Pharmacological Reports
2011, 63, 924–934
ISSN 1734-1140

Non-neutral nonsynonymous single nucleotide polymorphisms in human ABC transporters: the first comparison of six prediction methods

Da Cheng Hao[1], Yao Feng[1], Rongrong Xiao[1], Pei Gen Xiao[2]

[1]Laboratory of Biotechnology, College of Environment, Dalian Jiaotong University, Dalian 116028, China

[2]Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100193, China

Correspondence: Da Cheng Hao, e-mail: hao@djtu.edu.cn; Pei Gen Xiao, e-mail: xiaopg@public.bta.net.cn

Abstract:
Nonsynonymous single nucleotide polymorphisms (nsSNPs) in coding regions that can lead to amino acid changes may cause alteration of protein function and account for susceptibility to disease and altered drug/xenobiotic response. Abundant nsSNPs have been found in genes coding for human ATP-binding cassette (ABC) transporters, but there is little known about the relationship between the genotype and phenotype of nsSNPs in these membrane proteins. In addition, it is unknown which prediction method is better suited for the prediction of non-neutral nsSNPs of ABC transporters. We have identified 2,172 validated nsSNPs in 49 human ABC transporter genes from the Ensembl genome database and the NCBI SNP database. Using six different algorithms, 41 to 52% of nsSNPs in ABC transporter genes were predicted to have functional impacts on protein function. Predictions largely agreed with the available experimental annotations. Overall, 78.5% of non-neutral nsSNPs were predicted correctly as damaging by SNAP, which together with SIFT and PolyPhen, was superior to the prediction methods Pmut, PhD-SNP, and Panther. This study also identified many amino acids that were likely to be functionally critical but have not yet been studied experimentally. There was significant concordance between the predicted results of SIFT and PolyPhen. Evolutionarily non-neutral (destabilizing) amino acid substitutions are predicted to be the basis for the pathogenic alteration of ABC transporter activity that is associated with disease susceptibility and altered drug/xenobiotic response.

Key words:
phenotype, SNAP, PolyPhen, SIFT, Panther, Pmut, SNP, ABC transporter

Introduction

ATP-binding cassette (ABC) transporters are members of a protein superfamily that is one of the largest and most ancient families, with representatives in all extant phyla from prokaryotes to humans [19]. ABC transporters are transmembrane proteins that utilize the energy of adenosine triphosphate (ATP) hydrolysis to carry out certain biological processes, including translocation of substrates across membranes and non-transport-related processes such as translation of RNA and DNA repair [14, 42]. A wide variety of substrates are transported across extra- and intracellular membranes, including metabolic products, lipids and sterols, and drugs. Proteins are classified as ABC transporters based on the sequence and organization of their ABC domain(s).

There are 49 known ABC transporters in humans, which are classified into seven families by the Human Genome Organization (http://nutrigene.4t.com/human-abc.htm). Nonsynonymous SNPs (nsSNPs) of the human ABC transporter genes may cause absent or reduced transport activity, and polymorphisms of ABC transporters have been found to be closely related to altered drug clearance and/or drug response, cystic fibrosis, and various eye diseases and syndromes [21]. For example, nsSNPs of ABCB1 (P-glycoprotein) influence highly active antiretroviral therapy (HAART) efficacy and AIDS-free survival [17].

Human genetic variation may directly or indirectly influence responses to modern antiretroviral therapies for HIV. It is already known that some immunogenetic and other human genetic variations affect the natural history of HIV disease progression where individuals are untreated, but less information is available as to whether these differences are still relevant in the context of HAART [3]. Antiretroviral therapy adds additional opportunities for human genetic contributions to affect variable prognosis, in particular for those genes that influence pharmacokinetics and/or adverse events (e.g., ABC transporter genes). To date, the majority of studies investigating the influence of human genetic variation on HIV disease and treatment outcome have focused on a small number of SNPs, not including most of the ABC transporters.

The functional impact of most nsSNPs in human ABC transporter genes is still unknown. However, computational technologies aid the experimental exploration of nsSNPs and are indispensable in predicting the response to HAART in HIV-infected subjects. For example, Hao et al. [15] identified 923 nsSNPs from human phase II xenobiotic metabolizing enzyme genes and used SNAP, Panther, and PolyPhen to predict the impact of these nsSNPs on enzyme function. In addition, Wang et al. [41] predicted the phenotype of 1,632 nsSNPs of human ABC transporters using SIFT and PolyPhen. However, the numbers of nsSNPs in ABC transporter genes collected in public databases and publications are rapidly increasing, and new algorithms with better predictive power are becoming available. The present assessment of the presumably functionally essential residues of drug/xenobiotic transporters remains far less complete. In this study, we therefore investigated the potential effect of known human ABC transporter nsSNPs on protein function using six algorithms. The data set we compiled is the largest and most diverse one to date for the evaluation of prediction methods that require similar types of inputs.

Materials and Methods

Nonsynonymous SNP datasets

The data on human ABC transporter genes were collected from Ensembl (http://www.ensembl.org/Homo_sapiens/Search/) and Entrez Gene on the NCBI website (http://www.ncbi.nlm.nih.gov/sites/entrez). Expired and merged gene names were excluded from the study. The majority of the variants included in this analysis were identified during the screening of 12 human *ABCA*, 11 *ABCB*, 13 *ABCC*, 4 *ABCD*, 1 *ABCE*, 3 *ABCF*, and 5 *ABCG* genes from Ensembl (http://www.ensembl.org/Homo_sapiens/Gene/Variation_ Gene/). Ensembl integrates genetic variants from dbSNP, Uniprot, the personal genomes of Watson and Venter, and Illumina human genome sequencing results and links to information such as transcripts, population genetics, individual genotype, genomic context, phenotype data, and phylogenetic context. Information including gene symbol, gene name, mRNA accession number (ENST or NM), protein accession number (ENSP or NP), SNP ID, amino acid residue 1 (wild-type, wt), amino acid position, and amino acid residue 2 (missense) were collected. Supplementary variants were identified from Entrez Gene on NCBI and through PubMed literature searching and added to the dataset after cross-examination. The information on the effect of the nsSNPs on enzyme activity and the correlation between the nsSNPs and disease/adverse drug reaction/toxicant intake were extracted from *in vivo* and *in vitro* experiments (e.g., recombinant protein analysis) according to the literature.

Prediction of the phenotypes of nsSNPs in human ABC transporter genes

The effect of the variant amino acid substitution on protein function was predicted using PolyPhen (http://genetics.bwh.harvard.edu/pph/), Panther (http://www.pantherdb.org/tools/csnpScoreForm.jsp), SNAP (http://cubic.bioc.columbia.edu/services/SNAP/submit.html), SIFT (http://sift.jcvi.org/#), Pmut (http://mmb2.pcb.ub.es:8080/PMut/), and PhD-SNP (http://gpcr2.biocomp.unibo.it/cgi/predictors/PhD-SNP/PhD-SNP.cgi). The table with the above information is available upon request.

PolyPhen uses empirically derived rules based on previous research in protein structure, interaction, and