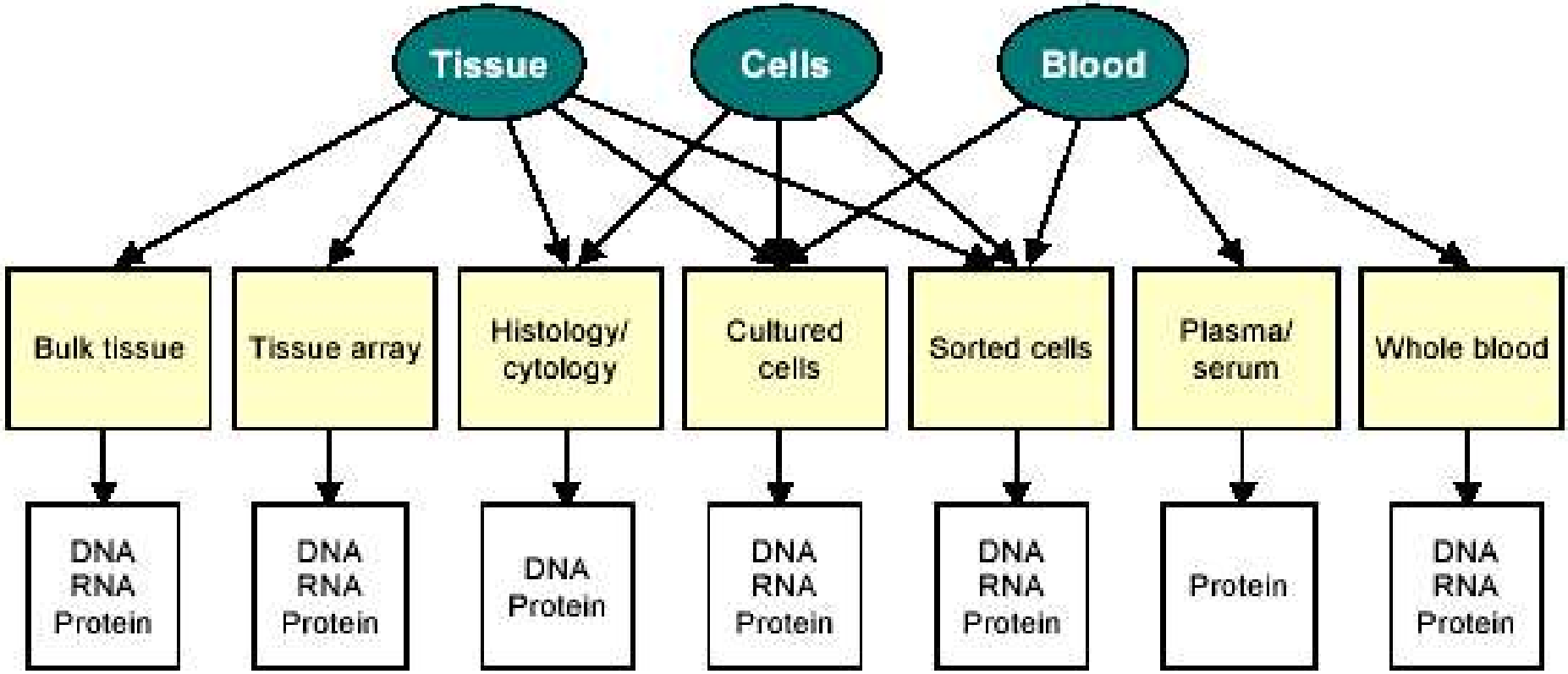
Biobanks: Accelerating molecular medicine

Dr. Csaba Szalai

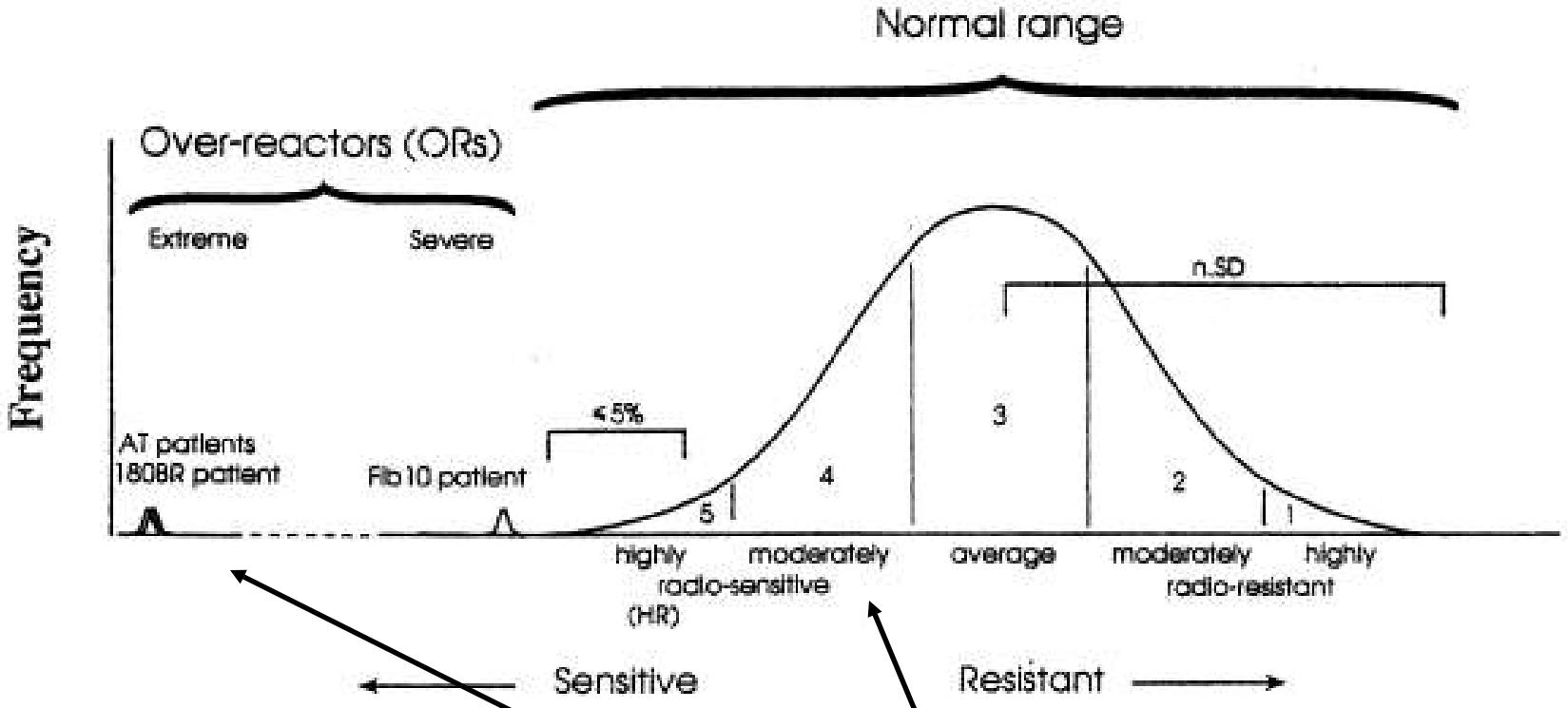
Illustration of the Types of Samples Collected by Repositories



- Why is it important to collect human DNA samples?

- How can SNP screening accelerate molecular medicine?

Responds of different individuals to environmental stimuli



Mutations with high vs. low penetrancy

Common diseases (multifactorial diseases)

- E.g.: CAD, asthma, hypertension, 2DM, obesity etc.
- They are orders of magnitude more common than individual mendelian disorders.
- E.g.: Asthma: 6-10%, HT: 25-35%, obesity: 20-75%
- CAD is responsible for 39% of all death
- Familial clustering
- Mainly unknown genetic background

Annual cost for multifactorial diseases

- CAD: $100 billion (USA)
- Asthma: $11 billion (USA)
- Obesity: 2-7% of total health cost (developed countries)

Why is it useful to identify the genomic background of a disease?

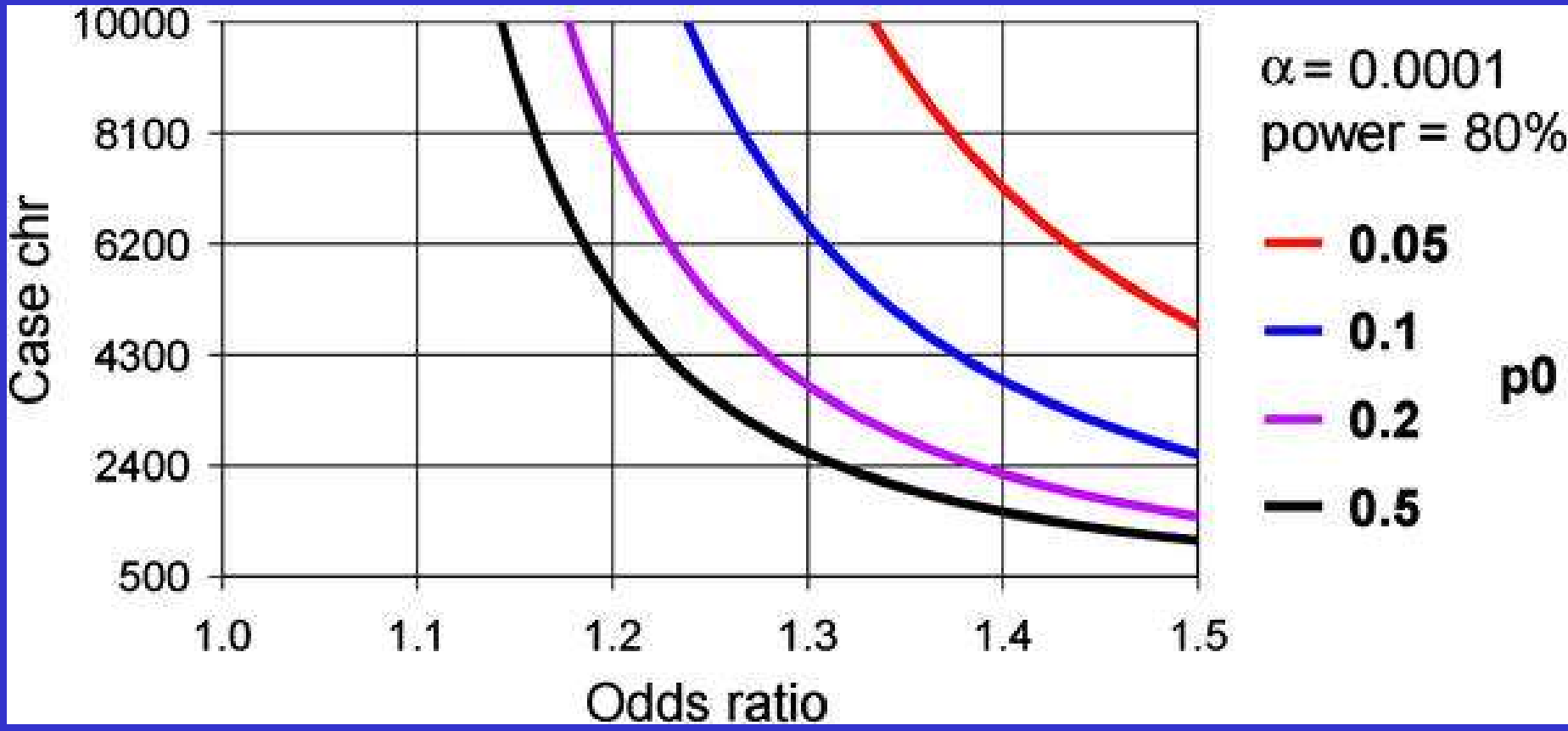
1. Identifying the molecular pathomechanism → new drug targets
2. Personalized therapy
3. Moving from "diagnose and treat" to → "predict and prevent"

1. Segít megismerni a molekuláris pathomechanimust: új gyógyszercélpontok azonosítása
2. Emberek közötti genetikai különbségek miatt más a kezelésre adott válasz: a genetikai háttér megismerésével lehetőség nyílik a személyre szabott kezelésre
3. Ki lehet szűrni a betegségre genetikailag hajlamos embereket: áttérés a 'diagnosztizáld és kezeld'-ről a 'jósold és előzd meg' stratégiára

Immense effort, little succes. Why?

- More than one gene in each individual might interact to produce the disease phenotype (**polygenic inheritance**)
- Different disease alleles might exist in different individuals (**genetic heterogeneity**),
- **Incomplete penetrance**
- By contrast with single-gene disorders, genes that predispose to these diseases will not usually contain mutations that lead to a gross aberration in function. Most often they will be **variants of normal genes**, the evolutionary advantages of which have become obscure.
- Odds ratio for each allele close to 1

Calculating the case number necessary to achieve significant results



European Consortium for IDDM Genome Studies:

A Genomewide Scan for Type 1-Diabetes Susceptibility in Scandinavian Families: Identification of New Loci with Evidence of Interactions

Am J Hum Genet. 2001 Dec;69 (6):1301-13.

Families of Patients Included in Genomewide Linkage Analysis

Category	Danish	Norwegian	Swedish	Total
Sib pairs:				
Affected	175	89	200	464
Unaffected	135	23	53	211
Discordant	287	127	277	691
Half-sib pairs:				
Affected	0	1	13	14
Unaffected	0	0	4	4
Discordant	0	3	20	23
Affected father-child pairs	41	3	20	64
Affected mother-child pairs	12	4	8	24
Avuncular pairs	0	8	1	9
First-cousin pairs	0	0	2	2
Families:				
Multiplex families	147	77	184	408
Simplex[a]	5	2	9	16

[a] These families were in the TDT analysis.

Lokusznév	Kapcsolt marker vagy gén	Citogenetikai lokusz
IDDM1*	HLA DRB1-DQA1-DQB1	6p21.3
IDDM2*	insulin gén 5' VNTR	11p15.5
IDDM3	D15S107 microsatellita	15q26
IDDM4*	FGF3/D11S1337 régió	11q13
IDDM5*	D6S476-D6S448 régió	6q25
IDDM6	D18S487 microsatellita	18q21
IDDM7	D2S152 microsatellita	2q31–33
IDDM8*	D6S281 microsatellita	6q27
IDDM9	D3S1303	3q21-q25
IDDM10	D10S193-D10S588 régió	10p11-q11
IDDM11	D14S67 microsatellita	14q24.3-q31
IDDM12*	CTLA-4 gén	2q33
IDDM13	IGFBP-2, -5génrégió	2q34
IDDM15	D6S283 microsatellita	6q21
1q	D1S1644	1q42
GCK	glukokináz gén	7q
Xp	DXS1068	Xp11–13

Searching for susceptibility alleles for 2DM

- Prevalence of 2DM:
  - 1985: 30 million
  - 2000: 160 million
  - 2025: 300 million (estimated)
- Concordance
  - in monozygotic twins: close to 100%
  - in dizygotic: 25-38%

Study design

UK Cambridge

Candidate Genes (71)

SNP Discovery; Genotyping (152 polymorphisms)

Case (N-517) - Control (N-517)
Single SNP Association Studies:
- test association of allele 2 with disease using dominant, additive, and recessive models

Case (N-517) - Control (N-517)
Haplotype Associations:
- identify SNPs contributing to associations
- reconstruct haplotypes using SNP selection
- test haplotype association with disease

Identify SNPs Associated with Disease

Identify Haplotypes Associated with Disease

Test SNPs and Haplotypes in QT study (N-1100) Using Genetic Model Showing Association with Disease in Case-Control Study

QT: body mass index (BMI); fasting glucose; 2hr glucose; fasting insulin; 30 min insulin incremental response

Genes with SNPs and/or Haplotypes Associated with Disease & QT

Gének és SNP-k melyek asszociáltak 2DM-mel

Gene	SNPID	SNP	Allele 2 Dominant (22 + 12)		Additive Effect		Allele 2 Recessive (11 + 12)	
			OR	*p* Value	OR	*p* Value	OR	*p* Value
SOS1	8	IVS17+53	0.58	0.0032	0.58	0.0020		
SLC2A2	21	IVS5–15	1.44	0.0117	1.35	0.0230		
	23	T198	1.39	0.0318	1.36	0.0323		
	24	T110I	1.49	0.0059	1.40	0.0114		
PPARGC1	30	T528			0.83	0.0378		
	31	G482S					0.67	0.0295
PIK3R1	42	IVS4+82	1.41	0.0090	1.34	0.0088		
INS	72	3p+9					2. O_2	0.0258
KCNJ11	74	3p+215	0.75	0.0299	0.76	0.0021	0.59	0.0027
	76	A190			0.79	0.0127	0.62	0.0260
	77	E23K					1.49	0.0333
ABCC8	79	IVS38+54			1.25	0.0131	1.68	0.0043
	81	A1369S			1.23	0.0256	1.68	0.0048
	84	IVS18–36					3.43	0.0163
	87	K649					3.90	0.0157
	89	IVS11–74					2.82	0.0480
ABCC9	100	IVS13–76					1.99	0.0339
LIPC	114	IVS1+49	0.76	0.0468	0.77	0.0291		
PYY	123	IVS3+68	1.47	0.0240	1.47	0.0157		
INSR	131	IVS6+43	1.48	0.0039	1.32	0.0119		

SNP identifiers (SNPID), OR, significance level (*p* value), and genetic model are shown. *p* values for the additive effect are for the test for a linear trend across the genotypes, which were coded as 0 = 11, 1 = 12, 2 = 22. Allele 2 dominant refers to a combination of 12 + 22 and allele 2 recessive refers to combination of 11 + 12.

DOI: 10.1371/journal.pbio.0000020.t002

- How can SNP screening be used for personalized medicine?

Gyógyszerre reagáló/nem reagáló emberek kiválasztása SNP-k segítségével

What is an SNP?

Different people can have a different nucleotide or base at a given location on a chromosome

. . . G G T A A C T G . . .

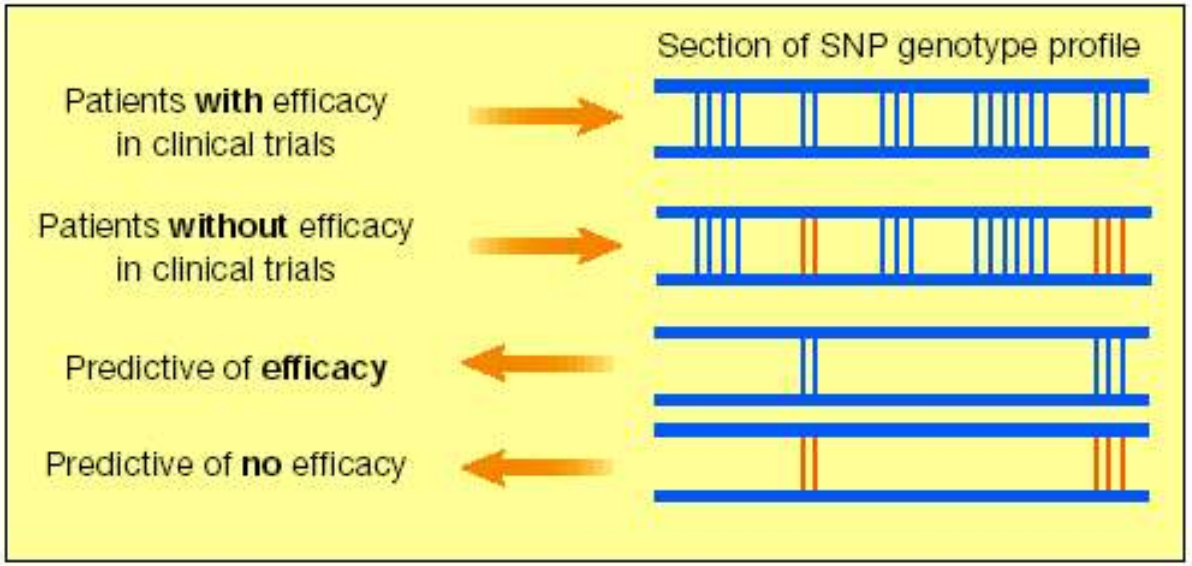
. . . G G C A A C T G . . .

What is an SNP map?

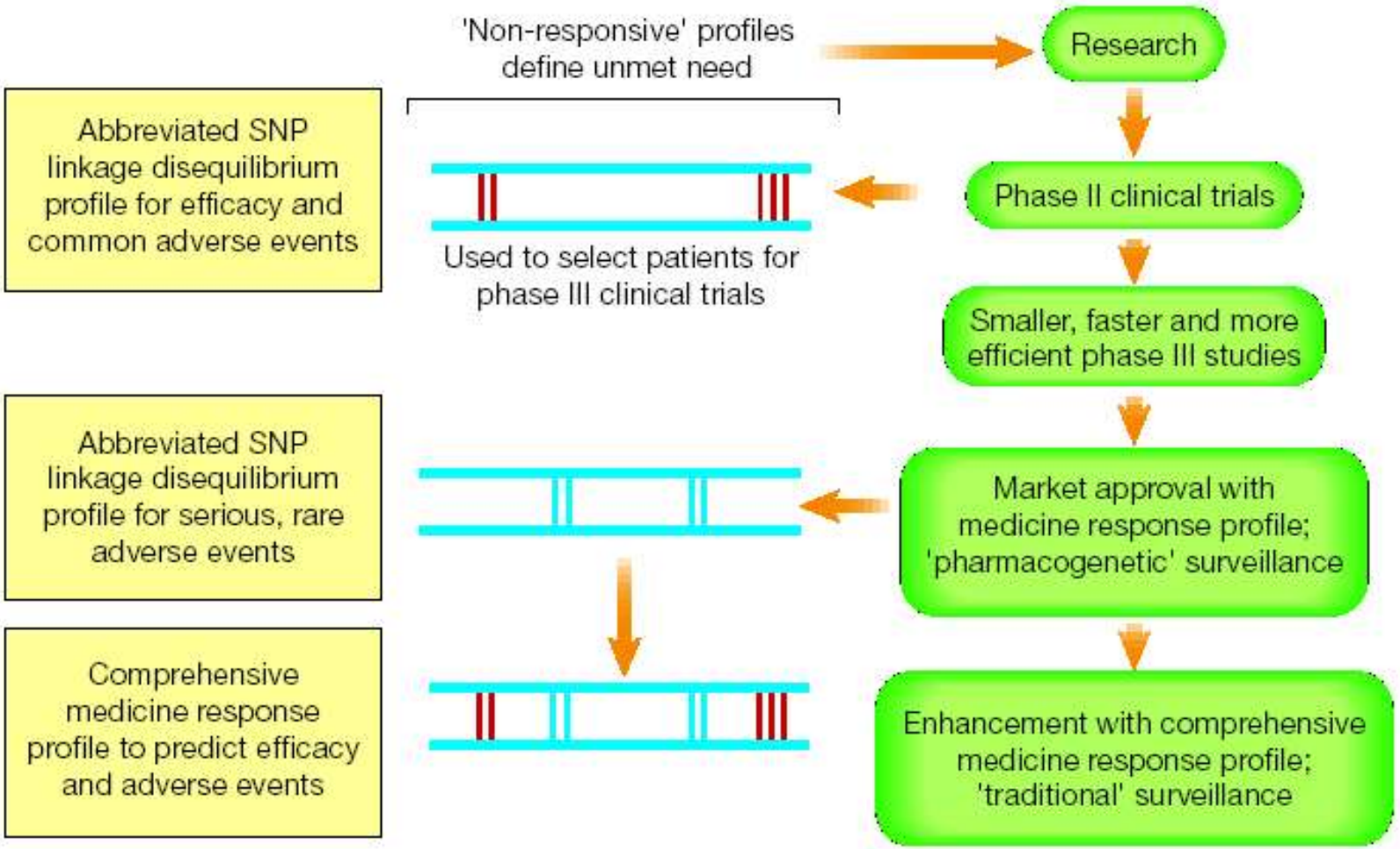
Location of SNPs on human DNA

Human DNA

How can an SNP map be used to predict medicine response?



SNP-k felhasználása a gyógyszer-jelölt klinikai tesztelésénél.



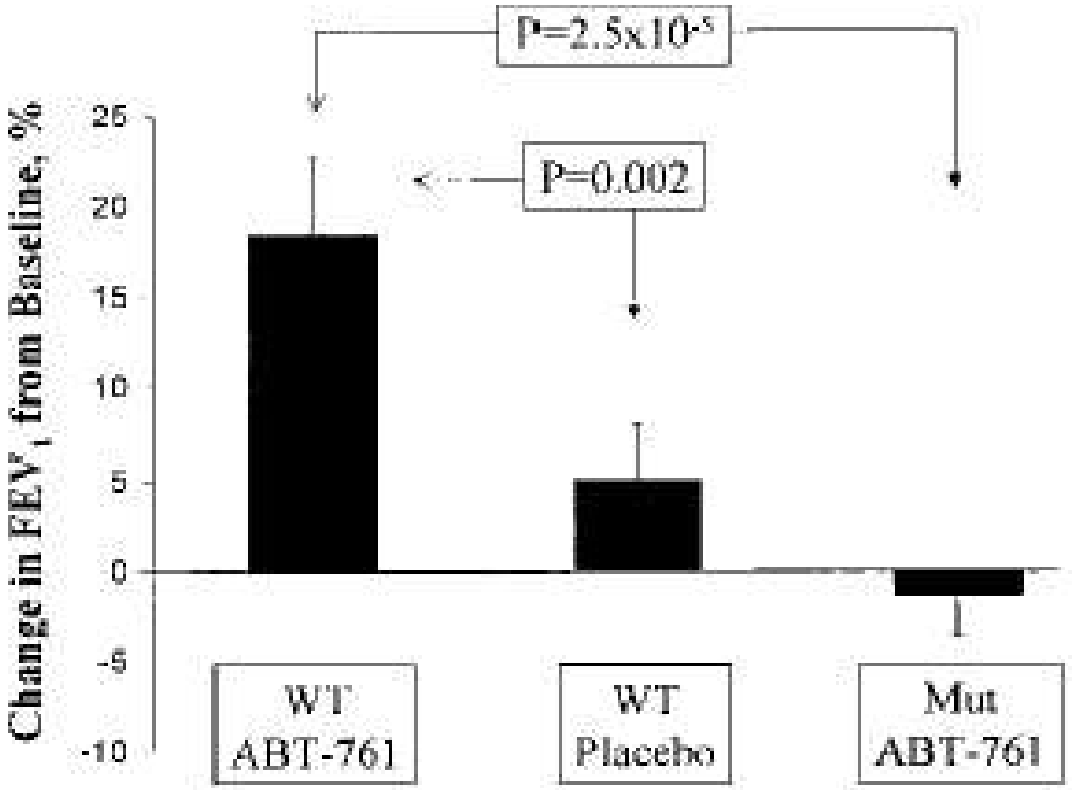
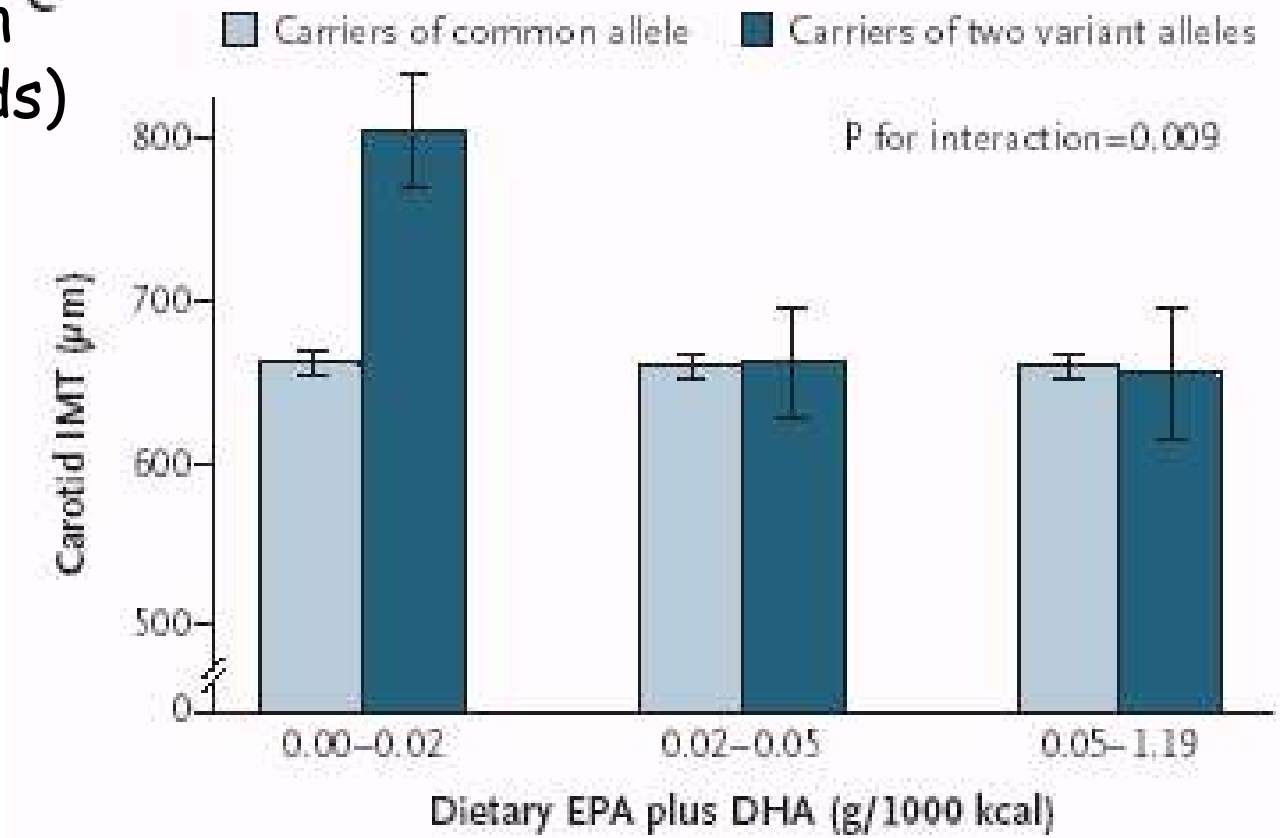


Figure 2. ALOX5 genotype predicts antileukotriene response. Shown is the outcome of a clinical trial of ABT-761, an ALOX5 inhibitor similar to zileuton, stratified by genotype. Improvement in FEV_1 from pretreatment baseline at 84 days of treatment was significantly greater for subjects possessing the wild-type (*WT*) genotype treated with ABT-761 (300 mg/day) compared with subjects possessing any ALOX5 mutant (*Mut*) allele. Modified from Drazen and coworkers (36).

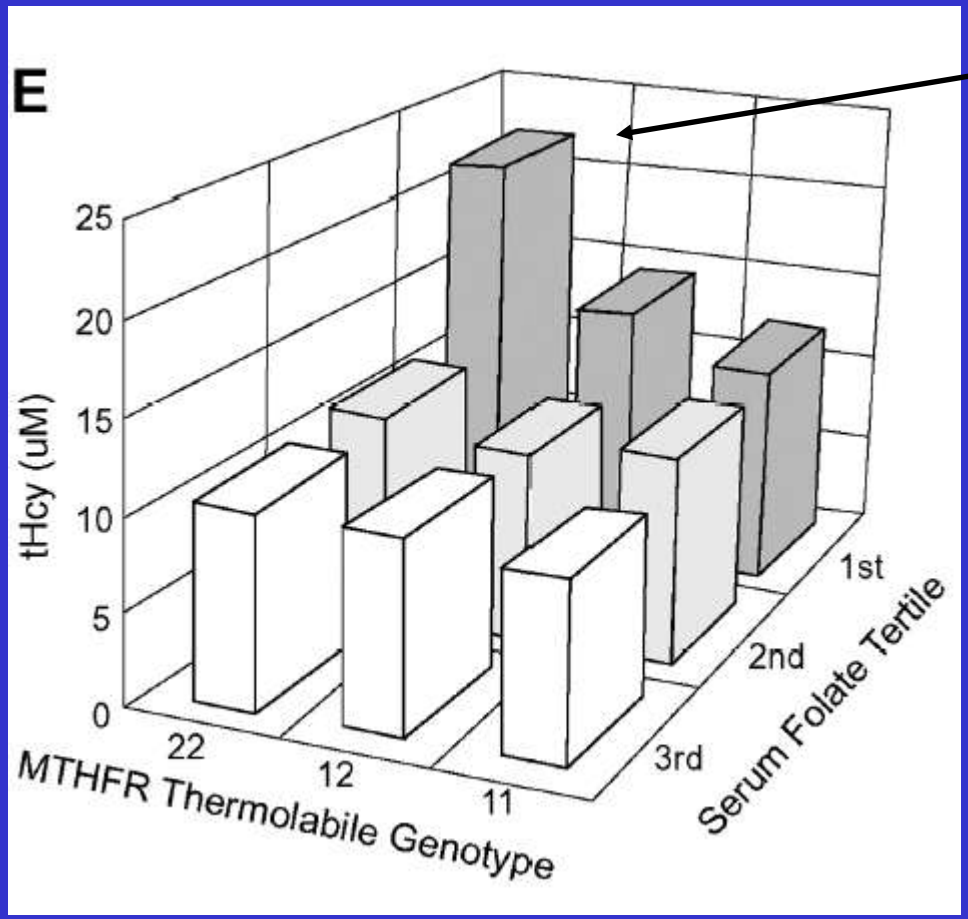
Többszörösen telítetlen omega-3 zsírsavakban gazdag táplálkozás (halolaj fogyasztása) az 5 lipoxigenáz gén promóter variáns homozigótákban megakadályozta a szív karotiszának szűkületét (szívinfarktus kialakulását)

Effect of fish oil consumption (rich in omega -3-fatty acids) on carotid intima media thickness in patients with different 5-lipoxygenase genotypes



MTHFR

- 5,10-@METHYLENETETRAHYDROFOLATE REDUCTASE; MTHFR: homocisztein lebontásának egyik kulcsenzimje.
- Termolabil változat: 677C→T (ALA222VAL) fólsav hiányosokban asszociál magas homocisztein szinttel és CAD-dal



Low folic acid consumption + MTHFR, 677TT genotype→ high serum homocystein level → high CAD risk

β2 adrenoreceptor gén polimorfizmus és az elszívott doboz cigaretta mennyiségének hatása az asztma kockázatára. Minél többet dohányzik az arg-16 homzigóta, annál nagyobb az esélye, hogy asztmás lesz. A többi genotípusúnál nem volt ilyen összefüggés. (*jelenti a szignifikáns különbséget)