

Probabilistic Orthology Analysis of the ATP-Binding Cassette Transporters: Implications for the Development of Multiple Drug Resistance Phenotype^S

Ciaran Fisher, Tanya Coleman, and Nick Plant

Centre for Toxicology, Faculty of Health and Medical Sciences, University of Surrey, Guildford, Surrey, United Kingdom (C.F., N.P.); and Clinical Pharmacology and Drug Metabolism and Pharmacokinetics, AstraZeneca Clinical Development, Alderley Park, Macclesfield, Cheshire, United Kingdom (T.C.)

Received February 13, 2012; accepted April 16, 2012

ABSTRACT:

Drug transporters are rapidly becoming recognized as central to determining a chemical's fate within the body. This action is a double-edged sword, protecting the body from toxicants, but also potentially leading to reduced clinical efficacy of drugs through multiple drug resistance phenotype. To examine the interrelationship of this superfamily, we have constructed phylogenetic trees over an extended evolutionary distance representing each of the seven subfamilies. In addition, using protein sequences from species important in the design and evaluation of novel chemicals, namely human, macaque, rat, mouse, and dog, we have undertaken probabilistic orthology analysis to examine speciation probabilities within this phylogeny. These data allow us to accurately

predict orthologous sequences across these species, an important confirmatory step with implications for cross-species extrapolation of data during drug safety testing. Finally, we present the first complete phylogeny for subfamilies within humans constructed using the entire coding sequences, at both the DNA and protein levels. We demonstrate for the first time that genes associated with the multiple drug resistance phenotype cluster separately from other genes within the same subfamily, suggestive of a conserved, fundamental, difference in these proteins. Such work may help guide future studies on the mechanisms underlying multiple drug resistance as well as the development of novel therapeutic approaches to mitigate against its development.

Introduction

The ATP-binding cassette (ABC) genes comprise a superfamily with representatives found in all characterized eukaryotic and prokaryotes; indeed, this superfamily encodes approximately 5% of the *Escherichia coli* genome (Fath and Kolter, 1993; Davidson and Chen, 2004). The majority of ABC genes encode membrane-bound transport proteins, which act to move polar molecules across the nonpolar lipid membrane, using the hydrolysis of ATP. As such, these transporters play an important role in the absorption, distribution, metabolism, and excretion (ADME) of chemicals (Glavinas et al., 2004). In prokaryotes, ABC transporters may act as both importer and exporter proteins (Fath and Kolter, 1993; Davidson and Chen, 2004). By contrast, in eukaryotes, these proteins act solely as export transporters, and this represents an important functional breakpoint within the superfamily. Such efflux is central to the removal of potentially harmful chemicals from cell systems; an action that undoubtedly underlies the biological survival advantage conferred by these pro-

teins and explains their conservation across evolutionary time (Dean and Annilo, 2005).

Although the ability to rapidly eliminate potentially harmful chemicals has obvious survival advantages, it also represents a challenge during long-term chemotherapy. Expression levels of a number of ABC transporters has been shown to contribute to the development of multidrug resistance (MDR) phenotype, whereby the ADME of administered chemicals is altered, usually resulting in altered pharmacokinetics and reduced clinical efficacy. MDR has been shown to have a negative impact on the treatment of a number of disease states, including cancer (Deeley et al., 2006; Gillet et al., 2007). Thus, much work has thus been undertaken to understand the molecular mechanisms underlying MDR and how this can be mitigated during long-term chemotherapy (Coley, 2008). However, translation of these mitigation strategies to the clinic has generally been poor, and MDR still represents a significant hurdle to successful chronic chemotherapy regimens (Coley, 2008; Tiwari et al., 2011).

Previous studies on the evolution of the ABC genes have not included all the species used in the preclinical testing of novel drugs; such a comprehensive analysis would be important for the robust extrapolation of data from preclinical test species to humans. In addition, phylogenetic analysis has often been restricted to only fragments of the total coding sequence [such as the ATP-binding domain, which is likely to be the least variable domain (Dean and Allikmets,

This work was supported by AstraZeneca/UK Biotechnology and Biological Sciences Research Council [Grant BB/E527671/1].

Article, publication date, and citation information can be found at <http://dmd.aspetjournals.org>.

<http://dx.doi.org/10.1124/dmd.112.045062>.

^S The online version of this article (available at <http://dmd.aspetjournals.org>) contains supplemental material.

ABBREVIATIONS: ABC, ATP-binding cassette; ADME, absorption, distribution, metabolism, and excretion; MDR, multidrug resistance; RefSeq, reference sequences; MPR, most parsimonious reconciliation; TMD, transmembrane domain.

1995; Saier and Paulsen, 2001; Dean and Annilo, 2005)] and has relied on distance-based methodologies that are generally accepted to not produce the most robust phylogenetic relationships across superfamilies (Koski and Golding, 2001).

In the current study, we have used protein alignments of all members of the ABC drug transporter family in humans and a number of important model animals for the testing of novel chemicals to undertake probabilistic orthology analysis. This allows the robust assignment of ortho- or paralog status to protein pairs, including probability values, thus providing important information for extrapolation of effects between species. In addition, we have expanded on previous phylogenetic studies by using full ABC nucleotide and protein sequences across a range of evolutionarily diverse species. Such an approach not only separates ABC subfamilies correctly, but it also clusters those ABC transporters associated with MDR, suggesting that this is a specifically evolved design function.

Materials and Methods

Sequence Identification and Alignment of ABC Genes. The reference sequences (RefSeq) for all 48 members of the human ABC superfamily were identified (Supplemental Table 1) along with their common alternate names and GenBank accession number. Subfamilies A to D and G represent those sequences that encode membrane-bound drug transport proteins, whereas subfamilies E and F encode ATP-binding cassette proteins that do not play a role in drug transport. Using these human sequences as the query, a cross-species megaBLAST was undertaken to identify similar sequences in 10 other species (Altschul et al., 1990; Zhang et al., 2000). Where multiple sequences against a single query were identified within a species, all sequences were taken forward, with duplicates trimmed when identified.

Multiple sequence alignments were performed for each subfamily using ClustalX, with an additional alignment being performed solely for the human sequences (Higgins and Sharp, 1988; Thompson et al., 1997): for each subfamily, an *E. coli* ABC gene sequence was included as the root for the subsequent phylogenetic tree. All alignments were assessed manually to ensure that gap insertions were sensible, with subsequent phylogenetic analysis also automatically removing regions containing significant gaps.

Phylogenetic Analysis of ABC Subfamilies. Distance matrices were produced from the ClustalX output using dnadist, including 100 replicates for bootstrap analysis (Felsenstein, 1997). For individual subfamily analysis, these matrices were analyzed with Neighbor using an unpaired grouping method with arithmetic mean to produce a consensus tree using the majority rule (Felsenstein, 1997).

For analysis of all subfamilies within humans, ClustalX-generated alignments were analyzed with Phylml, because this process is more computationally efficient for larger analysis (Felsenstein, 1989; Guindon and Gascuel, 2003). One hundred phylogenetic trees were constructed using a maximum-likelihood approach, and Consense was used to generate a consensus tree from the bootstrapped data (Felsenstein, 1989).

Probabilistic Orthology Analysis of ABC Transporter Proteins from Human, Mouse, Rat, Dog, and Macaque. Probabilistic orthology analysis is a reconciliation-based orthology methodology that uses the probabilistic gene evolution model described by Arvestad et al. (2003). The output determines the probability that any given divergence within a phylogenetic tree is the result of speciation as opposed to duplication. This allows the robust assignment of orthologs and paralogs within a tree. Protein sequences for each ABC drug transporter from human, dog, rat, mouse, and macaque were aligned using ClustalX2 (version 2.0.12), and ProTest was used to determine that the Le and Gascuel (LG) algorithm was the optimum amino acid replacement model for accurate phylogeny resolution (Abascal et al., 2005; Le and Gascuel, 2008). A maximum-likelihood phylogenetic tree was generated with Phylml, which was the input for the probabilistic orthology analysis. In addition to a phylogenetic relationship of genes, a phylogenetic relationship of species was also required for probabilistic orthology analysis, and this was constructed using species divergence times taken from Kumar and Hedges (1998) and Jacobs and Downs (1994). Probabilistic orthology analysis was undertaken using PrimeGEM, as described in Sennblad and Lagergren (2009), using 10,000 MCMC iterations,

MCMC-estimated duplication and loss rates, and with an output of posterior orthology probabilities. This output was analyzed using the MCMC_analysis perl script available from http://prime.sbc.su.se/primeGEM/downloads/perl/mcmc_analysis.

Results

Phylogenetic Analysis of ABC Subfamilies. Using a distance-based method, we were able to generate nucleotide-level phylogenetic trees for each subfamily (Supplemental Figs. 1–5). Table 1 represents a summary of these data, describing the total number of ABC genes identified in each species, plus the number of sequences for each species that could be clearly demonstrated to lie within a single subfamily through phylogenetic analysis.

Using distance-based methods, it is only appropriate to assign individual sequences within a subfamily, and not to predict orthologs. Assignment of orthologs based purely upon phylogenetic trees derived by most parsimonious reconciliation (MPR) has been demonstrated to have poor predictive value (Koski and Golding, 2001). A more appropriate analysis is probabilistic orthology analysis (Sennblad and Lagergren, 2009). In this method, gene evolution is set within the context of a species evolution tree and modeled under variable gene birth-death rate parameters (Rannala and Yang, 2007).

For each human ABC protein, we present the most probable ortholog for macaque, rat, mouse, and dog (Supplemental Table 2), along with the probability score for that match. In general, the given nomenclature for each protein is consistent with the indicated analysis; for example, the protein named ABCA6 in humans and mice is demonstrated to be the result of a speciation event, and hence are orthologs, with a probability of over 99%. However, some ortholog assignments are not fully supported by the probabilistic orthology analysis: for example, in the case of ABCD2, probabilistic orthology analysis assigns high probabilities to the sequence being an ortholog of either ABCD2 or ABCD3. In such cases, sequences are tentatively assigned as orthologs to the human sequence for which there is the higher probability, but alternate high probability matches are noted within the table (Supplemental Table 2) and phylogeny (Supplemental Fig. S6). Finally, we are able to expand the current knowledge base, assigning ortholog status to several “orphan” sequences, particularly from the macaque, which has been relatively poorly investigated until now.

In addition, probabilistic orthology analysis provides an estimate of the most probable gene loss/duplication rates that would result in the

TABLE 1

Total number of individual genes identified for each species

Using human ABC transcript RefSeq sequences as the query term, a cross-species megaBLAST was undertaken to identify homologs in each of the listed species. Significant hits, not inclusive of transcript variants or pseudogenes, were included in an initial phylogenetic analysis, allowing designation of sequences into most probable subfamilies. Numbers in the “?” column indicate sequences identified in the megaBLAST search that cannot be conclusively identified as belonging to a specific subfamily.

Species Name	Subfamily							?	Total
	A	B	C	D	E	F	G		
<i>Homo sapiens</i>	12	11	12	4	1	3	5		48
<i>Pan troglodytes</i>	11	11	11	4	1	3	5	2	48
<i>Macaca mulatta</i>	13	11	12	4	1	3	3		47
<i>Canis lupus familiaris</i>	13	9	12	4	1	2	5	2	46
<i>Mus musculus</i>	14	11	11	4	2	2	5	2	51
<i>Rattus norvegicus</i>	14	10	9	4	1	2	5	3	48
<i>Bos taurus</i>	11	9	10	3	1	3	5	3	45
<i>S. purpuratus</i>	1	4	6	1	1	2	2	12	29
<i>Danio rerio</i>	5	5	9	3	1	2	5	9	40
<i>C. elegans</i>			1	1	1	1		29	32
<i>D. melanogaster</i>				2	1	1	1	38	43

given phylogeny. For each combination of duplication and loss rates examined within the analysis, the probabilistic likelihood is determined, producing the three-dimensional Gaussian distribution seen in Fig. 1. The maximal posterior probability of the analysis represents the most likely duplication and loss rate estimates, which were 0.0095Myr^{-1} and 0.0122Myr^{-1} , respectively. It is important to note, that these estimates are relevant for only the species set examined (human, rat, mouse, dog, and macaque), with duplications and loss rates often being considerably different over larger evolutionary distances.

Phylogenetic Analysis of the ABC Superfamily in Humans.

After examination of phylogenetic relationships within subfamilies across a number of animal species relevant to preclinical chemical testing, we next examined the relationship between human subfamilies. Phyml was used to confer a logical consensus tree with acceptable bootstrap values at major nodes (Fig. 2). This analysis was able to successfully resolve the subfamilies, and it provides further insight into the evolution of this superfamily. Unsurprisingly, subfamilies E and F, which lack a transmembrane domain (TMD), cluster separately from other subfamilies. However, they diverge at different points within the phylogenetic tree, suggesting that these two TMD-lacking subfamilies have arisen as independent loss-of-function (TMD) events.

We have also observed that those sequences that encode proteins associated with the MDR phenotype appear to segregate within the phylogeny. Within the B subfamily, a clear separation of ABCBs 1, 4, 5, and 11 can be seen from the rest of the subfamily (Fig. 2), with the former group all being previously demonstrated to play a role in the development of MDR (Childs et al., 1998; Ambudkar et al., 1999; Smith et al., 2000; Huang et al., 2004). In addition, ABCG2 separates from the rest of the G subfamily, and it is the only subfamily member associated with the MDR phenotype (Cervenak et al., 2006). Finally, the C subfamily also divides into those sequences associated with the MDR phenotype and those that are not, although in the latter case the segregation is not as clear. It is of interest to note that the ABCC9 and ABCC8 sequences appear to segregate with the MDR phenotype

group, lending further weight to the suggestion that these genes may indeed contribute to a drug resistance phenotype (Deeley et al., 2006; Zhou et al., 2008).

To complement the nucleotide-level analysis, we also undertook a protein-level analysis for the human ABC transporters. In this analysis, we also included the protein sequences from the preclinical species rat, mouse, dog, and macaque. The derived phylogenetic tree is consistent with the ortholog assignment, with tight clustering of orthologs at the end of branches (Fig. S6). In addition, simplification of the phylogenetic tree to illustrate the overall structure (Fig. 3) is consistent with the conclusions drawn from the nucleotide-level analysis of human sequences. Although the multispecies amino acid-level analysis produces a different tree topology, clustering of MDR-associated sequences is still observed within the phylogeny (Fig. 3).

Discussion

Proteins encoded by the A to D and G subfamilies of the ABC transporter superfamily play a central role in chemical ADME, affecting the pharmacokinetic profile and potential clinical efficacy of a compound (Hembruff et al., 2008; Kalliokoski and Niemi, 2009). It is important to understand the phylogenetic relationships of this superfamily for two reasons. First, the testing of novel chemical entities for both efficacy and toxicity is routinely undertaken in nonhuman mammalian species, with the data extrapolated to humans (Barille, 2008). For such extrapolations to be undertaken, complex physiologically based pharmacokinetic models have been developed (Dressman et al., 2011). However, at present, the role of drug transport proteins is poorly represented in many of these models, often being either encompassed in a generic “active transport” term or limited to very few specific transporters (Pang et al., 2009; Fan et al., 2010). One reason for this limitation is that the relationship between transporters in preclinical test species and humans is still relatively poorly understood, and as such this work aids in the focus of experimental work to identify kinetic differences between orthologs. Once coupled with data on species differences in transporter expression (Takahashi et al., 2008; Cedernaes et al., 2011), this will allow far more robust cross-species extrapolation of drug ADME (Glavinas et al., 2004) in model species and humans. Second, understanding the mechanisms underlying MDR is an important step in identifying potential means to mitigate this important limitation to chemotherapeutic intervention (Coley, 2008).

Phylogenetic analysis over an extended evolutionary distance allows the clear assignment of sequences to subfamilies for mammalian species. However, for species with a larger divergence time from humans, such as *Strongylocentrotus purpuratus*, *Caenorhabditis elegans*, and *Drosophila melanogaster*, robust subclassification is not possible in the majority of cases. Work by Sheps et al. (2004) also attempted to identify human orthologs for ABC drug transporters in *C. elegans*, using the amino acid sequence of ABC proteins. They were able to assign orthologs to 8 of 43 human ABC drug transporters, and in the current study we confirm five of these using a different analysis methodology. The use of the entire coding sequence for phylogenetic analysis has several benefits over the use of only a selected region (e.g., TMD or ATP binding domain). First, use of only a portion of the coding sequence excludes any variability seen within the rest of the sequence, which may result in bias in the generated phylogenies. Second, although the TMD is most likely to be the most variable region, and hence main driver for the phylogenetic trees, use of this alone would exclude the non-TMD containing subfamilies (ABCE and ABCF), reducing the completeness of the analysis. Third, for robust assignment of orthologs through probabilistic orthology anal-

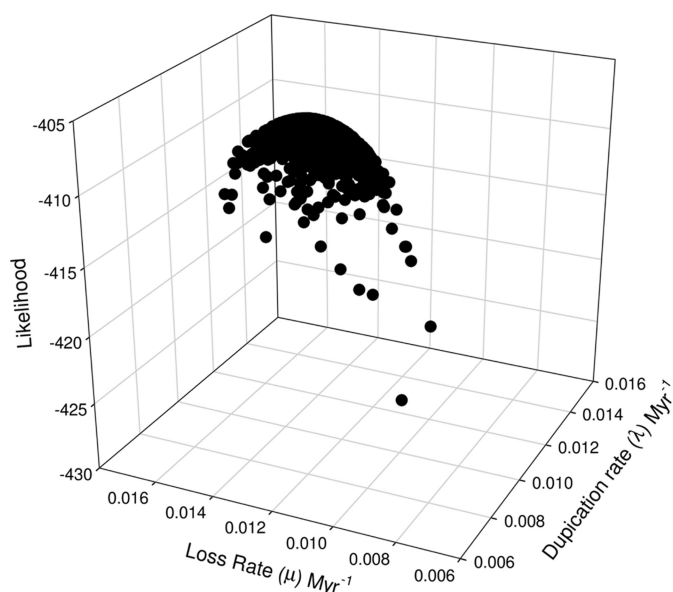


FIG. 1. Gene duplication and loss likelihood scores within the ABC transporter encoding subfamilies. Probabilistic orthology analysis was undertaken for ABC transporter proteins, using the gene evolution method of Arvestad et al. (2003). MCMC-estimated duplication and loss rates were calculated for every 10 iterations of a 10,000 interaction analysis.

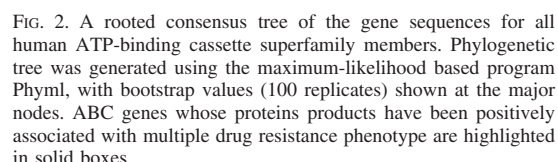


FIG. 2. A rooted consensus tree of the gene sequences for all human ATP-binding cassette superfamily members. Phylogenetic tree was generated using the maximum-likelihood based program Phym, with bootstrap values (100 replicates) shown at the major nodes. ABC genes whose proteins products have been positively associated with multiple drug resistance phenotype are highlighted in solid boxes

potential confounders in these studies, and, in general, the averages are still lower than the estimate for the ABC superfamily derived herein. These high duplication and loss rates support the assignment of ortholog status via probabilistic analysis, as opposed to a simple MPR-based approach; Sennblad and Langergren (2009) demonstrated that the rate of false orthology predictions from an MPR-based approach increased with the duplication and loss rates. The presence of significantly higher duplications and loss rates for the ABC superfamily could be reflective of a fluid phylogeny that can alter relatively rapidly, which would be logical for a protein family providing protection against chemicals in an ever-changing environmental milieu.

In comparison with previous publications, we have aligned the entire mRNA/protein sequence for phylogenetic analysis, rather than selected fragments. We demonstrate that robust phylogenies can be inferred from the alignment of full gene sequences (Supplemental

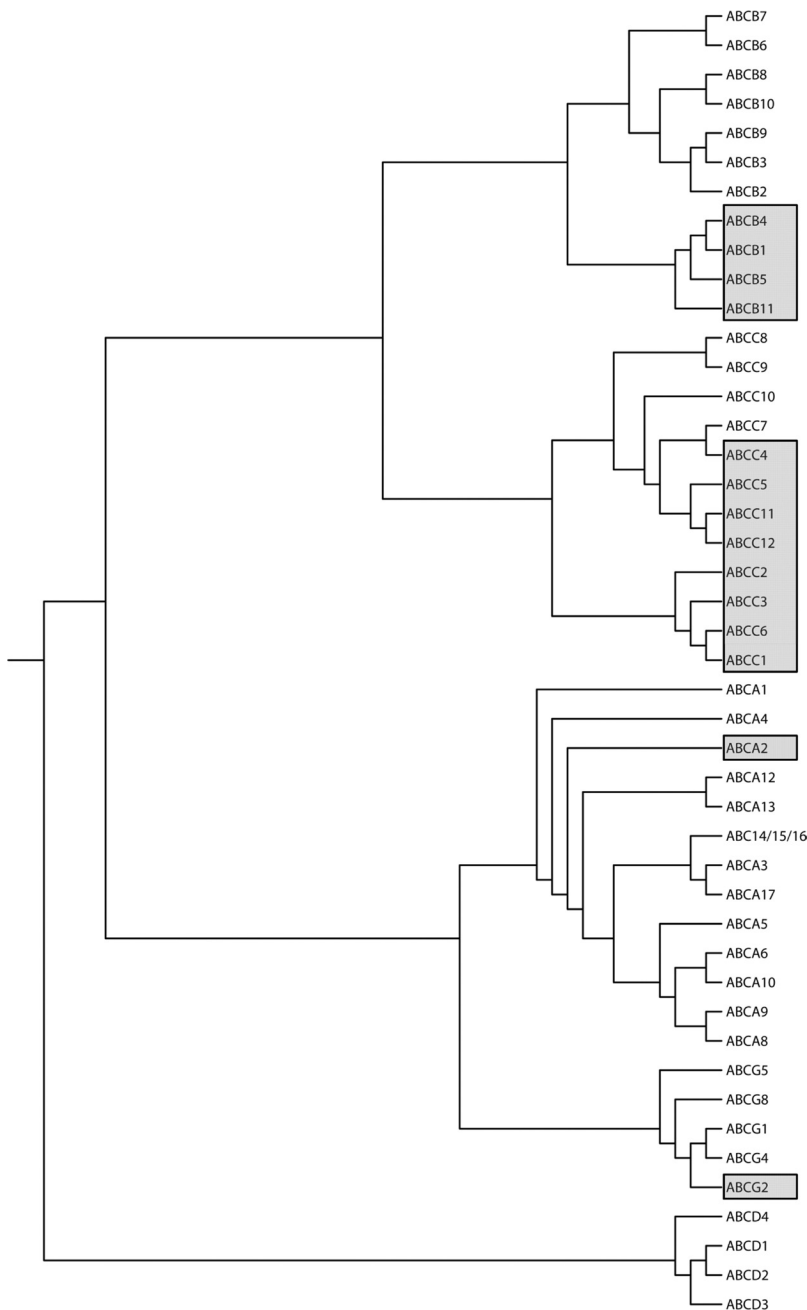


FIG. 3. A rooted consensus tree of the protein sequences subfamilies A, B, C, D, and G of the ATP-binding cassette superfamily in humans, macaques, rat, mouse, and dog. A multiple alignment was generated of ABC proteins from human, macaque, rat, mouse, and dog using ClustalX2. Optimum amino acid replacement model was determined by ProTest, and then an LG algorithm and Phym1 were used to generate a phylogenetic tree using a maximum-likelihood approach. The full tree is shown in Supplemental Fig. S6, with a simplified cartoon showing only the overall general structure shown here. ABC genes whose proteins products have been positively associated with multiple drug resistance phenotype are highlighted in solid boxes.

Figs. 1–5), distinguishing between full and half transporters within subfamilies. In addition, for a single species, human, we have reconstructed the entire superfamily. It is possible to successfully resolve the individual subfamilies, and some interesting implications arise from this analysis. As noted within the introduction, two subfamilies within the ABC superfamily do not encode drug transporters, and indeed it has been argued that these genes should be excluded from the superfamily (Rees et al., 2009). We demonstrate that these two subfamilies have arisen by independent events, most probably through loss of the TMD. This loss of TMD has obviously led to an altered localization of protein products from these subfamilies, whereas their retention of an ATP-binding domain allows them to undertake ATP-dependent processes. In the case of the ABCE subfamily, the sole gene encodes a ribonuclease L inhibitor, an important regulator of interferon action (Bisbal et al., 1995). In the case of the three ABCF gene products, these proteins are all members of the GCN20 family

and appear to play roles in tumor necrosis factor α -mediated signaling (Richard et al., 1998).

In addition, we provide data to support a clustering of those genes that encode transporters associated with MDR phenotype. This clustering, supported by both transcript and protein level analysis, could indicate that the MDR-associated genes/proteins have features that set them apart from other genes within their subfamilies. Because this relationship translates to the protein level, it also suggests that these features may be important in determining the molecular function(s) required to contribute to an MDR phenotype, although these features are as yet unelucidated. Whereas this separation from the main subfamily can be seen clearly in the B and G groups, it is undoubtedly less well defined within the C subfamily. This result may be of interest considering that the B and G subfamily members encode proteins that generally have parent chemicals as their substrates, whereas those transporters encoded by the C subfamily generally transport conju-

gated products of metabolism (Choi, 2005; Deeley et al., 2006). Further examination is required to fully understand the impact of these different roles in chemical ADME and on the development of a MDR phenotype.

In summary, the phylogenetic analyses contained herein extend current data on ABC gene orthologs in preclinical species, both identifying novel orthologs and correcting previous errors in annotation. Such information is important for the extrapolation of chemical effects in model organisms to humans and hence accurate risk assessment. In addition, we present the first complete human phylogeny across the entire superfamily, demonstrating segregation between sequences that encode ABC transporters evoking the MDR phenotype and those which do not, at both the gene and protein level. This leads to the exciting possibility of focusing further on those transporters most likely to result in MDR and the development of strategies to mitigate this process.

Authorship Contributions

Participated in research design: Plant, Fisher, and Coleman.
Conducted experiments: Plant and Fisher.
Contributed new reagents or analytic tools: Coleman.
Performed data analysis: Plant and Fisher.
Wrote or contributed to the writing of the manuscript: Plant, Fisher, and Coleman.

References

- Abascal F, Zardoya R, and Posada D (2005) ProtTest: selection of best fit models of protein evolution. *Bioinformatics* **21**:2104–2105.
- Altschul SF, Gish W, Miller W, Myers EW, and Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* **215**:403–410.
- Ambudkar SV, Dey S, Hrycyna CA, Ramachandra M, Pastan I, and Gottesman MM (1999) Biochemical, cellular, and pharmacological aspects of the multidrug transporter. *Annu Rev Pharmacol Toxicol* **39**:361–398.
- Arvestad L, Berglund AC, Lagergren J, and Sennblad B (2003) Bayesian gene/species tree reconciliation and orthology analysis using MCMC. *Bioinformatics* **19**:i7–i15.
- Barille FA (2008) *Principles of Toxicology Testing*. CRC Press, Boca Raton.
- Bisbal C, Martinand C, Silhol M, Lebleu B, and Salehzada T (1995) Cloning and characterization of a RNase L inhibitor. A new component of the interferon-regulated 2–5A pathway. *J Biol Chem* **270**:13308–13317.
- Cedernaes J, Olszewski PK, Almén MS, Stephansson O, Levine AS, Fredriksson R, Nylander O, and Schiöth HB (2011) Comprehensive analysis of localization of 78 solute carrier genes throughout the subsections of the rat gastrointestinal tract. *Biochem Biophys Res Commun* **411**:702–707.
- Cervenak J, Andrikovics H, Ozvegy-Laczka C, Tordai A, Nemet K, Váradi A, and Sarkadi B (2006) The role of the human ABCG2 multidrug transporter and its variants in cancer therapy and toxicology. *Cancer Lett* **234**:62–72.
- Childs S, Yeh RL, Hui D, and Ling V (1998) Taxol resistance mediated by transfection of the liver-specific sister gene of P-glycoprotein. *Cancer Res* **58**:4160–4167.
- Choi CH (2005) ABC transporters as multidrug resistance mechanisms and the development of chemosensitizers for their reversal. *Cancer Cell Int* **5**:30.
- Coley HM (2008) Mechanisms and strategies to overcome chemotherapy resistance in metastatic breast cancer. *Cancer Treat Rev* **34**:378–390.
- Cotton JA and Page RD (2005) Rates and patterns of gene duplication and loss in the human genome. *Proc Biol Sci* **272**:277–283.
- Davidson AL and Chen J (2004) ATP-binding cassette transporters in bacteria. *Annu Rev Biochem* **73**:241–268.
- Dean M and Allikmets R (1995) Evolution of ATP-binding cassette transporter genes. *Curr Opin Genet Dev* **5**:779–785.
- Dean M and Annilo T (2005) Evolution of the ATP-binding cassette (ABC) transporter superfamily in vertebrates. *Annu Rev Genomics Hum Genet* **6**:123–142.
- Deeley RG, Westlake C, and Cole SP (2006) Transmembrane transport of endo- and xenobiotics by mammalian ATP-binding cassette multidrug resistance proteins. *Physiol Rev* **86**:849–899.
- Dressman JB, Thelen K, and Willmann S (2011) An update on computational oral absorption simulation. *Expert Opin Drug Metab Toxicol* **7**:1345–1364.
- Fan J, Chen S, Chow EC, and Pang KS (2010) PBPK modeling of intestinal and liver enzymes and transporters in drug absorption and sequential metabolism. *Curr Drug Metab* **11**:743–761.
- Fath MJ and Kolter R (1993) ABC transporters: bacterial exporters. *Microbiol Rev* **57**:995–1017.
- Felsenstein J (1989) Phylip-Phylogeny Interface Package (Version 3.2). *Cladistics* **164**–166.
- Felsenstein J (1997) An alternating least squares approach to inferring phylogenies from pairwise distances. *Syst Biol* **46**:101–111.
- Gillet JP, Efferth T, and Remacle J (2007) Chemotherapy-induced resistance by ATP-binding cassette transporter genes. *Biochim Biophys Acta* **1775**:237–262.
- Glavinas H, Krajcsi P, Cserepes J, and Sarkadi B (2004) The role of ABC transporters in drug resistance, metabolism and toxicity. *Curr Drug Deliv* **1**:27–42.
- Guindon S and Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* **52**:696–704.
- Hembruff SL, Laberge ML, Villeneuve DJ, Guo BQ, Veitch Z, Cecchetto M, and Parissenti AM (2008) Role of drug transporters and drug accumulation in the temporal acquisition of drug resistance. *BMC Cancer* **8**:318.
- Higgins DG and Sharp PM (1988) CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. *Gene* **73**:237–244.
- Huang Y, Anderle P, Bussey KJ, Barbacioru C, Shankavaram U, Dai Z, Reinhold WC, Papp A, Weinstein JN, and Sadée W (2004) Membrane transporters and channels: role of the transporter in cancer chemosensitivity and chemoresistance. *Cancer Res* **64**:4294–4301.
- Jacobs LL and Downs WR (1994) The evolution of murine rodents, in *Rodents and Lagomorph Families of Asian Origin and Diversification* (Tomida Y, Li C, and Setoguchi T eds), pp 149–156, National Science Museum Monograph, Tokyo.
- Kalliokoski A and Niemi M (2009) Impact of OATP transporters on pharmacokinetics. *Br J Pharmacol* **158**:693–705.
- Koski LB and Golding GB (2001) The closest BLAST hit is often not the nearest neighbor. *J Mol Evol* **52**:540–542.
- Kumar S and Hedges SB (1998) A molecular timescale for vertebrate evolution. *Nature* **392**:917–920.
- Le SQ and Gascuel O (2008) An improved general amino acid replacement matrix. *Mol Biol Evol* **25**:1307–1320.
- Lynch M and Conery JS (2003) The evolutionary demography of duplicate genes. *J Struct Funct Genomics* **3**:35–44.
- Lynch M and Conery JS (2000) The evolutionary fate and consequences of duplicate genes. *Science* **290**:1151–1155.
- Pang KS, Maeng HJ, and Fan J (2009) Interplay of transporters and enzymes in drug and metabolite processing. *Mol Pharm* **6**:1734–1755.
- Rannala B and Yang Z (2007) Inferring speciation times under an episodic molecular clock. *Syst Biol* **56**:453–466.
- Rees DC, Johnson E, and Lewinson O (2009) ABC transporters: the power to change. *Nat Rev Mol Cell Biol* **10**:218–227.
- Richard M, Drouin R, and Beaulieu AD (1998) ABC50, a novel human ATP-Binding cassette protein found in tumor necrosis factor- α -stimulated synovial cells. *Genomics* **53**:137–145.
- Saier MH and Paulsen IT (2001) Phylogeny of multidrug transporters. *Semin Cell Dev Biol* **12**:205–213.
- Sennblad B and Lagergren J (2009) Probabilistic orthology analysis. *Syst Biol* **58**:411–424.
- Sheps JA, Ralph S, Zhao ZY, Baillie DL, and Ling V (2004) The ABC transporter gene family of *Caenorhabditis elegans* has implications for the evolutionary dynamics of multidrug resistance in eukaryotes. *Genome Biol* **5**:R15.
- Smith AJ, van Helvoort A, van Meer G, Szabo K, Welker E, Szakacs G, Varadi A, Sarkadi B, and Borst P (2000) MDR3 P-glycoprotein, a phosphatidylcholine translocase, transports several cytotoxic drugs and directly interacts with drugs as judged by interference with nucleotide trapping. *J Biol Chem* **275**:23530–23539.
- Takahashi M, Washio T, Suzuki N, Igeta K, Fujii Y, Hayashi M, Shirasaka Y, and Yamashita S (2008) Characterization of gastrointestinal drug absorption in cynomolgus monkeys. *Mol Pharm* **5**:340–348.
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, and Higgins DG (1997) The CLUSTAL-X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **25**:4876–4882.
- Tiwari AK, Sodani K, Dai CL, Ashby CR Jr, and Chen ZS (2011) Revisiting the ABCs of multidrug resistance in cancer chemotherapy. *Curr Pharm Biotechnol* **12**:570–594.
- Zhang Z, Schwartz S, Wagner L, and Miller W (2000) A greedy algorithm for aligning DNA sequences. *J Comput Biol* **7**:203–214.
- Zhou SF, Wang LL, Di YM, Xue CC, Duan W, Li CG, and Li Y (2008) Substrates and inhibitors of human multidrug resistance associated proteins and the implications in drug development. *Curr Med Chem* **15**:1981–2039.

Address correspondence to: Dr. Nick Plant, Centre for Toxicology, Faculty of Health and Medical Sciences, University of Surrey, Guildford, GU2 7XH, UK.
 E-mail: n.plant@surrey.ac.uk