

# The Human ATP-Binding Cassette (ABC) Transporter Superfamily

Michael Dean,<sup>1,6</sup> Andrey Rzhetsky,<sup>2,5</sup> and Rando Allikmets<sup>3,4</sup>

<sup>1</sup>Human Genetics Section, Laboratory of Genomic Diversity, National Cancer Institute-Frederick, Frederick, Maryland 21702, Departments of <sup>2</sup>Medical Informatics, <sup>3</sup>Ophthalmology, <sup>4</sup>Pathology, and <sup>5</sup>Columbia Genome Center, Columbia University, New York, New York, 10032 USA

The ATP-binding cassette (ABC) transporter superfamily contains membrane proteins that translocate a variety of substrates across extra- and intra-cellular membranes. Genetic variation in these genes is the cause of or contributor to a wide variety of human disorders with Mendelian and complex inheritance, including cystic fibrosis, neurological disease, retinal degeneration, cholesterol and bile transport defects, anemia, and drug response. Conservation of the ATP-binding domains of these genes has allowed the identification of new members of the superfamily based on nucleotide and protein sequence homology. Phylogenetic analysis is used to divide all 48 known ABC transporters into seven distinct subfamilies of proteins. For each gene, the precise map location on human chromosomes, expression data, and localization within the superfamily has been determined. These data allow predictions to be made as to potential functions or disease phenotypes associated with each protein. In this paper, we review the current state of knowledge on all human ABC genes in inherited disease and drug resistance. In addition, the availability of the complete *Drosophila* genome sequence allows the comparison of the known human ABC genes with those in the fly genome. The combined data enable an evolutionary analysis of the superfamily. Complete characterization of all ABC from the human genome and from model organisms will lead to important insights into the physiology and the molecular basis of many human disorders.

## ABC Protein and Gene Organization

The ABC genes represent the largest family of transmembrane proteins. These proteins bind ATP and use the energy to drive the transport of various molecules across all cell membranes (Higgins 1992; Childs and Ling 1994; Dean and Allikmets 1995). Proteins are classified as ABC transporters based on the sequence and organization of their ATP-binding domains, also known as nucleotide-binding folds (NBFs). The NBFs contain characteristic motifs, Walker A and B, separated by ~90–120 amino acids, found in all ATP-binding proteins (Fig. 1). ABC genes also contain an additional element, the signature C motif, located just upstream of the Walker B site (Hyde et al. 1990). The functional protein typically contains two NBFs and two transmembrane (TM) domains (Fig. 1). The TM domains contain 6–11 membrane-spanning  $\alpha$ -helices and provide the specificity for the substrate. The NBFs are located in the cytoplasm and transfer the energy to transport the substrate across the membrane. ABC pumps are mostly unidirectional. In bacteria, they are predominantly involved in the import of essential compounds that cannot be obtained by diffusion (e.g., sugars, vitamins, metal ions, etc.) into the cell. In eukaryotes, most ABC genes move compounds from the cytoplasm to the outside of the cell or into an extracellular compartment (endoplasmic reticulum, mitochondria, peroxisome). Most of the known functions of eukaryotic ABC transporters involve the shuttling of hydrophobic compounds either within the cell as part of a metabolic process or outside the cell for transport to other organs, or secretion from the body.

### Corresponding author.

E-MAIL dean@ncifcrf.gov; FAX (301) 846-1909.

Article and publication are at <http://www.genome.org/cgi/doi/10.1101/gr.184901>.

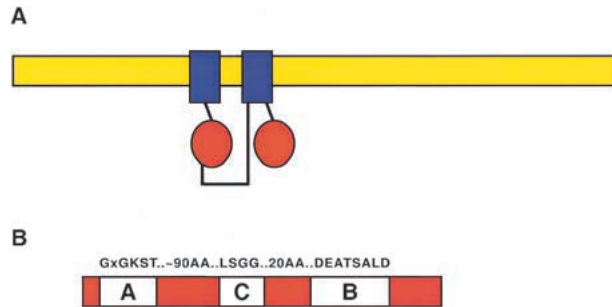
The eukaryotic ABC genes are organized either as full transporters containing two TMs and two NBFs, or as half transporters (Hyde et al. 1990). The latter must form either homodimers or heterodimers to constitute a functional transporter. ABC genes are dispersed widely in eukaryotic genomes and are highly conserved between species, indicating that most of these genes have existed since the beginning of eukaryotic evolution. The genes can be divided into subfamilies based on similarity in gene structure (half vs. full transporters), order of the domains, and on sequence homology in the NBF and TM domains. There are seven mammalian ABC gene subfamilies, five of which are found in the *Saccharomyces cerevisiae* genome.

## Overview of Human ABC Gene Subfamilies

A list of all known human ABC genes is displayed in Table 1. This list includes an analysis of recently released genome sequences (Lander et al. 2001; Venter et al. 2001). Several sequences remain in the genome with homology to ABC genes, many of which represent pseudogenes. A comprehensive description of all ABC genes goes beyond the scope of this review. Therefore, only a concise summary of each subfamily is provided.

### ABCA (ABCI)

This subfamily comprises 12 full transporters (Table 1) that are further divided into two subgroups based on phylogenetic analysis and intron structure (Brocardo et al. 1999). The first group includes seven genes dispersed on six different chromosomes (*ABCA1*, *ABCA2*, *ABCA3*, *ABCA4*, *ABCA7*, *ABCA12*, *ABCA13*), whereas the second group contains five genes (*ABCA5*, *ABCA6*, *ABCA8*, *ABCA9*, and *ABCA10*) arranged in a cluster on chromosome 17q24. The ABCA subfamily contains some of the largest ABC genes, several of which are >2100



**Figure 1** Diagram of a typical ABC transporter protein. (A) A diagram of the structure of a representative ABC protein is shown with a lipid bilayer in yellow, the transmembrane domains in blue, and the nucleotide binding fold in red. Although the most common arrangement is a full-transporter with motifs arranged N-TM-NBF-TM-NBF-C, as shown, NBF-TM-NBF-TM, TM-NBF, and NBF-TM arrangements are also found. (B) The NBF of an ABC gene contains the Walker A and B motifs found in all ATP-binding proteins. In addition, a signature or C motif is also present. The most common amino acids found in these motifs are shown above the diagram; subfamilies often contain characteristic residues in these and other regions.

amino acids. Two members of this subfamily, the ABCA1 and ABCA4 (ABCR) proteins, have been studied extensively. The ABCA1 protein is involved in disorders of cholesterol transport and high-density lipoproteins (HDL) biosynthesis (see below). The ABCA4 protein transports vitamin A derivatives in the outer segments of photoreceptor cells and therefore performs a crucial step in the visual cycle.

#### ABCB (MDR/TAP)

The ABCB subfamily is unique in that it contains both full transporters and half transporters. Four full transporters and seven half transporters are currently identified as members of this subfamily. *ABCB1* (*MDR/PGY1*) is the first human ABC transporter cloned and characterized through its ability to confer a multidrug resistance phenotype to cancer cells. The functional sites of *ABCB1* include the blood-brain barrier and the liver. The *ABCB4* and *ABCB11* proteins are both located in the liver and are involved in the secretion of bile acids. The *ABCB2* and *ABCB3* (*TAP*) genes are half transporters that form a heterodimer to transport peptides into the endoplasmic reticulum, which are presented as antigens by the Class I HLA molecules. The closest homolog of the TAPs, the *ABCB9* half transporter, has been localized to lysosomes. The remaining four half-transporters, *ABCB6*, *ABCB7*, *ABCB8* and *ABCB10* localize to the mitochondria, where they function in iron metabolism and transport of Fe/S protein precursors.

#### ABCC (CFTR/MRP)

The ABCC subfamily contains 12 full transporters with a diverse functional spectrum that includes ion transport, cell surface receptor, and toxin secretion activities. The *CFTR* protein is a chloride ion channel that has a role in all exocrine secretions, and mutations in *CFTR* cause cystic fibrosis (Quinton 1999). *ABCC8* and *ABCC9* proteins bind sulfonylurea and regulate potassium channels involved in modulating insulin secretion. The rest of the subfamily is composed of nine MRP-related genes. Of these *ABCC1*, *ABCC2*, and *ABCC3* transport drug conjugates to glutathione and other organic anions. The *ABCC4*, *ABCC5*, *ABCC11* and *ABCC12* proteins are smaller than the other MRP1-like genes and lack an amino-terminal domain (Borst et al. 2000) that is not essential for transport

function (Bakos et al. 2000). The *ABCC4* and *ABCC5* proteins confer resistance to nucleosides including PMEA and purine analogs.

#### ABCD (ALD)

The ABCD subfamily contains four genes in the human genome and two each in the *Drosophila* and yeast genomes. The yeast *PXA1* and *PXA2* products dimerize to form a functional transporter involved in very long chain fatty acid oxidation in the peroxisome (Shani and Valle 1998). All of the genes encode half transporters that are located in the peroxisome, where they function as homo- and/or heterodimers in the regulation of very long chain fatty acid transport.

#### ABCE (OABP) and ABCF (GCN20)

The ABCE and ABCF subfamilies contain genes that have ATP-binding domains that are clearly derived from ABC transporters but have no TM domain and are not known to be involved in any membrane transport functions. The ABCE subfamily is comprised solely of the oligo-adenylate binding protein, a molecule that recognizes oligo-adenylate that is produced in response to infection by certain viruses. This gene is found in multicellular eukaryotes, but not in yeast, suggesting it is part of innate immunity. Each ABCF gene contains a pair of NBFs. The best-characterized member, the *S. cerevisiae* *GCN20* gene mediates the activation of the eIF-2  $\alpha$ -kinase (Marton et al. 1997) and a human homolog, *ABCF1*, is associated with the ribosome and appears to have a similar role (Tyzack et al. 2000).

#### ABCG (White)

The human ABCG subfamily is comprised of six 'reverse' half transporters that have an NBF at the amino terminus and a TM domain at the carboxyl terminus. The most intensively studied ABCG gene is the white locus of *Drosophila*. The white protein, along with brown and scarlet, transport precursors of eye pigments (guanine and tryptophan) in the eye cells of the fly (Chen et al. 1996). The mammalian *ABCG1* gene is involved in cholesterol transport regulation (Klucken et al. 2000). Other ABCG genes include *ABCG2*, a drug resistance gene; *ABCG5* and *ABCG8*, transporters of sterols in the intestine and liver; *ABCG3*, to date found exclusively in rodents; and the *ABCG4* gene that is expressed predominantly in the liver. The functions of the last two genes are unknown.

#### ABC Genes and Human Genetic Disease

Many ABC genes were originally discovered during the positional cloning of human genetic disease genes. To date, 14 ABC genes have been linked to disorders displaying Mendelian inheritance (Klein et al. 1999) (Table 2). As expected from the diverse functional roles of ABC genes, the genetic deficiencies that they cause also vary widely. Because ABC genes typically encode structural proteins, all of the disorders are recessive, and are attributable to a severe reduction or lack of function of the protein. Heterozygous variants in ABC gene mutations, however, are being implicated in the susceptibility to specific complex disorders.

#### Cystic Fibrosis and CFTR

Cystic fibrosis is the most common fatal childhood disease in Caucasian populations, reaching frequencies ranging from 1:900 to 1:2500. This corresponds to a carrier frequency of 1:15–1:25. The disease is much less common in African and Asian populations, where carrier frequencies of 1:100 to 1:200 have been estimated. The disease frequency correlates with

**Table 1.** List of Human ABC Genes, Chromosomal Location, and Function

Gene	Alias	Location	Subfamily	Expression	Function
ABCA1	ABC1	9q31.1	ABC1	Ubiquitous	Cholesterol efflux onto HDL
ABCA2	ABC2	9q34	ABC1	Brain	Drug resistance
ABCA3	ABC3, ABCC	16p13.3	ABC1	Lung	
ABCA4	ABCR	1p22.1-p21	ABC1	Photoreceptors	N-retinylidene-PE efflux
ABCA5		17q24	ABC1	Muscle, heart, testes	
ABCA6		17q24	ABC1	Liver	
ABCA7		19p13.3	ABC1	Spleen, thymus	
ABCA8		17q24	ABC1	Ovary	
ABCA9		17q24	ABC1	Heart	
ABCA10		17q24	ABC1	Muscle, heart	
ABCA12		2q34	ABC1	Stomach	
ABCA13		7p11-q11	ABC1	Low in all tissues	
ABCB1	PGY1, MDR	7p21	MDR	Adrenal, kidney, brain	Multidrug resistance
ABCB2	TAP1	6p21	MDR	All cells	Peptide transport
ABCB3	TAP2	6p21	MDR	All cells	Peptide transport
ABCB4	PGY3	7q21.1	MDR	Liver	PC transport
ABCB5		7p14	MDR	Ubiquitous	
ABCB6	MTABC3	2q36	MDR	Mitochondria	Iron transport
ABCB7	ABC7	Xq12-q13	MDR	Mitochondria	Fe/S cluster transport
ABCB8	MABC1	7q36	MDR	Mitochondria	
ABCB9		12q24	MDR	Heart, brain	
ABCB10	MTABC2	1q42	MDR	Mitochondria	
ABCB11	SPGP	2q24	MDR	Liver	Bile salt transport
ABCC1	MRP1	16p13.1	CF/MRP	Lung, testes, PBMC	Drug resistance
ABCC2	MRP2	10q24	CF/MRP	Liver	Organic anion efflux
ABCC3	MRP3	17q21.3	CF/MRP	Lung, intestine, liver	Drug resistance
ABCC4	MRP4	13q32	CF/MRP	Prostate	Nucleoside transport
ABCC5	MRP5	3q27	CF/MRP	Ubiquitous	Nucleoside transport
ABCC6	MRP6	16p13.1	CF/MRP	Kidney, liver	
CFTR	ABCC7	7q31.2	CF/MRP	Exocrine tissue	Chloride ion channel
ABCC8	SUR	11p15.1	CF/MRP	Pancreas	Sulfonylurea receptor
ABCC9	SUR2	12p12.1	CF/MRP	Heart, muscle	
ABCC10	MRP7	6p21	CF/MRP	Low in all tissues	
ABCC11	MRP8	16q11-q12	CF/MRP	Low in all tissues	
ABCC12	MRP9	16q11-q12	CF/MRP	Low in all tissues	
ABCD1	ALD	Xq28	ALD	Peroxisomes	VLCFA transport regulation
ABCD2	ALDL1, ALDR	12q11-q12	ALD	Peroxisomes	
ABCD3	PXMP1, PMP70	1p22-p21	ALD	Peroxisomes	
ABCD4	PMP69, P70R	14q24.3	ALD	Peroxisomes	
ABCE1	OABP, RNS41	4q31	OABP	Ovary, testes, spleen	Oligoadenylate binding protein
ABCF1	ABCS0	6p21.33	GCN20	Ubiquitous	
ABCF2		7q36	GCN20	Ubiquitous	
ABCF3		3q25	GCN20	Ubiquitous	
ABCG1	ABC8, White	21q22.3	White	Ubiquitous	Cholesterol transport?
ABCG2	ABCP, MXR, BCRP	4q22	White	Placenta, intestine	Toxin efflux, drug resistance
ABCG4	White2	11q23	White	Liver	
ABCG5	White3	2p21	White	Liver, intestine	Sterol transport
ABCG8		2p21	White	Liver, intestine	Sterol transport

the frequency of the major allele of the CF gene, a deletion of three base pairs ( $\Delta F508$ ). At least two other populations, however, have high frequency CF alleles. The W1282X allele is found on 51% of the alleles in the Ashkenazi Jewish population and the 1677delTA allele has been found at a high frequency in Georgians and is also present at elevated level in Turkish and Bulgarian populations. This has led several groups to hypothesize that these alleles arose through selection of an advantageous phenotype in the heterozygotes. It is through CFTR that some bacterial toxins such as cholera and *Escherichia coli* cause increased fluid flow in the intestine and result in diarrhea. Therefore, several researchers have proposed that the CF mutations have been selected for in response to these diseases. This hypothesis is supported by studies showing that CF homozygotes fail to secrete chloride ions in response to a variety of stimulants, and a study in mice in which heterozygous null animals showed reduced intestinal

fluid secretion in response to cholera toxin (Gabriel et al. 1993). CFTR is also the receptor for *Salmonella typhimurium* and implication in the innate immunity to *Pseudomonas aeruginosa* (Pier et al. 1998).

Patients with two severe CFTR alleles like  $\Delta F508$  typically display severe diseases with inadequate secretion of pancreatic enzymes leading to nutritional deficiencies, bacterial infections of the lung, and obstruction of the vas deferens, leading to male infertility. Patients with at least one partially functional allele display enough residual pancreatic function to avoid the major nutritional and intestinal deficiencies (Dean et al. 1990) and subjects with very mild alleles display only congenital absence of the vas deferens with none of the other symptoms of CF. Recently, heterozygotes of CF mutations have been found to have an increased frequency of pancreatitis (Cohn et al. 1998) and bronchiectasis (Pignatti et al. 1995). Therefore, there is a spectrum of severity in the phe-

**Table 2. Diseases and Phenotypes Caused by ABC Genes**

Gene	Mendelian disorder	Complex disease
<i>ABCA1</i>	Tangier disease, FHDLD	
<i>ABCA4</i>	Stargardt/FFM, RP, CRD, CD	AMD
<i>ABCB1</i>		Digoxin uptake
<i>ABCB2</i>	Immune deficiency	
<i>ABCB3</i>	Immune deficiency	
<i>ABCB4</i>	PFIC-3	ICP
<i>ABCB7</i>	XLSA/A	
<i>ABCB11</i>	PFIC-2	
<i>ABCC2</i>	Dubin-Johnson Syndrome	
<i>ABCC6</i>	Pseudoxanthoma elasticum	
<i>ABCC7</i>	Cystic Fibrosis, CBAVD	Pancreatitis, bronchiectasis
<i>ABCC8</i>	FPHHI	
<i>ABCD1</i>	ALD	
<i>ABCG5</i>	Sitosterolemia	
<i>ABCG8</i>	Sitosterolemia	

FHDLD, familial hypoapoproteinemia; FFM, fundus flavimaculatus; RP, retinitis pigmentosum 19; CRD, cone-rod dystrophy; CD, xxxx; AMD, age-related macular degeneration; PFIC, progressive familial intrahepatic cholestasis; ICP, intrahepatic cholestasis of pregnancy; XLSA/A, X-linked sideroblastosis and anemia; FPHHI, familial persistent hyperinsulinemic hypoglycemia of infancy; ALD, adrenoleukodystrophy; IDDM, insulin-dependent diabetes mellitus.

notypes caused by this gene that is inversely related to the level of CFTR activity. Clearly, other modifying genes and the environment also affect disease severity, particularly the pulmonary phenotypes.

**Adrenoleukodystrophy**

Adrenoleukodystrophy (ALD) is an X-linked recessive disorder characterized by neurodegenerative phenotypes with onset typically in late childhood (Mosser et al. 1993). Adrenal deficiency commonly occurs and the presentation of ALD is highly variable. AMN, childhood ALD and adult onset forms are recognized, but there is no apparent correlation to *ABCD1* alleles. ALD patients have an accumulation of unbranched, saturated fatty acids with a chain length of 24–30 carbons, in the cholesterol esters of the brain and in adrenal cortex. The ALD protein is located in the peroxisome, where it is believed to be involved in the transport of very long chain fatty acids.

**Sulfonylurea Receptor**

The *ABCC8* gene is a high-affinity receptor for the drug sulfonylurea. Sulfonylureas are a class of drugs widely used to increase insulin secretion in patients with non-insulin-dependent diabetes. These drugs bind to the *ABCC8* protein and inhibit an associated potassium channel. Familial persistent hyperinsulinemic hypoglycemia of infancy is an autosomal recessive disorder in which subjects display unregulated insulin secretion. The disease was mapped to 11p15-p14 by linkage analysis, and mutations in the *ABCC8* gene are found in PHHI families (Thomas et al. 1995). The *ABCC8* gene has also been implicated in insulin response

in Mexican-American subjects (Goksel et al. 1998) and in type II diabetes in French Canadians (Reis et al. 2000) but not in a Scandinavian cohort (Altshuler et al. 2000).

**Bile Salt Transport Disorders**

Several ABC transporters are specifically expressed in the liver, have a role in the secretion of components of the bile, and are responsible for several forms of progressive familial intrahepatic cholestasis (PFIC). PFICs are a heterogeneous group of autosomal recessive liver disorders, characterized by early onset of cholestasis that leads to liver cirrhosis and failure (Alonso et al. 1994). The *ABCB4* (PGY3) gene transports phosphatidylcholine across the canalicular membrane of hepatocytes (van Helvoort et al. 1996). Mutations in this gene cause PFIC3 (Deleuze et al. 1996; de Vree et al. 1998) and are associated with intrahepatic cholestasis of pregnancy (Dixon et al. 2000). The rat *Abcc2* gene was found to have a frame-shift mutation in the strain defective in canalicular multispecific organic anion transport, the TR<sup>-</sup> rat (Paulusma et al. 1996). The TR<sup>-</sup> rat is an animal model of Dubin-Johnson syndrome and mutations in *ABCC2* have been identified in Dubin-Johnson syndrome patients (Wada et al. 1998). The *ABCC2* protein is expressed on the canalicular side of the hepatocyte and mediates organic anion transport. The *ABCB11* gene was originally identified based on homology to *ABCB1* (Childs et al. 1995). *ABCB11* is highly expressed on the liver canalicular membrane and has been shown to be the major bile salt export pump. Mutations in *ABCB11* are found in patients with PFIC2 (Strautnieks et al. 1998).

**Retinal Degeneration and *ABCA4***

The *ABCA4* gene is expressed exclusively in photoreceptors where it transports retinol (vitamin A) derivatives from the photoreceptor outer segment disks into the cytoplasm (Allikmets et al. 1997). The chromophore of a visual pigment rhodopsin, retinal, or conjugates with phospholipids are the likely substrates for *ABCA4*, as they stimulate the ATP hydrolysis of the protein (Sun et al. 1999). Mice lacking *Abca4* show increased all-trans-retinaldehyde (all-trans-RAL) following light exposure, elevated phosphatidylethanolamine (PE) in outer segments, accumulation of the protonated Schiff base complex of all-trans-RAL and PE (N-retinylidene-PE), and striking deposition of a major lipofuscin fluorophore (A2-E) in retinal pigment epithelium (RPE) (Weng et al. 1999). These data suggest that ABCR is an outwardly directed flippase for N-retinylidene-PE.

Mutations in the *ABCA4* gene have been associated with multiple eye disorders (Allikmets 2000). A complete loss of

**Table 3. ABC Transporters Involved in Drug Resistance**

Gene	Substrates	Inhibitors
<i>ABCB1</i>	Colchicine, doxorubicin, VP16, adriamycin, vinblastine, digoxin, saquinivir, paclitaxel	Verapamil, PSC833, GG918, V-104, Pluronic L61
<i>ABCC1</i>	Doxorubicin, daunorubicin, vincristine, VP16, colchicines, etoposide, rhodamine	Cyclosporin A, V-104
<i>ABCC2</i>	Vinblastine, sulfinpyrazone,	
<i>ABCC3</i>	Methotrexate, etoposide	
<i>ABCC4</i>	Nucleoside monophosphates	
<i>ABCC5</i>	Nucleoside monophosphates	
<i>ABCG2</i>	Mitoxantrone, topotecan, doxorubicin, daunorubicin, CPT-11, rhodamine	Fumitremorgin C, GF120918

ABCA4 function leads to retinitis pigmentosa, whereas patients with at least one missense allele have Startgardt disease (STGD). STGD is characterized by juvenile to early adult onset macular dystrophy with loss of central vision. *ABCA4* mutation carriers are also increased in frequency in age-related macular degeneration (AMD) patients. AMD patients display a variety of phenotypic features, including the loss of central vision after the age of 60. The causes of this complex trait are poorly understood, but a combination of genetic and environmental factors have a role. The abnormal accumulation of retinoids, caused by *ABCA4* deficiency has been postulated to be one mechanism by which this process could be initiated. Defects in *ABCA4* lead to an accumulation of retinal derivatives in the retinal pigment epithelium behind the retina.

#### Mitochondrial Iron Homeostasis

Several half transporters of the MDR/TAP subfamily have been localized to the inner membrane of the mitochondria. The yeast ortholog of *ABCB7*, *Atm1*, has been implicated in mitochondrial iron homeostasis, as a transporter in the biogenesis of cytosolic Fe/S proteins (Kispal et al. 1997). Two distinct missense mutations in *ABCB7* are associated with the X-linked sideroblastic anemia and ataxia (XLSA/A) phenotype (Allikmets et al. 1999). Three more half transporters from this subfamily, *ABCB6*, *ABCB8* and *ABCB10* have also been localized to mitochondria (Table 1).

#### Sterol Transport Deficiencies

Tangier disease is characterized by deficient efflux of lipids from peripheral cells, such as macrophages, and a very low level of HDL. The disease is caused by alterations in the *ABCA1* gene, implicating this protein in the pathway of removal of cholesterol and phospholipids onto HDL (Young and Fielding 1999). Patients with hypolipidemia have also been described that are heterozygous for *ABCA1* mutations, suggesting that *ABCA1* variations may have a role in regulating the level of HDLs in the blood (Marcil et al. 1999).

Subsequently, the sterol-dependent regulation of *ABCA1* expression was shown (Langmann et al. 1999). Current models for *ABCA1* function place it at the plasma membrane where it mediates the transfer of phospholipid and cholesterol onto lipid-poor apolipoproteins to form nascent HDL particles. The *ABCA1*-mediated efflux of cholesterol is regulated by nuclear hormone receptors, such as oxysterol receptors (LXR) and the bile acid receptor (FXR), as heterodimers with retinoid X receptors (RXR) (Repa et al. 2000).

Recently, two half-transporter genes, *ABCG5* and *ABCG8* were characterized (Berge et al. 2000; Lee et al. 2001), located head-to-head on the human chromosome 2p15-p16, and regulated by the same promoter. These genes are both mutated in families with sitosterolemia, a disorder characterized by defective transport of plant and fish sterols and cholesterol. Most likely, the two half-transporters form a functional heterodimer. The *ABCG1* gene is also regulated by cholesterol (Klucken et al. 2000) and *ABCG4* is highly expressed in the liver, suggesting that these two genes may also be involved in cholesterol transport (Table 1).

#### Multidrug Resistance

Cells exposed to toxic compounds can develop resistance by a number of mechanisms including decreased uptake, increased detoxification, alteration of target proteins, or increased excretion. Several of these pathways can lead to multidrug resistance (MDR), in which the cell is resistant to several drugs

in addition to the initial compound. This is a particular limitation to cancer chemotherapy and the MDR cell often displays other properties, such as genome instability and loss of checkpoint control, which complicate further therapy. ABC genes have an important role in MDR and at least six genes are associated with drug transport (Table 3).

#### ABCB1

The best characterized ABC drug pump is the *ABCB1* gene, formerly known as MDR1 or P-glycoprotein. *ABCB1* was the first human ABC transporter cloned and characterized through its ability to confer a multidrug resistance phenotype to cancer cells that had developed resistance to chemotherapy drugs (Juliano and Ling 1976). *ABCB1* has been demonstrated to be a promiscuous transporter of hydrophobic substrates, hydrophobic drugs including drugs including colchicine, VP16, adriamycin and vinblastine as well as lipids, steroids, xenobiotics, and peptides (for review, see Ambudkar 1998). The gene is thought to have an important role in removing toxic metabolites from cells, but is also expressed in cells at the blood-brain barrier and presumably has a role in transporting compounds into the brain that cannot be delivered by diffusion. *ABCB1* also affects the pharmacology of the drugs that are substrates and a common polymorphism in the gene affects digoxin uptake (Hoffmeyer et al. 2000).

#### ABCC1

The *ABCC1* gene was identified in the small-cell lung carcinoma cell line NCI-H69, a multidrug resistant cell that did not overexpress *ABCB1* (Cole et al. 1992). The *ABCC1* pump confers resistance to doxorubicin, daunorubicin, vincristine, colchicines and several other compounds, a very similar profile to that of *ABCB1*. Unlike *ABCB1*, however, *ABCC1* transports drugs that are conjugated to glutathione by the glutathione reductase pathway (Borst et al. 2000). *ABCC1* can also transport leukotrienes, such as leukotriene C4 (LTC4). LTC4 is an important signaling molecule for the migration of dendritic cells. Migration of dendritic cells from the epidermis to lymphatic vessels is defective in *Abcc1*<sup>-/-</sup> mice (Robbiani et al. 2000).

#### ABCG2

Analysis of cell lines resistant to mitoxantrone that do not overexpress *ABCB1* or *ABCC1*, led several laboratories to identify the *ABCG2* (*ABCP*, *MXR1*, *BCRP*) gene as a drug transporter (Allikmets et al. 1998; Doyle et al. 1998; Miyake et al. 1999). *ABCG2* confers resistance to anthracycline anticancer drugs and is amplified or involved in chromosomal translocations in cell lines selected with topotecan, mitoxantrone, or doxorubicin treatment. It is suspected that *ABCG2* functions as a homodimer because transfection of the gene into cells confers resistance to chemotherapeutic drugs. *ABCG2* can also transport several dyes such as rhodamine and Hoechst 33462 and the gene is highly expressed in a subpopulation of hematopoietic stem cells (side population) that stain poorly for these dyes. The normal function of the gene in these cells, however, is unknown. *ABCG2* is highly expressed in the trophoblast cells of the placenta. This suggests that the pump is responsible either for transporting compounds into the fetal blood supply, or removing toxic metabolites. The gene is also expressed in the intestine and *ABCG2* inhibitors that could be useful in making substrates orally available.

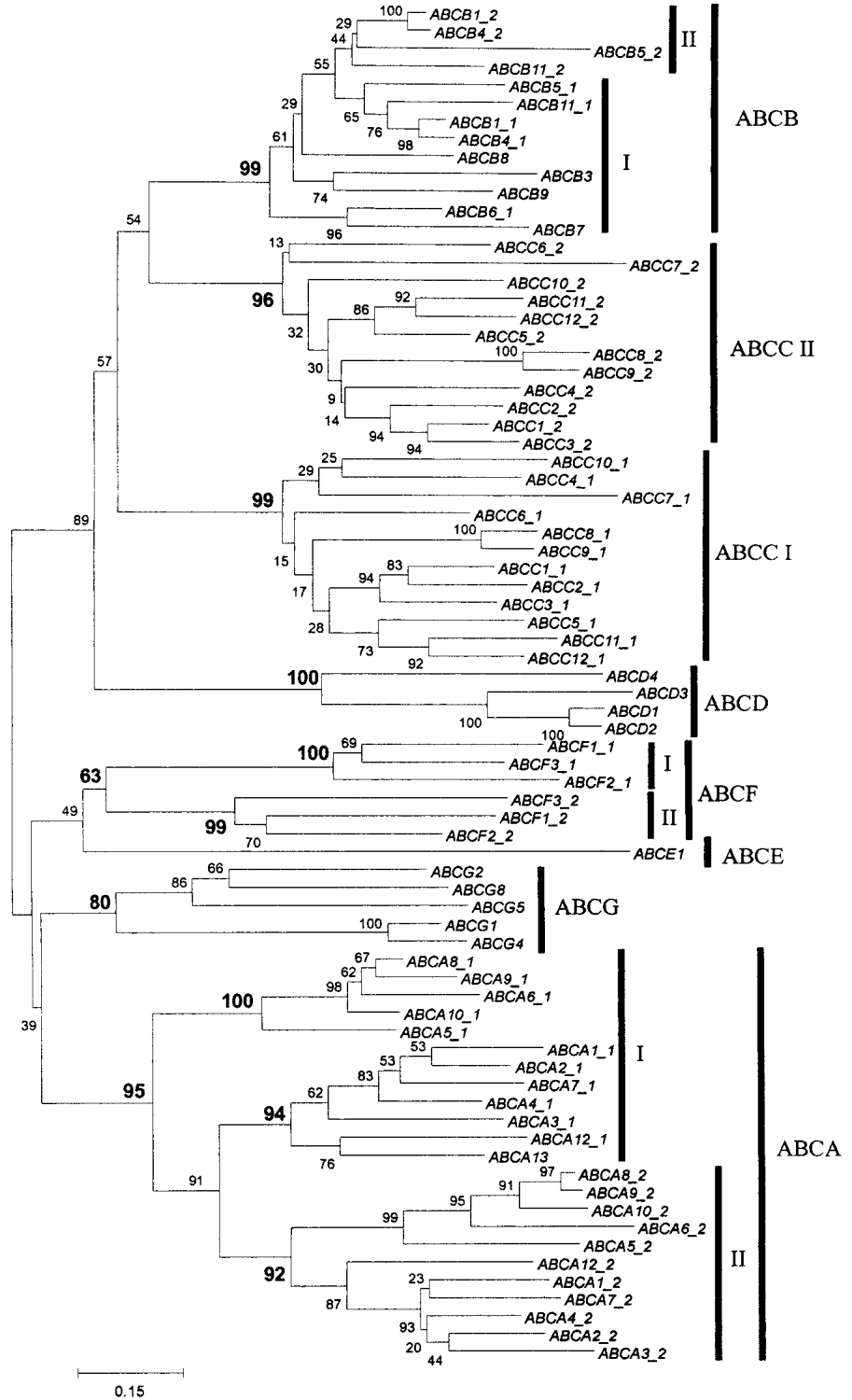
Phylogenetic Analysis of Human ABC Genes

The identification of the nearly complete set of human ABC genes allows a comprehensive phylogenetic analysis of the superfamily. To understand the organization of these genes within the subfamilies described previously, an alignment of the ATP-binding domains was generated and used for phylogenetic analysis. Figure 2 displays a neighbor-joining tree resulting from this analysis. The proposed nomenclature of ABC transporters is in excellent agreement with the phylogenetic trees obtained. In particular, all major ABC transporter families are represented in the human tree by stable clusters with high bootstrap values.

This analysis provides compelling evidence for frequent domain duplication of ATP-binding domains in ABC transporters. Virtually invariably, both ATP-binding domains within a gene are more

closely related to each other than to ATP-binding domains from ABC transporter genes of other subfamilies. This could be explained by a concerted evolution of domains within the same gene, but this seems unlikely because the two domains within each gene are substantially diverged. Therefore, it appears that duplication of ATP-binding domains within major

**Figure 2** Phylogenetic tree of the human ABC genes. ATP-binding domain proteins were identified using the model ABC\_tran (accession PF00005) of the pfam database (Bateman et al. 1999). The HMMSEARCH program from the HMMER package (Eddy 1998) and a set of custom-made service scripts were used to extract ATP-binding domains from all protein sequences of interest. Note that some proteins analyzed contain two ATP-binding domains (denoted on the figures as I and II), whereas others contained only one ATP-binding domain. Alignments were generated with the hidden Markov model (Eddy 1995) based on the HMMALIGN program using the ABC\_tran model. The resulting multiple alignment was analyzed with NJBOOT (N. Takezaki, pers. comm.) implementing the neighbor-joining tree-making algorithm (Saitou and Nei 1987) and the number at the branch of the nodes represents the value from 100 replications. The distance measure between sequences used for tree making was the Poisson correction for multiple hits (Zuckerandl and Pauling 1965). To verify the position of the previously unknown subgroup of *Drosophila* genes (CG6162, CG6162, and CG11147), they were aligned with a representative of each of the human subfamilies. Because some of the human proteins had two ATP-binding domains, the set contained three *Drosophila* and 12 human sequences. The JTT (Jones et al. 1992) model as defined in the MOLPHY package with the 'star decomposition' option was employed. The tentative best tree (the total number of possible trees for 15 sequences is too large for exhaustive search through all these trees) was then used for local maximum likelihood search through the surrounding tree topologies.



ABC families was attributable to several independent duplication events rather than a single ancestral duplication.

#### *Drosophila* ABC genes

To begin to understand the organization and evolution of the *Drosophila* ABC genes, the Celera (Myers et al. 2000) and Fly-Base databases were searched for sequences by a combination of BLAST searches and analysis of the annotation already present in the databases. Initial subfamily classifications were assigned based on homology and BLAST scores, and the loca-

tion of each gene was recorded (Table 4). In total, 56 genes were identified and there is at least one representative of each of the known mammalian subfamilies (Table 5). To confirm the subfamily groupings, the ATP-binding domain amino acid sequences were used to perform phylogenetic analyses. A representative tree is shown in Figure 3. As expected, genes from the same subfamily cluster together and confirm the initial assignments made by inspection.

As in the human and yeast genomes, the *Drosophila* ABC genes are largely dispersed in the genome. There are four clus-

**Table 4.** *Drosophila* ABC Genes

Gene	Alias	Protein acc.	DNA acc.	Size (AA), FAMILY	Location (chr. nuc)	Cyto. loc.	Mutants
CG3156		AAF45509	AE003417	609 B	X 252038-254671	1B4	
CG2759	w	AAF45826	AE003425	696 G	X 2545753-2539884	3B4	
CG1703		AAF48069	AE003486	901 E	X 11393813-11396731	10C10	
CG1824		AAF48177	AE003489	761 B	X 12363742-12360802	11B16	
CG9281		AAF48493	AE003500	611 E	X 15454374-15450765	13E14	
CG8473		AAF48511	AE003500	2556 A	X 15513659-15523896	13E18-F1	
CG12703		AE003513	AE003513	618 D	X 19494615-19497465	18F1-F2	bth
CG1819		AAF50847	AE003569	1500 A	X 20757531-20763638	19F1	fir, ms, mit(1)20
CG1718		AAF50837	AE003568	1713 A	X 20909795-20902146	19F2	
CG1801		AAF50836	AE003568	1511 A	X 20924492-20917580	19F2	
CG1494		AAF50838	AE003568	1197 A	X 20896205-20901578	19F2	
CG3164		AAF51548	AE003590	620 G	2L 123902-117541	21B	
CG4822		AAF51551	AE003590	643 G	2L 112000-116000	21B	
CG17646		AAF51341	AE003585	627 G	2L 1720498-1727693	22B3	
CG9892		AAF51223	AE003582	615 G	2L 2649300-2658596	23A6	
CG9664		AAF51131	AE003580	609 G	2L 3211844-3209624	23E4-23E5	
CG9663		AAF51130	AE003580	812 G	2L 3214000-3220000	23E4-23E5	
CG3327		AAF51122	AE003580	729 G	2L 3257267-325948	23F	
CG2969	Atet	AAF51027	AE003576	832 G	2L 4251813-4262480	24F8	
CG11147		AAF52284	AE003611	705 H	2L 5656028-5653232	26A1	
CG7806		AAF52639	AE003620	1487 C	2L 8212839-8218079	29A3-A4	
CG7627		AAF52648	AE003620	1327 C	2L 8262316-8256791	29B1	
CG5853		AAF52835	AE003626	689 G	2L 9854119-9847658	30E1-30E3	
CG5772	Sur	AAF52866	AE003627	2250 C	2L 10105357-10089272	31A2	
CG6214		AAF53223	AE003637	1896 C	2L 12619174-12641593	33F2	
CG7491		AAF53328	AE003641	324 A	2L 13675599-13676775	34D1	
CG17338		AAF53736	AE003661	1275 B	2L 18829742-18834099	37B9	pre, MR
CG10441		AAF53737	AE003661	1307 B	2L 18835157-18839979	37B9	
CG9270		AAF53950	AE003668	1014 C	2L 20741821-20738317	39A2	
CG8799		AAF58947	AE003833	1344 C	2R 4426560-4431236	45D1	
CG3879	Mdr49	AAF58437	AE003820	1279 B	2R 7940090-7934079	49E1	
CG8523	Mdr50	AAF58271	AE003815	1313 B	2R 9235904-9241222	50F1	
CG8908		AAF57490	AE003792	1382 A	2R 15203694-15208725	56F11	
CG10505		AAF46706	AE003453	1283 C	2R 16226805-16222698	57D2	
CG 17632	bw	AAF47020	AE003461	755 G	2R 18476505-18465883	59E3	
CG7955		AAF47526	AE003472	606 B	3L 1597621-1602155	62B1	
CG10226		AAF50670	AE003563	1320 B	3L 6180561-6175400	65A14	
Mdr65		AAF50669	AE003563	1302 B	3L 6186691-6181468	65A14	
CG5651		AAF50342	AE003553	611 E	3L 8895129-8892720	66E3-E4	
CG7346		AAF50035	AE003544	597 G	3L 11555624-11559309	68C10-C11	vin, cin, rose
CG4314	st	AAF49455	AE003527	666 G	3L 16398050-16400715	73A3	
CG5944		AAF49305	AE003522	1463 A	3L 17695681-17689489	74E3-E4	
CG6052		AAF49312	AE003523	1660 A	3L 17627439-17622025	74E3-E4	
CG9330		AAF49142	AE003516	708 E	3L 1971540-1947231	76B6	
CG14709		AAF54656	AE003692	1307 C	3R 7362645-7369141	86F1	
CG4225		AAF55241	AE003710	866 B	3R 11615803-11612420	89A11-A12	
CG4562		AAF55707	AE003728	1348 C	3R 15626899-15619809	92B9	
CG4794		AAF55726	AE003728	711 A	3R 15725586-15728807	92C1	
CG5789		AAF56312	AE003748	1239 C	3R 29281221-20277309	96A7	fs, l, aor,
CG18633		AAF56360	AE003749	702 G	3R 29625526-29622829	96B5	mar, mfs
CG11069		AAF56361	AE003749	602 G	3R 20635134-20637920	96B6	
CG6162		AAF56584	AE003756	535 H	3R 22087630-22088417	97B1	ird15, smi97B, Spn-D
CG9990		AAF56807	AE003766	808 H	3R 24409613-24429503	98F1	spg, lethal
CG11898		AAF56870	AE003768	1302 C	3R 24887241-24892598	99A	
CG11897		AAF56869	AE003768	1346 C	3R 24881629-24885998	99A	
CG2316		AAF59367	AE003844	730 D	4 154260-145146	101F	Scn, 5 lethals

**Table 5.** ABC Gene Subfamilies in Characterized Eukaryotes

Subfamily	Human	<i>Drosophila</i>	Yeast
A	12	19	0
B	11	10	4
C	12	12	7
D	4	2	2
E	1	1	0
F	3	3	5
G	5	15	10
H	0	3	0
Other	0	0	3 <sup>a</sup>
Total	48	56	31

<sup>a</sup>Includes YDR091c, YFL028c, and YDR061w.

ters of two genes and one cluster of four genes (Table 4). One of these clusters (on chromosome 2L, band 37B9) is composed of an ABCB and an ABCC gene indicating that this is a chance grouping of genes. The remaining clusters are composed of genes from the same subfamily and arranged in a head-to-tail fashion consistent with gene duplication. Because the clusters are themselves dispersed and involve different subfamilies, they presumably represent independent gene duplication events.

The best-studied *Drosophila* ABC genes are the eye pigment precursor transporters white, scarlet, and brown (*w*, *st*, and *br*, respectively). These genes are part of the ABCG subfamily and have a unique NBF-TM organization. Surprisingly, there are 15 ABCG genes in the fly genome, making this the most abundant ABC subfamily. This is in sharp contrast to the only five and six known ABCG genes in the human and mouse genomes, respectively. The *Drosophila* ABCG genes are highly dispersed in the genome with only two pairs of linked genes. In addition, they are quite divergent phylogenetically, suggesting that there were many independent and ancient gene duplication events.

Several *Drosophila* ABCB genes, *Mdr49*, *Mdr50*, and *Mdr65*, have also been well characterized. A fourth member of this group, *CG10226*, was identified that is clustered with *Mdr65* (Table 4). These genes are closely related to the human and mouse P-glycoproteins (ABCB1, ABCB4) and disruption of *Mdr49* results in sensitivity to colchicines (Wu et al. 1991).

To search for potential gene functions, each region containing a *Drosophila* ABC gene was searched for phenotypes that have been not assigned to a gene (Table 4). The most promising connection is the identification of several eye phenotypes (*vin*, *rose*, and *cln*) in the region of the *CG7346* gene. Because *CG7346* is part of the ABCG family and is therefore related to *w*, *st*, and *br*, it is tempting to speculate that mutations in *CG7346* cause one or more of these phenotypes. Because ABC genes perform very diverse functions and are associated with varied phenotypes it is hard to gather much additional insight from this analysis.

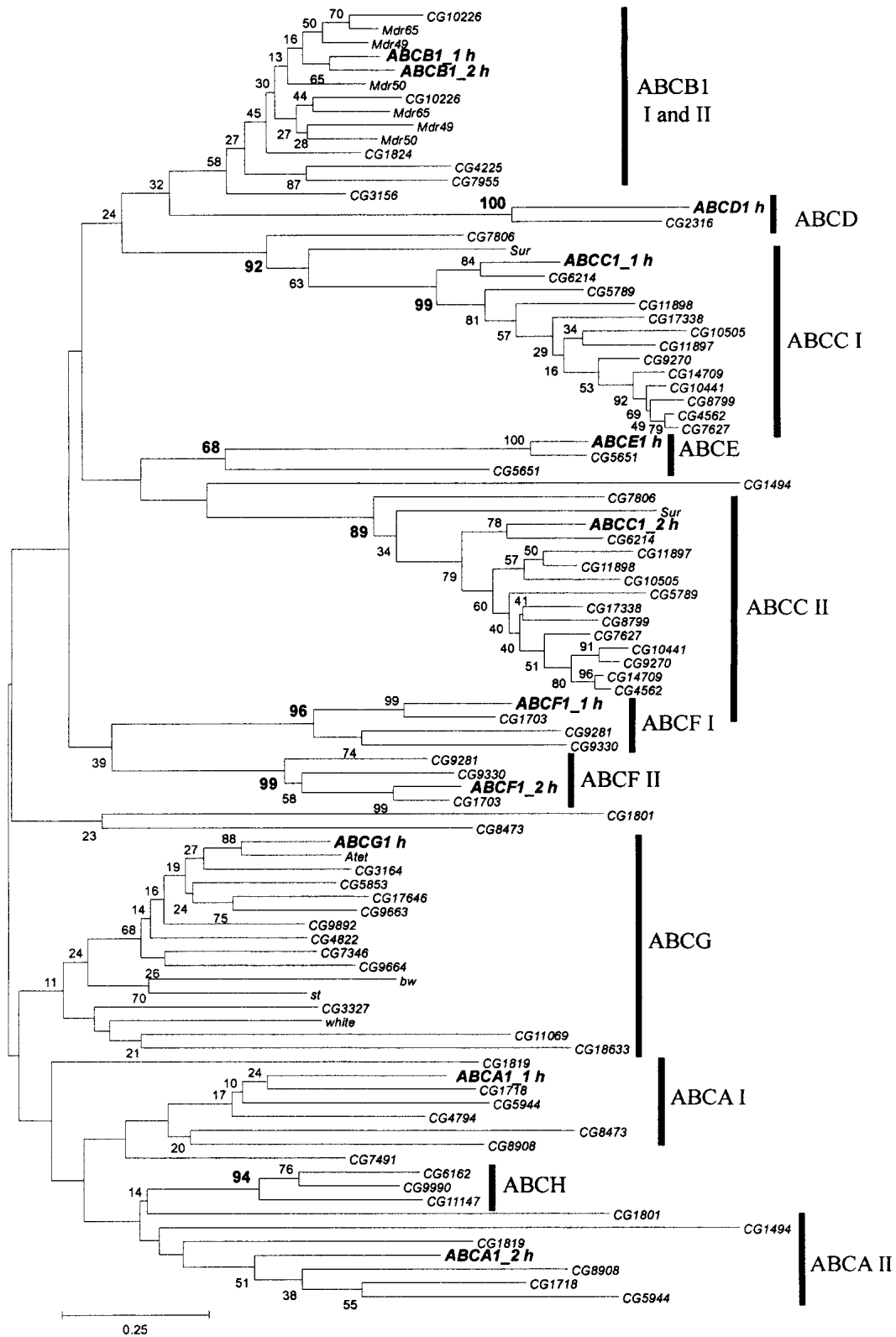
Three genes, *CG9990*, *CG6162*, and *CG11147*, were identified that do not fit into any of the known subfamilies and, in fact, are most closely related to ABC genes from bacteria. There are no close homologs to these genes in any other eukaryotic genome including worm and plants. These genes are within large contiguous sequences and have introns, therefore they do not represent contamination from bacterial sequences. This group forms a distinct cluster on the *Drosophila* tree. To elucidate the position of this group with respect to

human ABC transporter families, we produced two datasets with nine and 15 sequences, respectively, containing ATP-binding domains from the three *Drosophila* genes and from one representative from each of the human families. In the smaller subset, only one ATP-binding domain from each human gene was used. These two datasets were subjected to maximum likelihood analysis, with a heuristic algorithm for the larger dataset and with a rigorous exhaustive search for the smaller datasets. The result of the analysis clearly indicates that the new subfamily of ABC transporter in *Drosophila* is significantly different from all known families of ABC transporters (data not shown) and might have a yet unidentified functional role. We propose designating this new group of genes as subfamily H. DNA sequence searches showed that although significant homology is present between these three genes in both the NBF and TM regions, there is no homology in the TM regions with any other eukaryotic or prokaryotic ABC proteins. The closest related NBF sequences are all bacterial, such as the *E. coli* YH1H gene, a ribosomal ATPase. This gene shows 36% amino acid identity in the NBF domain.

The ATP-binding domains in a generic ABC transporter appear to be following a birth-and-death process that has been described before for other multigene families. In a typical birth-and-death process (Ota and Nei 1994; Nei et al. 1997), repeated genetic segments (in this case domains) experience nearly random fluctuation in their number. That is, genomic deletions decrease segment number, while genome duplications increase it. The phylogenetic trees that we present here for human and *Drosophila* give clear evidence of the birth-and-death evolution of ATP-binding domains in ABC transporters, and similar analyses of the yeast genes support this. In *S. cerevisiae*, there are several G subfamily ABC genes (e.g., YOR011w and SNQ2) that have two NBF and TM domains (Michaelis and Berkower 1995; Decottignies and Goffeau 1997). In contrast, all of the *Drosophila* and human ABCG family genes are half transporters. Therefore, it appears that the second ATP-binding domain in animal ABCG genes was lost in an ancestral lineage preceding the animal radiation but not the animal-fungi split. Similarly, ABCE genes in humans are one-domain genes, whereas both *Drosophila* and yeast have two-domain homologs (RLI1 in yeast and CG5651 in *Drosophila*). Therefore, the loss of the second domain appears to be a relatively recent event in human ABCE genes, probably in the vertebrate lineage.

## Perspectives

One of the most fascinating findings from the analysis of the ABC genes from the human, worm, and fly genomes is the remarkable similarity in the number of ABC genes. Although additional sequencing and annotation will undoubtedly change the exact number somewhat, it is clear that humans do not have substantially more ABC genes than much simpler eukaryotes. This suggests that there is a core of essential ABC genes that are required for all multicellular eukaryotes. These higher eukaryotes do have about twice the number of ABC genes as does *S. cerevisiae*, suggesting that the evolution of multicellularity was accompanied by the expansion in the number of ABC genes. It is clear, however, that certain species have expanded some subfamilies more than others. As mentioned above, flies have considerably more ABCG genes than humans (15 vs. five), whereas there are at least 12 ABCA genes in the human and only 10 in *Drosophila*. Despite the remarkable similarity in gene number, most *Drosophila* ABC genes do not have a clear ortholog in the human genome. The likely



**Figure 3** Phylogenetic tree of the *Drosophila* ABC genes. Analysis (see Fig. 2) was performed with all extracted *Drosophila* sequences and a representative of each human subfamily.

exceptions are the ABCD, ABCE, and ABCF subfamilies that contain nearly identical gene numbers in the two species. The only ABC genes in the human and yeast genomes that are documented to have similar function are the ABCD (ALD-like) genes that are involved in the transport of very long chain fatty acids, and the *ATM1/ABCB7* gene that is present in the mitochondria and has a role in iron metabolism.

A major limitation to the understanding of ABC genes is the difficulty in obtaining crystal structures. The three-dimensional structure of two bacterial NBF proteins has been obtained and has greatly improved our understanding of the organization and function of this portion of the protein (Hung et al. 1998). It is clear that conformational changes in the NBF influence substrate binding and transport, and to date, this can only be addressed by laborious mutagenesis and biochemical experiments (Hafkemeyer et al. 1998).

Ideally, the worm and fly ABC genes could be used to elucidate the function of their human counterparts. To date, however, we know considerably more about human ABC gene function that we do about either the *Drosophila* or *Caenorhabditis elegans* ABC genes. In fact, the only *Drosophila* genes initially identified based on their function are the eye pigment genes. Similarly, only one ABC gene in *C. elegans* (*ced-7*) was identified based on function (involvement in cell death; Wu and Horvitz 1998). It is not obvious why more ABC genes are not associated with visible or detectable phenotypes in *Drosophila* or *C. elegans* as many of the human ABC genes affect morphology, behavior, lifespan, and fertility and these are traits that have been observed and selected for repeatedly in these organisms. These model organisms, however, do provide an important resource for the future systematic study of ABC genes. Analysis of ABC gene expression combined with gene disruptions should yield important clues to gene function. In addition, the capability of designing suppressor screens may allow the identification of pathways that involve ABC genes.

## ACKNOWLEDGMENTS

We thank Kirby Smith for helpful comments on the manuscript and apologize to all whose primary papers could not be cited because of lack of space.

## REFERENCES

- Allikmets, R. 2000. Simple and complex ABCR: genetic predisposition to retinal disease. *Am. J. Hum. Genet.* **67**: 793–799.
- Allikmets, R., Singh, N., Sun, H., Shroyer, N.F., Hutchinson, A., Chidambaram, A., Gerrard, B., Baird, L., Stauffer, D., Peiffer, A., Rattner, A., Smallwood, P., Li, Y., Anderson, K.L., Lewis, R.A., Nathans, J., Leppert, M., Dean, M., and Lupski, J.R. 1997. A photoreceptor cell-specific ATP-binding transporter gene (*ABCR*) is mutated in recessive Stargardt macular dystrophy. *Nat. Genet.* **15**: 236–246.
- Allikmets, R., Schriml, L.M., Hutchinson, A., Romano-Spica, V., and Dean, M. 1998. A human placenta-specific ATP-binding cassette gene (*ABCP*) on chromosome 4q22 that is involved in multidrug resistance. *Cancer Res.* **58**: 5337–5339.
- Allikmets, R., Raskind, W.H., Hutchinson, A., Schueck, N.D., Dean, M., and Koeller, D.M. 1999. Mutation of a putative mitochondrial iron transporter gene (*ABC7*) in X-linked sideroblastic anemia and ataxia (XLSA/A). *Hum. Mol. Genet.* **8**: 743–749.
- Alonso, E.M., Snover, D.C., Montag, A., Freese, D.K., and Whittington, P.F. 1994. Histologic pathology of the liver in progressive familial intrahepatic cholestasis. *J. Pediatr. Gastroenterol. Nutr.* **18**: 128–133.
- Altschuler, D., Hirschhorn, J.N., Klannemark, M., Lindgren, C.M., Vohl, M.C., Nemes, J., Lane, C.R., Schaffner, S.F., Bolk, S., Brewer, C., et al. 2000. The common PPARgamma Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. *Nat. Genet.* **26**: 76–80.
- Ambudkar, S.V. and Gottesman, M.M. 1998. ABC Transporters: Biochemical, cellular, and molecular aspects. *Meth. Enzymol.* **292**: 1–853.
- Bakos, E., Evers, R., Calenda, G., Tusnady, G.E., Szakacs, G., Varadi, A., and Sarkadi, B. 2000. Characterization of the amino-terminal regions in the human multidrug resistance protein (MRP1). *J. Cell. Sci.* **113**: 4451–4461.
- Bateman, A., Birney, E., Durbin, R., Eddy, S.R., Finn, R.D., and Sonnhammer, E.L. 1999. Pfam 3.1: 1313 multiple alignments and profile HMMs match the majority of proteins. *Nucleic Acids Res.* **27**: 260–262.
- Berge, K.E., Tian, H., Graf, G.A., Yu, L., Grishin, N.V., Schultz, J., Kwiterovich, P., Shan, B., Barnes, R., and Hobbs, H.H. 2000. Accumulation of dietary cholesterol in sitosterolemia caused by mutations in adjacent ABC transporters. *Science* **290**: 1771–1775.
- Borst, P., Evers, R., Kool, M., and Wijnholds, J. 2000. A family of drug transporters: The multidrug resistance-associated proteins. *J. Natl. Cancer Inst.* **92**: 1295–1302.
- Brocardo, C., Luciani, M., and Chimini, G. 1999. The ABCA subclass of mammalian transporters. *Biochim. Biophys. Acta* **1461**: 395–404.
- Chen, H., Rossier, C., Lalioti, M.D., Lynn, A., Chakravarti, A., Perrin, G., and Antonarakis, S.E. 1996. Cloning of the cDNA for a human homologue of the *Drosophila* white gene and mapping to chromosome 21q22.3. *Am. J. Hum. Genet.* **59**: 66–75.
- Childs, S. and Ling, V. 1994. The MDR superfamily of genes and its biological implications. *Important Adv. Oncol.* : 21–36.
- Childs, S., Yeh, R.L., Georges, E., and Ling, V. 1995. Identification of a sister gene to P-glycoprotein. *Cancer Res.* **55**: 2029–2034.
- Cohn, J.A., Friedman, K.J., Noone, P.G., Knowles, M.R., Silverman, L.M., and Jowell, P.S. 1998. Relation between mutations of the cystic fibrosis gene and idiopathic pancreatitis. *N. Engl. J. Med.* **339**: 653–658.
- Cole, S.P.C., Bhardwaj, G., Gerlach, J.H., Mackie, J.E., Grant, C.E., Almqvist, K.C., Stewart, A.J., Kurz, E.U., Duncan, A.M.V., and Deeley, R.G. 1992. Overexpression of a transporter gene in a multidrug-resistant human lung cancer cell line. *Science (Washington, DC)* **258**: 1650–1654.
- de Vree, J.M., Jacquemin, E., Sturm, E., Cresteil, D., Bosma, P.J., Aten, J., Deleuze, J.F., Desrochers, M., Burdelski, M., Bernard, O., et al. 1998. Mutations in the MDR3 gene cause progressive familial intrahepatic cholestasis. *Proc. Natl. Acad. Sci.* **95**: 282–287.
- Dean, M. and Allikmets, R. 1995. Evolution of ATP-binding cassette transporter genes. *Curr. Opin. Genet. Dev.* **5**: 779–785.
- Dean, M., White, M.B., Amos, J., Gerrard, B., Stewart, C., Khaw, K.T., and Leppert, M. 1990. Multiple mutations in highly conserved residues are found in mildly affected cystic fibrosis patients. *Cell* **61**: 863–870.
- Decottignies, A. and Goffeau, A. 1997. Complete inventory of the yeast ABC proteins. *Nat. Genet.* **15**: 137–145.
- Deleuze, J.F., Jacquemin, E., Dubuisson, C., Cresteil, D., Dumont, M., Erlinger, S., Bernard, O., and Hadchouel, M. 1996. Defect of multidrug-resistance 3 gene expression in a subtype of progressive familial intrahepatic cholestasis. *Hepatology* **23**: 904–908.
- Dixon, P.H., Weerasekera, N., Linton, K.J., Donaldson, O., Chambers, J., Egginton, E., Weaver, J., Nelson-Piercy, C., de Swiet, M., Warnes, G., et al. 2000. Heterozygous MDR3 missense mutation associated with intrahepatic cholestasis of pregnancy: Evidence for a defect in protein trafficking. *Hum. Mol. Genet.* **9**: 1209–1217.
- Doyle, L.A., Yang, W., Abruzzo, L.V., Krogmann, T., Gao, Y., Rishi, A.K., and Ross, D.D. 1998. A multidrug resistance transporter from human MCF-7 breast cancer cells. *Proc. Natl. Acad. Sci.* **95**: 15665–15670.
- Eddy, S.R. 1995. Multiple alignment using hidden Markov models. *Ismb* **3**: 114–120.
- . 1998. Profile hidden Markov models. *Bioinformatics* **14**: 755–763.
- Gabriel, S.E., Clarke, L.L., Boucher, R.C., and Stutts, M.J. 1993. CFTR and outward rectifying chloride channels are distinct proteins with a regulatory relationship. *Nature* **363**: 263–268.
- Goksel, D.L., Fischbach, K., Duggirala, R., Mitchell, B.D., Aguilar-Bryan, L., Blangero, J., Stern, M.P., and O'Connell, P. 1998. Variant in sulfonylurea receptor-1 gene is associated with high insulin concentrations in non-diabetic Mexican Americans: SUR-1 gene variant and hyperinsulinemia. *Hum. Genet.* **103**: 280–285.
- Hafkemeyer, P., Dey, S., Ambudkar, S.V., Hrycyna, C.A., Pastan, I., and Gottesman, M.M. 1998. Contribution to substrate specificity and transport of nonconserved residues in transmembrane

- domain 12 of human P-glycoprotein. *Biochemistry* **37**: 16400–16409.
- Higgins, C.F. 1992. ABC transporters: From micro-organisms to man. *Annu. Rev. Cell. Biol.* **8**: 67–113.
- Hoffmeyer, S., Burk, O., von Richter, O., Arnold, H.P., Brockmoller, J., John, A., Cascorbi, I., Gerloff, T., Roots, I., Eichelbaum, M., et al. 2000. Functional polymorphisms of the human multidrug-resistance gene: multiple sequence variations and correlation of one allele with P-glycoprotein expression and activity in vivo. *Proc. Natl. Acad. Sci.* **97**: 3473–3478.
- Hung, L.W., Wang, I.X., Nikaido, K., Liu, P.Q., Ames, G.F., and Kim, S.H. 1998. Crystal structure of the ATP-binding subunit of an ABC transporter. *Nature* **396**: 703–707.
- Hyde, S.C., Emsley, P., Hartshorn, M.J., Mimmack, M.M., Gileadi, U., Pearce, S.R., Gallagher, M.P., Gill, D.R., Hubbard, R.E., and Higgins, C.F. 1990. Structural model of ATP-binding proteins associated with cystic fibrosis, multidrug resistance and bacterial transport. *Nature* **346**: 362–365.
- Jones, D.T., Taylor, W.R., and Thornton, J.M. 1992. The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* **8**: 275–282.
- Juliano, R.L. and Ling, V.A. 1976. A surface glycoprotein modulating drug permeability in Chinese hamster ovary cell mutants. *Biochim. Biophys. Acta* **455**: 152–162.
- Kispal, G., Csere, P., Guiard, B., and Lill, R. 1997. The ABC transporter Atm1p is required for mitochondrial iron homeostasis. *FEBS Lett.* **418**: 346–350.
- Klein, I., Sarkadi, B., and Varadi, A. 1999. An inventory of the human ABC proteins. *Biochim. Biophys. Acta* **1461**: 237–262.
- Klucken, J., Buchler, C., Orso, E., Kaminski, W.E., Porsch-Ozcuremez, M., Liebisch, G., Kapinsky, M., Diederich, W., Drobnik, W., Dean, M., Allikmets, R. et al. 2000. ABCG1 (ABC8), the human homolog of the Drosophila white gene, is a regulator of macrophage cholesterol and phospholipid transport. *Proc. Natl. Acad. Sci.* **97**: 817–822.
- Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Doyle, M., FitzHugh, W., et al. 2001. Initial sequencing and analysis of the human genome. International Human Genome Sequencing Consortium. *Nature* **409**: 860–921.
- Langmann, T., Klucken, J., Reil, M., Liebisch, G., Luciani, M.F., Chimini, G., Kaminski, W.E., and Schmitz, G. 1999. Molecular cloning of the human ATP-binding cassette transporter 1 (hABC1): Evidence for sterol-dependent regulation in macrophages. *Biochem. Biophys. Res. Commun.* **257**: 29–33.
- Lee, M.H., Lu, K., Hazard, S., Yu, H., Shulenin, S., Hidaka, H., Kojima, H., Allikmets, R., Sakuma, N., Pegoraro, R. et al. 2001. Identification of a gene, ABCG5, important in the regulation of dietary cholesterol absorption. *Nat. Genet.* **27**: 79–83.
- Marcil, M., Brooks-Wilson, A., Clee, S.M., Roomp, K., Zhang, L.H., Yu, L., Collins, J.A., van Dam, M., Molhuizen, H.O., Loubster, O., et al. 1999. Mutations in the ABC1 gene in familial HDL deficiency with defective cholesterol efflux. *Lancet* **354**: 1341–1346.
- Marton, M.J., Vazquez de Aldana, C.R., Qiu, H., Chakraborty, K., and Hinnebusch, A.G. 1997. Evidence that GCN1 and GCN20, translational regulators of GCN4, function on elongating ribosomes in activation of eIF2alpha kinase GCN2. *Mol. Cell. Biol.* **17**: 4474–4489.
- Michaelis, S. and Berkower, C. 1995. Sequence comparison of yeast ATP binding cassette (ABC) proteins. In *Cold Spring Harbor Symposium*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Miyake, K., Mickley, L., Litman, T., Zhan, Z., Robey, R., Cristensen, B., Brangi, M., Greenberger, L., Dean, M., Fojo, T., et al. 1999. Molecular cloning of cDNAs which are highly overexpressed in mitoxantrone-resistant cells: Demonstration of homology to ABC transport genes. *Cancer Res.* **59**: 8–13.
- Mosser, J., Douar, A.M., Sarde, C.O., Kioschis, P., Feil, R., Moser, H., Poustka, A.M., Mandel, J.L., and Aubourg, P. 1993. Putative X-linked adrenoleukodystrophy gene shares unexpected homology with ABC transporters. *Nature* **361**: 726–730.
- Myers, E.W., Sutton, G.G., Delcher, A.L., Dew, I.M., Fasulo, D.P., Flanigan, M.J., Kravitz, S.A., Mobarry, C.M., Reinert, K.H., Remington, K.A., et al. 2000. A whole-genome assembly of *Drosophila*. *Science* **287**: 2196–2204.
- Nei, M., Gu, X., and Sitnikova, T. 1997. Evolution by the birth-and-death process in multigene families of the vertebrate immune system. *Proc. Natl. Acad. Sci.* **94**: 7799–7806.
- Ota, T. and Nei, M. 1994. Divergent evolution and evolution by the birth-and-death process in the immunoglobulin VH gene family. *Mol. Biol. Evol.* **11**: 469–482.
- Paulusma, C.C., Bosma, P.J., Zaman, G.J.R., Bakker, C.T.M., Otter, M., Scheffer, G.L., Scheper, R.J., Borst, P., and Oude Elferink, R.P.J. 1996. Congenital jaundice in rats with a mutation in a multidrug resistance-associated protein gene. *Science* **271**: 1126–1128.
- Pier, G.B., Grout, M., Zaidi, T., Meluleni, G., Mueschenborn, S.S., Banting, G., Ratcliff, R., Evans, M.J., and Colledge, W.H. 1998. Salmonella typhi uses CFTR to enter intestinal epithelial cells. *Nature* **393**: 79–82.
- Pignatti, P.F., Bombieri, C., Marigo, C., Benetazzo, M., and Luisetti, M. 1995. Increased incidence of cystic fibrosis gene mutations in adults with disseminated bronchiectasis. *Hum. Genet.* **4**: 635–639.
- Quinton, P.M. 1999. Physiological basis of cystic fibrosis: A historical perspective. *Physiol. Rev.* **79**: S3–S22.
- Reis, A.F., Ye, W.Z., Dubois-Laforgue, D., Bellanne-Chantelot, C., Timsit, J., and Velho, G. 2000. Association of a variant in exon 31 of the sulfonylurea receptor 1 (SUR1) gene with type 2 diabetes mellitus in French Caucasians. *Hum. Genet.* **107**: 138–144.
- Repa, J.J., Liang, G., Ou, J., Bashmakov, Y., Lobaccaro, J.M., Shimomura, I., Shan, B., Brown, M.S., Goldstein, J.L., and Mangelsdorf, D.J. 2000. Regulation of mouse sterol regulatory element-binding protein-1c gene (SREBP-1c) by oxysterol receptors, LXRalpha and LXRBeta. *Genes & Dev.* **14**: 2819–2830.
- Robbiani, D.F., Finch, R.A., Jager, D., Muller, W.A., Sartorelli, A.C., and Randolph, G.J. 2000. The leukotriene C(4) transporter MRP1 regulates CCL19 (MIP-3beta, ELC)-dependent mobilization of dendritic cells to lymph nodes. *Cell* **103**: 757–768.
- Saitou, N. and Nei, M. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**: 406–425.
- Shani, N. and Valle, D. 1998. Peroxisomal ABC transporters. *Meth. Enzymol.* **292**: 753–776.
- Strautnieks, S., Bull, L.N., Knisely, A.S., Kocoshis, S.A., Dahl, N., Arnell, H., Sokal, E., Dahan, K., Childs, S., Ling, V., et al. 1998. A gene encoding a liver-specific ABC transporter is mutated in progressive familial intrahepatic cholestasis. *Nat. Genet.* **20**: 233–238.
- Sun, H., Molday, R.S., and Nathans, J. 1999. Retinal stimulates ATP hydrolysis by purified and reconstituted ABCR, the photoreceptor-specific ATP-binding cassette transporter responsible for Stargardt disease. *J. Biol. Chem.* **274**: 8269–8281.
- Thomas, P.M., Cote, G.J., Wohlk, N., Haddad, B., Mathew, P.M., Rabl, W., Aguilar-Bryan, L., Gagel, R.F., and Bryan, J. 1995. Mutations in the sulfonylurea receptor gene in familial persistent hyperinsulinemic hypoglycemia of infancy. *Science* **268**: 426–429.
- Tyzack, J.K., Wang, X., Belsham, G.J., and Proud, C.G. 2000. ABC50 interacts with eukaryotic initiation factor 2 and associates with the ribosome in an ATP-dependent manner. *J. Biol. Chem.* **275**: 34131–34139.
- van Helvoort, A., Smith, A.J., Sprong, H., Fritzsche, I., Schinkel, A.H., Borst, P., and van Meer, G. 1996. MDR1 P-glycoprotein is a lipid translocase of broad specificity, while MDR3 P-glycoprotein specifically translocates phosphatidylcholine. *Cell* **87**: 507–517.
- Venter, J.C., Adams, M.D., Meyers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A., et al. 2001. The sequence of the human genome. *Science* **291**: 1304–1351.
- Wada, M., Toh, S., Taniguchi, K., Nakamura, T., Uchiumi, T., Kohno, K., Yoshida, I., Kimura, A., Sakisaka, S., Adachi, Y., et al. 1998. Mutations in the canalicular multispecific organic anion transporter (cMOAT) gene, a novel ABC transporter, in patients with hyperbilirubinemia II/Dubin-Johnson syndrome. *Hum. Mol. Genet.* **7**: 203–207.
- Weng, J., Mata, N.L., Azarian, S.M., Tzekov, R.T., Birch, D.G., and Travis, G.H. 1999. Insights into the function of Rim protein in photoreceptors and etiology of Stargardt's disease from the phenotype in abcr knockout mice. *Cell* **98**: 13–23.
- Wu, C.T., Budding, M., Griffin, M.S., and Croop, J.M. 1991. Isolation and characterization of Drosophila multidrug resistance gene homologs. *Mol. Cell. Biol.* **11**: 3940–3948.
- Wu, Y.C. and Horvitz, H.R. 1998. The *C. elegans* cell corpse engulfment gene ced-7 encodes a protein similar to ABC transporters. *Cell* **93**: 951–960.
- Young, S.G. and Fielding, C.J. 1999. The ABCs of cholesterol efflux. *Nat. Genet.* **22**: 316–318.
- Zuckerkindl, E. and Pauling, L. 1965. Evolutionary divergence and convergence in proteins. In *Evolving Genes and Proteins* (ed. B.A. Vogel), pp. 97–166. Academic Press, New York.