PharmGED: Pharmacogenetic Effect Database

C. J. Zheng¹,², L. Y. Han¹,², B. Xie¹,², C. Y. Liew¹,², S. Ong¹,², J. Cui¹,², H. L. Zhang¹,², Z. Q. Tang¹,², S. H. Gan¹,², L. Jiang¹,² and Y. Z. Chen¹,²,*

¹Department of Pharmacy and ²Department of Computational Science, Bioinformatics and Drug Design Group, National University of Singapore, Blk S16, Level 8, 3 Science Drive 2, Singapore 117543

Received August 15, 2006; Revised and Accepted October 10, 2006

ABSTRACT

Prediction and elucidation of pharmacogenetic effects is important for facilitating the development of personalized medicines. Knowledge of polymorphism-induced and other types of drug-response variations is needed for facilitating such studies. Although databases of pharmacogenetic knowledge, polymorphism and toxicogenomic information have appeared, some of the relevant data are provided in separate web-pages and in terms of relatively long descriptions quoted from literatures. To facilitate easy and quick assessment of the relevant information, it is helpful to develop databases that provide all of the information related to a pharmacogenetic effect in the same web-page and in brief descriptions. We developed a database, Pharmacogenetic Effect Database (PharmGED), for providing sequence, function, polymorphism, affected drugs and pharmacogenetic effects. PharmGED can be accessed at http://bidd.cz3.nus.edu.sg/phg/ free of charge for academic use. It currently contains 1825 entries covering 108 disease conditions, 266 distinct proteins, 693 polymorphisms, 414 drugs/ligands cited from 856 references.

DATABASE STRUCTURE AND ACCESS

PharmGED has a web interface at http://bidd.cz3.nus.edu.sg/phg/. The entries of this database were derived from a comprehensive search of published literatures (via Medline) by using a similar search and evaluation procedure as we have used for developing other databases of drug-related proteins (15–18). Entries of this database are searchable by several methods. These methods include the search of protein name, drug/ligand name, disease name (extracted from the related terms described in the relevant publications) and drug class (derived based on the related terms described in the relevant publications). Full list of protein names, drug/ligand names, disease names and drug classes are separately provided in the PharmGED main web-page for facilitating the search of particular entries.

Moreover, keyword-based text search is also supported. The search is case insensitive and wildcards are supported. In a query, a user can specify full name or any part of the name in a text field. Wild character of ‘*’ and ‘?’ is allowed in text field. Here, ‘?’ represents any single character.
and '*' represents a string of characters of any length. For example, input of 'dehydrogenase' in the field of protein/gene name enables the finding of all entries containing 'dehydrogenase' in the protein name, such as 3-oxo-5-alpha-steroid 4-dehydrogenase 2, Alcohol dehydrogenase 1B, Alcohol dehydrogenase 1C, Fatty aldehyde dehydrogenase, Glucose-6-phosphate 1-dehydrogenase, NAD(P)H dehydrogenase (quinone) 1, etc. On the other hand, input of Fatty* dehydrogenase enables the finding of all dehydrogenases whose names start with 'Fatty'.

The result of a typical search is illustrated in Figure 1. In this interface, all entries that satisfy the specified search criteria are listed along with protein name, polymorphism rules, drug/ligand name, drug classification, disease name and links to other related entries in this database.

![Figure 1. The interface displaying a search result on PharmGED. All entries that satisfy the specified search criteria are listed along with protein name, polymorphism rules, drug/ligand name, drug classification, disease name and links to other related entries in this database.](image_url)
Bioinformatics & Drug Design group [BIDD]

PharmGED: Pharmacogenetic Effect Database

PharmGED is intended to specifically provide information about the effects of a particular protein polymorphism, non-coding region mutation, splicing alteration, or expression variation on the response of a particular drug. It currently contains 1,825 entries covering 266 distinct proteins, 693 polymorphisms, 414 drugs/ligands, and links to other related entries in this database. More detailed information of an entry can be obtained by clicking the corresponding protein name. The result is displayed in an interface shown in Figure 2. From this interface, one finds the accession number, name, sequence and function of protein, pharmacogenetic polymorphism, affected drugs and drug class, corresponding disease condition and pharmacogenetic effect. Moreover, the information about the related references and links to the literature database PUBMED (22) is also provided.

Figure 2. Interface displaying the detailed information of an entry in PharmGED.

POTENTIAL APPLICATIONS

Established links between polymorphisms of drug-related proteins and individual drug responses have been used in combination with genetic studies as indicators for predicting individual variations of drug response (23–27). Based on the statistical analysis of the data of polymorphisms and variation of drug response of the participating patients, simple rules may be derived in some cases for predicting individual variations of drug response from polymorphism data (23,24,26,28,29). These simple rules may be collected and used for developing...
<table>
<thead>
<tr>
<th>Protein</th>
<th>Drugs and treatment/action</th>
<th>Drug responses</th>
<th>Polymorphism rules and year of report</th>
<th>Number of patients with polymorphism</th>
<th>Reported percentage of patients who have a polymorphism and showed the expected effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cytochrome P450 1A2</td>
<td>Antipsychotic agents for schizophrenia patients</td>
<td>Tardive dyskinesia</td>
<td>Bsp120I (C→A) polymorphism in CYP1A2 gene, 2000 (38)</td>
<td>85</td>
<td>69</td>
</tr>
<tr>
<td>Cytochrome P450 2D6</td>
<td>Neuroleptic agents for chronic schizophrenic patients</td>
<td>Tardive dyskinesia</td>
<td>CYP2D6*4 genotype, 1998 (39)</td>
<td>13</td>
<td>81</td>
</tr>
<tr>
<td>UDP-glucuronosyltransferase</td>
<td>Capecitabine/irinotecan for the treatment of metastatic colorectal cancer</td>
<td>Greater antitumor response with low toxicity</td>
<td>UGT1A7*2/2 genotype, 2005 (27)</td>
<td>6</td>
<td>100</td>
</tr>
<tr>
<td>UDP-glucuronosyltransferase I</td>
<td>Tranilast for the prevention of restenosis following coronary revascularization</td>
<td>Hyperbilirubinemia</td>
<td>UGT1A7*3 genotype, 2005 (27)</td>
<td>7</td>
<td>100</td>
</tr>
<tr>
<td>N-acetyltransferase 2</td>
<td>Isoniazid for the prophylaxis and treatment of tuberculosis</td>
<td>ADRs such as peripheral neuritis, fever and hepatic toxicity</td>
<td>SA type (NAT2*6/<em>6, NAT2</em>6/<em>7 and NAT2</em>7/*7), 2002 (25)</td>
<td>6</td>
<td>83</td>
</tr>
<tr>
<td>Tryptophan hydroxylase</td>
<td>Fluvoxamine for the treatment of depression</td>
<td>Antidepressant response</td>
<td>A218C A/C phenotypes, 2001 (40)</td>
<td>107</td>
<td>76</td>
</tr>
<tr>
<td>Norepinephrine transporter</td>
<td>Milnacipran for the treatment of depression</td>
<td>Antidepressant response</td>
<td>A218C C/C phenotypes, 2001 (40)</td>
<td>70</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>A218C A/A phenotypes, 2001 (40)</td>
<td>40</td>
<td>65</td>
</tr>
<tr>
<td>Serotonin transporter</td>
<td>Serotonin reuptake inhibitors for the treatment of depression</td>
<td>Antidepressant response</td>
<td>s/s genotype of serotonin transporter gene promoter region, 2000–2004 (1,2,41,42)</td>
<td>11–72</td>
<td>54% at sixth week</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>s/l genotype of serotonin transporter gene promoter region, 2000–2004 (1,2,41,42)</td>
<td>2–47</td>
<td>55% at sixth week</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>l/l genotype of serotonin transporter gene promoter region, 2000–2004 (1,2,41,42)</td>
<td>4–16</td>
<td>48% at sixth week</td>
</tr>
<tr>
<td>Multidrug resistance-associated protein 1</td>
<td>Epileptic drugs for the treatment of epilepsy</td>
<td>Drug-resistant epilepsy</td>
<td>ABCB1 C3435T C/C genotype, 2003 (43)</td>
<td>73</td>
<td>75</td>
</tr>
<tr>
<td>Multidrug resistance-associated protein 1</td>
<td>Combination therapy of nelfinavir, efavirenz and nucleoside reverse transcriptase inhibitors for HIV-1 infected children</td>
<td>Virologic response by week 8</td>
<td>ABCB1 C3435T C/T genotype, 2003 (43)</td>
<td>169</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ABCB1 C3435T T/T genotype, 2003 (43)</td>
<td>73</td>
<td>53</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MDR1 C3435T C/C genotype, 2005 (44)</td>
<td>51</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MDR1 C3435T C/T genotype, 2005 (44)</td>
<td>33</td>
<td>91</td>
</tr>
</tbody>
</table>
a computer prediction system in a fashion similar to that of the HIV drug-resistant genotype interpretation systems (30).

Table 1 gives examples of the drug-related proteins in PharmGED with available information about pharmacogenetic polymorphism and drug-response variation from which a reasonably accurate rule have been derived in the literature for predicting responses to a specific drug or drug group. The reported percentage of patients who have a polymorphism and showed the expected effect is also given. Based on the test of the patients described in these reports, most of these rules are capable of predicting drug responses at success rates of 50–100%, which are not too much lower than and in many cases comparable to the accuracies of 81–97% for predicting HIV drug resistance mutations from the HIV-resistant genotype interpretation systems (30). This suggests that these simple rules have certain level of capacity for facilitating the prediction of pharmacogenetic effects and they may be used as the basis for developing more sophisticated interpretation systems similar to those of the HIV-resistant genotype interpretation systems (30).

PharmGED and other databases (14–21) can be potentially used for facilitating the generation of these rules. For instance, one entry of PharmGED describes that patients using classical neuroleptic such as fluphenazine, haloperidol have a higher incidence (81%) of Tardive dyskinesia if they are of the genotype CYP2D6*4. The corresponding polymorphism can be obtained from a link provided in the database. These data combined with other information in PharmaGED can be used to generate the rule for detecting this pharmacogenetic effect described in Table 1. In a second example, another entry of PharmGED describes that the polymorphism C3435T (Ile1145Ile) of protein MDR1-3435 variant is associated with different virologic response of nelfinavir in HIV-1 infected children. Fifty-nine percent of the 31 C/C genotype and 91% of the 33 C/T genotype show virologic response at eighth week, respectively. These data can then be used to generate the rule for this pharmacogenetic effect as described in Table 1.

CONCLUDING REMARKS

Knowledge about protein polymorphisms and drug responses appears to have reached a meaningful level for facilitating pharmacogenomic study and for predicting various types of individual variations of drug responses. Specialized pharmacogenetics databases serve as convenient resources for obtaining the relevant information. With the rapid development of genomics (31), pharmacokinetics (32–35) and pharmacogenomics (6,8,9), more information about drug-related proteins, polymorphisms and variations of drug responses are expected to become available. Moreover, progress in the study of proteomics (36) and pathways (37) related to drug-related proteins will further facilitate our understanding of the mechanism of individual variations in drug response.

ACKNOWLEDGEMENTS

This work was supported in part by grants from Singapore ARF R-151-000-031-112. The Open Access publication charges for this article were waived by Oxford University Press.

Conflict of interest statement. None declared.

REFERENCES


