Published in final edited form as:

Pharmacol Ther. 2017 July ; 175: 75–90. doi:10.1016/j.pharmthera.2017.02.036.

Personalized medicine: Genetic risk prediction of drug response★

Ge Zhang^{a,*} and Daniel W. Nebert^{a,b,**}

^aDivision of Human Genetics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH 45229-3039, United States

^bDepartment of Environmental Health and Center for Environmental Genetics, University of Cincinnati School of Medicine, Cincinnati, OH 45267-0056, United States

Abstract

Pharmacogenomics (PGx), a substantial component of “personalized medicine”, seeks to understand each individual's genetic composition to optimize drug therapy reactions (ADRs). Drug responses are highly variable because innumerable factors contribute to ultimate phenotypic outcomes. Recent genome-wide PGx studies have provided some insight into genetic basis of variability in drug response. These can be grouped into three categories. [a] *Monogenic* (Mendelian) *traits* include early examples mostly of inherited disorders, and some severe (idiosyncratic) ADRs typically influenced by single rare coding variants. [b] *Predominantly oligogenic traits* represent variation largely influenced by a small number of major pharmacokinetic or pharmacodynamic genes. [c] *Complex PGx traits* resemble most multifactorial quantitative traits — influenced by numerous small-effect variants, together with epigenetic effects and environmental factors. Prediction of *monogenic drug responses* is relatively simple, involving detection of underlying mutations; due to rarity of these events and incomplete penetrance, however, prospective tests based on genotype will have high false-positive rates, plus pharmacoeconomics will require justification. Prediction of *predominantly oligogenic traits* is slowly improving. Although a substantial fraction of variation can be explained by limited numbers of large-effect genetic variants, uncertainty in successful predictions and overall cost-benefit ratios will make such tests elusive for everyday clinical use. Prediction of *complex PGx traits* is almost impossible in the foreseeable future. Genome-wide association studies of large cohorts will continue to discover relevant genetic variants; however, these small-effect variants, combined, explain only a small fraction of phenotypic variance — thus having limited predictive power and clinical utility.

★Associate editor: Russell Prough

*Correspondence to: G. Zhang, Cincinnati Children's Hospital Medical Center, MLC 4006, 3333 Burnet Avenue, Cincinnati, OH 45229-3039, United States. ge.zhang@cchmc.org (G. Zhang). **Correspondence to: D. W. Nebert, University of Cincinnati Medical Center, P.O. Box 670056, Cincinnati, OH 45267-0056, United States. dan.nebert@uc.edu (D.W. Nebert).

Conflict of interest statement

Both authors declare that they have no actual, or potential, conflicts of interest.

Keywords

Pharmacogenetics; Pharmacogenomics (PGx); Genetic risk prediction; Monogenic (Mendelian) PGx traits; Predominantly oligogenetic PGx traits; Complex PGx traits; Genetic architecture

---

1. Introduction

How successful will we be in predicting drug responses caused by inherited genetic variation? This is the primary question addressed herein. *Inter-individual variability in response* is defined as an “effect of varying intensity occurring in different individuals receiving a specified drug dose”, or a “requirement of a range of doses (concentrations) in order to produce an effect of specified intensity in all patients.” Patients are well known to vary widely in their responses to drugs (Brunton, Chabner, & Knollman, 2011).

A substantial subset of *personalized medicine* comprises *personalized medication*. Over the past six decades, this field of study began by being called *pharmacogenetics*; however, since the Human Genome Project started in 1990, the term *pharmacogenomics* has become more popular. Although the terms are often used interchangeably — for many of us in the field, there are subtle differences, *i.e.* effect of individual genes (pharmacogenetics) *vs* total genomic expression (pharmacogenomics), in response to a drug. Pharmacogenomics aims to develop rational methods to *optimize drug therapy* with respect to the patient's genotype, as well as to *ensure maximum efficacy with minimal adverse effects in each individual*. Any drug capable of producing a desired therapeutic effect — can also produce unwanted, or adverse, side effects (Edwards & Aronson, 2000).

All the above-mentioned drug responses (therapeutic effect, adverse effect and toxic effect) are regarded in the field of genetics as *phenotypes*, or *traits*, which herein we collectively call “*pharmacogenetic or pharmacogenomic traits*” (PGx traits). Most drug responses are expressed as *multifactorial traits* — similar in many ways to human complex diseases (*e.g.* type-2 diabetes, schizophrenia, cancer), as well as quantitative traits such as height, blood pressure or serum lipid levels. One major difference between drug-response and complex-disease is obvious: any individual not challenged with a particular drug will never know his or her phenotype for that drug.

The degree of success in predicting outcome of a drug before treating the patient will depend on the “genetic basis of the PGx trait” (*genotype*), *i.e.* number of genetic variants contributing to that *phenotype*, allele frequency and *effect-size* of each contributing genetic variant, and interactions between them and with other environmental factors (Park et al., 2011). This review examines the *genetic architecture* (genetic basis) of various drug responses (PGx traits) and discusses statistical feasibility of genetic prediction.

2. Brief history of genetics

2.1. Gregor Mendel

Principles of “dominant-*versus*-recessive” classical genetics began in the 1860s by Mendel, using garden peas as his experimental model system (*e.g.* red flower color *dominant*, white

recessive). Whereas the F_1 cross yields all red flowers, the F_2 generation yields three red (two are R/r heterozygotes, and one R/R homozygote) and one white flower color, r/r ; this F_2 population represents the “Mendelian pattern of inheritance”, or distribution (Fig. 1A).

Recent advances have greatly expanded Mendel's studies, *e.g.* showing in the garden pea (*Pisum sativum*) at least two genes involved in red flower color: the *A* gene codes for a bHLH transcription factor that regulates anthocyanin pigmentation. The white-flowered mutant allele most likely used by Mendel is a simple G-to-A transition in a splice-donor site — resulting in a mis-spliced mRNA and premature stop codon; the *A2* gene encodes a WD40 protein, which is part of an evolutionarily highly conserved regulatory complex (Hellens et al., 2010).

2.2. Garrod's “inborn-errors-of-metabolism”

In the first decade of the 1900s, Sir Archibald Garrod described “in-born-errors-of-metabolism”: albinism, alkaptonuria, cystinuria and pentosuria. Each of these distinct clinical *autosomal recessive* traits show a pattern of inheritance similar to that of white flower color of Mendel's garden pea. Garrod is credited with ushering in the era of human genetics; the predominant underlying tenet was “one gene, one disease”, or “one wild-type (healthy) allele, and one disease allele”. For each pregnancy, two healthy parents, “carriers” heterozygous for a disease allele, bring a 25% chance of producing a child having both disease alleles and, hence, inheriting the unwanted disorder.

2.3. Single-coding mutations with severe effect

Garrod's four disorders (*described above*) are among the first single-coding mutations causing severe clinical disorders; these usually skip a generation. During the next several decades, many additional examples of autosomal recessive clinical diseases were described (*e.g.* maple syrup urine disease, phenylketonuria, cystic fibrosis, Gaucher disease, congenital adrenal hyperplasia); similarly, these diseases are autosomal recessive.

Autosomal dominant traits began to be discovered (*e.g.* Huntington disease, achondroplastic dwarfism, Marfan syndrome, neurofibromatosis, hereditary spherocytosis) in which the trait typically appears in each generation because heterozygotes exhibit the disease. In addition, random germline mutations (*de novo* mutations) — not present in somatic cells of either parent — can also occur in gametes, thereby giving the child a disease that had never before occurred in that genetic pedigree (“family tree”).

Traits transmitted as *X-linked recessives* were also described (*e.g.* red-green colorblindness, hemophilia — due to mutations on the X chromosome). Female carriers have a 50% chance of passing the defective allele to offspring, thereby causing the disorder in males but not females.

Further, *X-linked dominant* traits were identified (*e.g.* incontinentia pigmenti, Coffin-Lowry syndrome — also caused by mutations on the X chromosome). Affected females have a 50% chance of passing the mutant allele to all offspring, whereas affected males have a 100% chance of transmitting the defective allele to daughters. A single copy of the defective allele is sufficient to cause the disorder.

All the above-mentioned phenotypes follow bimodal distributions of Mendelian inheritance (Fig. 1A). Large-effect alleles can also result in a trimodal distribution (Fig. 1B), in which additive traits from both parents result in an intermediate phenotype. An additional caveat to virtually all Mendelian diseases includes “modifier genes”, *i.e.* any number of additional genes that can affect age of onset and/or degree of severity of the disease (Fig. 1C).

2.4. Biometric view of inheritance and reconciliation with Mendelian inheritance

In addition to these simple Mendelian traits having distinct patterns of inheritance, many traits such as body height and cognitive ability exhibit continuous variation in the population as well as strong resemblance within each family; however, the inheritance pattern of these continuous traits could not readily be explained by Mendelian laws. In the late 19th century and early 20th century, a group of statisticians (biometricians) — including Francis Galton, Karl Pearson and Raphael Weldon — argued against Mendelian inheritance, insisting that most phenotypic variation was continuous, and that a “blending model” seemed more appropriate to explain the inheritance pattern of continuous traits, because an offspring's phenotype is roughly the average of that of the two parents.

The debate between supporters of Mendelian inheritance (“Mendelians”) and the biometric school of inheritance lasted for two decades, until the disagreement was finally resolved by R.A. Fisher in 1918 in his publication of “*The Correlation between Relatives on the Supposition of Mendelian Inheritance*” (Fisher, 1918). In this monumental paper, Fisher showed that the “continuous variation” described by biometricians could represent the cumulative result of many discrete genetic loci (Fig. 1D), and that the family resemblance of continuous traits could be explained by Mendelian inheritance. Hence, it became appreciated that most *human complex traits and diseases* (*e.g.* height, body mass index, serum lipid levels, blood pressure, psoriasis, dementia, numerous types of cancer) represent *multifactorial traits* — *i.e.* the result of contributions from dozens or hundreds, perhaps thousands, of genes (polygenic), combined with additional modifying effects such as contribution of epigenetic effects and other environmental factors.

2.5. Beginning of the genomics era

Subsequent to advances in molecular biology, recombinant-DNA cloning and DNA sequencing that began in the 1970s — the field of genomics quickly helped us appreciate that “monogenic diseases” represent many more than one “disease allele.” Among the earliest breakthroughs was phenylketonuria (PKU), described as an *autosomal recessive* disorder caused by phenylalanine hydroxylase deficiency. After cloning the *PAH* gene from one chromosome of a “carrier” parent of a PKU child (Woo, Lidsky, Guttler, Chandra, & Robson, 1983), the critical *PAH* mutation represented a GT > AT transition at the 5′ splice donor-site of intron 12 (Marvit et al., 1987). At first, this discovery was hailed as “*the* disease allele” for PKU; however, within months, a second mutation (this one changing an amino acid) was reported. Three years later (Cotton, 1990), 18 distinct mutations had been identified, and the concept of *allelic heterogeneity* was widely accepted as the norm for most genes in which single mutations cause serious diseases.

Currently, 955 mutations (in and near the *PAH* gene), causing variable symptoms of PKU, have now been reported, worldwide [<http://www.biopku.org/home/pah.asp>], with ethnic differences in allelic frequencies. Similar stories can be told about many other Mendelian diseases.

2.6. Single-nucleotide polymorphisms (SNPs)

Following launch of the Human Genome Project in 1990, the field of genomics exploded with new knowledge. Yeast, worm and fly geneticists had been using the term “nucleotide substitutions” since the early 1980s, but in the mid-1990s several human genetics laboratories coined the term “*single-nucleotide polymorphisms*” (SNPs); “SNP fever” then began (“SNiPping through the DNA,” *etc.*). In retrospect, a better name for SNPs might have been “single-nucleotide variants”, but SNVs (“SNiVs”) doesn't sound as “catchy” as SNPs (“SNiPs”).

In the mid-1990s, many publications began to appear in top-tier journals, demonstrating “statistically significant” associations ($P < 0.05$) between one or a few SNPs and a complex disease — such as hypertension, Alzheimer disease, asthma, or schizophrenia. Just as quickly, however, experiments to corroborate those data failed to do so, and it was soon realized that the genetics of complex diseases is not nearly as simple as that of Mendelian diseases.

Methods such as linkage studies that had been successful in finding major genes were found to have limited power for detecting genes of modest effect or lower penetrance; thus, a different method of association studies — one that tested all candidate genes — would have greater power, even if this meant testing every gene in the genome. The landmark 2-page paper in 1996 by Risch and Merikangas (1996) proved to be visionary in this respect.

2.7. Recent rapid advances in our understanding of human genetics and genomics

Our appreciation of the concepts of human genetics and genomics has rapidly exploded during the past 25 years. We therefore believe that a brief review of past and present thinking in these fields, with regard to genome-wide association studies (GWAS), missing heritability and rare vs common variants, is necessary in order to acquaint the reader with the very complicated properties of multifactorial traits. As will be described, clearly the vast majority of drug responses (efficacy, adverse effect and toxic effect) involves — not Mendelian or predominantly oligogenic — but rather polygenic, multifactorial phenotypes.

2.7.1. Genome-wide association studies (GWAS)—The earliest GWAS was probably published in 2002, in which an association between the *LTA* gene and myocardial infarction was determined by typing almost 93,000 gene-based SNP markers (Ozaki et al., 2002). Another early GWAS — typing > 116,000 SNPs — reported an association between the *CFH* gene and age-related macular degeneration (Klein et al., 2005). More recently, easy-to-use DNA chips containing 1 million to 5 million SNPs have become readily available. The GWAS field has expanded exponentially; today, >24,000 SNP-trait associations have been reported in >2500 studies [<https://www.ebi.ac.uk/gwas/>]. These robust GWAS — having *P*-values ranging from $<10^{-8}$ to $<10^{-400}$ — underscore the value of

using more stringent statistical significance levels when one is studying 1 million SNPs in large cohorts containing thousands, or even hundreds of thousands, of samples.

GWAS quickly became far more reliable for genotype-phenotype association tests — when compared with studies involving one or several SNPs in small cohorts of several dozen or even several hundred individuals. These latter publications of (type-I and type-II error) artifacts have variously been called “the incidentalome” (Kohane, Masys, & Altman, 2006) and “the $P < 0.05$ false-positive studies” (Nebert, Zhang, & Vesell, 2008).

It should be recognized that multiple parameters (effect-size, allelic frequency, significance level, sample size) will all affect statistical power for any genotype-phenotype association study. Statistical power obviously improves with larger numbers of cases and controls. As any minor allele frequency (MAF) increases, fewer subjects are usually needed in the study group, and the level of detectable contribution by a genetic variant to a phenotype will be lower. If the MAF is low, greater numbers per group will be required, and the level of detectable contribution by a variant to a phenotype will be higher.

However, one cannot always embrace the above statements. For example, when the contribution (effect-size) of the risk variant is measured by heritability, or *variance explained*, influence of the MAF on statistical power of an association study is not significant. Nevertheless, if effect-size is measured by odds ratio (OR) or genotypic relative risk (GRR), the influence of a MAF will be greater — especially for MAFs < 0.05 . In addition to the MAF and prevalence of a disorder, the inheritance model (“*additive*,” “*dominant*,” or “*recessive*”) of a risk allele can also influence power of an association study.

Virtually never do GWAS have sufficient statistical power to detect epistasis (gene $\times$ gene interactions; $G \times G$) (Bhattacharjee et al., 2010; Sackton & Hartl, 2016) or gene-environment ($G \times E$) interactions (D. Thomas, 2010). Moreover, an additional underappreciated class of genetic interactions is *intergenic compound heterozygosity*, i.e. interactions between multiple rare variants contributing to a trait (Gibson, 2011). Such interactions encompass heterozygous combinations of multiple alleles. In a broad sense, if hundreds of mutations — each having a frequency of 0.1% — all contribute to the phenotype, then such an event could contribute substantially to individuals that are heterozygous at many of these loci.

2.7.2. “Missing heritability”: real or imagined?—Twenty years after launching the Human Genome Project, initial GWAS findings became frustrating to those who wanted clear-cut explanations into the etiology of a complex disease (Goldstein, 2009). For most complex diseases, even multiple GWAS variants considered together (e.g. using polygenic risk score) typically explain too little variability in disease occurrence to be of much predictive value (Manolio, 2013; Wray et al., 2013). However, more importantly, some GWAS data had identified potential novel therapeutic targets for treating a complex disease — without knowing its precise etiology. Similarly, some GWAS data might uncover potential therapeutic targets for treating an environmental disease, by learning something about its mode-of-action without necessarily understanding its precise mechanism-of-action.

Many GWAS often explain only a small proportion of heritability (defined as *additive genetic variance*). The absent proportion became known as “missing heritability” (Lander, 2011; Manolio et al., 2009), leading to renewed awareness about the *genetic architecture* of human complex disease and traits (Gibson, 2011; Zhang, 2015) — a topic that had been extensively debated in the early 20th century.

First, it was realized that heritability attributable to some common variants could be substantial. However, as GWAS cohort sizes continued to increase — thereby identifying additional variants contributing smaller effects to the trait — the “revealed heritability” continued to grow, albeit rarely reaching >20–25% for various diseases and traits (Lander, 2011).

In addition, current GWAS may overlook many variants of lower frequency (MAFs 1–5%), because existing SNP-typing arrays often lack a more useful marker. Many complex disease-related alleles are probably included in this frequency class. Also, new genotyping arrays and imputation methods, based on the 1000 Genomes Project (The 1000 Genomes Project Consortium, 2012, 2015) or The Haplotype Reference Consortium (McCarthy et al., 2016), are able to capture these less frequent variants. This topic is discussed in detail later.

GWAS also miss many common small-effect variants — due to limited sample-size and/or stringent statistical thresholds imposed to ensure reproducibility. Efforts seeking to infer contributions of loci that fall just short of statistical significance were then addressed (Park et al., 2010; Yang et al., 2010). Contributions of loci that fall even further short of statistical significance, likewise, will result in even smaller effect-sizes on phenotype. Although their individual contributions may be too small ever to detect by investigators designing feasible sample cohort sizes, these very-small-effect variants collectively will probably also explain a significant fraction of heritability (Gibson, 2010; Lander, 2011).

Furthermore, rare variants of large-effect will sometimes contribute substantially to common diseases, although their roles are just recently being explored. Increases in our understanding of rare variants have now advanced — *via* whole-exome sequencing (WES) (Bertier, Hetu, & Joly, 2016) and whole-genome sequencing (WGS); also called “next-generation” sequencing (NGS) (Cirulli & Goldstein, 2010; Pinto, Ariani, Bianciardi, Daga, & Renieri, 2016).

Whether rare variants will lead to discovery of a substantial number of new genes — is a question that perhaps can be answered by systematic WES or WGS. Given the background rate of rare variants, many thousands of samples will be required to achieve statistical significance. Correspondingly, how to quantify total heritability due to rare variants remains unclear. Although the inferred effect-sizes are larger, overall contribution to heritability may be small because of their low frequencies (Gibson, 2011; Zhang, 2015).

Lastly, some “missing heritability” might be purely an illusion (Lander, 2011) because heritability is estimated from epidemiological data by applying principles for inferring additive genetic effects. These approximations may be overestimated — due to methods that are not effective at excluding nonlinear contributions of $G \times G$ interactions, $G \times E$ interactions, or epistasis — which are likely to be important.

2.7.3. Rare versus common variants—Numerous explanations for “missing heritability” have been postulated (Gibson, 2011), including: the ‘*infinitesimal model*’ (large number of variants across the entire allele frequency spectrum of small effects); the ‘*rare variant model*’ (multiple large-effect rare variants that are poorly tagged by genotyping arrays); and the ‘*broad-sense heritability model*’ (contributions from $G \times G$, $G \times E$, and/or epigenetic interactions).

The *infinitesimal model* is a classic model for polygenic traits — first introduced by Fisher (1918). This model assumes a very large number of genetic loci, each with an infinitesimally small-effect. Those loci detected by GWAS are simply the largest effect-sizes (*i.e.* “low hanging fruits”). If a half dozen common variants explain 10% of the phenotype in a population, then the remaining 90% reflects innumerable variants in which each contributes very small effect on the phenotype (*i.e.* explaining $\ll 1\%$ variance) (Gibson, 2011).

Ultimately, every very-small-effect variant will contribute to every trait — but, with effect-sizes so small that it might take samples greater than the population size of the cohort, or even the species size on the planet, to detect! Examples include the massive meta-analyses of GWAS for height (Lango Allen et al., 2010; Wood et al., 2014) and body mass index (BMI) (Locke et al., 2015; Speliotes et al., 2010) in each of these cases, involving several hundred thousand people. In these studies, it has been shown that several hundred loci can contribute to some multifactorial traits; still, these loci often will not necessarily explain even one-fourth of all genetic variance. For example, using genome-wide data from $> 253,000$ individuals, 697 variants were identified at genome-wide significance, which, together, explain $\sim 20\%$ of the heritability for adult height (Wood et al., 2014). Therefore, the infinitesimal model suggests that heritability is not so much “missing”, but simply “hidden” beneath statistically-significance thresholds (Gibson, 2010). The largest-yet GWAS on human height has recently appeared (Marouli et al., 2017) in which almost 459,000 individuals, followed by an independent set of more than 252,000 participants, were studied using ExomeChip (Grove et al., 2013) to test low-frequency coding variants association. In addition to the previously reported ~ 700 common variants, 83 additional height-associated coding variants were found — having lower MAFs in the range of 0.1% to 4.8%. These data confirm further that increasingly larger sample sizes will continue to uncover additional rare and low-frequency SNPs having moderate-to-large effects associated with polygenic human traits. However, these newly identified rare and low-frequency variants only explained an additional small fraction (1.7%) of the heritability of height.

The *rare variant model* emphasizes contributions of multiple rare (MAFs $< 0.5\%$) or low-frequency (MAFs $< 5\%$) — as long as they are large-effect variants (Pritchard, 2001). In this model, *expressivity* can be modified by other loci or by the environment (Bodmer & Bonilla, 2008), but the idea is that the rare-susceptibility genotype is largely responsible for the trait. The rare-allele model would generally refer to dominant effects — caused either by *haploinsufficiency* or *gain-of-function* alleles. This would elevate risk > 2 -fold above background; under these conditions, *penetrance* need not be anywhere near 100%. In fact, the vast majority of unaffected individuals would be expected to carry one or more risk alleles. Thus, the notion is that a multifactorial trait is actually a combination of hundreds, even thousands, of similar but genetically heterogeneous conditions — attributable to rare

variants at individual loci. If each of these variants explains most of the phenotype in just a few people, their effects would be insufficient to explain variance in a total population, detectable by the usual GWAS. The total number of loci that may contribute to any multifactorial trait would be a function of the baseline incidence, number of rare variants per locus, and GRR (*i.e. effect-size*).

The *broad-sense heritability model* suggests that additive contributions, of common variants and large-effects of rare variants, are still insufficient to explain “missing heritability.” There are numerous studies detecting $G \times G$ interactions clinically and $G \times G$ interactions in model-organism quantitative genetic research. Moreover, an increasing number of studies have described epigenetic effects (Slatkin, 2009), parent-of-origin genetic contributions (Connolly & Heron, 2015), and inheritance of DNA-methylation patterns (Haque, Nilsson, Holder, & Skinner, 2016). The hypothesis of this third model is that, because GWAS measure only average effects of alleles across thousands of individuals, GWAS are incapable of capturing heterogeneity of effect-sizes at the family or the individual level — which would be a guarantee of the proposed broader components of genetic architecture. It cannot be ruled out that *broad-sense heritability* is likely to contribute some indeterminate amount of heritability to multifactorial traits.

3. Genetic studies of PGx traits

“Genetic differences in drug response” have long been recognized. Some suggest that Pythagoras — in southern Italy, about 510 BCE — was the first to recognize “dangers of some individuals, but not others, who eat fava beans” [*reviewed in* (Nebert, 1999)]; the adverse reaction in this case is hemolytic anemia due to glucose-6-phosphate dehydrogenase (G6PD) deficiency. During the 1940s, genetic differences in G6PD protein stability and function were discovered to be the cause [*reviewed in* (Motulsky, 1964)]. Others propose that Snyder's original study in 1932 — in which he described the “phenylthiourea nontaster” phenotype inherited as an autosomal recessive trait (Snyder, 1932) — represents the dawn of modern pharmacogenetics.

3.1. Early-day PGx examples

3.1.1. CYP2D6 gene polymorphism—The first PGx trait — elucidated by recombinant DNA methodology — was the debrisoquine/sparteine polymorphism (Gonzalez et al., 1988); the defective-enzyme disorder had been discovered independently in England (Mahgoub, Idle, Dring, Lancaster, & Smith, 1977) and Germany (Eichelbaum, Spannbrucker, Steincke, & Dengler, 1979). Dominant, additive, and recessive phenotypes were named “extensive-metabolizer” (EM), “intermediate-metabolizer” (IM), and “poor-metabolizer” (PM), respectively. Cloning the *CYP2D6* gene and identifying several variant alleles (Gonzalez et al., 1988) was the first time a polyallelic mechanism could be explained in molecular terms as a quantitative PGx trait. PM recessive alleles might encode an inactive CYP2D6 enzyme, unstable protein, incorrect splicing of the gene transcript, or even complete deletion of the gene; hence, clinical phenotypes resulted in lowered, or completely absent, CYP2D6 activity.

Subsequently, an “ultra-metabolizer” (UM) phenotype was discovered, caused by two to 13 *CYP2D6* duplicated genes (Bertilsson, Dahl, Dalen, & Al-Shurbaji, 2002). Thus, distribution of N individuals as a function of CYP2D6 phenotype is not simply bimodal or trimodal; it begins to look like a gradient with perhaps four distinct peaks (Fig. 2). Moreover, striking ethnic differences have been found among *CYP2D6* alleles responsible for PM, IM, EM and UM phenotypes (Teh & Bertilsson, 2012) — suggesting a genetic bottleneck perhaps triggered by some sort of environmental selective pressure.

Eliglustat is a potent selective inhibitor of glucosylceramide synthase, important for the biosynthesis of glucosylceramides that accumulate in Gaucher disease. Because CYP2D6 predominantly metabolizes this drug — CYP3A to a lesser extent — the patients' CYP2D6 metabolizer status, and use of concomitant drugs, must be carefully considered (Balwani et al., 2016; Belmatoug et al., 2016).

The “CYP2D6 panel” now comprises at least 25% of all commonly-prescribed drugs [(Lu, Wang, & Lin, 2003); <http://medicine.iupui.edu/clinpharm/ddis/>]. In the vast majority of cases, deficient CYP2D6 causes an *adverse effect* — the parent drug reaching toxic levels due to insufficient metabolism in PM patients. The opposite occurs in the case of opiates such as codeine, however. The high-CYP2D6-activity phenotype patient is at greater risk of an adverse effect, because the inactive parent drug requires CYP2D6 metabolism to form the active moiety, morphine (Linares, Fudin, Schiesser, Daly Linares, & Boston, 2015).

3.1.2. N-acetylator phenotype—Originally termed “isoniazid acetylation polymorphism”, this unusual drug response was identified clinically in the late 1940s when subjects — who had converted to a positive tuberculin test — were routinely treated with isoniazid. The clinical presentation was an unexpectedly high incidence of peripheral neuropathy. Using a 6-h plasma-level timepoint after an isoniazid standard dose, a *bimodal distribution* was reported (Price Evans, Manley, & McKusick, 1960). This was the first example of clearance of a parent drug that is genetically determined — *i.e.* the major enzyme that metabolizes isoniazid (*N*-acetyl-transferase) is defective in efficient removal of the parent drug. As the result of “slow” vs “rapid” plasma clearance, individuals were phenotyped as “slow acetylators” (*r/r*) or “rapid acetylators” (*R/R*, *R/r*); the autosomal recessive inheritance is consistent with the earlier findings of Mendel and Garrod. Three decades later (Blum, Grant, McBride, Heim, & Meyer, 1990), the *NAT2* gene, and not *NAT1*, was shown to be responsible for Mendelian inheritance of isoniazid *N*-acetylation.

The *NAT2* polymorphism is not only relevant to isoniazid, but also to dozens of other drug and environmental toxicant substrates of NAT2 (Grant, Goodfellow, Sugamori, & Durette, 2000; Walraven, Zang, Trent, & Hein, 2008). As is the case for virtually every drug-response gene due to the great human diasporas (Nebert & Dieter, 2000), striking inter-ethnic differences in frequency of *NAT2* alleles were discovered, *e.g.* among Japanese the PM phenotype subset is <10%, whereas among Caucasians PMs comprise >70% [reviewed in (Nebert, 1999)].

3.1.3. CYP2C19 gene polymorphism—CYP2C19 participates in metabolism of at least four dozen commonly-prescribed drugs [<http://medicine.iupui.edu/clinpharm/ddis/>]. The

story is similar to the two above-described examples: the EM phenotype appears to be dominant over the PM phenotype, and almost always the parent drug in PMs exhibits ADRs due to overdose. Ethnic differences are very striking for *CYP2C19* alleles; *e.g.* the East Asian PM subset is ~33% and Oceanian PM subset > 50%, whereas the Caucasian PM subset represents <6% (Mizutani, 2003).

CYP2C19 catalyzes bioactivation of clopidogrel, an anti-platelet prodrug. Loss-of-function alleles (*e.g.* *CYP2C19*2*) impair formation of active metabolites, resulting in decreased platelet inhibition and increased risk for adverse cardiovascular events, especially in patients undergoing percutaneous coronary intervention (PCI) (Johnson et al., 2012). Therefore, alternative anti-platelet therapy (*e.g.* prasugrel, ticagrelor) is now recommended for *CYP2C19* PM or IM patients — if there is no contraindication (Scott et al., 2013).

3.1.4. Abacavir-induced hypersensitivity—In an early example of identification of a PGx disorder by a candidate-region study, 185 patients treated with abacavir — an HIV-1 nu-cleoside-analog reverse-transcriptase inhibitor — resulted in a life-threatening hypersensitivity syndrome in 18 patients (Mallal, Nolan, et al., 2008). SNP-typing of loci in the major histocompatibility complex (MHC) region revealed a very strong association with the *HLA-B*57:01* allele (OR = 117) and also with the combination of *HLA-DR7* and *HLA-DQ3* loci (OR = 73), when compared with 167 abacavir-tolerant controls. A high association of abacavir-induced hypersensitivity with *HLA-B*57:01* was later confirmed in an elegant double-blind prospective, randomized study involving almost 2000 patients from 19 countries (Mallal, Phillips, et al., 2008).

3.1.5. Coumarin-associated gene polymorphisms—Optimizing warfarin, coumarin and acenocoumarol dosage for anticoagulation is an oligogenic example of a substantial contribution of several genes — resulting in rapid *vs* slow PK and/or PD of coumarins. *CYP2C9* (J. A. Goldstein & de Morais, 1994), *VKORC1* (D'Andrea et al., 2005), and *CYP4F2* (Caldwell et al., 2008) polymorphisms were independently discovered in candidate-gene studies and then later confirmed in GWAS (Cooper et al., 2008; Takeuchi et al., 2009). The combination of mutations in the *CYP2C9*, *CYP4F2* and *VKORC1* genes explains as much as 35–50% of the patient's total variation in drug response (Daly, 2009). This means the remaining genetic impact on coumarin response is derived by contributions of genes other than these three major genes.

3.1.6. Further examples of PGx traits—The above-described *NAT2*, *CYP2D6*, *CYP2C19*, *CYP2C9* and *CYP4F2* polymorphisms represent large-effect PK genes encoding enzymes involved in metabolism.

On the other hand, *HLA-B*57:01* represents an immune-response gene, whereas *VKORC1* and *IL28B* are considered to be large-effect PD genes. *HLA-B*57:01* encodes a particular cell-surface molecule responsible for presentation of endogenous peptides to immune system cells. *VKORC1* codes for vitamin K-epoxide reductase complex subunit-1 that is targeted directly by coumarins. Considered as vitamin K antagonists (Jackson & Suttie, 1977), coumarins are potent inhibitors of this reductase complex, leading to depletion of reduced vitamin K — essential for normal coagulation. This epoxide reductase is therefore

regarded as a “drug target” for coumarins, and, except for the vitamin K substrate, does not appear to metabolize any other drug.

Three additional large-effect PK genes showing substantial quantitative clinical differences in drug response include the thiopurine S-methyltransferase gene (*TPMT*), UDP glucuronosyltransferase-1A1 gene (*UGT1A1*), and dihydropyrimidine dehydrogenase gene (*DPYD*).

TPMT encodes an enzyme of critical importance in thiopurine drugs and therefore treatment of childhood leukemia (Eichelbaum, Ingelman-Sundberg, & Evans, 2006; Weinshilboum & Sladek, 1980). Of interest, a recent GWAS of > 1000 children with leukemia found *TPMT* as the only gene achieving genome-wide significance — top hit was rs1142345; 719A > G; $P = 8.6 \times 10^{-61}$ (Liu et al., 2016), thereby supporting the idea that differences in *TPMT* activity might reflect a Mendelian, or predominantly oligogenic, trait. These data further confirm the clinical usefulness of the red-cell *TPMT* enzyme assay, which is a phenotype (not genotype) test, because discovery of a completely new DNA variant that causes low *TPMT* enzyme activity is always possible.

UGT1A1 encodes the enzyme UGT1A1 that metabolizes, among many other drugs, irinotecan. Poor-metabolizer *UGT1A1**28 homozygosity is strongly associated with irinotecan-induced neutropenia; it is recommended — on the irinotecan product label for *UGT1A1**28 homozygotes — to decrease the starting dose (Kweekel, Guchelaar, & Gelderblom, 2008). Due to lack of sufficient prospective data, however, it remains uncertain whether dose-reduction leads to decreased toxicity or altered antitumor effect. Combined toxicity analysis indicates that most patients experiencing grade 3–4 diarrhea and/or neutropenia are not homozygous for *UGT1A1**28.

Frequently prescribed anticancer drugs, the fluoropyrimidines, are inactivated by hepatic dihydropyrimidine dehydrogenase (*DPYD*). As much as 5% of cancer populations exhibit *DPYD*-deficiency, which is considered having practical value (and even cost-effective) gene product for its anticancer drug substrates; *DPYD**2A is strongly associated with fluorouracil (5-FU) induced severe and life-threatening toxicity (Deenen et al., 2016). There is convincing evidence to implement prospective *DPYD* genotyping with an up-front dose adjustment in *DPYD*-deficient patients (Lunenburg et al., 2016). In each of these examples, if the PGx assay reveals diminished enzymatic activity or deficiency, lower doses or alternative drugs are usually recommended.

3.2. PGx genome-wide association studies (GWAS)

Advances in genomic technology — such as microarray-based genotyping platforms and NGS technologies, along with development of statistical packages for large-scale data analysis — have enabled researchers to perform genome-wide searches for novel and less obvious genes associated with drug responses. Obviously, it is more difficult to identify a large number of individuals being treated with a specific drug (Daly, 2010), compared with, for example, identifying a large number of individuals for height, blood pressure, or BMI. Since 2007, however, dozens of PGx GWAS with sufficient numbers in the cohorts have identified many genetic variants associated with various drug responses and adverse effects

(Table 1). GWAS are hypothesis-free and thus do not require a priori assumptions about genomic location of functional variants (McCarthy et al., 2008; Stranger, Stahl, & Raj, 2011); therefore, GWAS provide an unbiased and powerful tool for systematic discovery of genetic variants associated with PGx traits (Motsinger-Reif et al., 2013).

3.2.1. GWAS of ADRs—In the past few years, several GWAS have detected novel genetic loci associated with severe ADRs (Daly, 2012). Probably the first successful example is identification of a genetic risk variant for myopathy in patients treated with simvastatin (Link et al., 2008). Authors first screened ~300,000 markers in 85 cases and 90 controls selected from a clinical trial of ~12,000 participants, and then replicated the finding in ~20,000 subjects; they identified a *nonsynonymous variant* (rs4149056) in the *SLCO1B1* gene that confers an OR of 16.9 in the homozygous state (Table 1). Collectively, >60% of myopathy cases, receiving a high simvastatin dose (80 mg), can be attributed to the risk allele. The *SLCO1B1* gene codes for an organic anion transporter, suggesting that it functions in cellular uptake of statins (König, Seithel, Gradhand, & Fromm, 2006); however, this hypothesis was not firmly established, according to previous candidate-gene association studies (Morimoto et al., 2005). Subsequently, the correlation was reproduced in further clinical studies (Voora et al., 2009), and simvastatin dosing, based on the *SLCO1B1* genotype, has been recommended by the Clinical Pharmacogenomics Implementation Consortium (CPIC) (Ramsey et al., 2014). Intriguingly, association of the *SLCO1B1* rs4149056 variant with myopathy has been less compelling for other statins (Niemi, 2010).

In another PGx GWAS (Daly et al., 2009), the *HLA-B*57:01* allele was identified as a major risk factor for flucloxacillin-induced liver injury — a severe, but quite rare, ADR (~8.5 in every 100,000). In a small cohort (51 cases, 282 controls), authors identified a strong association at a SNP marker in complete linkage disequilibrium (LD) with *HLA-B*57:01* having an OR = 80.6 (Table 1). However, despite the strong association, only one in every 500 to 1000 flucloxacillin-treated individuals with this allele will develop liver injury, thereby limiting clinical utility (*i.e.* the test has a very high false-positive rate). This is an example in which *incomplete penetrance* makes it difficult for the physician or pharmacologist, who must make clinical decisions on a patient-by-patient basis.

Following this study, more GWAS have demonstrated strong associations between several *HLA* haplotypes and various ADRs (Table 1) — including lumiracoxib-induced liver injury (Singer et al., 2010), carbamazepine-induced hypersensitivity (McCormack et al., 2011), and lapatinib-induced liver injury (Parham et al., 2016). However, predictive values for typing each of these *HLA* alleles associated with these ADRs, unfortunately, are very low (Daly, 2014).

Other than the *HLA* region, PGx GWAS have also identified a number of associations with various ADRs (Fernandez et al., 2015; Karnes et al., 2015; Xu et al., 2015) — which would not have been suspected, based on established knowledge about each drug's PK and/or PD (Table 1). Although these discoveries may not have immediate clinical utility, they point to new biological pathways underlying each drug's mode-of-action and/or toxicity.

Unfortunately, not all GWAS of ADRs have yielded robust associations. For example, many GWAS of drug-induced QT-interval prolongation did not produce consistent results [*recently reviewed in* (Niemeijer, van den Berg, Eijgelsheim, Rijnbeek, & Stricker, 2015)].

Osteonecrosis of the jaw (ONJ) is a severe ADR, associated with chronic bisphosphonate treatment; initial GWAS (Sarasquete et al., 2008) revealed a correlation between ONJ and CYP2C8 variants, and this P450 enzyme is a plausible candidate for having a significant role in metabolism of several anticancer drugs; however, this finding has not been replicated in subsequent studies (English et al., 2010; Katz et al., 2011; Such et al., 2011). ADRs such as ONJ are well known to be extremely heterogeneous phenotypes, similar to many other complex diseases; therefore, unless one has a very large sample-size, loci that can be replicated will continue to be difficult to identify with certainty.

3.2.2. GWAS examples of drug efficacy and dosage—Warfarin is a widely used anticoagulant but its usage is complicated, due to wide inter-individual variability in dosing regimen and its narrow therapeutic “window”; this makes warfarin dosage a very active area of PGx studies. Candidate-gene association and molecular studies — well before the genomic era — had already revealed significant roles of two important genes (*CYP2C9* and *VKORC1*; *vide supra*). Multiple GWAS have now confirmed these known associations (Cooper et al., 2008; Takeuchi et al., 2009) and provided more accurate estimates of their effects in determining the observed variance (*VKORC1* = ~30% and *CYP2C9* = ~12%). Furthermore, a GWAS with larger sample-size (Takeuchi et al., 2009) also discovered an additional locus with relatively small-effect (*CYP4F2* = 1.1%); this *CYP4F2* polymorphism, however, has not been included in most dosing algorithms (Johnson et al., 2011; Pirmohamed, Kamali, Daly, & Wadelius, 2015). GWAS have also identified similar genetic determinants involving other coumarin anticoagulants (Teichert et al., 2009), as well as in several different ethnic groups (Cha et al., 2010; Parra et al., 2015; Perera et al., 2013).

Another early successful example of GWAS of drug efficacy came from studies of variable response to interferon- α treatment of hepatitis C virus (HCV) infection. In 2009, three independent studies (Ge et al., 2009; Suppiah et al., 2009; Tanaka et al., 2009) reported genetic variants near the *IL28B* gene associated with interferon- α treatment responses; the effect is measured by sustained virological response (SVR; *i.e.* absence of detectable virus at end of follow-up). The same variant was also reported to be correlated with spontaneous clearance of HCV (Thomas et al., 2009), suggesting the variant might influence host immune response — with or without drug intervention.

Also relevant, the advantageous allele (C allele at rs12979860) is more frequent in Caucasian and Asian populations than in Africans — a difference that very likely explains the previous report (Conjeevaram et al., 2006) of relatively poor outcomes among HCV-infected African Americans in trials of interferon- α treatment. Despite the large estimated effect-size (OR = 7.3), however, the imperfect predictive power, combined with lack of alternative treatments, had limited immediate clinical utilization of determining the *IL28B* genotype in personalized treatment decisions (Iadonato & Katze, 2009).

The new direct antiviral agents (DAAs), in association with the pegylated interferon/ribavirin combination, is the new standard of HCV treatment making it necessary to re-

evaluate the clinical utility of the *IL28B* genotype (Thompson & McHutchison, 2012). Although *IL28B* genotype testing continues to provide important clinical information regarding expected treatment response, *IL28B* genotyping will eventually lose its utility — as HCV therapy moves away from interferon-based regimens (Jensen & Pol, 2012; Liapakis & Jesudian, 2012).

3.2.3. GWAS of more complex PGx traits—GWAS have also been used to study efficacy of additional widely used medications for common complex disorders — such as type-2 diabetes (T2D), dyslipidemia, hypertension, and psychiatric disorders. The GoDarts Group (GoDarts et al., 2011) conducted GWAS for glycemic response to metformin in >1000 individuals with T2D and identified variants in the ataxia-telangiectasia-mutated gene (*ATM*) that were associated with treatment success; the association was replicated in an independent cohort with almost 1800 samples. The combined effect was 1.35 times higher for reaching the treatment goal (HbA1c <7%) for each minor allele (C) of rs11212617, at which the strongest association was observed.

Consequently, this association was replicated in multiple additional cohorts (van et al., 2012); however, this variant has not modified the metformin response in the Diabetes Prevention Program (DPP) (Florez et al., 2012). The rs11212617 variant thus appears to have no significant impact on the predictive value with regard to metformin action on glycemic control. Nevertheless, for reasons not clear, several observations appear to suggest a substantial role of *ATM* in glucose homeostasis. For example, ataxia telangiectasia patients can manifest a severe form of diabetes, and *Atm*($-/-$) knockout mice exhibit insulin resistance and abnormal adipose distribution (Takagi et al., 2015). More recently, in a large GWAS consisting of ~13,000 participants, the Metformin Genetics (MetGen) Consortium reported an association between an *SLC2A2* (facilitated glucose transporter) variant and metformin treatment response (Zhou et al., 2016).

Treatment response to statins has also been examined by multiple GWAS (Chasman et al., 2012; Deshmukh et al., 2012; Hopewell et al., 2013; Postmus et al., 2014) [*reviewed in* (Leusink, Onland-Moret, de Bakker, de, & AH, 2016)], with large sample-sizes ranging from ~2000 to >40,000 samples. These well-powered studies have identified common SNPs in five genes (*LPA*, *APOE*, *SLCO1B1*, *SORT1* and *ABCG2*) —albeit with modest effect-sizes. These variants are robustly associated with LDL-C response; however, none consistently modifies risk reduction for cardiovascular events, and therefore these findings have limited clinical utility.

3.2.4. Unsuccessful examples—Hypertension is the most prevalent modifiable risk factor for global disease burden, and many PGx studies have attempted to identify genetic variants responsible for the highly varying responses to one or another anti-hypertensive medication. Despite extreme efforts, these studies have yet to generate robust independently replicated findings (Lupoli, Salvi, Barcella, & Barlassina, 2015; Menni, 2015); such failures likely reflect the extreme heterogeneity and complexity of hypertensive disease. Elevated blood pressure has multiple etiologies and involves varying physiological changes throughout one's lifetime. Identified GWAS variants associated with blood pressure are far fewer than those associated with blood glucose or blood lipid levels, and identified blood

pressure variants also explain less phenotype variance (<3%) (Munroe, Barnes, & Caulfield, 2013) — which is probably further evidence consistent with a high degree of complexity for hypertensive disorders.

Moreover, the multiple classes of antihypertensive drugs (diuretics, ACE inhibitors, angiotensin II receptor blockers, calcium channel blockers, beta-blockers) — alone, and in combinations — further complicate “clean” PGx studies, not to mention variance due to within-person fluctuations of blood pressure (*e.g.* changes in diet, exercise, time-of-day, combinations of prescribed drugs, alcohol usage, *etc.*). Several PGx GWAS have reported variants in plausible genes associated with thiazide diuretics (Chittani et al., 2015; Turner et al., 2008; Turner et al., 2013), angiotensin II receptor blockers (Frau et al., 2014; Turner et al., 2012), and beta-blockers (Gong et al., 2015; Gong et al., 2016); yet, most of these studies involve relatively small discovery sample-sizes and are underpowered, as far as detecting associations having genome-wide significance (*i.e.* $P < 5 \times 10^{-8}$). Large international consortia have been established — aimed at increasing opportunities for discovery and replication by assembling large samples (Cooper-DeHoff & Johnson, 2016). However, even though large consortium studies might be able to identify genetic-variant associations in the future, it is hard to imagine that, even in combination, any of such newly identified small-effect variants will have important clinical predictive power.

Similarly, GWAS of antidepressant treatment response (Biernacka et al., 2015; Garriock et al., 2010; Ising et al., 2009; Uher et al., 2010) have not yielded promising results. One particular challenge in PGx studies of psychotropic medications is the evaluation of disease conditions and quantitative measurements of treatment responses, which — in contrast to blood pressure or serum cholesterol levels — are all based on multidimensional psychiatric diagnostic criteria. This method has long been criticized for its lack of reliability (*e.g.* degree of consensus on a diagnosis or treatment response as viewed by different clinicians; the wording and how questionnaires are answered by patients) [*reviewed in* (Aboraya, Rankin, France, El-Missiry, & John, 2006)]. This complexity, comparing the “equivocal phenotype” *vs* the “unequivocal phenotype”, has previously been emphasized (Nebert et al., 2008).

The same is especially true for major depressive disorder (MDD) (Freedman et al., 2013). Clinically-based psychiatric criteria do not directly reflect biologically homogeneous conditions — which makes phenotypic definition of psychiatric conditions highly heterogeneous and measurement of treatment responses very inaccurate. At least in part due to these difficulties, GWAS of MDD with reasonably large sample size (>9000 cases) have failed to identify any replicable association (Ripke et al., 2013). In contrast, a PGx study, evaluating parent drug (citalopram and escitalopram) and metabolite plasma concentrations in 435 MDD patients, successfully identified highly significant associations with *CYP2C19* and the *CYP2D6* variants (Ji et al., 2014). This important distinction — between diffuse psychiatric clinical phenotypes and quantitative measurements of blood or urine drug- or metabolite-level phenotypes — has previously been emphasized (Nebert et al., 2008). Therefore, it becomes problematic to discern how these pharmacometabolomic association studies [*e.g.* (Ji et al., 2014)] will aid in *genetic risk prediction* of many complex clinical responses.

4. Genetic architecture of PGx traits

Genotype-phenotype association studies, especially those recently described PGx GWAS (*vide supra*) have identified many genetic variants correlated with numerous drug-response phenotypes. These results have opened a window into the “*genetic architecture* of PGx traits.” Genetic architecture defines how genotypes map to phenotypes (*cf.* Glossary of Terms); specifically, it includes descriptions about the number of traits-associated variants, their allele frequencies and effect-size distribution, combined with patterns of pleiotropy and epistasis (Zhang, 2015). Knowledge about genetic architecture of PGx traits is not only important to our understanding the biology, but also is instrumental in PGx study-design and analysis. Genetic architecture primarily defines potential predictive power of PGx response using genomic data (Chhibber et al., 2014; Zhou & Pearson, 2013). In what follows, we have classified PGx traits into three broad categories (Mendelian, predominantly oligogenic, and complex) — based on their genetic architecture — and we discuss the implications in study design and clinical utility.

4.1. Monogenic (Mendelian) PGx traits

4.1.1. PGx response—Many (ordinarily rare) severe ADRs are usually expressed as binary traits (*i.e.* “present”, or “not present”). Occurrence of such rare/severe reactions depends almost exclusively on the presence of causative PGx variants (or mutations), whereas the ADR is less influenced by dosage in the therapeutic range. This class can be further categorized into three subgroups, based on molecular etiology:

1. Severe ADRs due to a major defect in drug metabolism or transporter gene (PK gene). Examples include statin-induced myopathy (*SLCO1B1*) and slow- vs rapid-acetylator phenotype (*NAT2*).
2. Dose-independent (idiosyncratic) ADRs due to inappropriate immune response — including severe drug-induced liver damage and hypersensitivity reactions (usually affecting skin) (Daly, 2013).
3. Genetic mutations in genes involved in mechanism-of-action of a therapeutic drug. For example, hereditary predisposition for venous or arterial thrombosis, such as mutations in *F5*, *PROC* and *PROS1*, have been associated warfarin-induced skin necrosis (Yang & Algazy, 1999).

4.1.2. Genetic basis—The genetic basis of Mendelian PGx traits usually involves nonsynonymous mutations in PK genes (*e.g.* *CYP2D6*, *SLCO1B1* or *NAT2*) or *HLA* genes (Daly, 2013). Because of the discrete nature of this type of binary PGx response and relatively large-effect of the causative genetic variant(s), consequence of genetic changes can be efficiently measured by OR or penetrance.

It should be emphasized that — even this simplest form of PGx trait is not simple, for two reasons. First, multiple mutations in a gene often influence gene function to varying degrees, *e.g.* to date, there are >180 different alleles reported for the *CYP2D6* gene [The Human Cytochrome P450 (*CYP*) Allele Nomenclature Database; <http://www.cypalleles.ki.se/>], which can cause a gradient (from mild to severe) of CYP2D6 enzyme deficiency. The same

can be said for *SLCO1B1* or *NAT2*. Second, although presenting with high OR, risk alleles may not always cause the ADR; as described earlier, this is an example of *incomplete penetrance*. For instance, despite the strong association (OR = 80) of flucloxacillin-induced toxicity with the *HLA-B*57:01* allele, only one in every ~500 to ~1000 individuals with this genotype, when treated with flucloxacillin, will develop drug-induced liver injury, (Daly et al., 2009).

4.1.3. Genetic study—Given the high OR (or penetrance), case/control studies with even small numbers of cases can be an efficient approach to identification of genetic variants responsible for this type of PGx trait. Especially for those idiosyncratic cases in which we have no plausible a priori biological hypothesis, GWAS with a small number of case/control samples can be a powerful approach to identifying causative variants (Daly, 2012). However, given the low occurrence of such rare ADRs — identification of even a small number of cases can be a daunting task — likely requiring international collaborations and a worldwide reporting system (Wilke et al., 2007).

4.1.4. Clinical utilization—Because the severity of such rare ADRs can sometimes be life-threatening, therefore, screening for these PGx variants that might cause this type of rare/severe ADR can be used as an exclusion criterion (contraindication). However, because of this rarity, sometimes it can be inefficient for pharmacoeconomical reasons, to screen every patient for such rare variants (Daly, 2013). Although individually less frequent, each individual has a higher chance of carrying at least one large-effect PGx variant that, in combination, can cause a major ADR, which can be prevented by preemptive PGx testing [*cf.* Section 7].

4.2. Predominantly oligogenic PGx traits

4.2.1. PGx response—This group of PGx traits usually involves quantitative variations in concentration of drug metabolites, drug efficacy, dosage, and/or toxic effects. Variability is expressed in a continuous spectrum among individuals in any population. Because of the existence of major contributing factors having predominantly large effect-size, distribution can exhibit multiple distinctive modes (*e.g.* CYP2D6-metabolizer groups; Fig. 2). However, these effects are usually seen only in well-controlled experimental studies. In clinical settings, *the distinction* between different modes (subgroups) will *usually be blurred*— due to other factors such as drug dosage, drug-drug interactions, and many environmental influences.

4.2.2. Genetic basis—The genetic basis of predominantly oligogenic PGx traits is usually influenced by common genetic variants in one or several major (PK or PD) genes having large-effect. In each of these genes, there may be multiple variants, coding or regulatory, that elicit a functional impact on drug response. Because of the quantitative nature of this group of PGx traits, effect-sizes of the associated genetic variants can be measured by *percentage of variance explained* (R^2). For example, common variants in *VKORC1* and *CYP2C9*, combined, can account for ~40% variance in warfarin dosing (Fung et al., 2012). Cigarette smoking-induced CYP1A2, as well as the heritable *CYP1A2* and *CYP2D6* polymorphisms, appear to explain ~20% of variance in serum fluvoxamine levels (Carrillo et al., 1996).

4.2.3. Genetic study—Given their relatively large effect-sizes, common variants having a major impact on predominantly oligogenic PGx traits can be identified by genetic association studies — using a regression-based approach for studying continuous variation in drug responses, or a case/control design based on dichotomization of a continuous drug response gradient, or the extreme discordant phenotype design (Zhang, Nebert, Chakraborty, & Jin, 2006).

Sample-size requirement for this type of study is principally defined by effect-size of the genetic variant and significance level (which needs to be adjusted for multiple-testing, if multiple genetic markers are examined). For example (Fig. 3), a GWAS with ~1000 samples could have ~80% power to identify variants that explain 4% variance at a significance level of $5e-8$ (adjusted for 1 million independent tests). After identification of the major genes or variants having large impact (say, $R^2 > 5\%$), increasing sample-size will continue to reveal more variants having minor effects. However, these small-effect variants — even in combination — will not substantially increase *variance explained* in drug response; hence, these variants will not substantially improve the predictive value. That is to say, those genetic variants that cannot be identified within a reasonably large sample-size (*e.g.* $N = 1000$) will have minimal predictive value in each individual patient. Using warfarin as an example, although larger GWAS did identify additional variants robustly associated with the phenotype, they added limited predictive power to the two major genes (*VKORC1* and *CYP2C9*).

4.2.4. Clinical utilization—Given the relatively large effect-size, it is possible to predict predominantly oligogenic PGx traits based on major contributing genetic variants. However, the true clinical utility might be complicated by several challenges:

1. In addition to major genetic determinants, this type of PGx trait is usually influenced by several or many environmental factors;
2. The predictive value can be different between ethnic populations. For example, *VKORC1* and *CYP2C9* variants account for as much as 40% of variability in warfarin-dosing among Europeans, but <10% among African Americans (Limdi et al., 2008; Limdi et al., 2010; Momary et al., 2007);
3. Predominantly oligogenic PGx traits under evaluation (such as dosage, drug or metabolite concentrations) may not be directly related to the ultimate treatment outcome (Zineh, Pacanowski, & Woodcock, 2013).

In addition, there might be a better alternative to genetic testing — a more accurate dosing or treatment option can often be achieved more rapidly by intensive monitoring. All these factors will require the true clinical utility and pharmacoeconomics of PGx tests to be established based on clinical trials (Ioannidis & Khoury, 2013; Rioux, 2000), with sample-sizes that are much larger than original PGx studies that identified the genetic association.

Using warfarin PGx dosing as an example, the initial association with *CYP2C9* and *VKORC1* was identified and validated by studies with < 100 samples (Wadelius & Pirmohamed, 2007); a dosing algorithm was later established by a large clinical study involving >5000 samples (International Warfarin Pharmacogenetics et al., 2009); however,

the true clinical utility has recently been challenged by three independent clinical trials (Kimmel et al., 2013; Pirmohamed et al., 2013; Verhoef et al., 2013) — each containing several hundreds to one thousand samples. Some fundamental differences in trial design were believed to cause the discordant findings (Pirmohamed et al., 2015; Zineh et al., 2013). This example underscores the need for more elegantly designed and methodologically rigorous clinical studies, to definitively clarify issues of public health relevance of pharmacogenetics.

4.3. Complex PGx traits – principally drug efficacy and ADRs

4.3.1. PGx response—The group of complex PGx traits usually involves variations in efficacy, side effect, toxicity, and common ADRs. Clinically, although one might simply dichotomize the trait as “efficacious” vs “inefficacious” or “causing” vs “not causing” an ADR — this type of variability is decidedly continuous in nature and usually follows a unimodal bell-shaped distribution (Fig. 1D), indicating involvement of multiple factors each having small-effect. Examples of variable responses to commonly used antidiabetic, lipid-lowering, antihypertensive, and antidepressant medications (*vide supra*) probably fall into this category.

4.3.2. Genetic basis—Although complex PGx GWAS are still sparse (Table 1), as far as providing a complete picture of the genetic architecture (Zhou & Pearson, 2013), it is reasonable to presume that the genetic basis of this group of PGx traits might resemble many other complex human traits, *i.e.* influenced by at least hundreds if not thousands of variants. And, similar to other complex traits, the genetic architecture of complex PGx traits might follow more closely the *infinitesimal model* described earlier (Gibson, 2011) — the variance being additively contributed by a very large number of causal variants, each having small-effect. Large-effect rare variants may exist, but will contribute little to *variance explained* in a population, due to their low allelic frequencies (Zhang, 2015). Genetic variants associated with *complex PGx traits* identified so far (Section 3.2.3 and Table 1) are mostly common variants (Zhou & Pearson, 2013).

4.3.3. Genetic study—Because of the complexity in genetic etiology and the uniformly small-effect sizes of underlying genetic variants, the only means to identify genetic variants associated with *complex PGx traits* is by carrying out well-powered (with at least thousands of samples) hypothesis-free GWAS. The resultant identified variants might help in revealing new biology behind drug action or even lead to identification of potential therapeutic targets; however, the results will have very limited predictive power or clinical utility, given their small effect-size. Large sample-sizes may enable identification of small-effect variants; however, it is unrealistic to imagine that ever-increasing sample sizes (to tens of thousands, or more) will reveal sufficient numbers of small-effect variants that, in combination, can explain a substantial fraction of variance and have predictive power.

4.3.4. Clinical utilization—Because precision of predicting PGx complex traits is primarily defined by the percentage of variance explained by predictors, it is difficult to imagine that there will be any chance to predict complex PGx traits by identifiable small-effect genetic variants, even in combination (Fig. 4).

It should be noted that the above categorization of PGx traits is not absolute: the genetic architecture of PGx traits should more or less follow a continuous spectrum, with overlapping among the three categories. For example, the slow-acetylator type — as measured by serum isoniazid levels in experimental subjects — might follow a Mendelian inheritance pattern (autosomal recessive); however, the concentration of isoniazid metabolites could be influenced by drug dosage, drug-drug interactions or other PK genes, such as *CYP2E1* and *GST1*, thereby changing the pattern to a predominantly oligogenic distribution. Moreover, even additional genes might contribute to isoniazid-induced hepatotoxicity.

5. Current status of clinical implementation of PGx – review of FDA PGx labels

We surveyed the current status of clinical utility of PGx by examining the Table of Pharmacogenomic Biomarkers in Drug Labeling, published by the FDA (<http://www.fda.gov/Drugs/ScienceResearch/ResearchAreas/Pharmacogenetics/ucm083378.htm>). The 204 records in that table annotate PGx-labeling information between 165 unique drugs and 67 unique genes or chromosomal abnormalities that influence action of these drugs.

Of the 165 drugs with FDA-approved PGx labeling information, the largest group is the antineoplastic agents ($N = 52$), 42 of which display target-specific somatic changes. The second largest group is the psychotropic drugs ($N = 23$), followed by central nervous system agents ($N = 19$) and anti-infectives ($N = 16$) (Fig. 5).

According to PharmGKB (<https://www.pharmgkb.org/view/drug-labels.do>) (Whirl-Carrillo et al., 2012), the level of action implied in each PGx label can be categorized into: [a] genetic testing required; [b] genetic testing recommended; [c] actionable PGx, and [d] informative PGx. Basically, the first two categories (*genetic testing required* or *recommended*) imply that some genetic testing should be conducted, or is recommended, before using the drug for its indication (or contraindication) of usage, dosage and administration, and/or use in a specific population. The *actionable PGx* category contains information about changes in drug efficacy or risk of ADRs due to genetic variations, with or without dosage recommendations. The actionable PGx information is usually established on relatively strong evidence. The last category (*informative PGx*) refers to *possible* involvement of a gene or protein in drug-action, but no clear evidence to date suggests a specific genetic variant associated with a variable drug response. Definitions of these PGx levels can be found at (<https://www.pharmgkb.org/page/drugLabelLegend>).

Following this categorization method, PharmGKB annotated FDA drug labels with the “PGx level” tag (<https://www.pharmgkb.org/view/drug-labels.do>). The majority (32 of 42) of antineoplastic drugs are known to target specific somatic changes that require genetic testing in order to guide usage; however, for drugs with PGx traits related to germline genetic variation, only 18 (14.6%) contain genetic-testing information. Among those 18, eight drugs target inborn-errors-of-metabolism (e.g. urea-cycle enzyme deficiency, *N*-acetylglutamate synthase deficiency), or genetic disorders (*CFTR*), or cancer (*BRCA1* or *BRCA2* mutations/

olaparib). The remaining ten drugs require genetic testing to detect major deficiencies in PK genes (*TPMT*, *CYP2D6* and *CYP2C19*) or *HLA* genes to avoid adverse reactions.

Of the 67 genes or chromosomal changes in the FDA PGx labels, 20 are somatic changes targeted by numerous ($N = 42$) antineoplastic drugs. For the remaining 47 genes with germline variants that have a PGx effect, 28 genes are actually those in which mutations cause rare metabolic disorders (such as *CYB5R1-4*: cytochrome *b*₅ reductase deficiency; *AKR1D1*, *AMACR*, *CYP27A1*, *CYP7A1*, *HSD3B7*: bile acid synthesis disorders; *HPRT1*: abnormal uric acid metabolism) or genetic diseases (*CASR*, *F2*, *F5*, *PROC*, *PROS1*, *SERPINC1*, *CFTR*, *BRCA1*, *BRCA2*). For understanding these diseases, cf. <https://www.ncbi.nlm.nih.gov/omim>.

These rare conditions involve either the treatment target or they can increase toxicity of drugs being administered for other purposes. After excluding these 28 “rare genetic disorder” genes, there are only 19 PGx genes in the conventional sense — ten of these are PK genes, including five phase I (*CYP2D6*, *CYP2C9*, *CYP2C19*, *CYP2B6*, *CYP3A5*), four phase II (*UGT1A1*, *TPMT*, *NAT2*, *NAT1*), one transporter gene (*SLCO1B1*) and one inborn-error-of-metabolism gene (*G6PD*). The remaining genes in the list of FDA PGx labels include four PD genes (*VKORC1*, *IFNL3*, *LDLR*, *RYR1*) and four *HLA* genes (*HLA-A*, *HLA-B*, *HLA-DQA1*, *HLA-DRB1*).

We might also categorize PGx genes based on the “PGx level” of associated PGx information in the FDA labels (annotated by PharmGKB). Most of the genes/chromosomal abnormalities (16 of 20) that are involved in somatic-cell alterations require genetic testing. However, only five of the “conventional” 19 PGx genes have genetic-testing information with regard to at least one associated drug; these include: *G6PD*, *CYP2D6*, *HLA-A*, *HLA-B* (Genetic testing recommended) and *TPMT* [Genetic testing recommended, but the red-cell enzyme assay (phenotype) is preferred]. Those five genes having genetic testing information are either PK genes with extreme metabolic phenotypes (i.e. *CYP2D6* PM or UM, *TPMT* deficiency), or inborn-error-of-metabolism genes (i.e. *G6PD* deficiency), or *HLA* genes responsible for severe ADRs. For the majority of PK or PD genes that show a more continuous variation-gradient in the population, although one can explain a portion of dissimilarity in drug response, there are no established genetic tests yet.

6. Genomic prediction of drug response—challenges and opportunities

6.1. Importance of the environment and epigenomics

Beyond the scope of this review is the role of the environment (including hepatic and renal function, diet, physical activity, mental stress, cigarette smoking, use of alcohol or recreational drugs, taking over-the-counter drugs concomitantly, occupational exposures to high doses of toxicants, etc.) which obviously can affect drug efficacy, intensity of ADRs, and drug toxicity; hence, the environment can interfere with genetic risk prediction of drug response.

Also, with each additional decade of life, we must keep in mind that environmental insults will cause alterations to our epigenomes. Epigenetic effects include DNA methylation,

RNA-interference (RNAi), histone modifications, and chromatin remodeling — all of which contribute overall to the *genetic architecture* of the individual. Moreover, among the ~70 trillion cells in the human body, there are ~210 distinct human cell types. Unlike DNA sequence, which is identical in all cells of the body, humans probably have ~210 distinct epigenomes, all of which are theoretically capable of changing, as a function of time and environmental stimuli. Environmental perturbations also contribute to transgenerational heritable phenotypes, presumed to be manifestations of epigenetic effects, but the mechanisms are poorly understood [reviewed in (Nebert, Zhang, & Vesell, 2013) and (van Otterdijk & Michels, 2016)].

With recent advances in high-throughput technologies, including WGS of a single cell, it will soon be possible to analyze many features of the epigenome in depth — allowing for the establishment of complete epigenomic profiles for basically every cell-type of interest (Martens, Stunnenberg, & Logie, 2011).

6.2. Challenges of clinical utility in risk prediction vs pharmacoeconomics

A recent initial assessment of the benefits of implementing PGx into the medical management of patients in a long-term-care facility (Saldivar et al., 2016) emphasized that healthcare costs associated with prescription drugs are enormous, particularly in patients with polypharmacy (taking more than five prescription medications). Applying PGx-guided prescription recommendations across this patient population — resulted in elimination and/or replacement of one to three drugs for 50% of the polypharmacy patient population tested — with an estimated annual savings of US\$621 per patient. Initial assessment of this study shows that there is a clear opportunity for concrete healthcare savings, based solely on careful prescription drug management when incorporating PGx testing.

Interestingly, the positive predictive value of *HLA-B*38:02:01* in anticipating carbimazole/methimazole-induced agranulocytosis is 0.07 and the negative predictive value is 0.999. Thus, it was extrapolated that ~211 cases need to be screened in order to prevent one case; screening for the risk allele will therefore be useful in preventing agranulocytosis — but only in populations in which frequency of the risk allele is high (Cheung et al., 2016). This is an excellent example of a potent effect of a rare-variant allele with, unfortunately, low penetrance.

A recent review (Abul-Husn, Owusu, Sanderson, Gottesman, & Scott, 2014) summarized the current state of implementing genetic testing for personalized medicine, with emphasis on clinical pharmacogenetic testing. This included a table listing company-based PGx tests approved by the US FDA for “*in vitro* diagnostic” use.

6.3. The field of “precision oncology”

Precision oncology (Relling & Evans, 2015), as discussed in Section 5, is not the same as PGx prediction of risk in therapeutic treatment of the whole patient. All cancers arise as a result of DNA mutations, which confer growth advantage of tumor cells. These specific mutations may be targets for treatment and might easily influence therapeutic responses significantly. Specific malignant cells — today, in fact, even a single cell — can be

genotyped and the most appropriate drug can then be selected, because it is quite certain to be tailored to inhibit growth successfully of that patient's tumor.

6.4. Consideration of cost-benefit ratios

ADRs are an important public health concern, accounting for ~5% of all hospital admissions and two-thirds of all complications occurring shortly after hospital discharge. There are often long delays between when a drug is approved, and when serious ADRs are identified. Recent and ongoing advances in drug safety surveillance include: [a] establishment of government-sponsored networks of population databases, [b] use of data-mining approaches, and [c] formal integration of diverse sources of drug safety information. These advances promise to decrease the delay-time in identifying drug-related risks, or in providing reassurance about the absence of such risks (Hennessy & Strom, 2015).

7. Conclusions and future perspectives

As we have indicated throughout this review, the promise of personalized medicine is far from being realized — especially in genetic prediction of responses to commonly used drugs, and in the treatment of extremely complex phenotypes such as T2D, hypertension and major depression disorder. Lack of evidence has been widely accepted as one of the most significant challenges of PGx implementation into clinical practice (Khoury et al., 2008; Swen et al., 2007). To overcome this barrier, a continuum of “translational research” agenda has been proposed (Khoury et al., 2007) to accelerate translation of genomic discoveries into healthcare benefits. However, as we have emphasized throughout this review, many valid but small-effect PGx associations, even in combinations, are not likely to improve genetic prediction of drug response in the individual. Among other major challenges of PGx implementation into clinical practice — is the lack of processes required to interpret genotype information, and to translate genetic information into clinically doing something (Relling & Klein, 2011).

Genomic studies with large sample-sizes, using GWAS or NGS technologies, will continue to identify genetic variants associated with countless drug responses, and this information might increase our understanding of the biology of a drug's mode-of-action or mechanism-of-action. However, the multifactorial nature of complex PGx traits may ultimately limit the possibilities for accurate prediction (Janssens & van Duijn, 2008; Nebert & Zhang, 2012). In view of this realization, it might be important to accept that there are “translational boundaries” in PGx prediction of drug response — usually restricted to simple (Mendelian or predominantly oligogenic) PGx traits influenced by genetic variants having large-effect.

As WGS technologies become more readily available and cheaper in clinical settings, *the cost-benefit balance* will gradually become more favorable to clinical implementation of genomic tests. Coupled with the development of more accurately annotated and clinically relevant resources — such as PharmGKB (Whirl-Carrillo et al., 2012), CPIC (The Clinical Pharmacogenetics Implementation Consortium) (Relling & Klein, 2011), as well as integration of genomic data into healthcare record systems — computer-aided interpretations of PGx test results will become more feasible and should be helpful in clinical decision-making (Relling & Evans, 2015). In the context of clinical implementation of PGx, a

preemptive approach (Dunnenberger et al., 2015) is particularly attractive, because the genotype does not change over time, *i.e.* a one-time genome-wide genotyping result can almost always be used to guide any prescribing decision throughout one's entire life. One exception to this generalization would be the patient's epigenomes of various cell-types, which are constantly changing. The contribution of epigenetics to genetic risk prediction is expected to become a very active area of investigation in the near future.

Although the overall health benefit has yet to be established, several studies (Ji et al., 2016; Van Driest et al., 2014) have already shown that the vast majority of patients carry at least one actionable PGx variant that could potentially be used to improve prescribed drug treatment in the clinic. Despite these promising aspects of PGx, however, we must recognize that — as we have attempted to convey throughout this review — genetic risk prediction of drug response works only for a few drugs, and it is usually based on the genotype of a few genetic variants having large-effect contributions to the phenotype (efficacy effect, adverse effect, and toxic effect). Hence, genomic prediction of complex PGx responses, based on the individual patient's whole-genome readout, is most likely unattainable in the foreseeable future.

Acknowledgments

We thank our colleagues for valuable discussions. This work was funded in part by NIH grant P30 ES006096 (D.W.N.) and U01 HG008666 (G.Z.).

Abbreviations

ADR	adverse drug reaction
GRR	genotypic relative risk
GWAS	genome-wide association study or studies
LD	linkage disequilibrium
MAF	minor allele frequency
OR	odds ratio
PD genes	those encoding proteins involved in pharmacodynamics
PGx	pharmacogenetics or pharmacogenomics
PK genes	those encoding proteins involved in pharmacokinetics
SNP	single-nucleotide polymorphism or variant
WES	whole-exome sequencing
WGS	whole-genome sequencing

References

- Aboraya A, Rankin E, France C, El-Missiry A, John C. Reliability of psychiatric diagnosis revisited: The clinician's guide to improve reliability of psychiatric diagnosis. *Psychiatry* (Edmont). 2006; 3:41–50. [PubMed: 21103149]
- Abul-Husn NS, Owusu OA, Sanderson SC, Gottesman O, Scott SA. Implementation and utilization of genetic testing in personalized medicine. *Pharmacogenomics and Personalized Medicine*. 2014; 7:227–240. [PubMed: 25206309]
- Balwani M, Burrow TA, Charrow J, Goker-Alpan O, Kaplan P, Kishnani PS, Weinreb N. Recommendations for the use of eliglustat in the treatment of adults with Gaucher disease type 1 in the United States. *Molecular Genetics and Metabolism*. 2016; 117:95–103. [PubMed: 26387627]
- Belmatoug N, Di Rocco M, Fraga C, Giraldo P, Hughes D, Lukina E, Cox TM. Management and monitoring recommendations for the use of eliglustat in adults with type 1 Gaucher disease in Europe. *European Journal of Internal Medicine*. 2016
- Bertier G, Hetu M, Joly Y. Unsolved challenges of clinical whole-exome sequencing: A systematic literature review of end-users' views. *BMC Medical Genomics*. 2016; 9:52. [PubMed: 27514372]
- Bertilsson L, Dahl ML, Dalen P, Al-Shurbaji A. Molecular genetics of *CYP2D6*: Clinical relevance with focus on psychotropic drugs. *British Journal of Clinical Pharmacology*. 2002; 53:111–122. [PubMed: 11851634]
- Bhattacharjee S, Wang Z, Ciampa J, Kraft P, Chanock S, Yu K, Chatterjee N. Using principal components of genetic variation for robust and powerful detection of gene-gene interactions in case-control and case-only studies. *American Journal of Human Genetics*. 2010; 86:331–342. [PubMed: 20206333]
- Biernacka JM, Sangkuhl K, Jenkins G, Whaley RM, Barman P, Batzler A, Weinshilboum R. The International SSRI Pharmacogenomics Consortium (ISPC): A genome-wide association study of antidepressant treatment response. *Translational Psychiatry*. 2015; 5:e553. [PubMed: 25897834]
- Blum M, Grant DM, McBride W, Heim M, Meyer UA. Human arylamine N-acetyltransferase genes: Isolation, chromosomal localization, and functional expression. *DNA and Cell Biology*. 1990; 9:193–203. [PubMed: 2340091]
- Bodmer W, Bonilla C. Common and rare variants in multifactorial susceptibility to common diseases. *Nature Genetics*. 2008; 40:695–701. [PubMed: 18509313]
- Brunton, L., Chabner, B., Knollman, B. Goodman & Gilman's the pharmacological basis of therapeutics. Printed in China: McGraw-Hill Companies, Inc.; 2011.
- Caldwell MD, Awad T, Johnson JA, Gage BF, Falkowski M, Gardina P, Burmester JK. CYP4F2 genetic variant alters required warfarin dose. *Blood*. 2008; 111:4106–4112. [PubMed: 18250228]
- Carrillo JA, Dahl ML, Svensson JO, Alm C, Rodriguez I, Bertilsson L. Disposition of fluvoxamine in humans is determined by polymorphic CYP2D6 and also by CYP1A2 activity. *Clinical Pharmacology and Therapeutics*. 1996; 60:183–190. [PubMed: 8823236]
- Cha PC, Mushiroda T, Takahashi A, Kubo M, Minami S, Kamatani N, Nakamura Y. Genome-wide association study identifies genetic determinants of warfarin responsiveness for Japanese. *Human Molecular Genetics*. 2010; 19:4735–4744. [PubMed: 20833655]
- Chasman DI, Giulianini F, MacFadyen J, Barratt BJ, Nyberg F, Ridker PM. Genetic determinants of statin-induced low-density lipoprotein cholesterol reduction. Justification for Use of statins in Prevention: An Intervention Trial Evaluating Rosuvastatin (JUPITER) trial. *Circulation. Cardiovascular Genetics*. 2012; 5:257–264. [PubMed: 22331829]
- Cheung CL, Sing CW, Tang CS, Cheng VK, Pirmohamed M, Choi CH, Kung A. *HLA-B*38:02:01* predicts carbimazole/methimazole-induced agranulocytosis. *Clinical Pharmacology and Therapeutics*. 2016; 99:555–561. [PubMed: 26599303]
- Chhibber A, Kroetz DL, Tantisira KG, McGeachie M, Cheng C, Plenge R, Pendergrass SA. Genomic architecture of pharmacological efficacy and adverse events. *Pharmacogenomics*. 2014; 15:2025–2048. [PubMed: 25521360]
- Chittani M, Zaninello R, Lanzani C, Frau F, Ortu MF, Salvi E, Barlassina C. TET2 and CSMD1 genes affect SBP response to hydrochlorothiazide in never-treated essential hypertensives. *Journal of Hypertension*. 2015; 33:1301–1309. [PubMed: 25695618]

- Cirulli ET, Goldstein DB. Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nature Reviews Genetics*. 2010; 11:415–425.
- Conjeevaram HS, Fried MW, Jeffers LJ, Terrault NA, Wiley-Lucas TE, Afdhal N, Howell CD. Peginterferon and ribavirin treatment in African-American and Caucasian-American patients with hepatitis-C genotype 1. *Gastroenterology*. 2006; 131:470–477. [PubMed: 16890601]
- Connolly S, Heron EA. Review of statistical methodologies for the detection of parent-of-origin effects in family trio genome-wide association data with binary disease traits. *Briefings in Bioinformatics*. 2015; 16:429–448. [PubMed: 24903222]
- Cooper GM, Johnson JA, Langaee TY, Feng H, Stanaway IB, Schwarz UI, Rieder MJ. Genome-wide scan for common genetic variants with a large influence on warfarin maintenance dose. *Blood*. 2008; 112:1022–1027. [PubMed: 18535201]
- Cooper-DeHoff RM, Johnson JA. Hypertension pharmacogenomics: In search of personalized treatment approaches. *Nature Reviews. Nephrology*. 2016; 12:110–122. [PubMed: 26592190]
- Cotton RG. Heterogeneity of phenylketonuria at the clinical, protein and DNA levels. *Journal of Inherited Metabolic Disease*. 1990; 13:739–750. [PubMed: 2246858]
- Daly AK. Pharmacogenomics of anticoagulants: Steps toward personal dosage. *Genome Medicine*. 2009; 1:10. [PubMed: 19348697]
- Daly AK. Genome-wide association studies in pharmacogenomics. *Nature Reviews. Genetics*. 2010; 11:241–246.
- Daly AK. Using genome-wide association studies to identify genes important in serious adverse drug reactions. *Annual Review of Pharmacology and Toxicology*. 2012; 52:21–35.
- Daly AK. Pharmacogenomics of adverse drug reactions. *Genome Medicine*. 2013; 5:5. [PubMed: 23360680]
- Daly AK. Human leukocyte antigen (HLA) pharmacogenomic tests: Potential and pitfalls. *Current Drug Metabolism*. 2014; 15:196–201. [PubMed: 24694233]
- Daly AK, Donaldson PT, Bhatnagar P, Shen Y, Pe'er I, Floratos A, Day CP. *HLA-B*57:01* genotype is a major determinant of drug-induced liver injury due to flucloxacillin. *Nature Genetics*. 2009; 41:816–819. [PubMed: 19483685]
- D'Andrea G, D'Ambrosio RL, Di PP, Chetta M, Santacroce R, Brancaccio V, Margaglione M. Polymorphism in the *VKORC1* gene associated with inter-individual variability in the dose-anticoagulant effect of warfarin. *Blood*. 2005; 105:645–649. [PubMed: 15358623]
- Deenen MJ, Meulendijks D, Cats A, Sechterberger MK, Severens JL, Boot H, Schellens JH. Upfront genotyping of DPYD*2A to individualize fluoropyrimidine therapy: A safety and cost analysis. *Journal of Clinical Oncology*. 2016; 34:227–234. [PubMed: 26573078]
- Deshmukh HA, Colhoun HM, Johnson T, McKeigue PM, Betteridge DJ, Durrington PN, Hitman GA. Genome-wide association study of genetic determinants of LDL-c response to atorvastatin therapy: Importance of Lp(a). *Journal of Lipid Research*. 2012; 53:1000–1011. [PubMed: 22368281]
- Dunnenberger HM, Crews KR, Hoffman JM, Caudle KE, Broeckel U, Howard SC, Relling MV. Preemptive clinical pharmacogenetics implementation: Current programs in five US medical centers. *Annual Review of Pharmacology and Toxicology*. 2015; 55:89–106.
- Edwards IR, Aronson JK. Adverse drug reactions: Definitions, diagnosis, and management. *Lancet*. 2000; 356:1255–1259. [PubMed: 11072960]
- Eichelbaum M, Spannbrucker N, Steincke B, Dengler HJ. Defective *N*-oxidation of sparteine in man: A new pharmacogenetic defect. *European Journal of Clinical Pharmacology*. 1979; 16:183–187. [PubMed: 499318]
- Eichelbaum M, Ingelman-Sundberg M, Evans WE. Pharmacogenomics and individualized drug therapy. *Annual Review of Medicine*. 2006; 57:119–137.
- English BC, Baum CE, Adelberg DE, Sissung TM, Kluetz PG, Dahut WL, Figg WD. A SNP in CYP2C8 is not associated with development of bisphosphonate-related osteonecrosis of the jaw in men with castrate-resistant prostate cancer. *Therapeutics and Clinical Risk Management*. 2010; 6:579–583. [PubMed: 21151627]
- Fernandez CA, Smith C, Yang W, Mullighan CG, Qu C, Larsen E, Relling MV. Genome-wide analysis links NFATC2 with asparaginase hypersensitivity. *Blood*. 2015; 126:69–75. [PubMed: 25987655]

- Fisher RA. The correlation between relatives on the supposition of mendelian inheritance. Transactions of the Royal Society of Edinburgh. 1918; 52:399–433.
- Florez JC, Jablonski KA, Taylor A, Mather K, Horton E, White NH. Diabetes Prevention Program Research, G. The C allele of ATM rs11212617 does not associate with metformin response in the Diabetes Prevention Program. Diabetes Care. 2012; 35:1864–1867. [PubMed: 22751958]
- Frau F, Zaninello R, Salvi E, Ortu MF, Braga D, Velayutham D, Glorioso N. Genome-wide association study identifies CAMKID variants involved in blood pressure response to losartan: The SOPHIA study. Pharmacogenomics. 2014; 15:1643–1652. [PubMed: 25410890]
- Freedman R, Lewis DA, Michels R, Pine DS, Schultz SK, Tamminga CA, Yager J. The initial field trials of DSM-5: New blooms and old thorns. The American Journal of Psychiatry. 2013; 170:1–5. [PubMed: 23288382]
- Fung E, Patsopoulos NA, Belknap SM, O'Rourke DJ, Robb JF, Anderson JL, Moore JH. Effect of genetic variants, especially *CYP2C9* and *VKORC1*, on the pharmacology of warfarin. Seminars in Thrombosis and Hemostasis. 2012; 38:893–904. [PubMed: 23041981]
- Garriock HA, Kraft JB, Shyn SI, Peters EJ, Yokoyama JS, Jenkins GD, Hamilton SP. A genomewide association study of citalopram response in major depressive disorder. Biological Psychiatry. 2010; 67:133–138. [PubMed: 19846067]
- Ge D, Fellay J, Thompson AJ, Simon JS, Shianna KV, Urban TJ, Goldstein DB. Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. Nature. 2009; 461:399–401. [PubMed: 19684573]
- Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, McVean GA. Genomes Project, C. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012; 491:56–65. [PubMed: 23128226]
- Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Abecasis GR. Genomes Project, C. A global reference for human genetic variation. Nature. 2015; 526:68–74. [PubMed: 26432245]
- Gibson G. Hints of hidden heritability in GWAS. Nature Genetics. 2010; 42:558–560. [PubMed: 20581876]
- Gibson G. Rare and common variants: Twenty arguments. Nature Reviews. Genetics. 2011; 13:135–145.
- Zhou K, Bellenguez C, Spencer CC, Pearson ER. GoDarts, Group, U. D. P. S., Wellcome Trust Case Control, C. Common variants near ATM are associated with glycemic response to metformin in type-2 diabetes. Nature Genetics. 2011; 43:117–120. [PubMed: 21186350]
- Goldstein DB. Common genetic variation and human traits. The New England Journal of Medicine. 2009; 360:1696–1698. [PubMed: 19369660]
- Goldstein JA, de Morais SM. Biochemistry and molecular biology of the human CYP2C subfamily. Pharmacogenetics. 1994; 4:285–299. [PubMed: 7704034]
- Gong Y, McDonough CW, Beitelshes AL, El RN, Hiltunen TP, O'Connell JR, Johnson JA. PTPRD gene associated with blood pressure response to atenolol and resistant hypertension. Journal of Hypertension. 2015; 33:2278–2285. [PubMed: 26425837]
- Gong Y, Wang Z, Beitelshes AL, McDonough CW, Langaee TY, Hall K, Johnson JA. Pharmacogenomic genome-wide meta-analysis of blood pressure response to β -blockers in hypertensive African Americans. Hypertension. 2016; 67:556–563. [PubMed: 26729753]
- Gonzalez FJ, Skoda RC, Kimura S, Umeno M, Zanger UM, Nebert DW, Meyer UA. Characterization of the common genetic defect in humans deficient in debrisoquine metabolism. Nature. 1988; 331:442–446. [PubMed: 3123997]
- Grant DM, Goodfellow GH, Sugamori K, Durette K. Pharmacogenetics of the human arylamine *N*-acetyltransferases. Pharmacology. 2000; 61:204–211. [PubMed: 10971207]
- Grove ML, Yu B, Cochran BJ, Haritunians T, Bis JC, Taylor KD, Boerwinkle E. Best practices and joint calling of the HumanExome BeadChip: The CHARGE Consortium. PloS One. 2013; 8:e68095. [PubMed: 23874508]
- Haque MM, Nilsson EE, Holder LB, Skinner MK. Genomic clustering of differential DNA methylated regions (epi-mutations) associated with epigenetic transgenerational inheritance of disease and phenotypic variation. BMC Genomics. 2016; 17:418. [PubMed: 27245821]

- Hellens RP, Moreau C, Lin-Wang K, Schwinn KE, Thomson SJ, Fiers MW, Ellis TH. Identification of Mendel's white flower character. *PLoS One*. 2010; 5:e13230. [PubMed: 20949001]
- Hennessy S, Strom BL. Improving postapproval drug safety surveillance: Getting better information sooner. *Annual Review of Pharmacology and Toxicology*. 2015; 55:75–87.
- Hopewell JC, Parish S, Offer A, Link E, Clarke R, Lathrop M, Collins R. Impact of common genetic variation in response to simvastatin therapy among 18,705 participants in the Heart Protection Study. *European Heart Journal*. 2013; 34:982–992. [PubMed: 23100282]
- Iadonato SP, Katze MG. Genomics: Hepatitis C virus gets personal. *Nature*. 2009; 461:357–358. [PubMed: 19759611]
- Klein TE, Altman RB, Eriksson N, Gage BF, Kimmel SE, Johnson JA. International Warfarin Pharmacogenetics, C. Estimation of the warfarin dose with clinical and pharmacogenetic data. *The New England Journal of Medicine*. 2009; 360:753–764. [PubMed: 19228618]
- Ioannidis JP, Khoury MJ. Are randomized trials obsolete or more important than ever in the genomic era? *Genome Medicine*. 2013; 5:32. [PubMed: 23673134]
- Ising M, Lucae S, Binder EB, Bettecken T, Uhr M, Ripke S, Muller-Myhsok B. A genomewide association study points to multiple loci that predict anti-depressant drug treatment outcome in depression. *Archives of General Psychiatry*. 2009; 66:966–975. [PubMed: 19736353]
- Jackson CM, Suttie JW. Recent developments in understanding the mechanism of vitamin K and vitamin K-antagonist drug action and consequences of vitamin K action in blood coagulation. *Progress in Hematology*. 1977; 10:333–359. [PubMed: 337364]
- Janssens AC, van Duijn CM. Genome-based prediction of common diseases: Advances and prospects. *Human Molecular Genetics*. 2008; 17:R166–R173. [PubMed: 18852206]
- Jensen DM, Pol S. IL28B genetic polymorphism testing in the era of direct-acting antiviral therapy for chronic hepatitis C: Ten years too late? *Liver International*. 2012; 32(Suppl. 1):74–78. [PubMed: 22212576]
- Ji Y, Schaid DJ, Desta Z, Kubo M, Batzler AJ, Snyder K, Weinshilboum RM. Citalopram and escitalopram plasma drug and metabolite concentrations: Genome-wide associations. *British Journal of Clinical Pharmacology*. 2014; 78:373–383. [PubMed: 24528284]
- Ji Y, Skierka JM, Blommel JH, Moore BE, VanCuyk DL, Bruflat JK, Black JL III. Preemptive pharmacogenomic testing for precision medicine: A comprehensive analysis of five actionable pharmacogenomic genes using next-generation DNA sequencing and a customized *CYP2D6* genotyping cascade. *The Journal of Molecular Diagnostics*. 2016; 18:438–445. [PubMed: 26947514]
- Johnson JA, Gong L, Whirl-Carrillo M, Gage BF, Scott SA, Stein CM. Clinical Pharmacogenetics Implementation, C. Clinical Pharmacogenetics Implementation Consortium Guidelines for *CYP2C9* and *VKORC1* genotypes and warfarin dosing. *Clinical Pharmacology and Therapeutics*. 2011; 90:625–629. [PubMed: 21900891]
- Johnson JA, Roden DM, Lesko LJ, Ashley E, Klein TE, Shuldiner AR. Clopidogrel: A case for indication-specific pharmacogenetics. *Clinical Pharmacology and Therapeutics*. 2012; 91:774–776. [PubMed: 22513313]
- Karnes JH, Cronin RM, Rollin J, Teumer A, Pouplard C, Shaffer CM, Roden DM. A genome-wide association study of heparin-induced thrombocytopenia using an electronic medical record. *Thrombosis and Haemostasis*. 2015; 113:772–781. [PubMed: 25503805]
- Katz J, Gong Y, Salmasinia D, Hou W, Burkley B, Ferreira P, Moreb JS. Genetic polymorphisms and other risk factors associated with bisphosphonate-induced osteonecrosis of the jaw. *International Journal of Oral and Maxillofacial Surgery*. 2011; 40:605–611. [PubMed: 21396799]
- Khoury MJ, Gwinn M, Yoon PW, Dowling N, Moore CA, Bradley L. The continuum of translation research in genomic medicine: How can we accelerate the appropriate integration of human genome discoveries into health care and disease prevention? *Genetics in Medicine*. 2007; 9:665–674. [PubMed: 18073579]
- Khoury MJ, Berg A, Coates R, Evans J, Teutsch SM, Bradley LA. The evidence dilemma in genomic medicine. *Health Affairs (Millwood)*. 2008; 27:1600–1611.

- Kimmel SE, French B, Kasner SE, Johnson JA, Anderson JL, Gage BF, Investigators C. A pharmacogenetic versus a clinical algorithm for warfarin dosing. *The New England Journal of Medicine*. 2013; 369:2283–2293. [PubMed: 24251361]
- Klein RJ, Zeiss C, Chew EY, Tsai JY, Sackler RS, Haynes C, Hoh J. Complement factor H polymorphism in age-related macular degeneration. *Science*. 2005; 308:385–389. [PubMed: 15761122]
- Kohane IS, Masys DR, Altman RB. The incidentalome: A threat to genomic medicine. *JAMA*. 2006; 296:212–215. [PubMed: 16835427]
- König J, Seithel A, Gradhand U, Fromm MF. Pharmacogenomics of human OATP transporters. *Naunyn-Schmiedeberg's Archives of Pharmacology*. 2006; 372:432–443.
- Kweekel D, Guchelaar HJ, Gelderblom H. Clinical and pharmacogenetic factors associated with irinotecan toxicity. *Cancer Treatment Reviews*. 2008; 34:656–669. [PubMed: 18558463]
- Lander ES. Initial impact of sequencing of the human genome. *Nature*. 2011; 470:187–197. [PubMed: 21307931]
- Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F, Hirschhorn JN. Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature*. 2010; 467:832–838. [PubMed: 20881960]
- Leusink M, Onland-Moret NC, de Bakker PI, de BA, AH M-vdZ. Seventeen years of statin pharmacogenetics: A systematic review. *Pharmacogenomics*. 2016; 17:163–180. [PubMed: 26670324]
- Liapakis A, Jesudian AB. Is there clinical utility to IL28B genotype testing in the treatment of chronic hepatitis C virus infection? *Pharmacogenomics*. 2012; 13:1317–1319. [PubMed: 22966879]
- Limdi NA, Arnett DK, Goldstein JA, Beasley TM, McGwin G, Adler BK, Acton RT. Influence of CYP2C9 and VKORC1 on warfarin dose, anticoagulation attainment and maintenance among European-Americans and African-Americans. *Pharmacogenomics*. 2008; 9:511–526. [PubMed: 18466099]
- Limdi NA, Wadelius M, Cavallari L, Eriksson N, Crawford DC, Lee MT. International Warfarin Pharmacogenetics, C. Warfarin pharmacogenetics: A single *VKORC1* polymorphism is predictive of dose across three racial groups. *Blood*. 2010; 115:3827–3834. [PubMed: 20203262]
- Linares OA, Fudin J, Schiesser WE, Daly Linares AL, Boston RC. CYP2D6 phenotype-specific codeine population pharmacokinetics. *Journal of Pain & Palliative Care Pharmacotherapy*. 2015; 29:4–15. [PubMed: 25562725]
- Link E, Parish S, Armitage J, Bowman L, Heath S, Matsuda F, Collins R. *SLCO1B1* variants and statin-induced myopathy—A genome-wide study. *The New England Journal of Medicine*. 2008; 359:789–799. [PubMed: 18650507]
- Liu C, Yang W, Pei D, Cheng C, Smith C, Landier W, Relling MV. A genome-wide approach validates that thiopurine methyltransferase activity is a mono-genic pharmacogenomic trait. *Clinical Pharmacology and Therapeutics*. 2016
- Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, Speliotes EK. Genetic studies of body mass index yield new insights for obesity biology. *Nature*. 2015; 518:197–206. [PubMed: 25673413]
- Lu AY, Wang RW, Lin JH. Cytochrome P450 in vitro reaction phenotyping: A re-evaluation of approaches used for P450 isoform identification. *Drug Metabolism and Disposition*. 2003; 31:345–350. [PubMed: 12642457]
- Lunenburg CA, Henricks LM, Guchelaar HJ, Swen JJ, Deenen MJ, Schellens JH, Gelderblom H. Prospective DPYD genotyping to reduce the risk of fluoropyrimidine-induced severe toxicity: Ready for prime time. *European Journal of Cancer*. 2016; 54:40–48. [PubMed: 26716401]
- Lupoli S, Salvi E, Barcella M, Barlassina C. Pharmacogenomics considerations in the control of hypertension. *Pharmacogenomics*. 2015; 16:1951–1964. [PubMed: 26555875]
- Mahgoub A, Idle JR, Dring LG, Lancaster R, Smith RL. Polymorphic hydroxylation of debrisoquine in man. *Lancet*. 1977; 2:584–586. [PubMed: 71400]
- Mallal S, Nolan D, Witt C, Masel G, Martin AM, Moore C, Christiansen FT. Association between presence of *HLA-B*57:01*, *HLA-DR7*, and *HLA-DQ3* and hypersensitivity to HIV-1 reverse-

- transcriptase inhibitor abacavir. *The New England Journal of Medicine*. 2008a; 358:568–579. [PubMed: 18256392]
- Mallal S, Phillips E, Carosi G, Molina J-M, Workman C, Tomažič J, Team P-S. *HLA-B*57:01* screening for hypersensitivity to abacavir. *Virology Journal*. 2008b; 5:88. [PubMed: 18671869]
- Manolio TA. Bringing genome-wide association findings into clinical use. *Nature Reviews. Genetics*. 2013; 14:549–558.
- Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, Visscher PM. Finding the missing heritability of complex diseases. *Nature*. 2009; 461:747–753. [PubMed: 19812666]
- Marouli E, Graff M, Medina-Gomez C, Lo KS, Wood AR, Kjaer TR, Lettre G. Rare and low-frequency coding variants alter human adult height. *Nature*. 2017; 542:186–190. [PubMed: 28146470]
- Martens JH, Stunnenberg HG, Logie C. The decade of the epigenomes? *Genes & Cancer*. 2011; 2:680–687. [PubMed: 21941622]
- Marvit J, DiLella AG, Brayton K, Ledley FD, Robson KJ, Woo SL. GT to AT transition at a splice donor site causes skipping of the preceding exon in phenylke-tonuria. *Nucleic Acids Research*. 1987; 15:5613–5628. [PubMed: 3615198]
- McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JP, Hirschhorn JN. Genome-wide association studies for complex traits: Consensus, uncertainty and challenges. *Nature Reviews. Genetics*. 2008; 9:356–369.
- McCarthy S, Das S, Kretschmar W, Delaneau O, Wood AR, Teumer A, Haplotype Reference C. A reference panel of 64,976 haplotypes for genotype imputation. *Nature Genetics*. 2016; 48:1279–1283. [PubMed: 27548312]
- McCormack M, Alfirevic A, Bourgeois S, Farrell JJ, Kasperaviciute D, Carrington M, Pirmohamed M. *HLA-A*31:01* and carbamazepine-induced hypersensitivity reactions in Europeans. *The New England Journal of Medicine*. 2011; 364:1134–1143. [PubMed: 21428769]
- Menni C. Blood pressure pharmacogenomics: Gazing into a misty crystal ball. *Journal of Hypertension*. 2015; 33:1142–1143. [PubMed: 25923729]
- Mizutani T. PM frequencies of major CYPs in Asians and Caucasians. *Drug Metabolism Reviews*. 2003; 35:99–106. [PubMed: 12959412]
- Momary KM, Shapiro NL, Viana MA, Nutescu EA, Helgason CM, Cavallari LH. Factors influencing warfarin dose requirements in African-Americans. *Pharmacogenomics*. 2007; 8:1535–1544. [PubMed: 18034618]
- Morimoto K, Ueda S, Seki N, Igawa Y, Kameyama Y, Shimizu A, Chiba K. (OATP-C)*SLCO1B1*15* is associated with statin-induced myopathy in hypercholester-olemic patients. *Clinical Pharmacology & Therapeutics*. 2005; 77:P21–P21.
- Motsinger-Reif AA, Jorgenson E, Relling MV, Kroetz DL, Weinshilboum R, Cox NJ, Roden DM. Genome-wide association studies in pharmacogenomics: Successes and lessons. *Pharmacogenetics and Genomics*. 2013; 23:383–394. [PubMed: 20639796]
- Motulsky AG. Pharmacogenetics. *Progress in Medical Genetics*. 1964; 23:49–74. [PubMed: 14110077]
- Munroe PB, Barnes MR, Caulfield MJ. Advances in blood pressure genomics. *Circulation Research*. 2013; 112:1365–1379. [PubMed: 23661711]
- Nebert DW. Pharmacogenetics and pharmacogenomics: Why is this relevant to the clinical geneticist? *Clinical Genetics*. 1999; 56:247–258. [PubMed: 10636440]
- Nebert DW, Dieter MZ. The evolution of drug metabolism. *Pharmacology*. 2000; 61:124–135. [PubMed: 10971198]
- Nebert DW, Zhang G. Personalized medicine: Temper expectations. *Science*. 2012; 337:910. [PubMed: 22923560]
- Nebert DW, Zhang G, Vesell ES. From human genetics and genomics to pharmacogenetics and pharmacogenomics: Past lessons, future directions. *Drug Metabolism Reviews*. 2008; 40:187–224. [PubMed: 18464043]
- Nebert DW, Zhang G, Vesell ES. Genetic risk prediction: Individualized variability in susceptibility to toxicants. *Annual Review of Pharmacology and Toxicology*. 2013; 53:355–375.

- Niemeijer MN, van den Berg ME, Eijgelsheim M, Rijnbeek PR, Stricker BH. Pharmacogenetics of drug-induced QT-interval prolongation: An update. *Drug Safety*. 2015; 38:855–867. [PubMed: 26108299]
- Niemi M. Transporter pharmacogenetics and statin toxicity. *Clinical Pharmacology and Therapeutics*. 2010; 87:130–133. [PubMed: 19890253]
- Ozaki K, Ohnishi Y, Iida A, Sekine A, Yamada R, Tsunoda T, Tanaka T. Functional SNPs in the lymphotoxin- α gene (LTA) that are associated with susceptibility to myocardial infarction. *Nature Genetics*. 2002; 32:650–654. [PubMed: 12426569]
- Parham LR, Briley LP, Li L, Shen J, Newcombe PJ, King KS, Spraggs CF. Comprehensive genome-wide evaluation of lapatinib-induced liver injury yields single genetic signal centered on known risk allele *HLA-DRB1*07:01*. *The Pharmacogenomics Journal*. 2016; 16:180–185. [PubMed: 25987243]
- Park JH, Wacholder S, Gail MH, Peters U, Jacobs KB, Chanock SJ, Chatterjee N. Estimation of effect-size distribution from genome-wide association studies and implications for future discoveries. *Nature Genetics*. 2010; 42:570–575. [PubMed: 20562874]
- Park JH, Gail MH, Weinberg CR, Carroll RJ, Chung CC, Wang Z, Chatterjee N. Distribution of allele frequencies and effect-sizes and their interrelationships for common genetic susceptibility variants. *Proceedings of the National Academy of Sciences of the United States of America*. 2011; 108:18026–18031. [PubMed: 22003128]
- Parra EJ, Botton MR, Perini JA, Krithika S, Bourgeois S, Johnson TA, Suarez-Kurtz G. Genome-wide association study of warfarin maintenance dose in a Brazilian sample. *Pharmacogenomics*. 2015; 16:1253–1263. [PubMed: 26265036]
- Perera MA, Cavallari LH, Limdi NA, Gamazon ER, Konkashbaev A, Daneshjou R, Johnson JA. Genetic variants associated with warfarin dose in African-American individuals: A genome-wide association study. *Lancet*. 2013; 382:790–796. [PubMed: 23755828]
- Pinto AM, Ariani F, Bianciardi L, Daga S, Renieri A. Exploiting the potential of next-generation sequencing in genomic medicine. *Expert Review of Molecular Diagnostics*. 2016; 16:1037–1047. [PubMed: 27574853]
- Pirmohamed M, Burnside G, Eriksson N, Jorgensen AL, Toh CH, Nicholson T, Group E-P. A randomized trial of genotype-guided dosing of warfarin. *The New England Journal of Medicine*. 2013; 369:2294–2303. [PubMed: 24251363]
- Pirmohamed M, Kamali F, Daly AK, Wadelius M. Oral anticoagulation: A critique of recent advances and controversies. *Trends in Pharmacological Sciences*. 2015; 36:153–163. [PubMed: 25698605]
- Postmus I, Trompet S, Deshmukh HA, Barnes MR, Li X, Warren HR, Caulfield MJ. Pharmacogenetic meta-analysis of genome-wide association studies of LDL-cholesterol response to statins. *Nature Communications*. 2014; 5:5068.
- Price Evans DA, Manley KA, McKusick VA. Genetic control of isoniazid metabolism in man. *British Medical Journal*. 1960; 2:485–491. [PubMed: 13820968]
- Pritchard JK. Are rare variants responsible for susceptibility to complex diseases? *American Journal of Human Genetics*. 2001; 69:124–137. [PubMed: 11404818]
- Ramsey LB, Johnson SG, Caudle KE, Haidar CE, Voora D, Wilke RA, Niemi M. The clinical pharmacogenetics implementation consortium guideline for *SLCO1B1* and simvastatin-induced myopathy: 2014 update. *Clinical Pharmacology and Therapeutics*. 2014; 96:423–428. [PubMed: 24918167]
- Relling MV, Evans WE. Pharmacogenomics in the clinic. *Nature*. 2015; 526:343–350. [PubMed: 26469045]
- Relling MV, Klein TE. CPIC: Clinical Pharmacogenetics Implementation Consortium of the Pharmacogenomics Research Network. *Clinical Pharmacology and Therapeutics*. 2011; 89:464–467. [PubMed: 21270786]
- Rioux PP. Clinical trials in pharmacogenetics and pharmacogenomics: Methods and applications. *American Journal of Health-System Pharmacy*. 2000; 57:887–898. (quiz 899–901). [PubMed: 10840530]

- Ripke S, Wray NR, Lewis CM, Hamilton SP, Weissman MM, Breen G, Sullivan PF. A mega-analysis of genome-wide association studies for major depressive disorder. *Molecular Psychiatry*. 2013; 18:497–511. [PubMed: 22472876]
- Risch N, Merikangas K. The future of genetic studies of complex human diseases. *Science*. 1996; 273:1516–1517. [PubMed: 8801636]
- Sackton TB, Hartl DL. Genotypic context and epistasis in individuals and populations. *Cell*. 2016; 166:279–287. [PubMed: 27419868]
- Saldivar JS, Taylor D, Sugarman EA, Cullors A, Garces JA, Oades K, Centeno J. Initial assessment of the benefits of implementing pharmacogenetics into the medical management of patients in a long-term care facility. *Pharmacogenomics and Personalized Medicine*. 2016; 9:1–6. [PubMed: 26855597]
- Sarasquete ME, Garcia-Sanz R, Marin L, Alcoceba M, Chillon MC, Balanzategui A, San Miguel JF. Bisphosphonate-related osteonecrosis of the jaw is associated with polymorphisms of cytochrome P450 CYP2C8 in multiple myeloma: A genome-wide single-nucleotide polymorphism analysis. *Blood*. 2008; 112:2709–2712. [PubMed: 18594024]
- Scott SA, Sangkuhl K, Stein CM, Hulot JS, Mega JL, Roden DM. Clinical Pharmacogenetics Implementation, C. Clinical Pharmacogenetics Implementation Consortium guidelines for CYP2C19 genotype and clopidogrel therapy: 2013 update. *Clinical Pharmacology and Therapeutics*. 2013; 94:317–323. [PubMed: 23698643]
- Singer JB, Lewitzky S, Leroy E, Yang F, Zhao X, Klickstein L, Paulding CA. A genome-wide study identifies *HLA* alleles associated with lumiracoxib-related liver injury. *Nature Genetics*. 2010; 42:711–714. [PubMed: 20639878]
- Slatkin M. Epigenetic inheritance and the missing heritability problem. *Genetics*. 2009; 182:845–850. [PubMed: 19416939]
- Snyder LH. Studies in human inheritance. IX. The inheritance of taste deficiency in man. *The Ohio Journal of Science*. 1932; 32:436–468.
- Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G, Jackson AU, Loos RJ. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nature Genetics*. 2010; 42:937–948. [PubMed: 20935630]
- Stranger BE, Stahl EA, Raj T. Progress and promise of genome-wide association studies for human complex trait genetics. *Genetics*. 2011; 187:367–383. [PubMed: 21115973]
- Such E, Cervera J, Terpos E, Bagan JV, Avaria A, Gomez I, de la Rubia J. *CYP2C8* gene polymorphism and bisphosphonate-related osteonecrosis of the jaw in patients with multiple myeloma. *Haematologica*. 2011; 96:1557–1559. [PubMed: 21685474]
- Suppiah V, Moldovan M, Ahlenstiel G, Berg T, Weltman M, Abate ML, George J. *IL28B* is associated with response to chronic hepatitis-C interferon- α and ribavirin therapy. *Nature Genetics*. 2009; 41:1100–1104. [PubMed: 19749758]
- Swen JJ, Huizinga TW, Gelderblom H, de Vries EG, Assendelft WJ, Kirchheiner J, Guchelaar HJ. Translating pharmacogenomics: Challenges on the road to the clinic. *PLoS Medicine*. 2007; 4:e209. [PubMed: 17696640]
- Takagi M, Uno H, Nishi R, Sugimoto M, Hasegawa S, Piao J, Mizutani S. ATM regulates adipocyte differentiation and contributes to glucose homeostasis. *Cell Reports*. 2015
- Takeuchi F, McGinnis R, Bourgeois S, Barnes C, Eriksson N, Soranzo N, Deloukas P. Genome-wide association study confirms *VKORC1*, *CYP2C9*, and *CYP4F2* as principal genetic determinants of warfarin dose. *PLoS Genetics*. 2009; 5:e1000433. [PubMed: 19300499]
- Tanaka Y, Nishida N, Sugiyama M, Kurosaki M, Matsuura K, Sakamoto N, Mizokami M. Genome-wide association of *IL28B* with response to pegylated interferon- α and ribavirin therapy for chronic hepatitis C. *Nature Genetics*. 2009; 41:1105–1109. [PubMed: 19749757]
- Teh LK, Bertilsson L. Pharmacogenomics of *CYP2D6*: Molecular genetics, in-terethnic differences and clinical importance. *Drug Metabolism and Pharmacokinetics*. 2012; 27:55–67. [PubMed: 22185816]
- Teichert M, Eijgelsheim M, Rivadeneira F, Uitterlinden AG, van Schaik RH, Hofman A, Stricker BH. A genome-wide association study of acenocoumarol maintenance dosage. *Human Molecular Genetics*. 2009; 18:3758–3768. [PubMed: 19578179]

- Thomas D. Gene-environment-wide association studies: Emerging approaches. *Nature Reviews. Genetics*. 2010; 11:259–272.
- Thomas DL, Thio CL, Martin MP, Qi Y, Ge D, O'Huigin C, Carrington M. Genetic variation in *IL28B* and spontaneous clearance of hepatitis C virus. *Nature*. 2009; 461:798–801. [PubMed: 19759533]
- Thompson AJ, McHutchison JG. Will IL28B polymorphism remain relevant in the era of direct-acting antiviral agents for hepatitis C virus? *Hepatology*. 2012; 56:373–381. [PubMed: 22511355]
- Turner ST, Bailey KR, Fridley BL, Chapman AB, Schwartz GL, Chai HS, Boerwinkle E. Genomic association analysis suggests chromosome 12 locus influencing anti-hypertensive response to thiazide diuretic. *Hypertension*. 2008; 52:359–365. [PubMed: 18591461]
- Turner ST, Bailey KR, Schwartz GL, Chapman AB, Chai HS, Boerwinkle E. Genomic association analysis identifies multiple loci influencing anti-hypertensive response to an angiotensin II receptor blocker. *Hypertension*. 2012; 59:1204–1211. [PubMed: 22566498]
- Turner ST, Boerwinkle E, O'Connell JR, Bailey KR, Gong Y, Chapman AB, Johnson JA. Genomic association analysis of common variants influencing anti-hypertensive response to hydrochlorothiazide. *Hypertension*. 2013; 62:391–397. [PubMed: 23753411]
- Uher R, Perroud N, Ng MY, Hauser J, Henigsberg N, Maier W, McGuffin P. Genome-wide pharmacogenetics of anti-depressant response in the GENDEP project. *The American Journal of Psychiatry*. 2010; 167:555–564. [PubMed: 20360315]
- Van Driest SL, Shi Y, Bowton EA, Schildcrout JS, Peterson JF, Pulley J, Roden DM. Clinically actionable genotypes among 10,000 patients with preemptive pharmacogenomic testing. *Clinical Pharmacology and Therapeutics*. 2014; 95:423–431. [PubMed: 24253661]
- van Otterdijk SD, Michels KB. Transgenerational epigenetic inheritance in mammals: How good is the evidence? *The FASEB Journal*. 2016; 30:2457–2465. [PubMed: 27037350]
- van LN, Nijpels G, Becker ML, Deshmukh H, Zhou K, Stricker BH, Pearson ER. A gene variant near *ATM* is significantly associated with metformin treatment response in type-2 diabetes: Replication and meta-analysis of five cohorts. *Diabetologia*. 2012; 55:1971–1977. [PubMed: 22453232]
- Verhoef TI, Ragia G, de Boer A, Barallon R, Kolovou G, Kolovou V, Group E-P. A randomized trial of genotype-guided dosing of acenocoumarol and phenprocoumon. *The New England Journal of Medicine*. 2013; 369:2304–2312. [PubMed: 24251360]
- Voora D, Shah SH, Spasojevic I, Ali S, Reed CR, Salisbury BA, Ginsburg GS. The *SLCO1B1**5 genetic variant is associated with statin-induced side effects. *Journal of the American College of Cardiology*. 2009; 54:1609–1616. [PubMed: 19833260]
- Wadelius M, Pirmohamed M. Pharmacogenetics of warfarin: Current status and future challenges. *The Pharmacogenomics Journal*. 2007; 7:99–111. [PubMed: 16983400]
- Walraven JM, Zang Y, Trent JO, Hein DW. Structure/function evaluations of single-nucleotide polymorphisms in human *N*-acetyltransferase-2. *Current Drug Metabolism*. 2008; 9:471–486. [PubMed: 18680467]
- Weinshilboum RM, Sladek SL. Mercaptopurine pharmacogenetics: Mono-genic inheritance of erythrocyte thiopurine methyltransferase activity. *American Journal of Human Genetics*. 1980; 32:651–662. [PubMed: 7191632]
- Whirl-Carrillo M, McDonagh EM, Hebert JM, Gong L, Sangkuhl K, Thorn CF, Klein TE. Pharmacogenomics knowledge for personalized medicine. *Clinical Pharmacology and Therapeutics*. 2012; 92:414–417. [PubMed: 22992668]
- Wilke RA, Lin DW, Roden DM, Watkins PB, Flockhart D, Zineh I, Krauss RM. Identifying genetic risk factors for serious adverse drug reactions: Current progress and challenges. *Nature Reviews. Drug Discovery*. 2007; 6:904–916. [PubMed: 17971785]
- Woo SL, Lidsky AS, Guttler F, Chandra T, Robson KJ. Cloned human phe-nylalanine hydroxylase gene allows prenatal diagnosis and carrier detection of classical phenylketonuria. *Nature*. 1983; 306:151–155. [PubMed: 6316140]
- Wood AR, Esko T, Yang J, Vedantam S, Pers TH, Gustafsson S, Frayling TM. Defining the role of common variation in the genomic and biological architecture of adult human height. *Nature Genetics*. 2014; 46:1173–1186. [PubMed: 25282103]

- Wray NR, Yang J, Hayes BJ, Price AL, Goddard ME, Visscher PM. Pitfalls of predicting complex traits from SNPs. *Nature Reviews. Genetics*. 2013; 14:507–515.
- Xu H, Robinson GW, Huang J, Lim JY, Zhang H, Bass JK, Yang JJ. Common variants in *ACYP2* influence susceptibility to cisplatin-induced hearing loss. *Nature Genetics*. 2015; 47:263–266. [PubMed: 25665007]
- Yang Y, Algaazy KM. Warfarin-induced skin necrosis in a patient with a mutation of the prothrombin gene. *The New England Journal of Medicine*. 1999; 340:735.
- Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, Visscher PM. Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics*. 2010; 42:565–569. [PubMed: 20562875]
- Zhang G. Genetic architecture of complex human traits: What have we learned from genome-wide association studies? *Current Genetic Medicine Reports*. 2015; 3:143–150.
- Zhang G, Nebert DW, Chakraborty R, Jin L. Statistical power of association using the extreme discordant phenotype design. *Pharmacogenetics and Genomics*. 2006; 16:401–413. [PubMed: 16708049]
- Zhou K, Pearson ER. Insights from genome-wide association studies of drug response. *Annual Review of Pharmacology and Toxicology*. 2013; 53:299–310.
- Zhou K, Yee SW, Seiser EL, van LN, Tavendale R, Bennett AJ, Pearson ER. Variation in the glucose transporter gene *SLC2A2* is associated with glycemic response to metformin. *Nature Genetics*. 2016; 48:1055–1059. [PubMed: 27500523]
- Zineh I, Pacanowski M, Woodcock J. Pharmacogenetics and coumarin dosing--Recalibrating expectations. *The New England Journal of Medicine*. 2013; 369:2273–2275. [PubMed: 24328463]

Fig. 1.

Phenotype distribution of different traits. A: a recessive Mendelian trait with two discrete phenotypes; B: a distinct codominant Mendelian trait with three discrete phenotypes; C: a quantitative trait — controlled predominantly by a large-effect gene, and undoubtedly additional modifiers showing continuous distribution, with three distinct modes; D: a quantitative trait influenced by numerous genetic and environmental factors, *i.e.* a polygenic trait that follows a normal distribution.

Fig. 2. Distribution of measurable CYP2D6 phenotypes. Four phenotypic groups can be distinguished: ultra-rapid (UM), extensive, sometimes called “efficient” (EM), intermediate (IM), and poor (PM) metabolizers [*modified from The Lancet* 2000; 356: 1667–71].

Fig. 3. Effect-size (measured by *variance explained*, R^2) that can be identified by GWAS as a function of different sample sizes [power = 80%, significance level $\alpha = 5 \times 10^{-6}$ (*green dashed line*) and 5×10^{-8} (*red solid line*)]. As shown by the figure, GWAS with ~1000 samples have appropriate power (>80%) to detect genetic variants that are likely to explain ~4% of *phenotypic variance*. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 4.

Performance of diagnostic test for a dichotomized outcome (evaluated by AUC, *area-under-curve*) is principally determined by percentage of *variance* in phenotype that can be explained by predictors. The three *curves* illustrate the AUC as a function of *phenotypic variance* explained by predictive variable(s), providing different prevalences of 0.01 (*red solid line*), 0.1 (*green dashed line*), and 0.50 (*blue dotted line*), of the dichotomized outcome. Given the same level of variance explained, greater discriminatory power (*i.e.* larger AUC) can be obtained for less frequent outcomes. Generally, in order to achieve good discriminative power (AUC >80%; *grey horizontal line*), the predictors should explain at least 20%–40% *phenotypic variance*; however, for most complex PGx traits, the identified genetic variants explain far less phenotypic variance than this amount. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 5.
Different classes of drugs having FDA PGx labels.

Table 1

List of selected PGx GWAS with genome-wide statistically significant findings ($P < 5\text{e-}8$).

Drug	Response	Study sample size		Associated genes and variants		Reference				
		Discovery	Replication	Genes	Variant type	Freq ^b	Effect	P-value	PMID	Year
<i>ADR (resembles Mendelian traits with incomplete penetrance)</i>										
Statin	Myopathy	85/90	20,000	<i>SLCO1B1</i>	<i>Non-synonymous</i>	0.15	OR = 4.5	4.00e-09	18650507	2008
Flucloxacillin	DILI ^a	51/282	23 cases	<i>HLA-B*57:01</i>	<i>HLA</i> haplotype	0.05	OR = 80.6	8.70e-33	19483685	2009
Lumiracoxib	DILI	41/176	98/405	<i>HLA-DRB1*15:01</i>	<i>HLA</i> haplotype	0.1	OR = 5.0	6.80e-25	20639878	2010
Carbamazepine	Hypersensitivity	22 + 43/3987	145	<i>HLA-DRB*15:02</i>	<i>HLA</i> haplotype	0.02-0.05	OR = 12.4	3.50e-08	21428769	2011
Heparin	Thrombo-cytopenia	67/884	94/178	<i>TDAG8</i>	Missense?	0.22	OR = 18.6 ^d	3.18e-09	25503805	2015
Anthracycline	Cardiotoxicity	280	96 + 80	<i>RARG</i>	Missense	0.25	OR = 4.7	5.90e-08	26237429	2015
Lapatinib	DILI	1192 (34/810)		<i>HLA-DRB1*07:01</i>	<i>HLA</i> haplotype	0.2	OR=14 ^d	7.8e-11	25987243	2016
Asparaginase	Hypersensitivity	3308(589)		<i>NEATC2</i>	Intronic	0.08	OR = 3.11	4.10e-08	25987655	2015
Cisplatin	Hearing loss	238	68	<i>ACYP2</i>	Intronic	0.07	HR = 4.5	3.90e-08	25665007	2015
<i>Dosage/efficacy (having major gene effect)</i>										
Warfarin	Dosage	181	374	<i>VKORC1</i> and <i>CYP2C9</i>	Multiple ^c		34%	6.20e-13	18535201	2008
Warfarin	Dosage	1053		<i>VKORC1</i> , <i>CYP2C9</i> and <i>CYP4F2</i>	Multiple		43.1%	<1e-78	19300499	2009
Acenocoumarol	Dosage	1451	287	<i>VKORC1</i> , <i>CYP2C9</i> , <i>CYP4F2</i> and <i>CYP2C18</i>	Multiple		48.80%	<5e-8	19578179	2009
PegIFN-alpha	Viral clearance	1137 + 1475		<i>IL28B</i>	Upstream	0.3	OR = 7.3	1.00e-38	19684573	2009
		293	555	<i>IL28B</i>		0.27	OR = 1.98	9.30e-09	19749758	2009
		142	172	<i>IL28B</i>		0.15	OR = 27.4	2.70e-32	19749757	2009
Citalopram and escitalopram	Concentration	435		<i>CYP2C19</i> and <i>CYP2D6</i>	Multiple		NA	<5e-8	24528284	2014
<i>Complex PGx traits (influenced by multiple genetic variants each with small effect)</i>										
Statin	LDL response	18,596	22,318	<i>SORT1</i> / <i>CELSR2</i> / <i>PSRC1</i> / <i>SLCO1B1</i> , <i>APOE</i> , <i>LPA</i>	Multiple		1.3%-5.2% ^e	<5e-8	25350695	2014
Metformin	HbA1c <7%	1024	1783 + 1113	<i>ATM</i>	Intronic	0.44	OR = 1.35	2.90e-09	21186350	2011
		10,577		<i>SLC2A2</i>	Intronic	0.29	0.17% greater	6.60e-14	27500523	2016
Antidepressant	Questionnaire, physician's opinion	865	1529	<i>HPRTP4</i>	Intergenic	0.4	OR = 1.36	5.03e-08	25897834	2015

^aDILI, drug-induced liver injury.

^bFreq, minor allele frequency (MAF) of the most significant or putative functional variant.

^cMultiple associated variants (probably due to allelic heterogeneity); variant types and frequencies are not shown.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

OR of homozygotes.
Percentage of extra LDL-C lowering in carriers versus non-carriers of the SNP.
p