Principles of Pharmacogenomics

Sherin Shaaban MD PhD FACMG
Assistant Professor of Pathology, University of Utah School of Medicine
Medical director of Pharmacogenomics, Molecular Genetics and Genomics, ARUP Laboratories

Learning objectives

Define what pharmacogenetics/genomics (PGx) is

Discuss pharmacogenes and possible genetic variations

Learn what molecular methodologies are used in the field

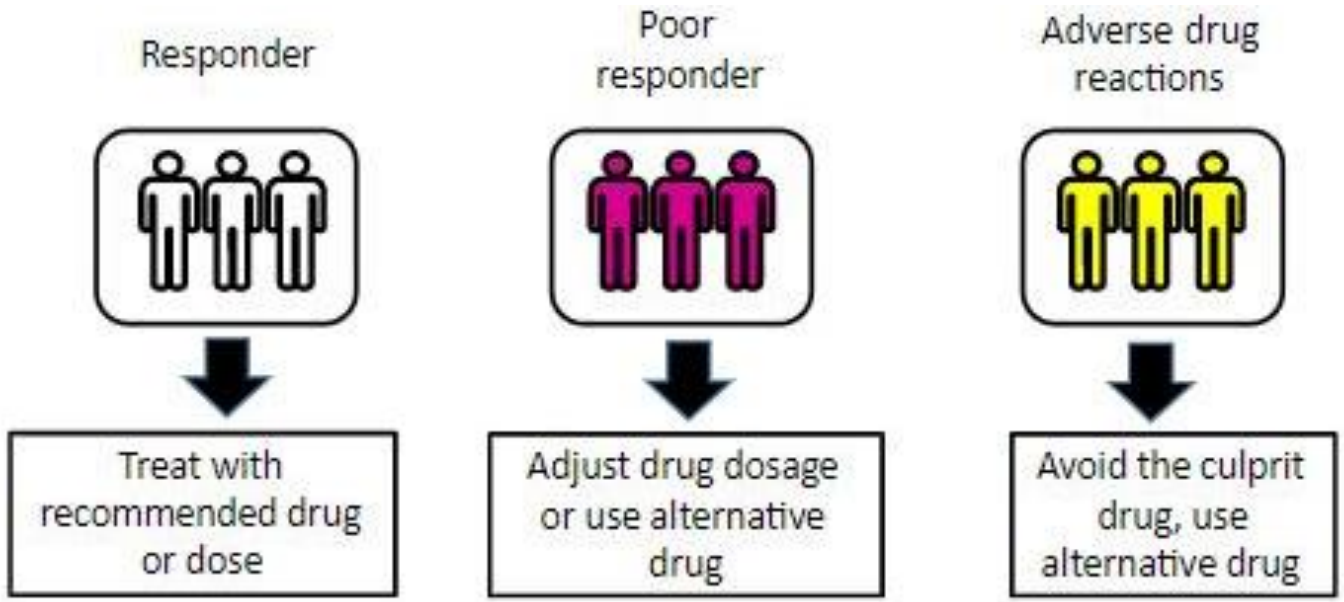
Discuss PGx Implementation challenges

What is Pharmacogenetics?

Layman's language

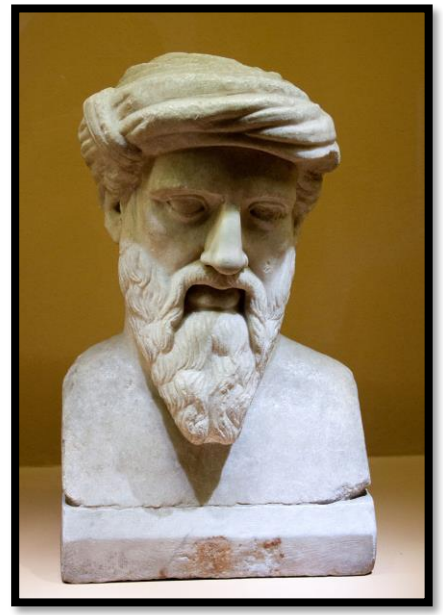
- Pharmacogenetics (sometimes called pharmacogenomics) is a field of research that studies how a person's genes affect how he or she responds to medications. Its long-term goal is to help doctors select the drugs and doses best suited for each person.

- It is part of the field of precision medicine, which aims to treat each patient individually.



National Institute of General Medical Sciences

History



Pythagoros (570-495 BC):

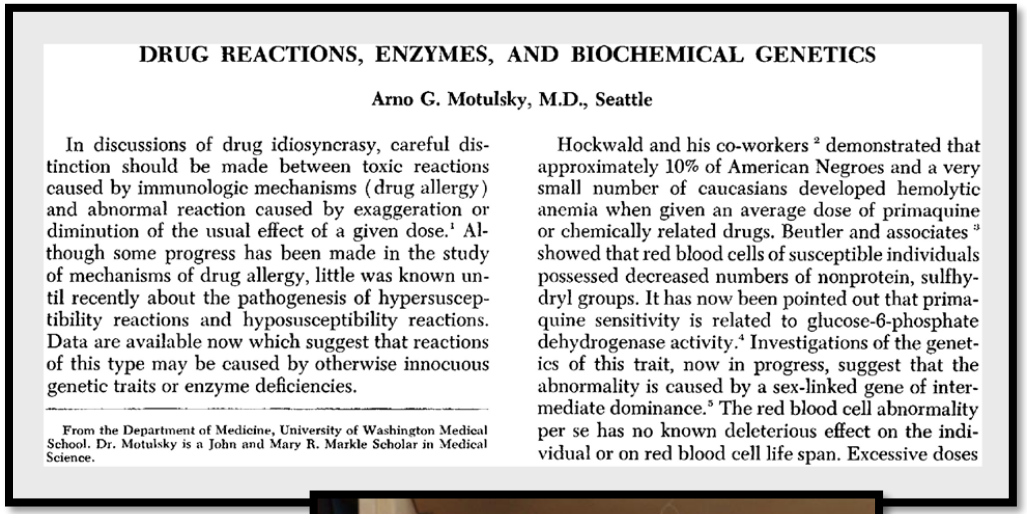
The first pharmacogenetics observation was made by
("be far from the fava beans consumption")
He noticed that fava beans caused a condition (*now known as acute hemolytic anemia*) in certain people but not in others

Archibald Garrod (1857-1936) discovered inherited alkaptonuria.

Established a concept: the metabolism of molecular compounds can be altered by inherited genetic factors and cause an abnormal accumulation of "intermediate" metabolites

Meletis J. Favism: a brief history from the "abstain from beans" of Pythagoras to the present. Arch Hellenic Med 2012; 29: 253–263

1957



DRUG REACTIONS, ENZYMES, AND BIOCHEMICAL GENETICS

Arno G. Motulsky, M.D., Seattle

In discussions of drug idiosyncrasy, careful distinction should be made between toxic reactions caused by immunologic mechanisms (drug allergy) and abnormal reaction caused by exaggeration or diminution of the usual effect of a given dose.[1] Although some progress has been made in the study of mechanisms of drug allergy, little was known until recently about the pathogenesis of hypersusceptibility reactions and hyposusceptibility reactions. Data are available now which suggest that reactions of this type may be caused by otherwise innocuous genetic traits or enzyme deficiencies.

From the Department of Medicine, University of Washington Medical School. Dr. Motulsky is a John and Mary R. Markle Scholar in Medical Science.

Hockwald and his co-workers [2] demonstrated that approximately 10% of American Negroes and a very small number of caucasians developed hemolytic anemia when given an average dose of primaquine or chemically related drugs. Beutler and associates [3] showed that red blood cells of susceptible individuals possessed decreased numbers of nonprotein, sulfhydryl groups. It has now been pointed out that primaquine sensitivity is related to glucose-6-phosphate dehydrogenase activity.[4] Investigations of the genetics of this trait, now in progress, suggest that the abnormality is caused by a sex-linked gene of intermediate dominance.[5] The red blood cell abnormality per se has no known deleterious effect on the individual or on red blood cell life span. Excessive doses

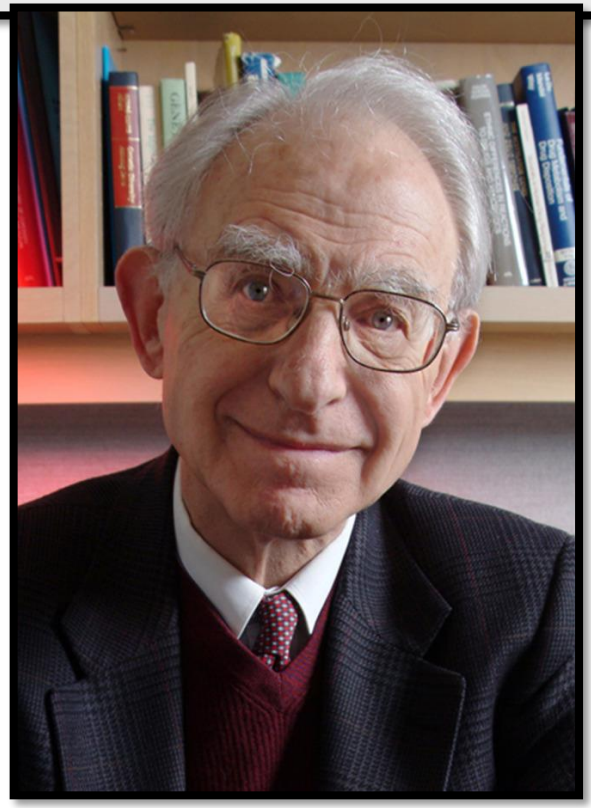
Arno Motulsky (1923–2018)

- Earned his title of the father of pharmacogenetics

"It is not unlikely that some drug sensitivity reactions … be produced by (genetic) mechanisms"

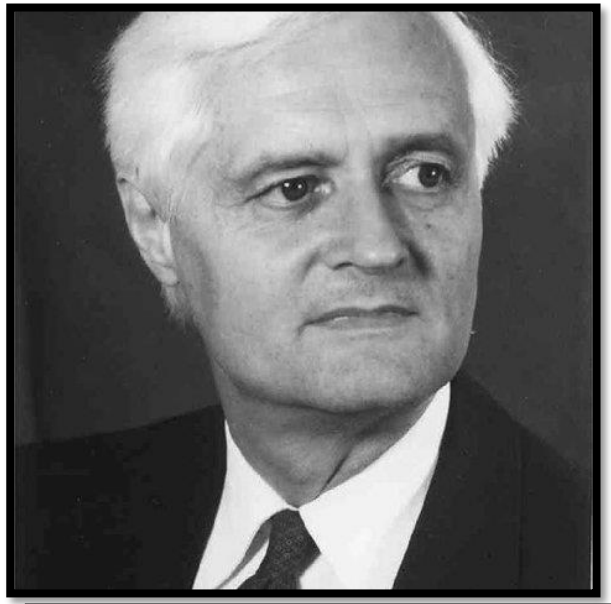
Two major observations:

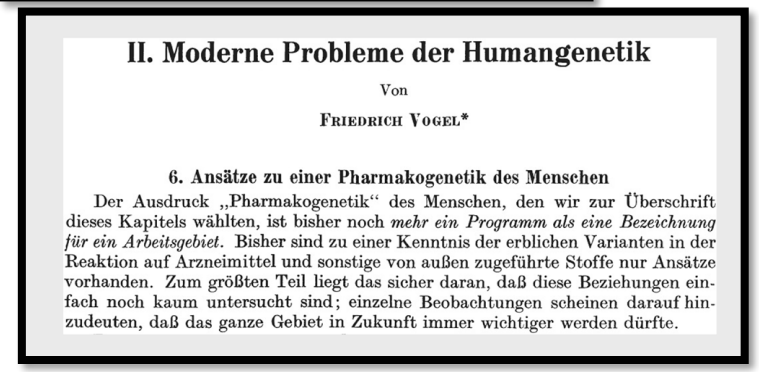
- Some soldiers developed hemolytic anemia when given Primaquine (antimalarial)
- Some patients developed prolonged apnea when given succinylcholine



Friedrich Vogel (1925–2006)

He developed the term “pharmacogenetics” in 1959, as the close contact between genetics and pharmacology





II. Moderne Probleme der Humangenetik

Von

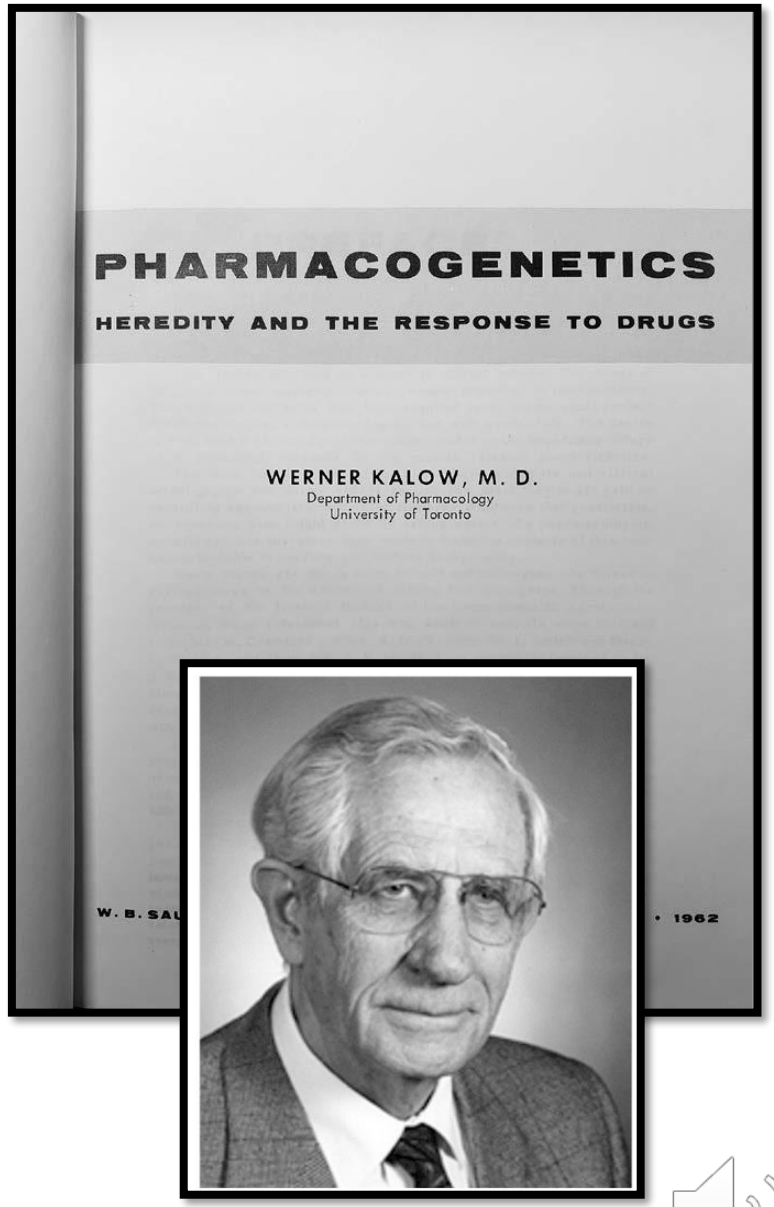
FRIEDRICH VOGEL*

6. Ansätze zu einer Pharmakogenetik des Menschen

Der Ausdruck „Pharmakogenetik“ des Menschen, den wir zur Überschrift dieses Kapitels wählten, ist bisher noch *mehr ein Programm als eine Bezeichnung für ein Arbeitsgebiet*. Bisher sind zu einer Kenntnis der erblichen Varianten in der Reaktion auf Arzneimittel und sonstige von außen zugeführte Stoffe nur Ansätze vorhanden. Zum größten Teil liegt das sicher daran, daß diese Beziehungen einfach noch kaum untersucht sind; einzelne Beobachtungen scheinen darauf hinzudeuten, daß das ganze Gebiet in Zukunft immer wichtiger werden dürfte.

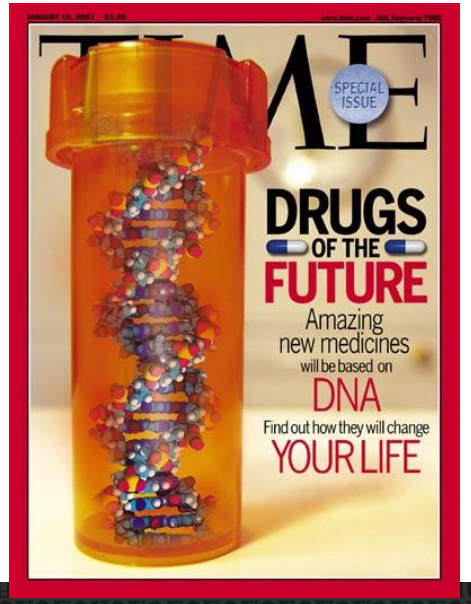
Werner Kalow (1917–2008)

- The first book entirely dedicated to the field.
- Summarizes all the work and available knowledge of that time.
- Helped change pharmacogenetics from a subspecialty to an entire field



What is Pharmacogenetics? Non-Layman's language

- Pharmacogenetics is the branch of pharmacology and genetics concerned with the inter-individual metabolic and therapeutic responses to a given medication.



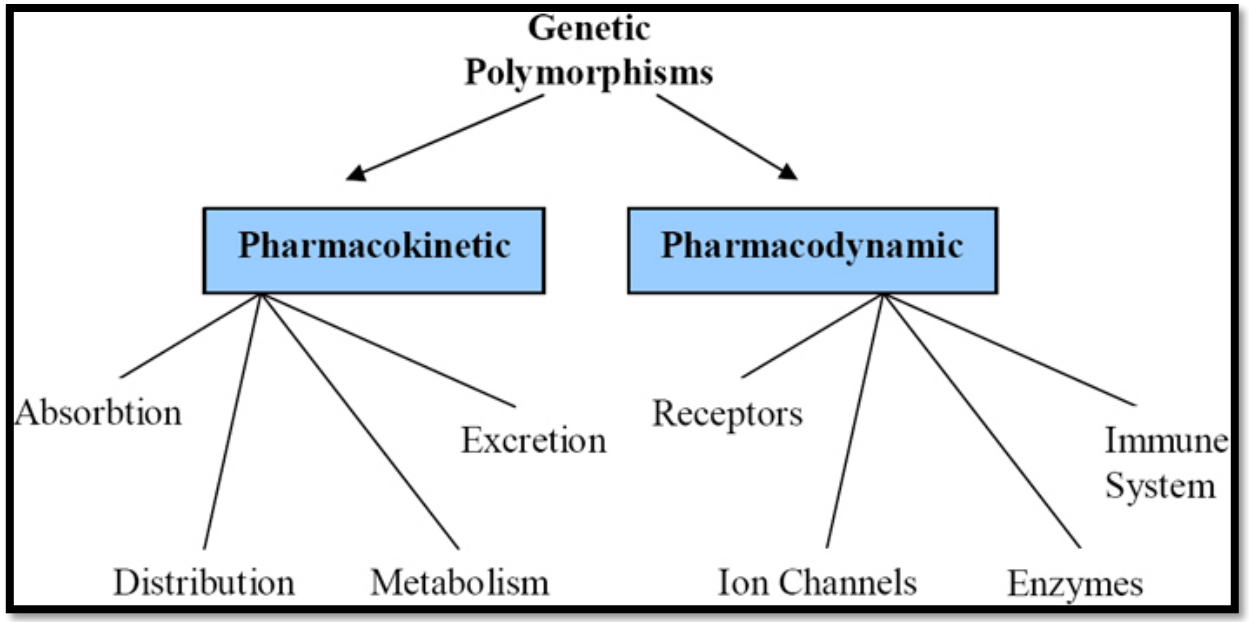
Encompasses two closely related fields:

Pharmacokinetics: Studies the absorption, distribution, metabolism, and excretion pathways of the drug.

'What the body does to the drug'. BODY → DRUG

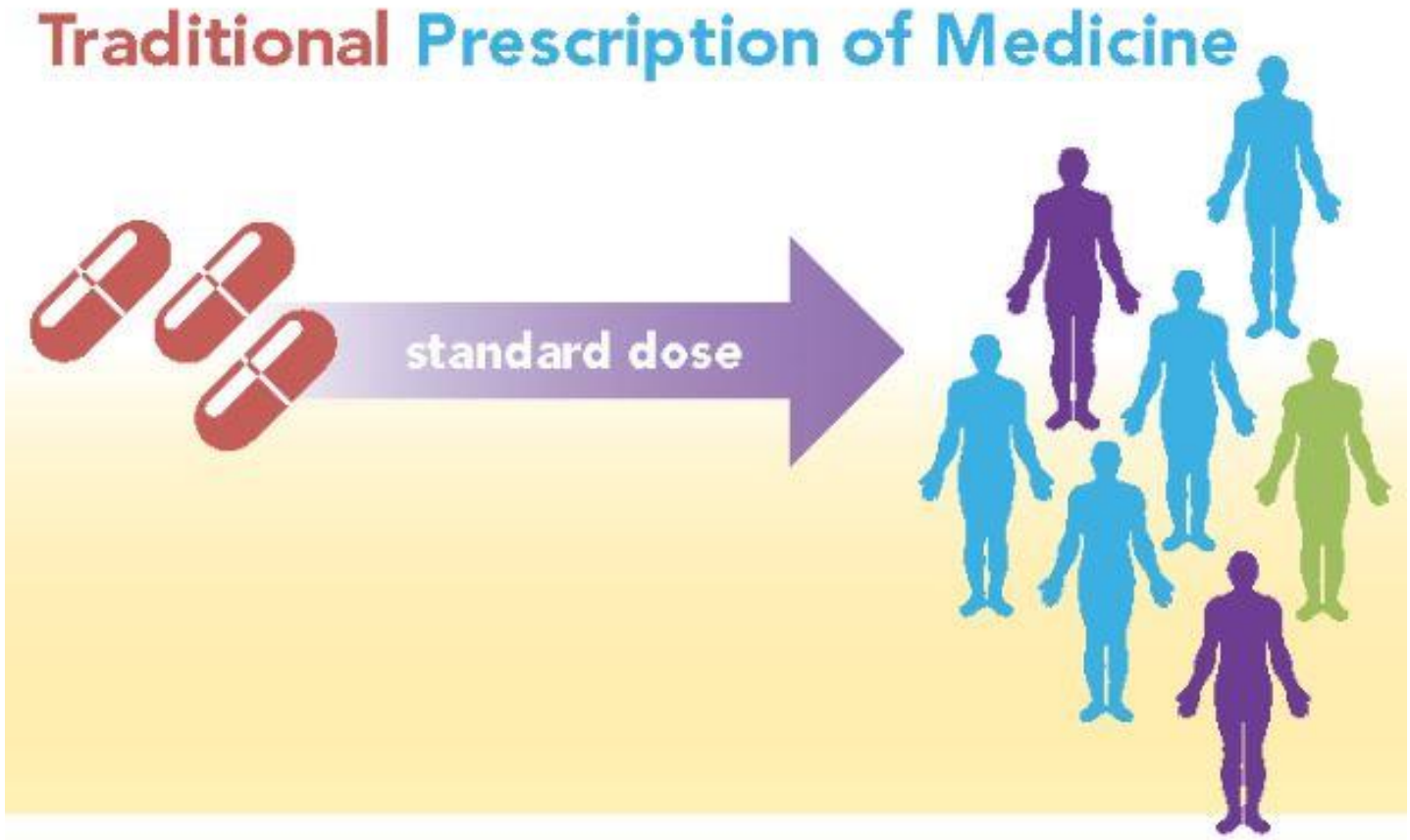
Pharmacodynamics: Concerned with the drug effects on the organism as a whole

'What the drug does to the body' DRUG → BODY



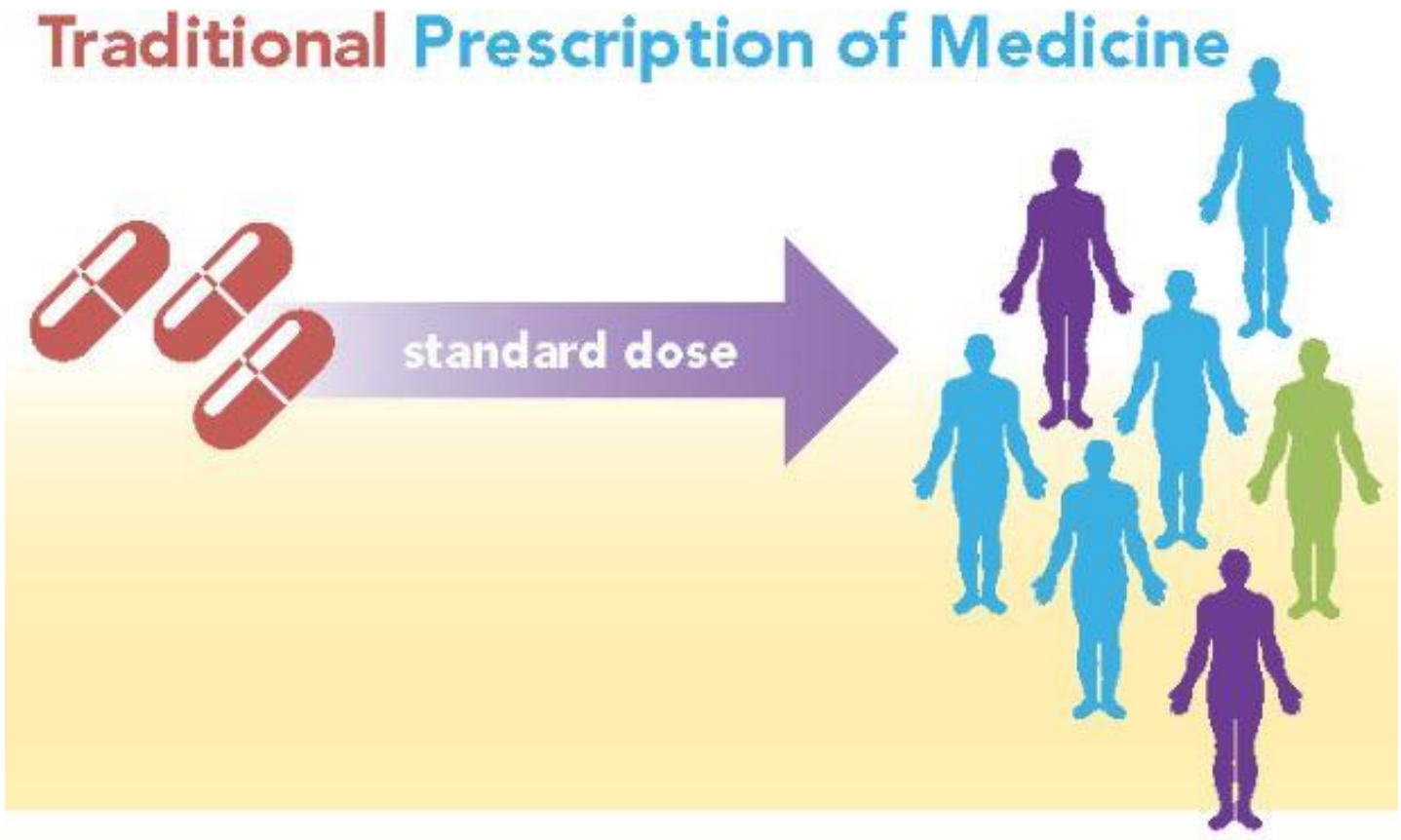
Current practice

One size fits all



Current practice

One size fits all



Dose works for many

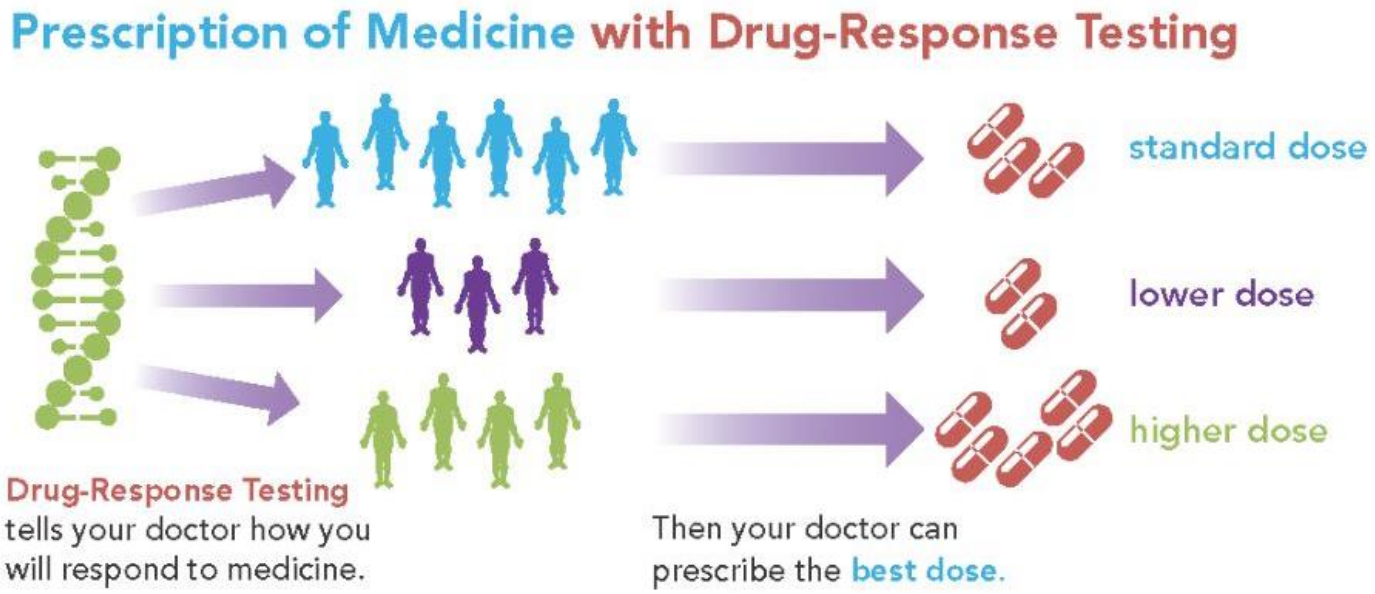
Need less dose

Need higher dose, or alternative drug

Goals of Pharmacogenetics

- Maximize drug efficacy
- Minimize drug toxicity
- Predict patients who will respond to intervention
- Aid in new drug development

Proactive Approach



Application

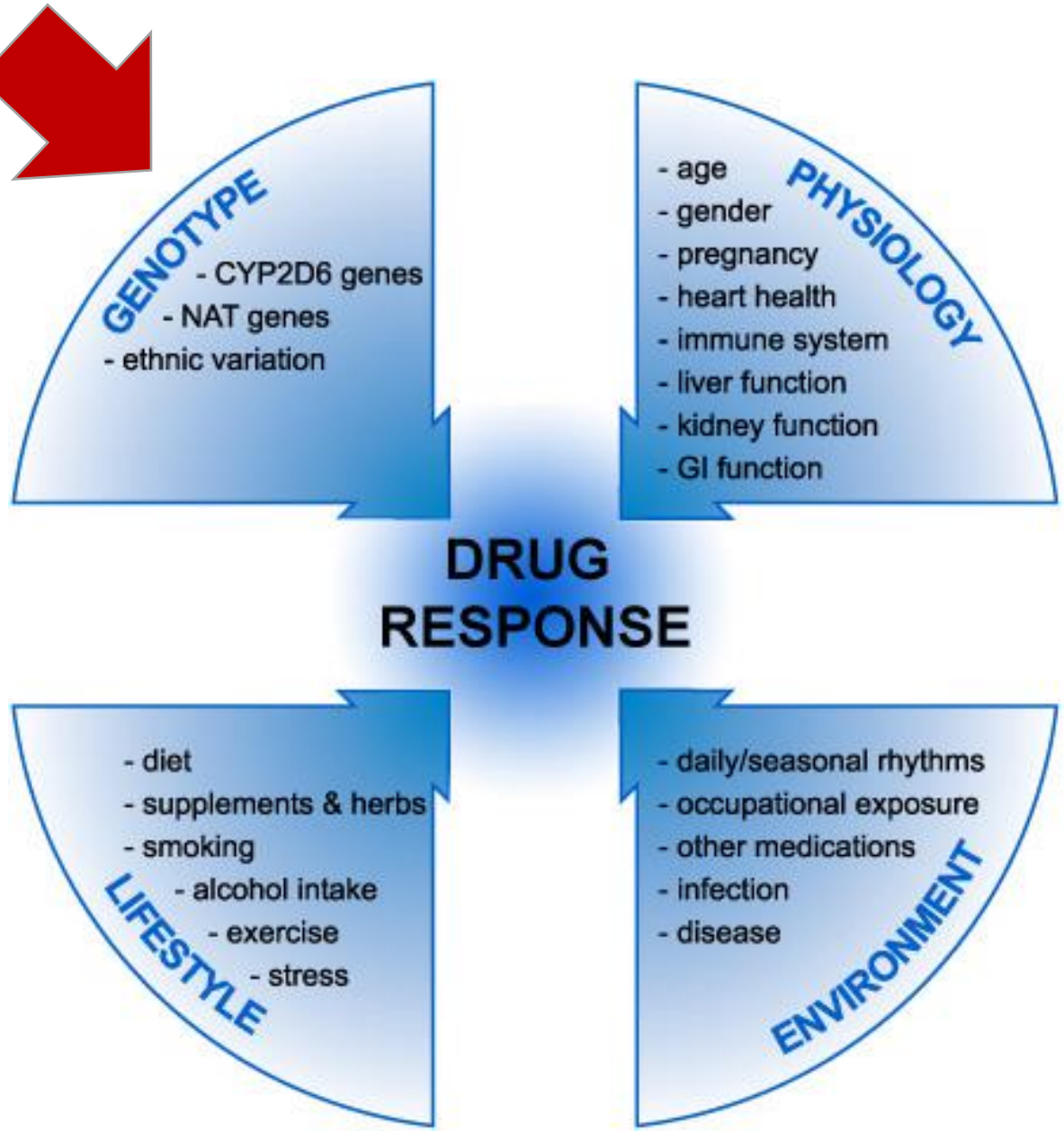
Drug selection/avoidance

- Who is at high risk of a serious ADR
- Who is not likely to respond

Dose optimization

- Who is likely to be sensitive or resistant to a drug
- What dose and what frequency is needed

Factors Influencing Drug Response



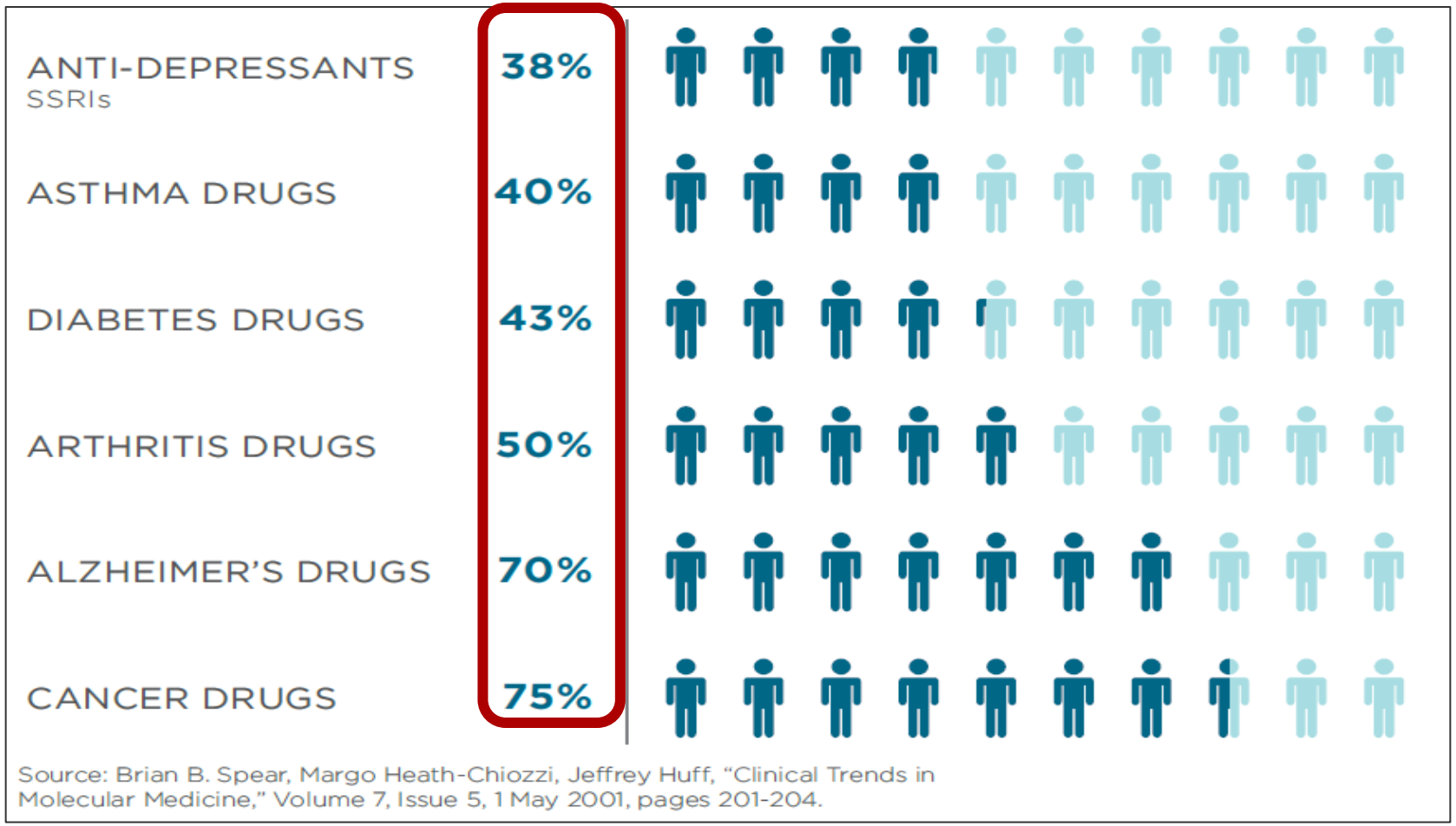
Frequency of PGx variants

Around 97–98% of people have at least one actionable variant in their drug-related genes.

The possibility of the presence of a genetic variant [mainly loss of function (LOF) variant] in pharmacogenes is 93% for every individual

Genetic variation in human drug-related genes (2017).

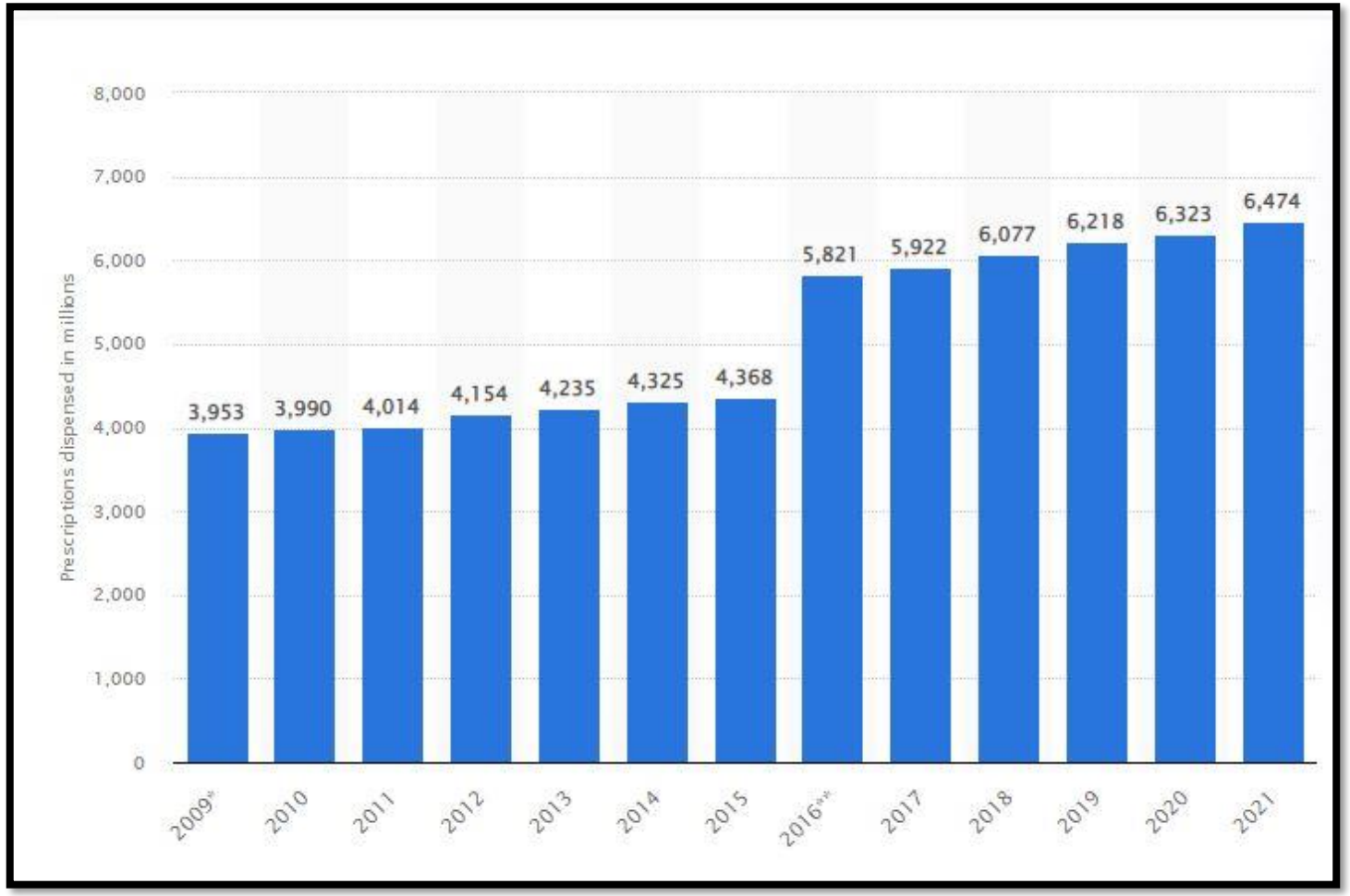
Percentage of patient population for which a drug class is effective



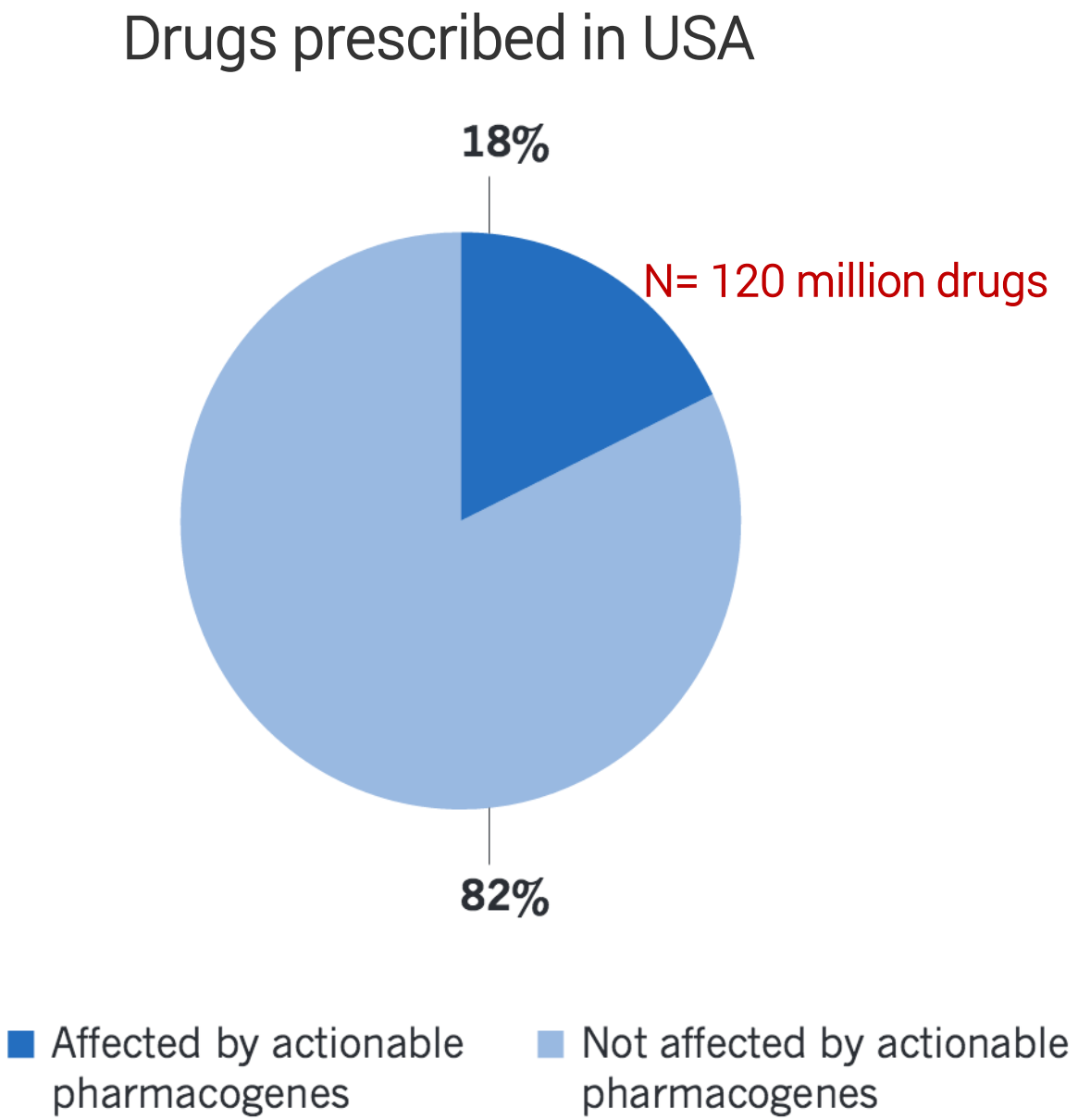
Source: Brian B. Spear, Margo Heath-Chiozzi, Jeffrey Huff, "Clinical Trends in Molecular Medicine," Volume 7, Issue 5, 1 May 2001, pages 201-204.

http://www.personalizedmedicinecoalition.org/

- In 2009 the number of prescriptions dispensed was near 3.95 billion.
- In 2021 the number of prescriptions dispensed was around 6.47 billion.

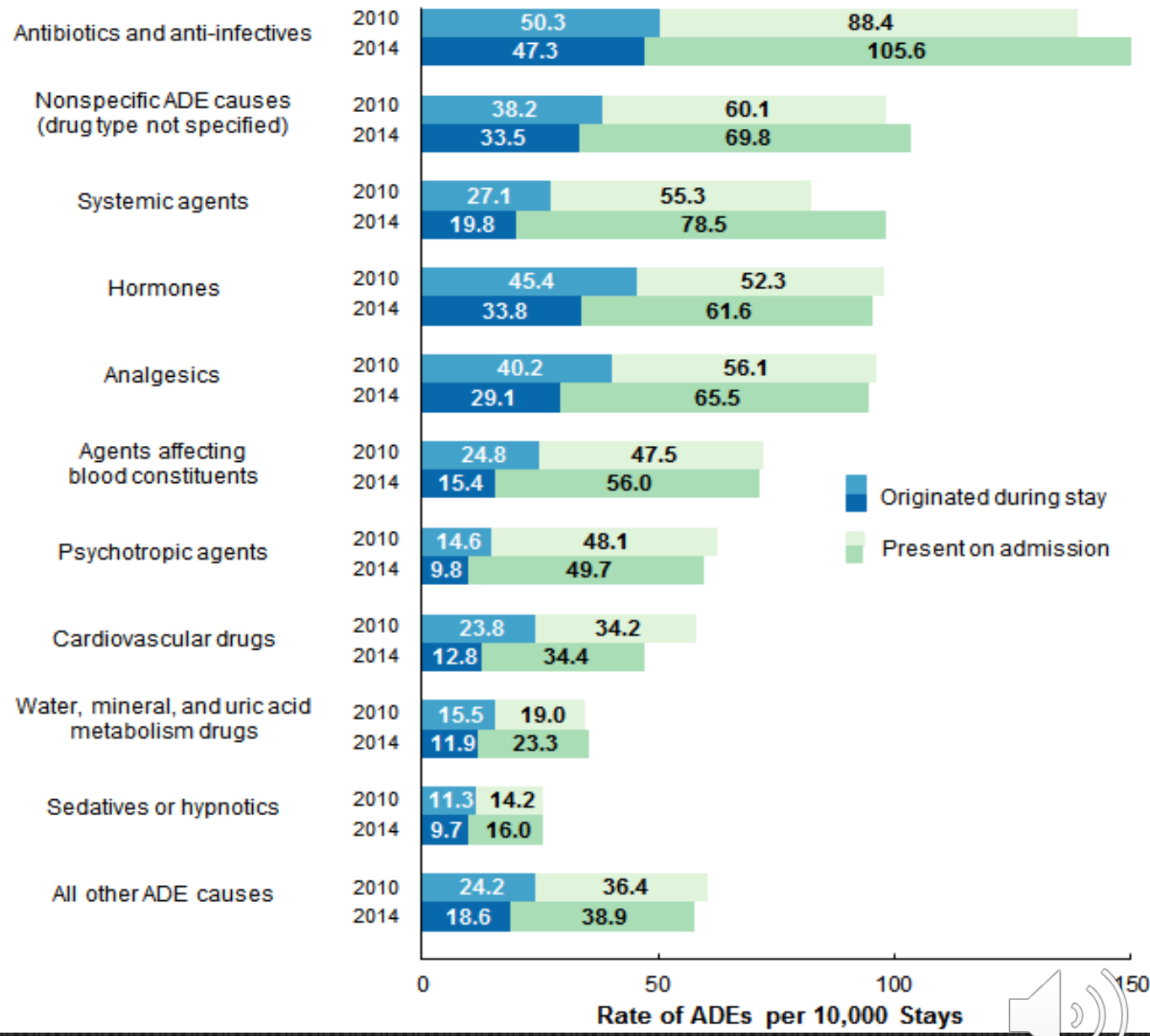


- In 2009 the number of prescriptions dispensed was near 3.95 billion.
- In 2021 the number of prescriptions dispensed was around 6.47 billion.



Adverse Drug Events (ADE)

- ADEs: leading type of nonsurgical adverse event occurring in hospitals in the US
- Some ADEs are the result of medication errors, but also may occur when medications are taken correctly.
- Patients hospitalized with an ADE have an increased length of stay, higher costs, and increased risk of in-hospital death.

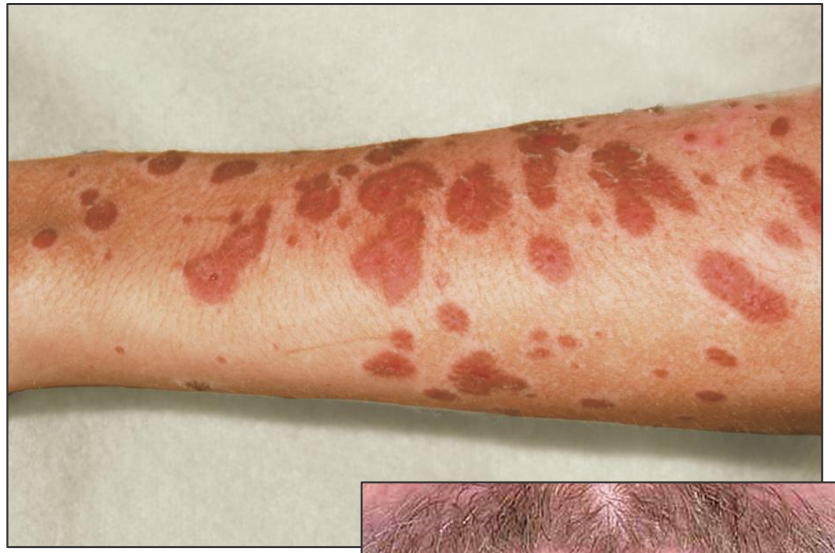


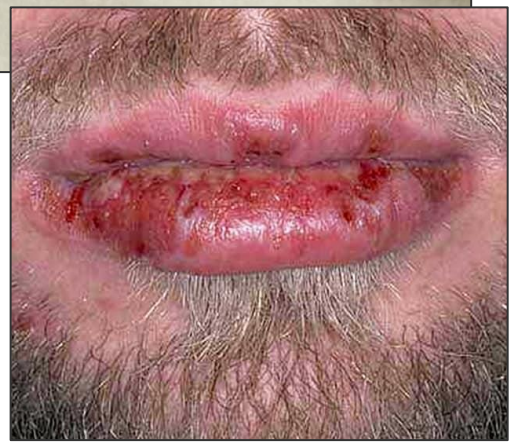
Adverse Drug Events: ADE

- 4th Leading cause of death ahead of pulmonary disease, diabetes, AIDS, pneumonia, accidents and automobile deaths.
- 100.000 deaths due to ADE per year
- $136 Billion Costs of ADEs per year
- Large percentage is <u>PREVENTABLE</u>

Adverse Drug Events: ADE

- Dose-related (Augmented)
- Non-dose-related (Bizarre)
- Dose-related and time-related (Chronic)
- Time-related (Delayed)
- Withdrawal (End of use)
- Failure of therapy (Failure)





- **Examples:** Rashes, jaundice, anemia, leucopenia, kidney damage, nerve injury or anaphylaxis

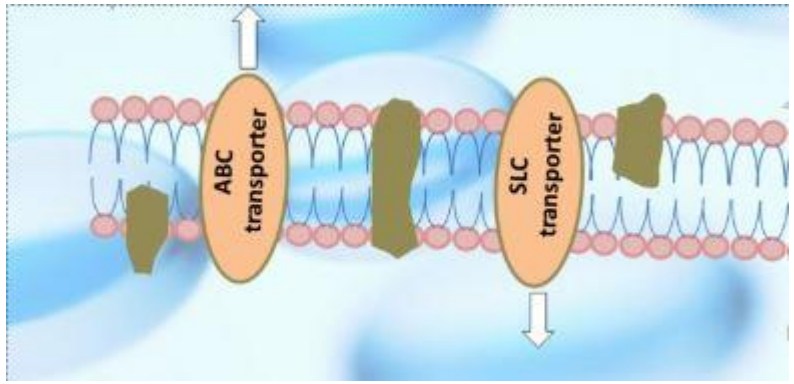
***Life threatening:** Stevens-Johnson syndrome or toxic epidermal necrolysis

How do we study pharmacogenetics?

Pharmacogenes (focus on functional variants)

Drug transporters: cell surface proteins

The solute carrier (SLC) transporters e.g SLCO1B1
The ATP-binding cassette (ABC) transporters e.g. ABCB1



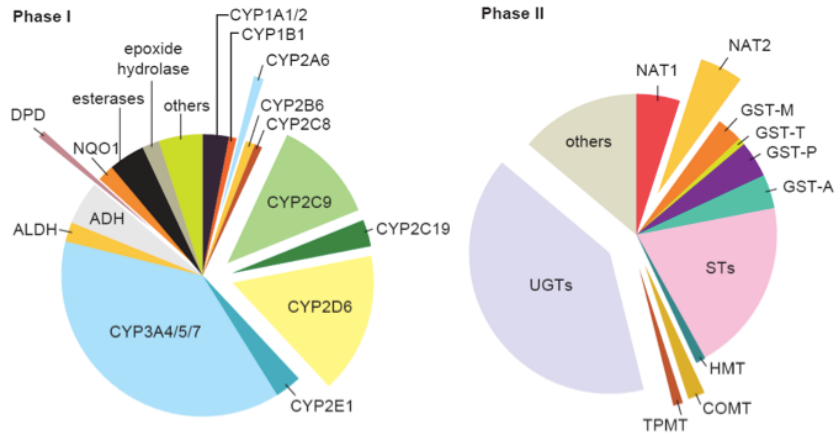
How do we study pharmacogenetics?

Pharmacogenes

Drug transporters: cell surface proteins
The solute carrier (SLC) transporters e.g SLCO1B1
The ATP-binding cassette (ABC) transporters e.g. ABCB1

Drug Metabolizing Enzymes

Phase 1 metabolizing enzymes: e.g. CYP2D6, CYP3A5
Phase 2 metabolizing enzymes: TPMT, UGT1A1



Phase 1 – Oxidation, Reduction, Methylation, Hydroxylation, Deamination
Phase 2 – Conjugation (D-glucuronidation, O-sulfation, N-acetylation, O-, N-, S-methylation, glutathione, amino acid conjugation)

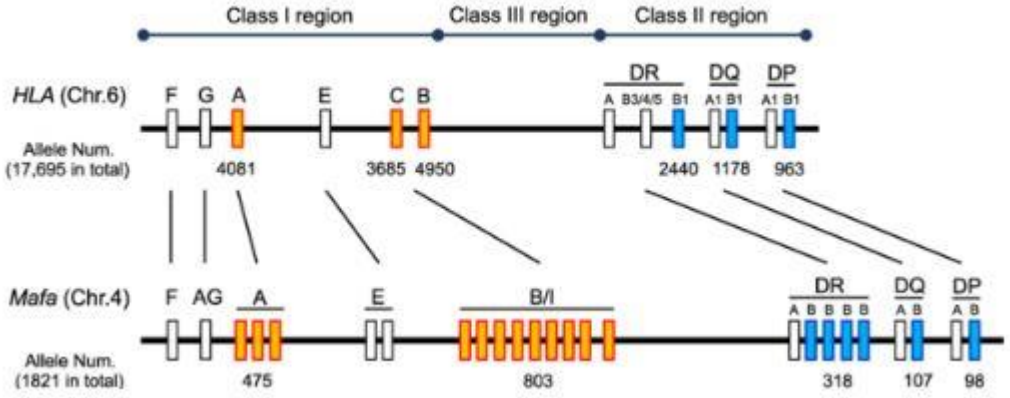
How do we study pharmacogenetics?
Pharmacogenes (focus on functional variants)

Drug transporters: cell surface proteins
The solute carrier (SLC) transporters e.g SLCO1B1
The ATP-binding cassette (ABC) transporters e.g. ABCB1

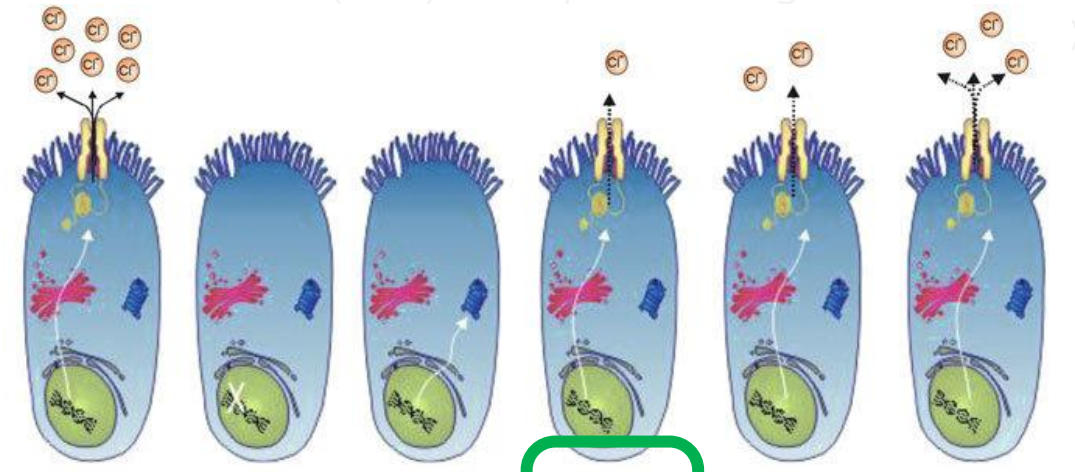
Drug Metabolizing Enzymes
Phase 1 metabolizing enzymes: e.g. CYP2D6, CYP3A5
Phase 2 metabolizing enzymes: TPMT, UGT1A1

Major histocompatibility complex genes (HLA alleles)

Some associated with increased risk for an allergic response to certain medications leading to Stevens-Johnson syndrome or toxic epidermal necrolysis.
HLA-B*57:01, HLA-B*15:02



How do we study pharmacogenetics?
Pharmacogenes (focus on functional variants)



Normal	Class I	Class II	Class III	Class IV	Class IV
Defect	Defective Synthesis	Defective processing or maturation	Defective regulation	Defective conductance	Reduced synthesis and stability
Therapy	Readthrough	Correctors (+ potentiators)	Potentiators	Potentiators	Potentiators Spicing modulators
Mutations (examples)	**G542X** W1282X R553X, E882X 621 + 1G → T	**ΔF508** N1303K ΔI507 R1066C	**G551D** G551S G178R G1244E	R117H R347P R1070W	3272 6A→G A455E D565G 3849+1kb C→T

Drug targets

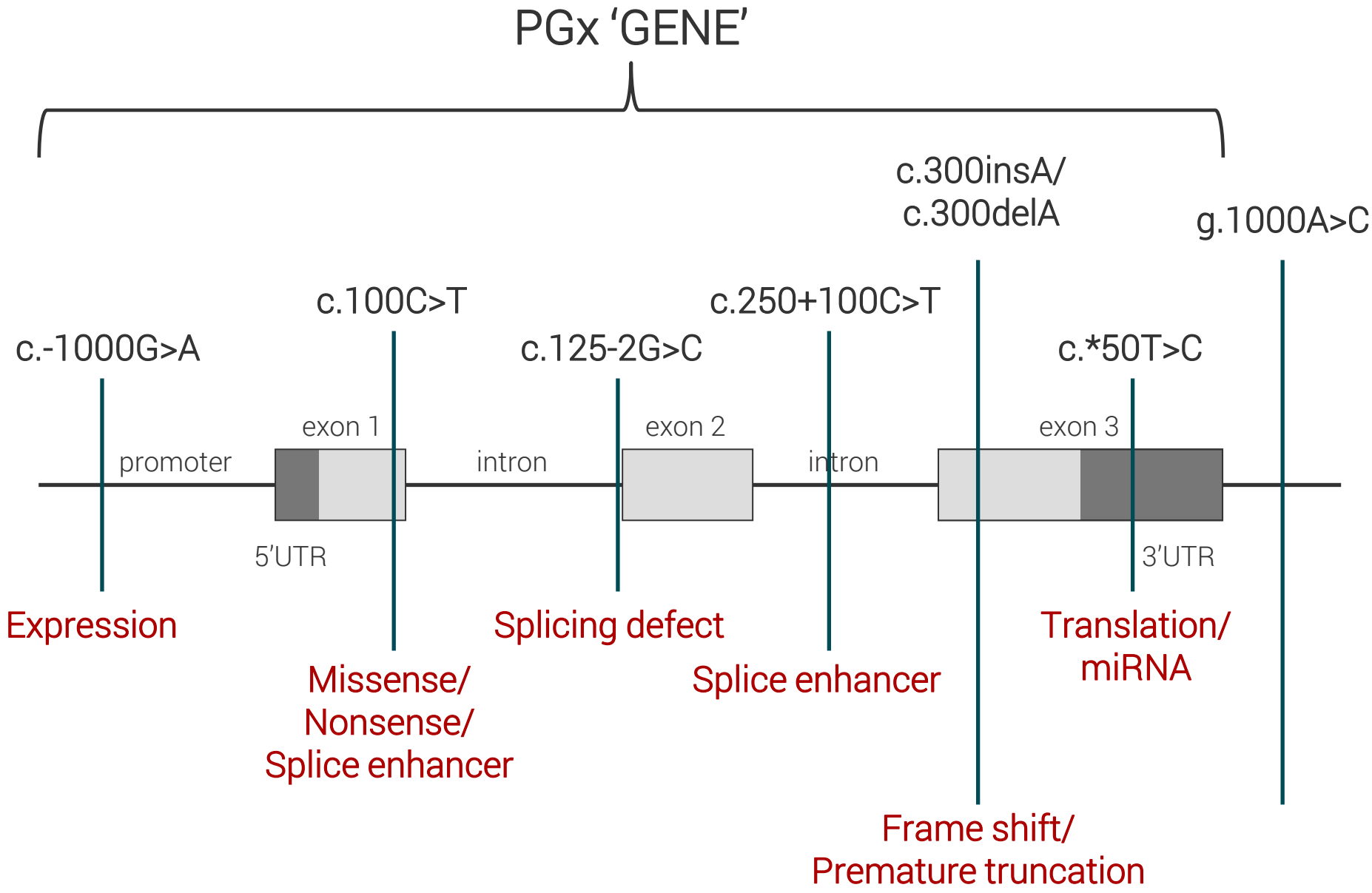
- Molecules or pathways that a drug is designed to affect in order to deliver therapy.
- Work by altering the amount of the target protein or by delivering therapy only to specific genetic variants.
- **e.g.** Ivacaftor to treat cystic fibrosis

Pharmacogenes variations

- SNPs
- Small Insertion and deletions
- Structural variants
- CNV: e.g. *CYP2D6*
- Tandem repeats e.g. *UGT1A1*

Table 2 Counts of Ensembl consequence type for variants mapped to canonical transcripts of PGRNseq captured genes

ENSEMBL consequence type	IN PGx	IN 1KG	IN EXAC
Upstream Gene Variant	6,094	2,122	23
Intron Variant	5,542	2,016	460
Missense Variant	4,806	1,485	1,792
3 Prime UTR Variant	4,245	1,539	65
Downstream Gene Variant	3,574	1,239	44
Synonymous Variant	3,147	1,335	1,255
5 Prime UTR Variant	931	287	59
Missense Variant, Splice Region Variant	147	48	62
Splice Region Variant, Intron Variant	142	60	49
Stop Gained	97	20	31
Splice Region Variant, Synonymous Variant	90	—	36
Splice Acceptor Variant	18	5	3
Splice Donor Variant	15	3	6
Splice Region Variant, 5 Prime UTR Variant	14	3	3
Initiator Codon Variant	11	2	2
Stop Gained, Splice Region Variant	3	1	1
Stop Lost	2	—	—
Stop Retained Variant	1	1	—
Splice Region Variant, 3 Prime UTR Variant	1	1	—
Total	28,880	10,167	3,891


PGx 'GENE'
c.-1000G>A
c.100C>T
c.125-2G>C
c.250+100C>T
c.300insA/
c.300delA
c.*50T>C
g.1000A>C
promoter
exon 1
intron
exon 2
intron
exon 3
5'UTR
3'UTR
Expression
Missense/
Nonsense/
Splice enhancer
Splicing defect
Splice enhancer
Frame shift/
Premature truncation
Translation/
miRNA

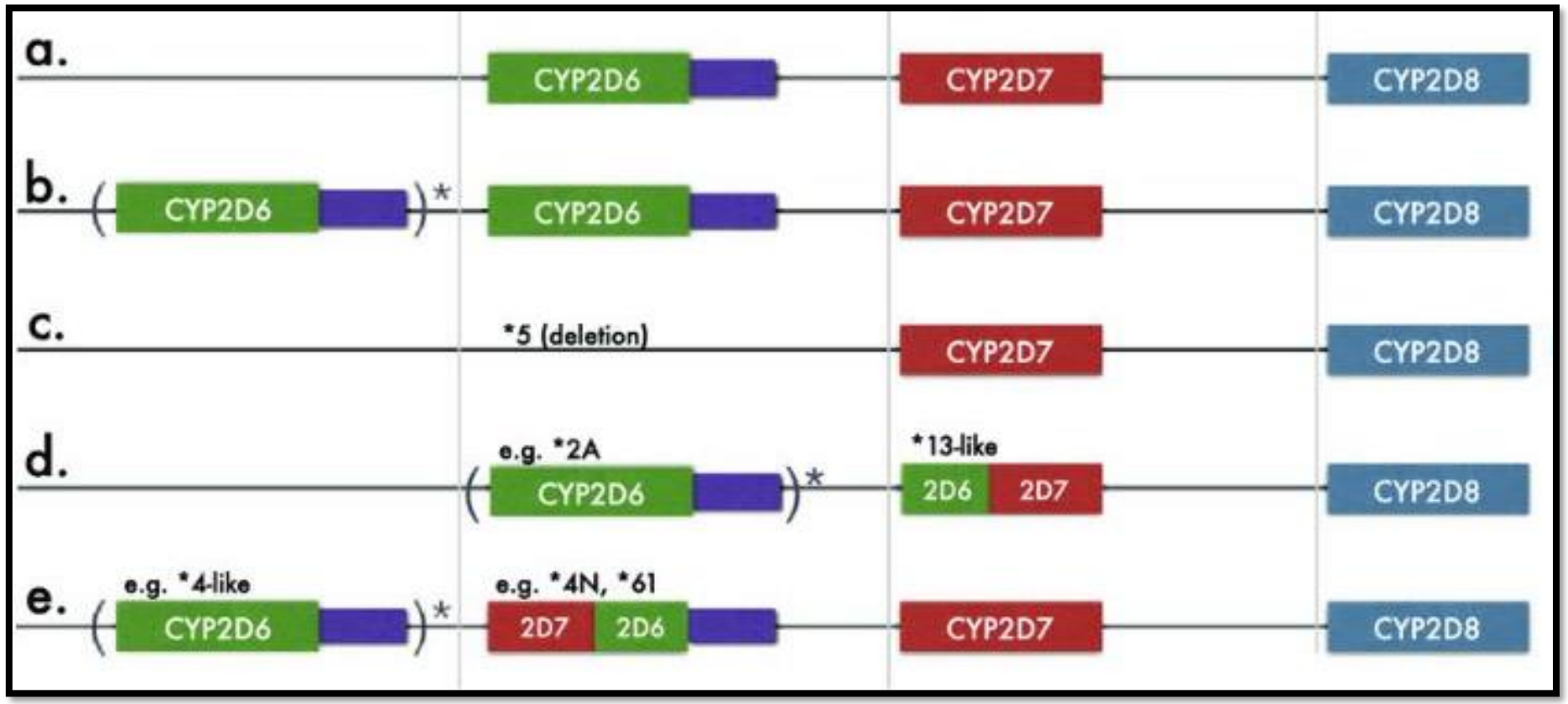
Example variants and predicted function

Cytochrome P450 2D6 (*CYP2D6*) alleles and their effects on CYP2D6 enzyme activity

CYP2D6 alleles	Allele designation	Enzyme activity	Allele abbreviation
*1, *2, *33, *35	Normal or wild type	Normal	EM
*3, *4, *5-*8, *11-*16, *18-*21, *36, *38, *40, *42, *44, *56, *62	Null	No protein, inactive or negligible	PM
*9, *10, *17, *29, *41, *59	Reduced activity	Decreased	IM
*22-*28, *30-*32, *34, *37, *39, *43, *45-*55	Unknown activity	Unknown	Not applicable

Hoskins et al. Nature Reviews Cancer 2009

Example copy number variation/re-arrangements e.g CYP2D6



Normal

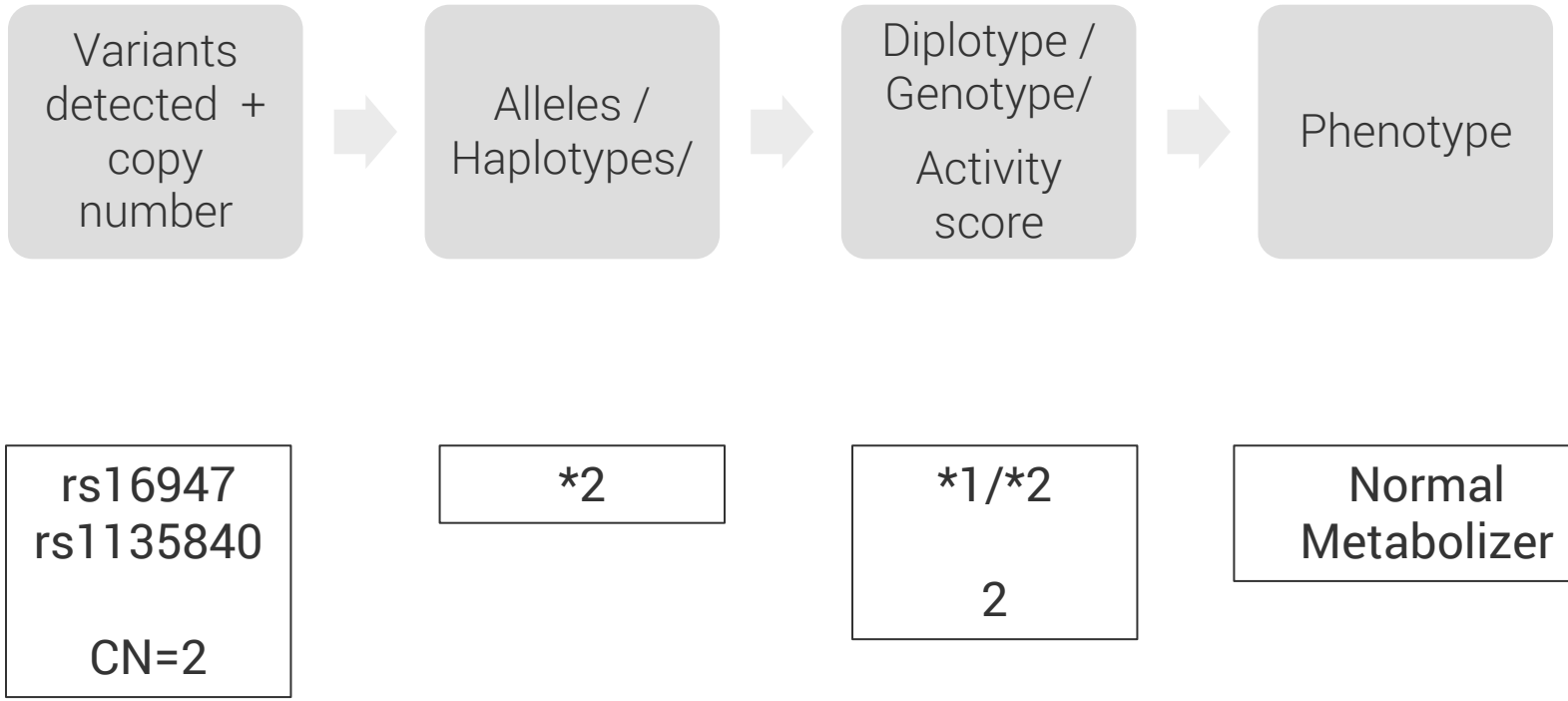
Duplication

Deletion

Hybrid allele

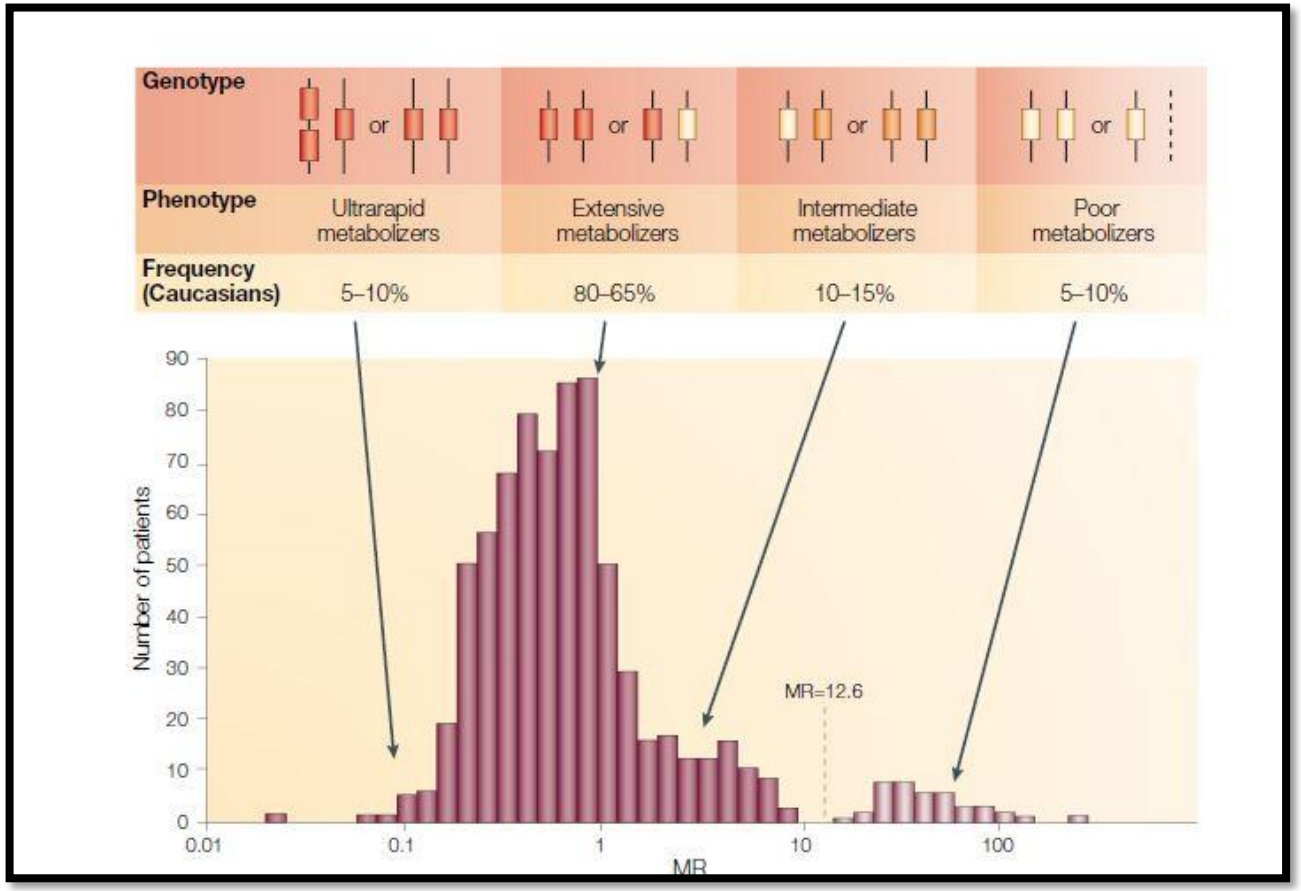
Duplication
+
Hybrid allele

How is the phenotype determined?



Genetic variation in pharmacogenes and its effect on protein function is translated into “Metabolizer phenotype”

Ultra-rapid	Extreme metabolic activity, which may result in poor efficacy and therapeutic failure of the drug
Extensive	Normal to high metabolic activity
Intermediate	Impaired or slow metabolic activity
Poor	Low to absent metabolic activity, which may result in a higher risk of toxicity



Resources for PGx Knowledge

- National and international drug agencies
  - » FDA, EMA, PMDA
- Professional consensus guidelines
  - » CPIC, DPWG, CPNDS
  - » Professional societies (AMP, ACMG)
- Peer-reviewed literature

Objective:

Create, curate, and post freely available, peer-reviewed, evidence-based, updatable, and detailed **gene/drug clinical practice guidelines**

CPIC®
Clinical Pharmacogenetics Implementation Consortium

CYP2C9, HLA-B and Phenytoin	fosphenytoin phenytoin	CYP2C9 HLA-B
CYP2C9, VKORC1, CYP4F2 and Warfarin	warfarin	CYP2C9 CYP4F2 VKORC1
CYP2D6 and Atomoxetine	atomoxetine	CYP2D6
CYP2D6 and Ondansetron and Tropisetron	ondansetron tropisetron	CYP2D6
CYP2D6 and Tamoxifen	tamoxifen	CYP2D6
CYP2D6, CYP2C19 and Selective Serotonin Reuptake Inhibitors	citalopram escitalopram fluvoxamine paroxetine sertraline	CYP2C19 CYP2D6

CPIC®
Clinical Pharmacogenetics Implementation Consortium

Table 1 Assignment of likely CYP2D6 phenotypes based on genotypes

Phenotype[a]		Genotype	Examples of *CYP2D6* diplotypes[b]
Metabolizer	Activity score		
CYP2D6 ultrarapid metabolizer	> 2.0	An individual carrying duplications of functional alleles	**1/*1xN, *1/*2xN, *2/*2xN*[c]
CYP2D6 normal metabolizer	1.5 and 2.0	An individual carrying two normal function alleles or one normal function and one decreased function allele	**1/*1, *1/*2, *1/*9, *1/*41, *2/*2,*
CYP2D6 normal metabolizer or intermediate metabolizer (controversy remains)[d]	1.0	An individual carrying two decreased function alleles or one normal function and one no function allele. *An activity score (AS) of 1.0 is associated with decreased tamoxifen metabolism to endoxifen compared to those with an AS of 1.5 or 2.*	**1/*4, *1/*5, *41/*41*
CYP2D6 intermediate metabolizer	0.5	An individual carrying one decreased function and one no function allele	**4/*10,*4/*41, *5/*9*
CYP2D6 poor metabolizer	0	An individual carrying only no functional alleles	**3/*4,*4/*4, *5/*5, *5/*6*

CPIC®
Clinical Pharmacogenetics Implementation Consortium

Table 2 Dosing recommendations for tamoxifen based on *CYP2D6* phenotype

Phenotype		Implications	Therapeutic recommendation[b]	Classification of recommendation[a]
Metabolizer status	Activity score			
CYP2D6 ultrarapid metabolizer	>2.0	Therapeutic endoxifen concentrations	Avoid moderate and strong CYP2D6 inhibitors. Initiate therapy with recommended standard of care dosing (tamoxifen 20 mg/day).	Strong
CYP2D6 intermediate metabolizer	0.5	Lower endoxifen concentrations compared to normal metabolizers; higher risk of breast cancer recurrence, event-free and recurrence-free survival compared to normal metabolizers.	Consider hormonal therapy such as an aromatase inhibitor for postmenopausal women or aromatase inhibitor along with ovarian function suppression in premenopausal women, given that these approaches are superior to tamoxifen regardless of *CYP2D6* genotype.[43] If aromatase inhibitor use is contraindicated, consideration should be given to use a higher but FDA approved tamoxifen dose (40 mg/day).[45] Avoid CYP2D6 strong to weak inhibitors.	Moderate

Methods in pharmacogenetics testing

- DNA microarray (allele-specific hybridization)
- Invader assay (cleavage-based with endonuclease enzyme)
- Mass spectrometry
- PCR-RFLP (restriction endonuclease)
- Pyrosequencing (primer extension)
- Sanger sequencing (chain termination)
- SNaPShot (primer extension)
- TaqMan (allele-specific hybridization)
- Next Generation Sequencing

Cost

Throughput

Required prior knowledge

SNV Panels

- SNV panel testing is the most used technology in PGx practice
- Commercial-panels or custom panels
- Typically contain a preselected set of SNVs: few variants in specific genes OR genome-wide
- Typically contain variants linked to drug response in PGx guidelines or on PharmGKB
- The evidence underlying the selected variants can vary from only the most strongly associated variants, to containing all variants potentially or theoretically associated with drug response
- Quick result at low costs
- Most have no CNV detection, no phasing, and no hybrid detection

Commerical vs Custom panels

Why the need?

- Many commercial arrays contain a high number of variants, making a fast turnaround time and interpretation challenging.
- Commercial arrays would include variants which may not be of direct interest in a clinical setting due to lack of evidence of clinical utility

Solution

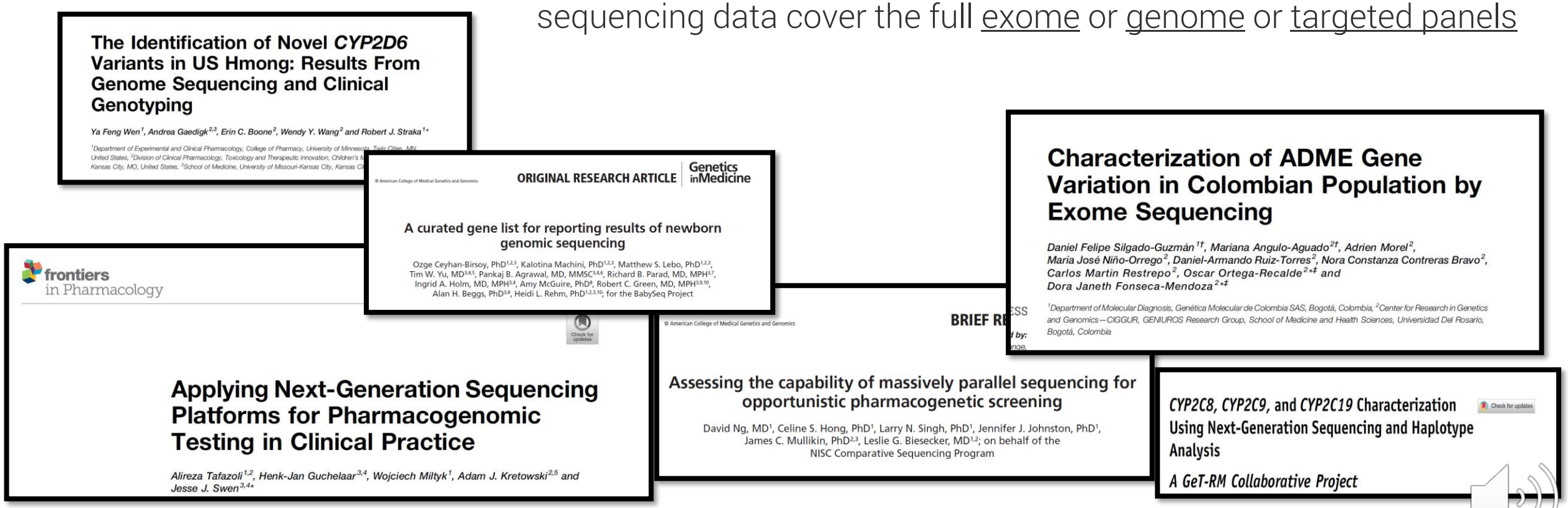
- Many institutes choose to customize a panel with genes/variants of interest.

Genome wide panels

- Offers genome wide coverage + PGx coverage
- Hundreds of thousands of markers
- Can miss specific alleles for PGx genes and CN which is not ideal

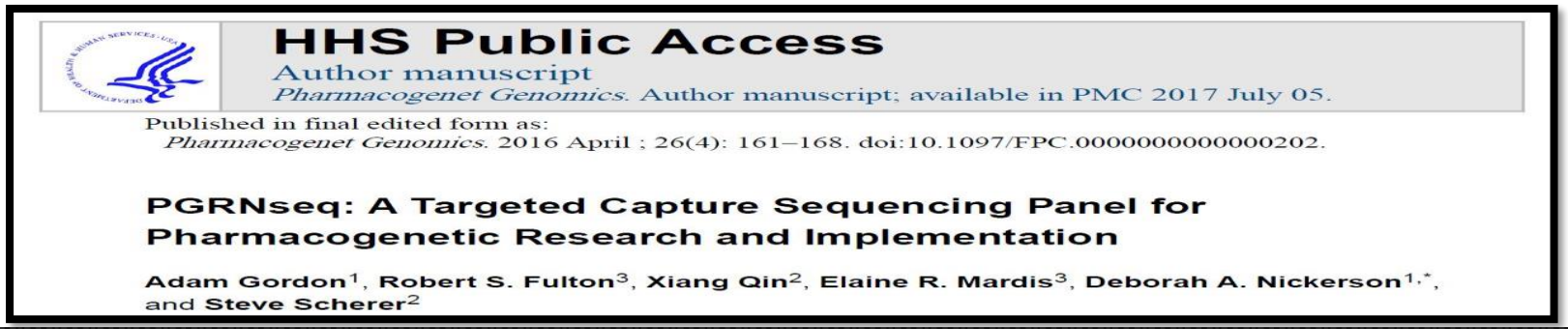
Next Generation Sequencing

- NGS technologies are not yet routinely applied in clinical PGx. e.g. in a recent market analyses out of 25 labs, only two used NGS
- Many research studies conducted using NGS for PGx
- While SNV panels only cover a limited set of selected variants, sequencing data cover the full exome or genome or targeted panels



NGS Panels

- Custom-capture panels of genes with associations to pharmacogenetic phenotypes.
- Generate deep coverage data
- >99% concordant with orthogonal datasets
- Identify novel, rare variants of interest. Value in research and clinical settings.
- Limitations:
  - ❖ Miss non-coding and complex structural variants for specific pharmacogenes (including CYP2A6, CYP2D6, and HLA-B)
  - ❖ Require better computational resources for data interpretation



HHS Public Access

Author manuscript

Pharmacogenet Genomics. Author manuscript; available in PMC 2017 July 05.

Published in final edited form as:

Pharmacogenet Genomics. 2016 April ; 26(4): 161–168. doi:10.1097/FPC.0000000000000202.

PGRNseq: A Targeted Capture Sequencing Panel for Pharmacogenetic Research and Implementation

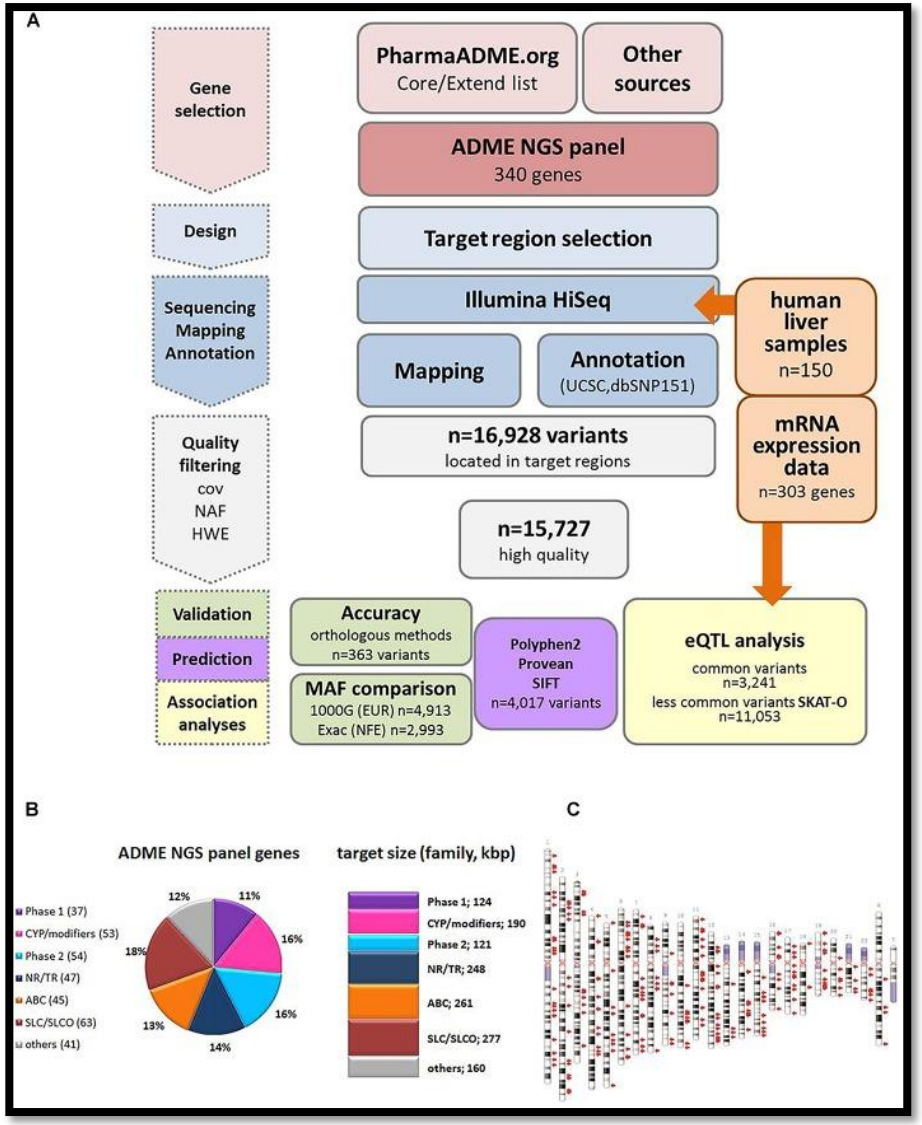
Adam Gordon[1], Robert S. Fulton[3], Xiang Qin[2], Elaine R. Mardis[3], Deborah A. Nickerson[1,*], and Steve Scherer[2]

NGS Panels

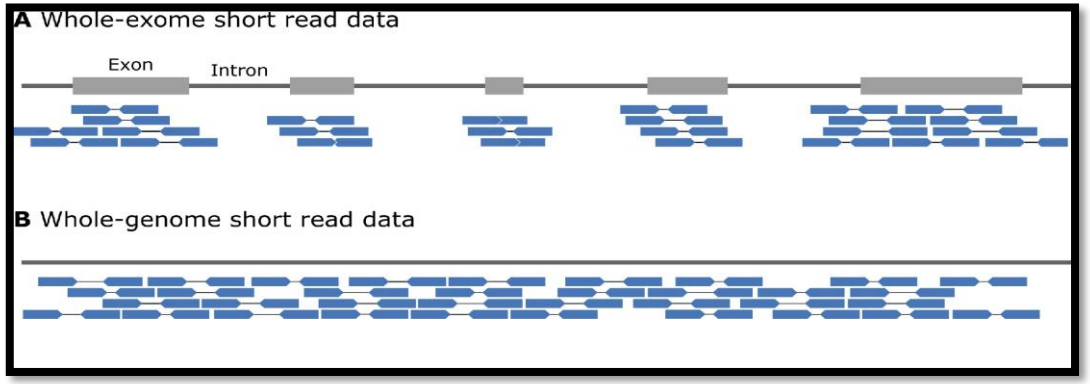
A New Panel-Based Next-Generation Sequencing Method for ADME Genes Reveals Novel Associations of Common and Rare Variants With Expression in a Human Liver Cohort

Kathrin Klein[1,2], Roman Tremmel[1,2], Stefan Winter[1,2], Sarah Fehr[3,4], Florian Battke[3,4], Tim Scheurenbrand[3,4], Elke Schaeffeler[1,2], Saskia Biskup[3,4], Matthias Schwab[1,2,5,6] and Ulrich M. Zanger[1,2]

- Genes coding for phase I and II enzymes, drug transporters and regulator/modifier genes
- Coding regions, adjacent introns, and 5' and 3' UTRs in flanking sequences
- >99% concordance.
- Very high read-depth
- Combined in-silico prediction with expression data, identified eQTLs .



WES and WGS (Short reads)



- Initial efforts: re-purpose already existent exomes or genomes to detect PGx variants.

- 94% concordance between PGx panel and WES
- 96% concordance between PGx panel and WGS.

- Some very important alleles could be missed be WES or WGS. e.g.

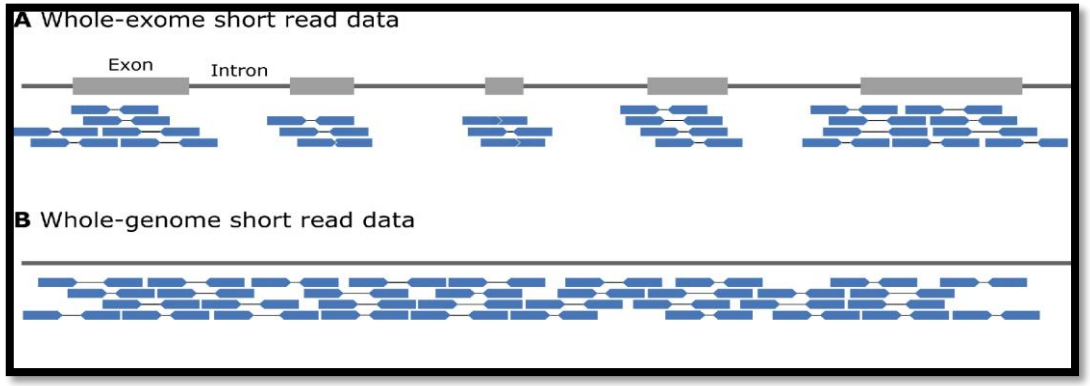
  Non-coding: CYP2C19*17 variants ; VKORC1, CYP2D6*4 and *41
  CYP2D6 copy number variation
  CYP2D6/2D7 hybrids
  HLA-genes

- Short reads pose a limitation: the identification of structural variants, repetitive regions, phasing of alleles and distinguishing highly homologous regions.

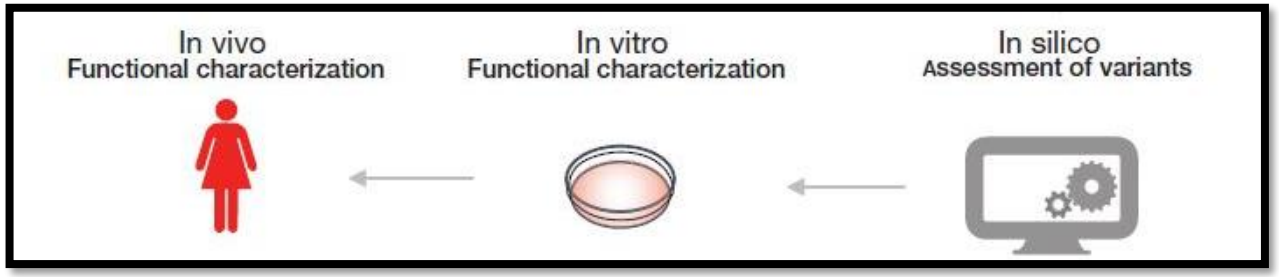
Cross-Comparison of Exome Analysis, Next-Generation Sequencing of Amplicons, and the iPLEX® ADME PGx Panel for Pharmacogenomic Profiling

Eng Wee Chua[1,2†], Simone L. Cree[1†], Kim N. T. Ton[1], Klaus Lehnert[3], Phillip Shepherd[4], Nuala Helsby[5] and Martin A. Kennedy[1]*

WES and WGS (Short reads)



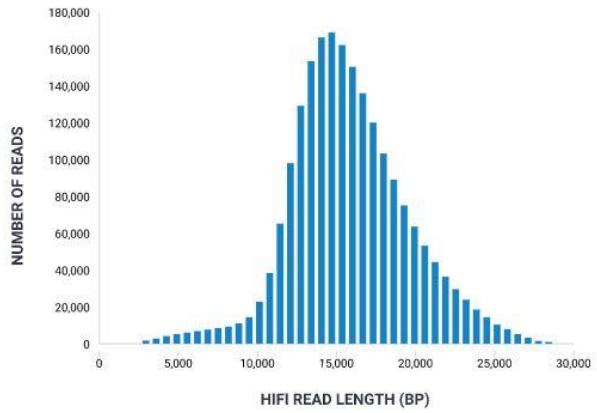
- May facilitate the discovery of novel loci (but will need a confirmative study or extensive invitro research to attribute potential, newly identified variants in a particular gene to drug response).
- WGS: Structural variants, non-coding, copy numbers..etc

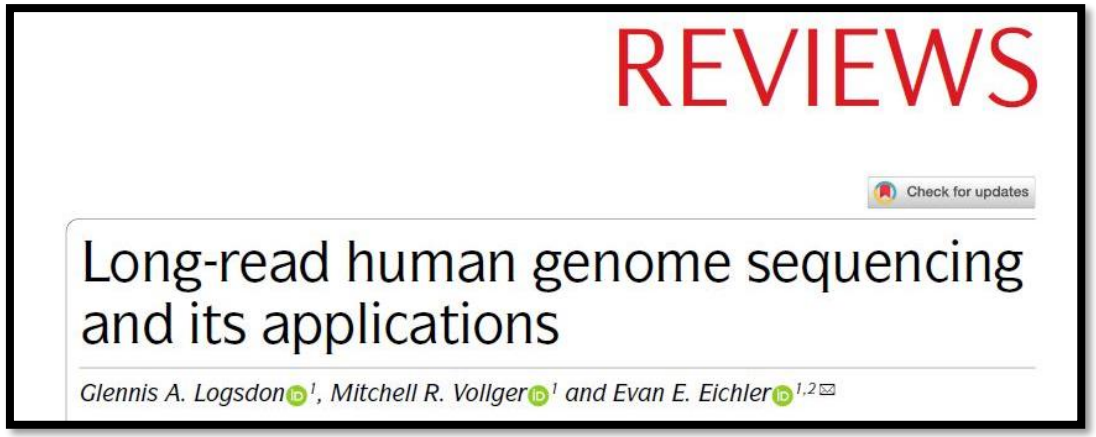


NGS (Long reads)

- Long-reads (>10 kilobase on average, sometimes tens to thousands of kb in length)
- Sequencing process occurs in real-time.
- Sequencing and library preparation without PCR amplification (no PCR bias)

- Two major technologies:
  - ✓ **Pacific Bioscience (PacBio)technology**: Uses SMRT (single molecule real-time)-sequencing
  - ✓ **Oxford Nanopore Technologies (ONT)**: Nanopores through which the DNA strand is pulled, the disruption in the current is specific to a codon, allowing for the full assembly of the DNA sequence



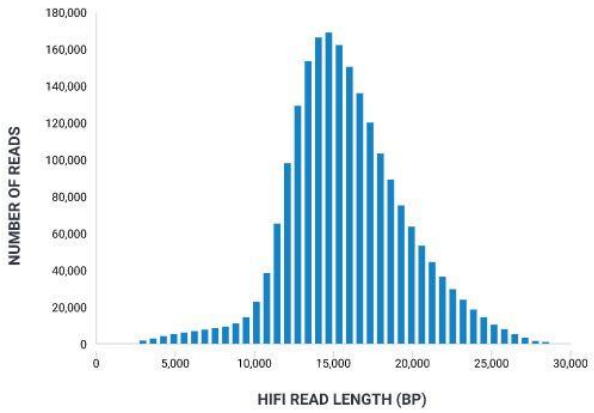

REVIEWS

Check for updates

Long-read human genome sequencing and its applications

Glennis A. Logsdon[1], Mitchell R. Vollger[1] and Evan E. Eichler[1,2]✉

NGS (Long reads)

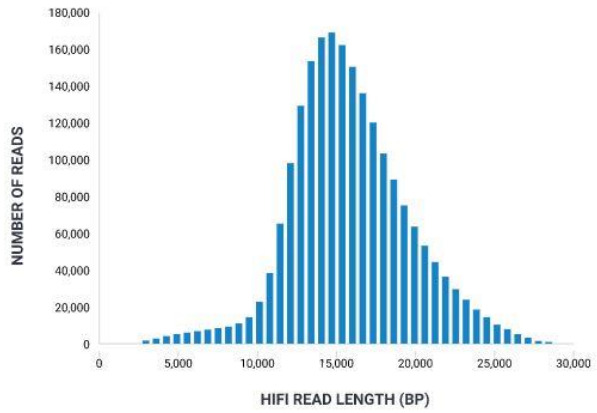


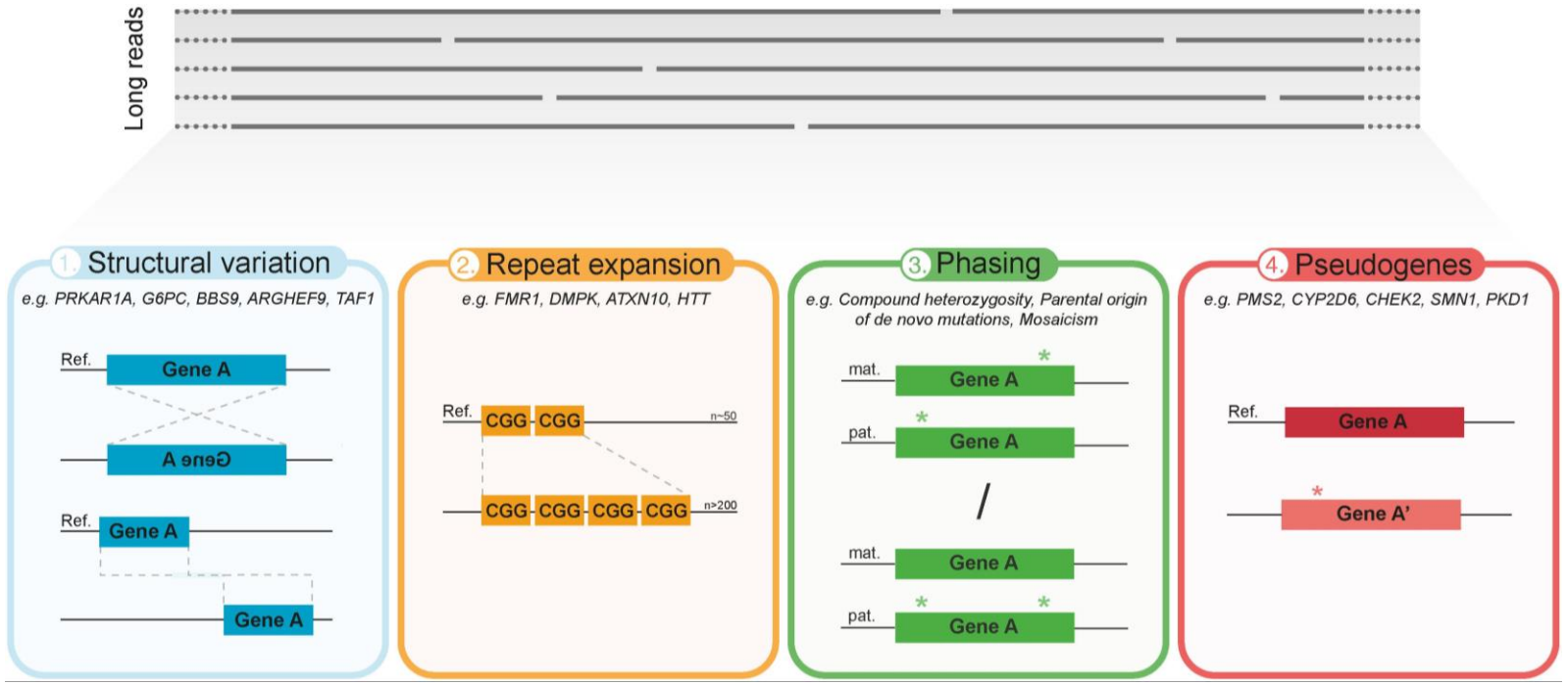
- ✓ **Pacific Bioscience (PacBio)technology**: Uses SMRT (single molecule real-time)-sequencing
- ✓ **Oxford Nanopore Technologies (ONT)**: Nanopores through which the DNA strand is pulled, the disruption in the current is specific to a codon, allowing for the full assembly of the DNA sequence

Sequencing technology	Platform	Data type	Read length (kb)		Read accuracy (%)	Throughput per flow cell (Gb)		Estimated cost per Gb (US$)	Maximum throughput per year (Gb)[a]
			N50	Maximum		Mean	Maximum		
Pacific Biosciences (PacBio)	RS II[b]	CLR	5–15	>60	87–92	0.75–1.5	2	333–933[c]	4,380
	Sequel	CLR	25–50	>100		5–10	20	98–195[d]	17,520
	Sequel II	CLR	30–60	>200		50–100	160	13–26[e]	93,440
		HiFi	10–20	>20	>99	15–30	35	43–86[e]	10,220
Oxford Nanopore Technologies (ONT)	MinION/ GridION	Long	10–60	>1,000	87–98	2-20	30	50–500[f]	21,900 (MinION) 109,500 (GridION)
		Ultra-long	100–200	>1,500		0.5–2	2.5	500–2,000[f]	913 (MinION) 4,563 (GridION)
	PromethION	Long	10–60	>1,000		50–100	180	21–42[f]	3,153,600
Illumina	NextSeq 550	Single-end	0.075–0.15	0.15	>99.9	16–30	>30	50–63[g]	>47,782
		Paired-end	0.075–0.15 (×2)	0.15 (×2)		32–120	>120	40–60[g]	>70,080
	NovaSeq 6000	Single-end	0.05–0.25	0.25		65–3,000	>3,000	10–35[h]	>1,194,545
		Paired-end	0.05–0.25 (×2)	0.25 (×2)					

NGS (Long reads)

- Expensive
- Data processing is significantly more intensive
- Throughput and accuracy lower compared to short-reads (Gb/year)

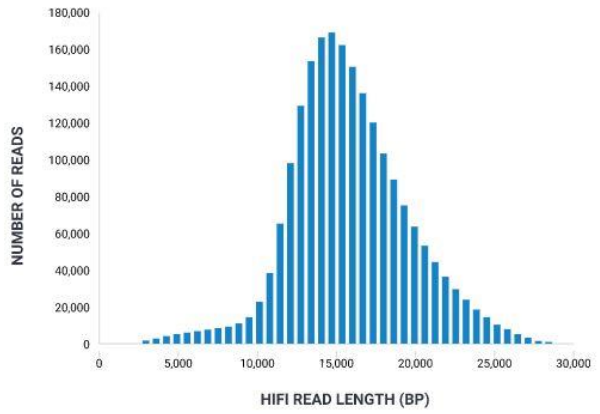


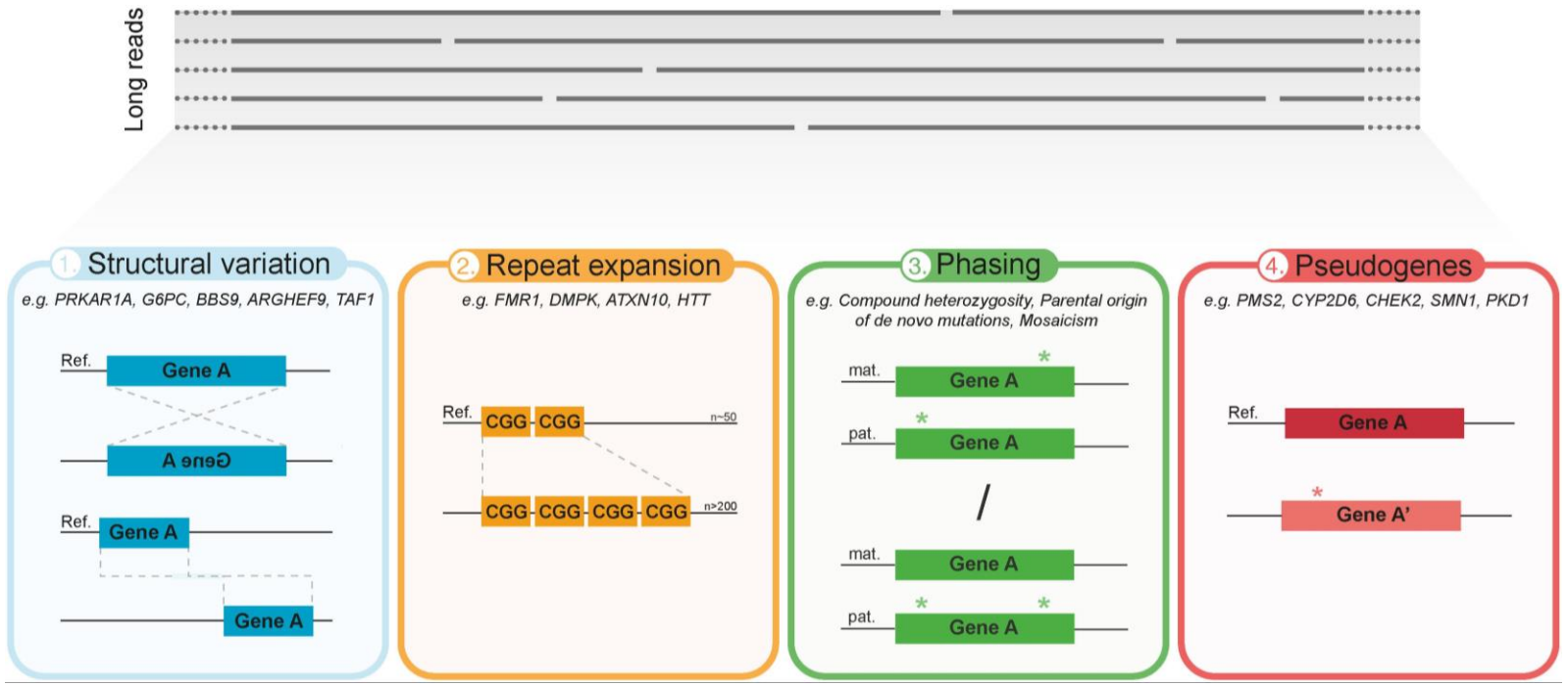


Promise:

NGS (Long reads)

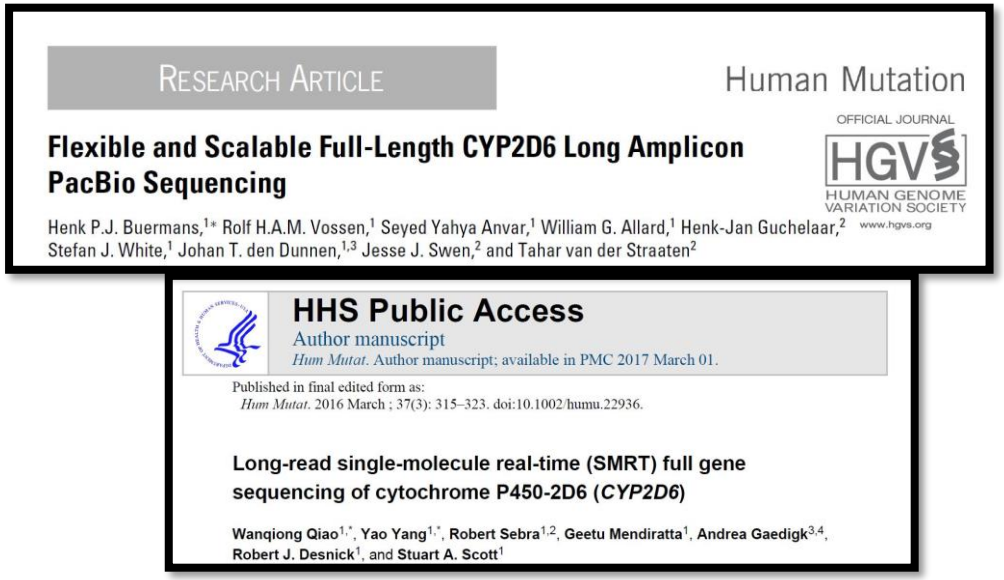
- Resolve some of the most challenging regions of the human genome
- Detect previously inaccessible structural variants
- Telomere-to- telomere assemblies of whole chromosomes





NGS (Long reads) In PGx

- No clinical adoption yet
- Few single gene studies
- Advantages: All apply to PGx implementation


RESEARCH ARTICLE

Human Mutation

OFFICIAL JOURNAL HGVS HUMAN GENOME VARIATION SOCIETY www.hgvs.org

Flexible and Scalable Full-Length CYP2D6 Long Amplicon PacBio Sequencing

Henk P.J. Buermans,[1*] Rolf H.A.M. Vossen,[1] Seyed Yahya Anvar,[1] William G. Allard,[1] Henk-Jan Guchelaar,[2] Stefan J. White,[1] Johan T. den Dunnen,[1,3] Jesse J. Swen,[2] and Tahar van der Straaten[2]

HHS Public Access

Author manuscript

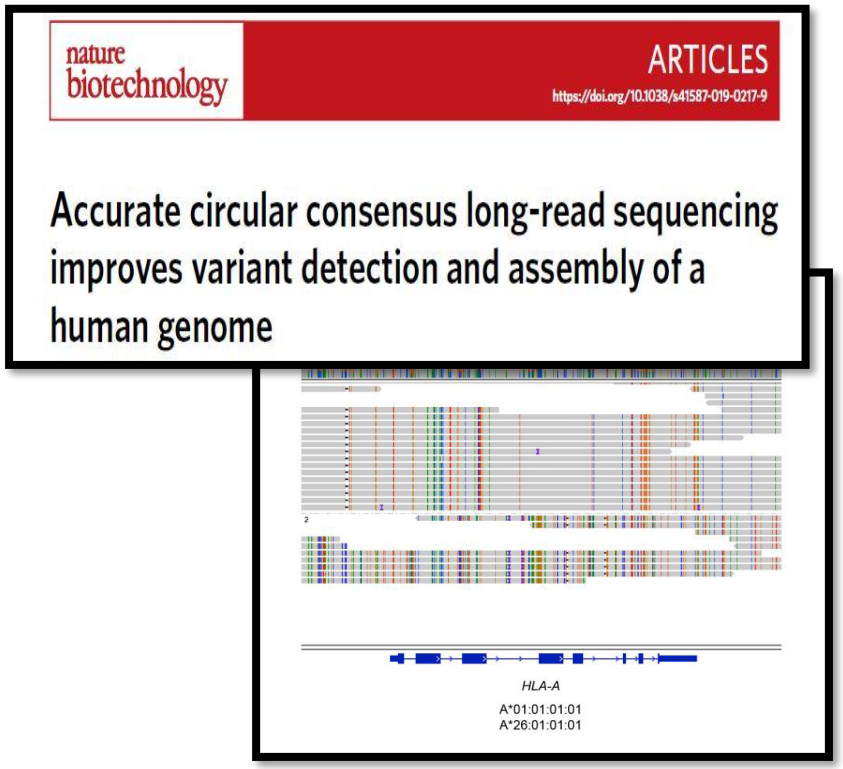
Hum Mutat. Author manuscript; available in PMC 2017 March 01.

Published in final edited form as:

Hum Mutat. 2016 March ; 37(3): 315–323. doi:10.1002/humu.22936.

Long-read single-molecule real-time (SMRT) full gene sequencing of cytochrome P450-2D6 (*CYP2D6*)

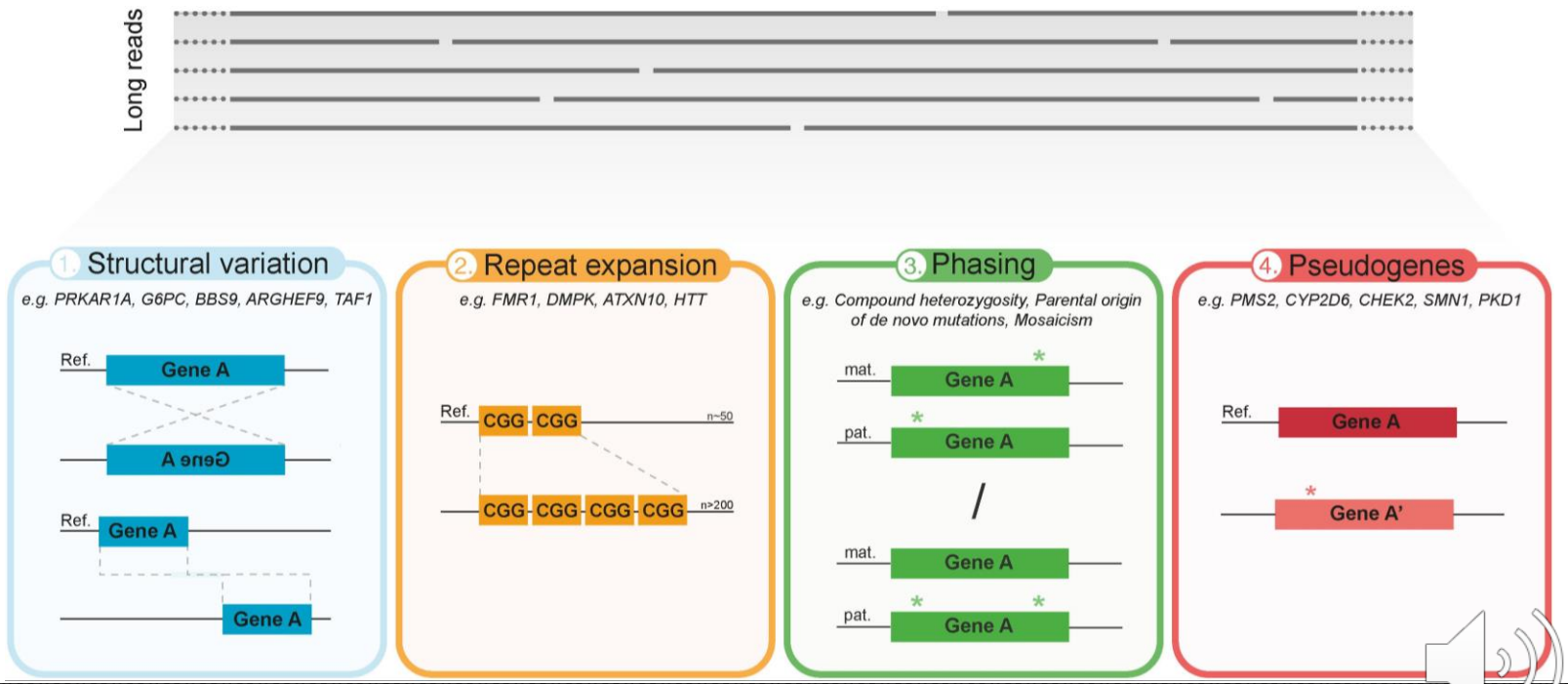
Wanqiong Qiao[1,*], Yao Yang[1,*], Robert Sebra[1,2], Geetu Mendiratta[1], Andrea Gaedigk[3,4], Robert J. Desnick[1], and Stuart A. Scott[1]


nature biotechnology

ARTICLES

https://doi.org/10.1038/s41587-019-0217-9

Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome



PGx Application Challenges

- Metabolizer Phenotype Inference
- Haplotype Phasing
- Structural Variants
- Variants of Unknown Effect
- Difficult genes
- Limitations in clinical implementation

PGx Application Challenges

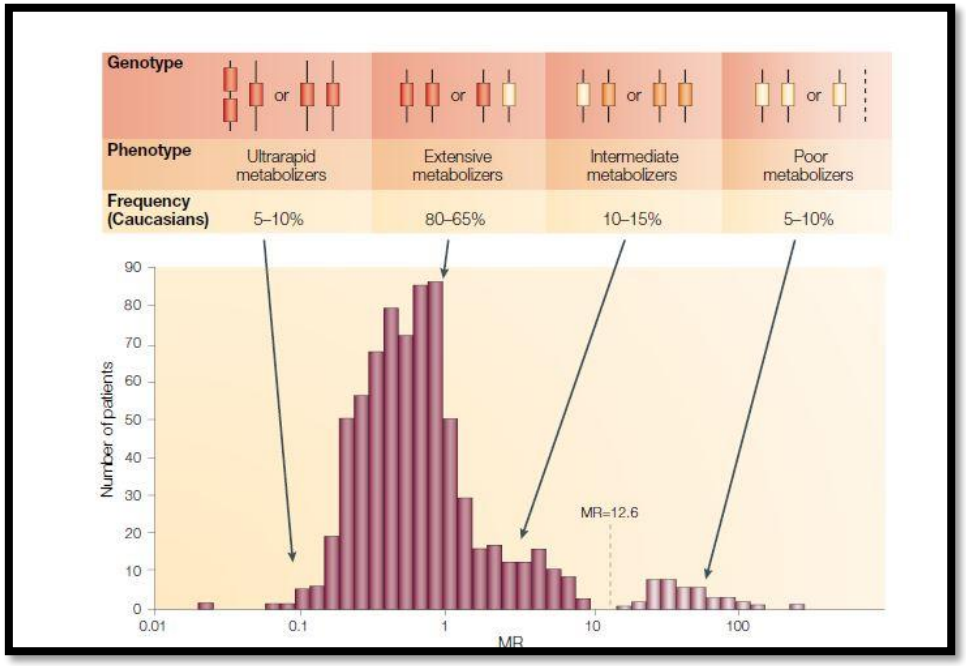
Metabolizer Phenotype Inference

Drug Metabolizer Phenotype Inference

- Genotypes are obtained
- For some genes (CYP genes) those genotypes make up haplotypes referred to as star alleles (*)
- The combination of the two * alleles makes a diplotype and translated into drug metabolizer phenotype.

PharmVar
Pharmacogene Variation Consortium

The Pharmacogene Variation (PharmVar) Consortium is a central repository for pharmacogene (PGx) variation that focuses on haplotype structure and allelic variation.



PharmVar
Pharmacogene Variation Consortium

Drug Metabolizer Phenotype Inference

CYP2C19*1	PV00598	80161A>G (I331V)
CYP2C19*2	PV00599	12662A>G (splice defect), 19154G>A (splice defect), 80161A>G (I331V)

CYP2D6*2 has 30 sub-alleles!!!

CYP2D6*2		PV00427	2851C>T (R296C), 4181G>C (S486T)
CYP2D6*2.001	CYP2D6*2A	PV00129	-1584C>G, -1235A>G, -740C>T, -678G>A, 214G>C, 221C>A, 223C>G, 227T>C, 232G>C, 233A>C, 245A>G, 310G>T, 745C>G, 842T>G, 1662G>C, 2851C>T (R296C), 3385A>C, 3585G>A, 3791C>T, 4181G>C (S486T), 4482G>A
CYP2D6*2.002	CYP2D6*2B	PV00150	1038C>T, 1662G>C, 2851C>T (R296C), 4181G>C (S486T)
CYP2D6*2.003	CYP2D6*2C	PV00149	1662G>C, 2471T>C, 2851C>T (R296C), 4181G>C (S486T)
CYP2D6*2.004	CYP2D6*2D	PV00152	2851C>T (R296C), 4181G>C (S486T)
CYP2D6*2.005	CYP2D6*2E	PV00836	-1584C>G, -1235A>G, -984G>A, -740C>T, -678G>A, 214G>C, 221C>A, 223C>G, 227T>C, 232G>C, 233A>C, 245A>G, 310G>T, 745C>G, 842T>G, 996C>G, 1662G>C, 2851C>T (R296C), 3385A>C, 3585G>A, 3791C>T, 4181G>C (S486T), 4482G>A

PharmVar
Pharmacogene Variation Consortium

Drug Metabolizer Ph… …ence

CYP3A4*14

CYP3A4*16

CYP2D6*2 has 30 sub…

CYP2D6 has over a 100 * alleles and over 2000 variants that define that

CYP2D6*2		PV00427	
CYP2D6*2.001	CYP2D6*2A	PV0012…	…842T>G, 1662G>C, 2851C>T (R296C),
CYP2D6*2.002	CYP2D…	PV0015…	
CYP2D6*2.003	CYP2D6*2C		
CYP2D6*2.004	CYP2D6*2D	PV00152	
CYP2D6*2.005	CYP2D6*2E	PV…836	-1584… …>A, 214G>C… …C>A, 223C>G, 227T>C, 232G>C, 233A>C, 245A>G, 310G>T, 745C>G, 842T>G, 996C>G, 1662G>C, 28… (R296C), 3… 3585G>A… …C>T, 4181G>C …8486T), 4482G>A

PharmVar
Pharmacogene Variation Consortium

Drug Metabolizer Phenotype Inference

- Several tools have been developed to assign *-allele and haplotypes based on sequencing data
- Performance is quite variable, in a study Aldy showed the least errors, compared to Stargazer and Astrolabe (2 compared to 9 and 10 respectively out of 21 alleles tested)
- All tools use PharmVar as their database. PharmVar is updated continuously leading to potential differences in assignments if not every tool is updated at the same time.
- Most tools require training datasets, and variation in those training sets can result in different sensitivity.

Stargazer

PharmCAT

Aldy

Astrolabe

Cypripi

g-Nomic

PHARMIP

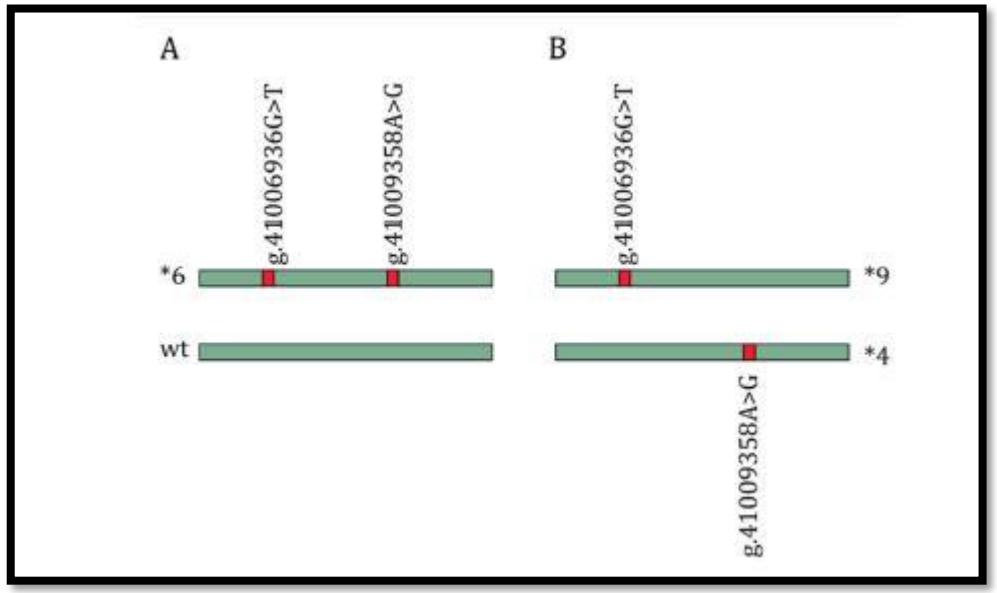
Cyrius

PGx Application Challenges

Haplotype Phasing

Haplotype phasing

- Determining if variants are located on the same allele or if they are on different alleles, leading to differences in phenotype assignment
- Given the polymorphic nature of many pharmacogenes, the likelihood of identifying multiple heterozygous variants within the gene locus of interest is highly likely
- While CPIC and the DPWG report which diplotype translate into which phenotypes; but no guidance on phasing
- Having two variants on one allele is different than having them on opposing alleles.



CYP2B6

Haplotype phasing

- Resolving phasing by short-read NGS:
  - » Linked-read sequencing: use of barcoded short fragments sequenced with conventional short-read methods.
  - » Using barcodes, every read can be linked back to the original position and artificial long input DNA can be reconstructed

- Resolving phasing by Long-read NGS:
  - » The length of the reads can be utilized for haplotype phasing

PGx Application Challenges

Structural Variants

Structural Variants

- Majority of pharmacogenes Are largely characterized by complex regions

- CNVs, structural rearrangements and repetitive regions

Protein	Gene	Related Drugs		Locus Size (bp)	Rare Variants, n (% of Known Variants)	Part of Locus Defined as Complex, %(bp)
		CPIC	DPWG			
CACNA1S	*CACNA1S*	7	-	73,055	2520 (98%)	33.3
CFTR	*CFTR*	1	-	250,187	1684 (99%)	42.2
CYP2B6	*CYP2B6*	1	1	27,149	761 (98%)	100.0
CYP2C9	*CYP2C9*	10	2	50,734	632 (98%)	72.0
CYP2C19	*CYP2C19*	15	10	90,525	712 (99%)	83.6
CYP2D6	*CYP2D6*	14	21	4408	992 (97%)	100.0
CYP3A5	*CYP3A5*	1	1	31,833	643 (98%)	49.4
CYP4F2	*CYP4F2*	1	-	20,098	766 (97%)	51.4
DPD	*DPYD*	2	4	917,258	1211 (98%)	40.0
FACT. V LEIDEN	*FACT. V LEIDEN*	-	1	72,423	1679 (97%)	41.9
G6PD	*G6PD*	1	-	16,183	465 (98%)	36.4
HLA-A	*HLA-A*	2	1	4625	423 (71%)	100.0
HLA-B	*HLA-B*	6	7	87,698	308 (78%)	62.1
IFNL3	*IFNL3*	2	-	1577	317 (95%)	100.0
IFNL4	*IFNL4*	2	-	3543	404 (97%)	100.0
NUDT15	*NUDT15*	3	3	9656	244 (99%)	64.7
RYR-1	*RYR1*	7	-	153,866	6584 (98%)	51.4
SLCO1B1	*SLCO1B1*	1	2	108,045	951 (96%)	69.6
TPMT	*TPMT*	3	3	26,764	346 (97%)	52.3
UGT1A1	*UGT1A1*	1	1	13,052	470 (99%)	40.3
VKORC1	*VKORC1*	1	3	5139	370 (98%)	41.8

Structural Variants by NGS

- Multiple tools designed to extract CNVs:
  XHMM , CoNIFER, Varseq, CNVnator
- Agreement between methods is low
- There is bias towards smaller CNVs vs. large CNVs
- Distinction between a pharmacogene and a pseudogene can be challenging (e.g. CYP2D6 and CYP2D7 share >98% of their sequence)
- Long-read sequencing can distinguish gene from pseudogene, can better assess large insertions and deletions, and structural variants
- Full characterization of the complexity of pharmacogenes is still in the research phase

Protein	Gene	Related Drugs		Locus Size (bp)	Rare Variants, n (% of Known Variants)	Part of Locus Defined as Complex, %(bp)
		CPIC	DPWG			
CACNA1S	*CACNA1S*	7	-	73,055	2520 (98%)	33.3
CFTR	*CFTR*	1	-	250,187	1684 (99%)	42.2
CYP2B6	*CYP2B6*	1	1	27,149	761 (98%)	100.0
CYP2C9	*CYP2C9*	10	2	50,734	632 (98%)	72.0
CYP2C19	*CYP2C19*	15	10	90,525	712 (99%)	83.6
CYP2D6	*CYP2D6*	14	21	4408	992 (97%)	100.0
CYP3A5	*CYP3A5*	1	1	31,833	643 (98%)	49.4
CYP4F2	*CYP4F2*	1	-	20,098	766 (97%)	51.4
DPD	*DPYD*	2	4	917,258	1211 (98%)	40.0
FACT. V LEIDEN	*FACT. V LEIDEN*	-	1	72,423	1679 (97%)	41.9
G6PD	*G6PD*	1	-	16,183	465 (98%)	36.4
HLA-A	*HLA-A*	2	1	4625	423 (71%)	100.0
HLA-B	*HLA-B*	6	7	87,698	308 (78%)	62.1
IFNL3	*IFNL3*	2	-	1577	317 (95%)	100.0
IFNL4	*IFNL4*	2	-	3543	404 (97%)	100.0
NUDT15	*NUDT15*	3	3	9656	244 (99%)	64.7
RYR-1	*RYR1*	7	-	153,866	6584 (98%)	51.4
SLCO1B1	*SLCO1B1*	1	2	108,045	951 (96%)	69.6
TPMT	*TPMT*	3	3	26,764	346 (97%)	52.3
UGT1A1	*UGT1A1*	1	1	13,052	470 (99%)	40.3
VKORC1	*VKORC1*	1	3	5139	370 (98%)	41.8

PGx Application Challenges

Variants of Unknown Effect

Variants of unknown effect

- The clinical utility is limited and not much is known about their function
- Solutions could be:
  - Cell-line models
  - In silico predictions (sequence conservation, the physiochemical and crystal structure of the protein, or on evolutionary scores)
  - Studying patients displaying the most extreme phenotypes

PGx Application Challenges

Difficult genes

Difficult genes

Gene	Challenges
CYP2D6	Structural variants and gene re-arrangements
	Pseudogenes
	CNVs
	Highly polymorphic
UGTA1A	Rare population specific variants
	Non-coding variants
VKORC1	Non-coding variants
HLA	Highly polymorphic regions
	Rare population specific variants

PGx Application Challenges

Clinical implementation

Challenges to clinical application

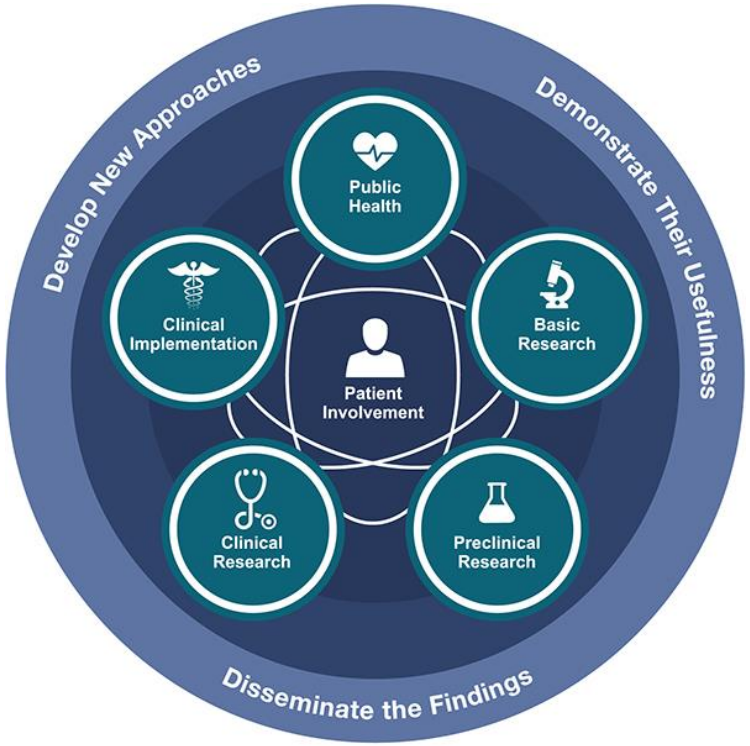
Translating basic science findings into clinical practice:

*Clinical Studies/Trials

*Clinical practice recommendations and guidelines

*Adoption of guidelines into evidence-based practice

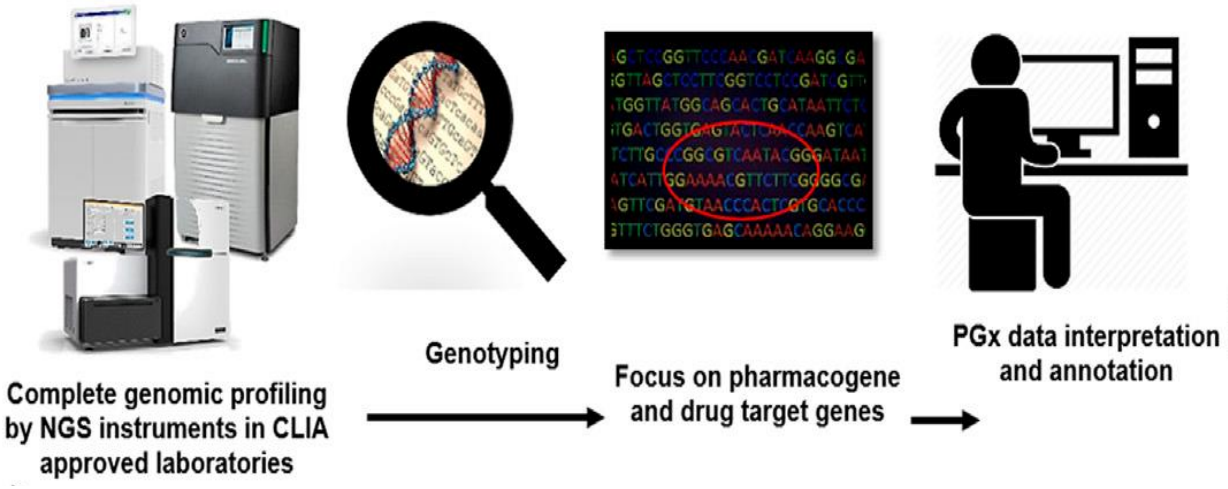
*Assessing efficacy, cost, outcomes.etc.



Credit: National Center for Advancing Translational Sciences

Challenges to clinical application

- Need results fast…point-of-care, pre-emptive testing
  - New molecular testing assays can be performed in ~1 hour
- Need for infrastructure:
  - High throughput sequencers and molecular Sequencing technologies
  - Specific software and computational tools



Challenges to clinical application

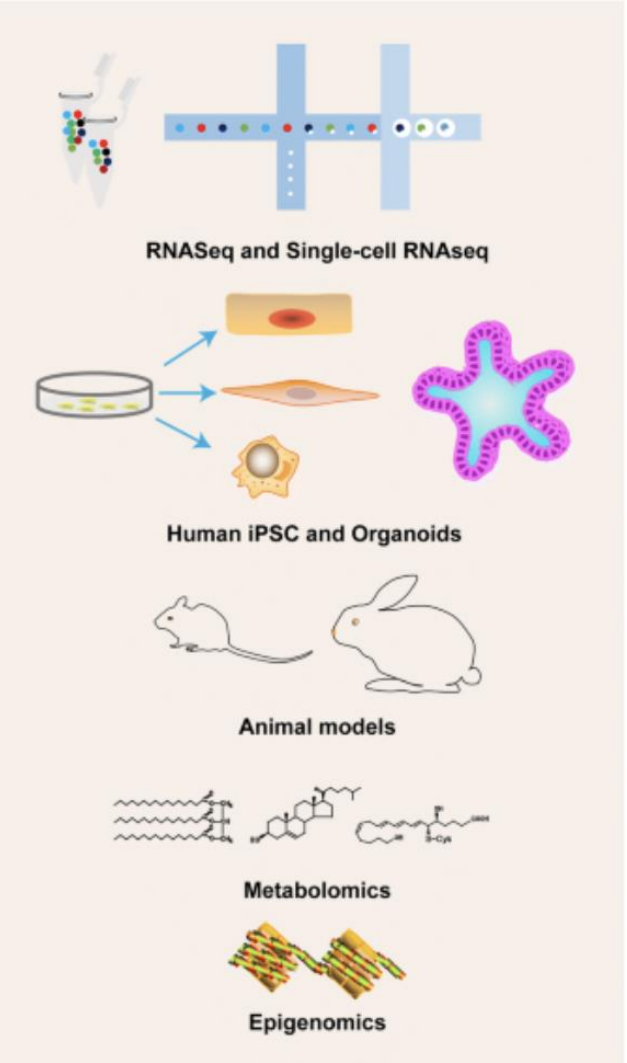
Functional characterization of variants:

*In silico tools

*In Vitro studies

*In Vivo models

*Non-coding elements and regulatory regions



Challenges to clinical application

- Physicians training

  Translating genotypes into phenotypes and using guidelines to guide managements is not common knowledge

- Laboratory consultations

  Laboratories offering tests should offer consultations

- PGx teams/ clinics

  Trained to use clinical support systems

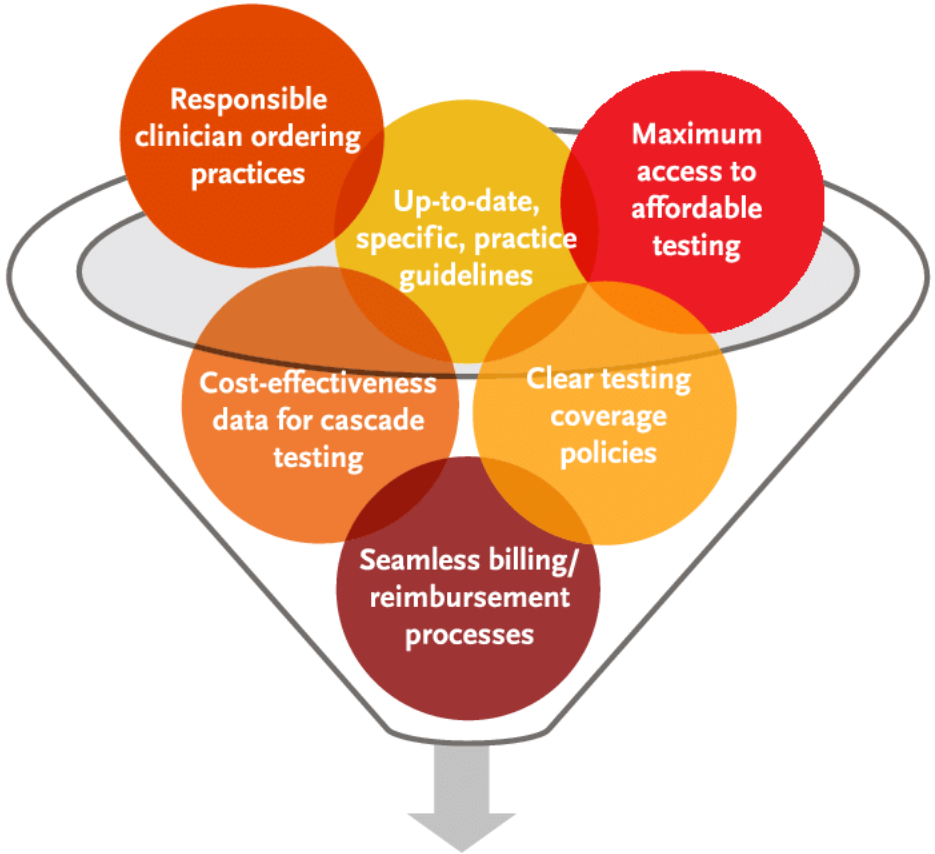
PGx trained clinicians

Challenges to clinical application

Research funding and Reimbursement

*While still low, reimbursement for PGx testing almost doubled in the last few years.

*Allies within legislative bodies needed to ensure adequate funding for both research as well as for efforts to build precision medicine and PGx programs



PHARMACOGENOMIC RESOURCES

1. *http://www.cypalleles.ki.se/*
   - Catalog of *CYP450* genetic variation and nomenclature.

2. *http://www.pharmGKB.org/*
   - A central PGx resource: collect, encode, and disseminate knowledge about the impact of human genetic variations on drug response.

3. *http://www.cpicpgx.org*
   - CPIC: continued guidelines on PGx-based clinical management.

4. *http://www.warfarindosing.org/*
   - Online tool to predict warfarin dose using both clinical and genetic variables.

5. *http://medicine.iupui.edu/clinpharm/ddis/*
   - CYP450 drug interaction database.

6. *http://www.fda.gov/Drugs/ScienceResearch/ResearchAreas/Pharmacogenetics/ucm083378.htm*
   - FDA pharmacogenetics biomarker table.

ARUP® LABORATORIES

A nonprofit enterprise of the University of Utah and its Department of Pathology