ANNUAL REVIEWS

Annual Review of Pharmacology and Toxicology

Measuring Pharmacogene Variant Function at Scale Using Multiplexed Assays

Renee C. Geck,[1] Gabriel Boyle,[1] Clara J. Amorosi,[1] Douglas M. Fowler,[1,2] and Maitreya J. Dunham[1]

[1]Department of Genome Sciences, University of Washington, Seattle, Washington 98195, USA; email: dfowler@uw.edu, maitreya@uw.edu

[2]Department of Bioengineering, University of Washington, Seattle, Washington 98195, USA

Annu. Rev. Pharmacol. Toxicol. 2022. 62:531–50

First published as a Review in Advance on September 13, 2021

The *Annual Review of Pharmacology and Toxicology* is online at pharmtox.annualreviews.org

https://doi.org/10.1146/annurev-pharmtox-032221-085807

Keywords

pharmacogenomics, multiplexed assay of variant effect, deep mutational scan, next-generation sequencing, precision medicine, variable drug response

Abstract

As costs of next-generation sequencing decrease, identification of genetic variants has far outpaced our ability to understand their functional consequences. This lack of understanding is a central challenge to a key promise of pharmacogenomics: using genetic information to guide drug selection and dosing. Recently developed multiplexed assays of variant effect enable experimental measurement of the function of thousands of variants simultaneously. Here, we describe multiplexed assays that have been performed on nearly 25,000 variants in eight key pharmacogenes (*ADRB2*, *CYP2C9*, *CYP2C19*, *NUDT15*, *SLCO1B1*, *TMPT*, *VKORC1*, and the *LDLR* promoter), discuss advances in experimental design, and explore key challenges that must be overcome to maximize the utility of multiplexed functional data.

INTRODUCTION

Pharmacogenomics and Variable Drug Responses

PGx: pharmacogenomics, the study of genetic variation contributing to variable drug response

VDR: variable drug response

PharmGKB: Pharmacogenomics Knowledgebase, a curated collection of information on clinically actionable gene-drug interactions

CPIC: Clinical Pharmacogenetics Implementation Consortium, a group that provides clinical practice guidelines based on gene-drug interactions

Pharmacogenomics (PGx) is the study of genetic variation that contributes to variable drug responses (VDRs) (1). Approximately 25–50% of patients experience VDRs that lead to toxicity or change the amount of drug required to reach therapeutic concentrations (2–4). VDRs arise from pharmacokinetics of absorption, distribution, metabolism, and excretion (ADME) or pharmacodynamics of drug targets; therefore, PGx focuses on genetic variation in drug-metabolizing enzymes, transporters, targets, and interacting factors (4).

Genetic variation accounts for 20–40% of VDRs (5, 6), and the prevalence of observed pharmacogene variants predicts that 80% of patients likely have at least one variant that could affect drug response (7). Actionable pharmacogenes are ones for which variants can inform dosing or the prescription of an alternate therapy (8). Therapeutic index, severity of toxicity, consequences of underprescribing, and availability of alternate therapies are all considered when designating a pharmacogene as actionable (9). Therefore, identifying a variant in an actionable pharmacogene that predisposes to a VDR can inform clinical decisions over the course of a patient's life. To facilitate the application of genetic data to clinical action, groups such as the Pharmacogenomics Knowledgebase (PharmGKB) and the Clinical Pharmacogenetics Implementation Consortium (CPIC) provide evidence-based guidelines to interpret variant effects (10, 11). CPIC lists 440 gene-drug pairs where variants affect drug responses and has issued guidelines with recommended prescribing actions for 23 pharmacogenes (**https://cpicpgx.org/guidelines/**).

Most variants with suggestions for clinical implementation have a minor allele frequency (MAF) over 5% (12). However, multiple studies have illuminated the importance of rare (MAF <1%) variants. Studies of pharmacogene variation in data from large-scale sequencing projects predict that each individual harbors 40–120 pharmacogene variants, with 10–40% of this variation due to rare variants (6, 13).

Next-generation sequencing (NGS) improves upon platforms that only genotype candidate variants by identifying novel rare variants. Although most individuals profiled by NGS harbor common variants, the information gained from NGS is still beneficial (14): In some NGS studies of pharmacogenes, 90% of rare variants were previously unreported (4).

With the increased speed and decreased cost of NGS, our ability to sequence genomes has outstripped our ability to interpret the phenotypic consequences of genomic variants (2, 12, 15). Variants of unknown function hamper the utility of clinical sequencing efforts because only decisively interpreted variants can be used to guide clinical care (15). In order to implement NGS data for clinical prevention of VDRs, experimental methods are necessary to assess variants of unknown function.

Challenges with Current Approaches to Variant Functional Analysis

Many methods can be applied to analyze variant function, each with strengths and weaknesses. Genome-wide association studies and pedigree analysis are gold standards, but the rarity of many pharmacogene variants often precludes their use (16). Biochemical assays are used to determine the activity of enzyme variants and drug transporters but are limited in scale (12, 15, 17). Functional tests can be performed on variants expressed in laboratory systems and for point-of-care diagnostics when enzymatic activity correlates with likelihood of VDR (9). However, these methods have historically been performed reactively, after a variant is identified in an individual, increasing costs and often not providing data rapidly enough to be useful for the original patient (8, 9).

Proactive methods to test all possible variants have the potential to lead to immediate implementation upon identification of a variant, decreasing the cost burden on patients and improving

outcomes (9). Computational tools can proactively predict the effect of sequence variation on function (2), but recent studies have highlighted discordance with biochemical and clinical data (1, 18). Thus, computational predictions are considered low evidence of pathogenicity in the clinic (12, 15). In order to provide more data for clinical application, we need an approach that increases both the scale and accuracy of proactive variant functionalization.

MAVE: multiplexed assay of variant effect, a functional genomics method that uses next-generation sequencing to track variants after selection for altered function

VAMP-seq: variant abundance by massively parallel sequencing, a MAVE with selection based on the abundance of a tagged protein

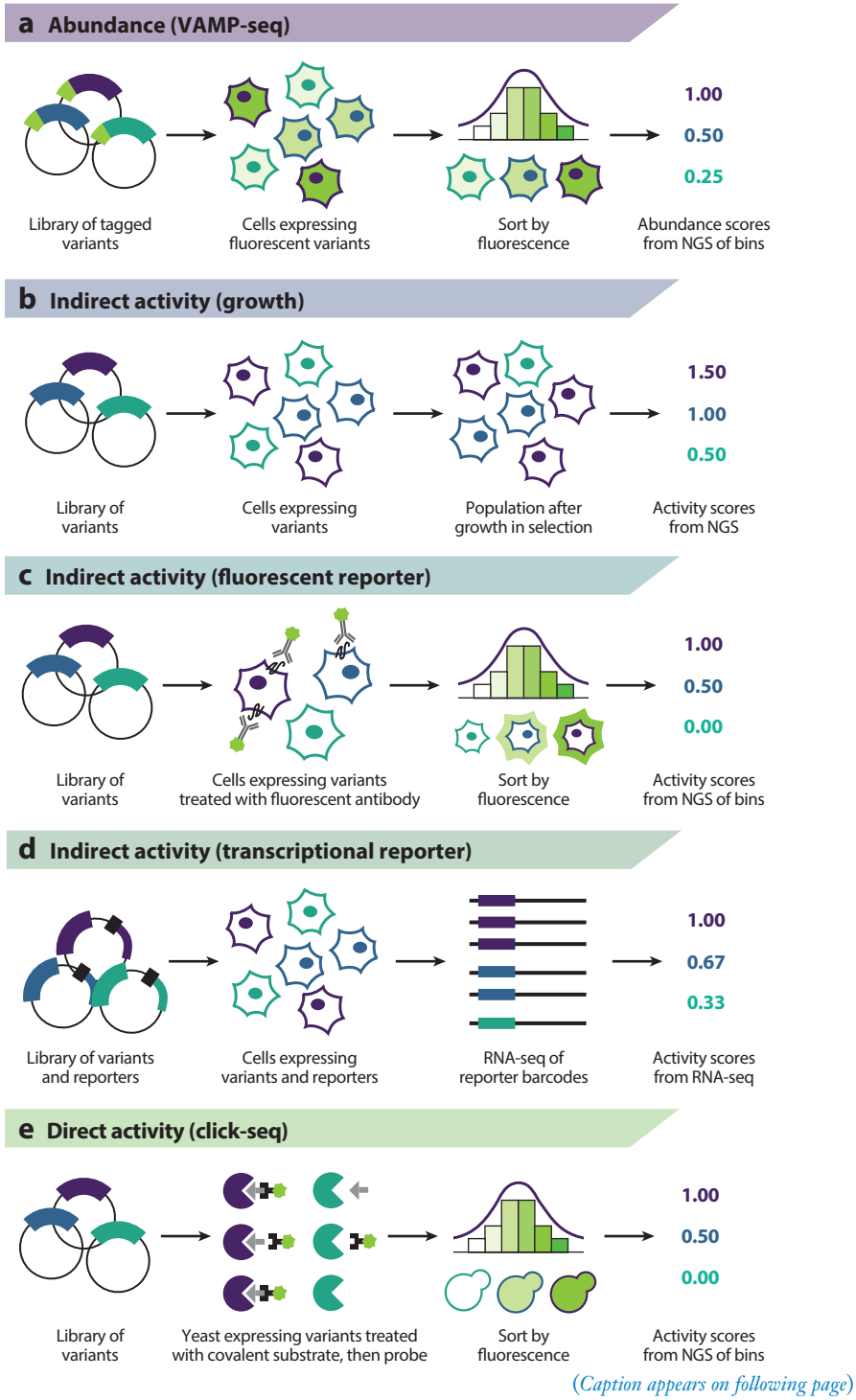
MAVEs

A multiplexed assay of variant effect (MAVE) assesses the function of hundreds to thousands of variants simultaneously by using NGS to track each variant throughout a selection, producing a functional score for every variant (19). MAVEs enable proactive measurement of the functional effects of all possible single-nucleotide variants in a target gene, offering a general, scalable way to resolve variants of uncertain significance (15).

Most MAVEs involve expressing a library of variants in a cell system so each cell expresses a single variant. The cells are selected based on a phenotype conferred by the expressed gene, and sequencing is used to determine the frequency of each variant in the selected populations (17). Strategies for library creation, sequencing, and data analysis have recently been reviewed elsewhere (15, 20, 21), so we focus here on the selection methods most relevant to pharmacogenes. MAVEs have been completed on eight pharmacogenes: *ADRB2*, *CYP2C9*, *CYP2C19*, *NUDT15*, *LDLR*, *SLCO1B1*, *TPMT*, and *VKORC1*. Most of these MAVEs are deep mutational scans, which measure the effect of amino acid substitutions on protein function (22). These pharmacogene MAVEs used five different selections: selection for steady-state protein abundance in cells; indirect selection for activity using growth, fluorescent reporters, or RNA sequencing (RNA-seq); and direct selection for enzyme activity using covalent substrates (**Figure 1**).

MAVEs for protein abundance are broadly applicable since they assess the ability of variants to fold and avoid degradation inside cells. To assess protein abundance, variant abundance by massively parallel sequencing (VAMP-seq) measures target protein variants expressed in cells as fusions to green fluorescent protein (**Figure 1*a***). Cells with high fluorescence express high-abundance variants that avoid degradation, whereas cells with low fluorescence express low-abundance variants. Cells are sorted into bins according to their fluorescence, and each bin is deeply sequenced to reveal the frequency of every variant at each level of fluorescence (23). Each variant's distribution across the bins is converted into an abundance score. VAMP-seq is particularly useful for pharmacogenes, since many deficiencies have been attributed to decreased abundance (18).

Selections for enzyme activity must be adapted to each pharmacogene because each selection relies on the specific biochemistry of each enzyme. One way to indirectly measure enzyme activity is cell growth or survival (**Figure 1*b***). Here, cells harbor a metabolic reporter for the target enzyme's activity or have a deletion of the genomic copies of the target enzyme. The variant-expressing cells are grown in conditions that select against loss-of-function variants, and cells are sequenced throughout growth to measure variant frequencies (16, 22). Since the target enzyme's activity is required for growth under selection, wild-type (WT)-like variants have higher frequency in the final population, and loss-of-function variants have lower frequency. Many growth-based MAVEs are conducted in the yeast *Saccharomyces cerevisiae*, where 2,696 human genes can functionally complement the yeast ortholog (15, 24). Functional complementation assays in yeast have already been developed for 179 human disease-associated genes (25). Beyond yeast, an explosion in gene essentiality screening data has identified approximately 2,000 genes important for growth in numerous human cell lines (26, 27), highlighting candidates for growth-based MAVEs in human cells (28).



(*Caption appears on following page*)

Figure 1 (*Figure appears on preceding page*)

Basic schematics of MAVEs. (*a*) To determine variant abundance, cells expressing a library of fluorescently tagged variants are sorted by fluorescence, and the resultant bins are sequenced. (*b*) To indirectly determine variant activity, cells expressing a library of variants are grown in conditions that select against loss-of-function variants. A decrease in frequency in the final sequencing data signifies decreased activity. (*c*) Variant activity can also be indirectly measured by treating cells expressing a library of variants with fluorescent antibody against a cell surface protein reporter of enzyme activity. Cells are sorted by fluorescence, and the resultant bins are sequenced. (*d*) Transcriptional reporters can be used to indirectly measure activity. Variants in a protein with a transcriptional effector are introduced into cells with unique barcodes under an effector-mediated binding site. For MAVES on promoters, the barcodes are placed directly under the control of promoter variants. RNA-seq to detect barcodes is used to group variants by transcriptional activity. (*e*) To directly measure variant activity, yeast cells expressing a library of variants are treated with a covalent substrate, excess is washed away, and substrate bound to active enzyme is then attached to a fluorophore using click chemistry. Cells are sorted by fluorescence, and the resultant bins are sequenced. Panel *e* adapted from Reference 31. Abbreviations: MAVE, multiplexed assay of variant effect; NGS, next-generation sequencing; RNA-seq, RNA sequencing; VAMP-seq, variant abundance by massively parallel sequencing.

Other indirect selection methods for enzyme activity involve additional types of reporters. For fluorescent reporters, cells are engineered to express a cell-surface marker that is driven by the enzyme activity of the target gene, which is then detected using a fluorescent antibody against the marker (15, 29) (**Figure 1*c***). Cells expressing the library of variants are sorted according to fluorescence levels and sequenced to score every variant's ability to drive the reporter. RNA-seq can also be a reporter for enzyme activity when the downstream output of a pathway is transcriptional activation. RNA barcodes are placed into transcribed regions regulated by binding motifs for the relevant transcription factor and integrated into cells with the library of variants (16) (**Figure 1*d***). The RNA barcodes are sequenced to determine transcriptional activity associated with each variant. A similar design can be used to assess promoter or enhancer activity (30).

Finally, target enzyme activity can be directly selected using a covalent substrate (**Figure 1*e***). Cells expressing a library of the target enzyme are treated with a substrate that reacts covalently with the enzyme. Unreacted substrate is washed away and click chemistry is used to attach a fluorophore to enzyme-bound substrate, hence this method is called click-seq (31). Cells are sorted according to their fluorescence and sequenced to score every variant's ability to react with the covalent substrate.

Every MAVE produces a set of scores related to variant function. These functional scores quantify how each variant performs in the assay, typically relative to the WT protein. Functional scores can be used to classify the molecular effect of variants (e.g., loss of function or WT-like). They can also be compared to variants with known clinical effects to translate loss of molecular function into impact on a clinical phenotype (23, 28). How functional scores translate into clinical phenotypes is different for each MAVE and affected by a variety of factors ranging from data quality to how well the functional assay correlates with the clinical phenotype. Thus, including as many known pathogenic and benign variants as possible in each library is important (32).

click-seq: a MAVE with selection based on enzyme activity using a covalent substrate and click chemistry

MaveRegistry: a catalog of published and ongoing MAVEs

MAVEs ON PHARMACOGENES

Recently, multiplexed functional data have emerged for a variety of genes, including a handful of pharmacogenes. As of March 2021, MAVEs have been performed on the coding sequences of pharmacogenes *ADRB2*, *CYP2C9*, *CYP2C19*, *NUDT15*, *SLCO1B1*, *TPMT*, *VKORC1*, and the *LDLR* promoter (**Table 1**). For each of these genes, we review the rationale for applying MAVEs, how MAVEs were used, and what was learned. MAVEs are also in progress on the coding sequences of *CYP2D6*, *G6PD*, *HMGCR*, *LDLR*, and *MTHFR* according to MaveRegistry.

Table 1 MAVEs on pharmacogenes

Gene	MAVE	Number of variants (% of possible)[a]	Decreased function (%)[b]	Reference(s)
ADRB2	Activity (reporter)	7,800 (99.6%)	ND	90
CYP2C9	Abundance	109 (1.17%)	17.4	18
	Abundance	6,370 (68.4%)	36.8	31
	Activity (direct)	6,142 (66.0%)	64.9	
CYP2C19	Abundance	121 (1.30%)	29.8	18
LDLR promoter	Activity (reporter)	945 (99.1%)	23.8	30
NUDT15	Abundance	2,923 (94.1%)	30.0	48, 52
	Activity (survival)	2,935 (94.2%)	24.0	
SLCO1B1	Abundance	137 (1.04%)	13.1	99
TPMT	Abundance	3,689 (79.2%)	21.1	23
VKORC1	Abundance	2,695 (87.0%)	25.0	29
	Activity (reporter)	697 (22.5%)		

[a]Numbers refer to single missense amino acid variants only, except for *LDLR* promoter, for which they refer to single missense base substitutions.
[b]What constitutes decreased function for each MAVE is defined in the referenced study.
Abbreviations: MAVE, multiplexed assay of variant effect; ND, not determined.

TPMT

The thiopurine drugs thioguanine, 6-mercaptopurine (6-MP), and the 6-MP prodrug azathioprine are used to treat leukemia and autoimmune diseases by inhibiting de novo purine synthesis and damaging DNA and RNA (33). Thiopurine drugs are administered as inactive precursors that are metabolized into thioguanine nucleotides. If thioguanine nucleotides are not inactivated by thiopurine methyltransferase (TPMT), they can accumulate to toxic levels that cause hepatotoxicity and severe myelosuppression (34, 35). Genetic variation in *TPMT* is primarily responsible for thiopurine intolerance in patients of European and African ancestry, while variation in nudix hydrolase 15 (*NUDT15*) explains most of the VDRs to thiopurines in people of Asian and Hispanic backgrounds (36).

TPMT is highly polymorphic with more than 40 reported variants (37–39). Up to 14% of the general population possess a *TPMT* variant that reduces thioguanine nucleotide metabolism and may thus experience VDRs to thiopurines (40). CPIC provides guidelines for decreasing thiopurine doses based on *TPMT* variant function, which they have assigned for 12 variants (36).

Previous studies identified three TPMT variants that comprise 95% of decreased-function *TPMT* alleles in individuals with European or African ancestry and showed that they decrease TPMT abundance (41–43). Therefore, Matreyek et al. (23) used VAMP-seq (**Figure 1*a***) to measure protein abundance of TPMT variants in HEK293T cells. They measured abundance scores for 3,689 (79.2%) of the possible TPMT missense variants. VAMP-seq-derived TPMT variant abundance scores correlated well with individually measured variant abundance ($n = 19$, Pearson's $r = 0.75$), and previously characterized variants had reduced abundance consistent with previous studies (42–44).

Ninety-six of 118 rare TPMT variants of unknown function recently identified by large-scale population sequencing (45) and by sequencing individuals with thiopurine intolerance (46) were scored. Fourteen (14.6%) had low abundance and 17 (17.7%) possibly low abundance, indicating that individuals with these variants may have increased risk for thiopurine toxicity. Overall, 777 (21.1%) of tested missense variants were classified as low abundance, most of which had not been previously implicated in thiopurine intolerance.

NUDT15

Variation in *NUDT15* also contributes to thiopurine VDRs. NUDT15 dephosphorylates thioguanosine triphosphate to the less toxic thioguanosine monophosphate; thus, loss of NUDT15 function elevates DNA-incorporated thioguanine (DNA-TG) and cytotoxic effects in response to thiopurines (36).

Currently, CPIC suggests decreasing thiopurine dose or using alternate drugs for seven common *NUDT15* variants (36). However, the functional effects of other *NUDT15* variants, including six recently discovered variants and 77 missense variants reported on the Genome Aggregation Database (gnomAD) (47), remained unknown (48, 49). Thus, Suiter et al. (48) conducted two complementary MAVEs on *NUDT15*, measuring abundance and thiopurine toxicity for 3,077 (98.7%) of the possible *NUDT15* missense variants expressed in HEK293T cells.

gnomAD: Genome Aggregation Database, a curated collection of exome and genome sequences compiled from disease-specific and population genetic studies

They first assayed abundance by VAMP-seq since NUDT15 variants with known clinical effects have reduced stability (50). Thermal stability assays correlated with fluorescence ($n = 14$, Spearman's $\rho = 0.98$), validating the assay. In total, 735 variants had lower abundance than the R139C variant, which is associated with decreased thiopurine tolerance (51), and were thus considered likely to also increase risk of thiopurine intolerance.

To directly assay thiopurine cytotoxicity, they treated cells expressing *NUDT15* variants with thioguanine for six days. Cells harboring loss-of-function *NUDT15* variants were unable to detoxify the thioguanine nucleotides and, consequently, grew more slowly (**Figure 1*c***). They validated the assay by measuring toxic DNA-TG accumulation, observing a negative correlation between activity and DNA-TG ($n = 9$, Spearman's $\rho = -0.72$). Combining multiple functional scores to make a conclusion about overall variant function remains a challenge. In this case, Suiter et al. (48) designated the lower of the abundance and thiopurine toxicity scores for each variant as its activity score and determined that 1,152 (40.5%) of tested variants were potentially damaging. Cagiada et al. (52) further analyzed the data and determined that 410 (14%) of the variants they analyzed lost both activity and abundance, 270 (10%) only activity, and 439 (16%) were low abundance but retained near-WT activity. In general, variants affecting fully buried residues led to low abundance (29/35), while low-activity variants were in residues that coordinate substrate and cofactor binding clustered around the active site (48, 52).

To assess the clinical utility of their functional data, Suiter et al. (48) used their activity scores to predict thiopurine toxicity in 2,398 patients. Ten patients had *NUDT15* missense variants; five experienced thiopurine toxicity and five did not. Activity scores predicted these responses with 100% sensitivity and specificity, while either abundance or thiopurine sensitivity scores alone lacked sensitivity, emphasizing the strength of combining scores from multiple assays that test different aspects of variant function. Suiter et al. also identified 108 unique *NUDT15* missense variants in 8,871 individuals (6.3%) in gnomAD. Surprisingly, 8,323 (93.8%) of these individuals had variants with likely damaging activity scores.

Future studies of *NUDT15* are still necessary to address potential substrate specificity and interactions with *TPMT* variants. Dosing strategies for individuals who harbor missense variants in both genes (36) would be better informed by MAVEs of *NUDT15* in the context of common *TPMT* variants, and the converse. By addressing the combined effects of genes that contribute to thiopurine toxicity, guidelines for administration of these important drugs could be further improved.

VKORC1

Vitamin K oxidoreductase (VKOR) is a transmembrane protein, encoded by the *VKORC1* gene, that drives the vitamin K cycle and is necessary to carboxylate vitamin K–dependent clotting

factors (53, 54). Variants with reduced abundance can cause vitamin K–dependent blood clotting deficiency 2 (53). VKOR is the target of the anticoagulant drug warfarin, which is prescribed to over 15 million people annually (55) who must be monitored to avoid hemorrhage or blood clots. *VKORC1* polymorphisms contribute 25% of the variation in warfarin dose (56), and another 15–20% is contributed by variants in CYP2C9, which metabolizes warfarin (57). Improper warfarin dosing is associated with substantial risk of hemorrhage (58–60), so understanding factors that contribute to warfarin VDRs is critical.

In 2007, the US Food and Drug Administration (FDA) updated the label of warfarin to include information about pharmacogenetic testing for *CYP2C9* and *VKORC1* (61), and cell-based assays were used to study the impact of a promoter variant on VKOR activity and abundance (53, 62–65). However, the only variant included in CPIC guidelines is the promoter variant (66), and only two variants were determined to be pathogenic by ClinVar (67), whereas 224 additional coding missense variants have been identified in gnomAD.

Thus, Chiasson et al. (29) executed MAVEs assessing VKOR abundance and activity. They measured the abundance of 2,695 (87.0%) missense variants via VAMP-seq (**Figure 1*a***). MAVE design was validated by testing individual fluorescence, which correlated well with abundance scores ($n = 10$, Pearson's $r = 0.96$); western blots also showed high concordance with abundance scores. The activity of 697 (22.5%) variants was measured using an adapted cell-based assay in which fluorescent antibodies signal successful carboxylation of a blood clotting cofactor secreted and retained on the cell surface (68) (**Figure 1*c***).

Based on the patterns of low variant abundance scores, Chiasson et al. (29) confirmed studies showing that VKOR has four transmembrane domains (63, 69), not three as others proposed (70, 71). Positions necessary for VKOR activity were identified by investigating variants with high abundance and low activity scores. Eleven functionally constrained positions defined part of the VKOR active site, six of which were previously identified in vitamin K docking simulations (62). Two conserved cysteines important for warfarin binding were also identified (63).

Chiasson et al. (29) curated a list of 215 variants previously reported to affect warfarin response and classified 193 according to their abundance. Of these, 129 (60%) had WT-like or possibly WT-like abundance, 30 (14.0%) were low abundance, and 12 (5.6%) were high abundance. Additionally, one variant had low activity and high abundance, indicating that loss of its activity is not a result of reduced abundance. Overall, they identified 54 previously uncharacterized clinical variants that may contribute to warfarin VDRs.

Seven warfarin resistance variants have been previously identified (53, 66, 72, 73) and were investigated to determine their effects on VKOR abundance and activity. Known resistance variants spanned a range of abundances, and all five of the resistance variants scored for activity were WT-like, indicating they may confer resistance by blocking warfarin binding.

Given that *VKORC1* promoter variation can alter warfarin sensitivity, future MAVEs focused on *VKORC1* expression are also warranted (74). Full understanding of the clinical implications of missense variations in *VKORC1* remains an important goal for treating blood clotting disorders.

CYP2C9

CYP2C9 encodes a cytochrome P450 enzyme that oxidizes endogenous and xenobiotic compounds, metabolizing approximately 15% of small-molecule drugs (75). CPIC lists 11 level A priority *CYP2C9* gene-drug pairs (66, 76, 77), including the anticoagulant warfarin, the anticonvulsant phenytoin, and several nonsteroidal anti-inflammatory drugs.

CYP2C9 variants account for 15–20% of the variation in warfarin dose (57), often leading to warfarin sensitivity: Patients with the loss-of-activity *CYP2C9*2* or **3* variants require lower

warfarin maintenance doses and are at higher risk of a serious or life-threatening bleeding event (78). Genotype-guided warfarin dosing can be effective in reducing VDRs in specific situations (79) but relies on *CYP2C9* and *VKORC1* variants of known function.

PharmVar: Pharmacogene Variation, a consortium that catalogs nomenclature and allelic variation in pharmacogenes

Two groups have conducted MAVEs on CYP2C9 at different scales and measuring different aspects of CYP2C9 function. Zhang et al. (18) constructed a variant library of 109 existing human CYP2C9 missense variants and measured the abundance of these variants using VAMP-seq (23) (**Figure 1*a***). They found that 19 (17.4%) variants had less than approximately 25% of WT protein abundance, indicating that these variants could lead to VDRs and should be of clinical interest. Fifteen (78.9%) of these variants had not been previously reported by the Pharmacogene Variation (PharmVar) Consortium.

Amorosi et al. (31) performed two MAVEs on over 8,000 total CYP2C9 variants, measuring both enzyme activity and abundance. One library was codon optimized for human expression to measure abundance and contained 6,370 (68.4%) of the possible missense variants, and one was codon optimized for yeast expression to measure enzyme activity and contained 6,142 (66.0%) variants. CYP2C9 variant abundance was measured using VAMP-seq (23) (**Figure 1*a***) and validated by individual fluorescence measurements, which correlated with abundance scores ($n = 12$, Pearson's $r = 0.94$). Variant enzyme activity was measured using click-seq with a CYP2C9-specific activity-based probe (an analog of tienilic acid, a covalent inhibitor of CYP2C9) coupled to activity-based protein profiling (**Figure 1*e***). The activity scores correlated highly with individual labeling ($n = 14$, Pearson's $r = 0.99$) and with gold-standard liquid chromatography–mass spectrometry measurement of warfarin turnover ($n = 14$, Pearson's $r = 0.87$).

Of these CYP2C9 variants, 2,347 (36.8%) had decreased abundance, and their abundance scores correlated well with abundance reported by Zhang et al. (Pearson's $r = 0.74$). Furthermore, 3,987 (64.9%) CYP2C9 variants showed decreased activity, and protein abundance was responsible for half of the variation in CYP2C9 function. Thus, as for *NUDT15* and *VKORC1*, measuring multiple protein functions can reveal a more comprehensive set of loss-of-function variants and yield insight into the mechanisms by which variants alter function. By comparing variant activity and abundance, Amorosi et al. (31) confirmed that the structural core of CYP2C9 was key for protein stability and thus activity, although mutations to many positions involved in heme coordination and binding had little effect on abundance yet ablated activity. Positions within the CYP2C9 active site also generally tolerated mutations.

To assess the clinical utility of the multiplexed functional data, Amorosi et al. (31) compared the activity and abundance scores to CPIC functional classes for 32 variants with clinical recommendations and found them largely concordant. They also calculated activity scores for 340 missense variants curated from gnomAD, 319 of which lacked functional annotations on CPIC. Additionally, 199 (58.5%) variants exhibited a significant loss of activity, emphasizing the potential clinical impact of these scores to identify individuals with a high likelihood of VDRs.

In the future, other substrates of CYP2C9 should be tested when measuring activity to determine the extent of substrate-dependent variant effects, which have been previously reported (80). Performing a MAVE of CYP2C9 with a warfarin-derived substrate would be ideal, but this substrate was not amenable to click-seq (A.E. Rettie, personal communication). Alternate approaches like droplet-based microfluidics technologies coupled with fluorogenic substrates are also promising (81), but these approaches are not without their own technical challenges (82).

CYP2C19

CYP2C19 encodes a cytochrome P450 enzyme that metabolizes many important drugs, including antidepressants, proton pump inhibitors, and the antiplatelet prodrug clopidogrel (83). CPIC

lists eight drugs with level A priority in a gene-drug pair with *CYP2C19*. As with other cytochrome P450 enzymes, variation in *CYP2C19* can lead to interindividual variation in response to CYP2C19-metabolized drugs.

Clopidogrel is metabolized into an active form primarily by CYP2C19, with several cytochrome P450 enzymes playing lesser roles. Individuals with loss-of-function *CYP2C19* variants have decreased platelet responsiveness to clopidogrel and are at increased risk for serious adverse cardiovascular events (84). Additionally, the common gain-of-function *CYP2C19*17* promoter variant results in increased activity, and individuals homozygous for this variant have an ultrarapid metabolizer phenotype and an increased risk of bleeding complications when taking clopidogrel (85, 86).

So far, a small MAVE has been performed on CYP2C19. Zhang et al. (18) constructed a variant library of 121 human CYP2C19 missense variants present in the general population and measured the abundance of these variants using VAMP-seq (23) (**Figure 1*a***). They reported that 36 (29.8%) of these variants had less than approximately 25% of WT protein abundance, and fluorescence of five variants with a range of abundance scores correlated with previously published western blot data (1). Thirty of the 36 low-abundance variants had not been previously reported in PharmVar.

Future studies could use a more complete *CYP2C19* library, consider testing additional phenotypes such as enzyme activity in addition to abundance, and measure the effects of noncoding variation to CYP2C19 function, given the importance of the *CYP2C19*17* promoter variant.

ADRB2

ADRB2 encodes the β_2-adrenergic receptor (β_2AR), a G protein–coupled receptor (GPCR), the target of agonist drugs used to treat asthma (albuterol and salmeterol) and antagonists used to treat cardiovascular disease (carvedilol and propranolol) (87). β_2AR signals through the heterotrimeric G_s protein to activate adenylyl cyclase and generate cyclic AMP.

Three *ADRB2* variants alter agonist response: R16G and Q27E enhance it, and T164I reduces it (87). However, due to conflicting results (87, 88), and because *ADBR2* variation accounts for only a small proportion of β_2-agonist VDRs (89), *ADRB2* is classified as having provisional level D gene-drug interactions on CPIC, indicating a lack of sufficient evidence for prescribing recommendations.

To characterize the functional effects of *ADRB2* variants, Jones et al. (90) constructed a reporter system in HEK293T cells containing a barcode under the control of cyclic AMP response elements. They integrated *ADRB2* variants into the reporter cells, treated them with four different concentrations of the β_2AR agonist isoproterenol, and measured reporter expression by RNA-seq (**Figure 1*d***). They scored 7,800 (99.6%) *ADRB2* missense variants, confirming that R16G, Q27E, and T164I behaved as expected in their MAVE and a luciferase reporter assay.

While Jones et al. (90) determined that 11 of the 180 *ADRB2* missense variants in gnomAD were potentially loss-of-function, they largely focused on structure-function relationships. Transmembrane domains were less tolerant of variation, and unbiased clustering by mutational tolerance revealed several important regions. Mutational tolerance correlated with residue conservation across GPCRs (Spearman's $\rho = -0.676$), suggesting that *ADRB2* functional data could apply to conserved regions of other closely related GPCRs, including those that are the targets of widely prescribed drugs (91).

The *ADRB2* MAVE provides a generalizable approach to studying the function of drug receptors. Approximately 34% of FDA-approved drugs target GPCRs (92), and drugs that target GPCRs or other transcriptional effector signaling pathways would be amenable to similar MAVEs.

SLCO1B1

SLCO1B1 also encodes a transmembrane protein, the organic ion transporter protein 1B1 (OATP1B1). OATP1B1 is important for hepatic uptake of estradiol, bilirubin, and statin drugs (93). Decreased OATP1B1 activity can lead to increased plasma concentrations of statins and myopathy (94). CPIC provides guidelines for simvastatin dosing, but they are specific to *SLCO1B1*5*, **15*, and **17*, which contain a V174A substitution and are most common in individuals of European descent (93, 95).

Fifteen other *SLCO1B1* variants have been individually characterized (93) and alter stability and translocation to the plasma membrane (96–98). Investigating the possibility that uncharacterized *SLCO1B1* variants may decrease stability, Zhang et al. (99), who conducted two of the CYP MAVEs described above, performed VAMP-seq (**Figure 1*a***) on a small library of 137 *SLCO1B1* variants identified in the general population. They found six variants (4.4%) with less than 25% WT abundance and 12 (8.8%) other variants with less than 50% abundance. They confirmed that all six of the very low abundance variants decreased OATP1B1 protein by western blotting and decreased uptake of radiolabeled 17-estradiol β-D-glucuronide. Further work is needed to determine the clinical predictive power of *SLCO1B1* functional scores, since there were no clinical data available on the six proposed severely damaging alleles.

To assess loss-of-function by mechanisms other than degradation, Zhang et al. (99) considered variants with mutations in the same TM4 domain as *SLCO1B1*5*, which leads to mislocalization and decreases activity but not abundance (100). They measured substrate uptake for eight variants in TM4. All had WT-like abundance, but two decreased and two increased substrate uptake, indicating that variants of *SLCO1B1* can have functional effects on activity but not abundance.

LDLR

Variants in the low-density lipoprotein (LDL) receptor gene *LDLR* are also important for the pharmacokinetics of statin drugs, since loss-of-function variants can lead to increased LDL and decrease response to lipid-lowering drugs (101). Variants in the promoter, introns, and coding regions of *LDLR* have been shown to increase risk of familial hypercholesterolemia (102–106), which is associated with VDRs that often lead to undertreatment (101). However, due to a lack of clinical support for variant-drug interactions, prescribing guidelines are not available for *LDLR* on CPIC. Five *LDLR* variants—two intronic and three in the 3′–untranslated region (UTR)—have been cataloged as level 3 in PharmGKB, indicating a lack of evidence or replication.

Kircher et al. (30) conducted MAVEs on the promoter sequences of ten genes, including *LDLR*. They expressed a library of all 954 variants in the LDLR promoter in HepG2 cells and quantified the expression of downstream barcode sequences by RNA tag-sequencing (**Figure 1*d***). Their reporter values were in agreement with previously published luciferase activity assays.

Sequencing of 945 (99.1%) variant reporters met their quality control requirements. A total of 288 (30.5%) variants led to significantly altered expression, with 225 (23.8%) decreased and 63 (6.7%) increased. Variants previously identified in individuals with familial hypercholesterolemia (105) significantly decreased expression. In general, Kircher et al. (30) observed that variants in the same transcription factor binding site led to similar functional effects. MAVEs in progress on the coding region of LDLR are likely to identify other functionally important variants, since over 1,600 variants have been described (101).

As the first MAVE on a noncoding region of a pharmacogene, this study of the LDLR promoter highlights the potential for multiplexed functional data to inform functional effects of noncoding regions. The selection of a gene for which promoter variants have been identified in patients

(103–105) and are known to alter pharmacological profiles increases the potential clinical utility of these data.

THE FUTURE OF MAVEs ON PHARMACOGENES

Best Practices for Future MAVEs

Important lessons have been learned from MAVEs conducted on diverse genes that should be applied to future MAVEs on pharmacogenes. As part of the Brotman-Baty Institute's Mutational Scanning Working Group, we previously published suggestions on MAVE design and interpretation (32).

In particular, when designing a MAVE, it is essential to consider the intrinsic limitations of the assay. Since each MAVE measures a specific function, one MAVE is unlikely to capture all functionally abnormal variants. This limitation is exemplified by the CYP2C9, NUDT15, and VKOR studies, which used MAVEs for abundance and also activity. Activity and abundance scores correlated well for CYP2C9 variants (Pearson's $r = 0.75$) but not for NUDT15 or VKOR variants (NUDT15 Pearson's $r = 0.38$, VKOR Pearson's $r = 0.26$), and even for CYP2C9, numerous variants were observed with reduced activity but WT-like abundance (31). Many completed MAVEs on pharmacogenes focus on only a single time point or a single drug, but time- and substrate-dependent effects of variants are important to consider when designing an assay (31, 90, 107). The dynamic range of each MAVE must be broad enough to distinguish known pathogenic from benign variants and validated by individually testing variants with a range of functional scores, preferably by an orthogonal gold-standard assay (32).

Adequate replication and consistent analysis methods are essential to compensate for noisy data (16, 17, 32); Kinney et al. (16) recommended protocols from the fields of experimental evolution and RNA-seq to produce values such as selection coefficients, which are more meaningful than commonly reported fold change–based functional scores. Selection coefficient–based scores can more readily be used to compare across genes and between data sets from different laboratories. Computing and reporting sensitivity and specificity of a MAVE for clinically characterized variants are also important for eventual implementation (32). For pharmacogenes with guidelines for prescription of alternate drugs, it may be best to compromise on specificity for the sake of sensitivity in order to identify individuals who may be at risk for VDR and can be prescribed an alternate drug (8).

Prioritizing Targets for Future MAVEs

Given the large number of pharmacogene variants of uncertain function, we need a systematic approach to prioritize targets of future MAVEs. We previously identified 31 pharmacogenes with high levels of evidence from CPIC that would benefit from functional analysis (17). In total, there remain 12 pharmacogenes denoted as having at least one level A drug interaction on CPIC and level 1A on PharmGKB that have not yet been investigated in MAVEs, published or in progress (108) (**Figure 2**).

Other methods for prioritization involve considering the information available on each pharmacogene and the adaptability of an assay to a MAVE. Of particular interest are pharmacogenes with notable variation in gnomAD but little functional information or clinical consequences noted in CPIC (17). Small proteins can be prioritized to simplify achievement of saturating mutagenesis, and assays of larger proteins could be simplified by focusing only on specific domains, as was done for *BRCA1* (28, 109, 110). Genes with functional assays amenable to multiplexing are also attractive: Extant reporter assays can be adapted to MAVEs, and genes essential for growth are good candidates for growth-based MAVEs.

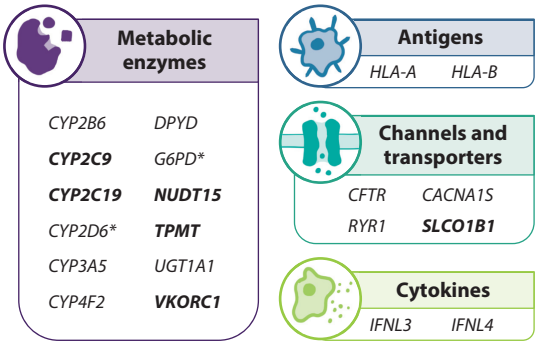


Figure 2

Pharmacogenes that significantly contribute to variable drug responses. Pharmacogenes listed have one or more level A (Clinical Pharmacogenetics Implementation Consortium) or level 1A (Pharmacogenomics Knowledgebase) drug interaction, signifying that genotyping can be used to inform drug prescription. Genes are sorted by biological function to consider the most suitable types of multiplexed assay of variant effect (MAVE). Genes in bold have been the subject of MAVEs; asterisks indicate a MAVE in progress, as listed on MaveRegistry.

Developing New MAVEs for Pharmacogenes

Since most pharmacogenes are enzymes, focusing efforts on designing robust, activity-based MAVEs is most likely to produce clinically relevant functional data. In order to study more complex questions, new assays and analysis methods must be developed. MAVES on pharmacogenes that encode secreted proteins, such as *IFNL3* (111), would require a method to link each variant to the cell that produced it while considering the effect of the variant on other cells in the population.

For pharmacogene variants, heterozygosity can affect risk of VDR, but current MAVEs do not faithfully recapitulate endogenous expression. Using phasing to define important haplotype interactions in pharmacogenes is likely to be informative, as observed for OATP1B1 N130D, which decreases plasma statin concentrations in the *SLCO1B1*1B* haplotype but increases plasma statins on the **15* haplotype with V174A (112). MAVEs could interrogate interactions between specific variants (8), which would be most beneficial for pharmacogenes with a small number of common haplotype backgrounds. One study found that most noncoding variant-drug associations on PharmGKB can be explained by haplotype associations with coding variants, supporting the utility of pharmacogene MAVEs on coding variants for common haplotypes (113).

There is also great potential to expand MAVEs on pharmacogenes beyond single-nucleotide variants in coding regions. A study of variants in 208 ADME genes in 62,402 individuals found that 97% of genes had at least one copy number variant (114). Copy number variants are common in GPCRs (115), and structural variants of cytochrome P450 gene *CYP2D6* are found in 30% of individuals with decreased CYP2D6 function (116). Recently, techniques using MuA transposase or clustered regularly interspaced short palindromic repeat–associated protein 9 have been used to make variant libraries of deletions (117, 118). Additionally, all of the current MAVEs on pharmacogenes have focused on coding regions, with the exception of the LDLR promoter. Massively parallel reporter assays have been successfully used to study the effects of regulatory elements such as promoters, enhancers, splice sites, and UTRs (16, 119). Other MAVE selection methods not yet applied to pharmacogenes include in vitro display, yeast two-hybrid screens, and transcription factor binding (15, 16), which could be used to interrogate effects of variants on drug-receptor binding or query the functions of regulatory factors. Applying these techniques to pharmacogenes would increase the scope of clinically interpretable genomic information.

Making Data from MAVEs Accessible

MaveDB: a repository for data sets from MAVEs

Before it can be applied clinically, the massive amount of data generated by MAVEs must be made available, accessible, and interpretable. The repository MaveDB contains 110 data sets from MAVEs on 65 targets (as of May 2021) and interfaces with applications for data analysis (120). MaveRegistry also catalogs published and ongoing MAVEs to encourage collaboration and decrease redundancy (108). As of May 2021, MaveRegistry lists published MAVEs on 98 targets and MAVEs in progress on an additional 43 targets.

Accessible data are essential for comparison of MAVEs and for development of tools with clinical applications. Multiplexed functional data can be used to build more accurate variant function prediction algorithms (8, 121, 122) and train machine learning algorithms to interpolate missing functional scores (123, 124).

Applying Multiplexed Functional Data in the Clinic

Clinical implementation of multiplexed functional data for PGx will require an increase in sequencing for precision medicine, which is promoted by variant functionalization: 85% of physicians do not routinely order PGx testing, mostly due to lack of guidelines for action based on test results (125). As more variants of actionable pharmacogenes are interpreted, the benefit of ordering PGx testing increases (9). Therefore, clinical action requires improved understanding of variant function and communication of functional data to clinicians.

There remains a need for a standardized system to incorporate variant functional scores into clinical databases. Thousands of pharmacogenetic biomarkers have been discovered, but relatively few have been implemented in the clinic due to challenges in validation and implementation (4). The American College of Medical Genetics and Genomics and the Clinical Genome Resource published guidelines that were designed for low-throughput functional assays and that can also be applied to MAVEs (32, 126–128). Reporting to repositories such as PharmVar will require determining the strength of evidence generated by each variant functional score, based on the sensitivity and specificity of the assay and on the reproducibility of the score itself (32). Eventually, we hope that a centralized database for MAVE reporting can be linked to relational databases curated by CPIC and PharmGKB, so functional scores can be directly communicated to clinicians and thereby be of greater use to the community (8). Standardized reporting and scoring guidelines for repositories, and educating clinicians on data availability and interpretation, are necessary to increase the impact of variant functionalization.

CONCLUSIONS

So far, MAVEs measuring a diversity of important protein and DNA sequence functions have been applied to eight pharmacogenes, covering nearly 25,000 variants. These data have the potential to contribute to pharmacogene variant interpretation from clinical sequencing data. By conducting multiplexed functional studies on more pharmacogenes, and making that information available to clinicians, we can improve precision medicine for safe and accurate drug dosing.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We would like to thank A.E. Rettie and members of the Dunham lab for critical advice and editing. The authors of this publication were supported by the National Institute of General Medical

Sciences (NIGMS) of the National Institutes of Health (NIH) under award R01 GM132162 to D.M.F. and M.J.D., and in part through the NIH National Cancer Institute Cancer Center Support grant P30 CA015704 to D.M.F. and M.J.D. R.C.G. was supported by the National Human Genome Research Institute of the NIH under award T32 HG00035 and the NIGMS under award F32 GM143852. The research of M.J.D. was supported in part by a Faculty Scholar grant from the Howard Hughes Medical Institute.

LITERATURE CITED

1. Devarajan S, Moon I, Ho M-F, Larson NB, Neavin DR, et al. 2019. Pharmacogenomic next-generation DNA sequencing: lessons from the identification and functional characterization of variants of unknown significance in CYP2C9 and CYP2C19. *Drug Metab. Dispos.* 47(4):425–35
2. Zhou Y, Fujikura K, Mkrtchian S, Lauschke VM. 2018. Computational methods for the pharmacogenetic interpretation of next generation sequencing data. *Front. Pharmacol.* 9:1437
3. Katara P, Yadav A. 2019. Pharmacogenes (PGx-genes): current understanding and future directions. *Gene* 718:144050
4. Schwarz UI, Gulilat M, Kim RB. 2019. The role of next-generation sequencing in pharmacogenetics and pharmacogenomics. *Cold Spring Harb. Perspect. Med.* 9(2):a033027
5. Ingelman-Sundberg M, Rodriguez-Antona C. 2005. Pharmacogenetics of drug-metabolizing enzymes: implications for a safer and more effective drug therapy. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360(1460):1563–70
6. Kozyra M, Ingelman-Sundberg M, Lauschke VM. 2017. Rare genetic variants in cellular transporters, metabolic enzymes, and nuclear receptors can be important determinants of interindividual differences in drug response. *Genet. Med.* 19(1):20–29
7. Schärfe CPI, Tremmel R, Schwab M, Kohlbacher O, Marks DS. 2017. Genetic variation in human drug-related genes. *Genome Med.* 9(1):117
8. Starita LM, Ahituv N, Dunham MJ, Kitzman JO, Roth FP, et al. 2017. Variant interpretation: functional assays to the rescue. *Am. J. Hum. Genet.* 101(3):315–25
9. Relling MV, Evans WE. 2015. Pharmacogenomics in the clinic. *Nature* 526(7573):343–50
10. Whirl-Carrillo M, McDonagh EM, Hebert JM, Gong L, Sangkuhl K, et al. 2012. Pharmacogenomics knowledge for personalized medicine. *Clin. Pharmacol. Ther.* 92(4):414–17
11. Relling MV, Klein TE. 2011. CPIC: Clinical Pharmacogenetics Implementation Consortium of the Pharmacogenomics Research Network. *Clin. Pharmacol. Ther.* 89(3):464–67
12. Russell LE, Schwarz UI. 2020. Variant discovery using next-generation sequencing and its future role in pharmacogenetics. *Pharmacogenomics* 21(7):471–86
13. Ingelman-Sundberg M, Mkrtchian S, Zhou Y, Lauschke VM. 2018. Integrating rare genetic variants into pharmacogenetic drug response predictions. *Hum. Genom.* 12(1):26
14. Caspar SM, Schneider T, Meienberg J, Matyas G. 2020. Added value of clinical sequencing: WGS-based profiling of pharmacogenes. *Int. J. Mol. Sci.* 21(7):2308
15. Weile J, Roth FP. 2018. Multiplexed assays of variant effects contribute to a growing genotype-phenotype atlas. *Hum. Genet.* 137(9):665–78
16. Kinney JB, McCandlish DM. 2019. Massively parallel assays and quantitative sequence-function relationships. *Annu. Rev. Genom. Hum. Genet.* 20:99–127
17. Chiasson M, Dunham MJ, Rettie AE, Fowler DM. 2019. Applying multiplex assays to understand variation in pharmacogenes. *Clin. Pharmacol. Ther.* 106(2):290–94
18. Zhang L, Sarangi V, Moon I, Yu J, Liu D, et al. 2020. CYP2C9 and CYP2C19: deep mutational scanning and functional characterization of genomic missense variants. *Clin. Transl. Sci.* 13(4):727–42
19. Gasperini M, Starita L, Shendure J. 2016. The power of multiplexed functional analysis of genetic variants. *Nat. Protoc.* 11(10):1782–87
20. Starita LM, Fields S. 2015. Deep mutational scanning: library construction, functional selection, and high-throughput sequencing. *Cold Spring Harb. Protoc.* 2015(8):777–80
21. Starita LM, Fields S. 2015. Deep mutational scanning: calculating enrichment scores for protein variants from DNA sequencing output files. *Cold Spring Harb. Protoc.* 2015(8):781–83

22. Fowler DM, Fields S. 2014. Deep mutational scanning: a new style of protein science. *Nat. Methods* 11(8):801–7
23. Matreyek KA, Starita LM, Stephany JJ, Martin B, Chiasson MA, et al. 2018. Multiplex assessment of protein variant abundance by massively parallel sequencing. *Nat. Genet.* 50(6):874–82
24. Laurent JM, Garge RK, Teufel AI, Wilke CO, Kachroo AH, Marcotte EM. 2020. Humanization of yeast genes with multiple human orthologs reveals functional divergence between paralogs. *PLOS Biol.* 18(5):e3000627
25. Sun S, Yang F, Tan G, Costanzo M, Oughtred R, et al. 2016. An extended set of yeast-based functional assays accurately identifies human disease mutations. *Genome Res.* 26(5):670–80
26. Chen H, Zhang Z, Jiang S, Li R, Li W, et al. 2020. New insights on human essential genes based on integrated analysis and the construction of the HEGIAP web-based platform. *Brief. Bioinform.* 21(4):1397–410
27. Boone C, Andrews BJ. 2015. The indispensable genome. *Science* 350(6264):1028–29
28. Findlay GM, Daza RM, Martin B, Zhang MD, Leith AP, et al. 2018. Accurate classification of BRCA1 variants with saturation genome editing. *Nature* 562(7726):217–22
29. Chiasson MA, Rollins NJ, Stephany JJ, Sitko KA, Matreyek KA, et al. 2020. Multiplexed measurement of variant abundance and activity reveals VKOR topology, active site and human variant impact. *eLife* 9:e58026
30. Kircher M, Xiong C, Martin B, Schubach M, Inoue F, et al. 2019. Saturation mutagenesis of twenty disease-associated regulatory elements at single base-pair resolution. *Nat. Commun.* 10(1):3583
31. Amorosi CJ, Chiasson MA, McDonald MG, Wong LH, Sitko KA, et al. 2021. Massively parallel characterization of CYP2C9 variant enzyme activity and abundance. *Am. J. Hum. Genet.* 108(9):1735–51
32. Gelman H, Dines JN, Berg J, Berger AH, Brnich S, et al. 2019. Recommendations for the collection and use of multiplexed functional data for clinical variant interpretation. *Genome Med.* 11(1):85
33. Coulthard S, Hogarth L. 2005. The thiopurines: an update. *Investig. New Drugs.* 23(6):523–32
34. Lennard L, Van Loon JA, Lilleyman JS, Weinshilboum RM. 1987. Thiopurine pharmacogenetics in leukemia: correlation of erythrocyte thiopurine methyltransferase activity and 6-thioguanine nucleotide concentrations. *Clin. Pharmacol. Ther.* 41(1):18–25
35. Szumlanski CL, Honchel R, Scott MC, Weinshilboum RM. 1992. Human liver thiopurine methyltransferase pharmacogenetics: biochemical properties, liver-erythrocyte correlation and presence of isozymes. *Pharmacogenetics* 2(4):148–59
36. Relling MV, Schwab M, Whirl-Carrillo M, Suarez-Kurtz G, Pui C-H, et al. 2019. Clinical Pharmacogenetics Implementation Consortium guideline for thiopurine dosing based on *TPMT* and *NUDT15* genotypes: 2018 update. *Clin. Pharmacol. Ther.* 105(5):1095–105
37. Wang L, Pelleymounter L, Weinshilboum R, Johnson JA, Hebert JM, et al. 2010. Very important pharmacogene summary: thiopurine S-methyltransferase. *Pharmacogenet. Genom.* 20(6):401–5
38. Katara P, Kuntal H. 2016. TPMT polymorphism: when shield becomes weakness. *Interdiscip. Sci.* 8(2):150–55
39. Schaeffeler E, Fischer C, Brockmeier D, Wernet D, Moerike K, et al. 2004. Comprehensive analysis of thiopurine *S*-methyltransferase phenotype-genotype correlation in a large population of German-Caucasians and identification of novel *TPMT* variants. *Pharmacogenetics* 14(7):407–17
40. Dean L. 2012. Azathioprine therapy and *TPMT* and *NUDT15* genotype. In *Medical Genetics Summaries*, ed. VM Pratt, SA Scott, M Pirmohamed, B Esquivel, MS Kane, et al. Bethesda, MD: Natl. Cent. Biotechnol. Inf.
41. Relling MV, Gardner EE, Sandborn WJ, Schmiegelow K, Pui C-H, et al. 2013. Clinical Pharmacogenetics Implementation Consortium guidelines for thiopurine methyltransferase genotype and thiopurine dosing: 2013 update. *Clin. Pharmacol. Ther.* 93(4):324–25
42. Tai HL, Krynetski EY, Schuetz EG, Yanishevski Y, Evans WE. 1997. Enhanced proteolysis of thiopurine S-methyltransferase (TPMT) encoded by mutant alleles in humans (*TPMT*3A*, *TPMT*2*): mechanisms for the genetic polymorphism of TPMT activity. *PNAS* 94(12):6444–49
43. Salavaggione OE, Wang L, Wiepert M, Yee VC, Weinshilboum RM. 2005. Thiopurine *S*-methyltransferase pharmacogenetics: variant allele functional and comparative genomics. *Pharmacogenet. Genom.* 15(11):801–15

44. Wu H, Horton JR, Battaile K, Allali-Hassani A, Martin F, et al. 2007. Structural basis of allele variation of human thiopurine-*S*-methyltransferase. *Proteins* 67(1):198–208
45. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, et al. 2016. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 536(7616):285–91
46. Appell ML, Berg J, Duley J, Evans WE, Kennedy MA, et al. 2013. Nomenclature for alleles of the thiopurine methyltransferase gene. *Pharmacogenet. Genom.* 23(4):242–48
47. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, et al. 2020. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* 581(7809):434–43
48. Suiter CC, Moriyama T, Matreyek KA, Yang W, Scaletti ER, et al. 2020. Massively parallel variant characterization identifies *NUDT15* alleles associated with thiopurine toxicity. *PNAS* 117(10):5394–5401
49. Yang JJ, Whirl-Carrillo M, Scott SA, Turner AJ, Schwab M, et al. 2019. Pharmacogene Variation Consortium gene introduction: *NUDT15*. *Clin. Pharmacol. Ther.* 105(5):1091–94
50. Moriyama T, Nishii R, Perez-Andreu V, Yang W, Klussmann FA, et al. 2016. *NUDT15* polymorphisms alter thiopurine metabolism and hematopoietic toxicity. *Nat. Genet.* 48(4):367–73
51. Yang JJ, Landier W, Yang W, Liu C, Hageman L, et al. 2015. Inherited *NUDT15* variant is a genetic determinant of mercaptopurine intolerance in children with acute lymphoblastic leukemia. *J. Clin. Oncol.* 33(11):1235–42
52. Cagiada M, Johansson KE, Valanciute A, Nielsen SV, Hartmann-Petersen R, et al. 2021. Understanding the origins of loss of protein function by analyzing the effects of thousands of variants on activity and abundance. *Mol. Biol. Evol.* 38(8):3235–46
53. Rost S, Fregin A, Ivaskevicius V, Conzelmann E, Hörtnagel K, et al. 2004. Mutations in *VKORC1* cause warfarin resistance and multiple coagulation factor deficiency type 2. *Nature* 427(6974):537–41
54. Li T, Chang C-Y, Jin D-Y, Lin P-J, Khvorova A, Stafford DW. 2004. Identification of the gene for vitamin K epoxide reductase. *Nature* 427(6974):541–44
55. Kane SP. 2020. *ClinCalc DrugStats database*. Database, ClinCalc. **http://www.clincalc.com/drugstats**
56. Owen RP, Gong L, Sagreiya H, Klein TE, Altman RB. 2010. VKORC1 pharmacogenomics summary. *Pharmacogenet. Genom.* 20(10):642–44
57. Duconge J, Cadilla CL, Windemuth A, Kocherla M, Gorowski K, et al. 2009. Prevalence of combinatorial CYP2C9 and VKORC1 genotypes in Puerto Ricans: implications for warfarin management in Hispanics. *Ethn. Dis.* 19(4):390–95
58. Rettie AE, Tai G. 2006. The pharmacogenomics of warfarin: closing in on personalized medicine. *Mol. Interv.* 6(4):223–27
59. Kamali F. 2006. Genetic influences on the response to warfarin. *Curr. Opin. Hematol.* 13(5):357–61
60. Wadelius M, Pirmohamed M. 2007. Pharmacogenetics of warfarin: current status and future challenges. *Pharmacogenomics J.* 7(2):99–111
61. Gage BF, Lesko LJ. 2008. Pharmacogenetics of warfarin: regulatory, scientific, and clinical issues. *J. Thromb. Thrombolysis.* 25(1):45–51
62. Czogalla KJ, Biswas A, Höning K, Hornung V, Liphardt K, et al. 2017. Warfarin and vitamin K compete for binding to Phe55 in human VKOR. *Nat. Struct. Mol. Biol.* 24(1):77–85
63. Shen G, Cui W, Zhang H, Zhou F, Huang W, et al. 2017. Warfarin traps human vitamin K epoxide reductase in an intermediate state during electron transfer. *Nat. Struct. Mol. Biol.* 24(1):69–76
64. Tie J-K, Jin D-Y, Tie K, Stafford DW. 2013. Evaluation of warfarin resistance using transcription activator-like effector nucleases-mediated vitamin K epoxide reductase knockout HEK293 cells. *J. Thromb. Haemost.* 11(8):1556–64
65. Gong IY, Schwarz UI, Crown N, Dresser GK, Lazo-Langner A, et al. 2011. Clinical and genetic determinants of warfarin pharmacokinetics and pharmacodynamics during treatment initiation. *PLOS ONE* 6(11):e27808
66. Johnson JA, Caudle KE, Gong L, Whirl-Carrillo M, Stein CM, et al. 2017. Clinical Pharmacogenetics Implementation Consortium (CPIC) guideline for pharmacogenetics-guided warfarin dosing: 2017 update. *Clin. Pharmacol. Ther.* 102(3):397–404
67. Landrum MJ, Chitipiralla S, Brown GR, Chen C, Gu B, et al. 2020. ClinVar: improvements to accessing data. *Nucleic Acids Res.* 48(D1):D835–44

68. Haque JA, McDonald MG, Kulman JD, Rettie AE. 2014. A cellular system for quantitation of vitamin K cycle activity: structure-activity effects on vitamin K antagonism by warfarin metabolites. *Blood* 123(4):582–89
69. Schulman S, Wang B, Li W, Rapoport TA. 2010. Vitamin K epoxide reductase prefers ER membrane-anchored thioredoxin-like redox partners. *PNAS* 107(34):15027–32
70. Wu S, Chen X, Jin D-Y, Stafford DW, Pedersen LG, Tie J-K. 2018. Warfarin and vitamin K epoxide reductase: a molecular accounting for observed inhibition. *Blood* 132(6):647–57
71. Tie J-K, Jin D-Y, Stafford DW. 2012. Human vitamin K epoxide reductase and its bacterial homologue have different membrane topologies and reaction mechanisms. *J. Biol. Chem.* 287(41):33945–55
72. Rost S, Pelz H-J, Menzel S, MacNicoll AD, León V, et al. 2009. Novel mutations in the VKORC1 gene of wild rats and mice—a response to 50 years of selection pressure by warfarin? *BMC Genet.* 10:4
73. Li W, Schulman S, Dutton RJ, Boyd D, Beckwith J, Rapoport TA. 2010. Structure of a bacterial homologue of vitamin K epoxide reductase. *Nature* 463(7280):507–12
74. Yuan H-Y, Chen J-J, Lee MTM, Wung J-C, Chen Y-F, et al. 2005. A novel functional VKORC1 promoter polymorphism is associated with inter-individual and inter-ethnic differences in warfarin sensitivity. *Hum. Mol. Genet.* 14(13):1745–51
75. Rettie AE, Jones JP. 2005. Clinical and toxicological relevance of CYP2C9: drug-drug interactions and pharmacogenetics. *Annu. Rev. Pharmacol. Toxicol.* 45:477–94
76. Theken KN, Lee CR, Gong L, Caudle KE, Formea CM, et al. 2020. Clinical Pharmacogenetics Implementation Consortium guideline (CPIC) for *CYP2C9* and nonsteroidal anti-inflammatory drugs. *Clin. Pharmacol. Ther.* 108(2):191–200
77. Karnes JH, Rettie AE, Somogyi AA, Huddart R, Fohner AE, et al. 2021. Clinical Pharmacogenetics Implementation Consortium (CPIC) guideline for *CYP2C9* and *HLA-B* genotypes and phenytoin dosing: 2020 update. *Clin. Pharmacol. Ther.* 109(2):302–9
78. Higashi MK, Veenstra DL, Kondo LM, Wittkowsky AK, Srinouanprachanh SL, et al. 2002. Association between CYP2C9 genetic variants and anticoagulation-related outcomes during warfarin therapy. *JAMA* 287(13):1690–98
79. Pirmohamed M, Kamali F, Daly AK, Wadelius M. 2015. Oral anticoagulation: a critique of recent advances and controversies. *Trends Pharmacol. Sci.* 36(3):153–63
80. Lee CR, Goldstein JA, Pieper JA. 2002. Cytochrome P450 2C9 polymorphisms: a comprehensive review of the in-vitro and human data. *Pharmacogenetics* 12(3):251–63
81. Huang M, Joensson HN, Nielsen J. 2018. High-throughput microfluidics for the screening of yeast libraries. *Methods Mol. Biol.* 1671:307–17
82. Woronoff G, El Harrak A, Mayot E, Schicke O, Miller OJ, et al. 2011. New generation of amino coumarin methyl sulfonate-based fluorogenic substrates for amidase assays in droplet-based microfluidic applications. *Anal. Chem.* 83(8):2852–57
83. Zanger UM, Schwab M. 2013. Cytochrome P450 enzymes in drug metabolism: regulation of gene expression, enzyme activities, and impact of genetic variation. *Pharmacol. Ther.* 138(1):103–41
84. Mega JL, Close SL, Wiviott SD, Shen L, Hockett RD, et al. 2009. Cytochrome p-450 polymorphisms and response to clopidogrel. *N. Engl. J. Med.* 360(4):354–62
85. Sim SC, Risinger C, Dahl M-L, Aklillu E, Christensen M, et al. 2006. A common novel *CYP2C19* gene variant causes ultrarapid drug metabolism relevant for the drug response to proton pump inhibitors and antidepressants. *Clin. Pharmacol. Ther.* 79(1):103–13
86. Zabalza M, Subirana I, Sala J, Lluis-Ganella C, Lucas G, et al. 2012. Meta-analyses of the association between cytochrome CYP2C19 loss- and gain-of-function polymorphisms and cardiovascular outcomes in patients with coronary artery disease treated with clopidogrel. *Heart* 98(2):100–108
87. Litonjua AA, Gong L, Duan QL, Shin J, Moore MJ, et al. 2010. Very important pharmacogene summary *ADRB2*. *Pharmacogenet. Genom.* 20(1):64–69
88. García-Menaya JM, Cordobés-Durán C, García-Martín E, Agúndez JAG. 2019. Pharmacogenetic factors affecting asthma treatment response. Potential implications for drug therapy. *Front. Pharmacol.* 10:520
89. Hizawa N. 2011. Pharmacogenetics of β_2-agonists. *Allergol. Int.* 60(3):239–46

90. Jones EM, Lubock NB, Venkatakrishnan AJ, Wang J, Tseng AM, et al. 2020. Structural and functional characterization of G protein-coupled receptors with deep mutational scanning. *eLife* 9:623108
91. Pacanowski MA, Johnson JA. 2007. PharmGKB submission update: IX. ADRB1 gene summary. *Pharmacol. Rev.* 59(1):2–4
92. Hauser AS, Attwood MM, Rask-Andersen M, Schiöth HB, Gloriam DE. 2017. Trends in GPCR drug discovery: new agents, targets and indications. *Nat. Rev. Drug Discov.* 16(12):829–42
93. Lee HH, Ho RH. 2017. Interindividual and interethnic variability in drug disposition: polymorphisms in organic anion transporting polypeptide 1B1 (OATP1B1; *SLCO1B1*). *Br. J. Clin. Pharmacol.* 83(6):1176–84
94. Voora D, Shah SH, Spasojevic I, Ali S, Reed CR, et al. 2009. The *SLCO1B1*5* genetic variant is associated with statin-induced side effects. *J. Am. Coll. Cardiol.* 54(17):1609–16
95. Ramsey LB, Johnson SG, Caudle KE, Haidar CE, Voora D, et al. 2014. The clinical pharmacogenetics implementation consortium guideline for *SLCO1B1* and simvastatin-induced myopathy: 2014 update. *Clin. Pharmacol. Ther.* 96(4):423–28
96. Tirona RG, Leake BF, Merino G, Kim RB. 2001. Polymorphisms in *OATP-C*: identification of multiple allelic variants associated with altered transport activity among European- and African-Americans. *J. Biol. Chem.* 276(38):35669–75
97. Wang X, Chen J, Xu S, Ni C, Fang Z, Hong M. 2019. Amino-terminal region of human organic anion transporting polypeptide 1B1 dictates transporter stability and substrate interaction. *Toxicol. Appl. Pharmacol.* 378:114642
98. Yao J, Hong W, Huang J, Zhan K, Huang H, Hong M. 2012. *N*-glycosylation dictates proper processing of organic anion transporting polypeptide 1B1. *PLOS ONE* 7(12):e52563
99. Zhang L, Sarangi V, Ho M-F, Moon I, Kalari KR, et al. 2021. *SLCO1B1*: application and limitations of deep mutational scanning for genomic missense variant function. *Drug Metab. Dispos.* 49(5):395–404
100. Kameyama Y, Yamashita K, Kobayashi K, Hosokawa M, Chiba K. 2005. Functional characterization of *SLCO1B1* (*OATP-C*) variants, *SLCO1B1*5*, *SLCO1B1*15* and *SLCO1B1*15+C1007G*, by using transient expression systems of HeLa and HEK293 cells. *Pharmacogenet. Genom.* 15(7):513–22
101. Santos PCJL, Pereira AC. 2015. Type of *LDLR* mutation and the pharmacogenetics of familial hypercholesterolemia treatment. *Pharmacogenomics* 16(15):1743–50
102. Kolovou GD, Kostakou PM, Anagnostopoulou KK. 2011. Familial hypercholesterolemia and triglyceride metabolism. *Int. J. Cardiol.* 147(3):349–58
103. Scholtz CL, Peeters AV, Hoogendijk CF, Thiart R, de Villiers JN, et al. 1999. Mutation −59c→t in repeat 2 of the LDL receptor promoter: reduction in transcriptional activity and possible allelic interaction in a South African family with familial hypercholesterolaemia. *Hum. Mol. Genet.* 8(11):2025–30
104. Mozas P, Galetto R, Albajar M, Ros E, Pocoví M, Rodríguez-Rey JC. 2002. A mutation (−49C>T) in the promoter of the low density lipoprotein receptor gene associated with familial hypercholesterolemia. *J. Lipid Res.* 43(1):13–18
105. De Castro-Orós I, Pampín S, Bolado-Carrancio A, De Cubas A, Palacios L, et al. 2011. Functional analysis of LDLR promoter and 5′ UTR mutations in subjects with clinical diagnosis of familial hypercholesterolemia. *Hum. Mutat.* 32(8):868–72
106. Reeskamp LF, Hartgers ML, Peter J, Dallinga-Thie GM, Zuurbier L, et al. 2018. A deep intronic variant in *LDLR* in familial hypercholesterolemia. *Circ. Genom. Precis. Med.* 11(12):e002385
107. Schwarz UI, Meyer zu Schwabedissen HE, Tirona RG, Suzuki A, Leake BF, et al. 2011. Identification of novel functional organic anion-transporting polypeptide 1B3 polymorphisms and assessment of substrate specificity. *Pharmacogenet. Genom.* 21(3):103–14
108. Kuang D, Weile J, Kishore N, Rubin AF, Fields S, et al. 2021. MaveRegistry: a collaboration platform for multiplexed assays of variant effect. *Bioinformatics* 2021:btab215
109. Starita LM, Young DL, Islam M, Kitzman JO, Gullingsrud J, et al. 2015. Massively parallel functional analysis of BRCA1 RING domain variants. *Genetics* 200(2):413–22
110. Starita LM, Islam MM, Banerjee T, Adamovich AI, Gullingsrud J, et al. 2018. A multiplex homology-directed DNA repair assay reveals the impact of more than 1,000 BRCA1 missense substitution variants on protein function. *Am. J. Hum. Genet.* 103(4):498–508

111. Muir AJ, Gong L, Johnson SG, Lee MTM, Williams MS, et al. 2014. Clinical Pharmacogenetics Implementation Consortium (CPIC) guidelines for *IFNL3* (*IL28B*) genotype and PEG interferon-α-based regimens. *Clin. Pharmacol. Ther.* 95(2):141–46
112. Nies AT, Niemi M, Burk O, Winter S, Zanger UM, et al. 2013. Genetics is a major determinant of expression of the human hepatic uptake transporter OATP1B1, but not of OATP1B3 and OATP2B1. *Genome Med.* 5(1):1
113. Park J, Lee SY, Baik SY, Park CH, Yoon JH, et al. 2020. Gene-wise burden of coding variants correlates to noncoding pharmacogenetic risk variants. *Int. J. Mol. Sci.* 21(9):3091
114. Santos M, Niemi M, Hiratsuka M, Kumondai M, Ingelman-Sundberg M, et al. 2018. Novel copy-number variations in pharmacogenes contribute to interindividual differences in drug pharmacokinetics. *Genet. Med.* 20(6):622–29
115. Lauschke VM, Ingelman-Sundberg M. 2020. Emerging strategies to bridge the gap between pharmacogenomic research and its clinical implementation. *NPJ Genomic Med.* 5(1):9
116. Del Tredici AL, Malhotra A, Dedek M, Espin F, Roach D, et al. 2018. Frequency of *CYP2D6* alleles including structural variants in the United States. *Front. Pharmacol.* 9:305
117. Morelli A, Cabezas Y, Mills LJ, Seelig B. 2017. Extensive libraries of gene truncation variants generated by in vitro transposition. *Nucleic Acids Res.* 45(10):e78
118. Guo X, Chavez A, Tung A, Chan Y, Kaas C, et al. 2018. High-throughput creation and functional profiling of DNA sequence variant libraries using CRISPR/Cas9 in yeast. *Nat. Biotechnol.* 36(6):540–46
119. Cheung R, Insigne KD, Yao D, Burghard CP, Wang J, et al. 2019. A multiplexed assay for exon recognition reveals that an unappreciated fraction of rare genetic variants cause large-effect splicing disruptions. *Mol. Cell.* 73(1):183–94.e8
120. Esposito D, Weile J, Shendure J, Starita LM, Papenfuss AT, et al. 2019. MaveDB: an open-source platform to distribute and interpret data from multiplexed assays of variant effect. *Genome Biol.* 20(1):223
121. Masso M. 2020. Functional analysis of BRCA1 RING domain variants: computationally derived structural data can improve upon experimental features for training predictive models. *Integr. Biol.* 12(9):233–39
122. Masso M, Bansal A, Bansal A, Henderson A. 2020. Structure-based functional analysis of BRCA1 RING domain variants: concordance of computational mutagenesis, experimental assay, and clinical data. *Biophys. Chem.* 266:106442
123. Gray VE, Hause RJ, Luebeck J, Shendure J, Fowler DM. 2018. Quantitative missense variant effect prediction using large-scale mutagenesis data. *Cell Syst.* 6(1):116–24.e3
124. Weile J, Sun S, Cote AG, Knapp J, Verby M, et al. 2017. A framework for exhaustively mapping functional missense variants. *Mol. Syst. Biol.* 13(12):957
125. Tafazoli A, Wawrusiewicz-Kurylonek N, Posmyk R, Miltyk W. 2020. Pharmacogenomics, how to deal with different types of variants in next generation sequencing data in the personalized medicine area. *J. Clin. Med. Res.* 10(1):34
126. Richards S, Aziz N, Bale S, Bick D, Das S, et al. 2015. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* 17(5):405–24
127. Abou Tayoun AN, Pesaran T, DiStefano MT, Oza A, Rehm HL, et al. 2018. Recommendations for interpreting the loss of function PVS1 ACMG/AMP variant criterion. *Hum. Mutat.* 39(11):1517–24
128. Brnich SE, Abou Tayoun AN, Couch FJ, Cutting GR, Greenblatt MS, et al. 2019. Recommendations for application of the functional evidence PS3/BS3 criterion using the ACMG/AMP sequence variant interpretation framework. *Genome Med.* 12(1):3

Annual Review of Pharmacology and Toxicology

Volume 62, 2022

Contents