Published in final edited form as:

Hum Mutat. 2011 May ; 32(5): 526–531. doi:10.1002/humu.21454.

Databases in the Area of Pharmacogenetics

Sarah C. Sim^{1,*}, Russ B. Altman², and Magnus Ingelman-Sundberg¹

¹Section of Pharmacogenetics, Department of Physiology and Pharmacology, Karolinska Institutet, Stockholm, Sweden

²Department of Bioengineering, Stanford University, Stanford, California

Abstract

In the area of pharmacogenetics and personalized health care it is obvious that databases, providing important information of the occurrence and consequences of variant genes encoding drug metabolizing enzymes, drug transporters, drug targets, and other proteins of importance for drug response or toxicity, are of critical value for scientists, physicians, and industry. The primary outcome of the pharmacogenomic field is the identification of biomarkers that can predict drug toxicity and drug response, thereby individualizing and improving drug treatment of patients. The drug in question and the polymorphic gene exerting the impact are the main issues to be searched for in the databases. Here, we review the databases that provide useful information in this respect, of benefit for the development of the pharmacogenomic field.

Keywords

pharmacogenetics; database; personalized medicine; adverse drug reaction; cytochrome P450; transporter; polymorphism

Introduction

There is a rapid development in genetic analyses and functional genomics covering next-generation sequencing methods and versatile arrays that rapidly analyze the most important mutations and gene copy number variations in the genome. Based on this progress, it becomes increasingly evident that much but not all of the drug therapy can be individualized based on genome sequence variation. These efforts can result in pharmacogenomic drug labels that advise the physician to use a genetic test prior to drug prescription. This test can be mandatory, recommended or for information purposes only (see [Sim and Ingelman-Sundberg, 2011] for a review). The genetic variation of interest mainly concerns genes encoding drug metabolizing enzymes, drug transporters, drug targets, and human leukocyte antigens (HLA). Only 30–60% of patients respond to common drug therapy using, for example, antidepressants, statins, and antipsychotics [Spear et al., 2001]. Furthermore, adverse drug reactions (ADRs) are an increasing problem [Moore et al., 2007] and apparently cost society as much as the drug therapy itself (cf. [Eichelbaum et al., 2006]). In addition, meta-analyses by Lazarou et al. [1998] revealed consequences of more than 100,000 deaths annually caused by ADRs in the United States, and Davies et al. [2009] have documented an increased length of hospitalization for more than 25% of patients

experiencing ADRs in the United Kingdom. Also, ADRs are shown to occur in 16% of hospitalized patients, frequently with severe consequences [Davies et al., 2009], and ADRs have in several studies been shown to be responsible for about 7% of hospitalizations in the average population [Davies et al., 2009; Lazarou et al., 1998], 20% of readmissions to the hospital [Davies et al., 2010] and more than 30% among elderly [Paul et al., 2008]. These data emphasize the need of updated databases that can provide fruitful guidance to both scientists, physicians, regulatory agencies, and industries to cope with this major problem in human health (Table 1).

The Pharmacogenomics Knowledge Base (PharmGKB)

The Pharmacogenomics Knowledgebase (PharmGKB, <http://www.pharmgkb.org/>) was established in 2000 and strives to provide information on how human genetic variation impacts drug response. Stable government-funded efforts such as GenBank, Gene Expression Omnibus, and dbGAP from the National Center for Biomedical Information (NCBI) at the National Institutes of Health (NIH) in the United States are the preferred databases for archiving data, whereas curation of genes (or groups of related homologous genes) and their variations are best handled by locus-specific databases such as the Human Cytochrome P450 (CYP) Allele Nomenclature Website (see later), and others focusing on transporters, receptors, and kinases. In contrast, PharmGKB focuses on four main activities: (1) curation of published literature in respect to relationships between genes, their variations, and drug response phenotypes, (2) creation of “Very Important Pharmaco-gene” (VIP gene) summaries that highlight specific pharmacogenomic significance of genes that are very important for drug response, (3) creation of pathway diagrams focusing on drug response (pharmacodynamics [PD]) and drug metabolism (pharmacokinetics [PK]) pathways, and finally (4) assistance to organize and lead data-sharing consortia within pharmacogenomics where multiple centers combine data and expertise to answer questions that cannot be answered by any center alone (usually because of statistical power and smaller data sets).

The core activity of PharmGKB curators is to survey the literature and capture reports of human genetic variation that impacts drug response phenotypes. These phenotypes may occur at the molecular (e.g., binding), cellular (e.g., expression), tissue (e.g., drug transformation) or organismal level (e.g., disease, symptoms), and all annotations are encoded with respect to these categories to assist in searches. The PharmGKB has a list of journals publishing pharmacogenomics observations, and also broadly searches the literature on a weekly basis to find associations between genes, their variations, drugs, and phenotypes. The PharmGKB uses the Human Genome Nomenclature Committee recommendations for genes and variations. Also, they use a combination of the RcNorm drug list with some modifications for drug, and Medical Subject Headings (MeSH) and Unified Medical Language System (UMLS) to encode phenotypes. The result of this curation activity is currently about 2,500 annotations that associate particular human variations, mostly single nucleotide polymorphisms (SNPs), but also including insertions or deletions (indels), copy number variations, structural changes, and haplotypes with drug-related phenotypes. These annotations were recently used to create a “pharmacogenomic annotation” of a complete human genome sequence [Ashley et al., 2010]. Interestingly, there were nearly 100 drugs for which risks or benefits could be predicted for this individual. Many of them were “low confidence” based on the level of published evidence, but several of them would have some clinical utility. The application of the PharmGKB annotations to this genome was manual, and so current efforts are aimed at representing the basic variant annotations and their potential clinical consequences in a more structured manner that will allow more automated application to individual genomes.

As the PharmGKB curators accumulate individual variant annotations, the genes may appear as either (1) repeatedly involved in drug response because of their role in general drug pharmacokinetics (e.g., cytochrome P450s, transporters, transferases) or (2) having an important role in the action of a single drug or class of drugs (e.g., pharmacodynamic drug targets). When the amount of information for these genes reaches a level that the curators find to be sufficiently large, they can designate these genes as VIP genes denoting genes that have a documented role in modulating the response to one or more drugs. The PharmGKB curators prioritize these genes for short written summary reviews of their role in drug response, and the key variations in these genes that have been shown to modulate drug response. These reviews are reexamined every 2 years or more frequently if the pace of new knowledge generation is rapid, and are written in a templated manner that includes names and alternate names, introductory information, key publications, key pathways, known drugs/substrates, associated phenotypes/diseases and summaries of key variants with their associated effects. There are currently 43 VIP genes on the PharmGKB Website, and their summaries are published regularly in a cooperative venture with the *Pharmacogenetics & Genomics* journal in order to draw attention to them and to provide a citation for their usage [Eichelbaum et al., 2009; Mi et al., 2010].

In addition to the VIP gene summaries, curators also create annotated graphical pathways depicting the relationships of genes and drugs. These pathways are the most popular feature at PharmGKB. Each pathway is the result of a PharmGKB curator gathering all information available for the drug of interest (e.g., by collecting the basic annotations and gene–drug relationships discussed above), and assembling a pathway summary. These pathways are often created collaboratively with experts working on these systems and they include a down-loadable graphical depiction that is linked to the individual genes and small molecules depicted. It also includes supporting data and references upon which the pathway is based. As with the VIP genes, the pathways are also published regularly in the *Pharmacogenetics & Genomics* journal [Eichelbaum et al., 2009; Sangkuhl et al., 2010]. In addition to PharmGKB-generated pathways, PharmGKB links to pathways from external resources, and they are made available to augment the PharmGKB pathways.

A somewhat unexpected activity for PharmGKB has been the convening of data sharing consortia. The mission of PharmGKB is to promote the creation and sharing of data about pharmacogenomics, and as a trusted third party, PharmGKB is in a good position to bring together groups of investigators interested in common scientific questions. After signing a memorandum of understanding that outlines the responsibilities and privileges of joining the consortium (and protects the rights of data generators to publish, as well as the rights of the consortium members to see data and participate in its analysis), the PharmGKB accepts data sets, brings them into a common format, and then participates in the statistical analysis to answer one or more shared scientific questions. The first example of a PharmGKB consortium was the International Warfarin Pharmacogenetics Consortium (IWPC), which brought together 21 groups internationally to create a globally relevant warfarin dosing algorithm, as well as an analysis of interethnicity differences in warfarin dosing [Klein et al., 2009; Limdi et al., 2010]. Other consortia have subsequently been established addressing the discovery of new genetic variations relevant to warfarin, the role of CYP2D6 variations in tamoxifen response, and the role of CYP2D6 in selective serotonin reuptake inhibitor (SSRI) response. A list of consortia and their members are available on the “Contributors” tab of PharmGKB.

In addition to a global search capability, users of PharmGKB are able to search for particular kinds of information in five more focused search boxes: genes, variants, pathways, drugs, and diseases. These search boxes can accept any text, but will only return results of one of these five specific types. In addition, the home page provides the PharmGKB curators

“Favorite Papers” that are updated every 2 weeks, and are available as an archive over the last 4 years. The home page also offers an automatically created pharmacogenomics news feed, access to tutorial information, introductions to pharmacogenetics, and access to tools.

Current new efforts at PharmGKB focus on:

- The development of automated natural language parsing tools to extract annotation information from the literature automatically [Coulet et al., 2010; Garten and Altman, 2009, 2010].
- The addition of chemoinformatics, such as drug structure similarity metrics to help predicting new gene-drug interactions [Garten et al., 2010; Hansen et al., 2009].
- The creation of systems biology methods to look at drug response with a more network-based view of the genes involved in producing complex drug responses [Hansen et al., 2009; Tatonetti et al., 2009].

The Human Cytochrome P450 (*CYP*) Allele Nomenclature Website

Cytochrome P450 enzymes are responsible for 75–80% of all phase I-dependent metabolism and for 65–70% of the clearance of clinically used drugs [Bertz and Granneman, 1997; Evans and Relling, 1999]. *CYP* enzymes are encoded by 57 *CYP* genes and divided into families, of which 1–3 (*CYP*1–3) are generally involved in the metabolism of exogenous substances including drugs, whereas members of other families mainly are involved in metabolism of endogenous compounds. Variation in *CYP* genes results in four different phenotypes classically defined as ultrarapid, extensive, intermediate, and poor metabolizers.

Characterization of the genetic polymorphism of *CYP* enzymes provides a basis for personalized medicine composing drug dose adjustment and selection of drug type [Ingelman-Sundberg, 2004; Ingelman-Sundberg et al., 2007; Kirchheiner et al., 2004; Pirmohamed and Park, 2001; Weinshilboum, 2003].

To unite and summarize efforts focusing on the characterization of *CYP* polymorphism, the Human Cytochrome P450 Allele (*CYP*-allele) Nomenclature Committee was established in 1999, and the efforts resulted in agreed *CYP* nomenclature guidelines and the launching of the *CYP*-allele Website (<http://www.cypalleles.ki.se/>). The main purpose of the *CYP*-allele Website is the management of new allele designations based on recognized nomenclature guidelines, facilitation of rapid publication, as well as providing a readily available summary of alleles and their associated effects [Ingelman-Sundberg and Oscarson, 2002; Oscarson and Ingelman-Sundberg, 2002; Sim and Ingelman-Sundberg, 2006, 2010].

At present, 29 polymorphic *CYP*s are included on the *CYP*-allele Website, with the addition in 2008 of NADPH cytochrome P450 reductase (POR) [Sim et al., 2009] due to its elementary function as electron donor to *CYP* enzymes. Every gene has its own Webpage, listing the associated alleles and their consequences on the genetic, molecular, in vitro and in vivo levels, all together with relevant references and links to the NCBI dbSNP database.

The nomenclature system designates numbers (such as *CYP2B6*4*) for alleles that have been thoroughly sequenced and well characterized in terms of linkage of the functionally relevant SNP with other polymorphic sites. Functionally relevant in this aspect is regarded as a variation causing measurable consequences such as amino acid substitutions, translation terminations, or splice defects. At a minimum, the open reading frame must have been sequenced. All sequence variations are numbered based on A in the translation initiation codon as the reference nucleotide (+1). Additional letter-designated suballeles of the same number (such as *CYP2B6*4B*) contain the same functional SNP but additional sequence variations not causing functional effects. A combination of functionally relevant variants are

designated a new allele number due to the potential of combinatorial effects, unless one of the variants is deleterious and thus defines the allele name. Computer-inferred haplotypes are generally avoided, although some predicted alleles have been assigned and included on the *CYP*-allele Website. Alleles published in the early beginning of the *CYP*-allele era do not fully comply with the current standard.

As an example of *CYP*-allele nomenclature, *CYP2C19*2* carries a SNP causing a splice defect of the transcribed RNA that is detrimental for enzyme function. This splice defect (c. 681G>A) exists in different combinations with other SNPs causing amino acid substitutions and the combination alleles have thus received allele names *CYP2C19*2A-D*. The effect of *CYP2C19* genotype on clopidogrel (Plavix®) antiplatelet treatment of cardiovascular disease has recently attained much focus because *CYP2C19* is responsible for bioactivation of the drug. A recent meta-analysis of nine studies and almost 10,000 patients showed a significantly increased risk of adverse cardiovascular events in carriers of defective (mainly *CYP2C19*2*) alleles due to impaired bioactivation of clopidogrel, especially in patients having undergone percutaneous intervention, and in particular for stent thrombosis as the major cardiovascular event [Mega et al., 2010].

Submission of new alleles is achieved by contacting the Webmaster of the *CYP*-allele Website, whereby the data characterizing the allele is reviewed for potential allele name designation (see inclusion criteria at <http://www.cypalleles.ki.se/criteria.htm>). All information is kept confidential and a manuscript in preparation can often serve as a good basis for review and discussion between the author and the Webmaster. Designation of allele names outside the *CYP*-allele nomenclature committee is not advised, due to the apparent risk of confusion in the literature. Submissions with respect to additional functional in vivo and in vitro characterization of alleles listed are also highly relevant, and aids in keeping the *CYP*-allele Website up to date.

The Human Arylamine N-Acetyltransferase (NAT) Gene Nomenclature Committee

Arylamine N-acetyltransferases (NATs) conjugate drugs and xenobiotics containing arylamine structures, and these enzymes were early described in terms of pharmacogenetic variation, resulting in slow and rapid acetylation phenotypes that may impact the risk of drug-induced toxicity [Sim et al., 2008]. There are two active genes in the *NAT* locus, *NAT1* and *NAT2* and in 1995, a nomenclature system was published [Vatsis et al., 1995]. However, a systematic procedure for naming new alleles was not in use at once and resulted in problems such as dual assignments. From 1998, further discussions resulted in the formation of the Human Arylamine N-Acetyltransferase (*NAT*) Gene Nomenclature Committee with the aim to reach a nomenclature consensus for *NAT* alleles [Hein et al., 2008, 2000]. A Website was launched for maintenance of the renewed system and to provide an update on *NAT* alleles (<http://n-acetyltransferasenomenclature.louisville.edu/>). New allele submissions are sent to the *NAT* Nomenclature Committee for review, and the *NAT* allele Website covers alleles of the *NAT1* and *NAT2* genes. The *NAT* allele nomenclature follows the star system (e.g., *NAT*5*) with special focus on functionally relevant sequence variations, whereby suballeles defined by an extra letter may contain additional variants that are functionally not relevant compared to the SNP causing the major effect. For example, alleles of the *NAT2*5a-p* series (*NAT2*5A-P*) all contain a SNP causing the I114T amino acid substitution that causes slow acetylation, which is associated with an increased risk of drug-induced liver injury in isoniazid-treated tuberculosis patients [Daly, 2010].

The NAT Website contains NAT allele information in the PDF format reporting nucleotide changes, amino acid changes, phenotypes (slow or rapid acetylation), NCBI dbSNP numbers, and relevant references, as well as information on allele inclusion criteria for designation of new alleles.

As opposed to other nomenclature systems, the “wild-type” allele is not designated *NAT*1* due to historical reasons. Thus, the reference allele for *NAT* nomenclature is *NAT*4* for both *NAT* genes. As a result of this, *NAT* alleles with numbers between **1* and **3* are excluded from the nomenclature system. At present, 29 alleles are presented for *NAT1* and 21 for *NAT2*.

The Pharmacogenetics of Membrane Transporters (PMT) Database

The Pharmacogenetics of Membrane Transporters (PMT) database (<http://pharmacogenetics.ucsf.edu/>) [Kroetz et al., 2010] is part of the PharmGKB (see section: The Pharmacogenomics Knowledge Base [PharmGKB]) and focuses on drug response in relation to genetic variation in membrane transporters. Because membrane transporters are involved in the transport of drugs, genetic variation affecting transport function may impact drug therapy in terms of both response and ADRs. The PMT project was initiated in 2000, and focuses on the two superfamilies of solute carrier (SLC) transporters and ATP binding cassette (ABC) transporters. The database covers coding and noncoding variations as well as both SNPs and indels, but does not provide allele or haplotype information. The list of transporter proteins contains the OMIM phenotype, nucleotide, and protein reference sequences, Entrez hits, and number of isoforms (splice variants) for every gene. The whole list is composed of 430 genes, although not all are transporter proteins because, for some unknown reason, genes such as monoamine oxidase A (MAOA), pregnane X receptor (PXR), and POR are also included in this list. Most genes are extensively presented on their own Web page, containing information such as main tissue expression sites, substrates, gene structure, and genetic variation. The genetic variants are reported by their genomic, transcript, and coding position, together with amino acid effects, array availability, PubMed and dbSNP links, and the frequency observed in different populations. Also, a secondary structure transmembrane prediction shows the position of amino acid substitutions, insertions, or deletions within the transporter protein. Additionally, the PMT database provides a search tool to access gene expression data with respect to a single gene in multiple tissues (Gene Plot), selected genes in selected tissues (Mixed Plot) and multiple genes in a single tissue (Tissue Plot).

Transporter Database (TP-search)

On the Transporter Database TP-search Website (<http://www.tp-search.jp/>) [Ozawa et al., 2004], transporters are summarized under the respective species and the human section contains 33 transporter proteins together with information such as the driving force used (e.g., ATP, Na⁺), the presence of SNPs, protein length, direction of transport and association with disease. In addition, searches can be performed by transporter name, compound name, drug–drug interaction, pathophysiology, genetic polymorphisms, and tissue name. Additional Web pages include gender differences, animal models, and genetic diseases. The Transporter Name search includes the choice of human, mouse, and/or rat species, and substrate, inhibitor, and/or inducer substances. The Transporter Name search outcome includes parameters associated with the requested transporter, that is, substance name and chemical structure, interaction type, K_m and K_i values, experimental setup used for producing the data and references to publications. Similar searches can be done with the Compound Name, where a list of transporters interacting with the substance is provided. The Drug–Drug Interaction search involves in vivo data, and displays drug–drug interactions

involving the searched compound, together with its effects on parameters such as clearance, $C_{\max}$ and area under the plasma concentration curve (AUC). Also, the interaction site where the effects are believed to occur is presented. The Tissue Distribution search provides schematic figures of the transporters present in intestine, kidney, or liver. In addition, a tissue distribution table lists the tissue and localization of the transporters, together with the experimental method used and the corresponding references. The Genetic Polymorphism search of functionality yields in vivo information on the respective transporter SNPs (but not alleles or haplotypes) with respect to nucleotide changes, protein changes, drug involved, type of drug administration, number and ethnicity of the study sample, parameters such as AUC or plasma concentration, and the relationship between the different genotypes. A similar genetic search with respect to frequency yields the distribution of the polymorphisms in different populations. However, a reference sequence with respect to the nucleotide change is not provided and the TP-search database has not been updated since 2007.

UGT Alleles Nomenclature Page

UGT enzymes (Phase II metabolism) link glycosyl groups to endogenous and exogenous substrates including drugs, to yield more hydrophilic products that can be more easily excreted from the body [Mackenzie et al., 2005]. The UDP-glucuronosyltransferase (UGT) Allele Nomenclature Page (http://www.pharmacogenomics.pha.ulaval.ca/sgc/ugt_alleles) contains UGT enzymes of subfamily UGT1A (nine enzyme Web pages) and UGT2B (five active Web pages). The inclusion criteria are based on the *CYP*-allele nomenclature system, and thus alleles designated by the Webmaster contains an asterisk and a number where the *1 allele (e.g., *UGT1A1*1*) is considered the wild-type allele. Similarly to *CYP*-alleles, the UGT Alleles Nomenclature Page focus on nucleotide changes that cause functional effects such as amino acid substitutions, splice defects, and frameshifts, and thus alleles with additional variations not causing any effects receives a sub (letter) allele name. Every enzyme has a “Haplotype” page and a “SNP” page, where the former contains a list of all alleles and their changes, whereas the latter contain a list of all SNPs together with their rs numbers and the reference sequence used. As an example, the UGT1A1 Haplotype Web page contains 113 different alleles that are presented with their allelic names, nucleotide changes, protein effects, exon numbers affected, consequences (e.g., frameshift), phenotypes with respect to causation of Crigler-Najjar syndrome or Gilbert's syndrome, enzyme activity in vitro and in vivo, and the respective references. The rare Crigler-Najjar syndrome is caused by a complete inability to glucuronize bilirubin due to UGT1A1 mutations that consequently causes jaundice and potential brain injury, whereas reduced but not depleted UGT1A1-mediated glucuronidation of bilirubin leads to the less severe phenotype Gilbert's syndrome. UGT1A1 is also involved in the metabolism of the anticancer drug irinotecan, because the active metabolite SN-38 is glucuronidated and inactivated by UGT1A1. Excessive levels of SN-38 can cause neutropenia, the risk of which is affected by *UGT1A1* genotype. The wild-type allele *UGT1A1*1* carries a 6-repeat TATA box (A(TA)₆TAA), whereas *UGT1A1*36* has 5, *UGT1A1*28* has 7 and *UGT1A1*37* has 8 repeats. An increased number of TA repeats causes reduced *UGT1A1* gene transcription and the common reduced activity *UGT1A1*28* allele associates with reduced clearance of SN-38 (Schulz 2009 PMID: 19770637). Due to the increased risk of neutropenia in *UGT1A1*28*/**28* subjects, the Food and Drug Administration in the United States recommends a lower starting dose of irinotecan in such patients.

PharmaADME

PharmaADME (<http://www.pharmaadme.org/>) is a database initiated by academic and industrial interests. The aim of this database is to compile lists of genes and genetic variation affecting Absorption, Distribution, Metabolism, and Excretion (ADME) of drugs. The

PharmaADME's main aim is to aid in selection of genotyping assays for product development and clinical trials. PharmaADME provides three gene lists, the ADME Gene List, the Core Marker Gene List, and the Related Gene List. The ADME Gene Lists contain genes associated with drug metabolism, either sorted under the more important Core List of 32 genes or under the Extended ADME Gene List of 267 genes. The ADME Gene Lists contains the gene symbol, full gene name, and the protein class, but no other information and no links to information on which the qualifications are based or which drugs the gene product is affecting. The Related Gene List contains 74 genes that are related to, but not involved in, drug metabolism such as drug targets and receptors. The Core Marker Gene List is composed of 184 genotypes in 32 genes, of which 11 are CYPs, 9 are transporters, and 9 are transferases, which by PharmaADME have been selected as most important. This list comprises the gene symbol, the primary name (e.g., *CYP2D6**5) and the NCBI dbSNP number.

Conclusions

The area of pharmacogenomics is rapidly developing. It is anticipated that novel, cheaper and more efficient high throughput techniques for genomic analyses as well as the discovery of interindividual variation in epigenomic control of gene expression and the identification of additional types of snRNA molecules, will result in a rapid expansion of the area. Consequently, the number of useful biomarkers for predicting drug response will rapidly increase in quantity and quality. Thus, existing and future databases, which are developed in a user-friendly manner, will provide relevant information on drugs and gene variants for physicians, scientists, regulatory agencies and industry, and constitute very important tools for the possibility to improve drug treatment and settle the basis for personalized medicine in society.

Acknowledgments

Contract grant sponsors: The Swedish Research Council; The Swedish Brain Foundation; The Swedish Cancer Fund; NIH; Contract grant number: GM61374.

References

- Ashley EA, Butte AJ, Wheeler MT, Chen R, Klein TE, Dewey FE, Dudley JT, Ormond KE, Pavlovic A, Morgan AA, Pushkarev D, Neff NF, Hudgins L, Gong L, Hodges LM, Berlin DS, Thorn CF, Sangkuhl K, Hebert JM, Woon M, Sagreiya H, Whaley R, Knowles JW, Chou MF, Thakuria JV, Rosenbaum AM, Zaranek AW, Church GM, Greely HT, Quake SR, Altman RB. Clinical assessment incorporating a personal genome. *Lancet*. 2010; 375:1525–1535. [PubMed: 20435227]
- Bertz RJ, Granneman GR. Use of in vitro and in vivo data to estimate the likelihood of metabolic pharmacokinetic interactions. *Clin Pharmacokinet*. 1997; 32:210–258. [PubMed: 9084960]
- Coulet A, Shah NH, Garten Y, Musen M, Altman RB. Using text to build semantic networks for pharmacogenomics. *J Biomed Inform*. Aug 17.2010 Epub ahead of print.
- Daly AK. Drug-induced liver injury: past, present and future. *Pharmacogenomics*. 2010; 11:607–611. [PubMed: 20415545]
- Davies EC, Green CF, Taylor S, Williamson PR, Mottram DR, Pirmohamed M. Adverse drug reactions in hospital in-patients: a prospective analysis of 3695 patient-episodes. *PLoS One*. 2009; 4:e4439. [PubMed: 19209224]
- Davies EC, Green CF, Mottram DR, Rowe PH, Pirmohamed M. Emergency re-admissions to hospital due to adverse drug reactions within 1 year of the index admission. *Br J Clin Pharmacol*. 2010; 70:749–755. [PubMed: 21039769]
- Eichelbaum M, Altman RB, Ratain M, Klein TE. New feature: pathways and important genes from PharmGKB. *Pharmacogenet Genomics*. 2009; 19:403. [PubMed: 20161212]

- Eichelbaum M, Ingelman-Sundberg M, Evans WE. Pharmacogenomics and individualized drug therapy. *Annu Rev Med*. 2006; 57:119–137. [PubMed: 16409140]
- Evans WE, Relling MV. Pharmacogenomics: translating functional genomics into rational therapeutics. *Science*. 1999; 286:487–491. [PubMed: 10521338]
- Garten Y, Altman RB. Pharmspresso: a text mining tool for extraction of pharmacogenomic concepts and relationships from full text. *BMC Bioinformatics*. 2009; 10(Suppl 2):S6. [PubMed: 19208194]
- Garten Y, Altman RB. Teaching computers to read the pharmacogenomics literature...so you don't have to. *Pharmacogenomics*. 2010; 11:515–518. [PubMed: 20350132]
- Garten Y, Tatonetti NP, Altman RB. Improving the prediction of pharmacogenes using text-derived drug-gene relationships. *Pac Symp Biocomput*. 2010:305–314. [PubMed: 19908383]
- Hansen NT, Brunak S, Altman RB. Generating genome-scale candidate gene lists for pharmacogenomics. *Clin Pharmacol Ther*. 2009; 86:183–189. [PubMed: 19369935]
- Hein DW, Boukouvala S, Grant DM, Minchin RF, Sim E. Changes in consensus arylamine N-acetyltransferase gene nomenclature. *Pharmacogenet Genomics*. 2008; 18:367–368. [PubMed: 18334921]
- Hein DW, Grant DM, Sim E. Update on consensus arylamine N-acetyltransferase gene nomenclature. *Pharmacogenetics*. 2000; 10:291–292. [PubMed: 10862519]
- Ingelman-Sundberg M. Pharmacogenetics of cytochrome P450 and its applications in drug therapy: the past, present and future. *Trends Pharmacol Sci*. 2004; 25:193–200. [PubMed: 15063083]
- Ingelman-Sundberg M, Oscarson M. Human CYP allele database: submission criteria procedures and objectives. *Methods Enzymol*. 2002; 357:28–36. [PubMed: 12424894]
- Ingelman-Sundberg M, Sim SC, Gomez A, Rodriguez-Antona C. Influence of cytochrome P450 polymorphisms on drug therapies: pharmacogenetic, pharmacoeigenetic and clinical aspects. *Pharmacol Ther*. 2007; 116:496–526. [PubMed: 18001838]
- Kirchheiner J, Nickchen K, Bauer M, Wong ML, Licinio J, Roots I, Brockmoller J. Pharmacogenetics of antidepressants and antipsychotics: the contribution of allelic variations to the phenotype of drug response. *Mol Psychiatry*. 2004; 9:442–473. [PubMed: 15037866]
- Klein TE, Altman RB, Eriksson N, Gage BF, Kimmel SE, Lee MT, Limdi NA, Page D, Roden DM, Wagner MJ, Caldwell MD, Johnson JA, International Warfarin Pharmacogenetics Consortium. Estimation of the warfarin dose with clinical and pharmacogenetic data. *N Engl J Med*. 2009; 360:753–764. [PubMed: 19228618]
- Kroetz DL, Yee SW, Giacomini KM. The pharmacogenomics of membrane transporters project: research at the interface of genomics and transporter pharmacology. *Clin Pharmacol Ther*. 2010; 87:109–116. [PubMed: 19940846]
- Lazarou J, Pomeranz BH, Corey PN. Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. *JAMA*. 1998; 279:1200–1205. [PubMed: 9555760]
- Limdi NA, Wadelius M, Cavallari L, Eriksson N, Crawford DC, Lee MT, Chen CH, Motsinger-Reif A, Sagreiya H, Liu N, Wu AH, Gage BF, Jorgensen A, Pirmohamed M, Shin JG, Suarez-Kurtz G, Kimmel SE, Johnson JA, Klein TE, Wagner MJ, International Warfarin Pharmacogenetics Consortium. Warfarin pharmacogenetics: a single VKORC1 polymorphism is predictive of dose across 3 racial groups. *Blood*. 2010; 115:3827–3934. [PubMed: 20203262]
- Mackenzie PI, Bock KW, Burchell B, Guillemette C, Ikushiro S, Iyanagi T, Miners JO, Owens IS, Nebert DW. Nomenclature update for the mammalian UDP glycosyltransferase (UGT) gene superfamily. *Pharmacogenet Genomics*. 2005; 15:677–685. [PubMed: 16141793]
- Mega JL, Simon T, Collet JP, Anderson JL, Antman EM, Bliden K, Cannon CP, Danchin N, Giusti B, Gurbel P, Horne BD, Hulot JS, Kastrati A, Montalescot G, Neumann FJ, Shen L, Sibbing D, Steg PG, Trenk D, Wiviott SD, Sabatine MS. Reduced-function CYP2C19 genotype and risk of adverse clinical outcomes among patients treated with clopidogrel predominantly for PCI: a meta-analysis. *JAMA*. 2010; 304:1821–1830. [PubMed: 20978260]
- Mi H, Thomas PD, Ring HZ, Jiang R, Sangkuhl K, Klein TE, Altman RB. PharmGKB summary: dopamine receptor D2. *Pharmacogenet Genomics*. Aug 23.2010 Epub ahead of print.
- Moore TJ, Cohen MR, Furberg CD. Serious adverse drug events reported to the Food and Drug Administration, 1998–2005. *Arch Intern Med*. 2007; 167:1752–1759. [PubMed: 17846394]

- Oscarson M, Ingelman-Sundberg M. CYPalleles: a web page for nomenclature of human cytochrome P450 alleles. *Drug Metab Pharmacokinet*. 2002; 17:491–495. [PubMed: 15618703]
- Ozawa N, Shimizu T, Morita R, Yokono Y, Ochiai T, Munesda K, Ohashi A, Aida Y, Hama Y, Taki K, Maeda K, Kusuha H, Sugiyama Y. Transporter database, TP-Search: a web-accessible comprehensive database for research in pharmacokinetics of drugs. *Pharm Res*. 2004; 21:2133–2134. [PubMed: 15587938]
- Paul E, End-Rodrigues T, Thylen P, Bergman U. Adverse drug reactions a common cause of hospitalization of the elderly. A clinical retrospective study. *Lakartidningen*. 2008; 105:2338–2342. [PubMed: 18831440]
- Pirmohamed M, Park BK. Genetic susceptibility to adverse drug reactions. *Trends Pharmacol Sci*. 2001; 22:298–305. [PubMed: 11395158]
- Sangkul K, Shuldiner AR, Klein TE, Altman RB. Platelet aggregation pathway. *Pharmacogenet Genomics*. Oct 8.2010 Epub ahead of print.
- Sim E, Lack N, Wang CJ, Long H, Westwood I, Fullam E, Kawamura A. Arylamine N-acetyltransferases: structural and functional implications of polymorphisms. *Toxicology*. 2008; 254:170–183. [PubMed: 18852012]
- Sim SC, Ingelman-Sundberg M. The human cytochrome P450 Allele Nomenclature Committee Web site: submission criteria, procedures, and objectives. *Methods Mol Biol*. 2006; 320:183–191. [PubMed: 16719391]
- Sim SC, Ingelman-Sundberg M. The Human Cytochrome P450 (CYP) allele nomenclature website: a peer-reviewed database of CYP variants and their associated effects. *Hum Genomics*. 2010; 4:278–281. [PubMed: 20511141]
- Sim SC, Ingelman-Sundberg M. Pharmacogenetic biomarkers: new tools in current and future drug therapy. *Trends Pharmacol Sci*. Dec 22.2011 Epub ahead of print.
- Sim SC, Miller WL, Zhong XB, Arlt W, Ogata T, Ding X, Wolf CR, Flück CE, Pandey AV, Henderson CJ, Porter TD, Daly AK, Nebert DW, Ingelman-Sundberg M. Nomenclature for alleles of the cytochrome P450 oxidoreductase gene. *Pharmacogenet Genomics*. 2009; 19:565–566. [PubMed: 19535965]
- Spear BB, Heath-Chiozzi M, Huff J. Clinical application of pharmacogenetics. *Trends Mol Med*. 2001; 7:201–204. [PubMed: 11325631]
- Tatonetti NP, Liu T, Altman RB. Predicting drug side-effects by chemical systems biology. *Genome Biol*. 2009; 10:238. [PubMed: 19723347]
- Vatsis KP, Weber WW, Bell DA, Dupret JM, Evans DA, Grant DM, Hein DW, Lin HJ, Meyer UA, Relling MV. Nomenclature for N-acetyltransferases. *Pharmacogenetics*. 1995; 5:1–17. [PubMed: 7773298]
- Weinshilboum R. Inheritance and drug response. *N Engl J Med*. 2003; 348:529–537. [PubMed: 12571261]

Table 1

Pharmacogenetic Databases

Database	Organization	Activities	Web address	Content
PharmGKB	The Pharmacogenomics Knowledgebase	Curation of pharmacogenetic effects, creation of gene and drug summaries, data sharing consortia	http://www.pharmgkb.org/	Genotype and phenotype of ADME genes and drug target genes
<i>CYP</i> -allele database	The Human Cytochrome P450 (CYP) Allele Nomenclature Website	Reviewing and designation of <i>CYP</i> alleles	http://www.cypalleles.ki.se/	CYP nomenclature and CYP allele characteristics
<i>NAT</i> -allele database	The Human Arylamine N-Acetyltransferase (NAT) Gene Nomenclature Committee	Reviewing and designation of <i>NAT</i> alleles	http://n-acetyltransferasenomenclature.louisville.edu/	NAT nomenclature and NAT allele characteristics
PMT database	The Pharmacogenetics of Membrane Transporters (PMT) database (part of PharmGKB)	Compilation of sequence variation in transporter genes (no alleles or haplotypes)	http://pharmacogenetics.ucsf.edu/	Genetic variation in SLC and ABCC transporters
TP-search database	The Transporter database	Summaries of transporters and their genetic variations (no alleles or haplotypes)	http://www.tp-search.jp/	Genotype and phenotype of SLC and ABCC transporters
<i>UGT</i> -allele database	The UDP-glucuronosyltransferase (UGT) Allele Nomenclature Page	Reviewing and designation of <i>UGT</i> alleles	http://www.pharmacogenomics.pha.ulaval.ca/sgc/ugt_alleles	UGT nomenclature and UGT allele characteristics
PharmaADME	Pharmacogenetics of Absorption, Distribution, Metabolism and Excretion genes	Compilation of ADME gene lists	http://www.pharmaadme.org/	Polymorphic ADME genes ranked in order of significance for causing variability in drug response