Pharmacogenetic Allele Nomenclature: International Workgroup Recommendations for Test Result Reporting

This article provides nomenclature recommendations developed by an international workgroup to increase transparency and standardization of pharmacogenetic (PGx) result reporting. Presently, sequence variants identified by PGx tests are described using different nomenclature systems. In addition, PGx analysis may detect different sets of variants for each gene, which can affect interpretation of results. This practice has caused confusion and may thereby impede the adoption of clinical PGx testing. Standardization is critical to move PGx forward.

Individuals vary considerably in their response to medications. Some patients show a substantial therapeutic response to a given drug, while others may not. In addition, certain patients may require considerably higher or lower doses of a drug to achieve maximum benefit or to avoid an adverse reaction. Many factors contribute to this variability, including sex, age, diet, environmental exposures (e.g., toxic chemicals or cigarette smoke), inflammation-induced phenocconversion, epigenetic

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention or the Agency for Toxic Substances and Disease Registry.

The US Food and Drug Administration (FDA) Center for Drug Evaluation and Research and the FDA Center for Radiological Health and Devices participated in this working group and contributed to drafting of the manuscript. This article reflects the views of the authors and should not be construed to represent the FDA’s views or policies.

Received 3 September 2015; accepted 14 October 2015; advance online publication 00 Month 2015. doi:10.1002/cpt.280
signatures, and drug interactions. The recognition that part of this variation in patient response may be genetic, and therefore potentially predictable, led to the development of numerous pharmacogenetic (PGx) tests to assess the presence or absence of known genetic variants to help predict an individual's response to drugs. The results of these tests can help healthcare providers to select the most effective drug(s) and doses for a patient, inform drug development, and/or stratify participants in clinical trials.

Research studies have identified over 1,000 human genes that may affect drug response. Clinically relevant PGx interactions typically involve genes related to absorption, distribution, metabolism, and excretion (ADME) of a drug, or genes that encode drug targets and other proteins involved in the drug’s mechanism of action. For example, non-ADME genes, such as CFTR and VKORC1, are relevant to ivacaftor and warfarin response, respectively, and may be tested to guide therapeutic decision-making. In addition, genes encoding the human leukocyte antigens (HLAs) are commonly tested for the variant HLA-B*57:01 to predict the likelihood of hypersensitivity reactions for abacavir in patients seeking human immunodeficiency virus antiretroviral therapy.

In 2015, about 150 different drugs that are approved by the US Food and Drug Administration (FDA) include pharmacogenetic information on the label, and only a few of them have recommendations for PGx testing. Similarly, 155 drugs have PGx information in their Summary of Product Characteristics as defined by the European Medicines Agency. Both the FDA label and the European Medicines Agency Summary of Product Characteristics contain information ranging from references to pharmacokinetic genes or drug targets to requirements for genetic testing. However, labels mentioning genetic testing are rare and are often associated with drugs for cancer treatment (e.g., EGFR/afatinib, ALK/crizotinib, and KRAS/panitumumab). Other examples of established gene/drug associations include HLA-B/abacavir, HLA-B/carbamazepine, CYP2C19/dlodigrel, CYP2D6/codeine, POLG/valproate, G6PD/rasburicase, TPMT/thiopurines, and others. Although many of these drugs are widely prescribed, for most of them PGx testing has yet to become common practice.

At the current time, clinical laboratories offer testing for a number of PGx genes, including CYP2C9, CYP2C19, CYP2D6, CYP3A4, CYP3A5, DPYD, HLA-B*57:01, SLCO1B1, VKORC1, and TPMT. The National Institutes of Health (NIH) Genetic Testing Registry lists over 280 PGx tests for about 160 drug responses. Although some laboratories that offer these tests have reported good success with respect to feasibility and acceptance of genetic testing, others have suggested that PGx tests are underutilized. Lack of acceptance can arise when clinicians do not have sufficient knowledge of PGx tests, are unable to obtain the test results within an appropriate turnaround time, are not sure whether there is sufficient evidence to support the use or reimbursement of PGx tests, or cannot interpret and translate the genotype information into clinical actions.

Ideally, test results and interpretations should be consistent regardless of which clinical laboratory performs the analysis. However, laboratories differ with respect to the PGx variants and haplotypes that are tested, and also the manner in which the results are interpreted and communicated to the prescriber. This variability can have a significant impact on clinical decision-making, particularly because many genes involved in ADME and other aspects of pharmacology and toxicology can vary considerably among individuals and populations. As such, standardized nomenclature and transparency of variants tested and used for PGx haplotype definitions are needed to advance the adoption of pharmacogenetic testing by clinicians and to ensure that prescribers have the essential information for appropriate treatment decisions.

Overview of nomenclature for PGx genes
Activity or function of enzymes, transporters, or drug targets (pharmacogenetic phenotype) can be predicted by testing for one or more known sequence variants in PGx genes. Variants or combinations of variants in a gene that are linked together on a single chromosome define haplotypes. The terms haplotype, allele, and allelic variation are often used interchangeably. Results of PGx tests are commonly reported as diplotypes (or haplotype pairs) because human genes are present in two copies, except for genes located on the nonhomologous parts of the X and Y chromosomes in men. However, haplotypes or diplotypes are typically assigned based on genotypes of tested genetic variants, and default assignments may be applied depending on whether these variants are detected. The summarization of observed variants into alleles/haplotypes facilitates the association of diplotypes with predicted phenotypes. Many PGx genes have greater clinical relevance when low-function variants are viewed as conferring recessive rather than codominant or dominant phenotypic effects, and prescribing guidelines often differ substantially for individuals carrying two variant alleles compared to those carrying only one dysfunctional allele.

Variants found during DNA sequence analysis or other types of genotyping tests used to diagnose inherited or somatic disorders are named using the Human Genome Variation Society (HGVS) nomenclature. The HGVS nomenclature had been designed specifically for use in clinical diagnostics and is currently the standard worldwide. Use of HGVS is recommended for clinical diagnostic reporting. This nomenclature system describes variants with respect to a reference sequence, making the genomic position of the variant and the changes to the DNA, RNA, and protein sequences comprehensible and less ambiguous to current and future users of the information. The HGVS nomenclature does not specify a specific reference sequence, thus, the same variant could be described using different reference sequences, which might cause confusion.

Unlike other genes, a variety of nomenclature systems have been developed to describe allelic variation and haplotypes of ADME genes. The most common is the “star” (∗) system, which was implemented in the 1990s and has been widely adopted in the field. In most cases, the “star” nomenclature indicates the default reference (wild type or fully functional) allele or haplotype, whereas
other designations (e.g., ‘2’ or ‘3’) define haplotypes carrying one or more variants. The ‘1’ allele definition is usually based on the subpopulation in which the gene was initially studied, and may not necessarily indicate the most common allele in all populations. In some cases, ‘1’ is not the reference allele; for example, NAT2*4 is the reference allele for the NAT2 gene as it is the most common functional allele across human populations.31,32

Laboratories report a variant if one or more allele-defining sequence variations are found and default to a reference allele (often, but not always ‘1’) assignment in their absence. It should be noted that the reference allele designation is assigned depending on which variants were assayed and consequently excluded, but does not consider variants not included in the assay (i.e., an assigned ‘1’ carrier may still have variants that the test was not designed to detect). In other words, a negative result for the alleles interrogated by the assay (and the designation of the ‘1’ haplotype or reference allele) does not exclude the possibility that other dysfunctional alleles may be present. The probability that a ‘1’ or reference allele default assignment is correct increases with the number of relevant sequence variations tested.

Lists of haplotypes and nomenclature for PGx genes can be found on a variety of gene or gene-family specific websites that are usually maintained by specific nomenclature committees (Table 1). Information about PGx haplotypes is also available through more comprehensive sites, including the Pharmacogenomics Knowledge Base (PharmGKB).8 For many PGx genes (e.g., VKORC1), there are no nomenclature committees. In some instances, such as SLC01B133 and ABCB1,34 tables summarizing variation have been published, but are not systematically maintained. Furthermore, variants of many PGx genes, including VKORC1, may be reported using more than one nomenclature system.35,36

Rationale for developing a standard PGx nomenclature

A major obstacle for the use of PGx information in clinical practice is the complexity of nomenclature systems and test designs, which can cause discrepancies in the predicted phenotype that is inferred from the genotype determined by the genetic test. Nomenclature differences contribute to difficulties reconciling genotype results for the same sample across laboratories37,38 and impede data analysis. Differences in test design can also lead to discordant diplotype results and ultimately incorrect phenotype predictions (Robert Freimuth, personal communication). As such, a standardized PGx nomenclature system that clearly describes the variants identified as well as increased transparency about the test design are crucial to advance the adoption of PGx testing by clinicians.

Comparing PGx testing results from different laboratories can be challenging. For example, the Centers for Disease Control and Prevention’s Genetic Testing Reference Material Coordination Program (GeT-RM)39 conducted projects to characterize genomic DNA samples from the Coriell Cell Repositories which can be used as reference materials for clinical PGx genetic testing.37,38 During the most recent study, nine volunteer clinical, research, and commercial laboratories were provided with blinded genomic DNA samples, which were tested for a number of PGx genes using a variety of different methods, including single nucleotide variant (SNV) genotyping, copy number variant (CNV) assessment, and DNA sequence analysis.

The results of the GeT-RM study illustrate many inconsistencies because of different nomenclature systems and PGx test designs. The participating laboratories used varying terms to describe SNV genotype results for 28 PGx genes. The star nomenclature system was used by some laboratories to describe alleles/haplotypes of many of the tested genes, whereas other laboratories utilized the predicted amino acid change or various other notations and/or nomenclatures. In many cases, laboratories described the detected sequence variations using several different nomenclature systems for the same genes (e.g., VKORC1; Table 2).

Furthermore, the GeT-RM study found inconsistencies in reporting of genotype results. Some of the DNA sequencing assays made genotype calls from the coding strand while others made calls based on the noncoding strand using the Human Genome Reference Assembly as a guide. One test did not use the Human Genome Reference Assembly, but instead used sequences representing the major alleles identified during the HapMap project (Figure 1). These differences may not affect the actual results, but could impact the way in which results are represented and the depiction of variants in databases and scientific literature. The tests performed in the GeT-RM study also showed little consistency in design. No two of the seven test panels detected the same set of SNVs/haplotypes for any of the 28 PGx genes.

Table 1 Examples of public PGx gene or gene family haplotype databases

<table>
<thead>
<tr>
<th>Gene or gene family</th>
<th>Resource</th>
<th>URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cytochrome P450 enzymes</td>
<td>Human CYP allele nomenclature database(^a)</td>
<td><a href="http://www.cypalleles.ki.se/">http://www.cypalleles.ki.se/</a></td>
</tr>
<tr>
<td>Cytochrome P450 enzymes</td>
<td>SuperCYP</td>
<td><a href="http://bioinformatics.charite.de/supercyp/">http://bioinformatics.charite.de/supercyp/</a></td>
</tr>
<tr>
<td>UGT</td>
<td>UGT nomenclature</td>
<td><a href="http://www.pharmacogenomics.pha.ulaval.ca/cms/ugt_alleles/">http://www.pharmacogenomics.pha.ulaval.ca/cms/ugt_alleles/</a></td>
</tr>
<tr>
<td>TPMT</td>
<td>TPMT nomenclature</td>
<td><a href="http://www.imh.liu.se/tpmtalleles/?l=en">http://www.imh.liu.se/tpmtalleles/?l=en</a></td>
</tr>
<tr>
<td>NAT</td>
<td>NAT allele nomenclature database</td>
<td><a href="http://nat.mbg.duth.gr">http://nat.mbg.duth.gr</a></td>
</tr>
<tr>
<td>HLA</td>
<td>HLA nomenclature</td>
<td><a href="http://hla.alleles.org/nomenclature/index.html">http://hla.alleles.org/nomenclature/index.html</a></td>
</tr>
</tbody>
</table>

CYP, cytochrome P450; HLA, human leukocyte antigens; NAT, Arylamine N-acetyltransferase; TPMT, Thiopurine methyltransferase; UGT, UDP-glucuronosyltransferase.

studied. Different haplotype calls were often made for the same allele in a given sample, because not all tests were designed to detect the same haplotypes, as illustrated in Table 2. In some cases, these differences in test design can cause incorrect genotype assignments and phenotype prediction. As exemplified in Figure 2, the CYP2C19 phenotype may not be accurately predicted when only a subset of relevant SNVs is tested. An intermediate metabolizer would be misclassified as an ultra-rapid or extensive metabolizer if rs28399504 (NG_008384.2:g.5001A>G, NM_000769.2:c.1A>T) is not tested.

Inconsistencies in nomenclature and design of assay panels create unnecessary complexity and make it difficult to compare data from different genotyping platforms. This has been observed in clinical proficiency testing surveys. Participants in the College of American Pathologists Pharmacogenetic Proficiency Survey often reported inaccurate genotypes and phenotypes because of differences in test design. For example, one of the CYP2D6 assays did not include any of the SNVs commonly used to identify the *4I allele. The results from this test could cause this allele to be classified as normal function (‘*1’) rather than decreased function (‘*4I’). In addition, approximately half of the tests in the study were not designed to detect the increased function allele CYP2C19*17, and incorrectly called the proficiency testing samples CYP2C19*1 (normal function).

Discrepancies may also arise because of the variable inclusion of methods to assess CNVs. For example, tests that do not interrogate structural variants of CYP2D6 may report an individual with a CYP2D6*1/*2 haplotype as an extensive metabolizer.
whereas another assay, that includes copy number assessment, may report the same genotype as an ultra-rapid metabolizer if a gene duplication is detected (e.g., CYP2D6*1/*2xN or *1xN/*2).

These differences make it very difficult not only to understand genotype results, but also to compare results generated by different laboratories and/or tests. Such discrepancies may cause confusion, misinterpretation, and incorrect results to be communicated; thus, it is clear that standardized nomenclature, transparency of variants tested, and unambiguous haplotype definitions are needed.

Discrepancies in PGx testing and reporting may also cause other wide-ranging effects. For example, inconsistencies in nomenclature or test design could affect clinicians’ ability to understand results from different laboratories and in the scientific literature, and may hamper decision-making. Correct and unambiguous results are also important for the patient’s medical record, as these results may follow a patient for a lifetime. Payers often consider genetic testing to be a singular event that will not be repeated for the same gene as they usually do not expect the genetic result to change with time. This can affect test reimbursement if tests need to be reordered because the results cannot be interpreted or used for subsequent drug selection/dosing. These barriers hinder clinical adoption of PGx testing. Standardization of nomenclature is also critical for the accurate accumulation of data in clinical databases, such as ClinVar, and gene variant databases, such as the Leiden Open Variation Database, and PharmGKB.

To date, most PGx tests have evaluated the presence or absence of a defined set of known variants. However, next generation sequencing is becoming more common in clinical and research laboratories and is also being applied to PGx testing. Sequence analysis will identify rare and/or novel variants with unknown or uncertain functions that complicate not only allele designation and genotype calling, but also the prediction of phenotype. Recent results from numerous whole genome and whole exome sequencing efforts (e.g., the NHLBI GO Exome Sequencing Project) revealed the presence of many rare variant alleles in genes relevant to drug metabolism and transport, often missense in nature, which are not included in the current PGx databases. In addition, many more rare variants are expected to be identified during ongoing large population sequencing studies, including the 100,000 Genomes project in the United Kingdom and the 1-Million-Genomes Project in the United States. Consideration of these rare variants will be required for the advancement of precision medicine initiatives. It will be difficult to develop new haplotype designations for these recently identified variants using the existing star or other nomenclature systems, therefore, it may be desirable to modify or discontinue the use of star allele nomenclature and instead describe PGx variants with the same naming conventions and systems used for other genes.

Figure 2  CYP2C19*2, *4, and *17 alleles are used to illustrate how metabolizer phenotypes will be miscalled if CYP2C19*2 and *17, but not CYP2C19*4, are tested. Four possible scenarios are shown (Cases A–D). Panel (a) shows the phenotypes based on testing the key SNVs for CYP2C19*2 and *17. Panel (b) shows the phenotypes based on testing the key variants for CYP2C19*2, *4, and *17. Phenotypes shown in red in Panel (b) would have been miscalled if rs28399504 was present in the sample, but not included in the assay. These alleles are used as examples, demonstrating that all alleles affecting CYP2C19 functionality should be interrogated in order to accurately call phenotypes. EM, extensive metabolizer; IM, intermediate metabolizer; PM, poor metabolizer; UM, ultra-rapid metabolizer.
that DNA sequencing of PGx-associated genes will eventually become the standard method for genotype determination, which makes the creation of a standard variant naming format even more critical.

This article describes consensus recommendations from an international workgroup composed of a variety of stakeholders to standardize the description and reporting of PGx variants. The group recommends the use of HGVS nomenclature to describe PGx variants and makes suggestions regarding the use of reference sequences, reference SNP cluster identifiers (rs IDs), and clinical reporting of variants and test descriptions. These recommendations are applicable to stakeholders, including clinical laboratories and researchers who generate and report the results of PGx testing.

**CONSENSUS RECOMMENDATIONS FOR STANDARDIZATION OF PGx (ADME) NOMENCLATURE**

To address the standardization of PGx nomenclature, the Centers for Disease Control and Prevention organized an international workgroup to review current PGx nomenclature practices used to describe allelic variation of ADME PGx genes and facilitate translation between different nomenclature systems. Non-ADME genes, such as \( HLA \), were excluded from the discussions. Ideas for mechanisms to increase the transparency of the PGx test design were also discussed. Workgroup participants (summarized in Table 3) included a wide variety of stakeholders, many of whom also hold membership in relevant professional organizations and groups.

The workgroup held a series of conference calls and developed consensus recommendations for clinical and research laboratories to standardize PGx nomenclature and test result reporting. The workgroup discussed problems and needs in the following five areas: (1) problems and limitations of the current PGx nomenclature systems; (2) use of HGVS nomenclature to standardize and provide unambiguous descriptions of PGx variants; (3) development of tools to assist transition from current PGx nomenclatures to HGVS; (4) ideas to standardize and archive description of PGx assays in laboratory reports and electronic medical records; and (5) development of a standardized panel of variants that should be included in PGx test panels. The workgroup decided to develop recommendations for the description and reporting of variants that could be applied to all PGx genes. Although topics related to the prediction and reporting of phenotype to physicians were not discussed, the issues addressed by the workgroup were considered relevant to the current ability of clinical laboratories to accurately infer phenotypes from PGx test results. Consensus was determined by oral agreement by the majority of workgroup members.

**Workgroup recommendations**

Current PGx nomenclature systems have been developed by domain experts, are widely accepted in PGx research communities, and have been commonly reported in the scientific literature. The workgroup agreed that the use of different nomenclature systems to describe PGx variants hinders communication and is not intuitive to clinicians, patients, or researchers outside of the field. The group also acknowledged that PGx variants would be better reported in the context of the human genome assembly (as is the case with most non-PGx genes), especially as next generation sequencing applications are becoming more common. Because the current nomenclature is widely used in the PGx community, the workgroup did not recommend creating new nomenclature systems or abandoning current ones. Instead, they preferred to facilitate harmonization by clarifying sequence definitions of alleles. The workgroup recommended utilization of the HGVS nomenclature for describing variants and haplotypes, and requiring transparency when reporting test results and describing underlying test methodologies. The workgroup suggested that for continuity and consistency with the literature, laboratories could also report star (\( * \)) alleles, or other legacy nomenclature, possibly as a side note on a report.

The following nine recommendations (summarized in Table 4) take advantage of universally recognized and utilized systems for describing allelic variation.

**Recommendations for naming sequence variants (recommendations 1–4)**

(1) Use the HUGO Gene Nomenclature Committee nomenclature to specify the gene. (2) Report variants using HGVS variant nomenclature. (3) Use a Locus Reference Genomic (LRG), RefSeqGene, and/or a specific Human Genome Reference Assembly as a reference sequence. Both the reference sequence accession number and version number of the sequence should be indicated. (4) Report rs IDs from dbSNP, when available.

The workgroup recommended utilizing widely accepted systems that are already in place for human genes, in particular the HUGO Gene Nomenclature Committee gene nomenclature and HGVS nomenclature and the Human Genome Reference Assembly. This will allow for transparent reporting of PGx variants to the clinical and research communities, facilitating incorporation of the data into existing human variation databases, such as ClinVar and Leiden Open Variation Database in a standardized manner. Use of fully qualified (reference sequence accession and version number indicated) HGVS nomenclature is recommended for submission of data to ClinVar as well as other databases.

The workgroup discussed ways to facilitate transition to the HGVS format from other nomenclature systems. The HGVS system requires description of variants relative to a reference sequence, which should be selected based on ability to support explicit (rather than inferred) representation, stable public access, comprehensiveness, and ease of use. Genomic sequences are recommended because intronic locations can be represented unambiguously. However, the exact genomic coordinates of such sequences are frequently updated as new information about the human genome is incorporated into the Human Genome Reference Assembly. These changes are versioned, and may alter the HGVS name of a variant, which could cause confusion when results are compared over time. In order to reduce such misunderstandings, laboratories can include the latest HGVS nomenclature and the familiar or alternative names for the same variant in parenthesis.
Table 3  Groups, organizations, and stakeholders represented by the workgroup members

Pharmacogenetics community:
- Pharmacogenomics Knowledge Base (PharmGKB)
- Pharmacogenomics Research Network (PGRN)
- Clinical Pharmacogenetics Implementation Consortium (CPIC)\(^a\)
- European Pharmacogenetic Implementation Consortium (Eu-PIC)
- Ubiquitous Pharmacogenomics (U-PGx)
- European Society for Pharmacogenomics and Personalized Therapy (ESPT)
- International Federation for Clinical Chemistry (IFCC) Task Force –Pharmacogenetics
- ClinGen PGx Working Group

Regulatory/governmental agencies:
- Food and Drug Administration (FDA)\(^b\)
- European Medicine Agency (EMA) PGx Working Group
- College of American Pathologists (CAP)
- Centers for Disease Control and Prevention (CDC)
- National Institutes of Health (NIH)

Genetic nomenclature committees:
- HUGO Gene Nomenclature Committee (HGNC)
- Human Genome Variation Society (HGVS)

Gene variant databases:
- National Center for Biotechnology Information (NCBI)
- ClinVar
- Genetic Testing Registry (GTR)
- RefSeqGene
- MedGen
- European Bioinformatics Institute (EBI)
- Locus Reference Genomic (LRG)
- Leiden Open Variation Database (LOVD)

PGx gene-specific nomenclature committees and databases:
- Human Cytochrome P450 (CYP) Allele Nomenclature Database
- UDP-Glucuronosyltransferase (UGT) Alleles Nomenclature Page
- Thiopurine methyltransferase (TPMT) Nomenclature Committee
- Arylamine N-acetyltransferase (NAT) Gene Nomenclature Committee

Professional societies and standards development organizations:
- International Federation for Clinical Chemistry (IFCC)
- Association for Molecular Pathology (AMP)
- HL7 Clinical Genomics Working Group

PGx test developers

Clinical and research laboratories


\(^b\)FDA supports efforts toward PGx nomenclature standardization but the use of standardized nomenclature is not currently a regulatory requirement for clearance or approval of in vitro diagnostic tests.

The workgroup recognized that the genomic coordinates of the Human Genome Reference Assembly can be long, tedious, and require noting the accession version number. Because of these challenges, the group recommended that HGVS notations could be made using an LRG\(^{50,51}\) or RefSeqGene sequence\(^{52}\) as the reference. LRG sequences are stable genomic reference sequences of clinically important genes. LRG sequences do not change, even when new versions of the Human Genome Reference Assembly are adopted; this allows for a constant and unambiguous reference, a stable variant nomenclature, and shorter, more practical, coordinate values. LRG sequences are not currently available for a number of clinically actionable genes, but their creation is encouraged. Investigators from different domains of PGx research may thus consider requesting the creation of LRG sequences for specific PGx genes of interest in agreement with community consensus.

RefSeqGene sequences, which are similar to LRGs, can also be used as reference sequences. RefSeqGene and LRG are tightly coordinated, and when an LRG is created it is equivalent to a version of a RefSeqGene. Chromosome sequences from the Human Genome Reference Assembly and gene-specific sequences from RefSeqGene are based on the current version of the human genome assembly. RefSeqGenes and chromosome sequences are updated and versioned as novel genomic information becomes available; thus, the version number must be reported for the unambiguous definition of these sequences.

Inclusion of dbSNP\(^{53}\) rs IDs, each of which corresponds to a variant in a specific genomic location, also provides a definition of each variant and allows reference to databases, such as PharmGKB and ClinVar. The rs IDs are available for the majority of commonly tested PGx variants and can be created upon request.\(^{54}\) In some cases, the variant described by the rs ID can refer to more than one sequence variant and may result in more than one haplotype designation. For example, the triallelic SNP rs5030865 (NG_008376.3:g.5959G>T, NM_000106.5:c.505G>A, NG_008376.3:g.5959G>A, NM_000106.5:c.505G>A), defines haplotypes CYP2D6*8 and CYP2D6*14 depending on whether the coding strand G>T or G>A substitution occurred. Hence, the HGVS annotation must also be stated to provide an unambiguous description.
Table 4  Summary of International Workgroup Recommendations

Naming sequence variants:

1. Use HUGO Gene Nomenclature Committee (HGNC) nomenclature to specify the gene.


3. Use Locus Reference Genomic (LRG), RefSeqGene, and/or a specific Human Genome Reference Assembly as a reference sequence. Both the reference sequence accession number and version number of the sequence should be indicated.

4. Report rs IDs from dbSNP, when available.

5. Use haplotype translation tables to convert star (*) alleles or other legacy nomenclature to fully-specified HGVS nomenclature for each variant in the haplotype.

Test report:

6. Indicate each variant and/or haplotype observed in the test report.

7. List variants and haplotypes that can be detected by the test (specific sites for genotyping tests, or regions for sequencing-based tests).

8. Describe the test, including limitations, such as types of variants that cannot be detected.

9. The test description should be made publicly available on the laboratories’ website and/or by registration of the test in the National Institutes of Health (NIH) Genetic Testing Registry (GTR).

Recommendation for naming variants (recommendation 5).

(5) Use haplotype translation tables to convert star (*) alleles or other legacy nomenclature to fully-specified HGVS nomenclature for each variant in the haplotype.

Naming a variant using HGVS can be time-consuming and any one of many reference sequences can be used, resulting in a large number of possible names for a given variant. The workgroup developed a series of tables based on the haplotype definition tables hosted by PharmGKB, which can be used to convert the star or other nomenclature to HGVS in a standardized manner, as shown in Figure 3. These conversion tables are available as “Reference Haplotype” lists on the Leiden Open Variation Database 3.0 website.42 The tables provide the HGVS nomenclature of each rs ID variant using the Human Genome Reference Assembly, the RefSeqGene sequence, and LRG when available. They allow for easy and standardized naming of variants of known PGx haplotypes using HGVS. Similar tables will have to be developed for other commonly tested PGx genes in order to create a uniform system. HGVS nomenclature for variants with rs IDs is also provided by theClinVar database and can be generated using the Mutalyzer SNP converter tool."55,56 The nomenclature provided by Mutalyzer may need adjustment if the rs ID can refer to more than one variant, as described above. When pharmacogenetic genotypes are submitted to databases like ClinVar and Leiden Open Variation Database, the HGVS expression for the genotype (e.g., combinations of haplotypes or rsSNPs) will also be provided in chromosome, RefSeqGene, and LRG coordinates.

Recommendations for the test report (recommendations 6–9).

(6) Indicate each variant and/or haplotype observed in the test report. (7) List variants and haplotypes that can be detected by the test (specific sites for genotyping tests, or regions for sequencing-based tests). (8) Describe the test, including limitations, such as types of variants that cannot be detected. (9) The test description should be made publicly available on the laboratories’ website, and/or by registration of the test in the NIH Genetic Testing Registry (GTR).

It is important to record not only the genotype determined for the patient, but also the exact variants and haplotypes that were tested. Current tests for PGx genes vary considerably between clinical and research laboratories. Hence, alleles/haplotypes are often called based on different combinations of tested variants. Without a common understanding of which sequence variants are tested and which were found to be nonreference (variant), it is difficult to comprehend the meaning of a genotype result and compare genotypes reported by different laboratories or research studies.

The content of clinical test reports is specified by regulatory agencies and professional guidelines.57–62 The reports should include information, such as patient name, name/address of testing laboratory, test report date, test performed, test results, interpretation, and other information, as required by local regulations. Clinical laboratories in the United States are required to provide, when requested, a description of the test, the established performance specifications, and technical limitations.57,60 Ideally, this information should be provided in the test report given to clinicians,58 although this is not mandatory and does not always occur.

Interpretation of PGx test results requires consideration of all variants that were tested. Because PGx haplotypes may be composed of one or more variants on the same chromosome, it is important to know which variants were tested and found in the sample and which were not. Laboratories typically report only those variants that were identified in a sample, because listing results for all variant loci tested would require a much larger report format. Descriptions of the variants included in a test are needed for result interpretation when the patient is tested initially and also if the test result will be used for future patient care decisions. It is important to archive the test description, because this information might not be transmitted to electronic medical records.

The unambiguous delineation of tested variants is supported by the GTR.14 By registering a genetic test in the GTR and reporting specific variants that are assayed by the test, a report
Panel a

**Figure 3** Example of a tool that can be used to convert between star allele and HGVS nomenclature for CYP2C19. (a) Portion of a haplotype – HGVS conversion table for CYP2C19. Coordinates from the Human Genome Reference Assembly (e.g., NC_000010.10 [GRCh37], NC_000010.11 [GRCh38]), RefSeqGene (NG_008384.2), reference transcript (NM_000769.2), and reference protein (NP_000760.1) are provided for each of the rs IDs indicated, as applicable. An LRG sequence is not currently available for this gene. The relative exonic location of each rsSNP is indicated. The haplotype status of each rsSNP is indicated as nonvariant (black font in the reference *1 haplotype) or variant (red font in the polymorphic haplotypes). The combination of variants that compose each haplotype is also shown. Cells with gold shading indicate the defining SNP for each haplotype. Cells with red shading indicate rsSNPs with more than one possible variant nucleotide. HGVS nomenclature was derived from the dbSNP rs ID using the Mutalyzer SNP Converter tool (https://www.mutalyzer.nl/). (b) Detail of CYP2C19 haplotype – HGVS conversion table showing variants defining CYP2C19*4, *4A, and *4B haplotypes.

### Panel b

<table>
<thead>
<tr>
<th>Haplotype id</th>
<th>CYP2C19 Tag(s)</th>
<th>rs11188072</th>
<th>rs12248560</th>
<th>rs28395084</th>
<th>rs1785098</th>
<th>rs3758581</th>
</tr>
</thead>
<tbody>
<tr>
<td>PA165980634</td>
<td>*1</td>
<td>C</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>A</td>
</tr>
<tr>
<td>PA165980637</td>
<td>*4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA165945750</td>
<td>*4A</td>
<td></td>
<td>G</td>
<td>T</td>
<td>G</td>
<td></td>
</tr>
<tr>
<td>PA165945751</td>
<td>*4B</td>
<td>T</td>
<td>T</td>
<td>G</td>
<td>T</td>
<td>G</td>
</tr>
</tbody>
</table>

Note: This position has 2 possible variants. Only A>G is associated with the haplotype. Note: NC_000010.10 had a T at this position, NC_000010.11 has an A.
generated for that test can provide a permanent record of what was assayed and therefore what variants were (or were not) explicitly detected. GTR provides a stable identifier (accession), which can be hyperlinked on the test report to the GTR website describing the test details, such as methodology, indications/conditions, targets (genes; variants), test performance characteristics, limitations, and laboratory and test certifications. Next generation sequencing methodology has the capability to detect many, but not all variants within the genomic region(s) sequenced. It would not be feasible to list all variants that are potentially detectable by a next generation sequencing test; however, those variants that are key determinants of a haplotype should be itemized to support transparency and reinterpretation. Thus, mechanisms, such as GTR, which support review of a test result by providing information on what alleles were potentially detectable by the assay and therefore what haplotypes have or have not been excluded, can inform the interpretation and resolution of discordant test results on the basis of assay capability. Furthermore, the test description within the GTR improves standardization and facilitates comparison of different tests based on test method and detectable variants.

The workgroup also discussed the importance of clearly stating in the report which rs IDs or genomic regions, as well as other variant types, such as CNVs, the test is designed to detect, how the alleles/haplotypes are determined, and which variants were successfully tested in the specific analyses. In addition to highlighting the variants that are covered by the test, laboratories should also clearly describe the limitations of the test, possibly including clinically relevant variants, haplotypes, and variant types, such as CNV, that cannot be detected. This information could be included in the “assay limitations” section of the laboratory report or included in a footnote or appendix. It is also important to state assay limitations related to haplotypes that cannot be differentiated because of variant phasing or test design. Information about test design and limitations should also be made publicly accessible. Laboratories may want to consider adding a disclaimer to the report stating that observed phenotype may be different than what is predicted based on the genotype because of the presence of untested variants, drug interactions, or other genetic and environmental factors.

The workgroup discussed whether laboratories should consider reporting genotypes for loci that were tested but found to be nonvariant (i.e., report genotypes for reference sequence at all nonvariant but tested locations). This information is crucial for understanding how the reported haplotypes were determined, which haplotypes cannot be excluded, and the limitations of the test. It may be possible to simplify these data by providing the results of all rsSNPs and other variant types, such as CNVs, that can be detected by the test and are found to be nonreference and then stating in the report that all other rsSNPs and variants that can be detected by the test (and listed in the test description) were found to be nonvariant (reference). This information could be provided in a footnote or as an appendix to the report. Variants should be described using HGVS nomenclature, preferably referenced to RefSeqGene or LRG coordinates, as stated above.

This detailed description of the test should inform clinicians and other professionals (current and future) about the capabilities and limitations of the test to allow full evaluation and comprehension of the test results. The workgroup recognized that, although it may not be feasible to list all possible variants and haplotypes that cannot be identified by the test, some context should be given to the results so that they can be interpreted and used for care of the patient in the future.

Additional considerations

There are a number of challenges that must be addressed and overcome before many of these recommendations can be implemented. LRG sequences need to be designated for all relevant PGx genes. Identification of a reference sequence for many of these genes can be difficult because of their highly polymorphic nature and substantial differences between populations. For example, for the VKORC1 variant rs9923231 (NG_011564.1: g.3588G>A, NM_206824.1: c.-1639G>A), A is the major allele in East Asians but G is the more common allele in other population groups. Thus, it is difficult to decide which one should be designated as the reference sequence for this gene, complicating LRG assignment.

The National Center for Biotechnology Information has multiple databases that provide identifiers for variant locations and specific alleles. The dbSNP assigns an rs ID to locations where variants <50 bp have been seen in one or more genomes. Ideally, the rs ID defines a unique genomic location. It is convenient to reference common PGx variants using the rs ID, however, caveats do exist. Some rsSNPs map to more than one genomic location because of sequence similarities, such as to a pseudogene or a functional gene within a family; therefore, the location of the variant being interpreted must be established. In addition, not all variants, such as duplications, deletions, or gene conversions that are larger than 50 base pairs are captured by dbSNP. Thus, other resources, such as dbVar,63 may also be used. The National Center for Biotechnology Information variation databases encourage feedback by a variety of stakeholders, including clinical laboratories, to prevent ambiguous variant descriptions based on its identifiers.

In PGx, it is critical to know which variants were identified and which were tested but found to be nonvariant (i.e., wild-type, reference). In the example shown in Table 2, laboratory 5 could not distinguish a VKORC1*1 from a *3 or *4, because it only tested for rs9934438. This may have been apparent if the laboratory report had indicated the result from each tested rsSNP. Although listing this information could make patient test reports very long, it provides transparency to the test and its interpretation. This information could be included as a supplement or hyperlink to the laboratory’s current patient test report to keep the report brief.

Details about the test, including alleles tested and test limitations, should be included in the electronic medical/health record together with the patient’s results. As of 2015, some electronic medical/health records can store the test results, but not the test description, as structured data. The test description is usually provided as unstructured text (often in pdf format) as part of the
laboratory test report, which may be available electronically or as a paper copy. In some cases, only the predicted phenotype is reported when results are shared with other laboratories. The storage and display of genetic testing results is currently being addressed by a number of groups, including the Institute of Medicine Action Collaborative DIGITize: Displaying and Integrating Genetic Information Through the EHR and the Electronic Medical Records and Genomics (eMERGE) Network.

PGx test standardization, a “Recommended Test Panel”? As discussed above, the variants included in clinical PGx tests are not standardized. Without exception, no two tests that examined a particular PGx gene included in the GeT-RM study detected the same set of variants and/or haplotypes. In addition, some tests used different combinations of variants to define the haplotypes, leading to discrepancies between platforms in the reported genotype of many samples. Differences in test design, including inconsistent inclusion of variants and/or gene copy number detection, may impact allele calling, genotype assignment, and ultimately test interpretation. The workgroup recommendations presented above assist in elucidating the differences between tests, but do not prevent them from occurring.

The variants tested and methods utilized to assign haplotypes would also need to be standardized to ensure that different tests provide consistent results. The nomenclature workgroup discussed the possible creation of a list of variants that should be tested, at a minimum, for clinically relevant PGx genes, analogous to the standardization of cystic fibrosis carrier screening. In 2001 and 2004, the American College of Medical Genetics and the American College of Obstetricians and Gynecologists developed a panel of 23 mutations that are recommended for cystic fibrosis carrier screening in the US population. Cystic fibrosis carrier screening assays offered by clinical laboratories typically include other alleles in addition to the 23 recommended cystic fibrosis alleles. Adoption of the American College of Medical Genetics/American College of Obstetricians and Gynecologists carrier panel has helped to standardize cystic fibrosis carrier testing and make the development and evaluation of assays more transparent.

“Recommended Test Panels” could be developed for clinical PGx assays, which may similarly help to standardize pharmacogenetic testing. SNVs/alleles for each gene could be selected using criteria, such as population frequency, level of supporting evidence for phenotypic outcome, clinical utility, and severity of adverse drug reaction. For example, a well-validated four-SNV set of rs1801280 (NG_012246.1:g.14100T>C, NM_000015.2:c.341T>C), rs1779930 (NG_012246.1:g.14349G>A, NM_000015.2:c.590G>A), rs1779931 (NG_012246.1:g.14616G>A, NM_000015.2:c.857G>A), and rs1801279 (NG_012246.1:g.13950G>A, NM_000015.2:c.191G>A) is considered sufficient for prediction of the fairly common poor metabolizer ‘S’, ‘6’, ‘7’, and ‘14’ allelic groups respectively of the NAT2 gene according to the current nomenclature.

On the other hand, complete DYPD deficiency is considered relatively rare (<1%) with estimates of the frequency of the 2*A allele (the most common variant associated with DYPD deficiency) ranging from <0.005 to 3.5% in different populations. Because patients who are homozygous for DYPD*2A are at highest risk for severe or even fatal 5-fluorouracil or capecitabine toxicity, use of both drugs is not recommended for these patients. Thus, although these variants are very rare, the severity of the adverse reaction would justify inclusion in the “Recommended Test Panel” for DYPD.

Standardized panels would enable physicians, pharmacists, researchers, and other stakeholders to understand PGx test results without extensive scrutiny of the alleles included in the assay, and would provide assurance that the panels include a core set of variants considered most important for clinical utility. This is especially important when test results, originally obtained to assist selection and dosing of one drug, are later used for selection and dosing of a different drug for the same patient. Assays would be directly comparable and yield comparable results for the same sample. As with cystic fibrosis, laboratories could add additional alleles to their assays for research and other purposes.

The workgroup members did not reach consensus regarding the feasibility of creating PGx “Recommended Test Panel” lists. Although all acknowledged the clear benefits of having such panels, many felt that such an undertaking fell outside of the scope of this particular workgroup. The European Pharmacogenetic Implementation Consortium, a workgroup of the International Federation for Clinical Chemistry, is currently developing a list of PGx alleles that should be included in clinical tests. It can be argued that determining a “Recommended Test Panel” for pharmacogenes is even more difficult than for Mendelian diseases. Additionally, allele function in in vitro models may not be equivalent to clinical phenotype, and allele function can be substrate (drug), or drug concentration-dependent. Although an allele can be extremely rare in the general population, it can be frequent in another population or in a phenotypically selected group of patients (e.g., those with side effects on specific drugs). Therefore, the question arises, if the allele is rare, but unequivocally affects function, should it be included in a test panel?

Additionally, a “Recommended Test Panel” would require periodic reevaluation as novel variants are found, characterized, and new insights into genotype-phenotype relationships are identified. Each association will need to be evaluated by the same criteria to assess importance and added to the “Recommended Test Panel” if the criteria are met. In some cases, further research may show that an allele previously on the “Recommended Test Panel” list no longer meets the required criteria. A plan for routine evaluation of the “Recommended Test Panel” would be critical.

Conclusion and future plans

The PGx Nomenclature Workgroup has developed recommendations to standardize the way pharmacogenetic variants are described and reported. These recommendations, together with those currently being developed to standardize phenotype inference, will make PGx test results more transparent and will harmonize PGx testing with the broader field of genetic testing. One important purpose of this article was to link existing, non-standard variant descriptions and reporting to existing standards like the HGVS nomenclature. Clinicians and others who are familiar with the star nomenclature might find the HGVS...
nomenclature to be tedious or hard to understand. To address this issue, the workgroup suggests that laboratory reports can include common or familiar names in parenthesis along with the HGVS notation. In addition, this group is currently developing HGVS conversion tables for pharmacogenes that will allow easy translation between the recommended nomenclature and those notations currently in use.

The workgroup is proposing these changes with the expectation that they will be considered by organizations, including gene variation databases, scientific journals, regulatory agencies, and professional societies, for the creation of policies, guidelines, and recommendations. After publication of these recommendations, the workgroup plans to meet with representatives of relevant professional societies, PGx gene nomenclature committees, and other groups, some of whom have begun to address these issues to explore these ideas further.

The workgroup hopes that these proposed recommendations will be discussed and ultimately adopted by the PGx community. Although some ideas described may seem controversial, these recommendations are intended to lead to a more standardized approach to testing and reporting of PGx test results. Ultimately, standardization will enhance clinical interpretation of PGx associations, accelerating the implementation of pharmacogenetics into routine clinical practice.

ACKNOWLEDGMENTS
The authors thank Rosane Charlab Orbach and Anuradha Ramamoorthy for thoughtful comments on the manuscript and also Alain Silk, Kellie Kelm and Zivana Tezak from FDA’s Center for Devices and Radiological Health for participating in the working group and providing comments on the manuscript.

SOURCE OF FUNDING
J.A.G.A. acknowledges financial support from RD12/0013/0002, ISCIII, and FEDER. T.E.K., K.S., and M.W.C. acknowledge financial support from NIH NIGMS R24 GM61374. This work was funded in part by the NIH/NIGMS (U19 GM61388; the Pharmacogenomics Research Network) (R.R.F.) and R24GM115264 (M.V.R., K.E.C., T.E.K., and M.W.C.). V.M.L. was supported by a MarieCurie IEF fellowship for career development in the context of the European FP7 framework programme. S.A.S. was supported by the National Institute of General Medical Sciences (NIGMS) of the National Institutes of Health (NIH) through grant K23 GM104401. M.S. and U.M.Z. were supported by the Robert Bosch Foundation, Stuttgart, Germany. A.G. acknowledges financial support from 2 R01 GM088076-05 and R01 DA035736. This research was supported (in part) by the Intramural Research Program of the NIH, National Library of Medicine. E.B. was funded by the National Human Genome Research Institute (NHGRI) grant U41HG003345 and Wellcome Trust grant 099129/Z/12/Z. This work was also supported by the IGNITE project grant U01HG007762 (V.M.P.). V.M.P. was supported by the Indiana University Health – Indiana University School of Medicine Strategic Research Initiative. PharmGKB is supported by the NIH/NIGMS R24 GM61374.

CONFLICT OF INTEREST
J.L.B. owns stock in AssureX and Oneome. Grants: NIH. Royalties: AssureX and Oneome. Patent for psychiatric pharmacogenomic selection algorithms licensed to AssureX. C.B. is an employee of Affymetrix. A.L.T. is an employee of Millennium Health. R.E. is an employee and option holder of Agena Biosciences. A.G. is a paid consultant for Millennium Health. H.H. is the Chief Science Officer at Translational Software, an interpretive service company and owns company’s stock. T.H. is an employee of Thermo Fisher Scientific. T.E.K. is a Scientific Consultant at Personalis. H.L.M. is on the Board of Directors at Cancer Genetics. V.M.P. is an employee of a fee-for-service clinical laboratory. M.V.R., her husband, and hospital get royalties from TPMT genetic test. A.R. is an employee of Aegis Sciences Corporation. S.A.S. receives support from the NIH for antiplatelet pharmacogenomics research and is an Associate Director of a clinical laboratory that performs pharmacogenetic testing. R.K.T. is an employee of Genelex. L.H.T. is employed by the Coriell Institute for Medical Research. R.T. consulted for Apexus and is an associated editor for CPT. U.M.Z. is co-inventor on several patent applications. All other authors reported no conflicts.

AUTHOR CONTRIBUTIONS
L.V.K. wrote the manuscript, all other authors contributed to discussions, provided figures, references, and text, and commented and edited the text.

Published 2015. This article is a U.S. Government work and is in the public domain in the USA.

14. Rubinstein, W.S. et al. The NIH genetic testing registry: a new, centralized database of genetic tests to enable access to...


