Pure Appl. Chem., Vol. 74, No. 11, pp. 2057–2081, 2002.

ABCG subfamily of human ATP-binding cassette proteins*

Stefan Lorkowski[1,‡] and Paul Cullen[1,2]

[1]*Institut für Arterioskleroseforschung, Domagkstraße 3, 48149 Münster, Germany;*
[2]*Ogham GmbH, Mendelstraße 11, 48149 Münster, Germany*

Abstract: ATP-binding cassette (ABC) proteins form one of the largest known protein families and have been found in all known organisms. Most members of the human ABC protein family are membrane-spanning transporters that use energy derived from the hydrolysis of ATP to transport specific substrates across cell membranes. Mutations in certain human ABC transporters of the subfamilies A, B, C, and D have been shown to cause a wide variety of inherited diseases such as the lung condition cystic fibrosis, the nervous degenerative condition adrenoleukodystrophy (of *Lorenzo's Oil* fame), hereditary macular degeneration of the eye (Stargardt's disease), and inherited deficiency of circulating high-density lipoproteins (Tangier disease or familial hypoalphalipoproteinemia). Very recent studies showed that mutations in two members of the subfamily G of human ABC transporters (ABCG5 and ABCG8) cause a condition called sitosterolemia in which plant sterols accumulate in the body and may be responsible for influencing total body sterol homeostasis. In addition, other members of the subfamily G, namely ABCG1 and ABCG4, have also been shown to be involved in cellular lipid trafficking and are thought to play important roles during foam cell formation of human macrophages. By contrast, ABCG2 is a multidrug resistance transporter. In this review, we focus on the current knowledge and physiological background of the members of the subfamily G. We also present new insights on the evolutionary relationship of human and nonhuman ABCG proteins.

ABBREVIATIONS USED IN THIS MANUSCRIPT

ABC	ATP-binding cassette (transporter/protein)
ABCP	placenta-specific ATP-binding cassette transporter
AML	acute myeloid leukemia
apoE	apolipoprotein E
BCRP	breast cancer resistance protein
CETP	cholesterol ester transfer protein
CYP7A1	cholesterol 7α-hydroxylase
EST	expressed sequence tag
HDL	high-density lipoprotein
LAL	lysosomal acid lipase (or acid cholesterol ester hydrolase)
LDL	low-density lipoprotein
LXR	liver X receptor
Lxra	mouse liver x receptor α

Pure Appl. Chem.* **74, 2021–2081 (2002). A collection of invited, peer-reviewed articles by the winners of the 2002 IUPAC Prize for Young Chemists.
‡Corresponding author: Institut für Arterioskleroseforschung, Domagkstraße 3, 48149 Münster, Germany; Tel.: +49 (251) 83 5 20 72; Fax: +49 (251) 83 5 20 62; E-mail: stefan.lorkowski@uni-muenster.de

Lxrb	mouse liver x receptor β
MRX	mitoxantrone resistance protein
PC	phosphatidylcholine
PE	phosphatidylethanolamine
PLTP	phospholipid transfer protein
PPARγ	peroxisome proliferator-activated receptor γ
RXR	retinoid X receptor
SNP	single nucleotide polymorphism
SREBP-1c	sterol regulatory element-binding protein-1c
TGFβ1	transforming growth factor β1
TM	transmembrane helix
TMD	transmembrane domain
ZNF202	zinc finger protein 202

INTRODUCTION

ABC proteins form one of the largest known protein families and are found in all eukaryotes and prokaryotes [1]. So far, approximately 1100 different proteins belonging to this family have been described in the literature and public sequence databases. For example, the genomes of the five eukaryotes *Homo sapiens*, *Mus musculus*, *Drosophila melanogaster*, *Saccharomyces cerevisiae*, and *Caenorhabditis elegans* contain about 230 different ABC proteins with different protein domain arrangements (Spitzer et al., manuscript in preparation).

ABC proteins are characterized by a highly conserved and well-defined cytoplasmic domain, the eponymous ATP-binding cassette (denoted ABC or sometimes NBD, nucleotide-binding domain) that is often associated with a hydrophobic TMD. The ABC consists of three highly conserved sequences, the Walker A and B motifs and a so-called ABC signature, or Walker C motif [2–4]. The signature motif is specific to ABC proteins and distinguishes them from other ATP-binding proteins. These three domains are required for the binding and hydrolysis of ATP, which in turn provides energy for the transport of a large variety of substrates across the membranes of the cell and of subcellular organelles.

The TMD of ABC proteins consists of between four and eight—usually six—consecutive membrane-spanning helices [5–7]. It is noteworthy that some ABC proteins lack a transmembrane-spanning region and are therefore thought to be involved in processes other than transport, as is the case with some DNA-unwinding enzymes (helicases) and inhibitors of the ubiquitous enzyme RNase that degrades RNA within the cells [5–7]. Typically, ABC transporter proteins pump substrates in a single direction.

Owing to their modular construction, single ABC proteins show wide variety in their domain architecture. So far, the following arrangements of domains in ABC transporters have been described: single ABC and TMD (both only found in prokaryotes), ABC-ABC, TMD-TMD (only found in prokaryotes), ABC-TMD, TMD-ABC, ABC-TMD-ABC-TMD, TMD-ABC-TMD-ABC, and TMD_0-TMD-ABC-TMD-ABC, of which the last three are exclusively found in eukaryotes [7].

An active ABC transporter protein complex is thought to consist of two ABC domains and at least two transmembrane regions encoded either by a single gene or by multiple genes encoding different domains [7]. Thus, in eukaryotes, ABC full transporters display a domain arrangement of ABC-TMD-ABC-TMD or TMD-ABC-TMD-ABC. ABC half transporters also exist with either the ABC-TMD or TMD-ABC arrangement; these form either homo- or heterodimers with other half transporter proteins to form functional transporter complexes. ABC proteins consisting of only two ABC domains have also been found in eukaryotes [4]. Functional phylogenetic studies have indicated that prokaryotic ABC transporter proteins can either be importer or exporter systems, whereas all eukaryotic ABC transporter proteins are exporters [8,9].

In view of the large number of ABC proteins that exist, it is not surprising that they play many roles in metabolism. To date, ABC proteins have been shown to function as multidrug resistance proteins and as transporters of metal ions, lipids, peptides, and several other molecules [1]. Interest in these proteins has been increasing steadily, as demonstrated by more than 6500 citations in PubMed, because a number of inherited diseases of humans are caused by mutations in ABC proteins. These diseases include HDL deficiency or Tangier disease (mutations in ABCA1), the neural wasting disorder adrenoleukodystrophy that was the subject of the film *Lorenzo's Oil* (mutations in ABCD1), and the lung disease cystic fibrosis (mutations in ABCC7), which affects about one person in about 2000 (for an overview, see Table 1) [5,6,10–14].

Despite this interest, little information is available on substrate and ligand specificity, three-dimensional structure, transport, and energizing mechanisms. The physiological role of these proteins is also far from being understood, in particular, in the case of eukaryotic or human ABC proteins. These deficiencies in our knowledge are important, because a detailed understanding of ABC proteins might contribute to the more effective treatment of important human diseases [5,6,10–14].

HUMAN ABC PROTEIN FAMILY

The human ABC protein family currently strictly comprises 48 members (recently, a 49th member of the human ABC protein family, ABCC13, has been identified that lacks the ATP-binding cassette) and can be divided into seven distinct subfamilies, based on structural organization of domains and amino acid homology [5,6,13,15]. An overview on the principle domain architecture of the different human ABC protein subfamilies is shown in Fig. 1, and a list of the human ABC genes is displayed in Table 1.

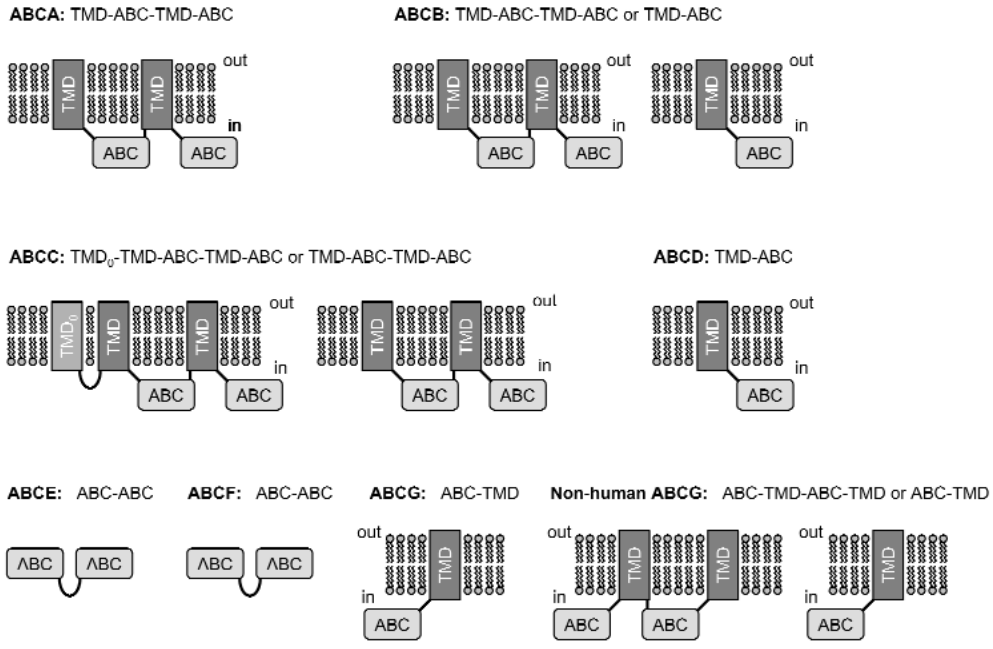


Fig. 1 Domain arrangement of human ABC proteins. TMDs are shown as dark gray rectangles, and ABC domains are shown as light gray rectangles with rounded corners. The basic system of active ABC transporters is a four-subunit system consisting of two TMDs and two ABC domains. Single ABC proteins that do not contain these four domains are formed by different proteins encoded either by the same gene or by different genes. It is noteworthy that some human ABC transporters in eukaryotes can contain additional insertions, for example, between the first ABC domain and the second TMD (e.g., ABCA1 or CFTR).

Table 1 List of all human ABC proteins, their chromosomal location, domain architecture, and putative function [5,6,13,15].

Subfamily	Gene	Alias[a]	Location	Domain arrangement[b]	Expression	Putative function	Phenotype
ABCA	*ABCA1*	*ABC1*	9q31.1	TMD-ABC-TMD-ABC	Ubiquitous	Phospholipid efflux	Tangier disease[c]
(ABC1)	*ABCA2*	*ABC2*	9q34	TMD-ABC-TMD-ABC	Brain	Lipid transport	–
	ABCA3	*ABC3, ABC-C*	16p13.3	TMD-ABC-TMD-ABC	Lung	–	–
	ABCA4	*ABCR, ABC10*	1p22.1–p21	TMD-ABC-TMD-ABC	Retina	N-retinylidene–PE efflux	Stargardt's disease 1[d]
	ABCA5	–	17q24.3	TMD-ABC-TMD-ABC	Muscle, heart, testes	–	–
	ABCA6	–	17q24.3	TMD-ABC-TMD-ABC	Liver	–	–
	ABCA7	–	19p13.3	TMD-ABC-TMD-ABC	Spleen, thymus	–	–
	ABCA8	–	17q24	TMD-ABC-TMD-ABC	Ovary	Drug resistance	–
	ABCA9	–	17q24	TMD-ABC-TMD-ABC	Heart	–	–
	ABCA10	–	17q24	TMD-ABC-TMD-ABC	Muscle, heart	–	–
	ABCA12	–	2q35	TMD-ABC-TMD-ABC	Stomach	–	–
	ABCA13	–	7p11–q11	TMD-ABC-TMD-ABC	Low in all tissues	–	–
ABCB	*ABCB1*	*PGY1, MDR* 7	p21.1	TMD-ABC-TMD-ABC	Adrenal, kidney, brain	Multidrug resistance	Ivermectin susceptibility
(MDR)	*ABCB2*	*TAP1, ABC17*	6p21.3	TMD-ABC	All cells	Peptide transport	–
	ABCB3	*TAP2, ABC18*	6p21.3	TMD-ABC	All cells	Peptide transport	Bare lymphocyte syndrome type I[e]
	ABCB4	*PGY3, ABC21*	7q21.1	TMD-ABC-TMD-ABC	Brain, liver	PC transport	Progressive familial intrahepatic cholestasis[f]
	ABCB5	–	7p14	TMD-ABC-TMD-ABC	Ubiquitous	–	–
	ABCB6	*MTABC3, ABC14*	2q36	TMD-ABC	Mitochondria	Iron transport	–
	ABCB7	*ABC7*	Xq12–q13	TMD-ABC	Mitochondria	Fe/S cluster transport	Sideroblastic anemia and spinocerebellar ataxia
	ABCB8	*M-ABC1*	7q36	TMD-ABC	Mitochondria	–	–
	ABCB9	–	12q24	TMD-ABC	Heart, brain	–	–
	ABCB10	*MTABC2*	1q42	TMD-ABC	Mitochondria	Peptides?	–
	ABCB11	*SPGP, ABC16*	2q24	TMD-ABC-TMD-ABC	Liver	Bile salt transport	Progressive familial intrahepatic cholestasis 2
ABCC	*ABCC1*	*MRP1*	16p13.1	TMD_0-TMD-ABC-TMD-ABC	Lung, testes	Drug resistance	–
(CFTR/MRP)	*ABCC2*	*MRP2*	10q24	TMD_0-TMD-ABC-TMD-ABC	Liver, intestine, kidney	Organic anion efflux	Dubin–Johnson syndrome[g]
	ABCC3	*MRP3*	17q21.3	TMD_0-TMD-ABC-TMD-ABC	Lung, intestine, liver	Drug resistance	–
	ABCC4	*MRP4*	13q32	TMD-ABC-TMD-ABC	Prostate	Nucleoside transport	–
	ABCC5	*MRP5*	3q27	TMD-ABC-TMD-ABC	Ubiquitous	Nucleoside transport	–
	ABCC6	*MRP6*	16p13.1	TMD_0-TMD-ABC-TMD-ABC	Kidney, liver	Peptides?	Pseudoxanthoma elasticum

Table 1 (*Continued*).

Subfamily	Gene	Alias[a]	Location	Domain arrangement[b]	Expression	Putative function	Phenotype
	ABCC7	*CFTR*	7q31.2	TMD_0-TMD-ABC-TMD-ABC	Exocrine tissue	Chloride ion channel	Cystic fibrosis (or mucoviscidosis)[h]
	ABCC8	*SUR*	11p15.1	TMD_0-TMD-ABC-TMD-ABC	Pancreas	Sulfonylurea receptor	Nesidioblastosis of pancreas[i]
	ABCC9	*SUR2*	12p12.1	TMD_0-TMD-ABC-TMD-ABC	Heart, muscle	Sulfonylurea receptor	–
	ABCC10	*MRP7*	6p21	TMD_0-TMD-ABC-TMD-ABC	Low in all tissues	–	–
	ABCC11	*MRP8*	16q12.1	TMD-ABC-TMD-ABC	Low in all tissues	–	–
	ABCC12	*MRP9*	16q12.1	TMD_0-TMD-ABC-TMD-ABC	Low in all tissues	–	–
	ABCC13	–	21q11.2	TMD	Fetal liver, bone marrow	Hematopoiesis	–
ABCD	*ABCD1*	*ALD*	Xq28	TMD-ABC	Peroxisomes	Fatty acid transport	Adrenoleukodystrophy
(ALD)	*ABCD2*	*ALDL1, ALDR*	12q11–q12	TMD-ABC	Peroxisomes	Fatty acid transport	–
	ABCD3	*PXMP1, PMP70*	1p22–p21	TMD-ABC	Peroxisomes	Fatty acid transport	Zellweger syndrome 2
	ABCD4	*PMP69, P70R*	14q24.3	TMD-ABC	Peroxisomes	Fatty acid transport	–
ABCE	*ABCE1*	*OABP, RNS41*	4q31	ABC-ABC	Ovary, testes, spleen	Oligoadenylate binding	–
(OABP)							
ABCF	*ABCF1*	*ABC50*	6p21.33	ABC-ABC	Ubiquitous	–	–
(GCN20)	*ABCF2*	–	7q35–7q36	ABC-ABC	Ubiquitous	–	–
	ABCF3	–	3q25.1–3q25.2	ABC-ABC	Ubiquitous	–	–
ABCG	*ABCG1*	*ABC8, White*	21q22.3	ABC-TMD	Ubiquitous	Lipid transport?	–
(White)	*ABCG2*	*ABCP, MXR, BCRP*	4q22	ABC-TMD	Placenta, intestine	Drug resistance	–
	ABCG3	–	8p12 ?	ABC-TMD	Thymus, spleen	No human gene?	–
	ABCG4	*White2*	11q23.3	ABC-TMD	Brain, thymus, lung	Lipid transport?	–
	ABCG5	*White3, sterolin1*	2p21	ABC-TMD	Liver, intestine	Sterol transport	Sitosterolemia
	ABCG8	*White4, sterolin2*	2p21	ABC-TMD	Liver, intestine	Sterol transport	Sitosterolemia

[a]Due to limited space, not all aliases are listed. For further details, see <http://www.humanabc.org/> or <http://www.gene.ucl.ac.uk/nomenclature/genefamily/abc.html>.
[b]Based on SMART predictions [116–118].
[c]Familial hypoalphalipoproteinemia, familial hypoapoproteinemia.
[d]Fundus flavimaculatis, retinitis pigmentosa, age-related macular degeneration, cone-rod dystrophy.
[e]Immune deficiency.
[f]Recurrent intrahepatic cholestasis of pregnancy, immune deficiency.
[g]Hyperbilirubinemia 2.
[h]Congenital bilateral absence of the vas deferens, pancreatitits, bronchiectasis.
[i]Autosomal recessive hyperinsulinism, familial persitent hyperinsulinemic hyopglycemia of infancy.

The 7 human subfamilies are termed ABCA (12 members), ABCB (11 members), ABCC (12 members, plus ABCC13 that lacks the ATP-binding cassette), ABCD (4 members), ABCE (1 member), ABCF (3 members), and ABCG (5 members) using the HUGO Gene Nomenclature for human and mouse ABC transporter genes [16]. Several other sequences exist in the genome with homology to ABC genes, many of which are thought to represent pseudogenes [5].

A complete description of all human ABC proteins goes beyond the scope of this article. The reader is referred to the excellent reviews provided by Dean et al. [5,6] and the Web site on human ABC transporters [13].

HUMAN ABCG (WHITE) SUBFAMILY

Compared to the members of the other human ABC transporter subfamilies, the members of the ABCG subfamily are unique in their domain architecture. All human ABCG proteins are composed of ABC-TMD-like half transporters, although full transporters have been described in other organisms such as *S. cerevisiae* and *Arabidopsis thaliana* [17–19].

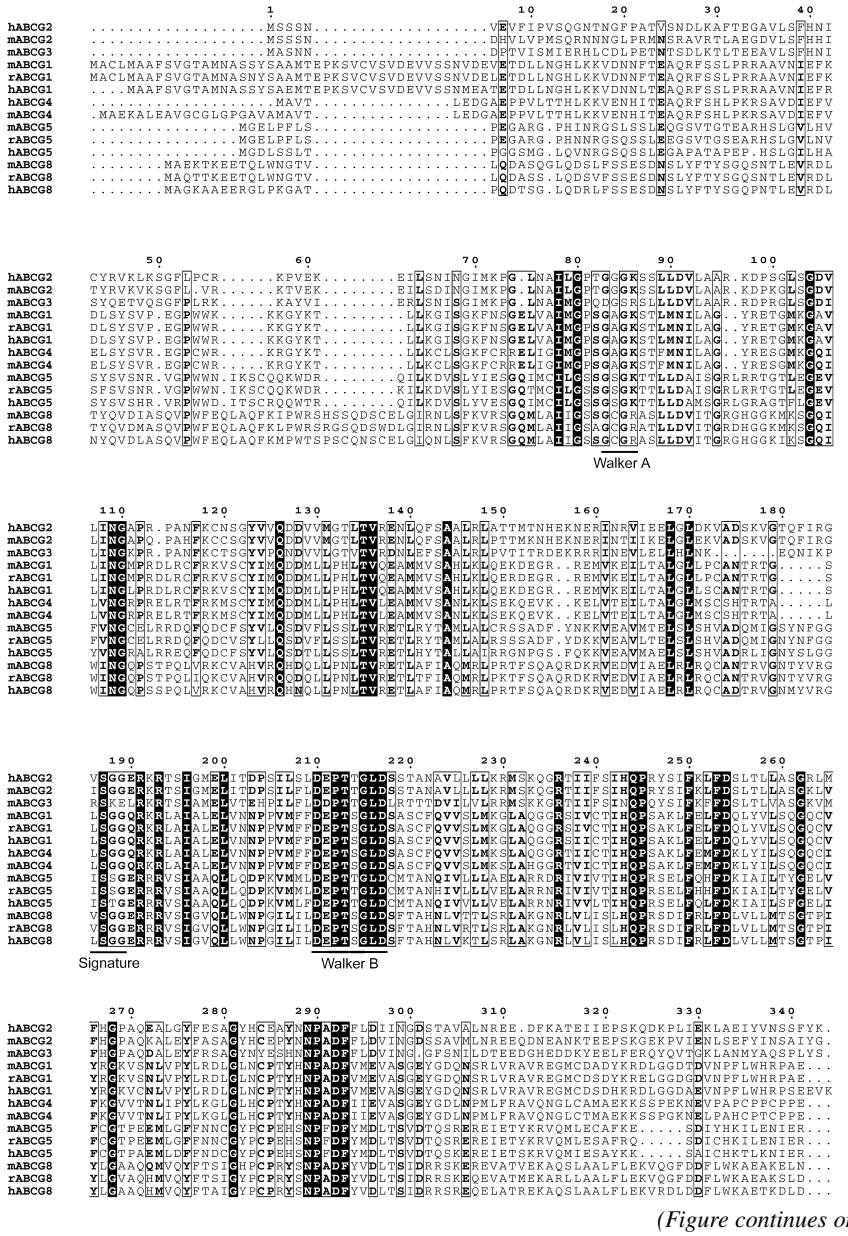
The human ABCG subfamily comprises five members (ABCG1, ABCG2, ABCG4, ABCG5, and ABCG8). A further mammalian ABCG protein, Abcg3, has as yet been found only in rodents and is closely related to the rodent Abcg2 protein and the human ABCG2 protein. Some have suggested that Abcg3 may also be contained within the human genome on chromosome 8p12 [20], although no sequence information is available reporting an *ABCG3* gene at this location.

Most of the human ABCG proteins seem to be involved in lipid and/or sterol metabolism. The ubiquitously expressed ABCG1 is thought to be involved in cellular cholesterol transport, and the closely related ABCG4 may have a similar function in the brain, where it is predominantly expressed. ABCG5 and ABCG8 are transporters of plant sterols in the intestine and liver. *ABCG2* is a drug resistance gene. To illustrate the degree of homology among ABCG transporters (Table 2), an alignment of all RefSeq protein sequences of human, mouse, and rat ABCG proteins is depicted in Fig. 2.

Most of the human ABCG proteins show less than 30 % identity and less than 50 % homology (Table 2). Several amino acid residues in the ABCG proteins of *H. sapiens*, *M. musculus*, and *Rattus norvegicus* are highly conserved in all 14 known ABCG proteins of these organisms and are different from both the Walker A and B motifs and the signature motif (see amino acids marked with black boxes in Fig. 2). These include the NP(A/F)DF (asn-pro-ala/phe-asp-phe) motif located between the Walker B motif and the first TM. It is noteworthy that most of these residues are located within or near the ATP-binding cassettes with only a few located within the TMD. However, we know nothing of the role of these residues in respect to protein function or structure. Our group is currently performing detailed

Table 2 Grade of homology and identity of human ABCG half transporter proteins (modified from [30]).

	ABCG1	ABCG2	Abcg3	ABCG4	ABCG5	ABCG8	
ABCG1		54	48	84	49	45	Homology
ABCG2	30		70	52	51	48	
Abcg3	27	55		46	43	43	
ABCG4	72	28	26		47	44	
ABCG5	26	29	24	27		50	
ABCG8	25	27	24	25	29		
	Identity						



(Figure continues on next page.)

Fig. 2 Alignment of the protein sequences of the currently known human, mouse, and rat ABCG proteins. Protein sequences were aligned using the ClustalW tool [113,114]. Alignments were formatted and visualized using ESPRIPT 2.0 [115]. Amino acids are shown in single-letter code, gaps due to optimal alignment are represented by dots. Identical amino acids are displayed in black boxes, homologous amino acids are boxed. Walker A, B, signature sequence motifs and the six membrane-spanning regions are underlined. TMs were predicted using the SMART 3.4 tool [116–118]. Note that only the core of the TMs is marked; the TMs for an individual protein shown may vary over about three amino acids. It is of note that the amino acid region of TMs 5 and 6 of ABCG8 and ABCG5, respectively, differ from the other ABCG proteins. TM 6 of ABCG5 is marked as a dotted line, and TMs 5 and 8 of ABCG8 are marked with dashed lines. It is noteworthy that several amino acid residues are highly conserved in vertebrate ABCG proteins. Interestingly, most of the conserved residues are located within the ABC, although they are different from the well-known Walker A and B, as well as the signature motif. The role of these conserved residues is not yet known.

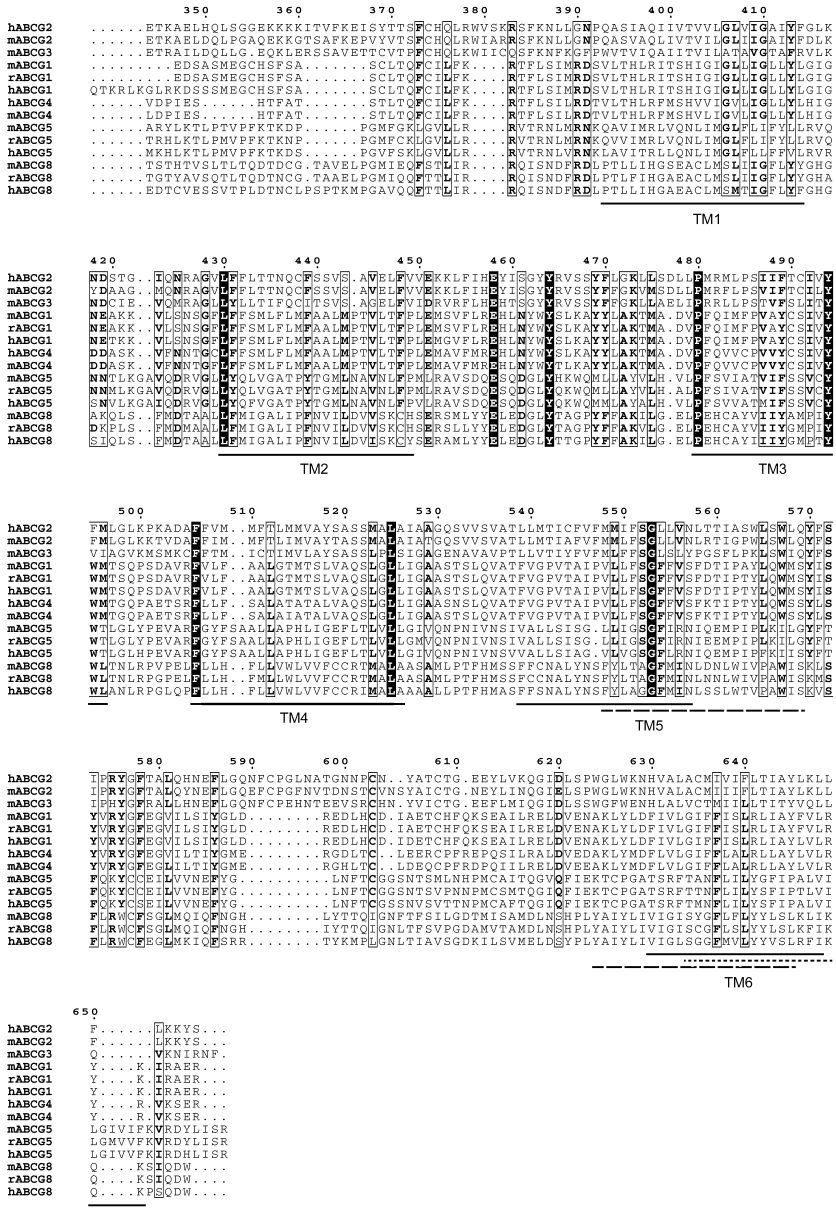


Fig. 2 *(Continued).*

systematic studies on the degree of conservation of these amino acid residues in all known ABCG-like proteins, in particular among the vertebrate ABCG homologs that differ from those mentioned above.

Evolution of ABCG proteins

Compared to other members of the ABC transporter subfamilies, the ABCG transporter subfamily is composed of "reverse" proteins that have an ATP-binding cassette at the amino terminus and a TM at the carboxy terminus. This is of interest in respect to the evolution of this protein subfamily. Two possible routes have been suggested by which ABCG protein may have arisen [21]. One is that the origi-

nal ABCG gene in which the ABC domain precedes the TMD arose from the fusion of independent regions encoding such domains. Based on new proteome data predicted from the genome sequence of *Dictyostelium discoideum*, Anjard et al. formulated a second alternative, namely that the ABCG gene may have arisen from the central region of a full transporter member of the A, B, or C subfamilies (TMD-ABC-TMD-ABC) by loss of the first TMD and the second ABC domain. Tandem duplication and fusion of ABCG half transporter genes may then have led to the full transporter proteins of the ABCG subfamily.

ABCG1/Abcg1—a lipid/cholesterol transporter?

ABCG1 (synonym: ABC8) was independently described by Chen et al., Croop et al., and Savary et al. as the homolog of the *Drosophila* white protein [22–24]. As an interesting historical point, the *Drosophila* white gene was the first locus mapped by Thomas Hunt Morgan, the founder of modern experimental genetics, in 1910 [25]. For this reason, the ABCG subfamily of the ABC protein family is sometimes called the white subfamily. However, Kuwana et al. described ATET, an ABC transporter expressed in the trachea of *D. melanogaster*, which shows higher similarity to ABCG1 than the *Drosophila* white protein [26]. Sequencing of the complete genome of *D. melanogaster* revealed evidence for at least two more ABCG-like proteins that show a higher degree of similarity to ABCG1 than the white protein. Chen et al. isolated the coding region of the human *ABCG1* gene and reported evidence of alternative splicing in the coding region (exon 20) for the first time (Fig. 3A) [22].

An avenue we used to investigate the role of ABCG1 was to determine the structure of the *ABCG1* gene and, in particular, to characterize the promoter, the region that regulates gene activity. In a first study, an ABCG1 transcript comprised of 15 exons was identified (transcript I in Fig. 3A) and a promoter lacking a TATA box was characterized [27]. This promoter is GC-rich and contains silencing elements that can mediate transcriptional repression [6]. Two further studies uncovered an unusual degree of complexity of the human *ABCG1* gene [28,29]. We found that the *ABCG1* gene contains five coding regions or exons that had not previously been described, and that it codes for a very large number of splice variants with putatively different amino acid sequences (transcripts I to V in Fig. 3A). In addition, we found that these alternatively spliced ABCG1 transcripts have different core promoters [28]. A further alternatively spliced ABCG1 transcript with a core promoter different from the ones previously described was identified by Kennedy et al. (transcript VI in Fig. 3A) [29]. Recent studies performed in our laboratory showed that only the wild-type transcript of the seven splice variants currently known is expressed at abundant levels in human monocyte-derived macrophages, suggesting that this is the most important transcript of the human *ABCG1* gene (unpublished results). It is noteworthy that this transcript is nearly identical to the murine *ABCG1* ortholog [28] and is very similar to the major transcript of the *ABCG4* gene (transcript I in Fig. 3B) [30]. At present, we are carrying out further studies to elucidate the details of the unusually high degree of complexity in the expression of the human *ABCG1* gene, including the tissue-specific expression of the different ABCG1 splice variants.

The human *ABCG1* gene has been mapped to chromosome 21q22.3 [22–24], in which several diseases including affective disorders such as depression and autosomal recessive nonsyndromic deafness DFNB10 have been located [31].

Several studies have focused on *ABCG1* as a positional candidate of affective disorders. A first report on variations in the *ABCG1* gene was provided by Chen et al. [22], who identified a DNA polymorphism with 62 % heterozygosity due to variation of a poly(T) region in the 3′-untranslated region of the *ABCG1* gene. Nakamura et al. described a G2457A polymorphism in the 3′-untranslated region of the human *ABCG1* gene and found significant associations between this polymorphism and the poly(T) polymorphism first described by Chen et al. with mood and panic disorders in Japanese males, but not in Japanese females [32]. In a study performed by Rujescu et al., the distribution of the G2457A polymorphism in patients with affective disorders, suicide attempters with various psychiatric diag-

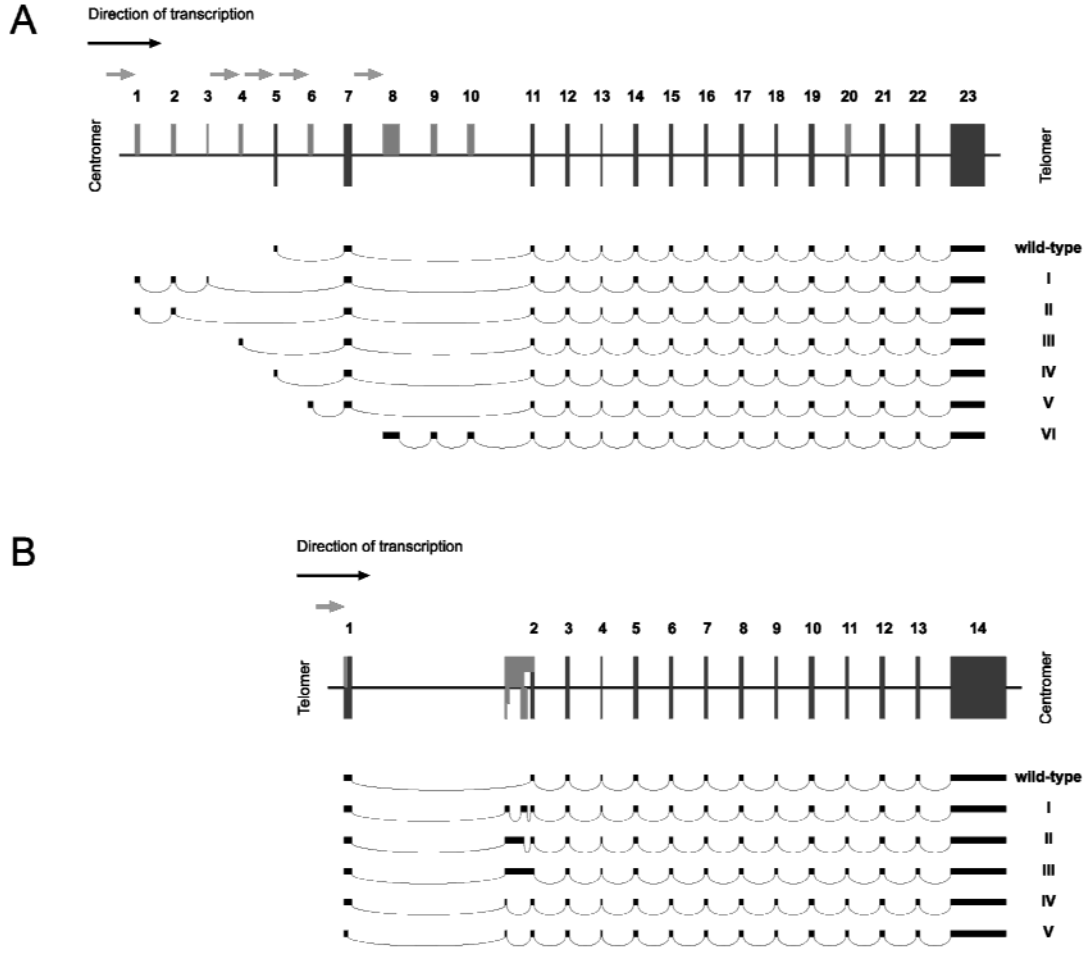


Fig. 3 Genomic structure of the human *ABCG1* and *ABCG4* genes. The human *ABCG1* gene **(A)** is comprised of 23 exons on chromosome 21q23.3 and spans about 97 kb. The wild-type mRNA is comprised of exons 5, 7, and 11–23. In human macrophages, several alternatively spliced transcripts (I to VI) have been identified [27–29]. Interestingly, these transcripts all have different core promoters that, in addition, also differ from the core promoter of the wild-type transcript. All known splice variants of the ABCG1 gene are expressed less abundantly than the wild-type transcript. The human *ABCG4* gene **(B)** consists of 14 exons, ten of which (exons 2–7 and 10–13) are identical in size to the corresponding exons of the ABCG1 gene (exons 5, 7, 11, 16, and 19–22). Two other exons have a similar size (exons 8 and 9 of the ABCG4 gene and exons 17 and 18 of the *ABCG1* gene, respectively). Exons 1 and 14 of the *ABCG4* gene and exons 5 and 23 of the ABCG1, respectively, contain the 5′- and 3′-untranslated regions, which are the major part of the mRNAs that are different in both genes. Similar or identical exons of both genes are aligned. So far, five different alternatively spliced transcripts encoded by the human *ABCG4* gene have been identified that are also expressed at lower levels compared to the wild-type transcript [94]. Expression of these transcripts seems to be tissue-specific: wild-type, brain; transcript I, lung; transcripts II to V, thymus. Despite the high degree of homology in the coding region of ABCG1 and ABCG4, no similarity between the thymus-specific transcripts of ABCG4 and the 5′-terminal exons of ABCG1 was found. Note that the size of introns is not scaled, whereas the size of exons is. Alternatively spliced mRNAs are shown below the genomic structure of each gene. Included exons are shown as black rectangles connected by black lines. Exons that only appear in alternatively splice transcripts are colored gray and shown in half-size or quarter-size of the exons forming the wild-type mRNAs. Relative positions of putative core promoters of wild-type and alternatively spliced transcripts are marked with gray arrows.

noses, and healthy subjects was investigated [33]. Initially, a trend toward a modest association with affective disorders in males was found. A repeat study with independent patients and controls, however, found no association with affective disorders, either in the replication or in the combined group. Furthermore, there was no association with suicidal behavior. These findings do not support the hypothesis that *ABCG1* is a susceptibility gene for affective disorders or suicidal behavior. In addition, a study performed by Johansson et al. revealed no association between the G2457A polymorphism and seasonal affective disorders in the Northern European population [34].

Kirov et al. screened the human *ABCG1* gene for polymorphisms and association with bipolar affective disorder 1 [35]. In this study, a total of 13 SNPs, a GCC repeat within exon 1 and two novel intronic variable number of tandem repeat polymorphisms were identified. Eight of the SNPs—the two variable number of tandem repeat polymorphisms, the GCC repeat, and two known microsatellite markers within the gene—were tested for association with bipolar affective disorder 1. No alleles or haplotypes were significantly preferentially transmitted from parents to affected offspring. However, the trend for preferential transmission of markers in the 3′-untranslated region was in the same direction as in the previous report of Nakamura et al. for association with mood and panic disorders and therefore warrant an attempt at replication. Marker-to-marker linkage disequilibrium showed that strong linkage disequilibrium was present over relatively short distances of up to 20 kb and was present for SNPs as well as for polymorphic repeats.

Besides affective disorders, the *ABCG1* gene was also screened as a positional candidate of nonsyndromic deafness. Berry et al. refined the locus of autosomal recessive nonsyndromic deafness DFNB10 to an area of less than 1 Mb between markers 1016E7.CA60 and 1151C12.GT45 on human 21q22.3 [31]. Due to the fact that among five other genes (trefoil factor 1, 2, and 3, phosphodiesterase 9A, and NADH-ubiquinone oxidoreductase flavoprotein 3), the *ABCG1* gene is localized in the redefined region, the authors determined the genomic structure of *ABCG1* in order to facilitate mutation analysis. All six genes in this region were screened and eliminated as candidates for DFNB10. Currently, therefore, there is no evidence that mutations or SNPs in the *ABCG1* gene may affect any of the diseases described above.

Expression of the *ABCG1* gene, as well as expression of many other ABC genes, was shown by us and others to be regulated during acetylated or oxidized LDL-mediated cholesterol loading (foam cell formation) and HDL-mediated lipid release in human macrophages [36–38] using an expression-profiling technique called differential display RT-PCR [39–41]. Recently, Venkateswaran et al. showed that ABCG1 mRNA levels in human macrophages are regulated by the LXR/RXR pathway [38], which recently has been shown to play an important role in lipid metabolism [42–46]. These two nuclear receptors bind naturally occurring substrates such as 22*R*-hydoxycholesterol (LXR ligand) and 9 *cis*-retinoic acid (RXR ligand). Binding of the ligands results in formation of an LXR/RXR heterodimer. The LXR/RXR dimer binds to specific regulator DNA sites, thus resulting in an induction of mRNA transcription. So far, LXR and RXR have been shown to be involved in the regulation of mRNA expression of the ABC transporters ABCA1 [45,47], ABCG1 [28,29,38], ABCG4 [48], ABCG5, and ABCG8 [48,49], and other proteins involved in lipid metabolism such as apoE [50,51], CETP [52], CYP7A1 [53,54], SREBP-1c [55], lipoprotein lipase [56], fatty acid synthase [57], and PLTP [58].

In vitro, foam cell formation of human monocyte-derived macrophages using acetylated or oxidized LDL results in an increase in the expression of ABCG1 [36–38], as well as several other ABC proteins such as ABCA1 and ABCG4 [48,59]. Peter A. Edwards' group showed that this increase in the expression of ABCG1 is mediated via oxysterols, known ligands of LXR. Therefore, it was not entirely surprising when expression of ABCG1 was found to be upregulated in macrophages isolated from individuals with Tangier disease, which is characterized by the accumulation of cholesterol within macrophages, even without loading of these cells with cholesterol [37].

In addition to these findings, we found that the expression of ABCG1 differed depending on the variant of apoE produced by the cells (unpublished results). The latter findings are of interest because (i) apoE is a major component of, for example, HDL, which is important for the reverse cholesterol

transport [46] and (ii) von Eckardstein et al. showed that, like ABCA1, ABCG1 may be involved in the secretion of apoE [60].

The current findings are clues that, like ABCA1, AGCG1 affects how the cell deals with cholesterol and lipids. Support for this conjecture is provided by the observation that the gene for *ABCG1* is constitutively switched on during the differentiation of the monocyte into the macrophage [36,37]. A first hint of the physiological role of ABCG1 was given in a study by Klucken et al. that showed that inhibition of ABCG1 may result in a decrease in the efflux cholesterol and PC [36]. Currently, the details of how ABCG1 may be involved in this process are not known, although it is of note that ABCG1 is concentrated mainly in the outer cell membrane and in membrane structures around the cell nucleus of macrophages [36,37]. This is consistent with the hypothesis that ABCG1 may be directly involved in cholesterol and PC efflux. However, in fibroblasts—the principal cells of connective tissue—by contrast, ABCG1 seemed to be diffusely scattered throughout the cell, although these cells are also able to promote HDL-mediated efflux of cholesterol and PC [37].

In the section that follows, we will describe further findings consistent with the hypothesis that ABCG1 is involved in cellular lipid trafficking. In atherosclerotic arteries, ABCG1 is highly expressed in foamy macrophages and in bundles of nerves coursing in the outermost layer of the artery, the adventitia [37]. *ZNF202*, located within a susceptibility locus for hypoalphalipoproteinemia (i.e., low HDL plasma levels) on chromosome 11q23 is a transcriptional repressor of various genes involved in lipid metabolism [61]. Porsch-Özcürümez et al. showed that the full-length ZNF202m1 isoform binds to GnT repeats within the promoters of ABCA1 and ABCG1 [62]. ZNF202m1 expression in HepG2 cells dose-dependently repressed the promoter activities of ABCA1 and ABCG1. Argmann et al. showed that TGFβ1 enhances cholesterol efflux and upregulates expression of ABCA1 and ABCG1, which suggests an anti-atherogenic role for this cytokine [63]. Conditional disruption of the PPARγ gene in mice results in lowered expression of ABCA1, ABCG1, and apoE in macrophages and reduced efflux of cholesterol from these cells [64].

Owing to the high degree of homology between ABCG1 and ABCG4, and because these proteins are both regulated by cholesterol and implicated in cholesterol transport, Annilo et al. suggested that they may form a heterodimer and function coordinately [65]. However, in human monocyte-derived macrophages, expression of most ABCG proteins except ABCG2 is much lower than that of ABCG1 (unpublished results). It is therefore unlikely, at least in macrophages, that ABCG1 forms heterodimers with ABCG4, ABCG5, or ABCG8. Owing to these findings and to the different roles of ABCG2 and ABCG1, we suggest that ABCG1 may act as a homodimeric transporter, although at present there is no direct evidence to support this hypothesis.

At the time of writing, the function of ABCG1 is unknown, and no inherited disease caused by mutations in the *ABCG1* gene has been identified. Nevertheless, it seems reasonable to suggest that ABCG1 is somehow involved in cellular lipid trafficking. This hypothesis is supported by recent findings of Bryan Brewer's group at the National Institutes of Health in Bethesda. This group showed that in transgenic mice, overexpression of ABCG1 results in a significant reduction of HDL cholesterol in the plasma, while overexpression of ABCA1 results in an increase of plasma levels of HDL cholesterol [66–68]. This may suggest that ABCG1 plays a greater role in lipid uptake than in lipid efflux, although this is not the only possible explanation.

ABCG2/Abcg2—a multidrug resistance transporter

Allikmets et al. characterized an ABC transporter gene, which they designated ABCP (synonyms: BCRP or MRX), which is highly expressed in the placenta [69]. The ABCP gene produces two transcripts that differ at the 5′-end, but encode the same protein consisting of 655 amino acids that is now designated ABCG2 by the HUGO Gene Nomenclature Committee [16].

The *ABCG2* gene was mapped to human chromosome 4q22 between markers D4S2462 and D4S1557, and the mouse *Abcg2* gene was mapped to chromosome 6, 28 to 29 cM from the centromer

[69]. Bailey-Dell et al. found that the *ABCG2* gene contains 16 exons spanning more than 66 kb and is abundantly expressed in placenta, as well as in liver, intestine, and stem cells. Sequence analysis revealed that the promoter region has a CCAAT box but no TATA box, a potential CpG island, and putative binding sites for SP1, AP1, and AP2 [70]. The promoter does not have a serum response element, suggesting that ABCG2 is not a lipid transporter [5]. Further studies regarding this point showed that expression of ABCG2 is not upregulated during foam cell formation in human monocyte-derived macrophages or by treatment of macrophages with LXR/RXR agonists [48].

Doyle et al. found that ABCG2 is overexpressed in the multidrug-resistant human breast cancer cells of the subline MCF-7/AdrVp relative to parental MCF-7 cells [71]. MCF-7/AdrVp cells display an ATP-dependent reduction in the intracellular accumulation of anthracycline anticancer drugs in the absence of overexpression of known multidrug resistance transporters such as P-glycoprotein (MIM 171050). In addition, ABCG2 is overexpressed in human colon carcinoma cells selected for mitoxantrone resistance [22]. When overexpressed in cell lines, ABCG2 has the ability to confer high levels of resistance to mitoxantrone, anthracyclines, doxorubicin, daunorubicin, bisantrene, and the camptothecins topotecan and SN-38, reduces daunorubicin accumulation and retention, and causes an ATP-dependent enhancement of the efflux of rhodamine-123. Thus, ABCG2 is a xenobiotic transporter of the drugs mentioned above and appears to play a major role in the multidrug resistance phenotype of a specific form of human breast cancer [72,73]. Because overexpression of ABCG2 in Sf9 insect cells revealed a high-capacity, vanadate-sensitive ATPase activity that was stimulated by compounds known to be transported by this protein, Ozvegy and colleagues concluded that ABCG2 probably functions as a homodimer in these expression systems because it is unlikely that putative transport partners in the Sf9 insect cells would be overexpressed at sufficiently high levels [74].

Two types of phenotypes were observed in *Abcg2*-overexpressing fibroblasts derived from mice lacking the multidrug transporters *Abcb1*, *Abcb4*, and *Abcc1* and selected to identify new cell lines resistant to topotecan, mitoxantrone, or doxorubicin. Of these fibroblasts, the doxorubicin-selected lines had much higher relative resistance to doxorubicin, etoposide, and bisantrene than did the lines selected with mitoxantrone or topotecan [75]. In addition to these findings, Robey et al. reported that some cells overexpressing ABCG2 transport rhodamine-123, whereas others do not [76]. These two ABCG2-mediated phenotypes were explained by the identification of polymorphisms in the *ABCG2* gene [77]. The ABCG2 wild-type protein has an arginine at position 482, which is associated with the inability to transport rhodamine-123, low resistance to anthracyclines, and high resistance to mitoxantrone and topotecan. The phenotype including rhodamine transport and doxorubicin resistance was associated with replacement of arginine at amino acid 482 in ABCG2 by glycine or threonine.

Several reports have suggested that ABCG2 is associated with drug resistance in acute myeloid leukemia (AML) [78,79]. According to Abbott et al., it is unclear if the ABCG2 mRNA expression measured in the samples used in these two studies was physiologically significant or due to contamination by normal hematopoietic cells in low-expressing cases. Because ABCG2 expression is expressed in several types of normal hematopoietic cells, including erythroid progenitors, natural killer cells, and primitive stem cells [74,80], expression in AML blast samples might have been due to contamination by normal hematopoietic cells [81]. In contrast to the reports presented by Ross et al. and Sargent et al. [78,79], Abbott et al. found that high levels of ABCG2 mRNA expression in adult AML blast specimens are relatively uncommon and that ABCG2 expression may be limited to a small cell subpopulation in some cases [81].

A complete understanding of the mechanism and biological function of ABCG2 will be important for understanding its normal physiology as well as for the development of novel chemotherapeutic treatment strategies. For this reason, inhibitors of the ABCG2 protein are of great interest as chemosensitizers for clinical drug resistance, for improving the pharmacokinetics of substrate chemotherapeutic drugs, and in functional assays of ABCG2 activity for tailoring chemotherapy (see, e.g., [83,84]).

The normal function of ABCG2 is not known. ABCG2 may serve a protective function by preventing toxins from entering cells as well as potentially playing a role in regulating stem cell differen-

tiation [73,82] through export of an endogenous metabolite capable of inducing such differentiation. However, the ABCG2 protein is present in the plasma membrane of cultured cells [85,86], and the murine Abcg2 protein routes to the apical membrane in polarized cells [87]. Expression of the ABCG2 mRNA is not limited to the tissues mentioned above, but is also present in venous and capillary endothelial cells of almost all tissues. ABCG2 protein has been found in the apical membranes of placental syncytiotrophoblasts, hepatocytes, the epithelial lining of the small intestine and colon, and the ducts and lobules of the mammary gland [88]. Based on the localization of ABCG2, it has been suggested that this protein may have a defensive function, for example, by limiting the uptake of amphipathic drugs from the gut and preventing entry of such drugs into the fetus [14]. Investigations of an *Abcg2*(-/-) mouse currently underway in Alfred H. Schinkel's group in Amsterdam may help to clarify these issues [14].

ABCG3/Abcg3—a regulator of ABCG protein activity?

Computer searches in databases of ESTs led to the identification of several mouse and rat sequences that had high homology to mouse and rat Abcg2 mRNAs, but that appeared to encode another gene [20]. Mickley et al. used the EST sequences to clone the entire coding region of this unknown gene. The sequence revealed a single open reading frame encoding a protein of 650 amino acids, designated Abcg3. Analysis of the genomic structure revealed that the *Abcg3* gene has 16 exons, including one 5′-noncoding exon. Abcg3 is most closely related to Abcg2 with 55 % overall amino acid identity and 72 % overall similarity.

The *Abcg3* gene is localized to a single region on mouse chromosome 5, band E3-4, 59 cM from the centromere. Mickley et al. reported that a murine BAC clone containing *Abcg3* gave a single signal on human chromosome 8p12. However, experiments and database searches have failed to reveal a human *Abcg3*-related gene ([20] and personal unpublished results) so that the human genome has no Abcg3 ortholog, although other rodents appear to have orthologous proteins [20].

The highest levels of expression for mouse *Abcg2* were found in the kidney, lung, and small intestine [20]. For mouse *Abcg3*, the highest levels of expression were found in thymus and spleen. Except for the relatively higher levels of expression in the small intestine for both *Abcg2* and *Abcg3*, there was no concordance in expression patterns.

Interestingly, the Abcg3 protein contains several unusual residues in the Walker A and signature motifs (see Fig. 2) [20]. The Walker A consensus is GAGKST (gly-ala-gly-lys-ser-thr), and the Abcg3 sequence is DGSRSL (asp-gly-ser-arg-ser-lys). The signature consensus is LSGG (lys-ser-gly-gly), and the Abcg3 sequence in this region is RSKE (arg-ser-lys-glu). Many of these residues are highly conserved in all ABC genes, and mutations in these regions typically lead to non-functional proteins. For this reason, Mickley et al. suggested that Abcg3 may not be able to bind and/or hydrolyze ATP and speculated that Abcg3 may be an inhibitory subunit that can dimerize and inactivate other ABC half transporters [20].

ABCG4/Abcg4—an ABC transporter involved in cellular lipid trafficking in the brain?

Analysis of sequence database showed that five members of the ABCG subfamily are present within the human genome, one of which is annotated as ABCG4 [5,15]. The human ABCG4 mRNA has a length of about 3.5 kb, and the identified open reading frame encodes for a polypeptide of 646 amino acids in length [48]. Comparison of the predicted ABCG4 peptide sequence with the remaining known human ABCG half-size transporters revealed the highest degree of homology and identity to human ABCG1 (84 and 72 %, respectively), whereas amino acid identities with the other ABCG subfamily members were far lower (human ABCG2, 28 %; mouse Abcg3, 26 %; human ABCG5, 24 %; human ABCG8, 24 %). The strikingly high homology between ABCG4 and ABCG1 raises the possibility that both genes may have coevolved more recently than the other members of the ABCG subfamily. The close evolutionary relationship of ABCG1 and ABCG4 is confirmed by phylogenetic analyses [89,90] that

demonstrate with high confidence a relatively recent duplication of the ABCG1 and ABCG4 genes and their close evolutionary relationship to the *ATET* gene in *D. melanogaster* [26].

The human *ABCG4* gene consists of at least 14 exons (Fig. 3B) and spans a region of 12.7 kb on the chromosomal region 11q23.3 between markers D11S939 and D11S924. This chromosomal region is of special interest, since it contains the locus for many other genes involved in cholesterol and lipid metabolism such as the repressor of ABCA1 and ABCG1 expression, ZNF202, already mentioned above [62,90]. Several inherited diseases of lipoprotein metabolism, such as primary hypoalphalipoproteinemia (MIM 605201; [91]) and lathosterolosis [92], have been mapped to this region.

All exons of the *ABCG4* gene are relatively small, ranging between 51 and 250 nucleotides, except for the last exon, which spans 1.8 kb. Most introns are shorter than 220 nucleotides; only four introns are larger than 1000 base pairs. Thus, the genomic size is the smallest of all reported human ABCG proteins (ABCG1, 96 or 40 kb; ABCG2, 66 kb; ABCG5 and ABCG8 together, 80 kb [48]). Comparison of intron and exon size of the ABCG1 and ABCG4 gene reveals no conservation of intron size, whereas the size of 10 out of 14 exons (exons 2–7 and 10–13 of the ABCG4 gene, respectively) is conserved. None of these exons is conserved between ABCG1 or ABCG4 and the other members of the ABCG subfamily [48]. As is the case with ABCG1, analysis of the genomic sequence upstream of the 5′-end of the transcript reveals no transcription initiation elements such as a TATA-box or enhancer elements within 300 bp upstream of the start of transcription [48].

It is noteworthy that TMs 2 and 5 are identical in ABCG1 and ABCG4, whereas the other TMs harbor several amino acid exchanges [48]. TMs 2 and 5 are probably critical for the transporter's structure and function, whereas exchanges within the other TMs may affect substrate specificity. Compared to the ABCG1 protein sequence, amino acids are missing within the ABCG4 sequence at three positions (Fig. 2). These positions are located within regions of lower homology and may indicate domains that are critical for the transporters' function. A high degree of sequence variability between the ABCG1 and ABCG4 protein sequence is observed at the amino terminus, the ABC, the first transmembrane-spanning helix, and TMs 5 and 6.

The mouse Abcg4 mRNA also encodes a protein consisting of 646 amino acids, including one ABC and six TMDs [93]. Fluorescent in situ hybridization with mouse and rat chromosomes has revealed that the *Abcg4* gene is located on chromosomes 9A5.3 and 8q22 distal in mouse and rat, respectively. In these loci, conserved linkage homologies were hitherto identified with human chromosome 11q23, which involves the human *ABCG4* gene.

The highest level of expression of ABCG4 mRNA was detected in the brain [90] and in neural but not in choroid retina [94]. An abundant ABCG4 transcript was apparent in all brain tissues, with the exception of the spinal cord. Oldfield et al. reported that the only other mouse tissue expressing Abcg4 at high levels was spleen, in contrast to the human data [94]. Yoshikawa et al. also found high transcript levels of mouse Abcg4 mRNA in the eye and in bone marrow [93].

Using RT-PCR with gene-specific primers, Oldfield et al. identified four different ABCG4 splicing variants in thymus (transcripts II to V in Fig. 3B) and a unique transcript in lung (transcript I in Fig. 3B) [94]. All these splicing isoforms have protein-coding capability although they utilize three different putative ATG start codons, resulting in proteins that vary in their amino-terminal regions. Despite the high degree of homology in the coding region of ABCG1 and ABCG4, there is no similarity between the transcript variants of ABCG4 and the transcript variants of ABCG1 [28].

Screening of ABCG transporters, which, like ABCG1 and ABCA1, are induced by the heterodimeric transcription factors LXR and RXR, revealed that the ABCG4 is the second most strongly inducible ABCG transporter upon treatment of human monocyte-derived macrophages with 9-*cis* retinoic acid (RXR agonist) and 22*R*-hydroxycholesterol (LXR agonist) [48]. The most strongly inducible member is ABCG1, whereas ABCG2 was the only noninducible ABCG protein. Removal of cholesterol from macrophages mediated by the synthetic cholesterol acceptor cyclodextrin decreased ABCG4 message levels in a similar fashion to the decrease in message levels of ABCG1 and ABCA1. This behavior is common to oxysterol-responsive genes.

The abundant expression of ABCG4 in the brain suggests a potential role for this gene in cholesterol/lipid transport in this tissue, but there is no direct evidence available at this time. However, scientists at Active Pass Pharmaceuticals (San Francisco, CA, USA) have discovered that transfection of cells with either ABCG4 or ABCB9, transporters that are enriched in the brain, increases extracellular levels of the β-amyloid peptide [95]. This phenomenon is due either to a change in the rate at which the β-amyloid peptide detaches from the cell membrane, or to changes in expression of the amyloid precursor protein. Although not fully validated as targets for Alzheimer's disease, ABCG4 and ABCB9 may be of use in indirect target validation [95].

Although details of the physiology of ABCG4 are not known at the present time, we would not be surprised if ABCG4 turned out to have a physiological role similar to that of ABCG1, but that its action is limited to its major tissue of expression, the brain. This thesis is mainly based on two features: (i) the high similarity of ABCG4 and ABCG1 at the mRNA level, at the protein level, and at the level of genomic structure, suggesting that both genes evolved from a very recent gene duplication event in vertebrates, and (ii) the similar induction of expression of the ABCG1 and ABCG4 genes by LXR and RXR antagonists.

ABCG5/Abcg5 and ABCG8/Abcg8—(plant) sterol transporters

ABCG5 (synonyms: sterolin 1 or white 3) and *ABCG8* (synonyms: sterolin 1 or white 4) are both located on chromosome 2p21 between markers D2S177 and D2S119 and are tandemly arrayed in a head-to-head orientation separated by 374 bp. Both genes contain 13 exons and span about 28 kb [96]. The *ABCG5* and *ABCG8* genes encode deduced proteins of 651 and 673 amino acids, respectively, that share 29 % amino acid sequence identity and 50 % amino acid sequence similarity. Both proteins are expressed at a high level in the liver and at lower levels in the small intestine and colon [96–100].

Mutations in the *ABCG5* and *ABCG8* genes can cause a condition termed sitosterolemia [96–99,101]. Sitosterolemia is a rare autosomal recessive disorder characterized by intestinal accumulation of all sterols, including cholesterol and plant and shellfish sterols, and impaired ability to excrete sterols into the bile. In a healthy individual, 50 to 60 % of the dietary cholesterol is absorbed and retained, whereas less than 5% of the noncholesterol sterols are retained. In sitosterolemia, affected individuals absorb not only cholesterol but also all other sterols, including plant sterols, in particular sitosterol (24-ethyl cholesterol), and shellfish sterols, from the intestine [102–105]. Patients with sitosterolemia have high levels of plant sterols in the plasma and develop fatty nodules in the tendons and skin known as xanthomas, accelerated atherosclerosis, and premature coronary artery disease [103,107].

It is likely that ABCG5 and ABCG8 cooperate to limit intestinal absorption and to promote excretion of sterols into bile. This means that the ABCG5/ABCG8 heterodimer is an outward pump for plant sterols that in the intestine regulates the absorption and in the liver mediates the canalicular secretion of unwanted sterols. Therefore, mutated forms of these transporters predispose to sterol accumulation and atherosclerosis [96].

A detailed characterization of the molecular defects in a very large multiethnic cohort of patients with sitosterolemia was presented by Lu et al. and is consistent with this hypothesis [108]. In this study, all affected individuals in 37 families carried a mutation in either *ABCG5* or *ABCG8* but not in both. There were no clear clinical features that differentiated individuals with mutations in *ABCG5* from persons with mutations in *ABCG8*. Lu and colleagues therefore concluded that the two proteins either act as functional heterodimers or are tightly coupled along a regulatory pathway of dietary sterol absorption [106]. Thus, complete loss of activity of any of the two ABC half transporters causes a functional deficiency.

Repa et al. and others studied the control of expression of *Abca1*, *Abcg5*, and *Abcg8* by the nuclear receptors *Lxra* and *Lxrb* [107]. In normal mouse sections, Repa et al. found that expression of *Abcg5* and *Abcg8* was localized to hepatocytes of the liver and showed a uniform distribution across the

hepatic lobule. In jejunal sections, expression of *Abcg5* and *Abcg8* was detected exclusively in enterocytes lining the villi. Expression of *Abca1* was found predominantly in lamina propria and occasionally in enterocytes. In normal mice fed with cholesterol or Lxr agonists, hepatic and jejunal expression of *Abcg5*, *Abcg8*, and *Abca1* was increased. Cholesterol feeding resulted in upregulation of *Abcg5* and *Abcg8* expression in *Lxrb*-deficient mice, but not in *Lxra*-deficient or *Lxra/Lxrb*-double knockout mice. Repa et al. concluded that *Abca1*, *Abcg5*, and *Abcg8* are expressed in absorptive enterocytes and that all three ABC transporters have a role in regulating cholesterol flux in the intestine.

Owing to the fact that nearly all Asian sitosterolemia patients have mutations in *ABCG5*, while most Caucasian sitosterolemia patients have mutations in ABCG8, it has been suggested that the two proteins may function independently. Although the ABCG5/ABCG8 heterodimer is thought to be the major transporter of sitosterol, it has also been proposed that the homodimers may transport some of the other sterols [6].

Graf et al. have shown that the ABCG5 and ABCG8 proteins undergo amino-linked glycosylation and that the mature glycosylated forms of ABCG5 and ABCG8 immunoprecipitated together, consistent with the hypothesis that ABCG5 and ABCG8 form a heterodimer [108]. Using immunoelectron microscopy, Graf et al. demonstrated that ABCG5 and ABCG8 are located on the plasma membrane of hepatocytes. In polarized WIF-B cells, ABCG5 localized to the apical (canalicular) membrane when coexpressed with ABCG8, but not when expressed alone. The formation of an ABCG5/ABCG8 heterodimer and its localization on the plasma membrane is consistent with the hypothesis that ABCG5 and ABCG8 act as a heterodimer to transport sterols across plasma membranes.

To test the hypothesis that ABCG5 and ACBG8 limit sterol absorption and promote biliary sterol excretion, Yu et al. generated transgenic mice overexpressing ABCG5 and ABCG8 [109]. In the transgenic mice, the expression pattern of the transgenes was similar to that of the endogenous genes. The results of Yu et al. clearly showed that increased expression of ABCG5 and ABCG8 selectively drives biliary neutral sterol secretion and reduces intestinal cholesterol absorption, leading to a selective increase in neutral sterol excretion and a compensatory increase in cholesterol synthesis.

Very recently, Helen H. Hobbs and her group disrupted Abcg5 and Abcg8 in mice [110]. Using this mouse model of sitosterolemia, these authors found that ABCG5 and ABCG8 are required for efficient secretion of cholesterol into bile and that disruption of these genes increases dramatically the responsiveness of plasma and hepatic cholesterol levels to changes in dietary cholesterol content.

HOMOLOGS OF HUMAN ABCG PROTEINS

Recognizing the close evolutionary relationship of ABCG1 and ABCG4, which is also confirmed by phylogenetic analyses, helped us to identify a mechanism of regulation of expression of *ABCG4* (see above). Phylogenetic analyses not only indicated a relatively recent duplication of the ABCG1 and *ABCG4* genes and their close evolutionary relationship to the *Drosophila ATET* gene, but also allowed to estimate the evolutionary relationship of all ABC proteins. Such phylogenetic analyses may also help us to identify regulatory mechanisms, motifs conserved throughout evolution, substrates, and binding partners.

At present, we are developing a rigorous procedure called RiPE (rigorous phylogeny environment) that automatically performs an evolutionary analysis of a protein subfamily [89]. We are using this procedure to calculate phylogenetic trees of the whole ABC protein family and of ABC protein subfamilies. These trees have better statistical support and include more ABC proteins than previously published trees (Spitzer et al., manuscript in preparation).

Using our RiPE procedure, we searched the proteomes of human (*H. sapiens*), mouse (*Mus musculus*), pufferfish (*F. rubripes*), fly (*D. melanogaster*), worm (*C. elegans*), mustard weed (*A. thaliana*), rice (*Oryza sativa*), yeast (*S. cerevisiae*), and a bacterium (*Escherichia coli*). We retrieved ABC proteins and calculated phylogenetic trees. In addition, we performed phylogenetic analyses using our RiPE procedure based on protein data included in the NCBI nonredundant peptide sequence database to calcu-

late trees including almost all ABC proteins known so far [111]. Details on the trees and how they were calculated will be available elsewhere (Spitzer et al., manuscript in preparation).

A simplified phylogenetic tree of all ABCG-like proteins found in the proteomes of the vertebrates *H. sapiens*, *M. musculus*, and *F. rubripes*; the invertebrates *D. melanogaster*, *Anopheles gambiae*, and *C. elegans*; the fungus *S. cerevisiae*; the common progenitor of plants animals and fungi, *Dictyostelium discoideum*; the microsporidian *Encephalitozoon cuniculi*; and the plants *A. thaliana* and *O. sativa* is shown in Fig. 4. This tree is the essence of the analyses mentioned above and is based on the alignment of homologous regions of ABCG proteins derived from the organisms listed above.

In the genomes of most of the organisms studied, there are more members of the ABCG subfamily than in the human genome (Table 3). While the ABCG proteins of the vertebrates, the fly, and the worm are all half transporters with an ABC-TMD domain architecture, most of the ABCG proteins of *S. cerevisiae* and some ABCG proteins of *A. thaliana* and *O. sativa* are full transporters with an ABC-TMD-ABC-TMD domain arrangement.

Table 3 Number of ABCG half and full transporters in proteomes assembled from completely sequenced genomes of selected eukaryotes.[a]

Organism	Number of ABCG half transporters	Number of ABCG full transporters	Total number of ABCG proteins
Anopheles gambiae	12	–	12
Arabidopsis thaliana	28	12	40
Caenorhabditis elegans	8	–	8
Dictyostelium discoideum	6	15	21
Drosophila melanogaster	14	–	14
Encephalitozoon cuniculi	5	–	5
Fugu rubripes	8	–	8
Guillardia theta	–	–	–
Homo sapiens	5	–	5
Mus musculus	6	–	6
Oryza sativa	20	10	30
Saccharomyces cerevisiae	1	9	10
Schizosaccharomyces pombe	–	2	2

[a]Numbers of genes are based on previous and own analyses of available proteome assemblies of these organisms. For this reason, the numbers of ABCG proteins in these organisms may vary if a new release of a proteome assembly is published.

The analysis of the phylogenetic trees suggests an early divergence of ABCG proteins. In the following, we will describe the clearly separated clusters in the detailed phylogenetic tree (not shown) that are the basis for the simplified tree in Fig. 4:

i. *White cluster*: This cluster is comprised of the white, brown, and scarlet proteins of *D. melanogaster* and *A. gambiae* and orthologs in *C. elegans*.
ii. *G1/G4/ATET cluster*: The proteins that form this cluster are found in the insects *D. melanogaster*, *A. gambiae* (including one of the closest ABCG1/ABCG4 orthologs called ATET) as well as in the vertebrates *H. sapiens*, *M. musculus*, *R. norvegicus*, and *F. rubripes* (including ABCG1/Abcg1 and ABCG4/Abcg4).
iii. *G5/G8 cluster*: This cluster is formed by the two halves of a yeast ABCG-like full transporter, worm, and fly/gnat proteins, and vertebrate proteins (ABCG5 and ABCG8 orthologs in human, mouse, and pufferfish).

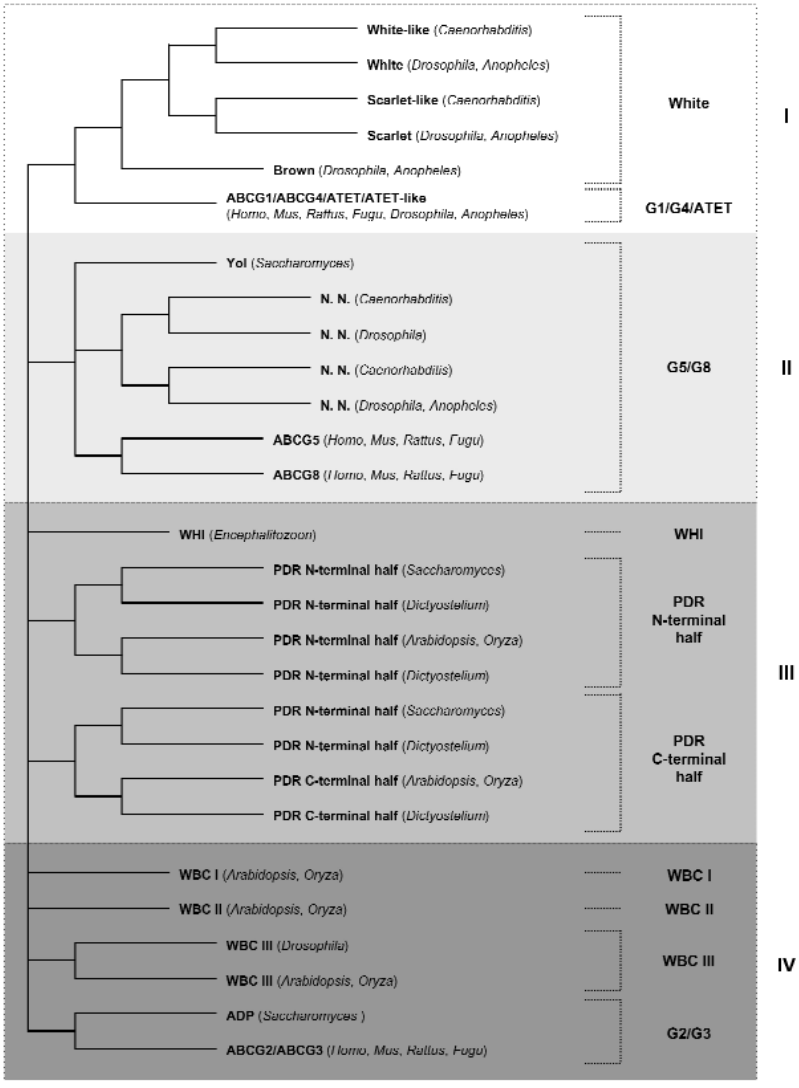


Fig. 4 Simplified phylogenetic tree of eukaryotic ABCG-like proteins. See text for details.

iv. *PDR N-terminal half cluster*: This cluster is comprised solely of the amino-terminal halves of the yeast and plant ABCG-like full transporters.
v. *PDR C-terminal half cluster*: Similar to the cluster described before, this cluster contains the carboxy-terminal halves of the yeast and plant ABCG-like full transporters.
vi. *WBC I cluster*: The WBC I cluster is exclusively comprised of plant proteins.
vii. *WBC II cluster*: This small cluster contains only proteins of mustard weed and rice.
viii. *WBC III cluster*: This cluster is also nearly exclusively formed by plant sequences. One single fly protein is included.
ix. *G2/G3 cluster*: This cluster is formed nearly exclusively by vertebrate proteins including ABCG2/Abcg2 and Abcg3 and their orthologs in *F. rubripes*. In addition, a single yeast half transporter belongs to this cluster.

Looking at the tree derived from the nonredundant database, which includes proteins from more organisms than the ones used for the tree described above (data not shown), we found that the clusters are mainly consistent with the clusters described here. Interestingly, the proteome data of the fungus *E. cuniculi* contains ABCG-structured half transporters clustering together with the full transporter proteins of yeast and plants (WHI cluster). Further analysis of these sequences may help us to determine how ABCG half and full transporters evolved.

Based on these clusters and on preliminary analyses not shown here, we suggest the simplified tree of ABCG proteins in Fig. 4. We speculate that the clusters described above may form four clusters of a higher order. In the figure, these four clusters are stained from light to dark gray. This is clearly the case for the white and G1/G4 clusters, which form a distinct cluster denoted cluster I. A further distinct cluster is comprised of the G5/G8 cluster and is denoted cluster II. Cluster III may be formed by the two PDR clusters and by some of the ABCG half transporters found in *E. cuniculi*. Note that the statistical support for this association is weak in our current analysis. A fourth cluster, denoted cluster IV, may be formed by the WBC and G2/G3 clusters. However, this association is also weak in our current analysis. At this point, the reader should keep in mind that most of the protein sequences used for the analyses described above are predicted sequences often lacking experimental verification. Inclusion of proteome data from newly sequenced organisms may help to retrieve trees with better statistical support and higher resolution at the roots of the clusters, thus allowing more precise conclusions on how ABCG proteins evolved.

CONCLUDING REMARKS

It now seems likely that not only the *ABCA1* gene, which was recently linked to the inherited lipid disorder known as Tangier disease, but also the other ABC transporter genes, in particular the members of the ABCG subfamily, are involved in cellular lipid uptake, lipid efflux, or lipid trafficking within the cell. Overall, however, very little is known on how the products of the ABCG gene family are regulated, or on how they function. Elucidation of these questions will require detailed studies of gene regulation, substrate specificity, and structural analysis. It will also be important to determine the binding partners of the half transporters. Defects in ABCG1 and ABCG4 may turn out to cause inherited diseases, as is the case for defects in many of the other 48 members of the human ABC transporter family. The putative role of the ABCG transporters in lipid transport may also indicate that these proteins play a role in the development of the atherosclerotic plaque and that ABCG1, ABCG4, ABCG5, and ABCG8 may be potential targets for anti-atherosclerotic therapies. Phylogenetic analysis to identify orthologous or paralogous* genes may provide clues to help in unraveling the phenotypes associated with mutations in these genes.

ACKNOWLEDGMENTS

The authors are participants in a project entitled "Macrophage Function and Stability of the Atherosclerotic Plaque" (MAFAPS, QLG2-CT-1999-01007) supported by the European Union as part of the Fifth Framework Program [112]. We thank Georg Fuellen and Michael Spitzer for providing phylogenetic tree analyses of ABC proteins. The authors also wish to express their appreciation to Prof. Erwin Arno Galinski for his advice and criticism in the preparation of this manuscript. Prof. Galinski was also a supervisor of Dr. Lorkowski's doctoral thesis.

*Orthologs and paralogs are two types of homologous sequences. Orthology describes genes in different species that derived from a common ancestor. Ortholog genes may or may not have the same function. Paralogy describes homologous genes within a single species that diverged by gene duplication.

REFERENCES

1. M. M. Gottesman and S. V. Ambudkar. *J. Bioenerg. Biomembr*. **33**, 453–458 (2001).
2. C. F. Higgins. *Annu. Rev. Cell Biol*. **8**, 67–113 (1992).
3. J. E. Walker, M. Saraste, M. J. Runswick, N. J. Gay. *EMBO J*. **1**, 945–951 (1982).
4. M. Dean and R. Allikmets. *Curr. Opin. Genet. Dev*. **5**, 779–785 (1995).
5. M. Dean, A. Rzhetsky, R. Allikmets. *Genome Res*. **11**, 1156–1166 (2001).
6. M. Dean. In *The Human ATP-Binding Cassette (ABC) Transporter Superfamily*, <http://www.ncbi.nlm.nih.gov/books/bv.fcgi?call=bv.View..ShowTOC&rid=mono_001.TOC&depth=2>.
7. I. B. Holland and M. A. Blight. *J. Mol. Biol*. **293**, 381–399 (1999).
8. E. Dassa and P. Bouige. *Res. Microbiol*. **152**, 211–229 (2001).
9. P. Bouige, D. Laurent, L. Piloyan, E. Dassa. *Curr. Protein Pept. Sci*. **3**, 541–559 (2002).
10. M. Dean and R. Allikmets. *J. Bioenerg. Biomembr*. **33**, 475–479 (2001).
11. I. Klein, B. Sarkadi, A. Varadi. *Biochim. Biophys. Acta* **1461**, 237–262 (1999).
12. T. Efferth. *Curr. Mol. Med*. **1**, 45–65 (2001).
13. <http://www.humanabc.org/>
14. P. Borst and R. O. Elferink. *Annu. Rev. Biochem*. **71**, 537–592 (2002).
15. M. Dean, Y. Hamon, G. Chimini. *J. Lipid Res*. **42**, 1007–1017 (2001).
16. <http://www.gene.ucl.ac.uk/nomenclature/genefamily/abc.html>
17. B. E. Bauer, H. Wolfger, K. Kuchler. *Biochim. Biophys. Acta* **1461**, 217–236 (1999).
18. A. Decottignies and G. Goffeau. *Nat. Genet*. **15**, 137–145 (1997).
19. R. Sanchez-Fernandez, T. G. Davies, J. O. Coleman, P. A. Rea. *J. Biol. Chem*. **276**, 30231–30244 (2001).
20. L. Mickley, P. Jain, K. Miyake, L. M. Schriml, K. Rao, T. Fojo, S. Bates, M. Dean. *Mamm. Genome* **12**, 86–88 (2001).
21. C. Anjard and W. F. Loomis. *Eukaryot. Cell* **1**, 643–652 (2002).
22. H. M. Chen, C. Rossier, M. D. Lalioti, A. Lynn, A. Chakravarti, G. Perrin, S. E. Antonarakis. *Am. J. Hum. Genet*. **59**, 66–75 (1996).
23. J. M. Croop, G. E. Tiller, J. A. Fletcher, M. L. Lux, E. Raab, D. Goldenson, D. Son, S. Arciniegas, R. L. Wu. *Gene* **185**, 77–85 (1997).
24. S. Savary, F. Denizot, M. Luciani, M. Mattei, G. Chimini. *Mamm. Genome* **7**, 673–676 (1996).
25. T. H. Morgan. *Science* **32**, 120–122 (1910).
26. H. Kuwana, K. Shimizu-Nishikawa, H. Iwahana, D. Yamamoto. *Biochim. Biophys. Acta* **1309**, 47–52 (1996).
27. T. Langmann, M. Porsch-Özcürümez, U. Unkelbach, J. Klucken, G. Schmitz. *Biochim. Biophys. Acta* **1494**, 175–180 (2000).
28. S. Lorkowski, S. Rust, T. Engel, E. Jung, K. Tegelkamp, E. A. Galinski, G. Assmann, P. Cullen. *Biochem. Biophys. Res. Commun.* **280**, 121–131 (2001).
29. M. A. Kennedy, A. Venkateswaran, P. T. Tarr, I. Xenarios, J. Kudoh, N. Shimizu, P. A. Edwards. *J. Biol. Chem*. **276**, 39438–39447 (2001).
30. T. Engel, S. Lorkowski, A. Lueken, S. Rust, B. Schlüter, G. Berger, P. Cullen, G. Assmann. *Biochem. Biophys. Res. Commun.* **288**, 483–488 (2001).
31. A. Berry, H. S. Scott, J. Kudoh, I. Talior, M. Korostishevsky, M. Wattenhofer, M. Guipponi, C. Barras, C. Rossier, K. Shibuya, J. Wang, K. Kawasaki, S. Asakawa, S. Minoshima, N. Shimizu, S. Antonarakis, B. Bonne-Tamir. *Genomics* **68**, 22–29 (2000).
32. M. Nakamura, S. Ueno, A. Sano, H. Tanabe. *Mol. Psychiatry* **4**, 155–162 (1999).
33. D. Rujescu, I. Giegling, N. Dahmen, A. Szegedi, I. Anghelescu, A. Gietl, M. Schafer, F. Muller-Siecheneder, B. Bondy, H. J. Moller. *Neuropsychobiology* **42**, 22–25 (2000).

34. C. Johansson, C. Smedh, T. Partonen, P. Pekkarinen, T. Paunio, J. Ekholm, L. Peltonen, D. Lichtermann, J. Palmgren, R. Adolfsson, M. Schalling. *Neurobiol. Dis*. **8**, 351–357 (2001).
35. G. Kirov, C. A. Lowry, M. Stephens, S. Oldfield, M. C. O'Donovan, S. L. Lightman, M. J. Owen. *Mol. Psychiatry* **6**, 671–677 (2001).
36. J. Klucken, C. Buchler, E. Orso, W. E. Kaminski, M. Porsch-Özcürümez, G. Liebisch, M. Kapinsky, W. Diederich, W. Drobnik, M. Dean, R. Allikmets, G. Schmitz. *Proc. Natl. Acad. Sci. USA* **97**, 817–822 (2000).
37. S. Lorkowski, M. Kratz, C. Wenner, R. Schmidt, B. Weitkamp, M. Fobker, J. Reinhardt, J. Rauterberg, E. A. Galinski, P. Cullen. *Biochem. Biophys. Res. Commun*. **283**, 821–830 (2001).
38. A. Venkateswaran, J. J. Repa, J. M. Lobaccaro, A. Bronson, D. J. Mangelsdorf, P. A. Edwards. *J. Biol. Chem*. **275**, 14700–14707 (2000).
39. P. Liang and A. B. Pardee. *Science* **257**, 967–971 (1992).
40. S. Lorkowski, P. Ellinghaus, E. A. Galinski, G. Assmann, P. Cullen. *Clin. Chim. Acta* **299**, 199–204 (2000).
41. J. S. Fisher, J. Meade, J. Walden, Z. Guo, P. Liang. In *Analysing Gene Expression. A Handbook of Methods: Possibilities and Pitfalls*, S. Lorkowski and P. Cullen (Eds.), pp. 494–501, Wiley-VCH, Weinheim (2002).
42. J. J. Repa, S. D. Turley, J. A. Lobaccaro, J. Medina, L. Li, K. Lustig, B. Shan, R. A. Heyman, J. M. Dietschy, D. J. Mangelsdorf. *Science* **289**, 1524–1529 (2000).
43. P. Costet, Y. Luo, N. Wang, A. R. Tall. *J. Biol. Chem*. **275**, 28240–28245 (2000).
44. K. Schwartz, R. M. Lawn, D. P. Wade. *Biochem. Biophys. Res. Commun*. **274**, 794–802 (2000).
45. P. A. Edwards, H. R. Kast, A. M. Anisfeld. *J. Lipid Res*. **43**, 2–12 (2002).
46. S. Lorkowski and P. Cullen. *Curr. Opin. Lipidol*. **13**, 689–692 (2002).
47. A. Venkateswaran, B. A. Laffitte, S. B. Joseph, P. A. Mak, D. C. Wilpitz, P. A. Edwards, P. Tontonoz. *Proc. Natl. Acad. Sci. USA* **97**, 12097–12102 (2000).
48. T. Engel, S. Lorkowski, A. Lueken, S. Rust, B. Schluter, G. Berger, P. Cullen, G. Assmann. *Biochem. Biophys. Res. Commun.* **288**, 483–488 (2001).
49. J. J. Repa, K. E. Berge, C. Pomajzl, J. A. Richardson, H. Hobbs, D. J. Mangelsdorf. *J. Biol. Chem.* **277**, 18793–18800 (2002).
50. B. A. Laffitte, J. J. Repa, S. B. Joseph, D. C. Wilpitz, H. R. Kast, D. J. Mangelsdorf, P. Tontonoz. *Proc. Natl. Acad. Sci. USA* **98**, 507–512 (2001).
51. P. A. Mak, B. A. Laffitte, C. Desrumaux, S. B. Joseph, L. K. Curtiss, D. J. Mangelsdorf, P. Tontonoz, P. A. Edwards. *J. Biol. Chem*. **277**, 31900–31908 (2002).
52. Y. Luo and A. R. Tall. *J. Clin. Invest.* **105**, 513–520 (2000).
53. J. Y. Chiang, R. Kimmel, D. Stroup. *Gene* **262**, 257–265 (2001).
54. S. Gupta, W. M. Pandak, P. B. Hylemon. *Biochem. Biophys. Res. Commun*. **293**, 338–343 (2002).
55. J. J. Repa, G. Liang, J. Ou, Y. Bashmakov, J. M. Lobaccaro, I. Shimomura, B. Shan, M. S. Brown, J. L. Goldstein, D. J. Mangelsdorf. *Genes Dev*. **14**, 2819–2830 (2000).
56. Y. Zhang, J. J. Repa, K. Gauthier, D. J. Mangelsdorf. *J. Biol. Chem*. **276**, 43018–43024 (2001).
57. S. B. Joseph, B. A. Laffitte, P. H. Patel, M. A. Watson, K. E. Matsukuma, R. Walczak, J. L. Collins, T. F. Osborne, P. Tontonoz. *J. Biol. Chem*. **277**, 11019–11025 (2002).
58. G. Cao, T. P. Beyer, X. P. Yang, R. J. Schmidt, Y. Zhang, W. R. Bensch, R. F. Kauffman, H. Gao, T. P. Ryan, Y. Liang, P. I. Eacho, X. C. Jiang. *J. Biol. Chem*. **277**, 39561–39565 (2002).
59. T. Langmann, J. Klucken, M. Reil, G. Liebisch, M. F. Luciani, G. Chimini, W. E. Kaminski, G. Schmitz. *Biochem. Biophys. Res. Commun*. **275**, 29–33 (1999).
60. A. Von Eckardstein, C. Langer, T. Engel, I. Schaukal, A. Cignarella, J. Reinhardt, S. Lorkowski, Z. Li, X. Zhou, P. Cullen, G. Assmann. *FASEB J*. **15**, 1555–1561 (2001).
61. S. Wagner, M. A. Hess, P. Ormonde-Hanson, J. Malandro, H. Hu, M. Chen, R. Kehrer, M. Frodsham, C. Schumacher, M. Beluch, C. Honer, M. Skolnick, D. Ballinger, B. R. Bowen. *J. Biol. Chem*. **275**, 15685–15690 (2000).

62. M. Porsch-Özcürümez, T. Langmann, S. Heimerl, H. Borsukova, W. E. Kaminski, W. Drobnik, C. Honer, C. Schumacher, G. Schmitz. *J. Biol. Chem*. **276**, 12427–12433 (2001).
63. C. A. Argmann, C. H. Van Den Diepstraten, C. G. Sawyez, J. Y. Edwards, R. A. Hegele, B. M. Wolfe, M. W. Huff. *Arterioscler. Thromb. Vasc. Biol*. **21**, 2011–2018 (2001).
64. T. E. Akiyama, S. Sakai, G. Lambert, C. J. Nicol, K. Matsusue, S. Pimprale, Y. H. Lee, M. Ricote, C. K. Glass, H. B. Brewer Jr., F. J. Gonzalez. *Mol. Cell. Biol*. **22**, 2607–2619 (2002).
65. T. Annilo, J. Tammur, A. Hutchinson, A. Rzhetsky, M. Dean, R. Allikmets. *Cytogenet. Cell. Genet*. **94**, 196–201 (2001).
66. T. Ito, S. L. Sabol, M. Amar, C. Knapper, C. Duarte, R. D. Shamburek, S. Meyn, S. Santamarina-Fojo, H. B. Brewer. *Circulation* **102**, 311 (2000).
67. R. R. Singaraja, V. Bocher, E. R. James, S. M. Clee, L. H. Zhang, B. R. Leavitt, B. Tan, A. Brooks-Wilson, A. Kwok, N. Bissada, Y. Z. Yang, G. Liu, S. R. Tafuri, C. Fievet, C. L. Wellington, B. Staels, M. R. Hayden. *J. Biol. Chem*. **276**, 33969–33979 (2001).
68. B. L. Vaisman, G. Lambert, M. Amar, C. Joyce, T. Ito, R. D. Shamburek, W. J. Cain, J. Fruchart-Najib, E. D. Neufeld, A. T. Remaley, H. B. Brewer Jr., S. Santamarina-Fojo. *J. Clin. Invest*. **108**, 303–309 (2001).
69. R. Allikmets, L. M. Schriml, A. Hutchinson, V. Romano-Spica, M. Dean. *Cancer Res.* **58**, 5337–5339 (1998).
70. K. J. Bailey-Dell, B. Hassel, L. A. Doyle, D. D. Ross. *Biochim. Biophys. Acta* **1520**, 234–241 (2001).
71. L. A. Doyle, W. Yang, L. V. Abruzzo, T. Krogmann, Y. Gao, A. K. Rishi, D. D. Ross. *Proc. Natl. Acad. Sci. USA* **95**, 15665–15670 (1998).
72. K. Miyake, L. Mickley, T. Litman, Z. Zhan, R. Robey, B. Cristensen, M. Brangi, L. Greenberger, M. Dean, T. Fojo, S. E. Bates. *Cancer Res*. **59**, 8–13 (1999).
73. K. F. K. Ejendal and C. A. Hrycyna. *Curr. Protein. Pept. Sci*. **3**, 503–511 (2002).
74. C. Ozvegy, T. Litman, G. Szakacs, Z. Nagy, S. Bates, A. Varadi, B. Sarkadi. *Biochem. Biophys. Res. Commun*. **285**, 111–117 (2001).
75. J. D. Allen, R. F. Brinkhuis, J. Wijnholds, A. H. Schinkel. *Cancer Res*. **59**, 4237–4241 (1999).
76. R. W. Robey, W. Y. Medina-Perez, K. Nishiyama, T. Lahusen, K. Miyake, T. Litman, A. M. Senderowicz, D. D. Ross, S. E. Bates. *Clin. Cancer Res*. **7**, 145–152 (2001).
77. M. Brangi, T. Litman, M. Ciotti, K. Nishiyama, G. Kohlhagen, C. Takimoto, R. Robey, Y. Pommier, T. Fojo, S. E. Bates. *Cancer Res*. **59**, 5938–5946 (1999).
78. D. D. Ross, J. E. Karp, T. T. Chen, L. A. Doyle. *Blood* **96**, 365–368 (2000).
79. J. M. Sargent, C. J. Williamson, M. Maliepaard, A. W. Elgie, R. J. Scheper, C. G. Taylor. *Br. J. Haematol*. **115**, 257–262 (2001).
80. C. W. Scharenberg, M. A. Harkey, B. Torok-Storb. *Blood* **99**, 507–512 (2002).
81. B. L. Abbott, A. M. Colapietro, Y. Barnes, F. Marini, M. Andreeff, B. P. Sorrentino. *Blood* **100**, 4594–4601 (2002).
82. S. Zhou, J. D. Schuetz, K. D. Bunting, A. M. Colapietro, J. Sampath, J. J. Morris, I. Lagutina, G. C. Grosveld, M. Osawa, H. Nakauchi, B. P. Sorrentino. *Nat. Med*. **7**, 1028–1034 (2001).
83. J. D. Allen, A. van Loevezijn, J. M. Lakhai, M. van der Valk, O. van Tellingen, G. Reid, J. H. Schellens, G. J. Koomen, A. H. Schinkel. *Mol. Cancer Ther*. **1**, 417–425 (2002).
84. R. L. Shepard, J. Cao, J. J. Starling, A. H. Dantzig. *Int. J. Cancer* **1**, 121–125 (2003).
85. E. Rocchi, A. Khodjakov, E. L. Volk, C. H. Yang, T. Litman, S. E. Bates, E. Schneider. *Biochem. Biophys. Res. Commun*. **271**, 42–46 (2000).
86. G. L. Scheffer, M. Kool, M. Heijn, M. de Haas, A. C. Pijnenborg, J. Wijnholds, A. van Helvoort, M. C. de Jong, J. H. Hooijberg, C. A. Mol, M. van der Linden, J. M. de Vree, P. van der Valk, R. P. Elferink, P. Borst, R. J. Scheper. *Cancer Res*. **60**, 5269–5277 (2000).
87. J. W. Jonker, J. W. Smit, R. F. Brinkhuis, M. Maliepaard, J. H. Beijnen, J. H. Schellens, A. H. Schinkel. *J. Natl. Cancer Inst*. **92**, 1651–1656 (2000).

88. M. Maliepaard, G. L. Scheffer, I. F. Faneyte, M. A. van Gastelen, A. C. Pijnenborg, A. H. Schinkel, M. J. van De Vijver, R. J. Scheper, J. H. Schellens. *Cancer Res*. **61**, 3458–3464 (2001).
89. G. Fuellen, M. Spitzer, P. Cullen, S. Lorkowski. *In Silico Biol.* (2002). Submitted for publication.
90. C. Monaco, M. Helmer Citterich, E. Caprini, I. Vorechovsky, G. Russo, C. M. Croce, G. Barbanti-Brodano, M. Negrini. *Genomics* **52**, 358–362 (1998).
91. E. N. Kort, D. G. Ballinger, W. Ding, S. C. Hunt, B. R. Bowen, V. Abkevich, K. Bulka, B. Campbell, C. Capener, A. Gutin, K. Harshman, M. McDermott, T. Thorne, H. Wang, B. Wardell, J. Wong, P. N. Hopkins, M. Skolnick, M. Samuels. *Am. J. Hum. Genet*. **66**, 1845–1856 (2000).
92. N. Brunetti-Pierri, G. Corso, M. Rossi, P. Ferrari, F. Balli, F. Rivasi, I. Annunziata, A. Ballabio, A. D. Russo, G. Andria, G. Parenti. *Am. J. Hum. Genet*. **71**, 952–958 (2002).
93. M. Yoshikawa, H. Yabuuchi, A. Kuroiwa, Y. Ikegami, Y. Sai, I. Tamai, A. Tsuji, Y. Matsuda, H. Yoshida, T. Ishikawa. *Gene* **293**, 67–75 (2002).
94. S. Oldfield, C. Lowry, J. Ruddick, S. Lightman. *Biochim. Biophys. Acta* **1591**, 175–179 (2002).
95. <http://www.activepass.com/AlzheimerProfile.pdf>
96. K. E. Berge, H. Tian, G. A. Graf, L. Yu, N. V. Grishin, J. Schultz, P. Kwiterovich, B. Shan, R. Barnes, H. H. Hobbs. *Science* **290**, 1771–1775 (2000).
97. M. H. Lee, K. Lu, S. Hazard, H. Yu, S. Shulenin, H. Hidaka, H. Kojima, R. Allikmets, N. Sakuma, R. Pegoraro, A. K. Srivastava, G. Salen, M. Dean, S. B. Patel. *Nat. Genet*. **27**, 79–83 (2001).
98. K. Lu, M. H. Lee, S. Hazard, A. Brooks-Wilson, H. Hidaka, H. Kojima, L. Ose, A. F. Stalenhoef, T. Mietinnen, I. Bjorkhem, E. Bruckert, A. Pandya, H. B. Brewer Jr., G. Salen, M. Dean, A. Srivastava, S. B. Patel. *Am. J. Hum. Genet.* **69**, 278–290 (2001).
99. J. A. Hubacek, K. E. Berge, J. C. Cohen, H. H. Hobbs. *Hum. Mutat*. **18**, 359–360 (2001).
100. S. Shulenin, L. M. Schriml, A. T. Remaley, S. Fojo, B. Brewer, R. Allikmets, M. Dean. *Cytogenet. Cell. Genet*. **92**, 204–208 (2001).
101. G. Salen, S. Patel, A. K. Batta. *Cardiovasc. Drug. Rev*. **20**, 255–270 (2002).
102. J. M. Ordovas and E. S. Tai. *Nutr. Rev*. **60**, 30–33 (2002).
103. M. H. Lee, K. Lu, S. B. Patel. *Curr. Opin. Lipidol*. **12**, 141–149 (2001).
104. S. B. Patel, A. Honda, G. Salen. *J. Lipid Res*. **39**, 1055–1061 (1998).
105. G. Salen, S. Shefer, L. Nguyen, G. C. Ness, G. S. Tint, A. K. Batta. *Subcell. Biochem*. **28**, 453–476 (1997).
106. K. Lu, M. H. Lee, S. B. Patel. *Trends Endocrinol. Metab*. **12**, 314–320 (2001).
107. J. J. Repa, K. E. Berge, C. Pomajzl, J. A. Richardson, H. Hobbs, D. J. Mangelsdorf. *J. Biol. Chem*. **277**, 18793–18800 (2002).
108. G. A. Graf, W. P. Li, R. D. Gerard, I. Gelissen, A. White, J. C. Cohen, H. H. Hobbs. *J. Clin. Invest*. **110**, 659–669 (2002).
109. L. Yu, J. Li-Hawkins, R. E. Hammer, K. E. Berge, J. D. Horton, J. C. Cohen, H. H. Hobbs. *J. Clin. Invest*. **110**, 671–680 (2002).
110. L. Yu, R. E. Hammer, J. Li-Hawkins, K. Von Bergmann, D. Lutjohann, J. C. Cohen, H. H. Hobbs. *Proc. Natl. Acad. Sci. USA* **99**, 16237–16292 (2002).
111. <http://www.ncbi.nlm.nih.gov>
112. S. Bellosta, F. Bernini, G. Chinetti, A. Cignarella, P. Cullen, A. von Eckardstein, A. Exley, J. Freeth, M. Goddard, M. Hofker, E. Kanters, P. Kovanen, S. Lorkowski, M. Pentikäinen, J. Printen, J. Rauterberg, A. Ritchie, B. Staels, B. Weitkamp, M. de Winther. *Nutr. Metab. Cardiovasc. Dis.* **12**, 3–11 (2002).
113. <http://www.ebi.ac.uk/clustalw/>
114. J. D. Thompson, D. G. Higgins, T. J. Gibson. *Nucleic Acids Res.* **22**, 4673–4680 (1994).
115. <http://www-pgm1.ipbs.fr:8080/cgi-bin/nph-ESPript_exe.cgi>
116. <http://smart.embl-heidelberg.de/>
117. J. Schultz, R. R. Copley, T. Doerks, C. P. Ponting, P. Bork. *Nucleic Acids Res.* **28**, 231–234 (2000).

118. I. Letunic, L. Goodstadt, N. J. Dickens, T. Doerks, J. Schultz, R. Mott, F. Ciccarelli, R. R. Copley, C. P. Ponting, P. Bork. *Nucleic Acids Res.* **30**, 242–244 (2002).