3,4,*] Michael N. Weedon,[1,2,*] Eleftheria Zeggini,[3,5,*] Rachel M. Freathy,[1,2] Cecilia M. Lindgren,[3,5] John R. B. Perry,[1,2] Katherine S. Elliott,[3] Hana Lango,[1,2] Nigel W. Rayner,[3,5] Beverley Shields,[2] Lorna W. Harries,[2] Jeffrey C. Barrett,[3] Sian Ellard,[2,6] Christopher J. Groves,[5] Bridget Knight,[2] Ann-Marie Patch,[2,6] Andrew R. Ness,[7] Shah Ebrahim,[8] Debbie A. Lawlor,[9] Susan M. Ring,[9] Yoav Ben-Shlomo,[9] Marjo-Riitta Jarvelin,[10,11] Ulla Sovio,[10,11] Amanda J. Bennett,[5] David Melzer,[1,12] Luigi Ferrucci,[13] Ruth J. F. Loos,[14] Inês Barroso,[15] Nicholas J. Wareham,[14] Fredrik Karpe,[5] Katharine R. Owen,[5] Lon R. Cardon,[3] Mark Walker,[16] Graham A. Hitman,[17] Colin N. A. Palmer,[18] Alex S. F. Doney,[19] Andrew D. Morris,[19] George Davey Smith,[4] The Wellcome Trust Case Control Consortium,[†] Andrew T. Hattersley,[1,2,‡§] and Mark I. McCarthy[3,5,‡]

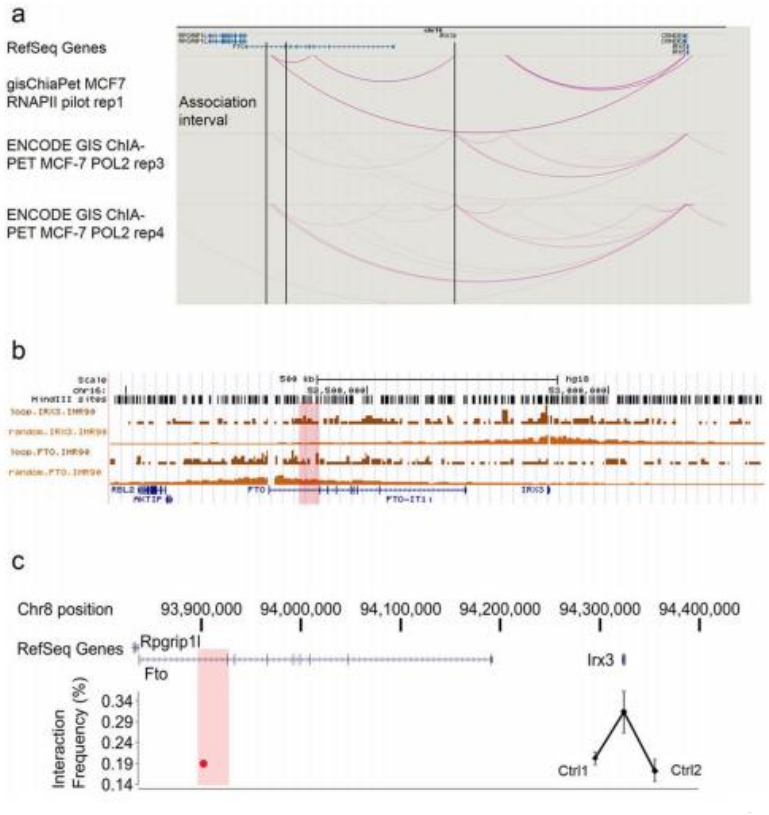


…remember association identifies loci- not genes!

Obesity-associated variants within *FTO* form long-range functional connections with *IRX3*

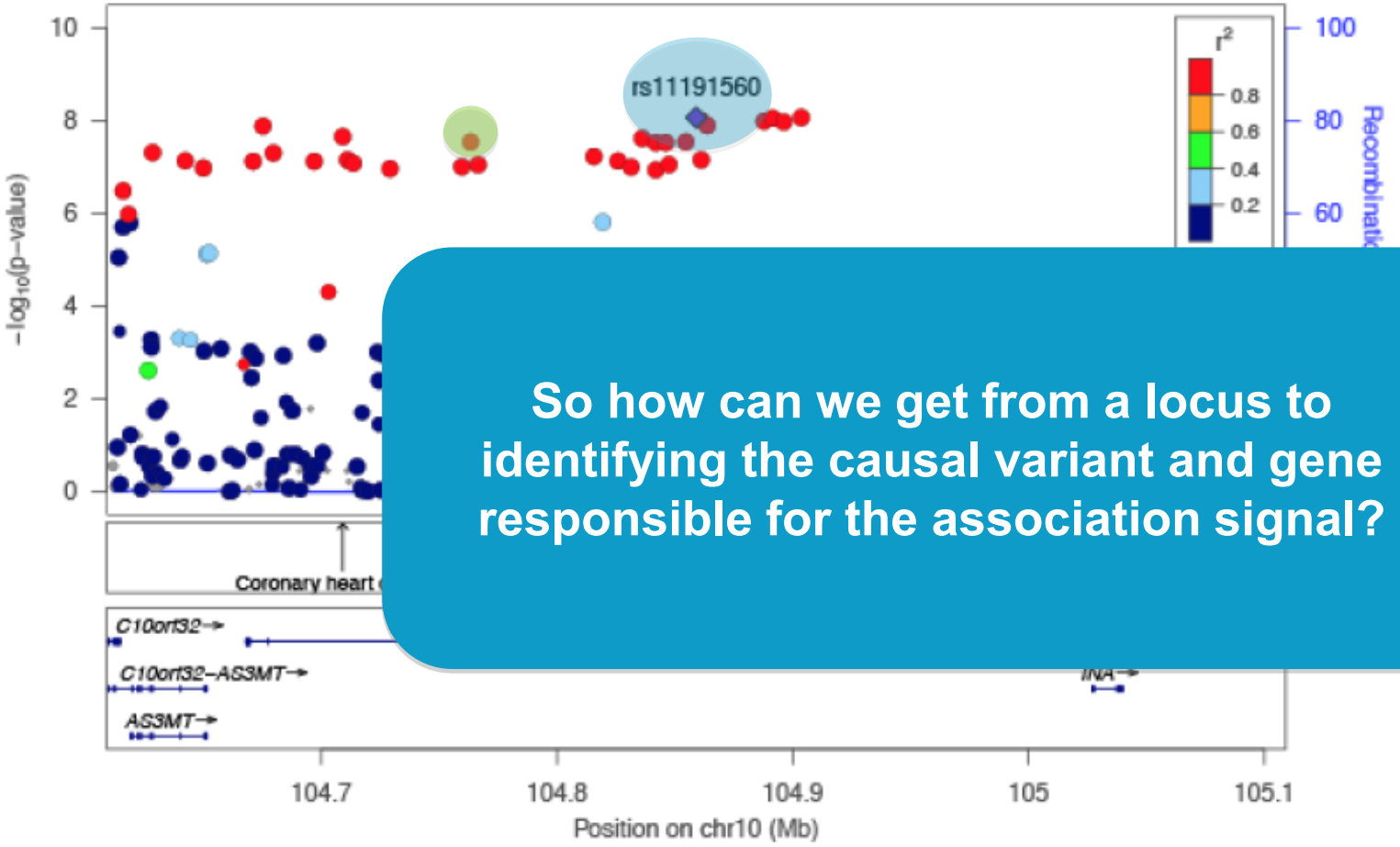
Scott Smemo[1,*], Juan J. Tena[2,*], Kyoung-Han Kim[3,*], Eric R. Gamazon[4], Noboru J. Sakabe[1], Carlos Gómez-Marín[2], Ivy Aneas[1], Flavia L. Credidio[1], Débora R. Sobreira[1], Nora F. Wasserman[1], Ju Hee Lee[3], Vijitha Puviindran[3], Davis Tam[3], Michael Shen[1], Joe Eun Son[5], Niki Alizadeh Vakili[3], Hoon-Ki Sung[5], Silvia Naranjo[2], Rafael D. Acemel[2], Miguel Manzanares[6], Andras Nagy[5], Nancy J. Cox[1,4], Chi-Chung Hui[3], Jose Luis Gomez-Skarmeta[2], and Marcelo A. Nóbrega[1]





Limitations of GWAS studies

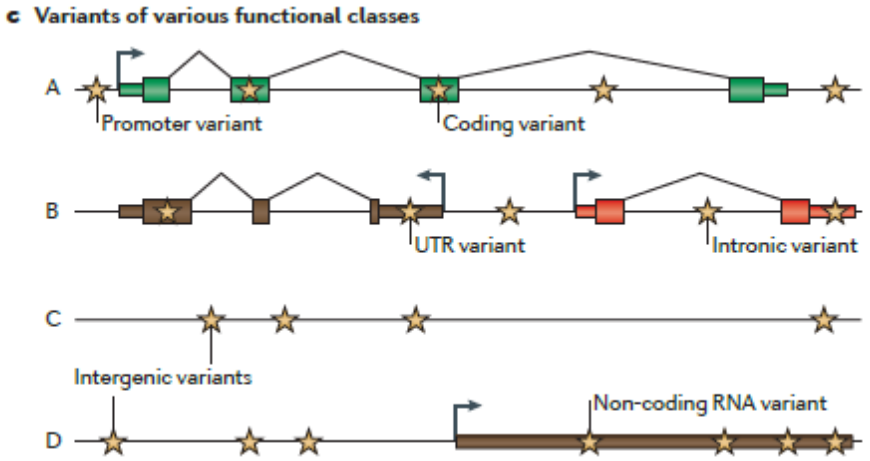
GWAS studies identify **genetic loci** NOT **genes**



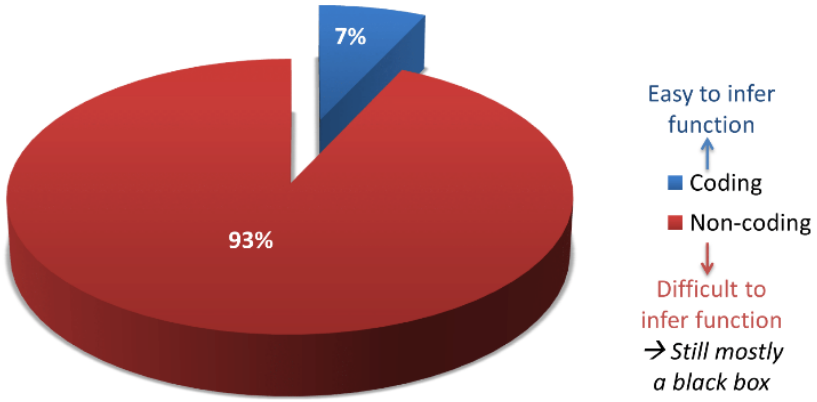
So how can we get from a locus to identifying the causal variant and gene responsible for the association signal?

How to estimate deleteriousness of SNVs to prioritize disease-causal variants ?

"Needles in stacks of needles: finding disease-causal variants in a wealth of genomic data" (Cooper & Shendure, Nature Rev. Genet. 2011):



>1,200 GWAS → 6,500 disease- or trait-predisposing SNPs



Hindorff et al., PNAS, 2010

How to estimate deleteriousness of SNVs to prioritize disease-causal variants ?

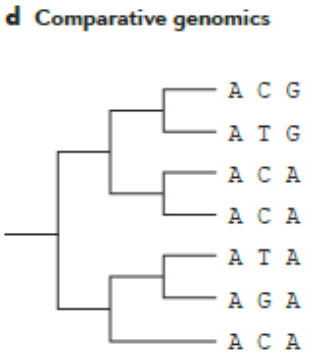
"Needles in stacks of needles: finding disease-causal variants in a wealth of genomic data (Cooper & Shendure, Nature Rev. Genet. 2011):

- Comparative genomics / cross-species conservation

- Reporter assays

- **Experimental** or **publicly** available data on functional variants:
  - Transciption Factor binding sites TFBS (*in silico* / Chromatin Immunoprecipation (ChIP)

  - Epigenomic marks (histone modifications, Dnase seq, ATACseq)



SNP2TFBS
Mapping SNPs to Transcription Factor Binding Sites

WASHU EPIGENOME BROWSER

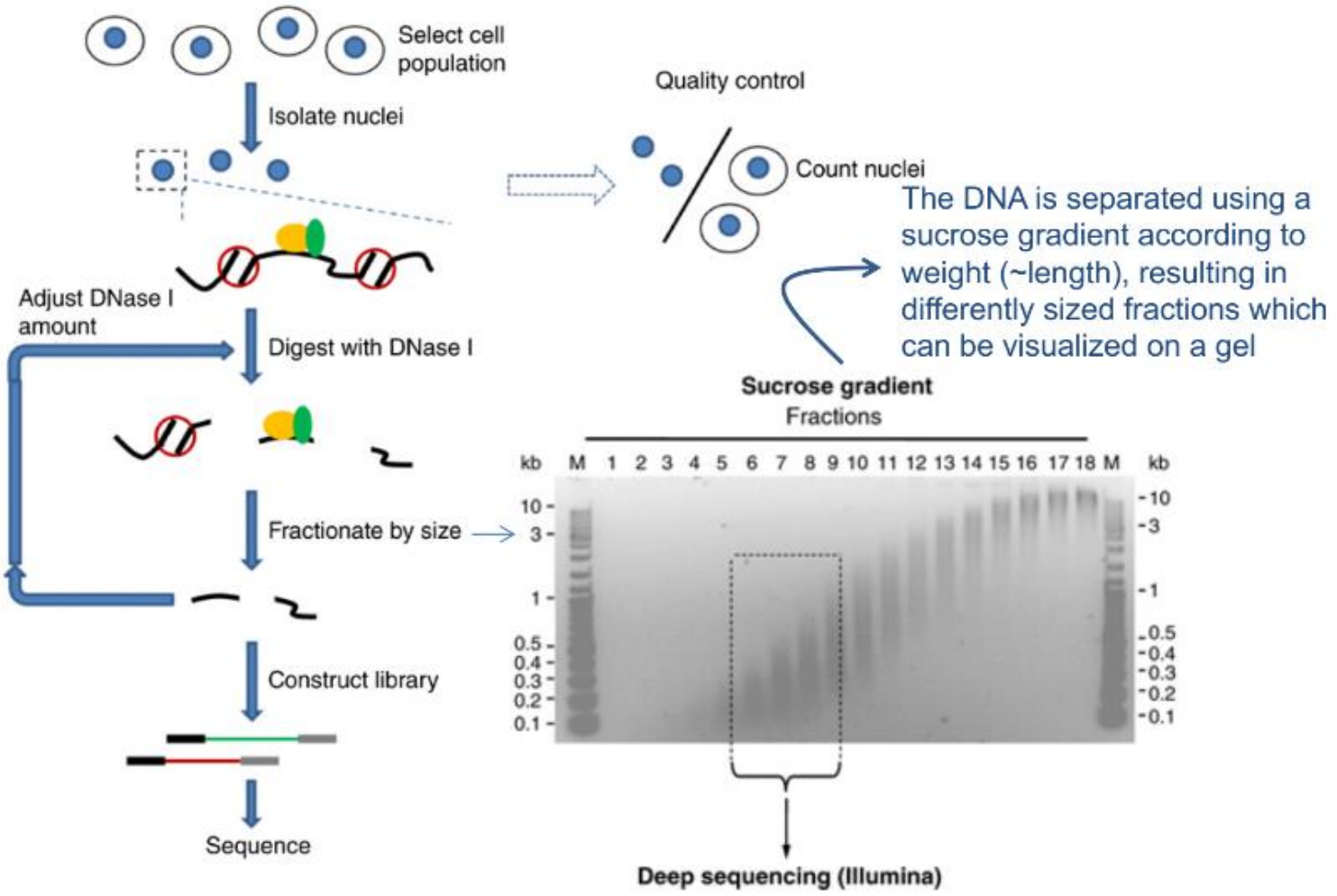
The ENCODE Project

ENCODE

• ENCyclopedia Of Dna Elements

• GOAL: Compile a comprehensive encyclopedia of all functional elements in the human genome

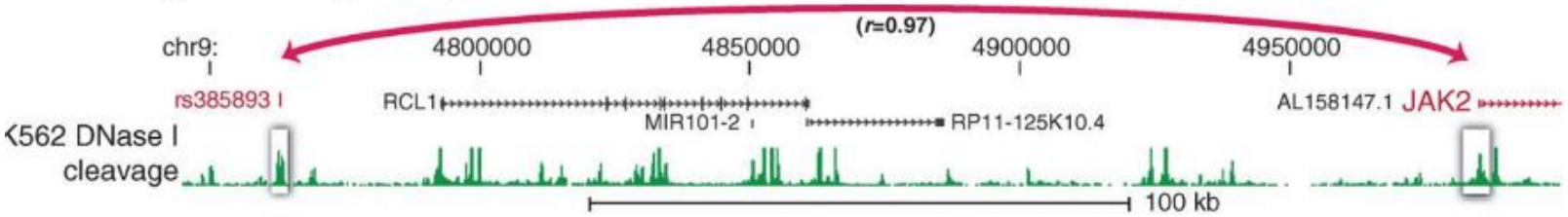
High-throughput DNAse I hypersensitive assays: concept



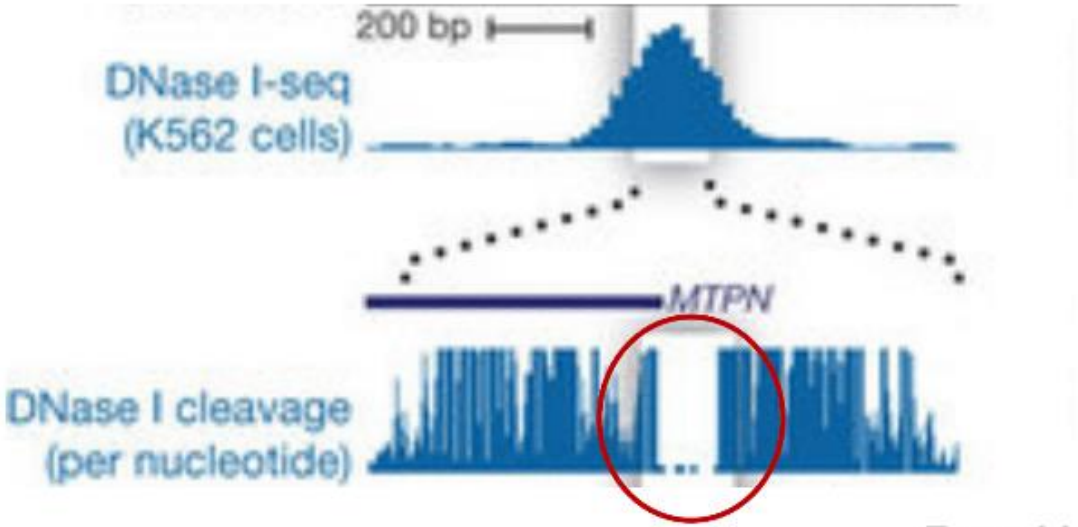
Zeng and Mortazavi, Nature Immunology, 2012

High-resolution DNase I cleavage patterns:

- provide a genome-wide high-resolution snapshot of "open", and thus presumed "active" or relevant "regulatory" regions

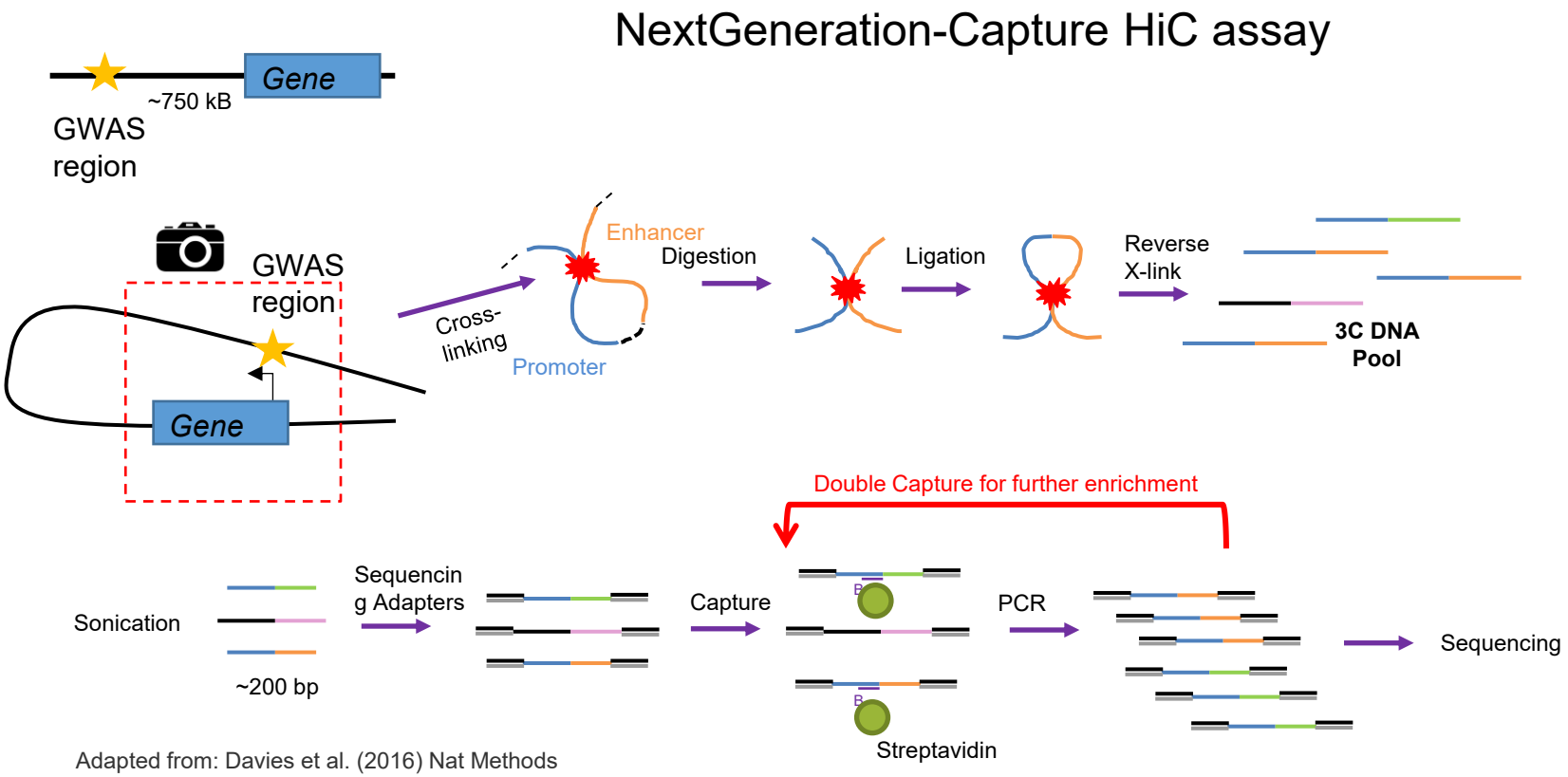


Green peaks are enriched for Dnase I sensitive DNA, thus they likely point to "interesting" regulatory regions.



From Maurano et al., Science, 2012

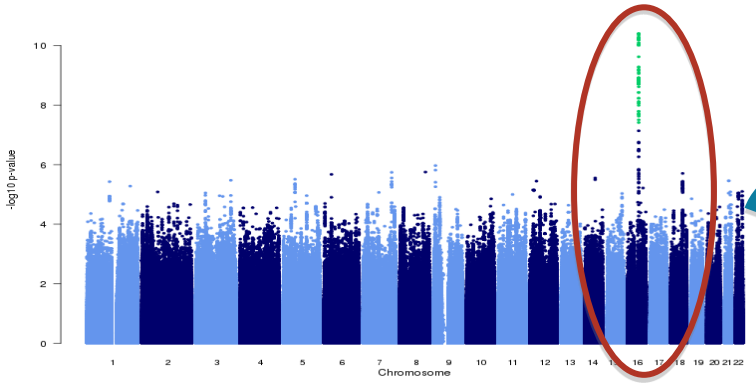
Chromosome Conformation Capture (3C): 3D architecture of a GWAS Region



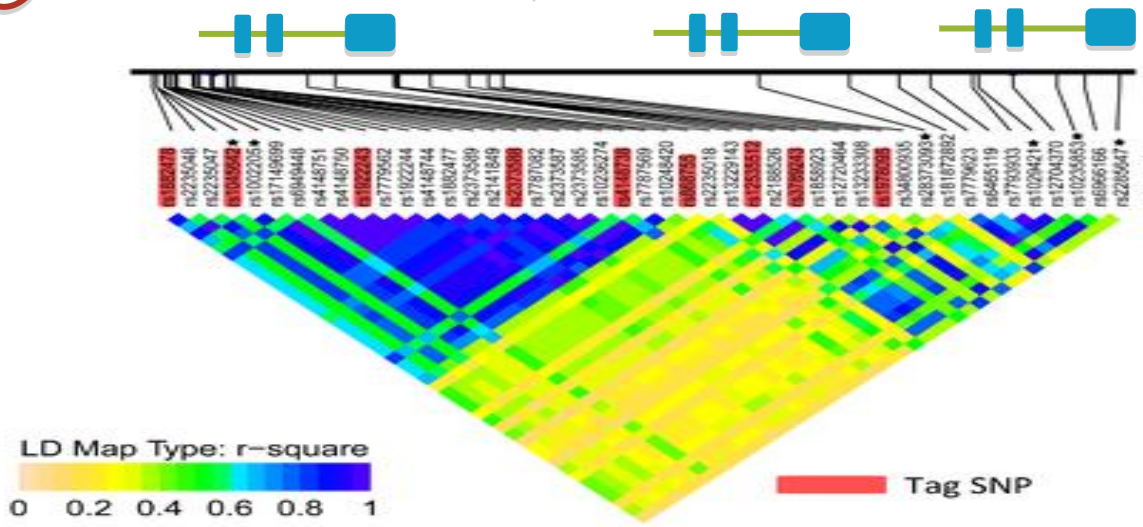
Adapted from: Davies et al. (2016) Nat Methods

Functional validation of GWAS hits using high throughput animal model techniques

GWAS region of interest



Determine all genes in interval of interest e.g. by LD



Functional validation of GWAS hits using Drosophila melanogaster ko models

GWAS hits (obesity)

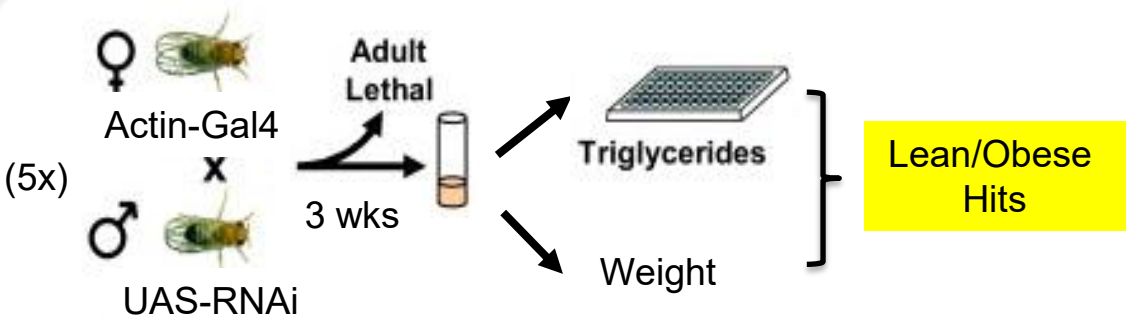
↓

Target genes (90)

↓

75 fly homologues + 20 positive controls

Phase 1



Lean/Obese Hits

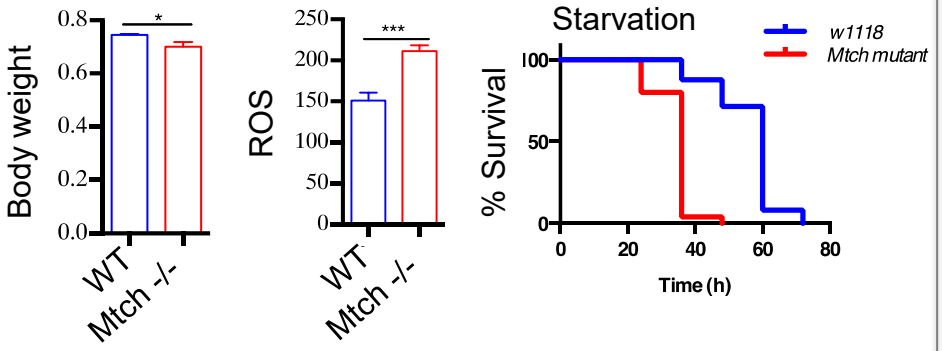
Phase 2

Nutritional Response
Fly knockdown lines
- High sucrose
- Starvation

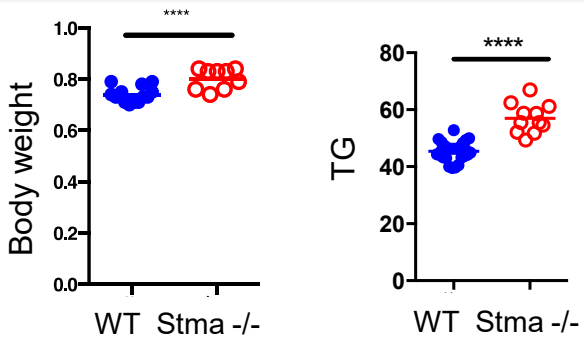
Example: Functional validation of GWAS hits using Drosophila melanogaster ko models

Determine mutants with GWAS related phenotypes

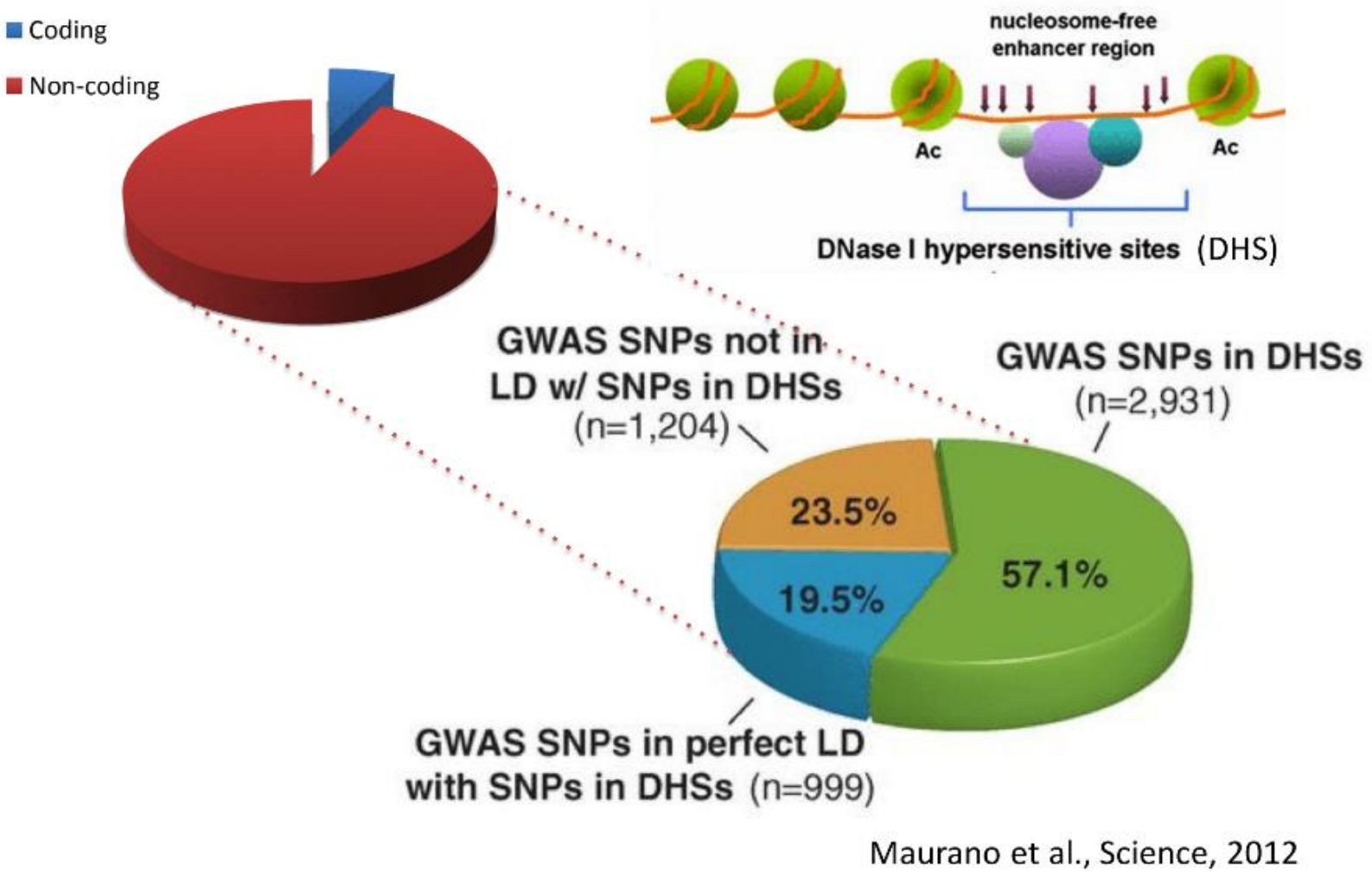
- **Example: MTCH2 (mt carrier 2):**
  - Lower expressed in AT in Cases vs Controls
  - OXPHOS repressor
  - associated with increased BMI in humans



- **Example: EFR3B (regulator of PI4K):**
  - Intronic SNP associated with weight phenotype
  - Brain-specific expression



The majority of trait-associated variants (SNPs) map to *functional* non-coding regions



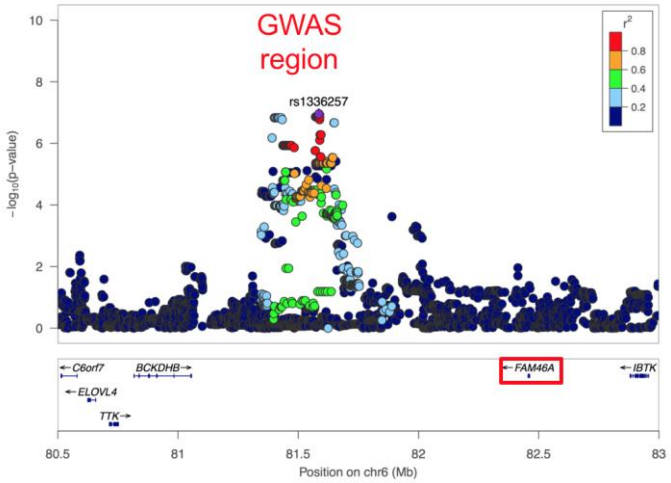
Maurano et al., Science, 2012

Expression quantitative trait locus (eQTL) analysis: identify genetic variants that significantly contribute to the variation in RNA expression abundance

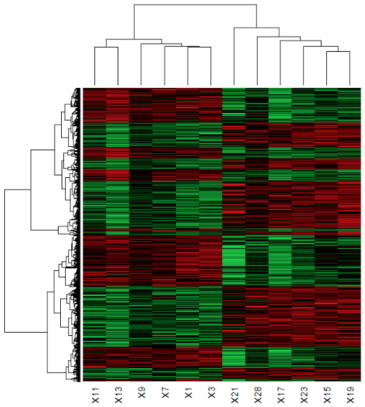
- eQTL analysis combines gene expression data with the SNP data for an associated locus
- In an eQTL analysis we perform an association analysis between the SNPs in a genetic location and the mRNA expression levels for the genes surrounding that locus.
- If the alleles for a given SNP show different level of expression for a surrounding gene, we can assume that this SNP is directly having an effect on the gene expression and hence has a functional impact as is a likely functional candidate for the initial association signal.

eQTL analysis of

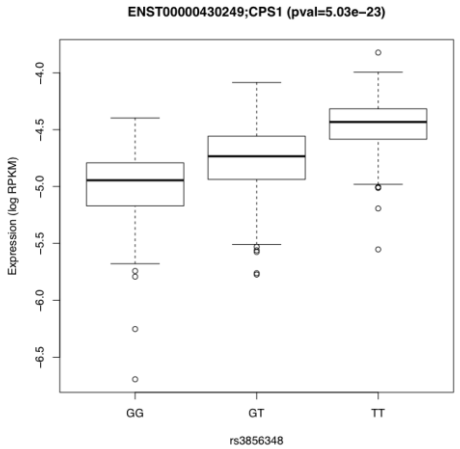
Select SNP's from positive GWAS signal



Determine gene expression levels in genes surrounding the GWAS locus



Combine data and perform association analysis between SNPs and expression levels



SNP alleles associated with differential expression of a specific gene in the GWAS region

=

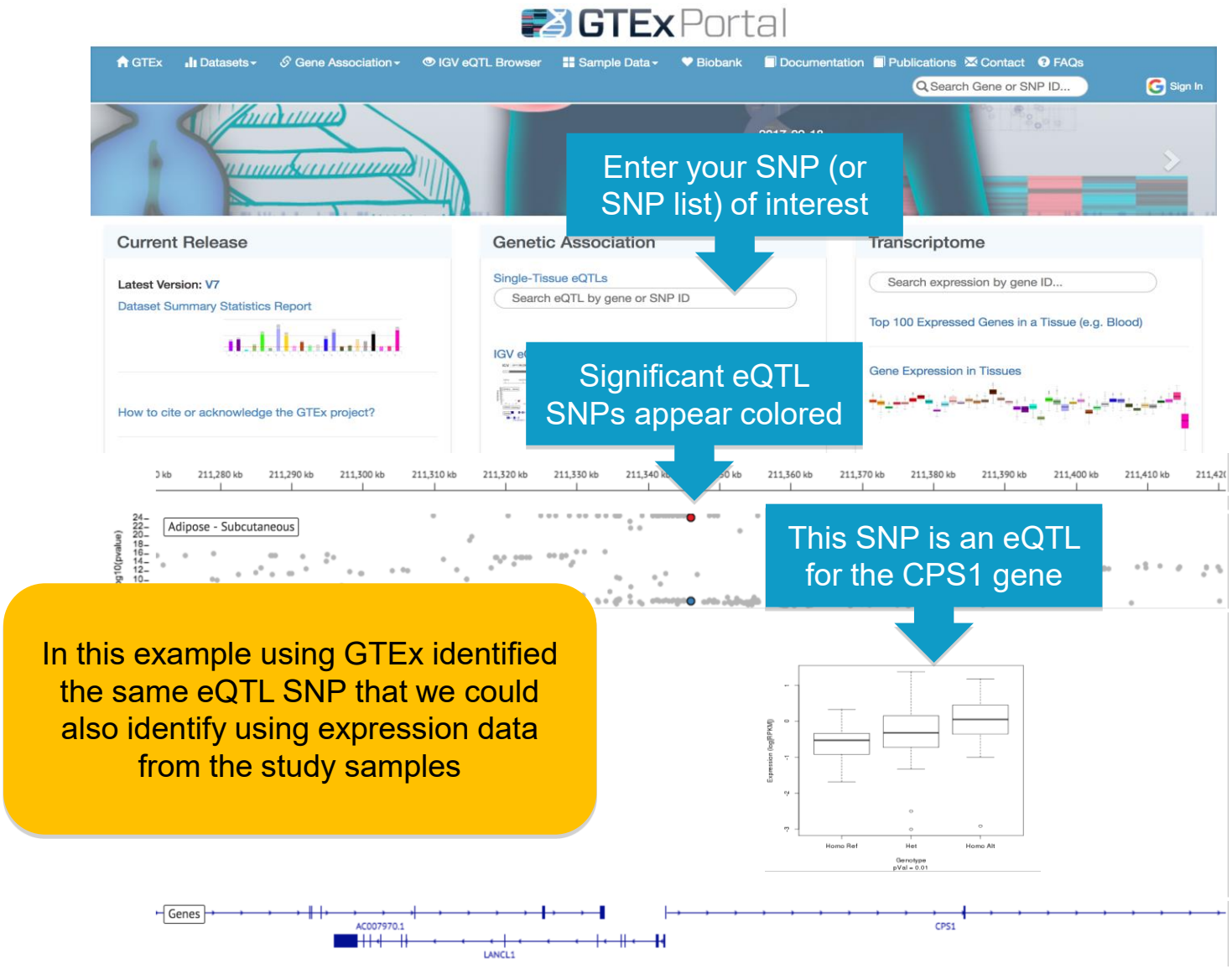
Functional SNP for that gene

Expression quantitative trait locus (eQTL) analysis: Limitations & *in silico* approaches

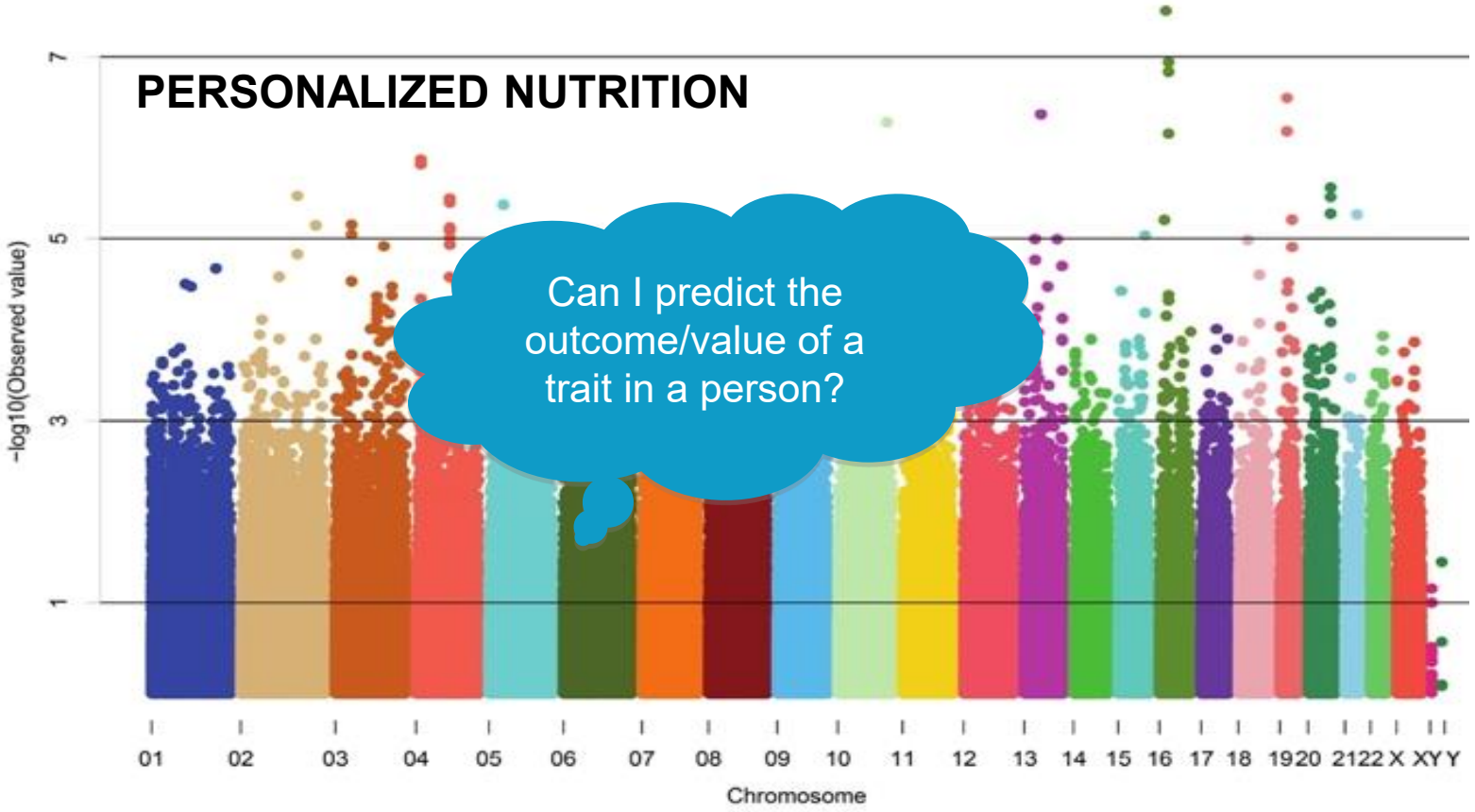
- In human intervention studies we often have tissue samples only from few, easily accessible tissues like blood, muscle, adipose tissue. Thus we rely on the functional GWAS gene being expressed in the available tissue.
- Often we have no tissue samples at all available to perform gene expression experiments on the study samples.

  In those cases we can try to use in-silico approaches to identify potential SNPs that are eQTLs for the genes in a GWAS region.

Example: in silico eQTL mapping using GTEx (https://www.gtexportal.org/home/)



In this example using GTEx identified the same eQTL SNP that we could also identify using expression data from the study samples



In Nutrigenetics we would like to predict:

- The reaction of a person to food or nutrients (e.g. food intolerances like gluten intolerance;
- The change in interaction of nutrinets (and/or other molecules) due to genetic variation (e.g. changes in bio-availability, nutrient-drug interactions;
- The quantitative impact of genetic variation on nutrient status

How do we study nutrigenetic effects in humans?

- Observational studies (case-con
  - Large sample sizes, **but**
  - Heterogenous information about dynamics in the population,
  - Environmental factor including diet intake parameters not well controlled
  - Can be cross-sectional or longitudinal

Can identify genetic variants that influence nutrient and nutrient-metabolite levels

- Clinical Intervention (experimental) studies (RCT, cross-over)
  - Well controlled for nutritional pa
  - Capture changes during the int
  - Duration limited (long term),
  - Difficult to obtain large sample sizes
  - Can be homeostatic or responsive

Can elaborate the effects of genetic variants on nutrient intake

Genetics influences the abundance, metabolism and function of…..

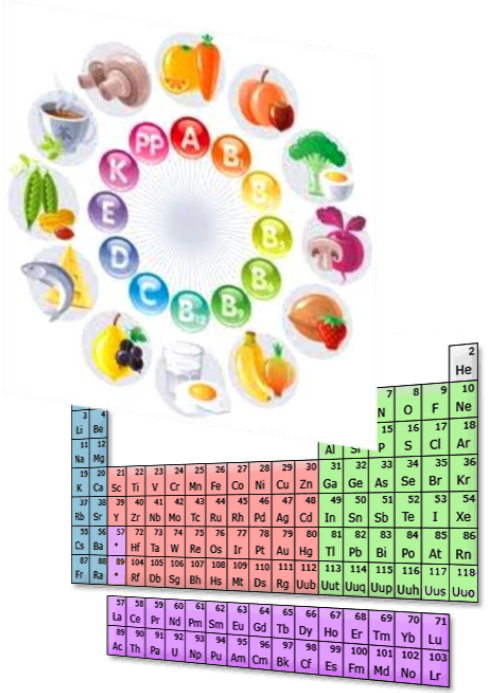
Lifestyle ➡ ***Diet*** ➡ ***Nutrients*** ➡ ***MACROnutrients***
➡ ***microNUTRIENTS***

Vitamins – essential nutrients, cofactors

Fatty Acids- they store energy form compartment structures and send signals

Amino Acids – signalling, building blocks

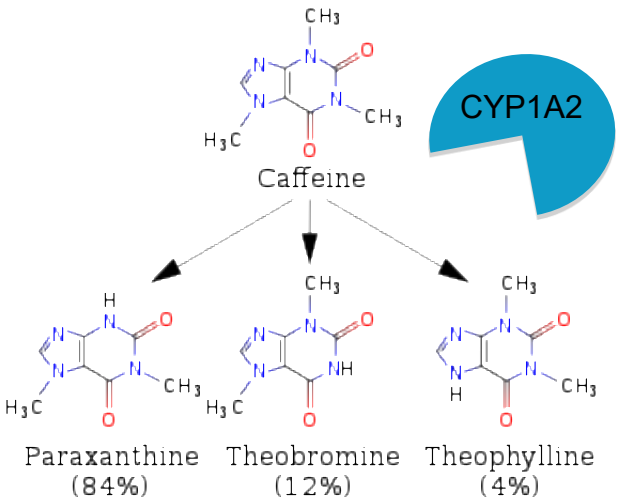
Minerals – trace elements, cofactors

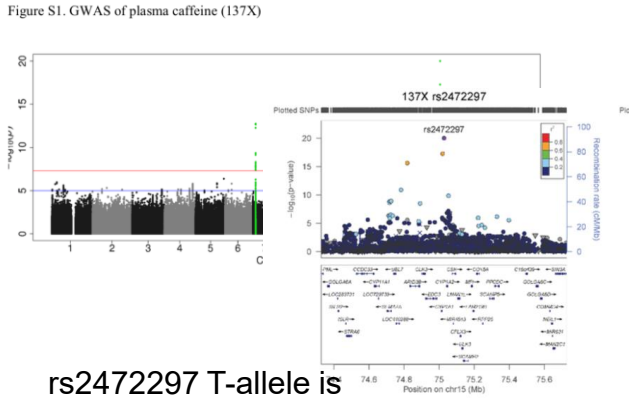


Genetic variants influence susceptibility to food sensitivities

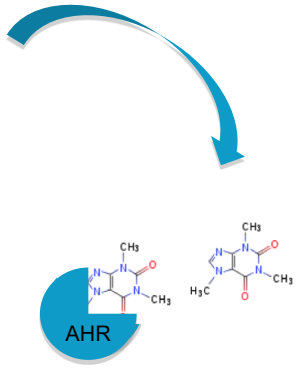
- Gluten;
- Lactose;
- Caffeine….

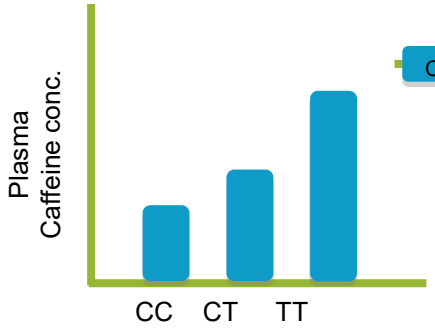
Genetic variation effects on nutrient metabolism: caffeine





rs2472297 T-allele is associated with slow caffeine metabolism





*AHR response element

*Aryl-hydrocarbon receptor

Environmental Heterogeneity: Lactose intolerance

N. European

Indian children

Afr American children

Indian adults

Mex American - adult

Cretans

Cypriots

N. American Jews

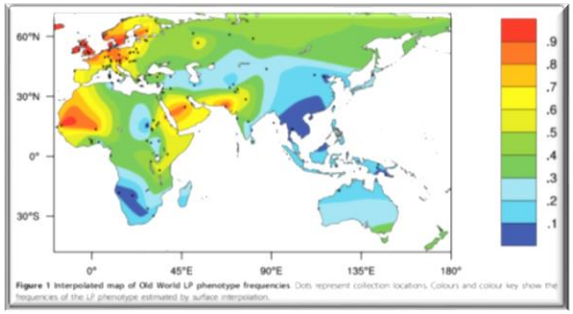
Mexicans - rural

Afr American - adult

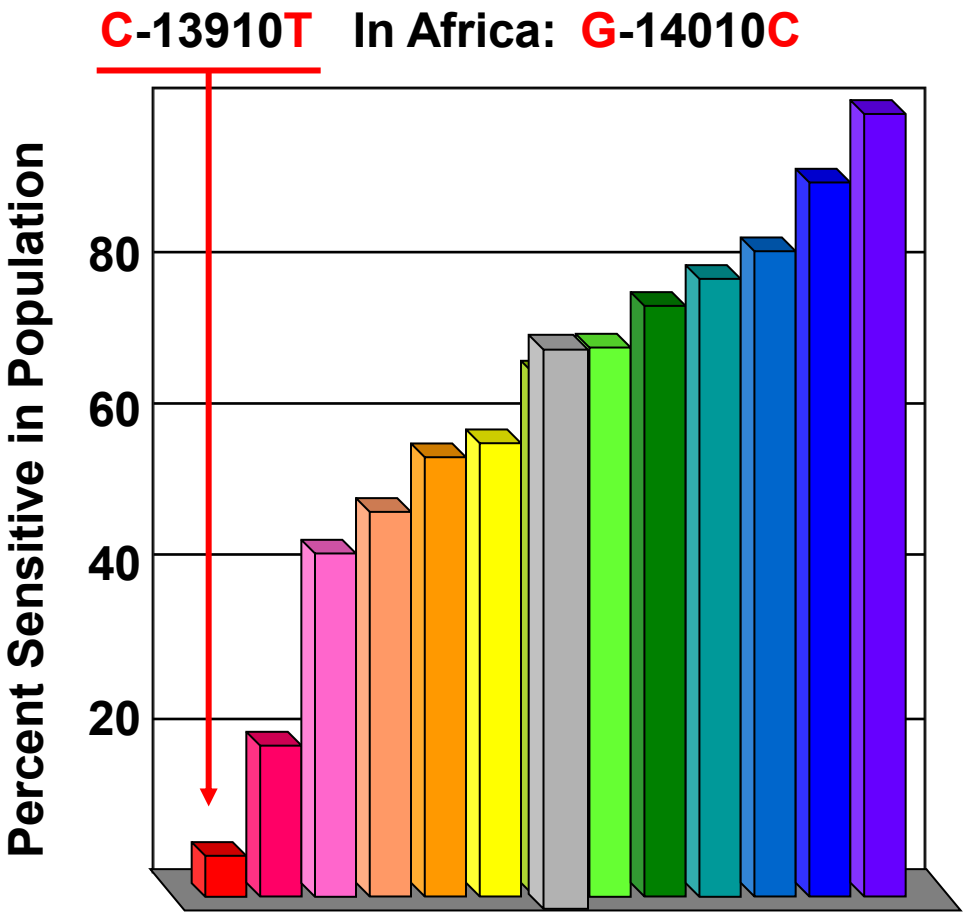
Eskimo

Asian Americans

SE Asians



Itan *et al. BMC Evolutionary Biology* 2010, **10**:36
http://www.biomedcentral.com/1471-2148/10/36



Kaput and Rodriguez, *Physiological Genomics* 16, 166 (2004)

Environment dependent genotype effects: vitamin D levels

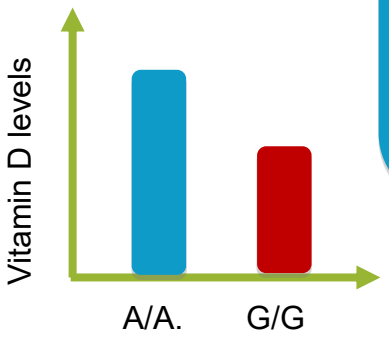
CYP2R1
rs10741657

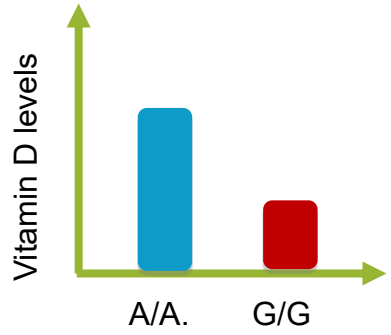
G/G

A/A

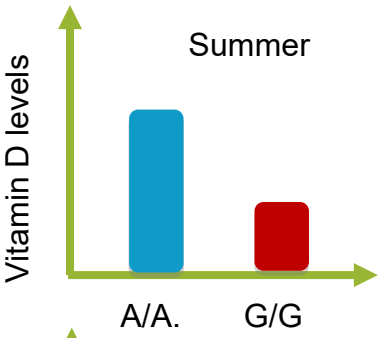
G/G

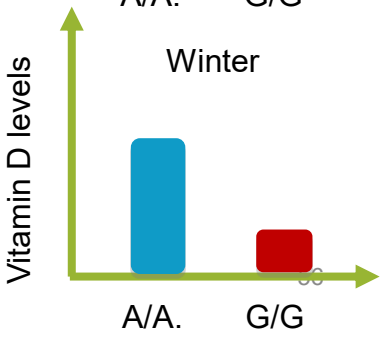
A/A



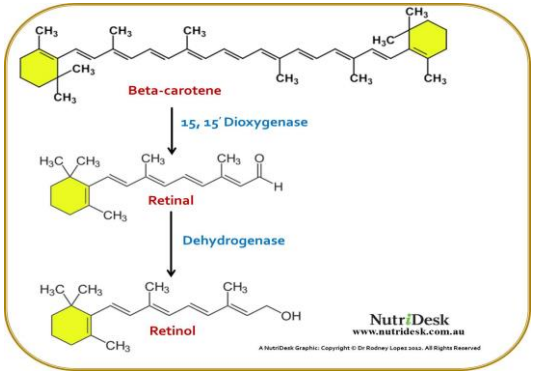


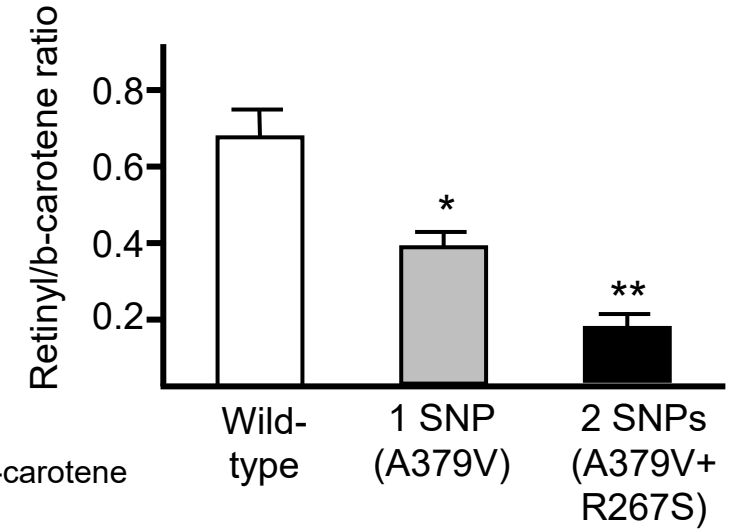
Current GWAS status- identified >140 loci hat influence vitamin D levels





Complex genetic effects: Vitamin A metabolism



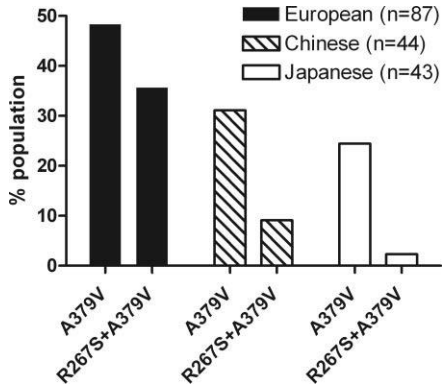


Leung et al. FASEB, 2009: Two missense mutations associated with b-carotene conversion levels

BCMO1 variant	K_m [μM]	V_{max} [nmol·mg protein^{-1}·min^{-1}]
Wild type	18.3	2.3
267S (rs12934922)	17.9	2.2
379V (rs7501331)	12.3	2.0
267S+379V	19.5	1.0 *

TABLE 5. *BCMO1 R267S and A379V SNP frequencies (n= 131)*

SNP	Genotype	*n*	%
R267S (rs12934922)	A/A	50	38.2
	A/T	52	39.7
	T/T	29	22.1
A379V (rs7501331)	C/C	74	56.5
	C/T	51	38.9
	T/T	6	4.6
R267S + A379V (combined)	AA/CC, AA/CT, AA/TT, AT/CC, TT/CC	100	76.3
	AT/CT, AT/TT, TT/CT	31	23.7



The majority of nutrients are influenced by MANY independent genetic variants

GWAS Catalog

The NHGRI-EBI Catalog of human genome-wide association studies

vitamin D measurement

Examples: breast carcinoma, rs7329174, Yao, 2q37.1, HBS1L, 6:16000000-25000000

GWAS / Search / vitamin D measurement

Refine search results

T Traits 1

Catalog stats

- *Last data release on 2022-03-23*
- *5675 publications*
- *202969 SNPs*
- *353558 associations*
- *Genome assembly GRCh38.p13*
- *dbSNP Build 154*
- *Ensembl Build 105*

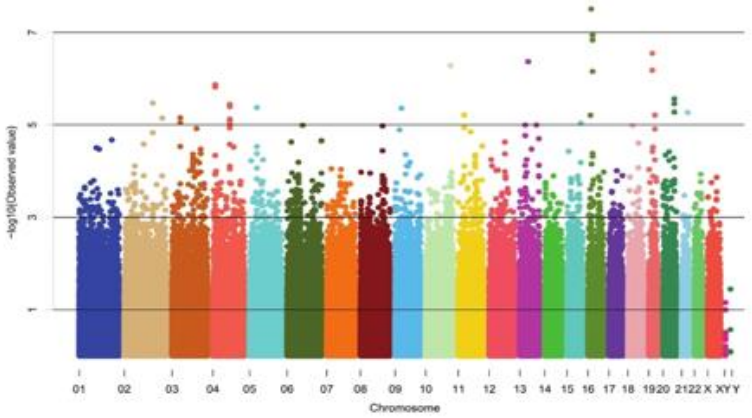
Search results for *vitamin D measurement*

T vitamin D measurement EFO_0004631

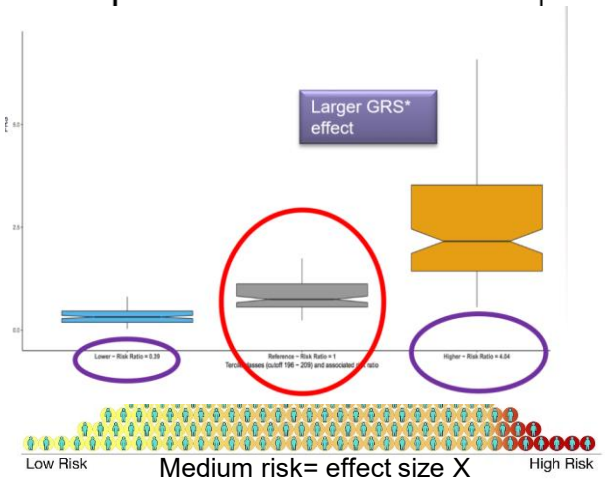
A quantification of Vitamin D levels, typically in blood. Vitamin D includes both CHOLECALCIFEROLS and ERGOCALCIFEROLS, which have the common effect of preventing or curing RICKETS in animals. It can ... Show more >

Associations 713 Studies 52

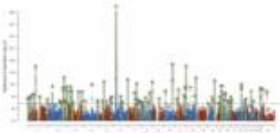
Using Polygenic risk scores to quantify the genetic effects



Sample / SNP	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	...
SNP_1	AA	AA	BB	NC	AA	AA	AA	AB	
SNP_2	BB	AB	BB	BB	AB	BB	BB	NC	
SNP_3	AA	AA	AB	AB	AA	BB	AA	AA	
SNP_4	AB	AA	AB	AB	BB	AB	NC	BB	...
SNP_5	BB	AB	BB	BB	NC	BB	BB	AA	
SNP_6	AB	BB	AB	AB	BB	AB	AA	AB	



Medium risk= effect size X



Independent Samples

BASE DATA — **TARGET DATA**

DATA

- Summary statistics
- Betas/ORs weights in PRS calculation

- Individual-level genotype and phenotype data
- Often small sample size

PROCESSING

Quality Control

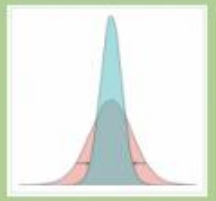
- Both data sets QC'ed as standard in GWAS
- Some QC requires special care in PRS eg. sample overlap, relatedness, population structure (see Section 2)
- Retain set of SNPs that overlap between base and target data

PRS CALCULATION

LD adjustment

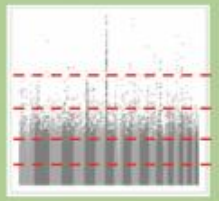
- eg. clumping

Beta Shrinkage



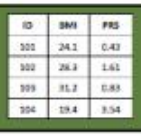
- eg. lasso/ridge

***P*-value thresholding**



- PRS at multiple *P*

TEST



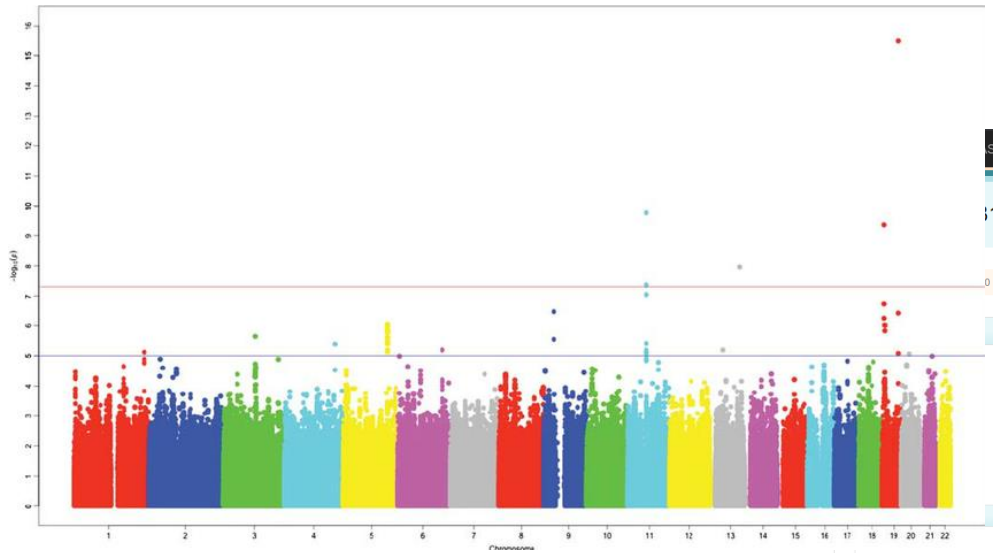
Generate PRS + Perform Association Testing

VALIDATE

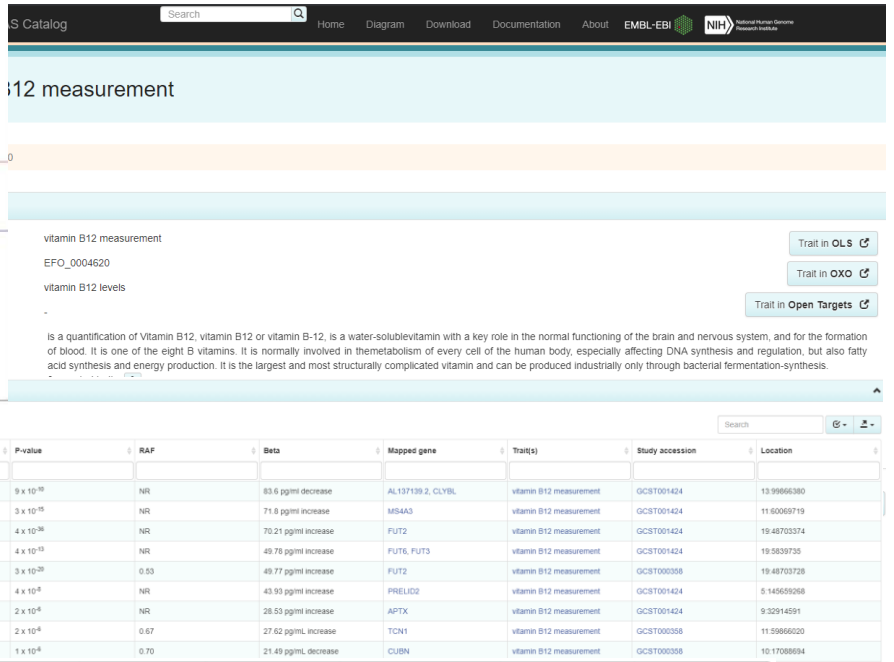
Out-of-sample PRS testing

- K-fold cross-validation
- Test in data separate from base/target

GWAS data to calculate quantitative polygenic risk scores for vitamin B12



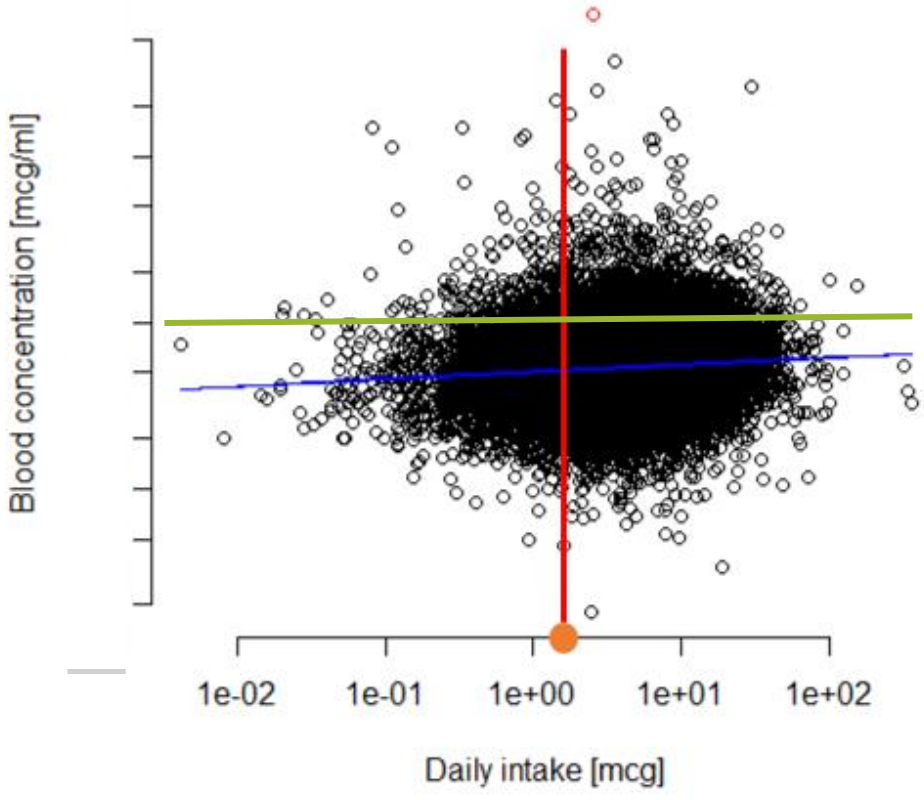
Xialing Lin Human Molecular Genetics · March 2012



Variant and risk allele	P-value	RAF	Beta	Mapped gene	Trait(s)	Study accession	Location
rs41281112-?	9 x 10^-10	NR	83.6 pg/ml decrease	AL137139.2, CLYBL	vitamin B12 measurement	GCST001424	13:99866380
rs2296585-?	3 x 10^-15	NR	71.8 pg/ml increase	MS4A3	vitamin B12 measurement	GCST001424	11:60069719
rs1047781-?	4 x 10^-36	NR	70.21 pg/ml increase	FUT2	vitamin B12 measurement	GCST001424	19:48703374
rs3760776-?	4 x 10^-13	NR	49.78 pg/ml increase	FUT6, FUT3	vitamin B12 measurement	GCST001424	19:5839735
rs602662-A	3 x 10^-20	0.53	49.77 pg/ml increase	FUT2	vitamin B12 measurement	GCST000358	19:48703728
rs10515552-?	4 x 10^-8	NR	43.93 pg/ml increase	PRELID2	vitamin B12 measurement	GCST001424	5:145659268
rs12377462-?	2 x 10^-6	NR	28.53 pg/ml increase	APTX	vitamin B12 measurement	GCST001424	9:32914591
rs526934-A	2 x 10^-6	0.67	27.62 pg/mL increase	TCN1	vitamin B12 measurement	GCST000358	11:59866020
rs11254363-A	1 x 10^-6	0.70	21.49 pg/mL decrease	CUBN	vitamin B12 measurement	GCST000358	10:17068694

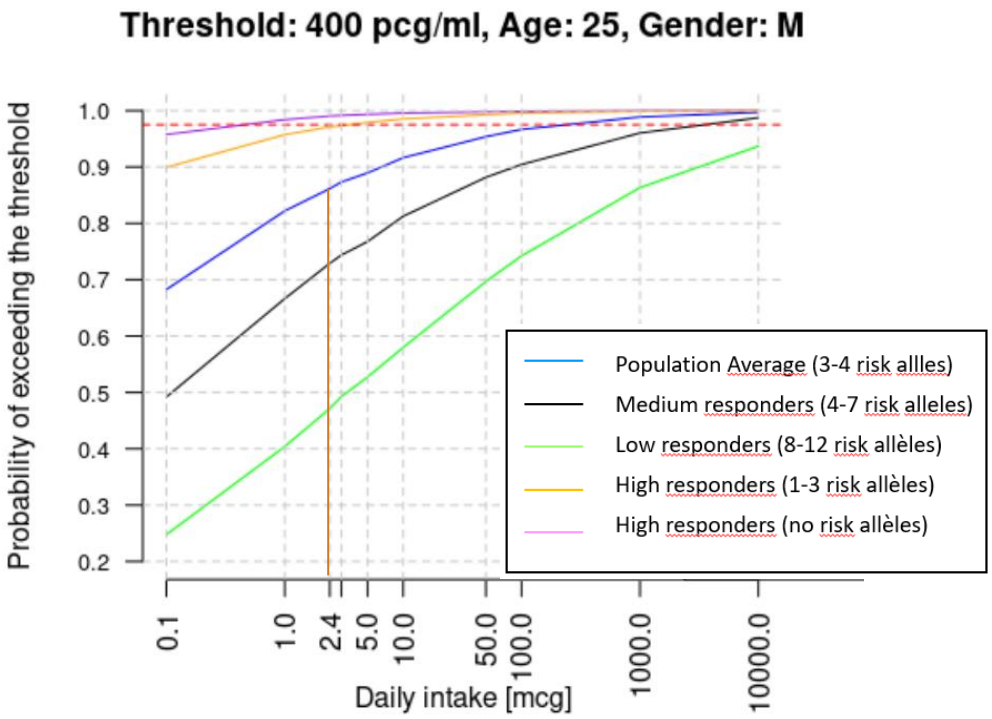
Trait	Measurment Method in Arivale	Units	Reference Lab Range	Range in Arivale Sample	Arivale Sample Size for PRS	Number of SNPs from GWAS	Number of SNPs included in the PRS	PRS Risk Allele Range in dataset	PRS Coefficient in Model	PRS P-value in Model	R2 of Model
B12	Methylmalonic Acid Serum	(nmol/L)	<270	53-874	1737	21	13	5-22	4.03431124	6.95E-05	0.3

Case example: Genetic risk score provides basis for quantitative vitamin B12 recommendations

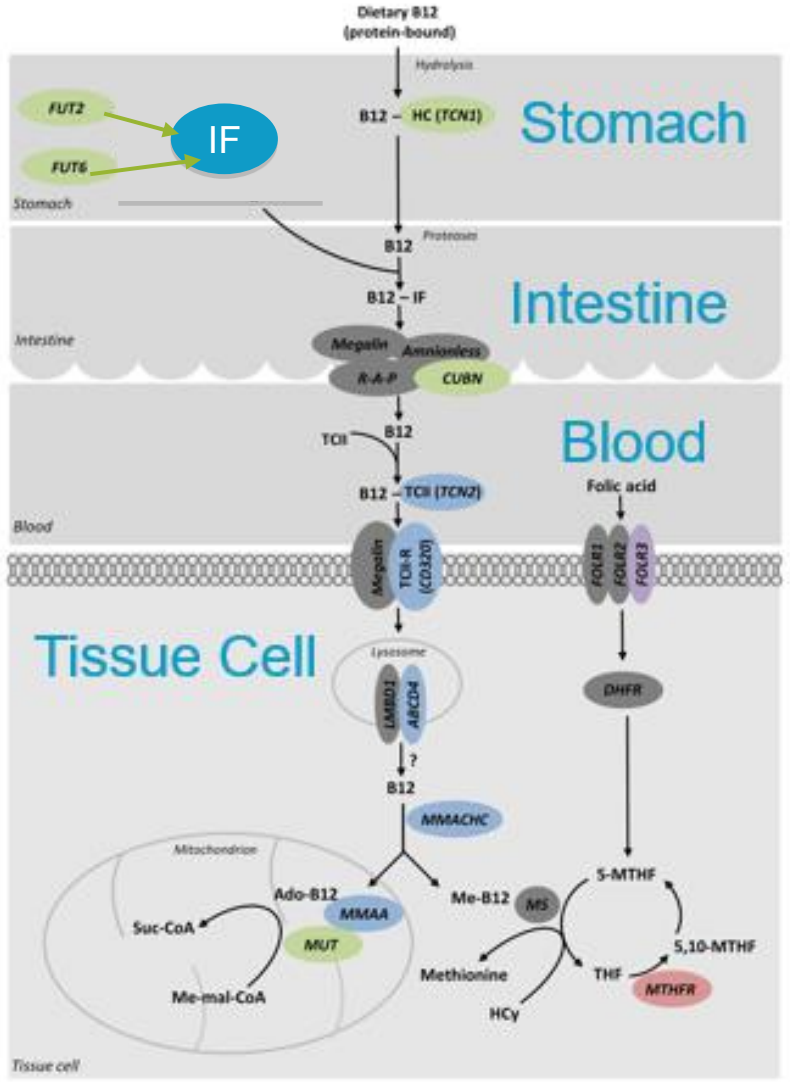


- 13 SNPs capture most of the variability in vitamin B12 levels
- Highest polygenic risk score associated with a app. **104 pg/ml** decrease in blood concentration.
- US population between 156 – 560 pg/ml (median 384 pg/ml). Sufficiency is defined as **400 pg/ml**. Deficiency if concentration is **<250 pg/ml**

Gene variants affect key proteins involved in vitamin B12 absorption



- Low responders with the highest polygenic risk score do not reach sufficiency, even at extreme high (single) doses of B12;
- Depending on the risk SNPs one carries the recommendations can be adapted to ensure efficient supplementation



Source: Grarup et al. Genetic architecture of vitamin B12 and folate levels. PLOS Genetics, 2013

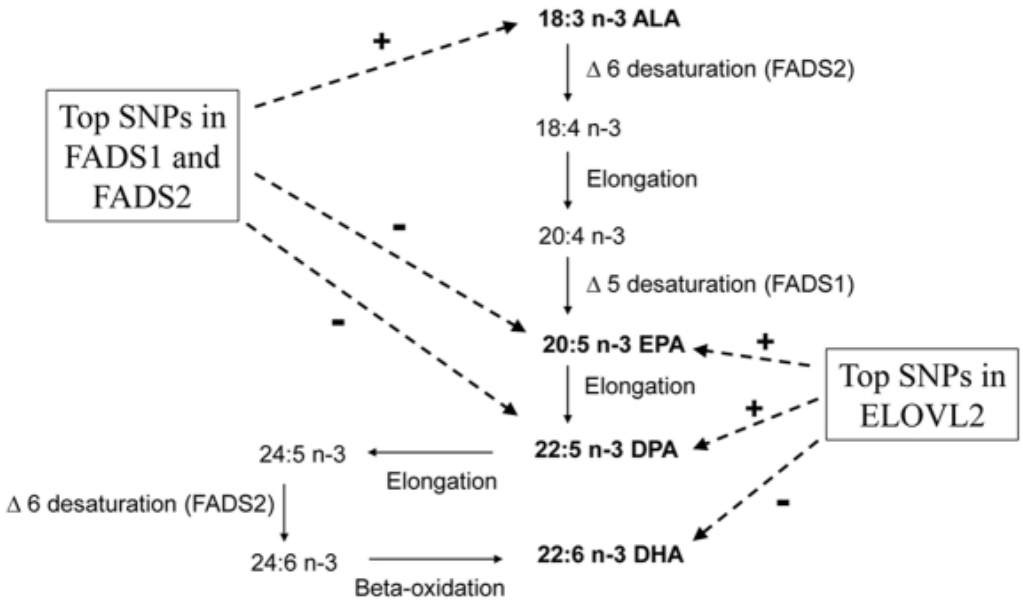
Genetically guided omega-3 supplementation

GWAS have idnetified94 SNPs associated with 9 genes: ELOVL2, FAD1, FAD2, SYCP2L, HERV-FRD, C15orf27, TRPA1, ME3, LRRC4C

Genetic effects on base-line omega-3 levels

Genetic effects on change (delta) for omega-3 Levels after supplementation with DHA

n-3 PUFA metabolic pathway and summary of GWAS in this pathway



What genetic variation can tell us about Nutrition?

Nutrigenetics can inform us about individuals that are at risk to develop nutritional deficiencies AND can estimate the nutritional needs for such an individual

Nutrigenetics can tell us something about geographical differences in nutrient efficiency- possibility to adapt regional recommendations;

Nutrigenetics & Nutrigenomics can help us identify other health related environmental factors that influence a phenotype;

Q & A

Time to take a break….