

Μαριάνθη Γεωργίτση, PhD



Επίκουρη Καθηγήτρια Γενετικής και Μοριακής βάσης ασθενειών

Εργαστήριο Γονιδιωματικής Ποικιλότητας και Γενετικής Επιδημιολογίας

Τμήμα Μοριακής Βιολογίας και Γενετικής - Σχολή Επιστημών Υγείας

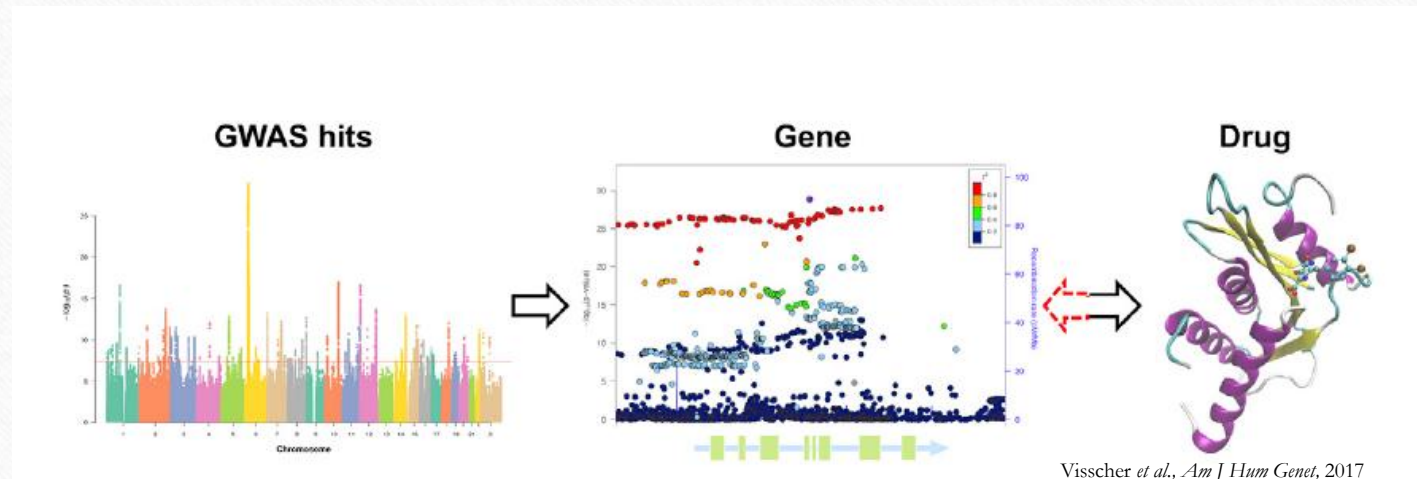
Δημοκρίτειο Πανεπιστήμιο Θράκης

E-mail: mgeorgit@mbg.duth.gr

Prospects and limitations in the use of genetic markers in translational research of complex diseases

«ΜΕΤΑΦΡΑΣΤΙΚΗ ΕΡΕΥΝΑ ΣΤΗ ΒΙΟΙΑΤΡΙΚΗ»

Μοριακή Διαγνωστική, Βιοδείκτες και Στοχευμένες Θεραπείες



Alexandroupoli, 09.05.2025

<https://reporter.nih.gov/>

Search terms: "genetic biomarker"

Parameters: (agency) NIH, (summary by) fiscal year, (plot by) funding



Chart Type

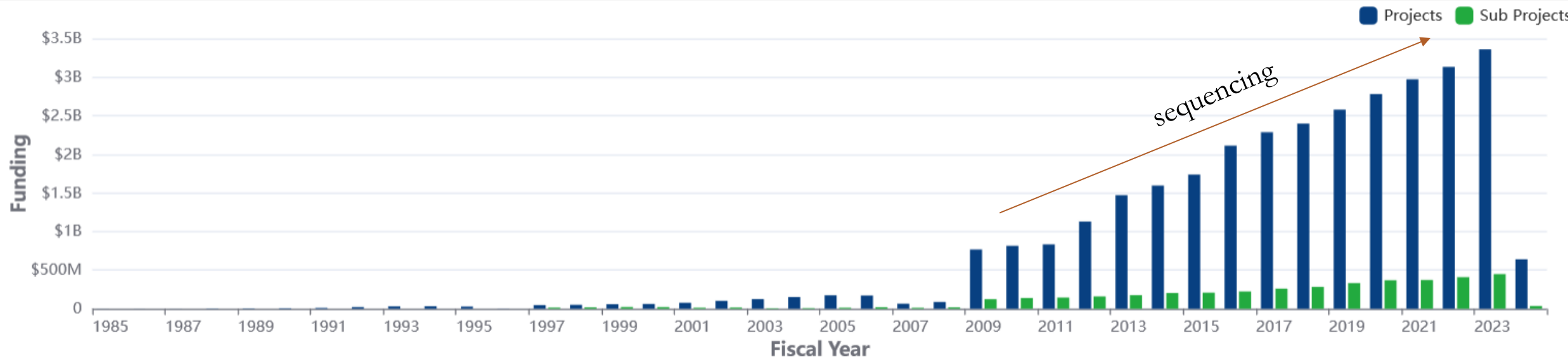
Summary By (Chart X Axis)

Plot By (Chart Y Axis)

Fiscal Year ▼

Projects

✓ Funding



Prospects and limitations in the use of genetic markers in translational research of complex diseases

O U T L I N E

- Introduction (basic concepts, tools, methodological approaches)
- Prospects & Limitations
- Type 2 Diabetes Mellitus
- Examples of translational success in T2D

Βιοδείκτες (Biomarkers)

- Ένας βιοδείκτης (ή βιολογικός δείκτης) είναι ένας μετρήσιμος δείκτης μιας βιολογικής κατάστασης ή ενός βιολογικού προβλήματος.
- Οι βιοδείκτες χρησιμοποιούνται για την αξιολόγηση φυσιολογικών βιολογικών διεργασιών, παθογόνων διεργασιών ή αποκρίσεων σε θεραπευτικές παρεμβάσεις. Μπορεί να είναι μόρια, **γονίδια**, γονιδιακά προϊόντα, ένζυμα ή συγκεκριμένα κύτταρα που παρέχουν πληροφορίες σχετικά με την κατάσταση υγείας ή ασθένειας ενός ατόμου.
- Το Εθνικό Ινστιτούτο Υγείας των ΗΠΑ (NIH) ορίζει ένα βιοδείκτη ως εξής:

«Ένα χαρακτηριστικό που μετράται και αξιολογείται αντικειμενικά ως δείκτης φυσιολογικών βιολογικών διεργασιών, παθογόνων διεργασιών ή φαρμακολογικών αποκρίσεων σε μια θεραπευτική παρέμβαση».

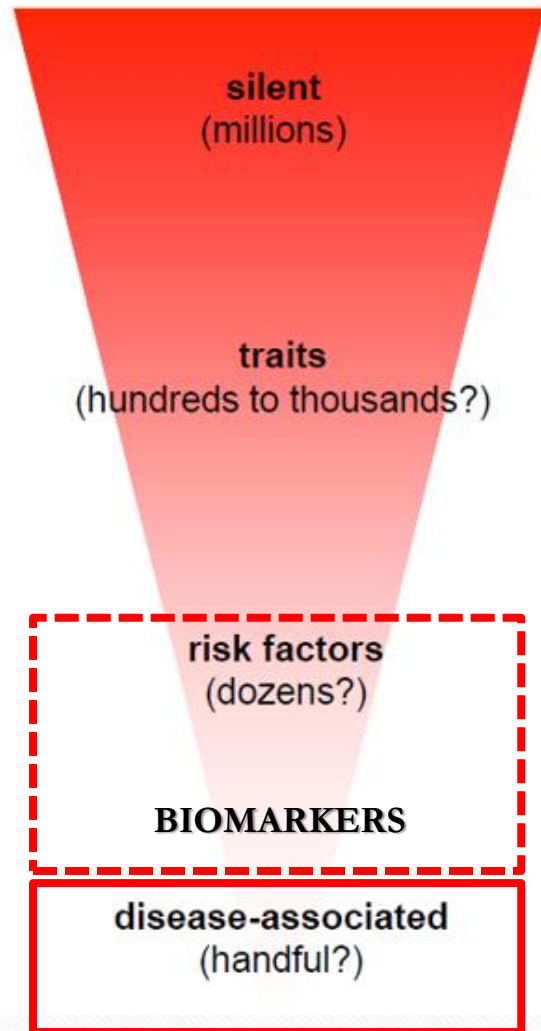
Τύποι βιοδεικτών

Οι βιοδείκτες μπορούν να ταξινομηθούν σε διάφορες **κατηγορίες** με βάση την εφαρμογή τους:

1. **Διαγνωστικοί βιοδείκτες:** Για την **ανίχνευση** ή την **επιβεβαίωση** της παρουσίας μιας νόσου ή κατάστασης.
2. **Προγνωστικοί βιοδείκτες:** Παρέχουν πληροφορίες σχετικά με την **πιθανή πορεία** ή **έκβαση** μιας νόσου.
3. **Προγνωστικοί βιοδείκτες:** Υποδεικνύουν την πιθανότητα **ανταπόκρισης** ενός ασθενούς σε μια συγκεκριμένη θεραπεία.
4. **Φαρμακοδυναμικοί βιοδείκτες/βιοδείκτες απόκρισης:** Χρησιμοποιούνται για να δείξουν ότι έχει εμφανιστεί βιολογική **απόκριση** ως **απάντηση** σε μια θεραπευτική παρέμβαση.
5. **Βιοδείκτες παρακολούθησης:** Χρησιμοποιούνται για την **παρακολούθηση της κατάστασης** μιας νόσου ή των αποτελεσμάτων της θεραπείας με την πάροδο του χρόνου.

Ένας ιδανικός βιοδείκτης πρέπει να έχει τα ακόλουθα **χαρακτηριστικά**:

- ☐ **Ειδικότητα:** Θα πρέπει να είναι **ειδικός για μια συγκεκριμένη** ασθένεια ή κατάσταση.
- ☐ **Ευαισθησία:** Θα πρέπει να είναι σε θέση **να ανιχνεύει ακόμη και μικρές αλλαγές** στη βιολογική κατάσταση.
- ☐ **Αναπαραγωγιμότητα:** Η μέτρηση του βιοδείκτη θα πρέπει να είναι συνεπής σε διαφορετικά εργαστήρια και συνθήκες δοκιμής.
- ☐ **Μη επεμβατικότητα:** Ιδανικά, οι βιοδείκτες θα πρέπει να είναι μετρήσιμοι σε εύκολα προσβάσιμα σωματικά υγρά όπως το αίμα, τα ούρα, το σάλιο ή ο μυελός των οστών.
- ☐ **Κλινική συνάφεια:** Ο βιοδείκτης θα πρέπει να παρέχει σημαντικές πληροφορίες που μπορούν να καθοδηγήσουν τη λήψη κλινικών αποφάσεων.

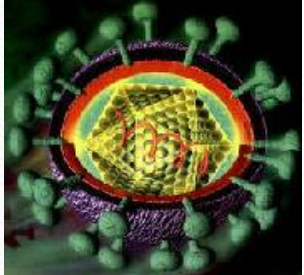


Which are medically relevant?

Genetic Biomarkers

! Genetic risk estimation as the earliest measurable contributor to any disease !

infections



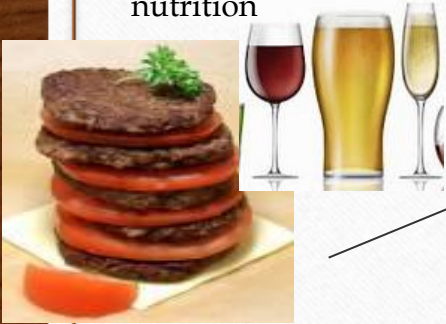
UV



radiation



nutrition



smoking



drugs

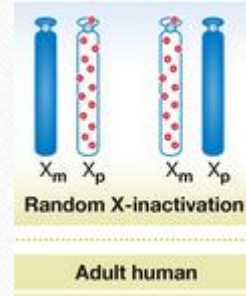


stress

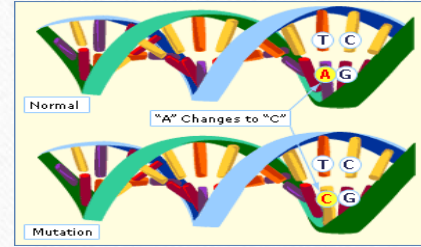


ENVIRONMENT

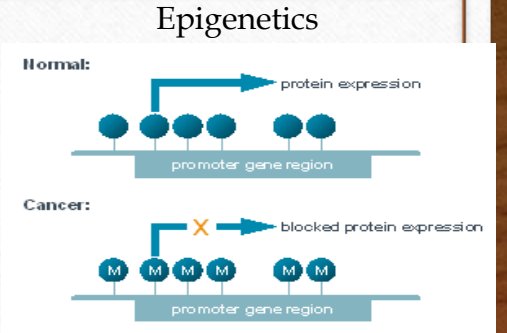
GENETICS & DNA



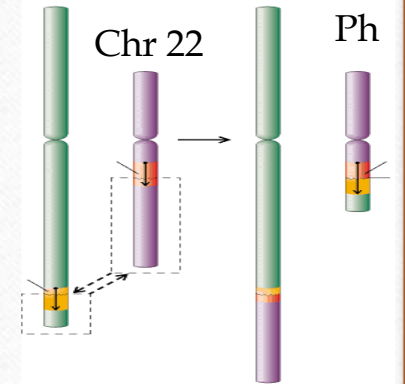
X-chr inactivation
(females)



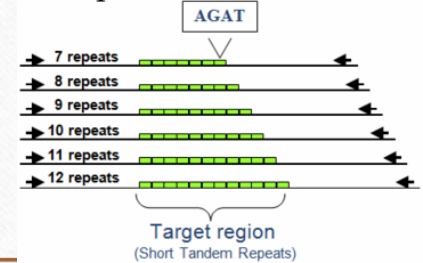
DNA variation



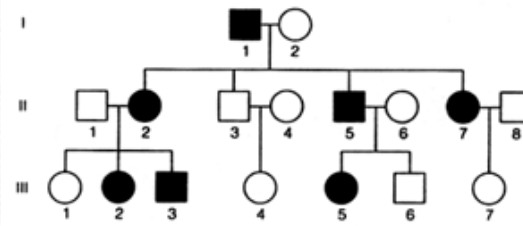
Chromosomal
rearrangements

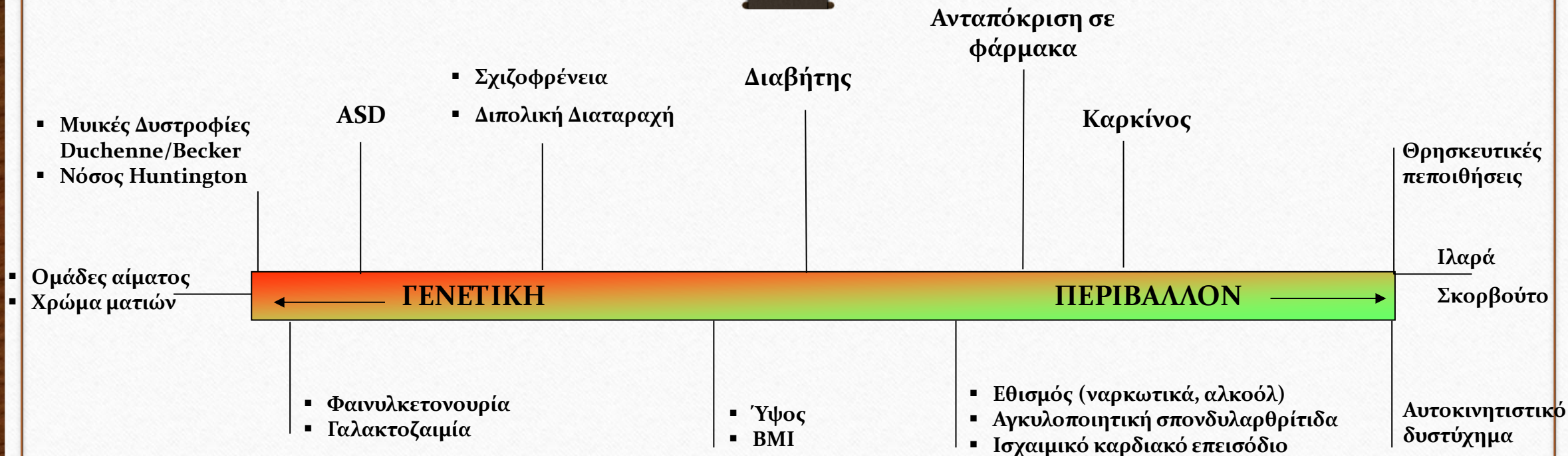


Repetitive DNA

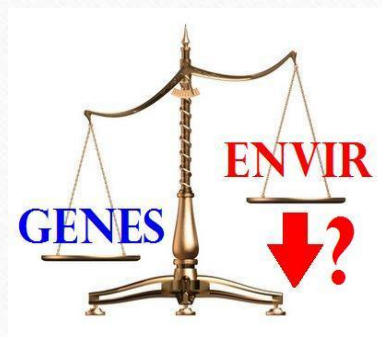


Family history





Rare diseases
Mendelian genetic basis
Mono-/Oligo-genic
“Rare disease-rare variant” hypothesis



“Common disease-common variant” hypothesis

Common diseases
Complex genetic basis
Polygenic

Single-gene disorders

- Low impact on public health cost
- One or a few gene(s)
- Mendelian inheritance (dominant/recessive)
- Rare variants of large effect → **BIOMARKERS**
- Classical genetics approaches
- Examples:
 - Huntington's disease/Myotonic dystrophies
 - Cystic fibrosis
 - Muscular dystrophies (Duchenne/Becker)
 - Rett Syndrome
 - Fragile X
 - Osteogenesis Imperfecta
 - Hereditary cancer syndromes

Multifactorial disorders (Complex traits)

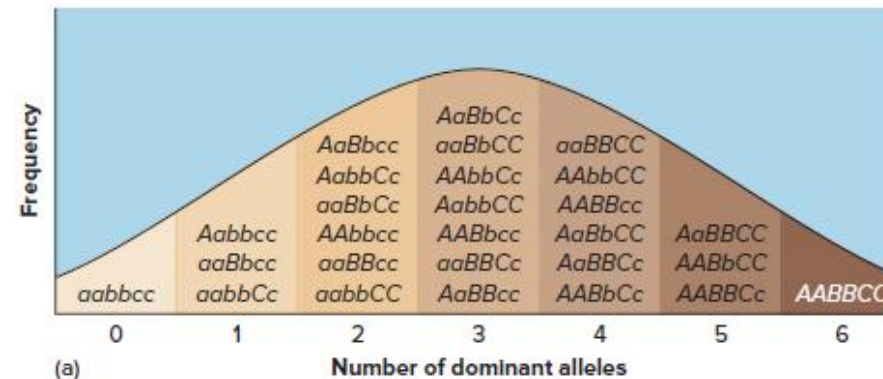
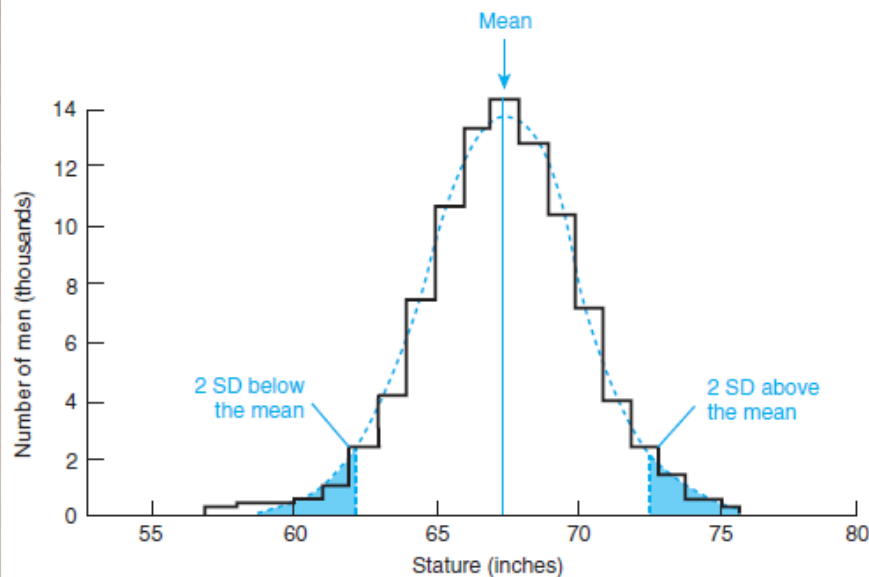
- Serious impact on public health cost
- Multiple genes and loci
- Complex pattern of inheritance (additive)
- Variable heritability (h^2)
- Common and rare genetic variants → **LACK OF BIOMARKERS**
- Whole-genome scans → new technologies/analytical tools
- Examples:
 - Stroke/CVD
 - Diabetes (Type 2)
 - Schizophrenia/Bipolar Disorder/OCD
 - Autism Spectrum Disorder (ASD)/ADHD/Language & Learning Disorders
 - Osteoarthritis
 - Alzheimer's/Dementia



Figure 7.3 The inheritance of height. Genetics students at the University of Notre Dame lined up by height in inches, revealing the continuously varying nature of height. *David Hyde/Wayne Falda/McGraw-Hill Education*

Quantitative traits and Quantitative Trait Loci (QTLs)

“All-or-none” vs “Shades of grey”



(b)

Figure 7.4 Variations in skin color. (a) A model of three genes, with two alleles each, can explain broad hues of human skin. In actuality, this trait likely involves many more than three genes. (b) Humans come in a great variety of skin colors. Skin color genes can assort in interesting ways. These beautiful young ladies, Marcia and Millie, are twins! Their father is Jamaican with dark skin and tight dark curls and their mother is European with fair skin and golden-brown hair. (b): *SWNS/South West News Service Ltd.*

Complex trait heritability

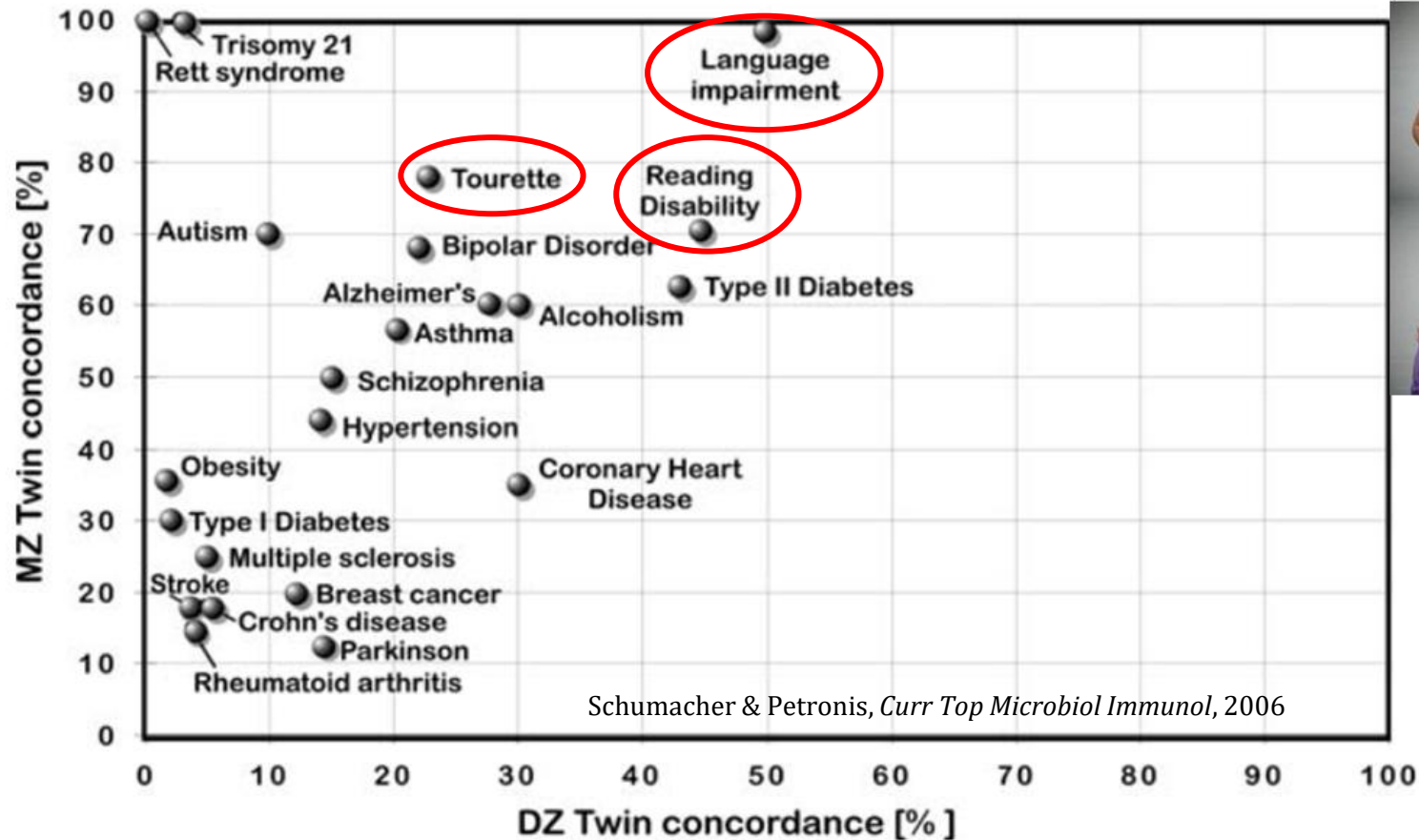


Fig. 1 Concordance of MZ and DZ twins for different disorders. As a rule, the degree of concordance in MZ twins is lower than 100% for nearly all complex diseases but substantially higher in comparison to the concordance rate in DZ twins

$$h^2 = \frac{\text{Variance in DZ pairs} - \text{Variance in MZ pairs}}{\text{Variance in DZ pairs}}$$

Photos by Jason Reed/Reuters



Human Genetics: Concepts and Applications,
McGraw Hill, 13th ed., 2021

$$P = G + E$$

$$V_P = V_G + V_E + V_{G \times E}$$

$$V_P = V_A + \underbrace{V_D + V_I}_{\text{Non-Additive Genetic Variance}} + V_E$$

Phenotypic Variance Additive Genetic Variance Non-Additive Genetic Variance Environmental Variance

$$H^2 = V_G / V_P$$

Broad-sense heritability

$$h^2 = \frac{V_A}{V_P}$$

Narrow-sense heritability

A heritability close to 1 indicates a large portion of the phenotypic variation is due to genetic factors

Heritability

(Κληρονομησιμότητα ή Κληρονομική Ικανότητα)

$$V_G = V_A + V_D + V_I$$

- The total *genetic* variance for a character (V_G) is a function of:
- **Additive genetic variance (V_A)** – variation due to the additive effects of alleles
- **Dominance genetic variation (V_D)** – variation due to dominance relationships among alleles
- **Epistatic genetic variation (V_I)** – variation due to interactions among loci

Complex trait heritability

$$h^2 = \frac{V_A}{V_P}$$

Narrow-sense
heritability

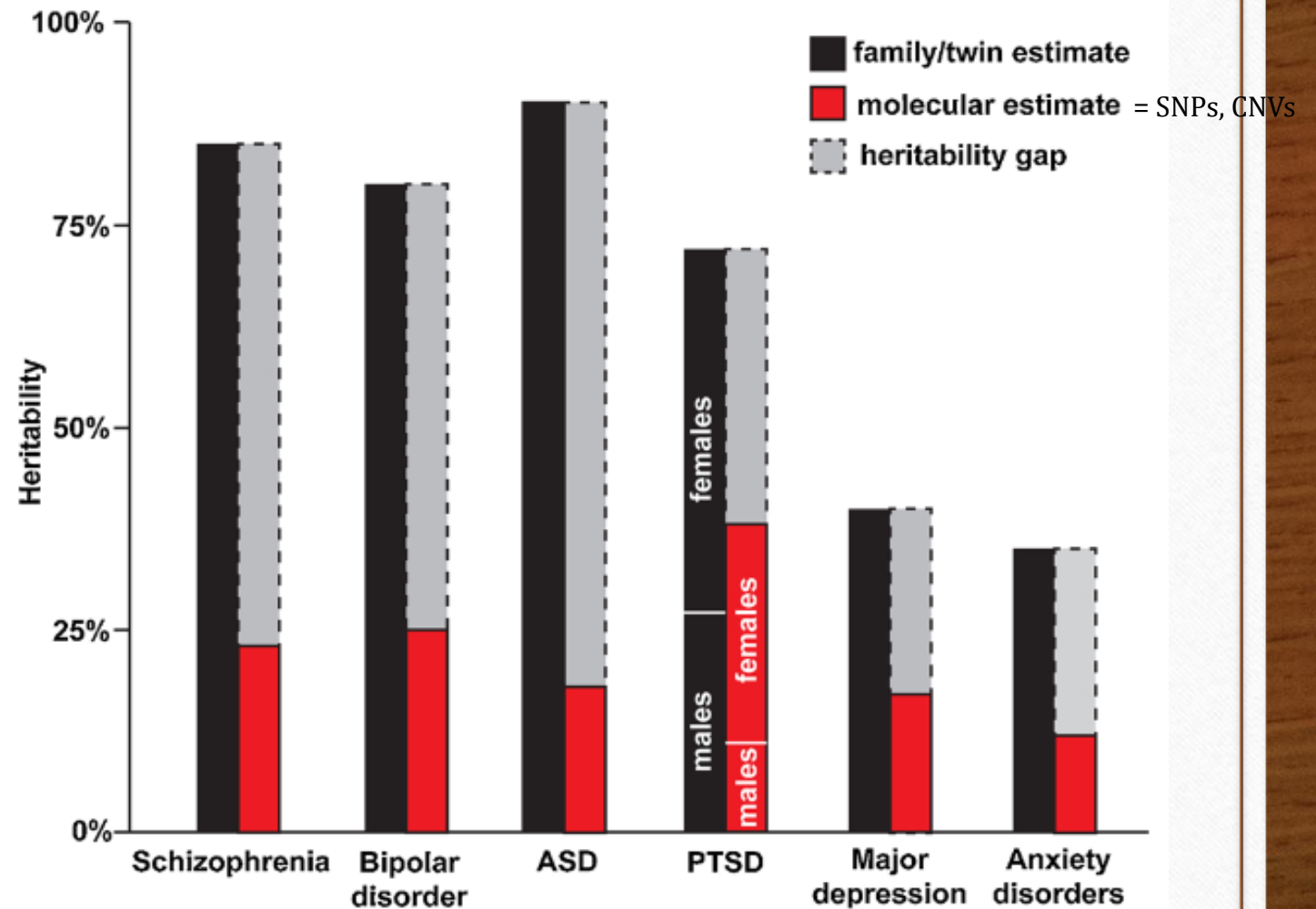
Table 7.2 Heritabilities for Some Human Traits

Trait	Heritability
Clubfoot	0.8
Height	0.8
Blood pressure	0.6
Body mass index	0.4–0.7
Verbal aptitude	0.7
Mathematical aptitude	0.3
Spelling aptitude	0.5
Total fingerprint ridge count	0.9
Intelligence	0.5–0.8
Total serum cholesterol	0.6

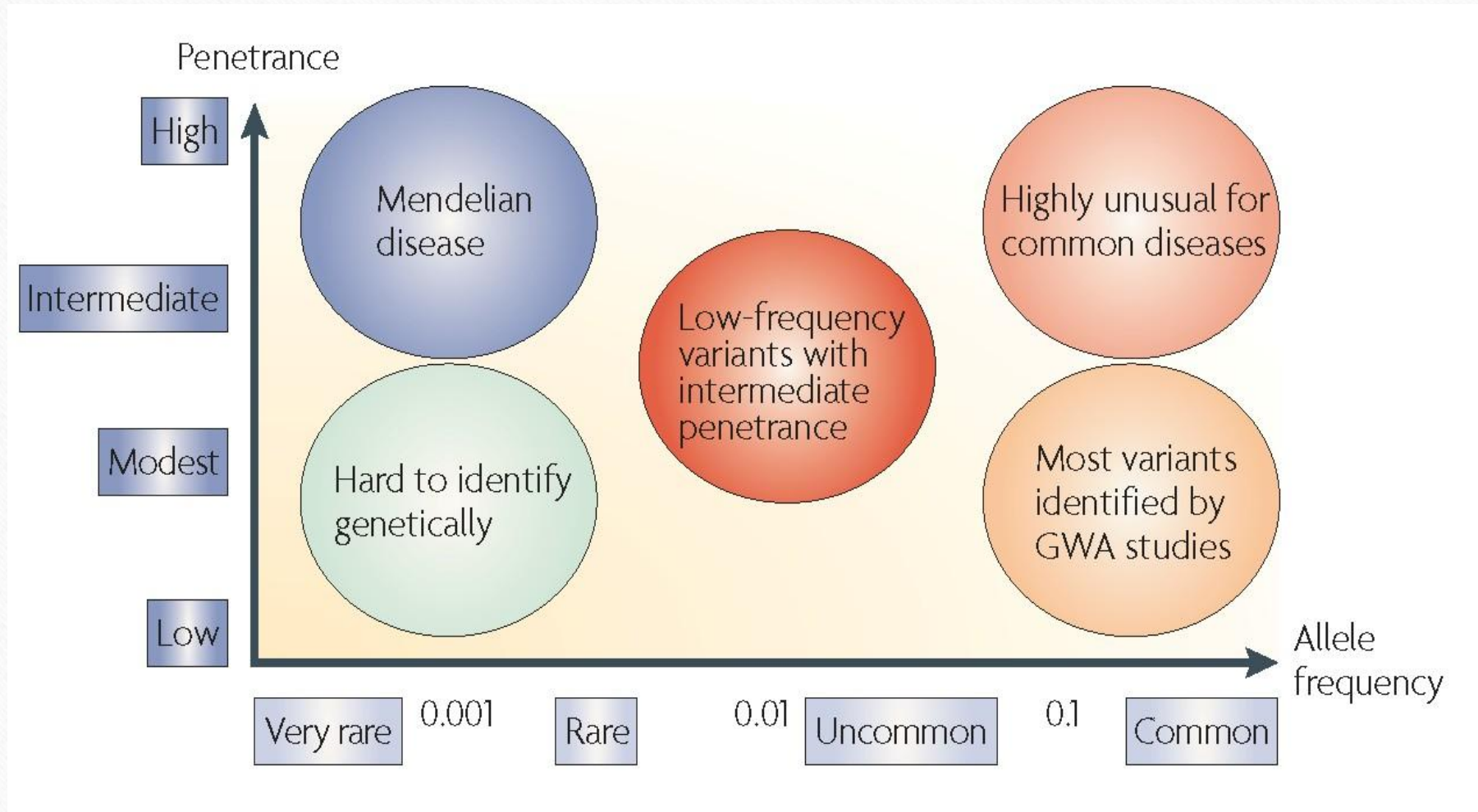
Human Genetics: Concepts and Applications, McGraw Hill, 13th ed., 2021

D. van Calker and T. Serchov

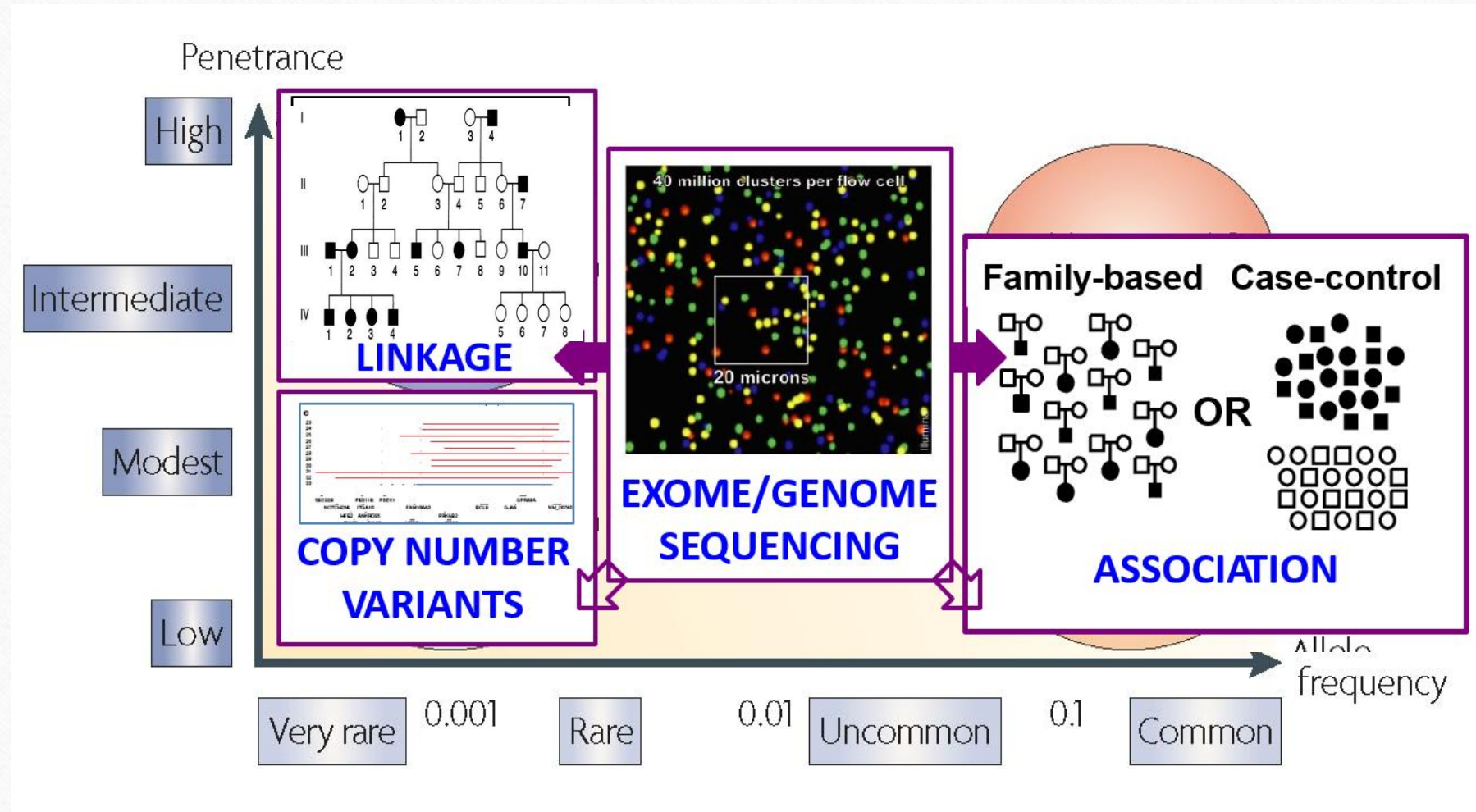
Neuroscience and Biobehavioral Reviews 126 (2021) 23–42

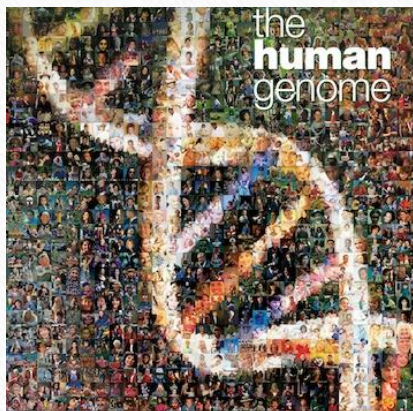


The landscape of human genome variation



The landscape of human genome variation





<https://www.genome.gov/10001772>

NOTE: The first genetic association studies focused on **candidate gene analysis** and, therefore, were not suited for novel genetic risk loci identification.

The completion of the **Human Genome Project**, the **HapMap Project** and **1000Genomes Project**, along with many technological and conceptual advances, have paved the way to array (chip)-based GWAS and Next-Generation Sequencing

→ **High-throughput targeted genotyping and NGS**



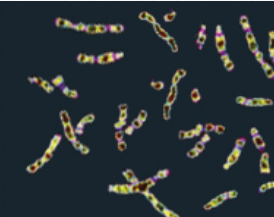
<https://www.genome.gov/10001688/international-hapmap-project>



500k UK people data

1000 Genomes

A Deep Catalog of Human Genetic Variation



<https://www.internationalgenome.org/>

Genome-wide association studies (GWAS)

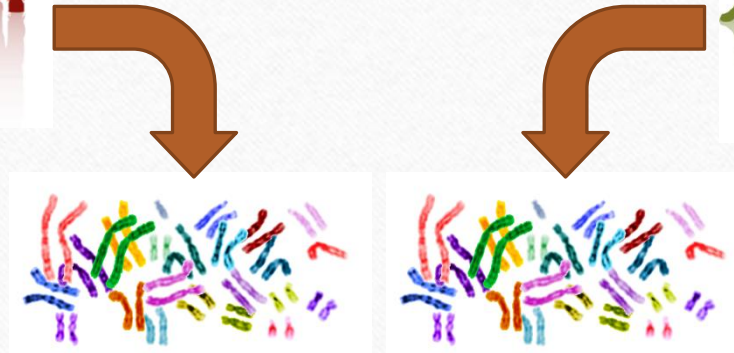
Patients



Controls



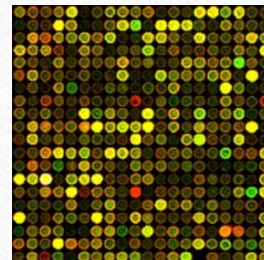
Thousands of samples
– The more the better !



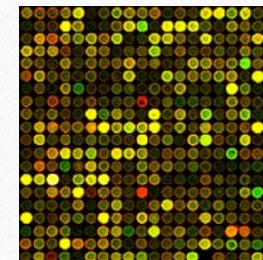
- ✓ Better phenotyping of cases – Inclusion/Exclusion criteria
- ✓ More careful selection of controls
- ✓ Investigators joining forces – Data sharing
- ✓ Funding of large-scale projects (multicentered – multiethnic)
- ✓ International data repositories – Free access

- ✓ Catalogued SNPs (Minor Allele Frequencies – MAFs) and tagSNPs
- ✓ Patterns of linkage disequilibrium (LD) per population
- ✓ Population-specific variation (AIMs)
- ✓ Frameworks to analyze enormous datasets – Development of bioinformatics tools
- ✓ Pathway analyses
- ✓ Meta-analyses

**Genotyping of
eg 100k-6M SNPs**



**Comparison of alleles
–
Statistical analysis**



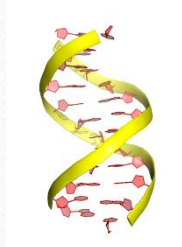
SNPs or SNVs = Single Nucleotide
Polymorphisms (Variants)

Workflow of a typical GWAS study



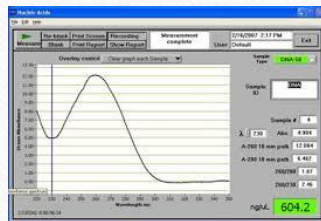
Sample collection

Peripheral blood
Buccal cells
...



DNA extraction and Quantification

Manual purification
Kit-based purification



Fluorespectrophotometry
Spectrophotometry

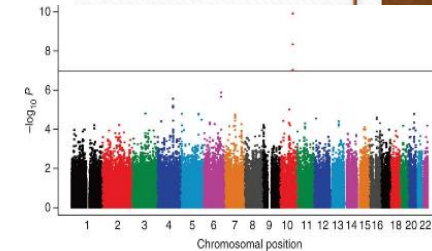
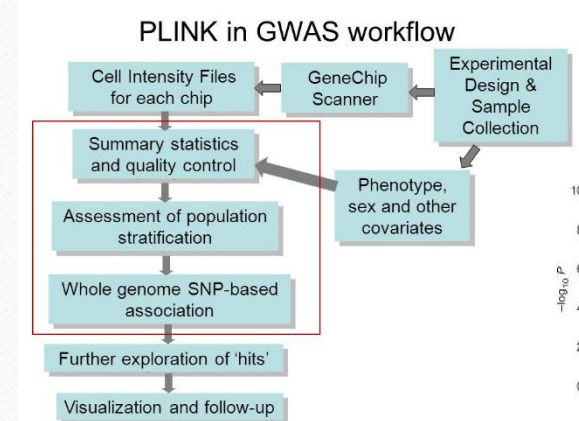


DNA Plating

96-well plates
384-well plates



Array run



Statistical analysis
(trait-dependent)

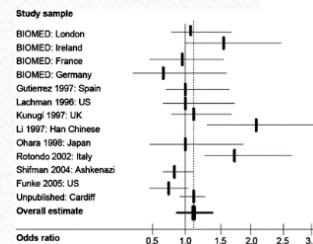
- ANOVA
- Chi-square
- Covariate adjustment
- Multiple testing correction
- Population stratification

Pathway analyses

Replication study
(stage 2)

- Sample size (large)
- Population
- Phenotypic criteria

Meta-analyses

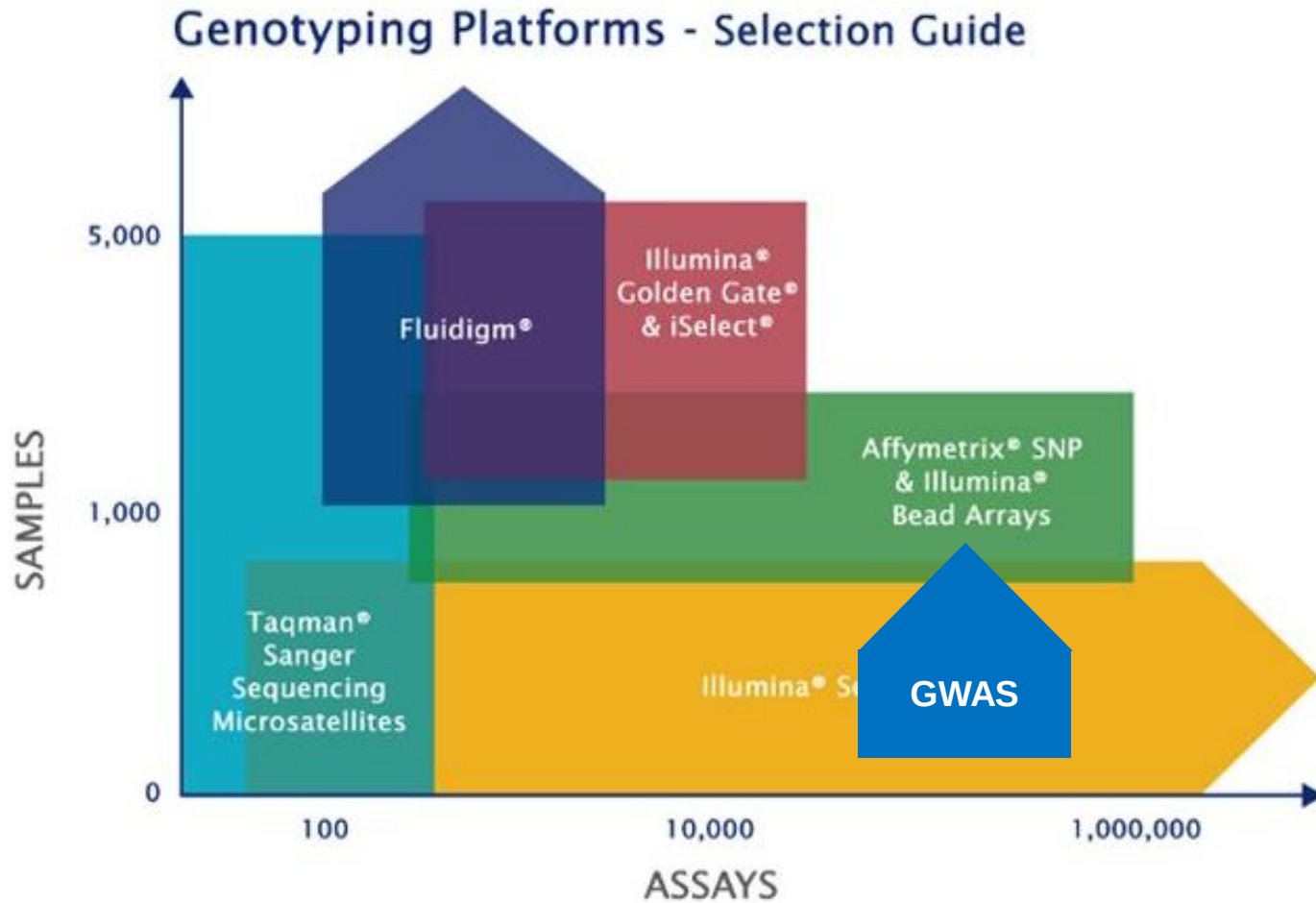


Validation
(technical, functional)

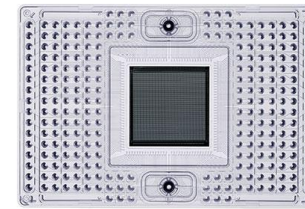
- Other genotyping platform
- In silico analysis
- Sequencing
- Gene expression analysis
- Cell-based analysis
- Animal models

Προς την ανάπτυξη τεχνικών γονιδιωματικής κλίμακας

1. Κόστος
2. Στόχος
3. Εργαλεία ανάλυσης



<http://www.lifesciences.sourcebioscience.com/media/426305/genotyping%20image.png>



Juno 96.96 Genotyping IF



WES
WGS

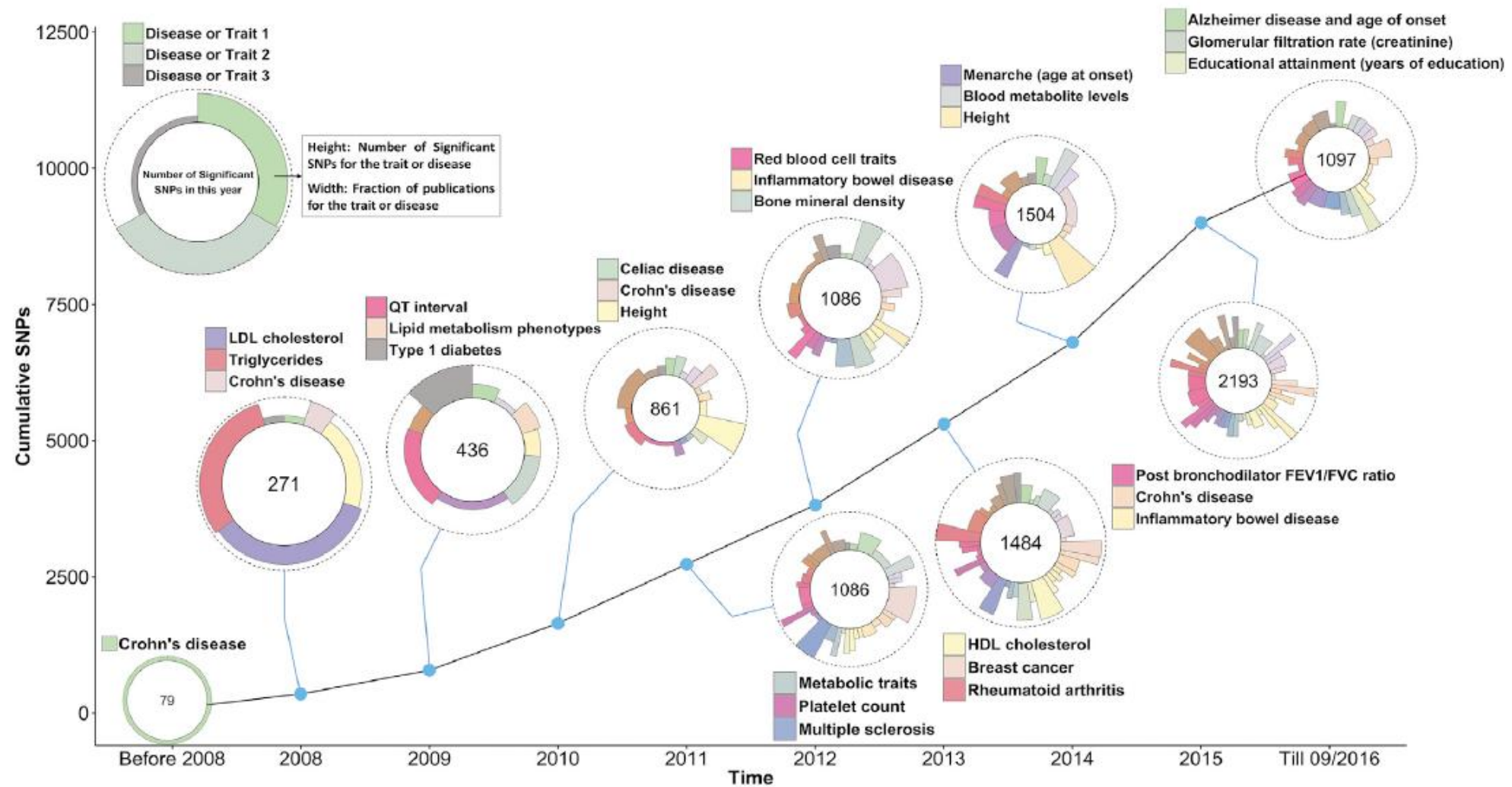


Figure 2. GWAS SNP-Trait Discovery Timeline

Data used for generating the graph were taken from the GWAS Catalogue.¹⁰ SNPs and traits were selected according to the following filters. SNPs were selected with a p value $< 5 \times 10^{-8}$. For each trait with two or more selected SNPs, SNPs were removed if they had an $LD\ r^2 > 0.5$ (calculated from 1000 Genomes phase 3 data) with another selected SNPs and their p value was larger. For each year of discovery, only the top three traits and diseases with the largest number of SNPs are labeled in the circle.

Genome-wide association studies (2005 – present day)

- ✓ Καλύτερος φαινοτυπικός χαρακτηρισμός των περιστατικών – Κριτήρια ένταξης/Αποκλεισμού
- ✓ Προσεκτικότερη επιλογή μαρτύρων
- ✓ Σύμπραξη ερευνητών – Διαμοιρασμός δεδομένων
- ✓ Χρηματοδότηση ερευνητικών έργων μεγάλης κλίμακας (πολυκεντρικές – πολυεθνικές μελέτες)
- ✓ Θέσπιση και λειτουργία διεθνών αποθετηρίων δεδομένων – Ελεύθερη (?) πρόσβαση ερευνητών

- 
- Συλλογή δειγμάτων
 - Θέσπιση και δραστηριότητα ερευνητικών κοινοπραξιών

- ✓ Καταλογογράφηση SNPs
- ✓ Πρότυπα ανισορροπίας σύνδεσης (Linkage Disequilibrium – LD)
- ✓ Ποικιλομορφία γονιδιώματος σε πληθυσμιακό επίπεδο (Ancestry-Informative Markers – AIMs)
- ✓ Πλαίσια (pipelines) ανάλυσης δεδομένων μεγάλου όγκου – Ανάπτυξη εργαλείων βιοπληροφορικής ανάλυσης
- ✓ Αναλύσεις μονοπατιών
- ✓ Μεταanalύσεις - PRSs

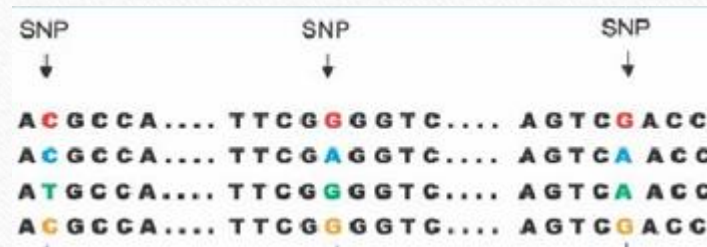
- 
- Γενετική αρχιτεκτονική
 - Αναλύσεις
 - Βιοπληροφορική

Γενετική Αρχιτεκτονική

Πληθυσμός



Γενετικές παραλλαγές



Είδος (ποιες)

Αριθμός (πόσες)

Συχνότητα
(πόσο συχνές)

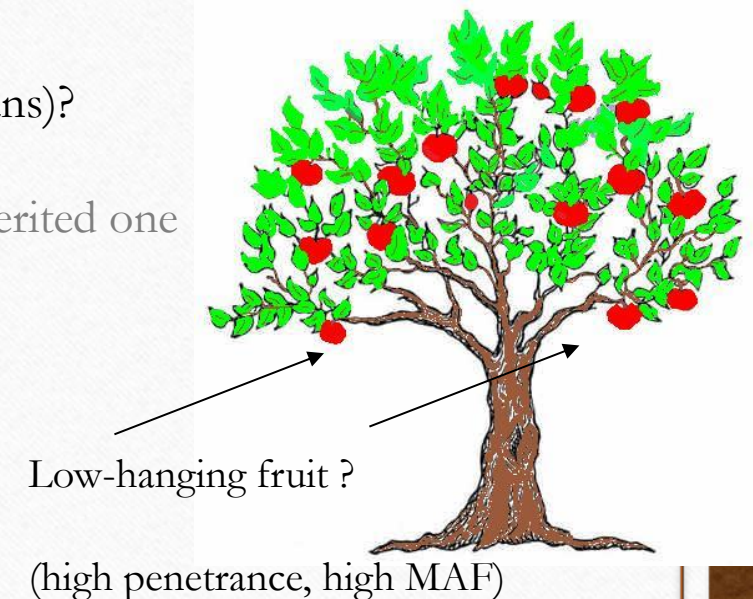
Αποτέλεσμα (πόσο σημαντικές)

Linkage Disequilibrium

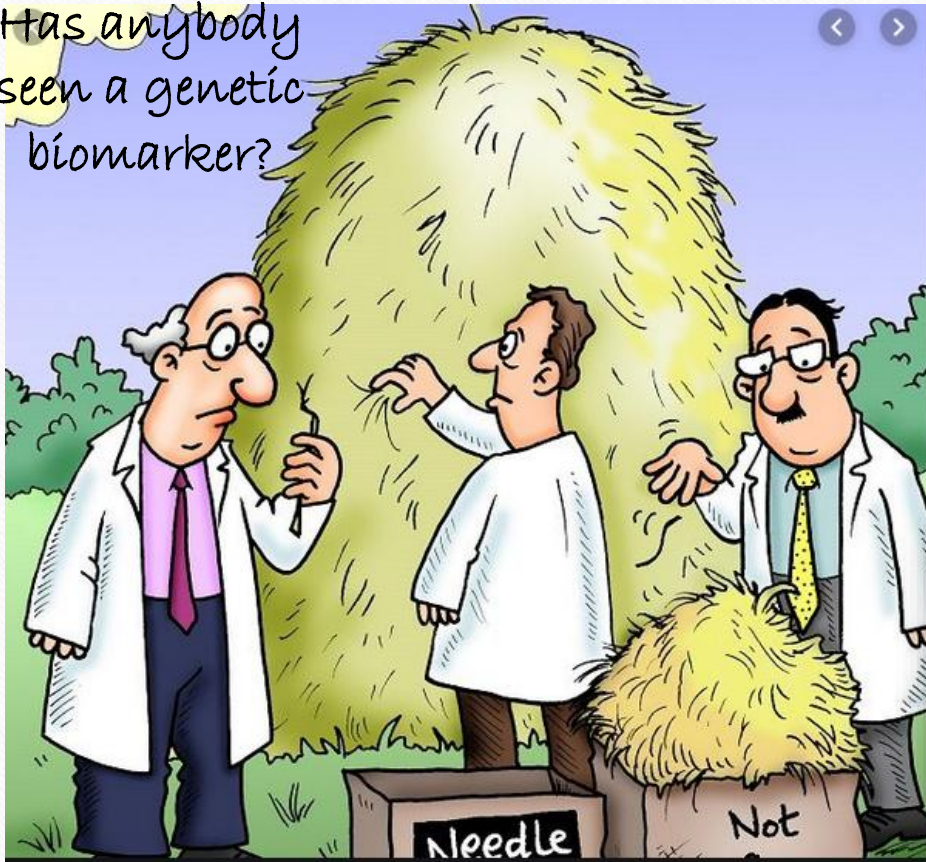
(με ποιες άλλες)

GWAS limitations

- Only associations revealed – Not causation (Positive or negative correlation may imply underlying mechanisms, but is not proof that A actually causes B)
- Sex chromosomes (X, Y) omitted from analysis (until recently)
- Patient selection (population-based vs selected group) and controls selection (are they truly free of any disease? → genetic correlation between disorders/cross-disorder analyses)
- Can results from one ancestry (eg Caucasians) be extrapolated to another (eg Asians)?
- Distinguish true cases from phenocopies (=a trait or illness that resembles an inherited one but is attributed to an environmental cause)
- Unable to uncover tissue-specific *gene × gene* interactions



Has anybody
seen a genetic
biomarker?



We should not expect common variants to have large effects, for evolutionary reasons.

Large effects are bad, as most variants with large effects are consequently selected against, thus never become fixed and common.

Missing heritability

(ελλείπουσα κληρονομησιμότητα)

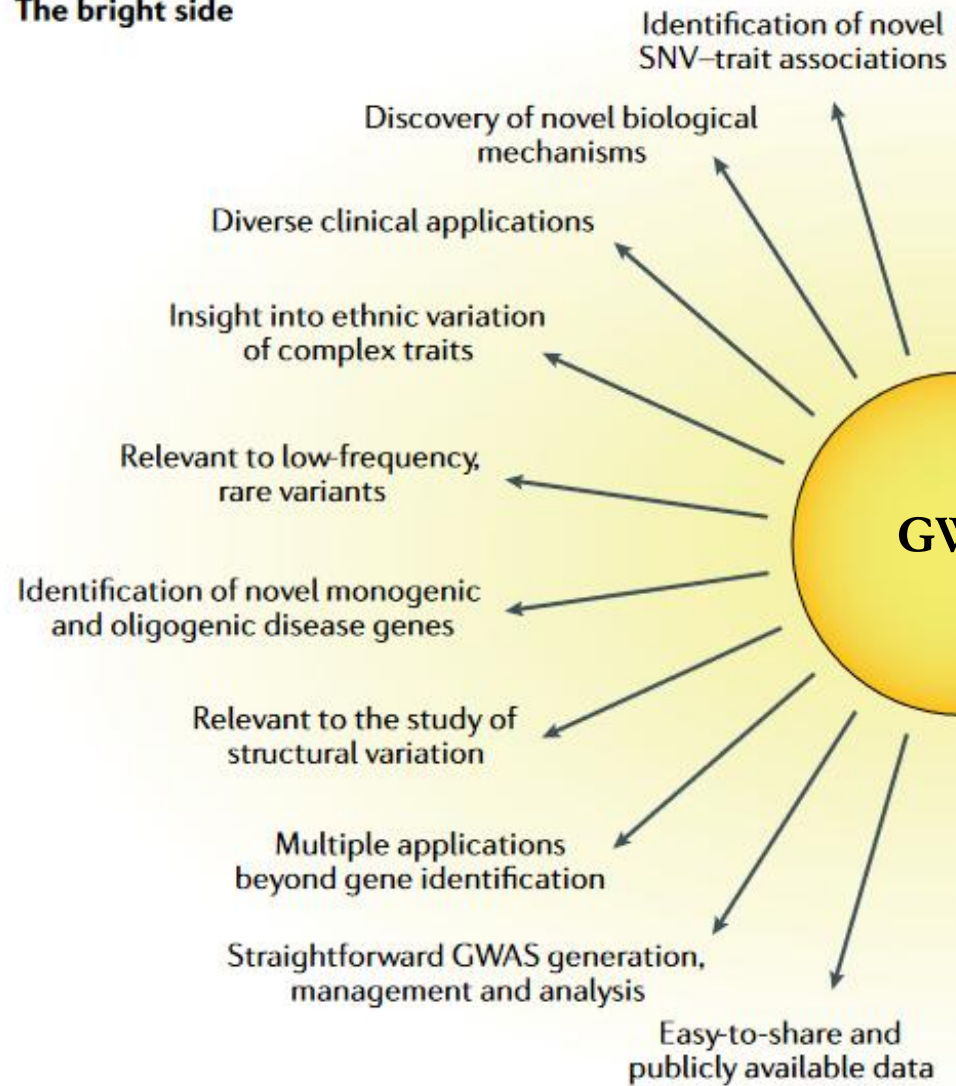
Rare variants (MAF<1%)

- Structural variants (CNVs)
- *Gene × gene* interactions (epistasis)
- Epigenetics
- *Gene × environment* interactions
- Something unknown?

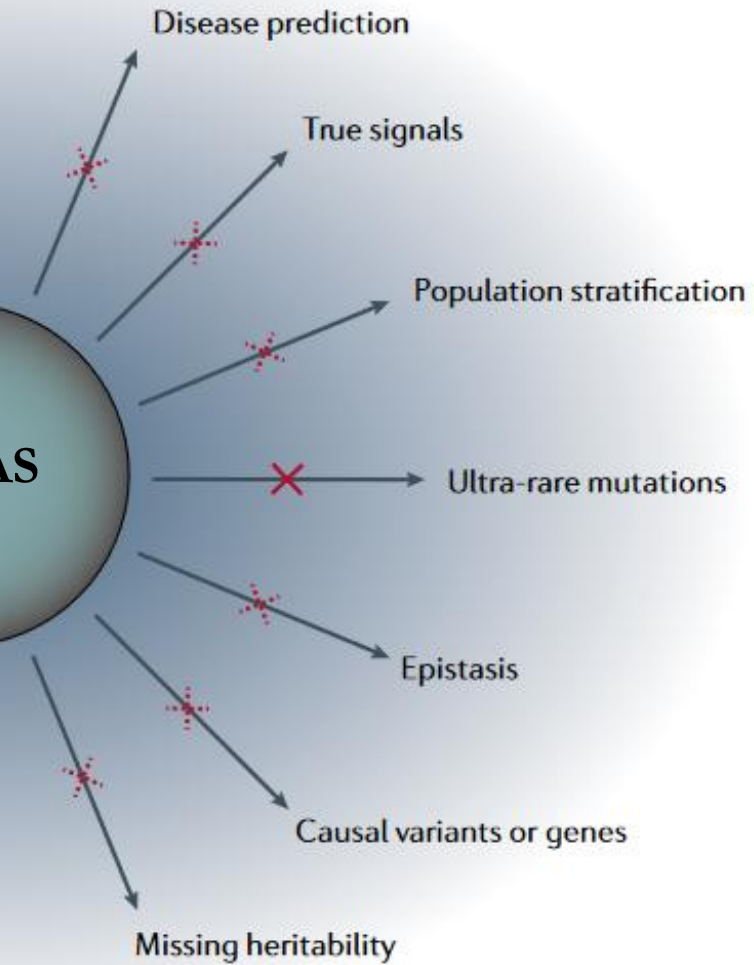
‘There could be scarier and more intractable reasons for unaccounted-for heritability that are not even being discussed. “It’s a possibility that there’s something we just don’t fundamentally understand,” Kruglyak says. “That it’s so different from what we’re thinking about that we’re not thinking about it yet.”

Still the mystery continues to draw its sleuths, for Kruglyak as for many other basic-research scientists. “You have this clear, tangible phenomenon in which children resemble their parents,” he says. “Despite what students get told in elementary-school science, we just don’t know how that works.”

The bright side



The dark side

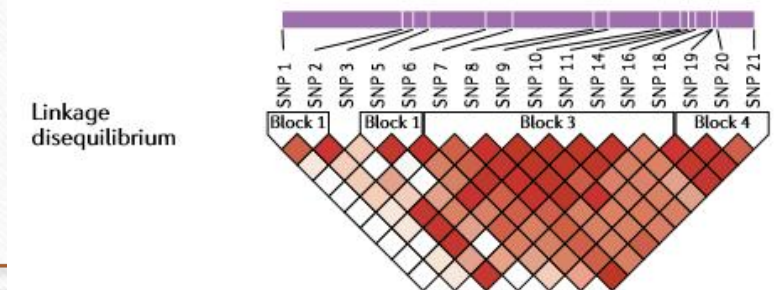
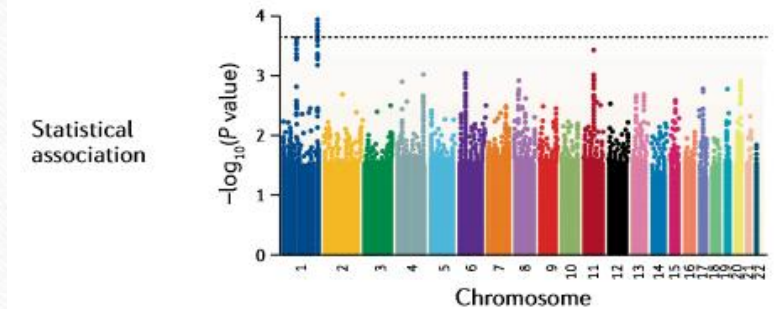
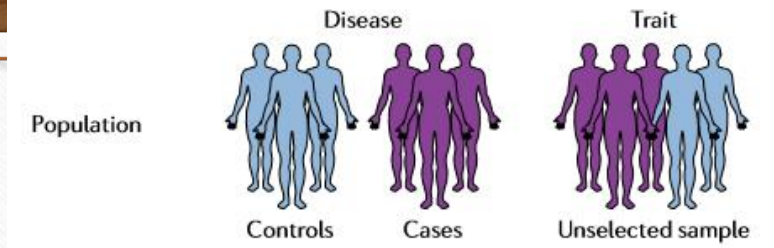
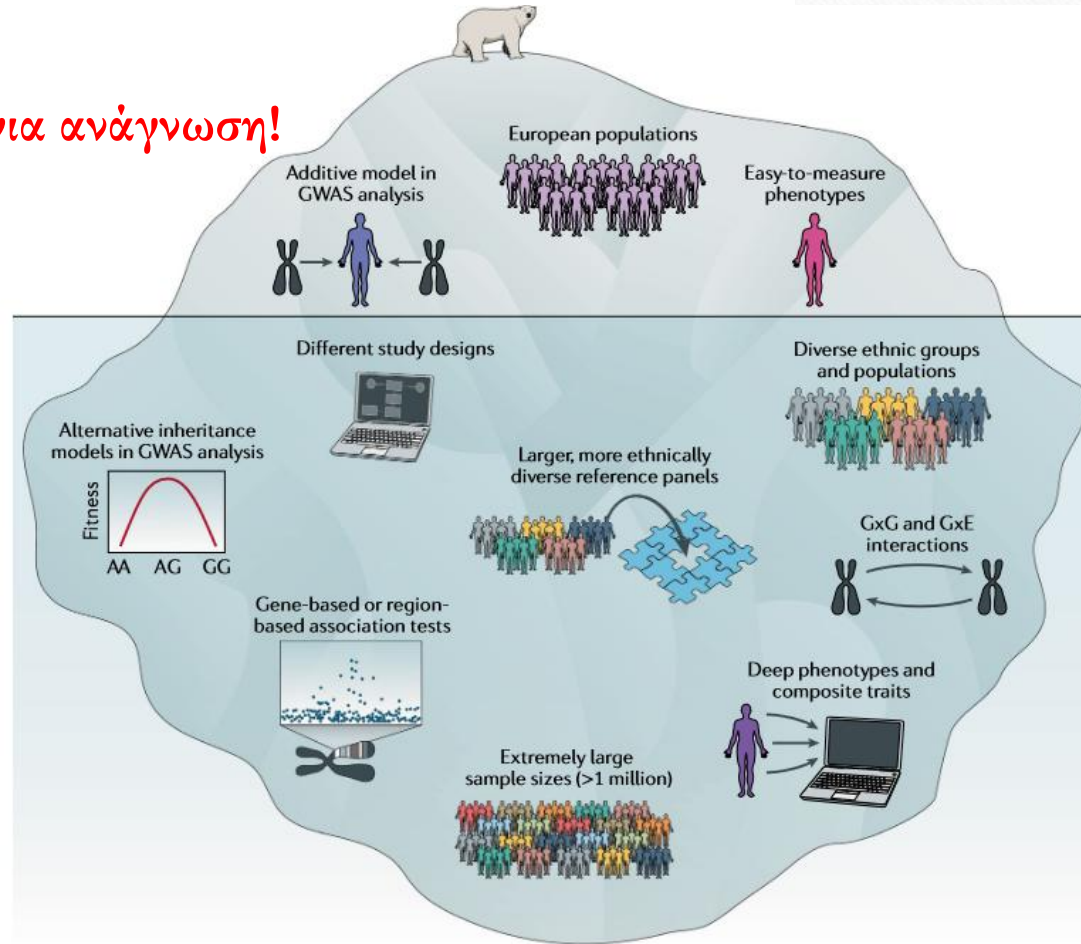


Tam et al., *Nat Rev Genet*, 2019

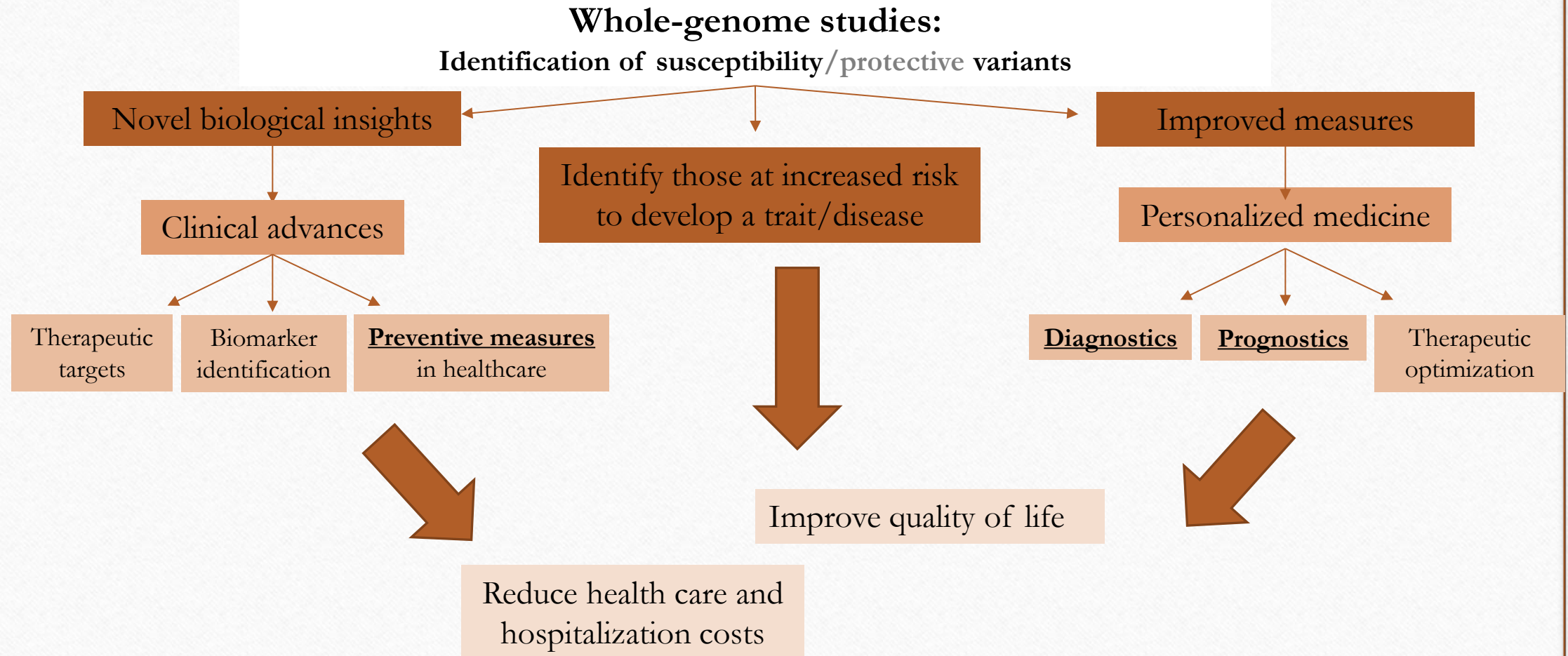
Benefits and limitations of genome-wide association studies

Vivian Tam¹, Nikunj Patel¹, Michelle Turcotte¹, Yohan Bossé^{2,3}, Guillaume Paré^{1,4}
and David Meyre^{1,4,5} *

Ισχυρή σύσταση για ανάγνωση!

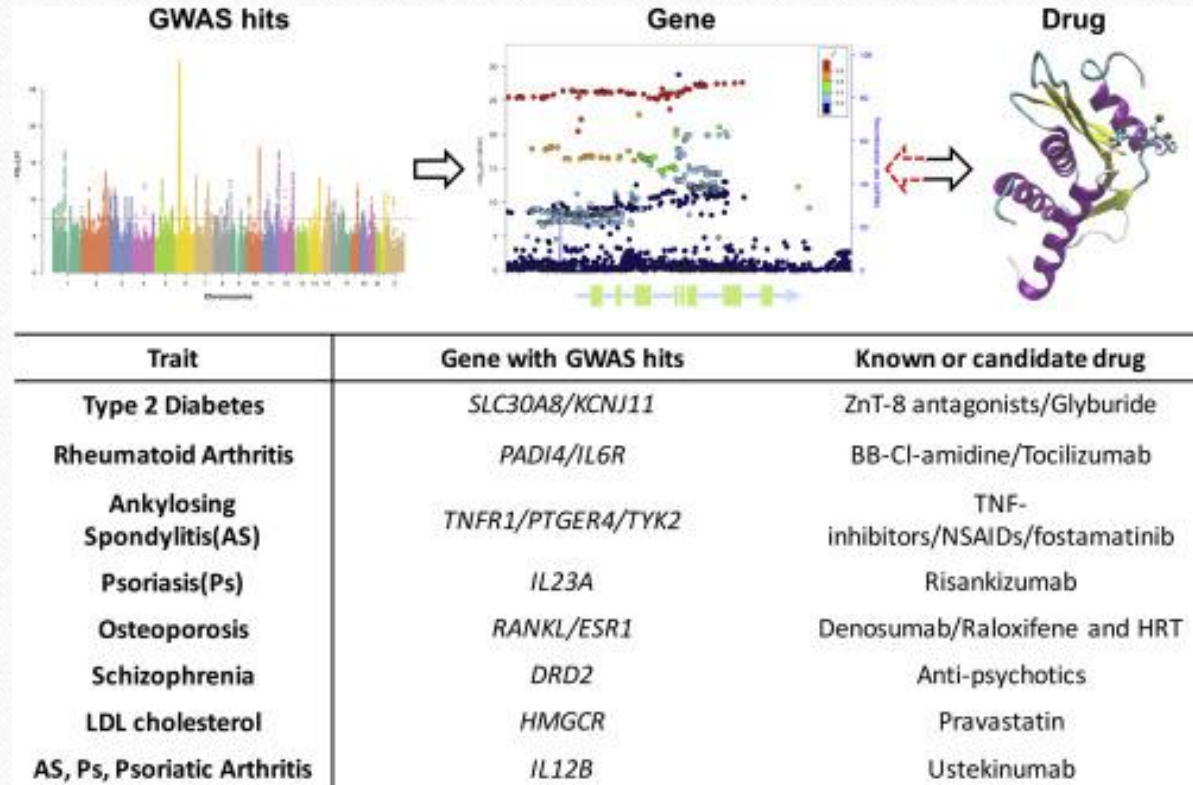


What whole-genome studies have to offer – Translational research



Examples of links between GWAS discoveries and drugs:

Genetically-informed translational research/Drug development



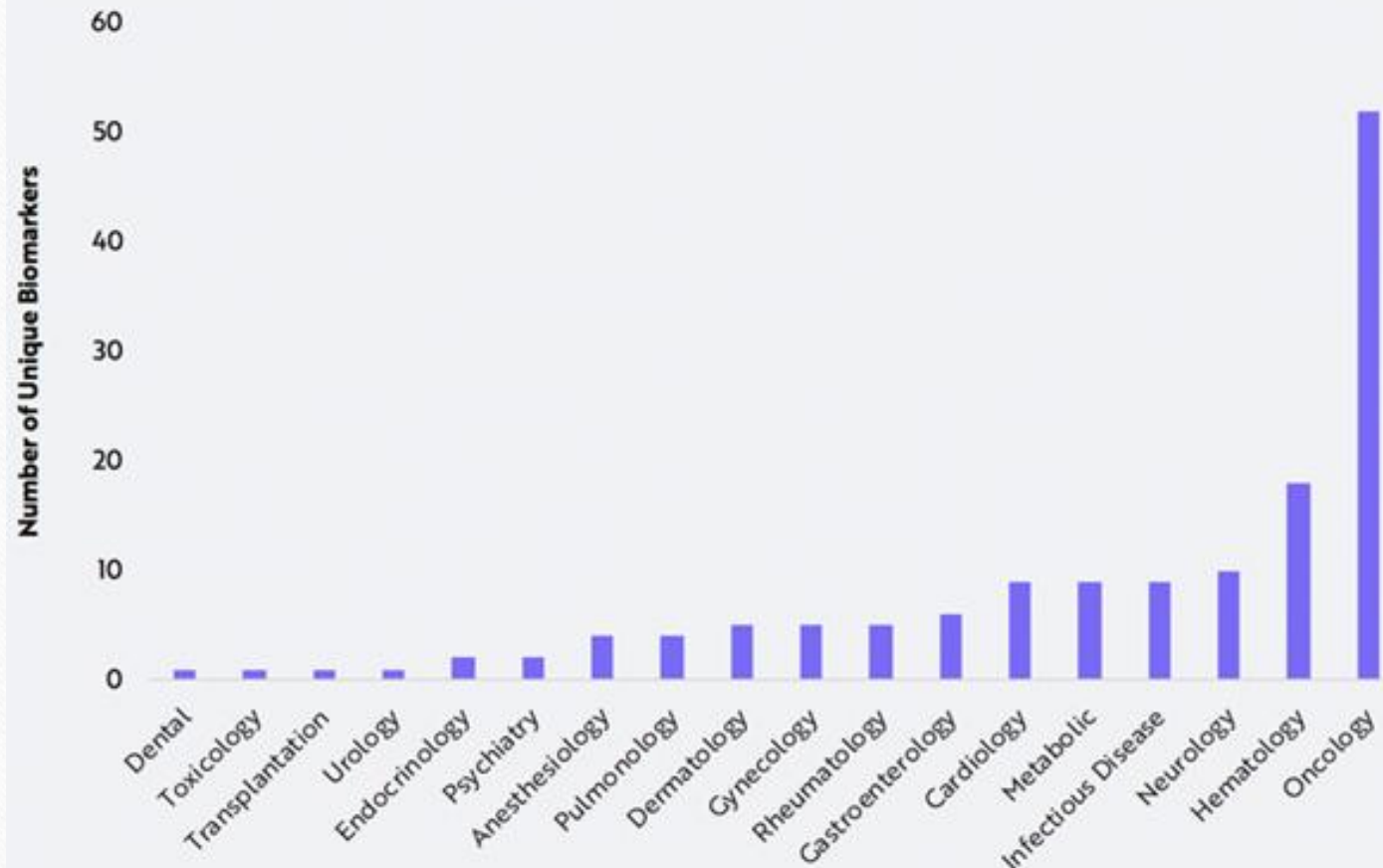
Identifying genetic markers with translational potential is crucial for medical **fields where:**

- a) **Diagnostic criteria are not based on biological markers**, such as psychiatry and behavioural traits, and
- b) There is **lack of effective treatment**
- c) The **phenotype** is quite **heterogeneous**
- d) Complex diseases are often **late-onset**, thus predisposed individuals could benefit from early detection and preventive testing/medical follow-up

Genetically-informed translational research/Drug development

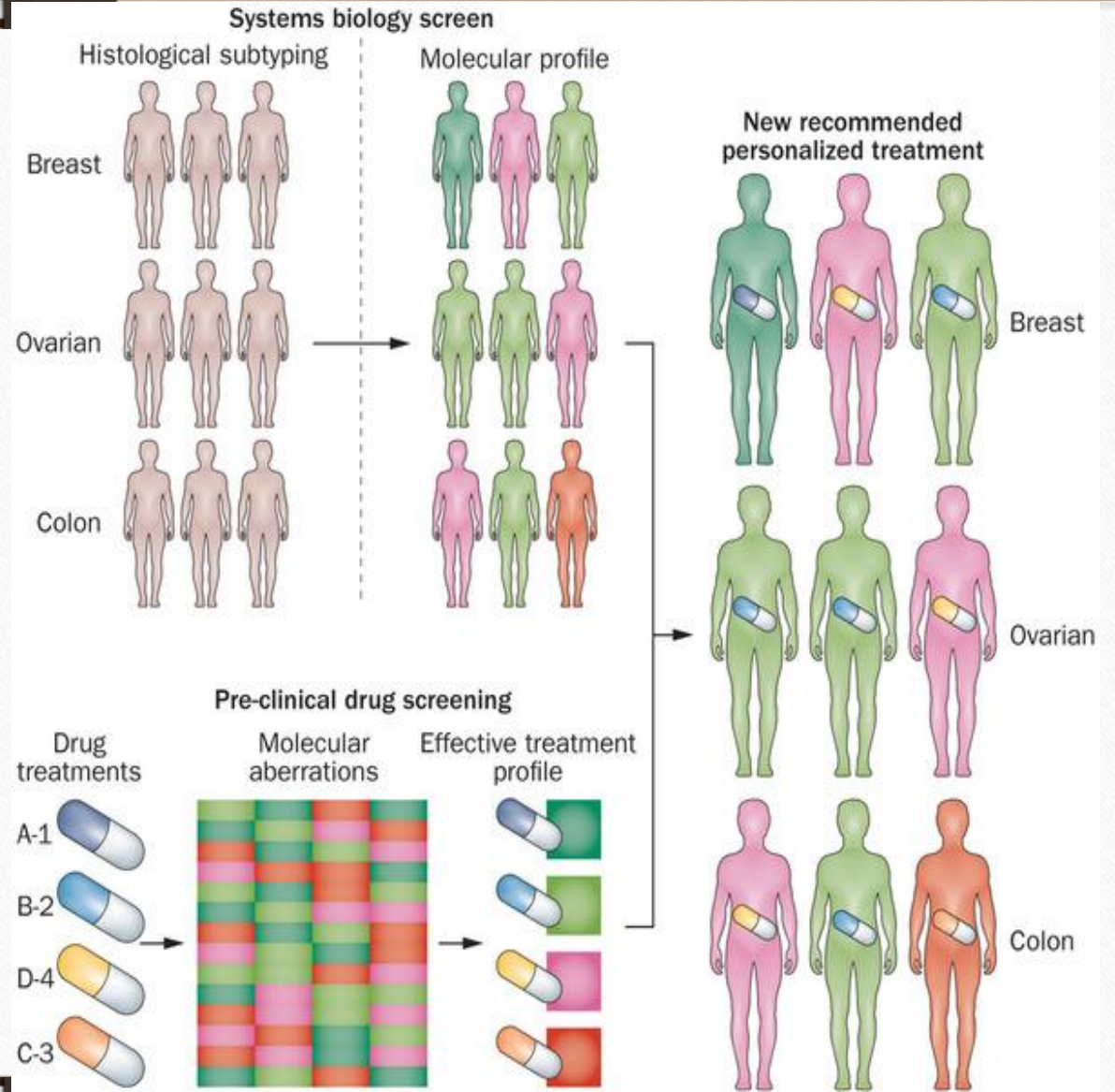
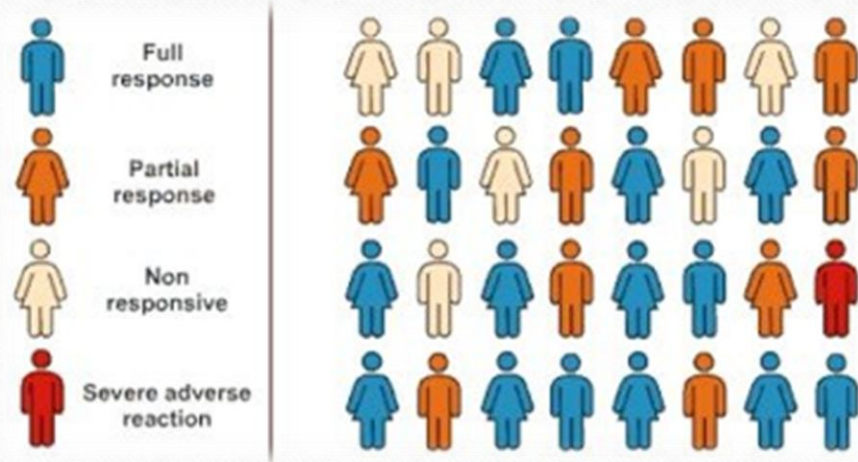
Distribution of Genetic Biomarkers With an FDA-Approved Therapy

Data Source: <https://www.fda.gov/drugs/science-and-research-drugs/table-pharmacogenomic-biomarkers-drug-labeling>



Genetically-informed translational research/ Drug development

Pharmacogenomics

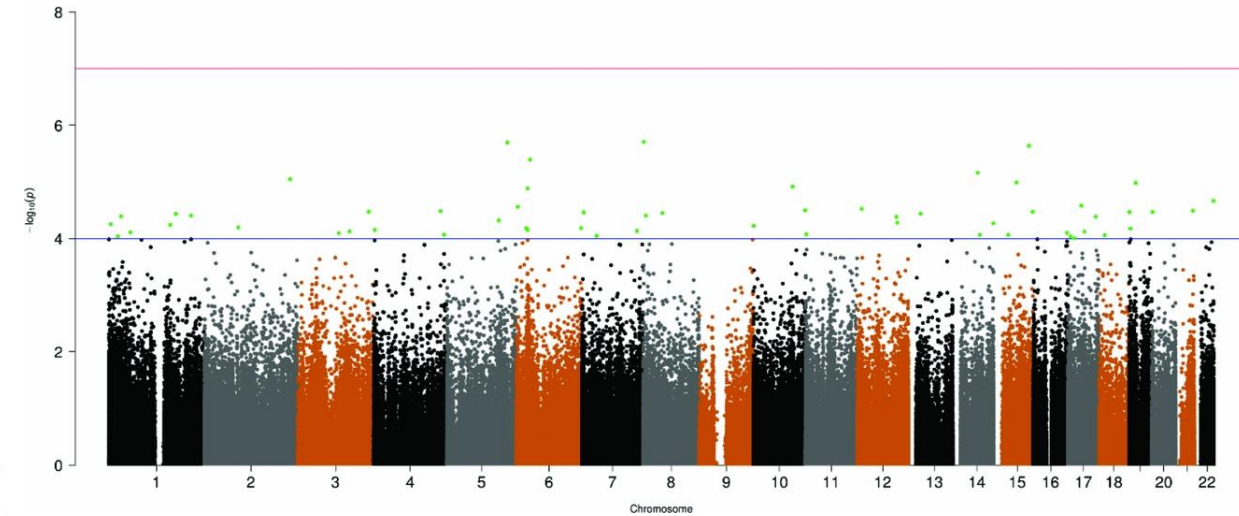
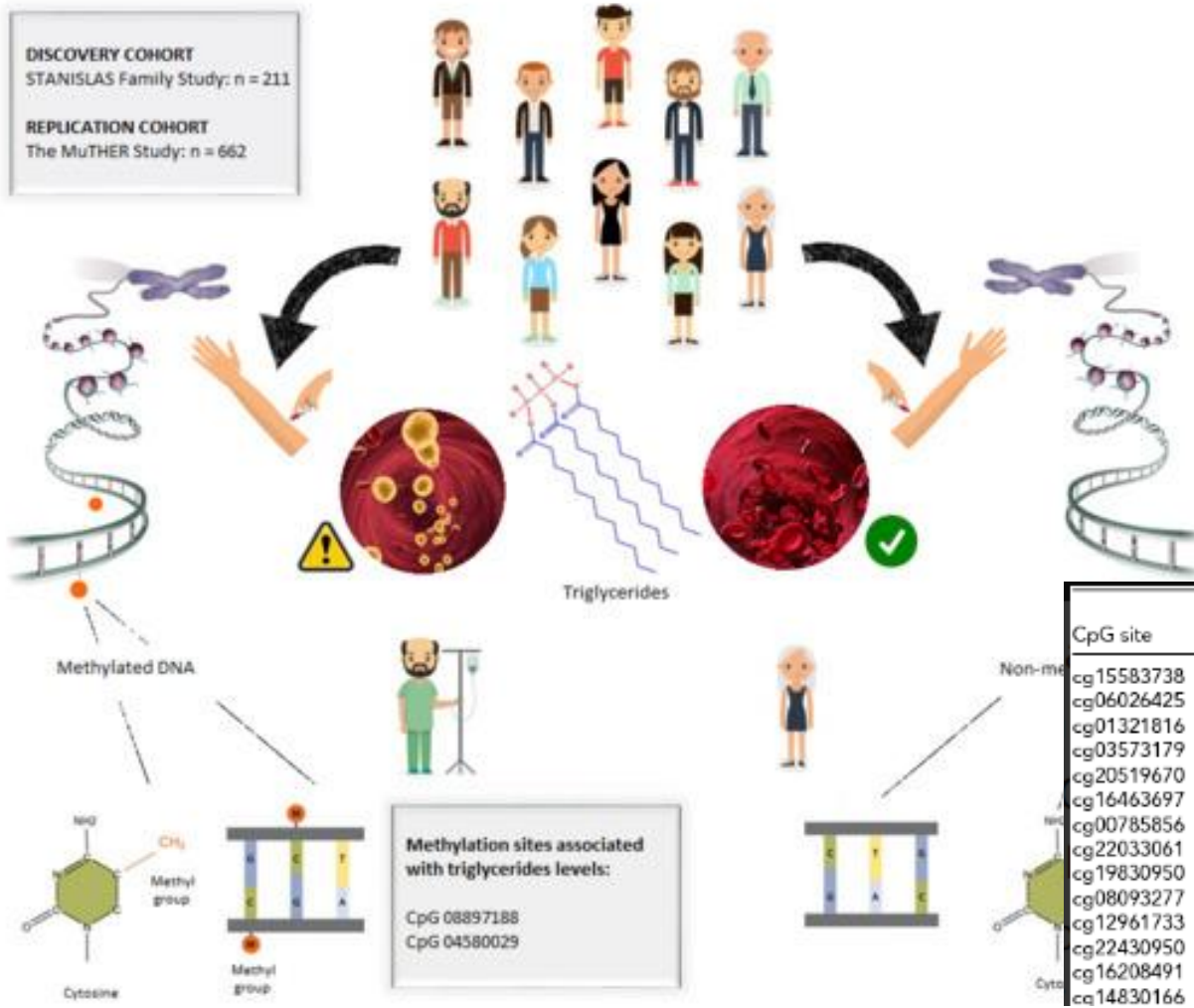


EWAS

(Epigenome-wide Association Study)

DISCOVERY COHORT
STANISLAS Family Study: n = 211

REPLICATION COHORT
The MuTHER Study: n = 662

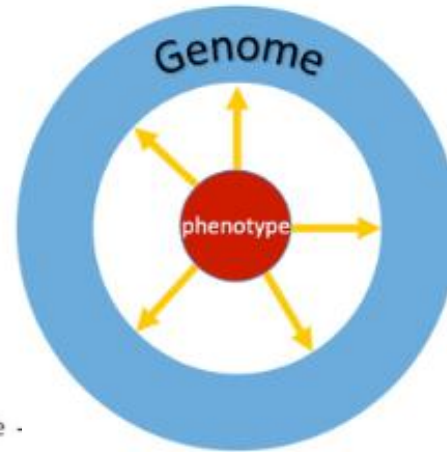


CpG site	CHR	CpG Location (bp)	p-value	$\Delta\beta$	[Nearest] gene	Gene location
cg15583738	8	2176944	$1.98e-06$	0.00408	—	—
cg06026425	5	157284650	$2.03e-06$	0.008678	CLINT1	5q33.3
cg01321816	15	91358514	$2.32e-06$	0.006536	BLM	15q26.1
cg03573179	6	36165382	$4.07e-06$	0.003791	BRPF3	6p21.31
cg20519670	14	65172006	$6.91e-06$	-0.00561	PLEKHG3	14q23.3
cg16463697	2	223886480	$8.97e-06$	0.006082	KCNE4	2q36.1
cg00785856	15	59041883	$1.03e-05$	0.005749	ADAM10	15q21.3
cg22033061	19	17531746	$1.04e-05$	0.003424	FAM125A/MVB12A	19p13.11
cg19830950	10	102729375	$1.21e-05$	0.002377	SEMA4G	10q24.31
cg08093277	6	29595299	$1.31e-05$	0.006787	GABBR1	6p22.1
cg12961733	22	50165244	$2.17e-05$	0.009437	BRD1	22q13.33
cg22430950	17	35166190	$2.62e-05$	-0.0146	—	—
cg16208491	6	4021748	$2.76e-05$	0.003249	PRPF4B	6p25.2
cg14830166	12	11821908	$3.01e-05$	-0.01192	ETV6	12p13.2

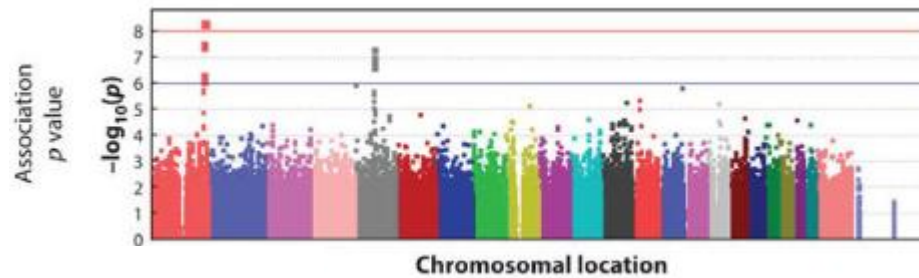
PheWAS

(Phenome-wide Association Study)

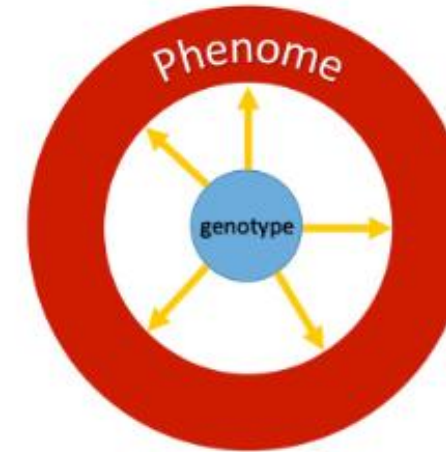
GWAS: examines associations between specific phenotypes and genetic variants across the genome



GWAS: Target phenotype -

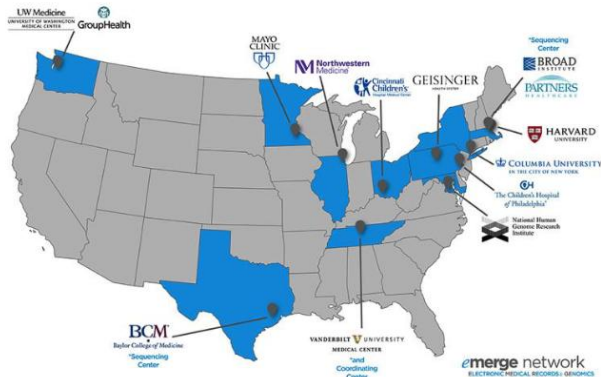
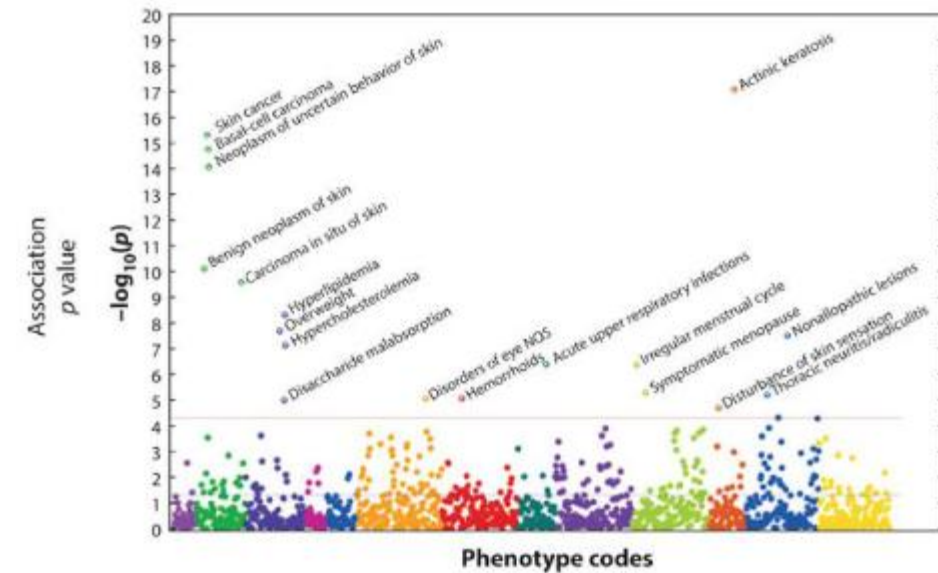


PheWAS: examines associations between specific genetic variants and a large number of different phenotypes (phenome)



- Ένα υποσύνολο των παραλλαγών του γονιδιώματος
- Μεγάλος αριθμός φαινοτύπων
- Πολλαπλές συσχετίσεις

PheWAS: Target genotype -
(or other input variable, e.g., a specific disease, trait, or exposure)



Ένα γνώρισμα/νόσημα
Όλο το γονιδίωμα
Συγκεκριμένες συσχετίσεις

Polygenic Risk Scores (PRSs)

(Πολυγονιδιακοί Δείκτες Κινδύνου)

- A count of the number of the risk variants across multiple genomic loci present in the person's DNA, weighted so that the presence of some risk variants is considered more important than others.

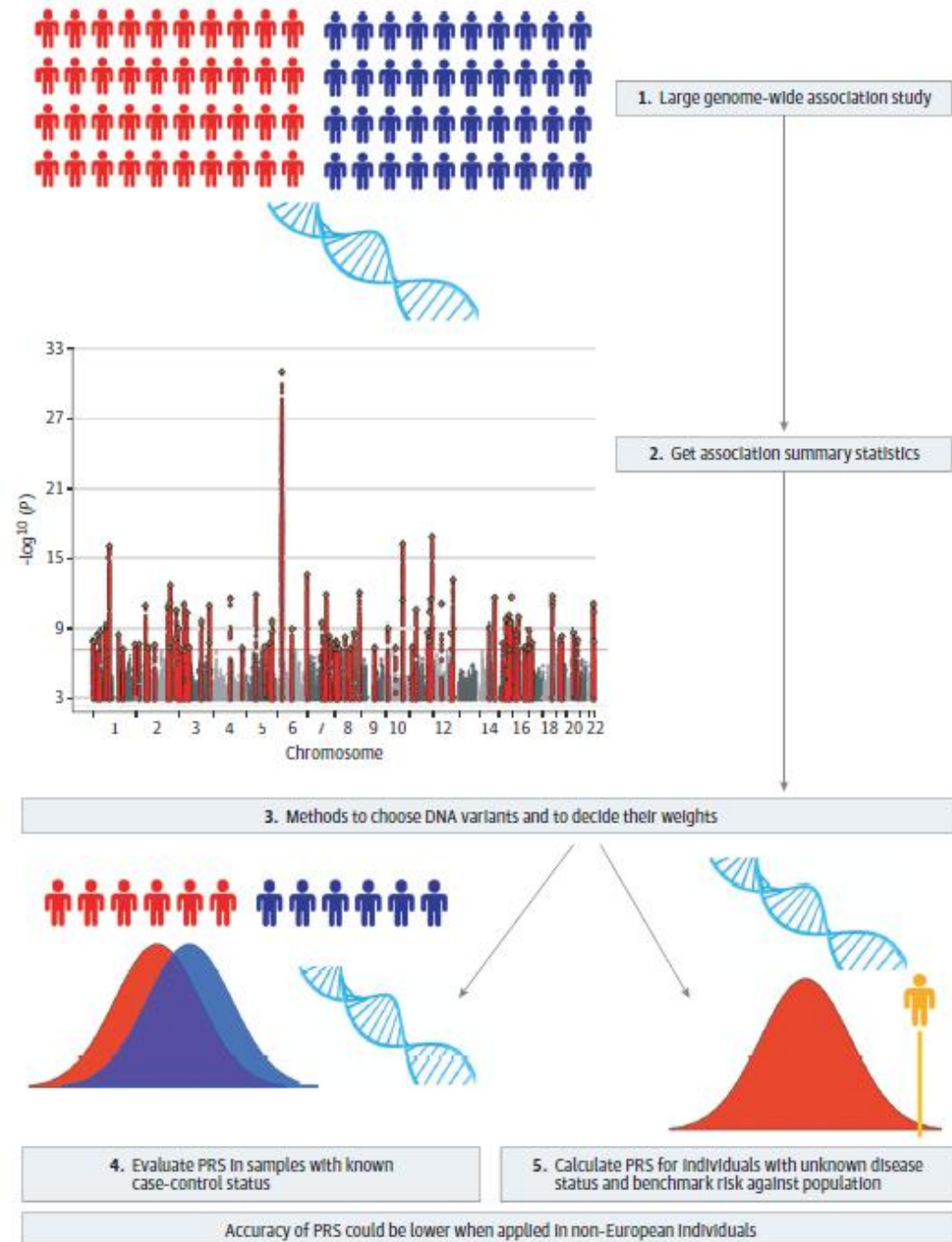
- ✓ IBD (Crohn's, UC)
- ✓ T1DM and autoimmune diseases
- ✓ T2DM and obesity
- ✓ CVD and hypertension
- ✓ AD and neurodegenerative disorders
- ✓ ADHD and ASD
- ✓ Schizophrenia
- ✓ Preterm delivery
- ✓ Drug response (PGx)
- ✓ Ancestry testing



PGS Catalog

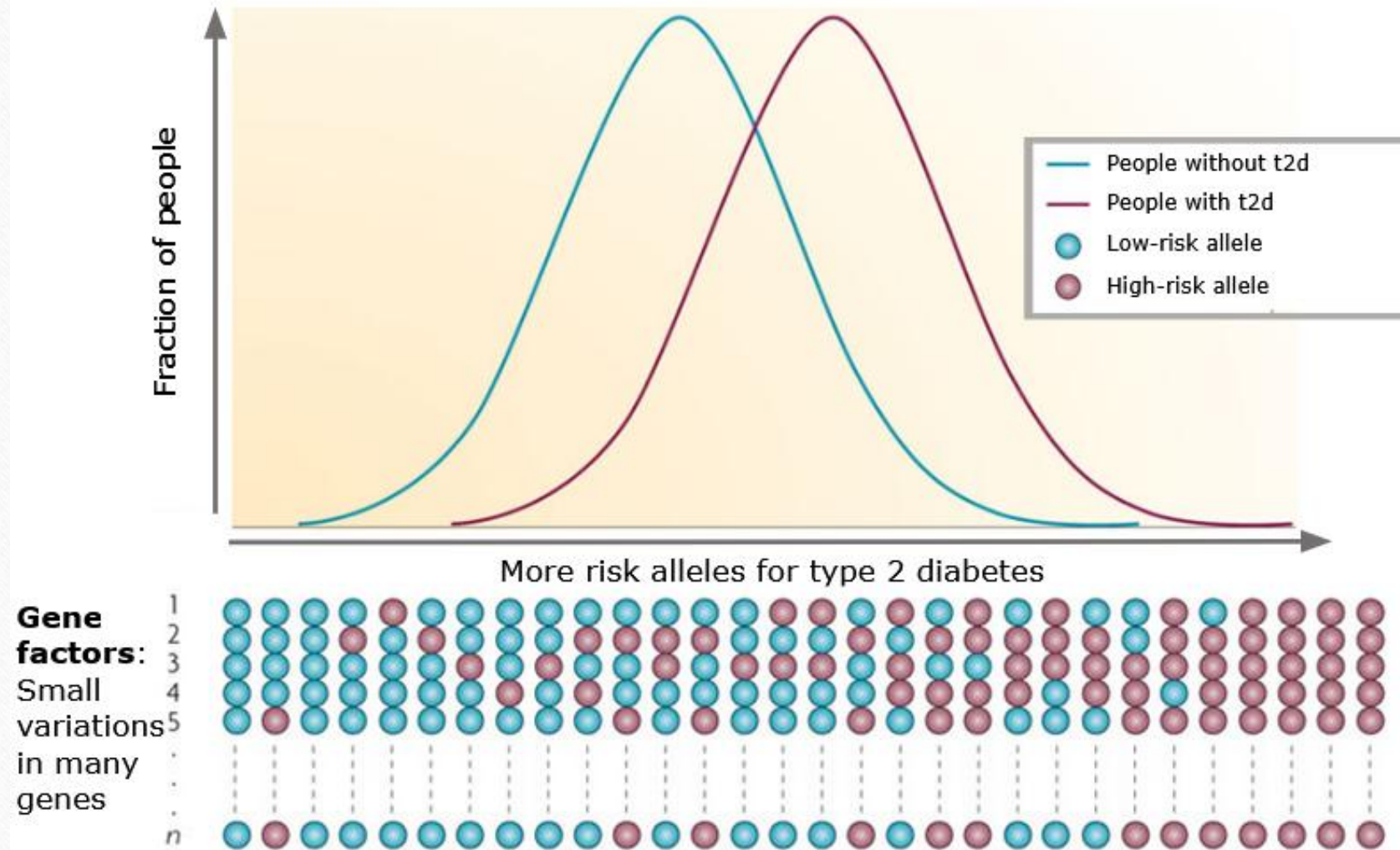
Wray et al., *JAMA Psychiatry*, 2020

Figure 2. Schematic of the Steps Needed to Generate and Validate Polygenic Risk Scores (PRS)



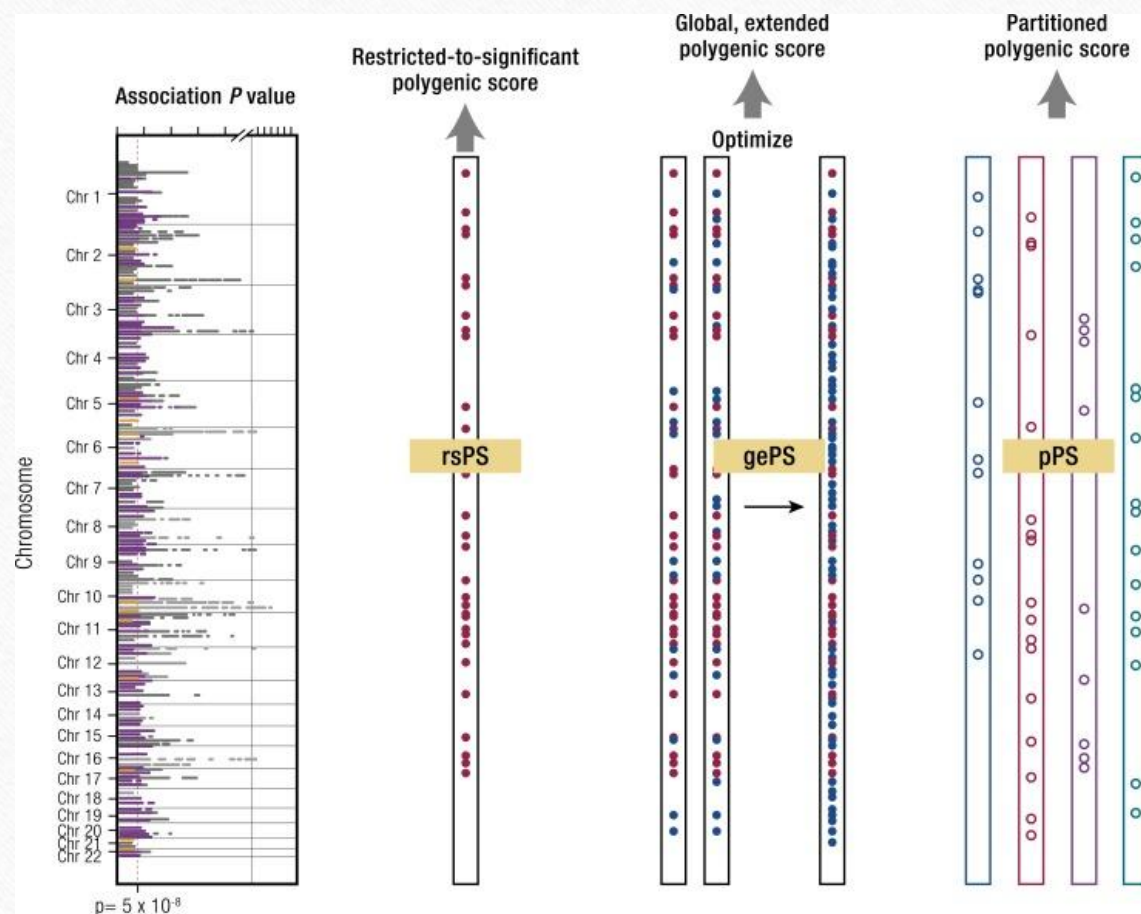
Polygenic Model of Disease

For complex diseases, variations in many genes may add to risk of getting that disease.



Polygenic Risk Scores (PRSs)

(Πολυγονιδιακοί Δείκτες Κινδύνου)



Restricted-to-significant polygenic scores (rsPSs): scores composed of variants at the extreme of a statistical distribution, most usually those that pass the genome-wide significant threshold ($p < 5 \times 10^{-8}$) for the trait concerned.

Global extended polygenic scores (gePSs):

scores generated from a deeper set of variants generated from genome-wide analyses, typically involving large numbers of sub-threshold significant variants ($p > 5 \times 10^{-8}$).

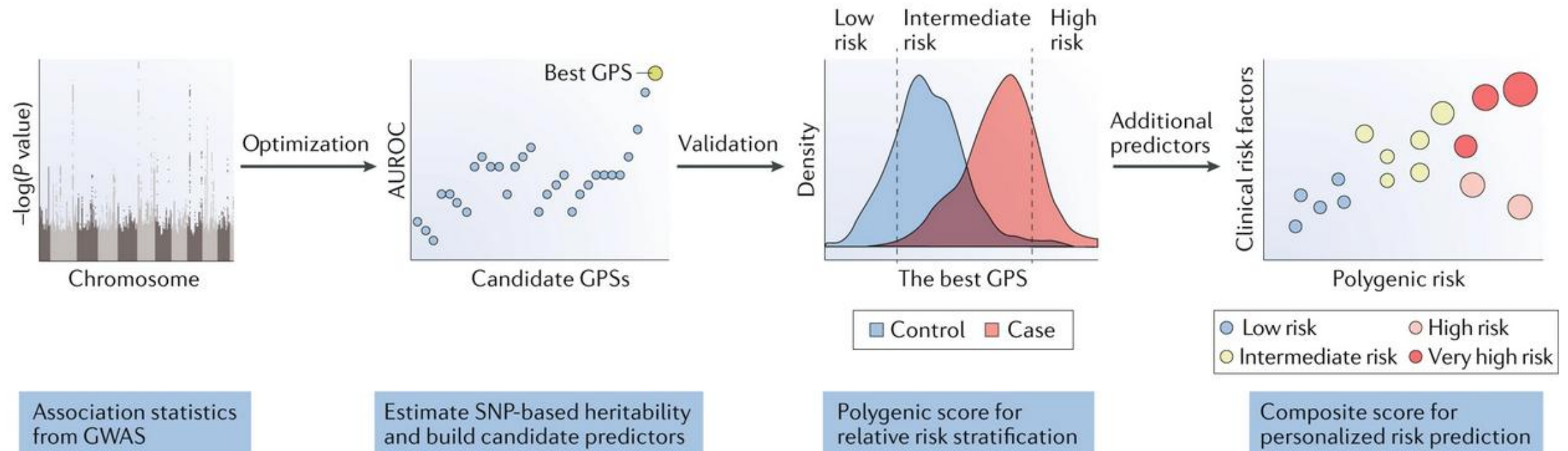
Partitioned (or process-specific) polygenic scores (pPSs):

scores composed of variants grouped according to some common biological process (e.g., association with a related endophenotype, tissue expression of related genes, chromatin state).

Alzheimer Disease (AD):

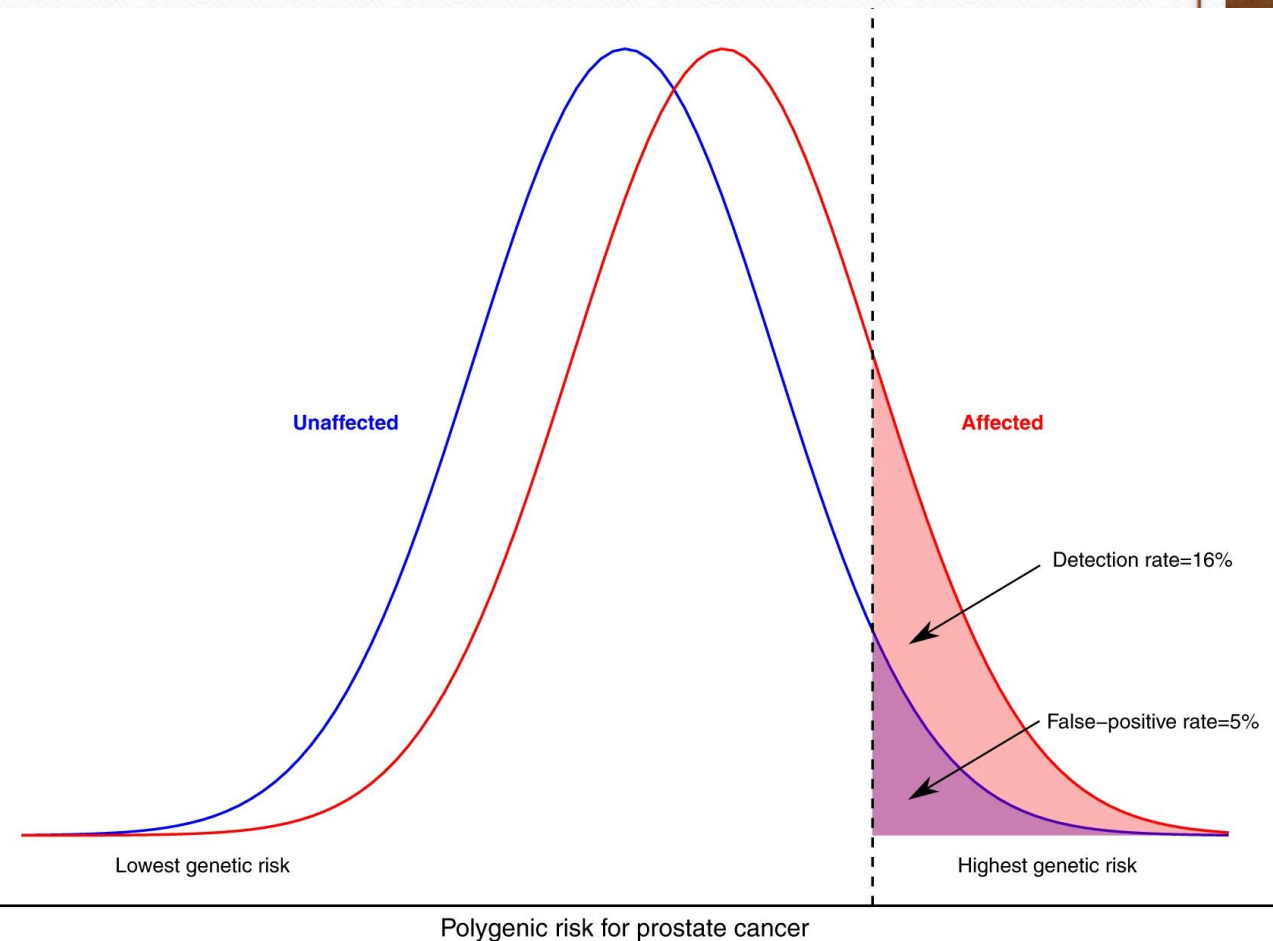
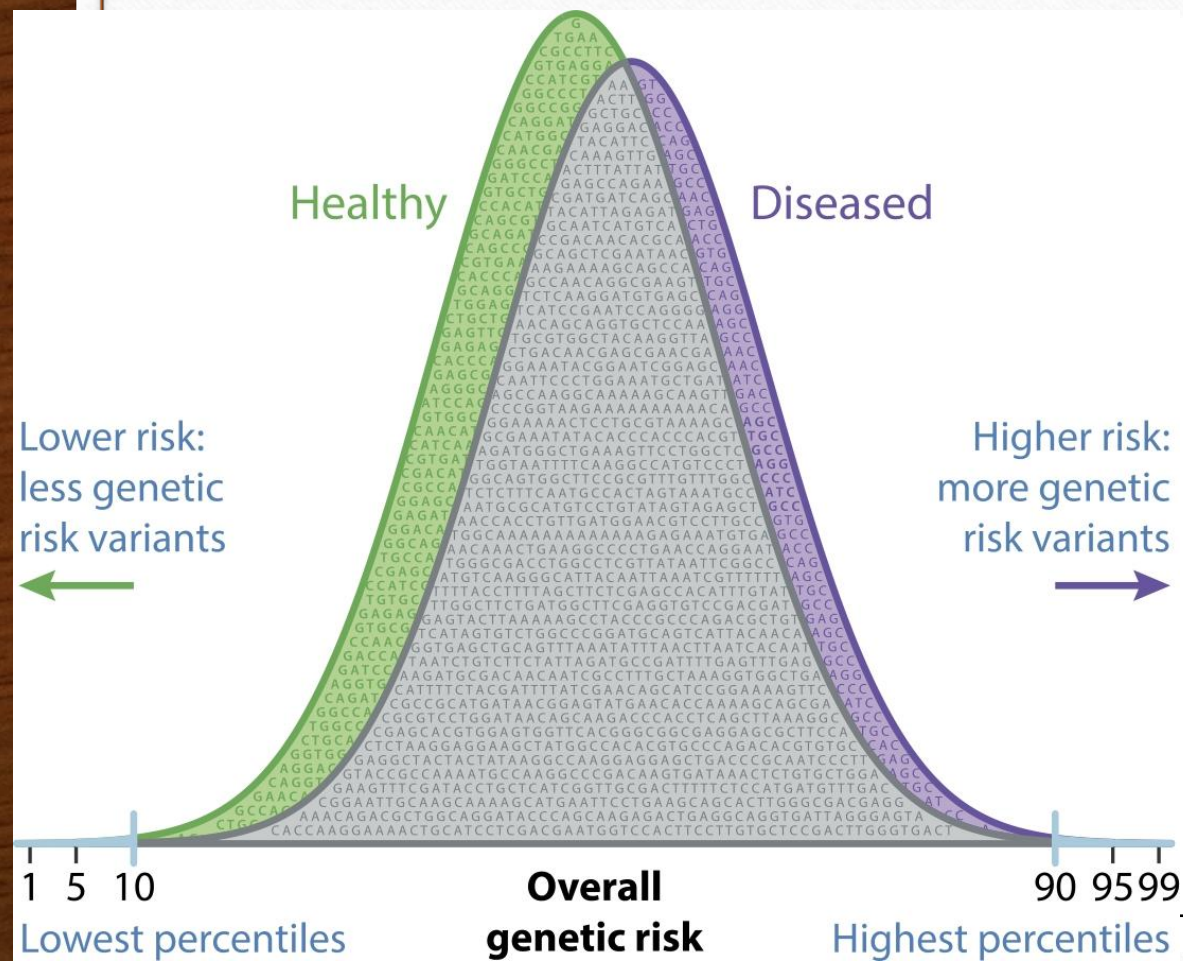
- AD risk, age-at-onset
- Progression (normal to MCI to AD)
- Cognitive function and memory
- Brain structure and function (MRI, PET)
- Biochemical changes in brain and periphery (CSF, blood, post-mortem)

Polygenic Risk Scores (PRSs)

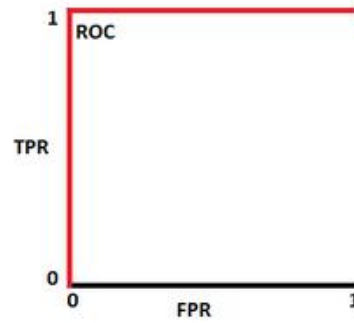
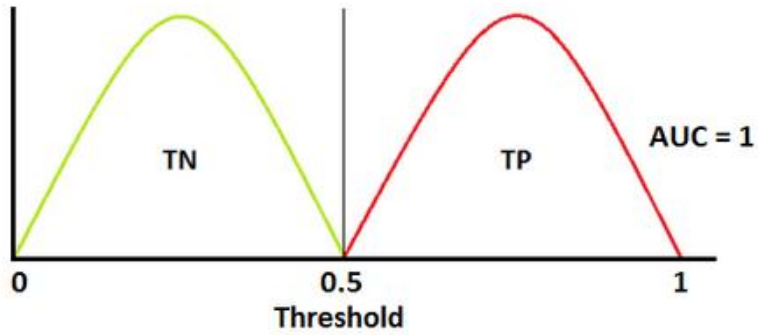


A genome-wide polygenic risk score (GPS) is based on genome-wide association study (GWAS) summary statistics. The optimization step enables selection of the best method according to the genetic architecture of a disease under study. The validation step requires an external cohort and is critical to obtaining reliable metrics of performance. Clinical predictors of absolute risk will require incorporation of additional demographic, clinical or lifestyle factors into composite risk models. AUROC, area under receiver operating characteristic. SNP, single-nucleotide polymorphism.

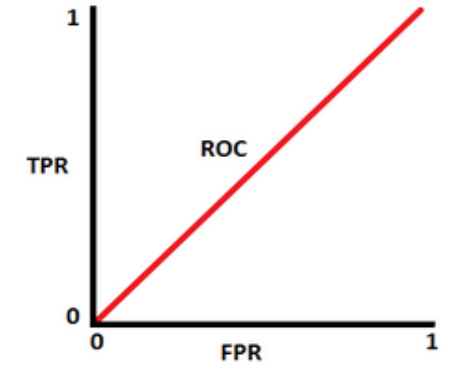
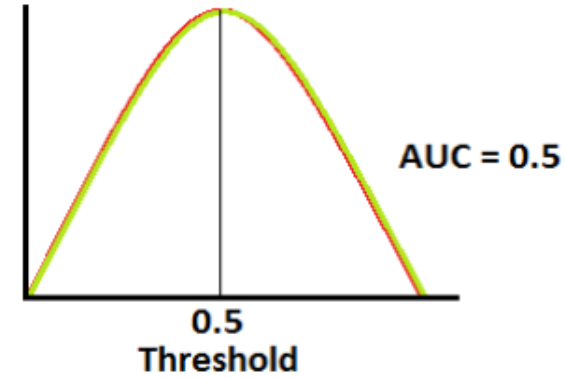
Polygenic Risk Scores (PRSs)



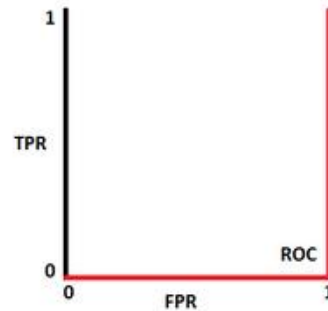
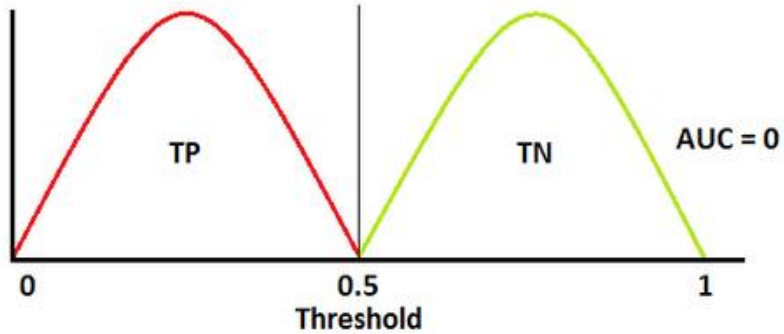
The perfect model



The non-informative model



The worst model



The current status in PRS research

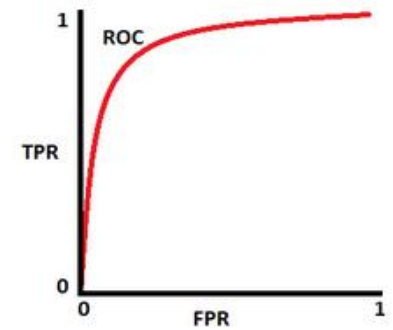
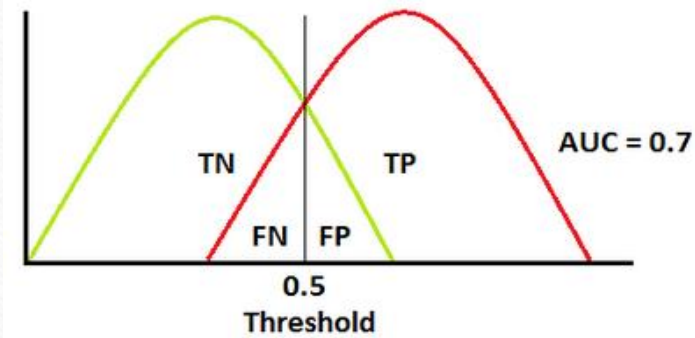
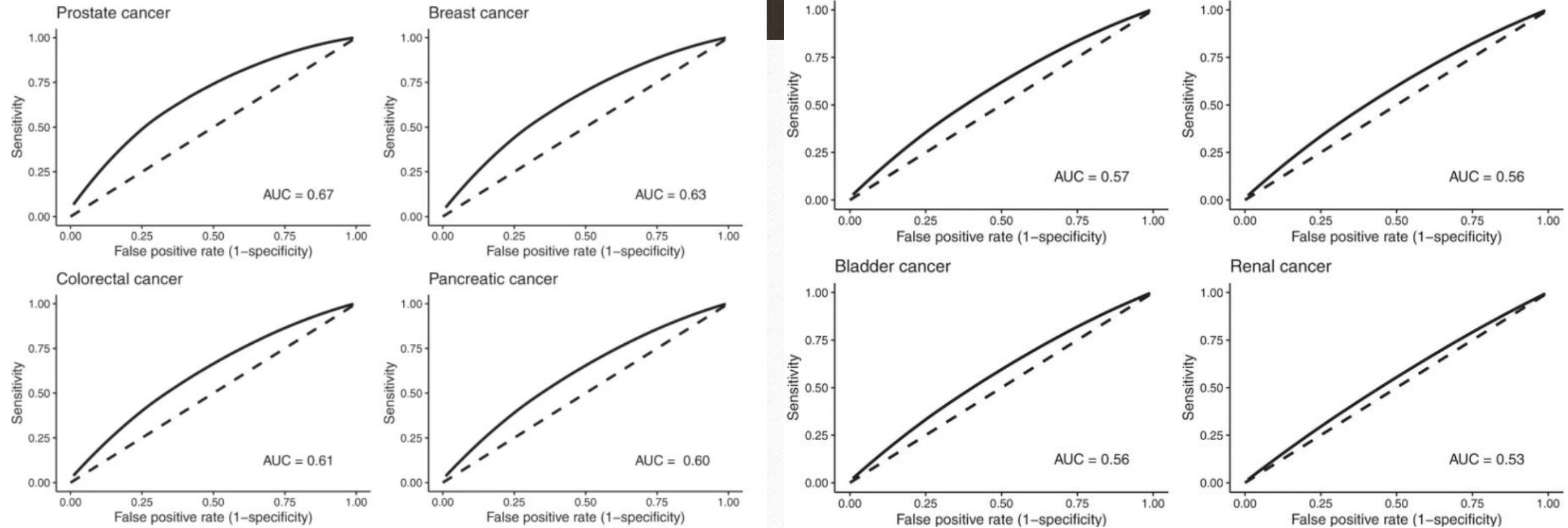
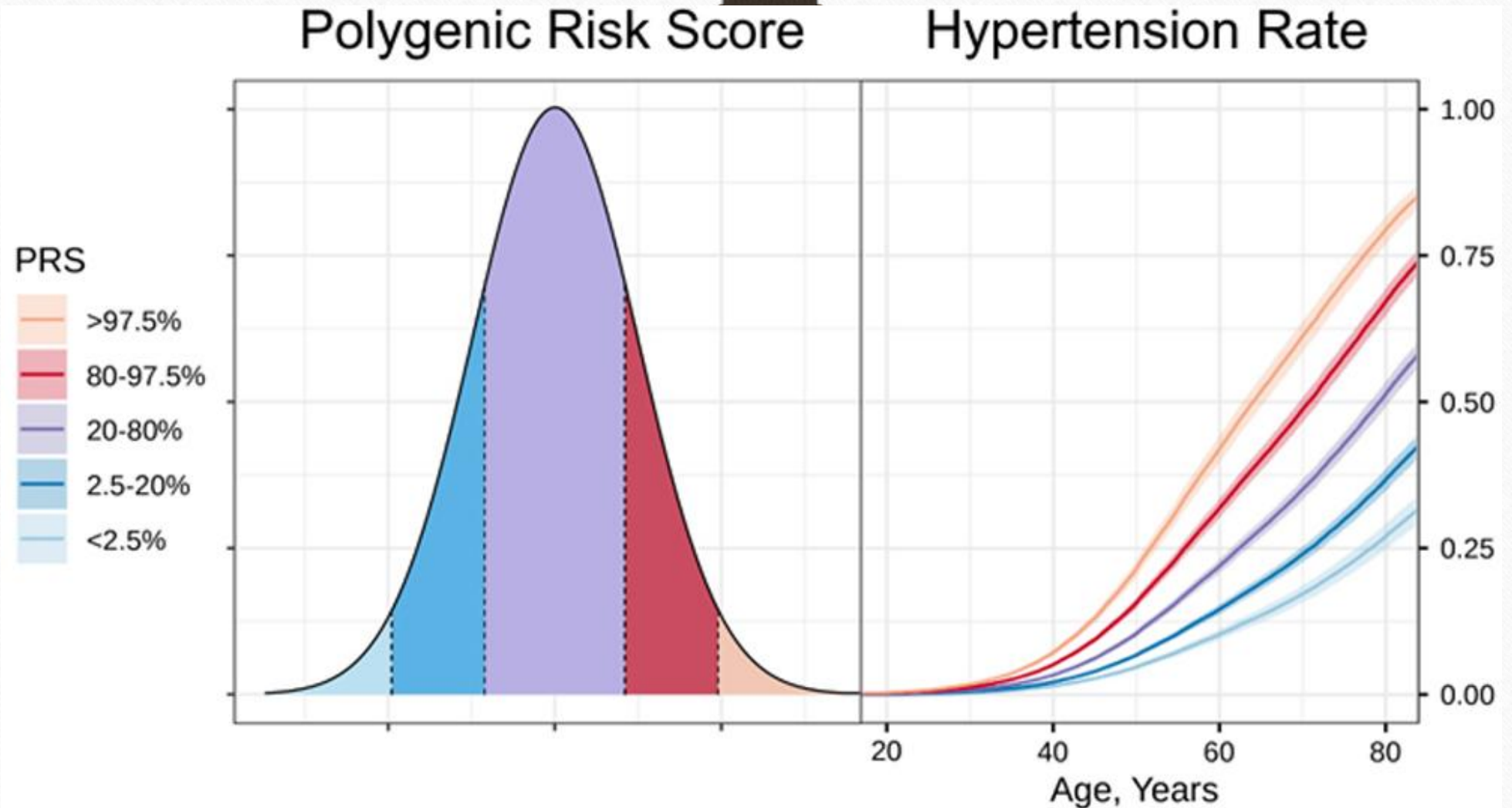


Fig. 2: Receiver operator plots of polygenic risk scores for eight common cancers.

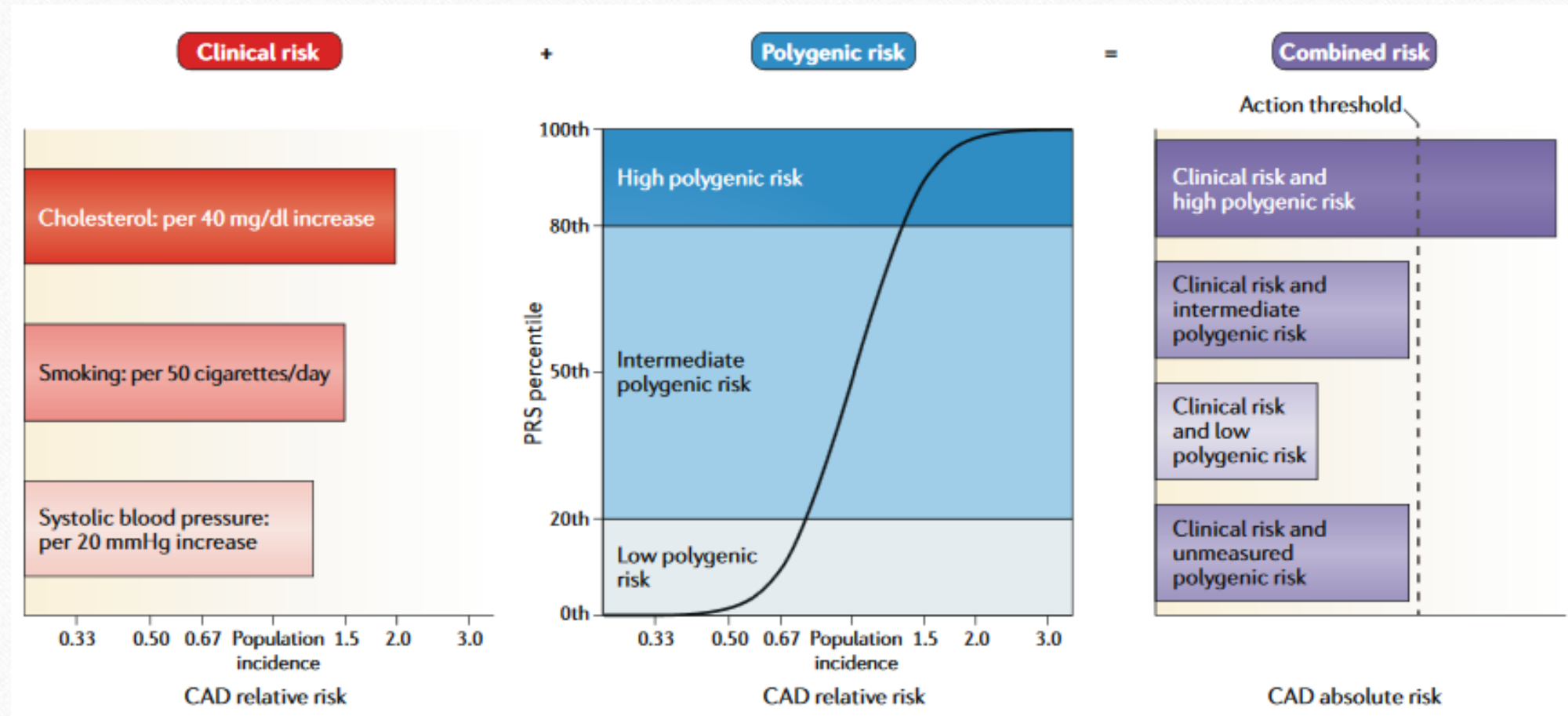


- ✓ **AUC values** of 0.7–0.8 are considered as acceptable and >0.8 as affording good discrimination
- ✓ Digital **mammography** for breast cancer screening has an AUC of 0.78!
- ✓ Breast cancer has the best characterized set of **non-genetic risk factors**: the AUC for these risk factors is 0.637 while the AUC of the current PRS is 0.631; and in combination the AUC is 0.683



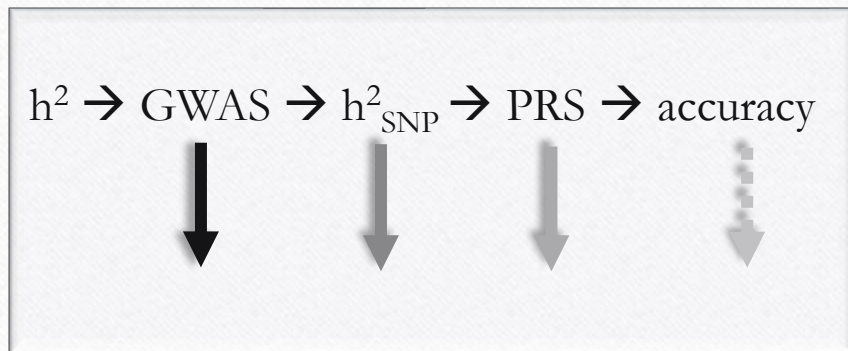
BP PRSs together with traditional risk factors could improve prediction of hypertension and particularly early-onset hypertension, which confers substantial CVD risk

Polygenic Risk Scores (PRSs) – The future prospects



PRSs pros

- DNA collection is performed once
- The costs of generating PRSs is very low
- Can be recalculated from the same genetic data if new information to improve PRS becomes available
- From a single DNA test, many PRSs can be generated
- Important for population-based screening and prevention programs (ex. CRC, BrCa, glaucoma)
- Specific diagnosis in early phase of illness when symptoms may be general and non-specific (ex. CKD)
- Contribution to treatment choices, response to treatment, adverse outcomes (PGx)



PRSs cons

- Prerequisite: Large GWAS sample sizes
 - Still limited predictive value (accuracy):
 - Genetic factors are not the only risk factors in common diseases (age, sex, environment, etc)
 - PRSs explain only part of the genetic contribution (based on common variants only, each with a small effect)
 - Prediction is not equally valid when using the same PRSs across ethnic ancestries → Need for ancestry-specific PRSs
 - The utility of the PRSs varies between conditions – dependent on their calculated heritability
-
- How should PRS-associated risk be communicated to the general public? (patient, family members, etc)
 - D-t-C companies already provide PRS for a number of diseases and traits – Are clinicians ready to/should they interpret “online PRS calculator” data?
 - Implications for health insurance?

Heritability of neuropsychiatric phenotypes

D. van Calker and T. Serchov

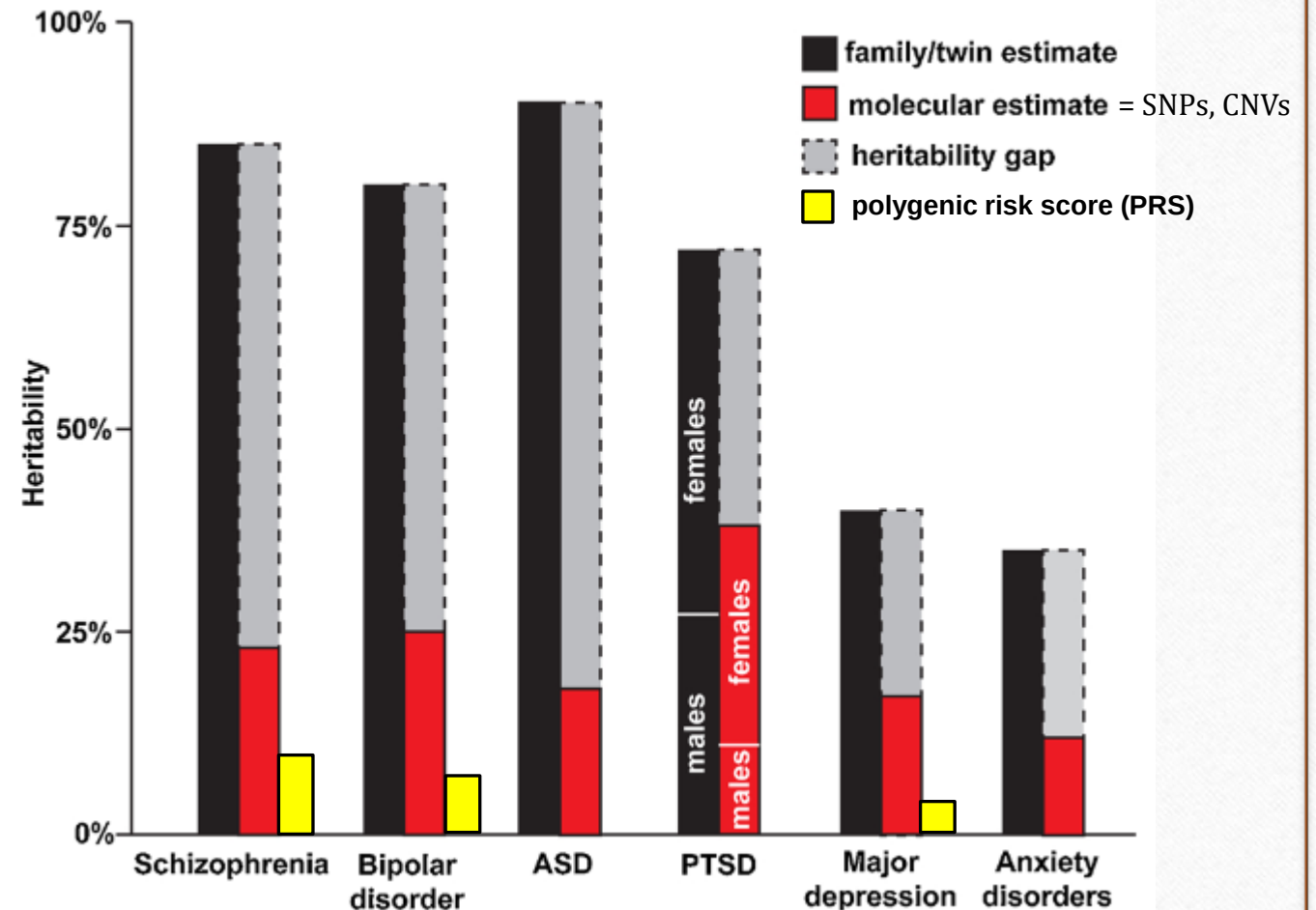
Neuroscience and Biobehavioral Reviews 126 (2021) 23–42

$$P = G + E$$

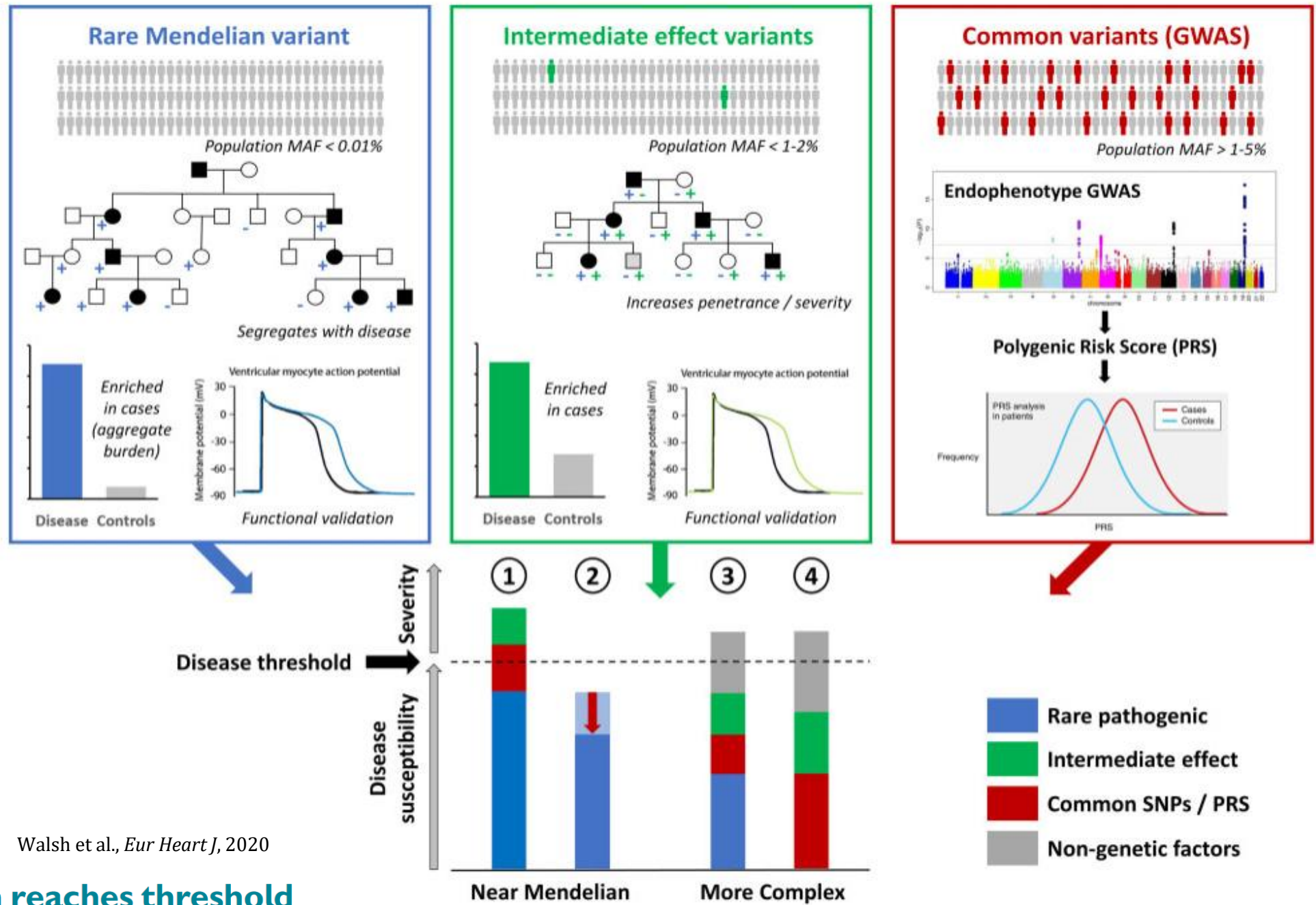
$$V_P = V_G + V_E + V_{G \times E}$$

$$V_P = V_A + \underbrace{V_D + V_I}_{\text{Non-Additive Genetic Variance}} + V_E$$

Phenotypic Variance Additive Genetic Variance Non-Additive Genetic Variance Environmental Variance



Σύνοψη

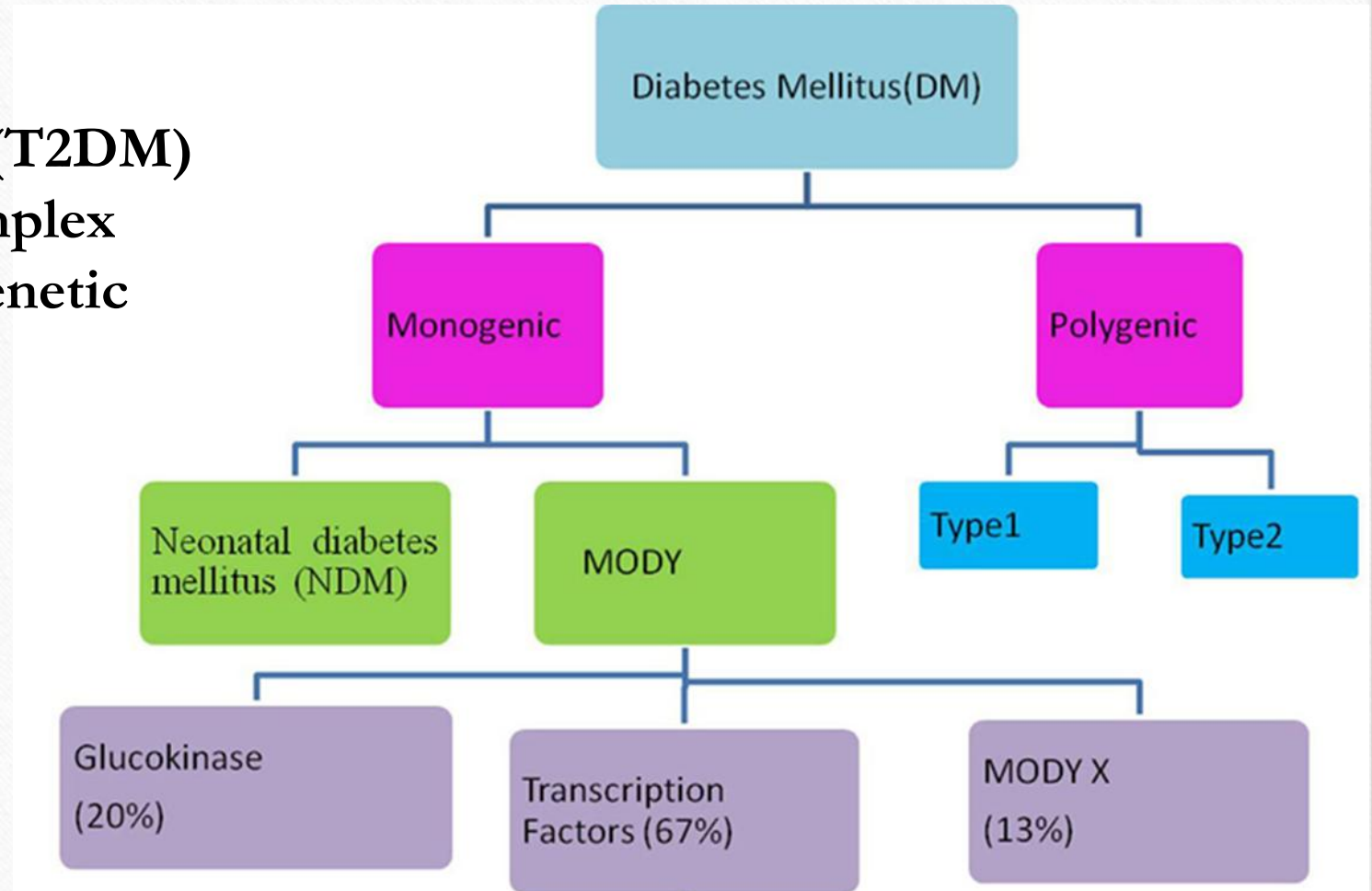


Walsh et al., *Eur Heart J*, 2020

When genetic burden reaches threshold

Part II

Type 2 Diabetes Mellitus (T2DM)
as a paradigm of a complex
disease with a strong genetic
component



Type 2 Diabetes Mellitus (T2DM)

- 80-90% of all diabetes cases
- NIDDM/T2DM (70-85%)
- 1-5% prevalence (6-8% in USA, 7% GR, 5-10% Malta/Finland, 15% Turkey, 25% N. Zealand aborigines, 39% Nauru-Pacific Ocean, 50% Pima Indians)
- ~500M cases, ~700M by 2050 (International Diabetes Federation)
- 1st degree relatives: ~3X higher risk of developing T2DM
- Macrovascular and microvascular complications
- Genes, family history, sex, ancestry, age → non-modifiable T2DM predisposing factors

COMPARISON OF TYPE 1 AND TYPE 2 DIABETES MELLITUS

Characteristic	Type 1 (IDDM)	Type 2 (NIDDM)
Sex	Female = male	Female > male
Age at onset	Childhood and adolescence	Adolescence through adulthood
Ethnic predominance	Whites	African Americans, Mexican Americans, Native Americans
Concordance		
Monozygotic twins	33%-50%	69%-90%
Dizygotic twins	1%-14%	24%-40%
Family history	Uncommon	Common
Autoimmunity	Common	Uncommon
Body habitus	Normal to wasted	Obese
Acanthosis nigricans	Uncommon	Common
Plasma insulin	Low to absent	Normal to high
Plasma glucagon	High, suppressible	High, resistant
Acute complication	Ketoacidosis	Hyperosmolar coma
Insulin therapy	Responsive	Resistant or responsive
Oral hypoglycemic therapy	Unresponsive	Responsive

The Staggering Costs of Diabetes

GROWING EPIDEMIC

Diabetes affects
30 million
children and adults
in the U.S.



That's
1 in 11
Americans.



84 million Americans
have prediabetes and are at risk
for developing type 2 diabetes.

**90% of them don't know
they have it.**



Every **21 seconds**
someone in the
U.S. is diagnosed
with diabetes.

1 in 3 US
adults by 2050

Admixed
populations

HUMAN COSTS

African Americans and
Hispanics are over
50%
more likely to have diabetes
than non-Hispanic whites.

People with diabetes are at higher risk of serious health complications:



STROKE



BLINDNESS



KIDNEY
DISEASE



HEART
DISEASE



LOSS OF
TOES, FEET,
OR LEGS

ECONOMIC COSTS



The total cost of diabetes
and prediabetes in the U.S. is
\$322 billion.



The average price of
insulin increased nearly
3x
between 2002 and 2013.



People with diabetes have
health care costs
2.3x greater
than those
without diabetes.

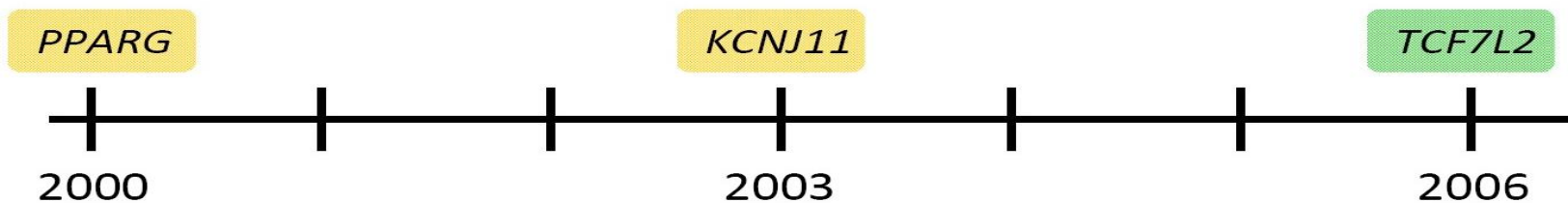
Prevention!!!

Learn more at diabetes.org

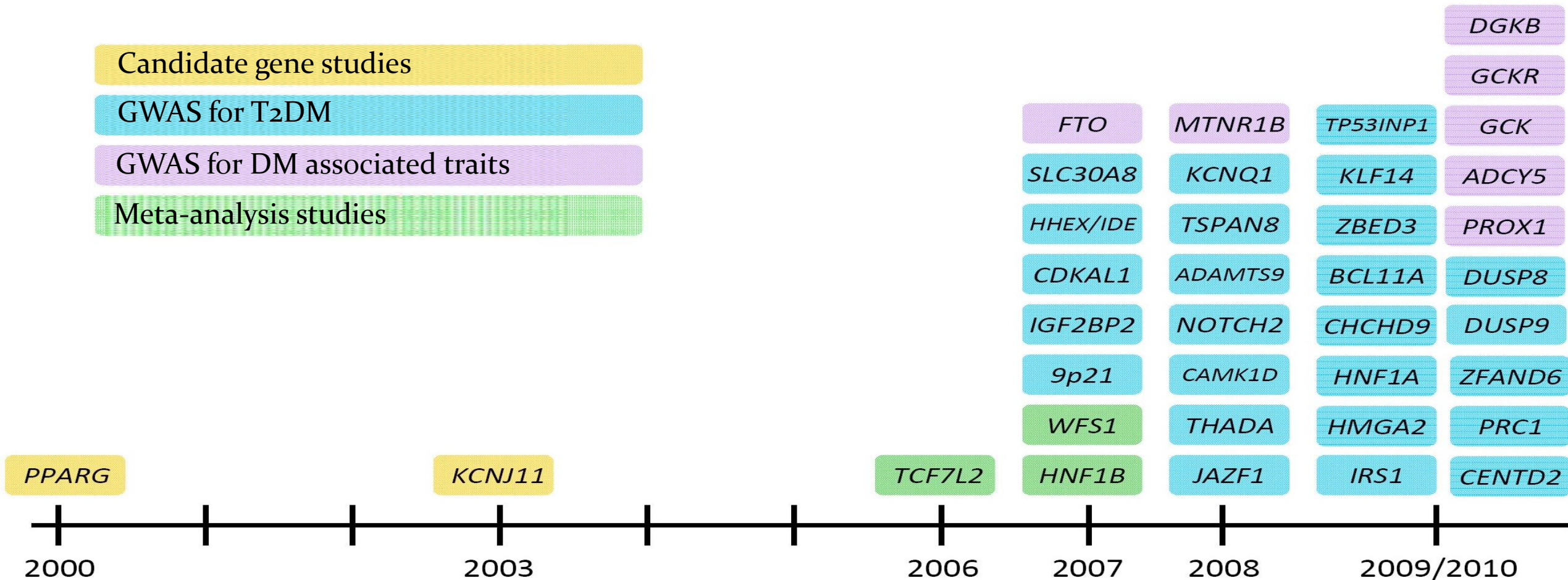


T2DM associated genes before 2007 – The pre-GWAS era

- Biological (functional) candidates (ignores less-obviously implicated genes)
- Positional candidates (small numbers of genes assessed)



T2DM associated genes after 2007 – The GWAS era

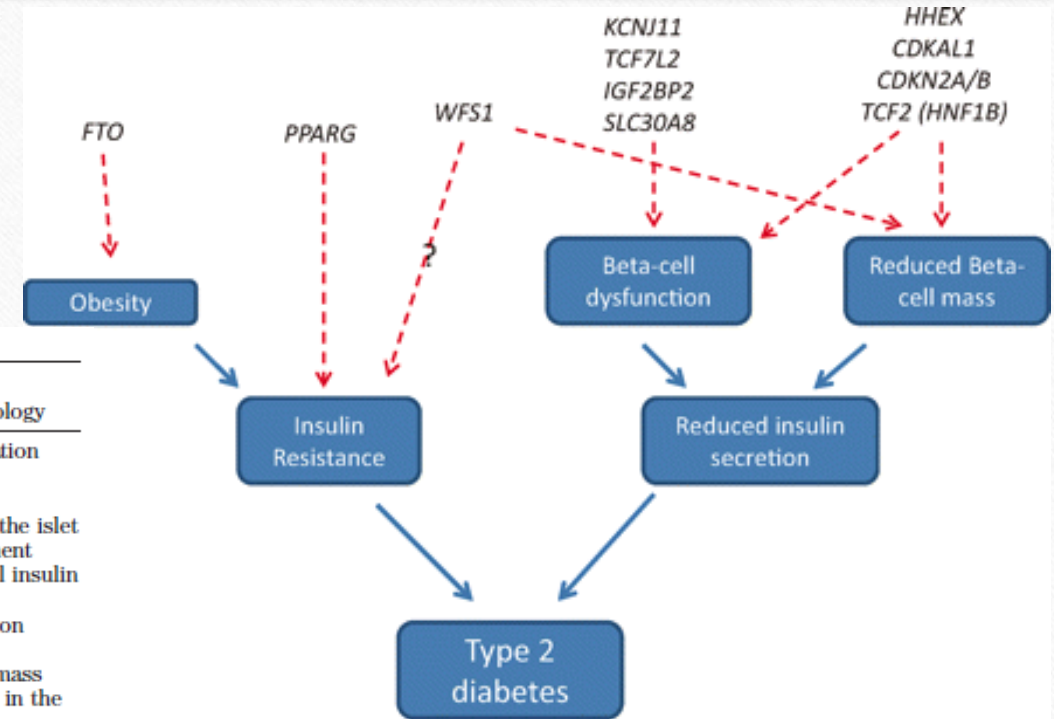


T2DM associated genes and molecular functions

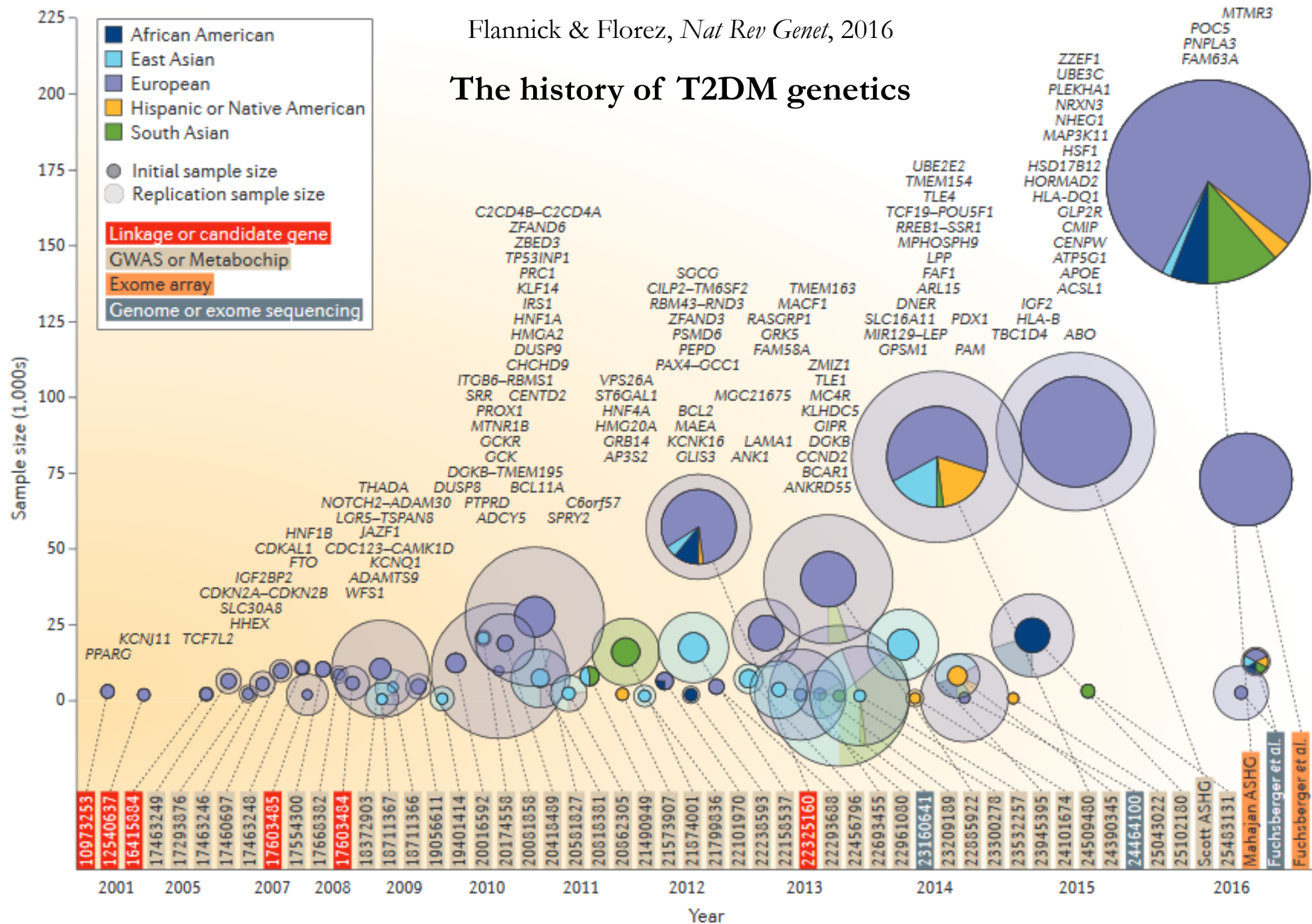
Summary details of the first 17 loci with a proven role in type 2 diabetes susceptibility

Signal	Chromosome	Representative SNP	Risk allele frequency	Effect size	How found	Hypothesized biology
<i>PPARG</i>	3	rs1801282	0.85	1.23	Candidate	Adipocyte differentiation and function
<i>KCNJ11</i>	11	rs5219	0.40	1.15	Candidate	β-Cell K _{ATP} channel
<i>TCF7L2</i>	10	rs7901695	0.40	1.37	Large-scale association	Incretin signaling in the islet
<i>HHEX</i>	10	rs5015480	0.63	1.13	GWA	Pancreatic development
<i>SLC30A8</i>	8	rs13266634	0.72	1.12	GWA	Zn transport in β-cell insulin granules
<i>FTO</i>	16	rs8050136	0.45	1.23	GWA	Hypothalamic effect on weight regulation
<i>CDKAL1</i>	6	rs10946398	0.36	1.16	GWA	β-Cell function and mass
<i>CDKN2A/B</i>	9	rs10811661	0.86	1.19	GWA	Cell cycle regulation in the β-cell
<i>IGF2BP2</i>	3	rs4402960	0.35	1.11	GWA	mRNA processing in the β-cell
<i>WFS1</i>	4	rs10010131	0.60	1.11	Large-scale association	Endoplasmic reticulum stress
<i>TCF2/HNF1B</i>	17	rs757210	0.43	1.08	Large-scale association	β-Cell development and function
<i>JAZF1</i>	7	rs864745	0.50	1.10	GWA	Transcriptional repression in the islet
<i>CDC123/CAMK1D</i>	10	rs12779790	0.18	1.09	GWA	Cell cycle regulation (<i>CDC123</i>)
<i>TSPAN8</i>	12	rs7961581	0.27	1.09	GWA	Cell surface glycoprotein
<i>THADA</i>	2	rs7578597	0.90	1.12	GWA	Apoptosis
<i>ADAMTS9</i>	3	rs4607103	0.76	1.06	GWA	Metalloprotease
<i>NOTCH2</i>	1	rs10923931	0.11	1.11	GWA	Pancreatic development

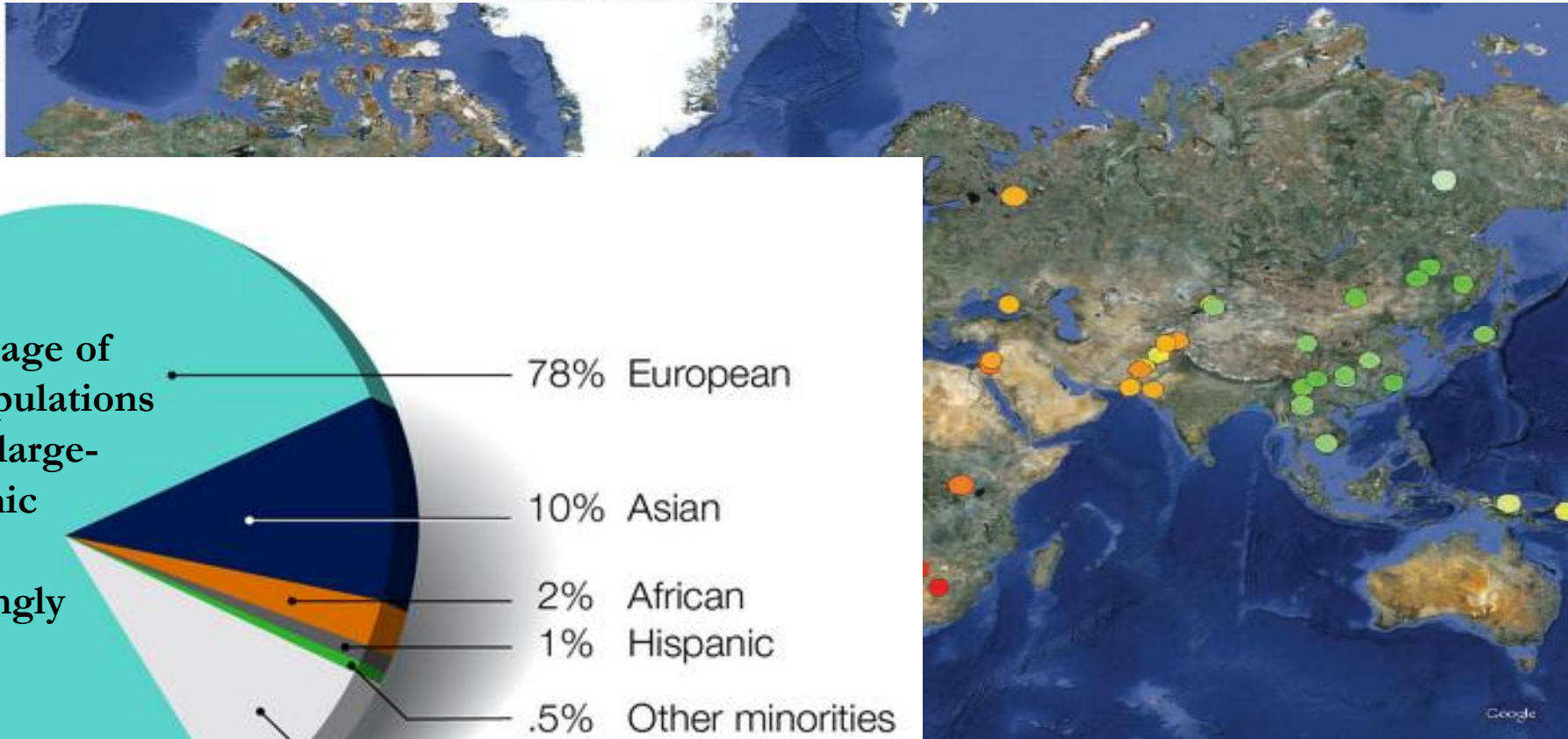
All loci have been shown to attain significance levels consistent with genome-wide significance in European populations. Note that in most cases the single nucleotide polymorphisms (SNPs) denoted are unlikely to be causal. Effect size is given as the estimated OR per copy of the risk allele. The biological processes listed are based on best available knowledge, but empirical data confirming these are not yet available



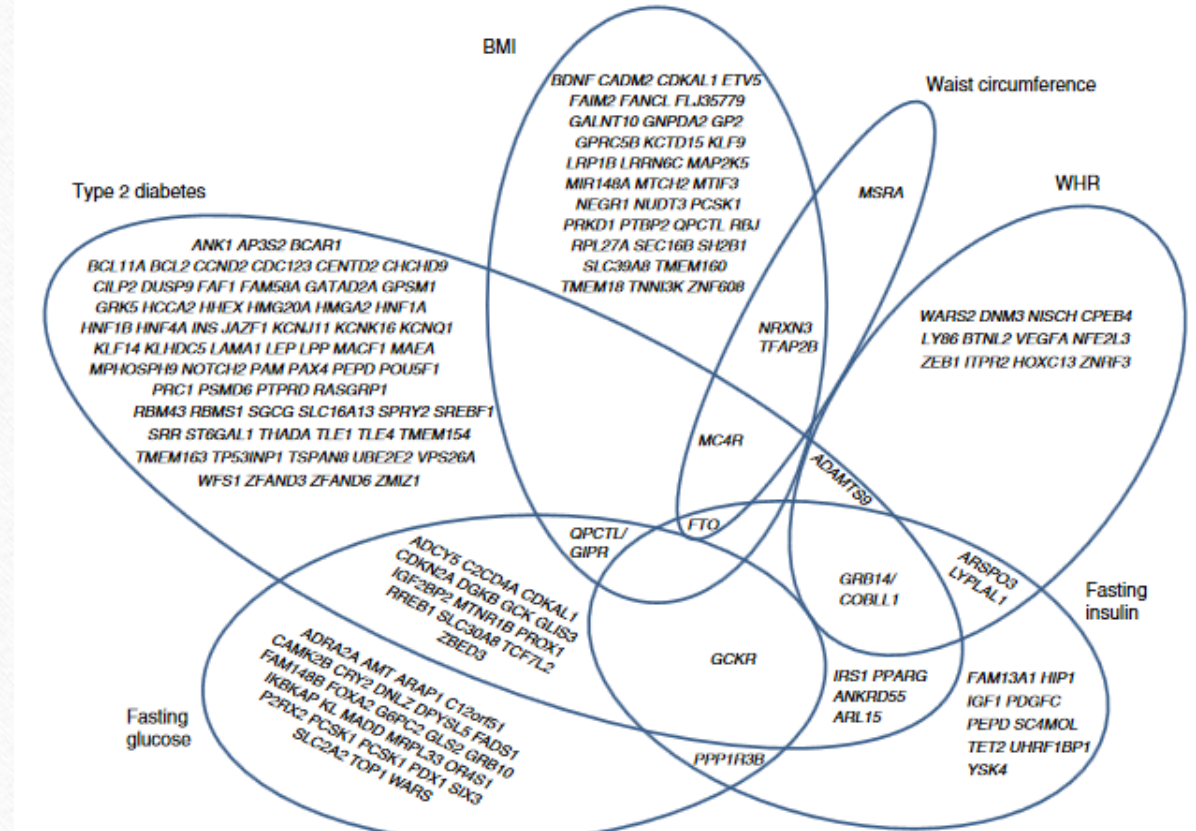
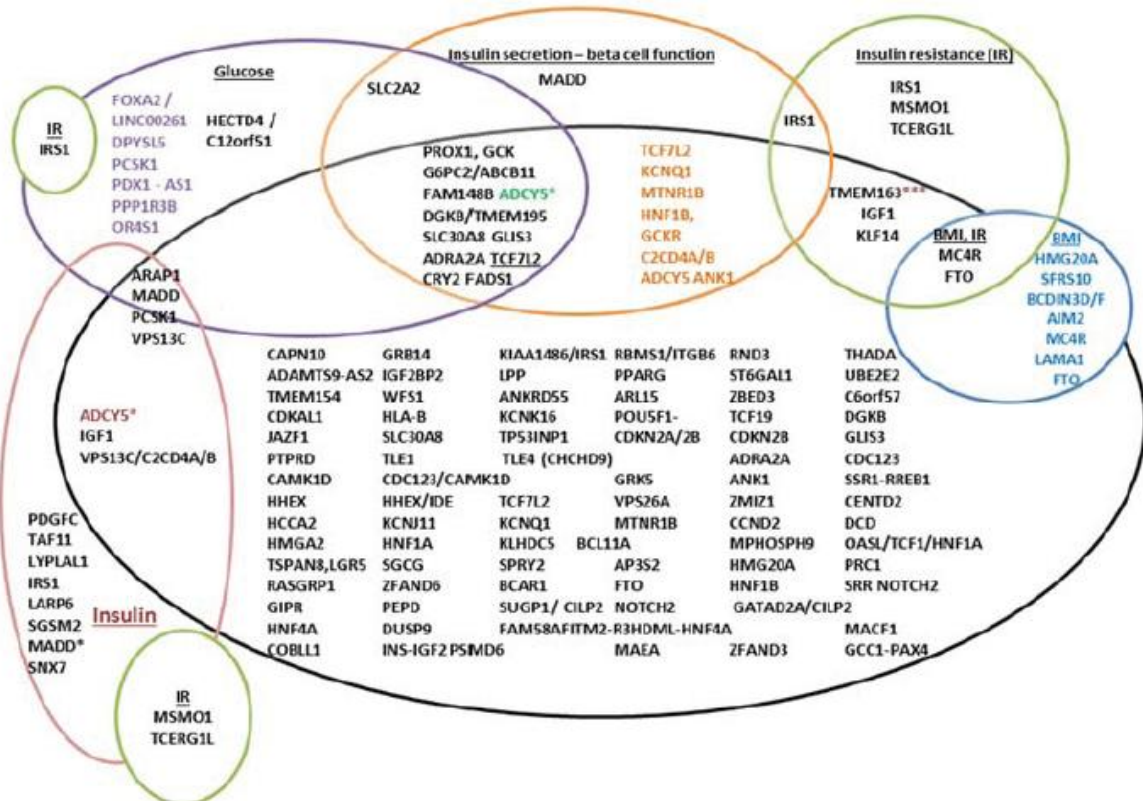
The history of T2DM genetics



Alleles associated with complex disorders differ in frequency around the world



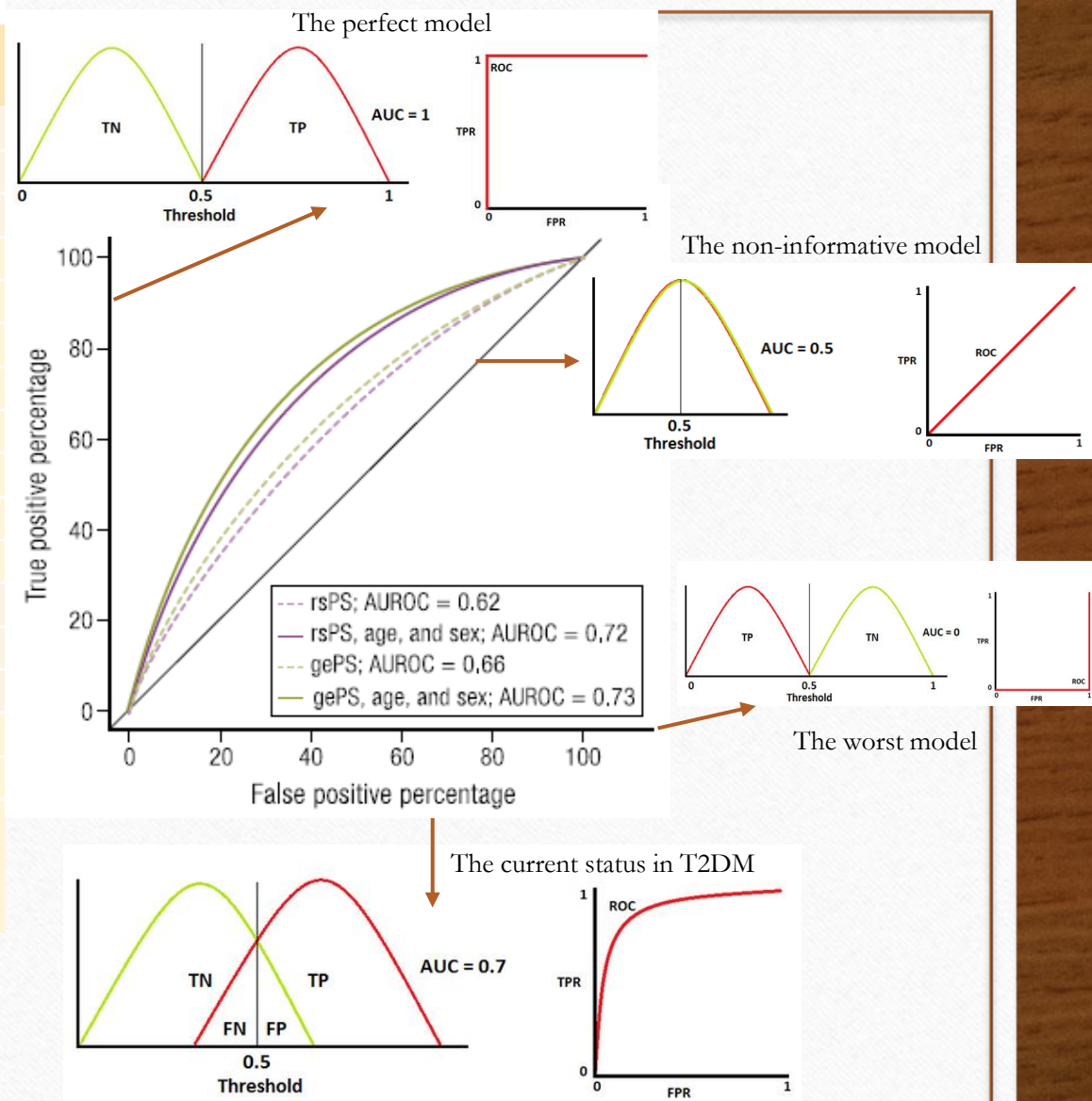
T2DM and glycemic traits – The GWAS era



PRSs and diabetes

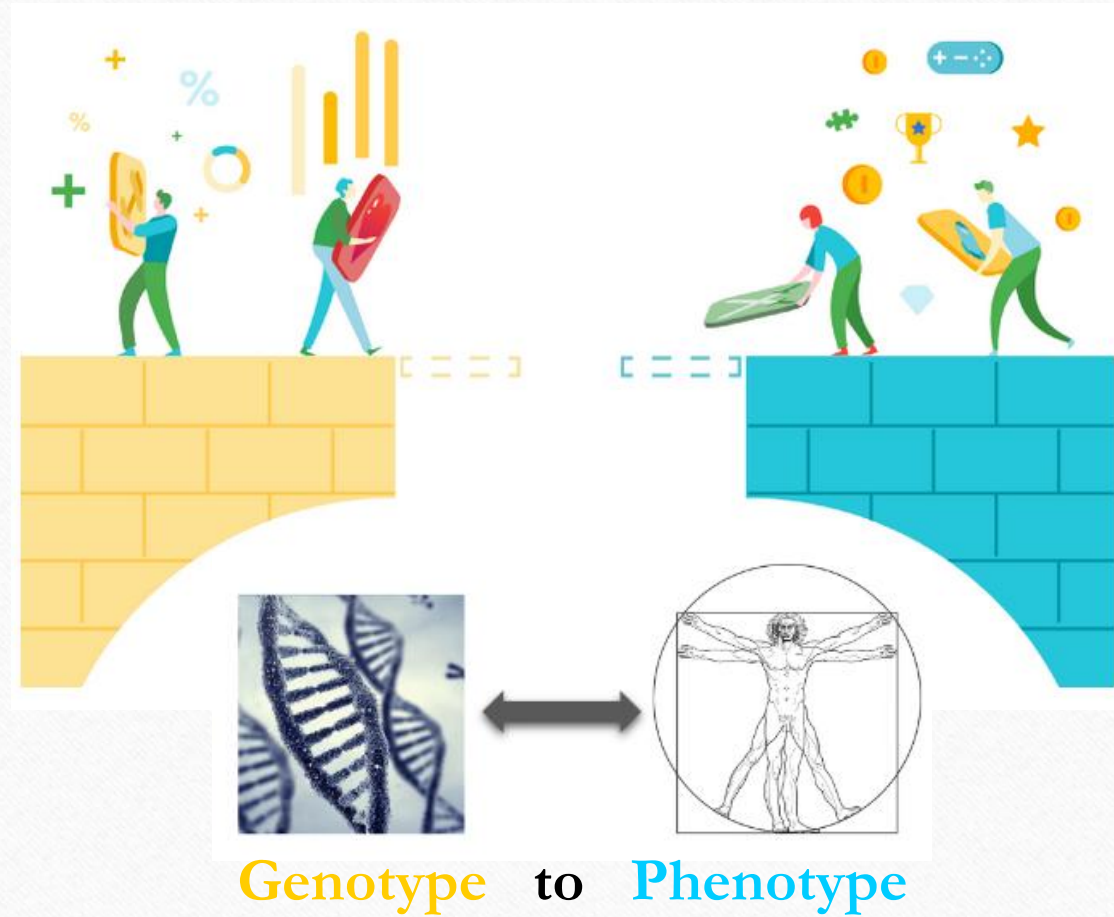
Table 1. Comparison of Three Published Global, Extended Polygenic Scores for T2D

		Study		
		Khera <i>et al.</i> , 2018 (13)	Mahajan <i>et al.</i> , 2018 (9)	23andMe (43)
Discovery GWAS	Number of cases	26,676	55,005	80,792
	Number of controls	132,532	400,308	1,479,116
	Reference	Scott <i>et al.</i> , 2017 (44)	Mahajan <i>et al.</i> , 2018 (9) ^a	Multhaup <i>et al.</i> , 2019 (43)
Optimization data set	Methods	LDpred	Pruning and thresholding	Predetermined cutoffs
	Number of cases	2785	5639	48,028
	Number of controls	120,280	112,307	893,692
	P value threshold	—	0.1	1×10^{-5}
	LD pruning threshold	—	$r^2 > 0.6$	50-kb window
	Tuning parameter	$\rho = 0.01$	—	—
	Polymorphisms in risk score	6917,436	171,249	1244
	Reference	UK Biobank	UK Biobank	23andMe ^b
Testing data set	Number of cases	5853	13,480	9008
	Number of controls	288,978	311,390	167,622
	Reference	UK Biobank	UK Biobank	23andMe
AUROC in testing data set (Europeans)	Not adjusted for age and sex ^c	0.64 ^d	0.66	0.65
	Adjusted for age and sex	0.73	0.73	—
	OR of top 5% bin vs remainder population	2.75	2.75 without age and sex adjustment	2.76 ^d
			4.52 with age and sex adjustment	



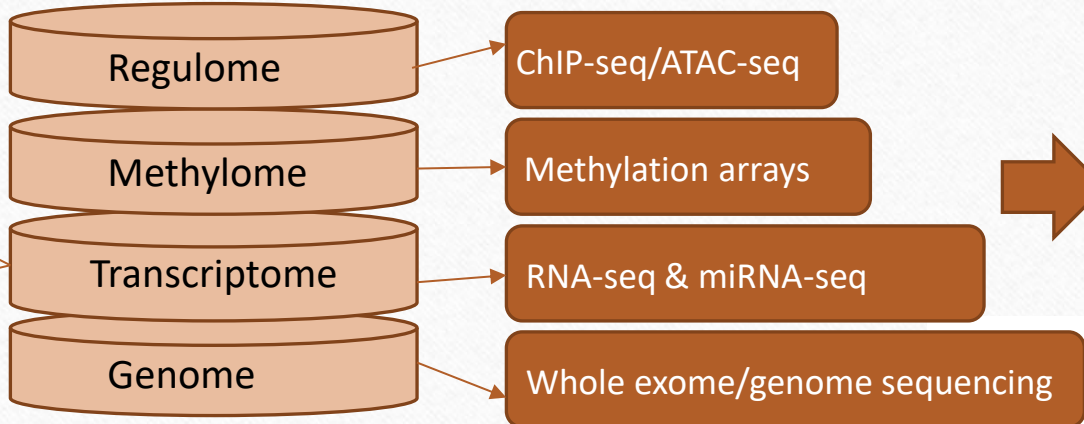
If the estimates of relative risk seen in UK Biobank participants in recent studies generalize to the population level, then there are likely to be >1M individuals in the UK, who, on the basis of their PRS alone, have a lifetime risk of T2D!

We need to bridge the knowledge gap from **sequence** to **consequence**



Combining different methods to explore beyond the genome ... the way forward

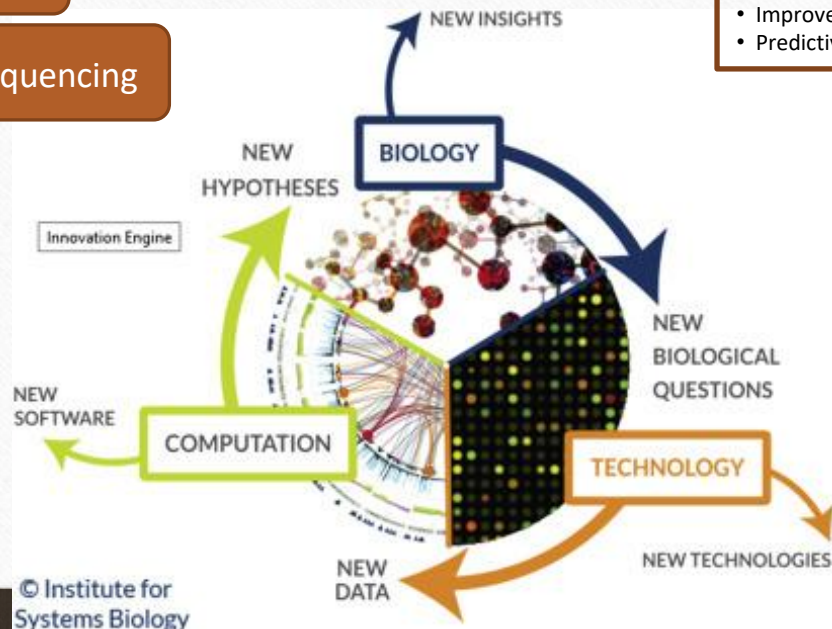
'Omics' approaches



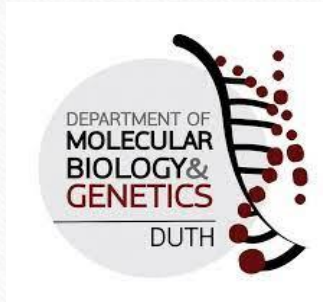
Comprehensive molecular signatures via integrated understanding of disease

Integrated 'omics' key benefits

- Disease-associated pathways
- Disease biomarkers
- Sub-phenotyping/disease subtyping
- Disease diagnosis/prognosis/staging/survival
- Patient stratification
- Novel clinical interventions/ treatments
- Improved drug design
- Predictive/Precision medicine



Prospects and limitations in the use of genetic markers in translational research of complex diseases



«ΜΕΤΑΦΡΑΣΤΙΚΗ ΕΡΕΥΝΑ ΣΤΗ ΒΙΟΙΑΤΡΙΚΗ»
Μοριακή Διαγνωστική, Βιοδείκτες και Στοχευμένες Θεραπείες



Μαριάνθη Γεωργίτση

Επίκουρη Καθηγήτρια Γενετικής και Μοριακής Βάσης Ασθενειών

E-mail: mgeorgit@mbg.duth.gr

Alexandroupoli, 09.05.2025