doi:10.1006/geno.2001.6663, available online at http://www.idealibrary.com on IDEAL **Article**

A Family of 12 Human Genes Containing Oxysterol-Binding Domains

Cynthia J. Jaworski,† Ernesto Moreira,* Aiguo Li, Randy Lee, and Ignacio R. Rodriguez

National Eye Institute, National Institutes of Health, Bethesda, Maryland 20892, USA

**Present address: The Wilmer Eye Institute, The Johns Hopkins University, School of Medicine, 522 Maumenee, 600 N. Wolfe St., Baltimore, Maryland 21287, USA*

†*To whom correspondence and reprint requests should be addressed. Fax: (301) 496-1396. E-mail: jaworski@helix.nih.gov.*

Oxysterol-binding proteins (OSBPs) have been described in a wide range of eukaryotes, and are often found to be part of a multi-gene family. We have used bioinformatics and data mining as a starting point for identifying new family members in humans based on the presence of the OSBP signature EQVSHHPP. In addition to *OSBP* and the recently reported *OSBP2*, we have found 10 other genes encoding oxysterol-binding domains. Here, we report cDNA and deduced peptide sequences of the previously unknown OSBPs and compare the peptides and genes. All of the genes encode a pleckstrin homology domain, except *OSBPL2*. However, two of the peptides, OSBPL2 and OSBPL1A, consist of the OSBP domain only. A second *OSBPL1* transcript (*OSBPL1B*) contains 15 additional upstream exons, with a deduced peptide containing a pleckstrin homology domain. Cladistic analysis divides the human *OSBP* genes into five groups, whose members share similarities in sequence and gene structure; RT-PCR analysis indicates that expression patterns among group members vary widely.

Key Words: oxysterols, oxysterol-binding protein, multi-gene family, bioinformatics

INTRODUCTION

Oxysterols are oxygenated derivatives of cholesterol that occur as products of sterol oxidation or of the metabolic paths followed by sterols, steroids, and bile acids. Physiological functions for oxysterols are diverse. Through transcriptional inhibition of the low-density lipoprotein receptor, cholesterol synthase, and 3-hydroxy-3-methylglutaryl coenzyme A reductase, the oxysterols are part of the regulatory apparatus for cholesterol homeostasis [1]. They are involved in apoptosis [2–4], calcium uptake [5], and cell differentiation [6], and have been implicated in various pathophysiological states such as atherosclerotic plaque formation [7,8]. The full range of oxysterol actions is not known, and the nature of their protein intermediaries is only beginning to be understood.

Oxysterol-binding protein (OSBP; GenBank acc. no. P22059), a cytosolic mammalian protein that binds oxysterols, has been extensively characterized. It has been cloned and sequenced in humans [9] and rabbits [10], and regions of the protein have been identified that mediate ligand binding [11] and interactions with Golgi membranes [12]. Mammalian OSBPs seem to be members of a large family of OSBPs that share salient sequence similarity in evolutionarily diverse species, with representatives in yeast [13], nematodes [14], insects [15], and plants [16]. The high degree of evolutionary conservation has allowed the definition of a PROSITE signature sequence (Prosite ID OSBP: PS01013, E-[KQ]-x-S-H-[HR]-P-P-x-[STACF]-A). The signature sequence is present in all OSBP family members documented at present, and no false positives have been recorded.

A second OSBP was recently characterized in humans (OSBP2) that shares 63% overall peptide sequence identity with OSBP [17]. *OSBP* and *OSBP2* share almost identical gene structures, indicating they are products of relatively recent gene duplication on the evolutionary time scale. Despite this high degree of similarity, OSBP and OSBP2 seem to have different binding affinities for oxysterols [17]. In contrast to the highly similar OSBP and OSBP2 in humans, many species express several very different proteins that nevertheless are members of the OSBP family. For example, there are seven OSBP family members encoded in the yeast genome (*Saccharomyces cerevisiae*) [13]. The yeast genes, designated *OSH-1* through *OSH-7* (oxysterol-binding protein homologue), can be classified into four separate subgroups based on sequence similarity and shared protein features. Peptides from two of the subgroups consist almost entirely of the OSBP domain, which spans approximately 400 amino acids. Peptides from the other two subgroups are nearly twice as

0888-7543/01 $35.00

doi:10.1006/geno.2001.6663, available online at http://www.idealibrary.com on IDEAL

A

```
OSBPL3 MMSDEKNLGVSQKLVSPSRSTSSCS..SKQGSRQ.................................DSWEVVEGLRGEM.NYTQEPPVQ
OSBPL7 MDFQERD.........PPFLPESAQSSKPSSAQQASEL.............................WEVVEEPRVRLGTEGVMPERQ
OSBPL6 MSSDEKGISPAHKTSTPTHRSASSSTSSQRDSRQSIHILERTASSSTEPSVSRQLLEPEPVPLSKEADSWEIIEGLKIGQ.TNVQKPDKH

OSBPL3 EGFLLKKRKWPLKGWHKRFFYLDKGILKYAKSQTDIEREKLHGCIDVGLSVMSVKKSSKCIDLDTEEHIYHLKVKSEEVFDEWVSKLRHH
OSBPL7 EGHLLKRRKWPLKGWHKRYFVLEDGILHYATTRQDITKGKLHGSIDVRLSVMSINKKAQRIDLDTEDNIYHLKIKSQDLFQSWVAQLRAH
OSBPL6 EGFMLKKRKWPLKGWHKRFFVLDNGMLKYSKAPLDIQKGKVHGSIDVGLSVMSIKKKARRIDLDTEEHIYHLKVKSQDWFDAWVSKLRHH

OSBPL3 RMYRQNEIAMFPHEVN.HFFSGSTITDSSSGVFDSI..SSRKRSSISKQNLFQTGSNVSFSCG.GETRVPLWLQSSEDMEKCSKDLAHCH
OSBPL7 RLAERLDMPRGS.......LPSTAHRKVPGAQLPTAATASALPGLGPRE.................KVSSWLRDSDGLDRCSHELSECQ
OSBPL6 RLYRQNEIVRSPRDASFHIFPSTSTAESSPAANVSVMDGKMQPNSFPWQSPLPCSNSLPATCTTGQSKVAAWLQDSEEMDRCAEDLAHCQ

OSBPL3 AYLVEMSQLLQSMDVLHRTYSAPAINAIQGG.SFESPKKEKRSHRRWRSRAIGKDAKGTLQVP.KPFSGPVRLHSSNPNLST.LDFGEE.
OSBPL7 GKLQELHRLLQSLESLHRIPSAPVIPTHQASVTTERPKKGKRTSRMWCTQSFAKDDTIG.........RVGRLHGSVPNLSRYLESRDSS
OSBPL6 SNLVELSKLLQNLEILQRTQSAPNFTDMQAN.CVDISKKDKRVTRRWRTKSVSKDTKIQLQVPFSATMSPVRLHSSNPNLCADIEFQTPP

OSBPL3 KNYSDGSETSSEFSKMQEDLCHIAHKVYFTLRSAFNIMSAEREKLKQLM.EQDASSPSAQVIGLKNALSSALAQNTDLKERLRRIHAES
OSBPL7 ...GTRGLPPTDYAHLQRSFWALAQKVHSSLSSVLAALTMERDQLRDM........HQGSELSRM......GVSEASTGQRRLHSLSTSS
OSBPL6 SHLTDPLESSTDYTKLQEEFCLIAQKVHSLLKSAFNSIAIEKEKLKQMVSEQDHSKGHSTQMARLRQSLSQALNQNAELRSRLNRIHSES

OSBPL3 LL...LDSPAVAKSGDNLAEENSRDENRALVHQLSNESRLSITDSLSEFFDAQEVLLSPSSSENEISDDD.SYVSDISDNLSLDNLSDDL
OSBPL7 .....DTTADSFSSLNPEEQEALYMKGRELTPQLSQTSILSLADSHTEFFDACEVLLSASSSENEGSEEEESCTSEITTSLSEEML..DL
OSBPL6 IICDQVVSVNIIPSPDEAGEQ..IHVSLPLSQQVANESRLSMSESVSEFFDAQEVLLSASSSENEASDDE.SYISDVSDNISEDNTSVAD

OSBPL3 DNERQTLGPVLDSGREAKSRRRTCLPAPCPSSSNISLWNILRNNIGKDLSKVAMPVELNEPLNTLQRLCEELEYSELLDKAAQIPSPLER
OSBPL7 KGAERCQKGGCVPGRPMGPPRRRCLPAASGPGADVSLWNILRNNIGKDLSKVSMPVQLNEPLNTLQRLCEELEYSSLLDQASRIADPCER
OSBPL6 MISRQILNGELTGG.AFRNGRRACLPAPCPDTSNINLWNILRNNIGKDLSKVSMPVELNEPLNTLQHLCEEMEYSELLDKASETDDPYER

OSBPL3 MVYVAAFAISAYASSYYRAGSKPFNPVLGETYECIREDKGFQFFSEQVSHHPPISACHAESRNFVFWQDVRWKNKFWGKSMEIVPIGTTH
OSBPL7 MVYIAAFAVSAYSSTYHRAGCKPFNPVLGETYECERPDRGFRFISEQVSHHPPISACHAESENFAFWQDMKWKNKFWGKSLEIVPVGTVN
OSBPL6 MVLVAAFAVSGYCSTYFRAGSKPFNPVLGETYECIREDKGFRFFSEQVSHHPPISACHCESKNFVFWQDIRWKNKFWGKSMEILPVGTLN

OSBPL3 VTLPVFGDHFEWNKVTSCIHNILSGQRWIEHYGEIVIKNLHDDSCYCKVNFIKAKYWSTNAHEIEGTVFDRSGKAVHRLFGKWHESIYCG
OSBPL7 VSLPRFGDHFEWNKVTSCIHNVLSGQRWIEHYGEVLIRNTQDSSCHCKITFCKAKYWSSNVHEVQGAVLSRSGRVLHRLFGKWHEGLYRG
OSBPL6 VMLPKYGDYYVWNKVTTCIHNILSGRRWIEHYGEVTIRNTKSSVCICKLTFVKVNYWNSNMNEVQGVVIDQEGKAVYRLFGKWHEGLYCG

OSBPL3 GGSSSACVWRANPMPKGYEQYYSFTQFALELNEMDPSSKSLLPPTDTRFRPDQRFLEEGNLEEAEIQKQRIEQLQRERRRVLEENHVEHQ
OSBPL7 PTPGGQCIWKPNSMPPDHERNFGFTQFALELNELTAELKRSLPSTDTRLRPDQRYLEEGNIQAAEAQKRRIEQLQRDRRKVMEENNIVHQ
OSBPL6 VAPSAKCIWRPGSMPTNYELYYGFTRFAIELNELDPVLKDLLPPTDARFRPDQRFLEEGNLEAAASEKQRVEELQRSRRRYMEENNLEHI

OSBPL3 PRFFRK....SDDDSWVSNGTYLELRKDLGFSKLDHPVLW
OSBPL7 ARFFRRQTDSSGKEWWVTNNTYWRLRAEPGYGNMDGAVLW
OSBPL6 PKFFKKVIDANQREAWVSNDTYWELRKDPGFSKVDSPVLW
```

B

```
OSBPL2   MNGEEEFFDAVTGFDSDNSSGEFSEANQKVTGMIDLDTSKNNRIGKTGERPSQENGIQKHRTSLPAPM.FSRSDFSVWTILKKCVGLELS
OSBPL1A  MSEEKD.....................................CGGGDALSNGIKKHRTSLPSPMMFSRNDFSIWSILRKCIGMELS

OSBPL2   KITMPIAFNEPLSFLQRITEYMEHVYLIHRASCQPQPLERMQSVAAFAVSAVASQWERTGKPFNPLLGETYELIREDLGFRFISEQVSHH
OSBPL1A  KITMPVIFNEPLSFLQRLTEYMEHTYLIHKASSLSDPVERMQCVAAFAVSAVASQWERTGKPFNPLLGETYELVRDDLGFRLISEQVSHH

OSBPL2   PPISAFHSEGLNHDFLFHGSIYPKLKFWGKSVEAEPRGTITLELLKHNEAYTWTNPTCCVHNVIIGKLWIEQYGTVEILNHRTGHKCVLH
OSBPL1A  PPISAFHAEGLNNDFIFHGSIYPKLKFWGKSVEAEPKGTITLELLEHNEAYTWTNPTCCVHNIIVGKLWIEQYGNVEIINHKTGDKCVLN

OSBPL2   FKPCGLFGKELHKVEGHIQDKNKKKLFMIYGKWTECLWGIDPVSYESFKKQERRGDHLRKAKLDEDSGKADSDVADDVPVAQ.ETVQVIP
OSBPL1A  FKPCGLFGKELHKVEGYIQDKSKKKLCALYGKWTECLYSVDPATFDAYKKNDKKNTEEKK....NSKQMSTSEELDEMPVPDSESVFIIP

OSBPL2   GSKLLWRINTRPPNSAQMYNFTSFTVSLNELETGMEKTLPPTDCRLRPDIRGMENGNMDLASQEKERLEEKQREARRERAKEEAEWQTRW
OSBPL1A  GSVLLWRIAPRPPNSAQMYNFTSFAMVLNEVDKDMESVIPKTDCRLRPDIRAMENGEIDQASEEKKRLEEKQRAARKNRSKSEEDWKTRW

OSBPL2   FYPGNNPYTGTPDWLYAGDYFERNFSDCPDIY
OSBPL1A  FHQGPNPYNGAQDWIYSGSYWDRNYFNLPDIY
```

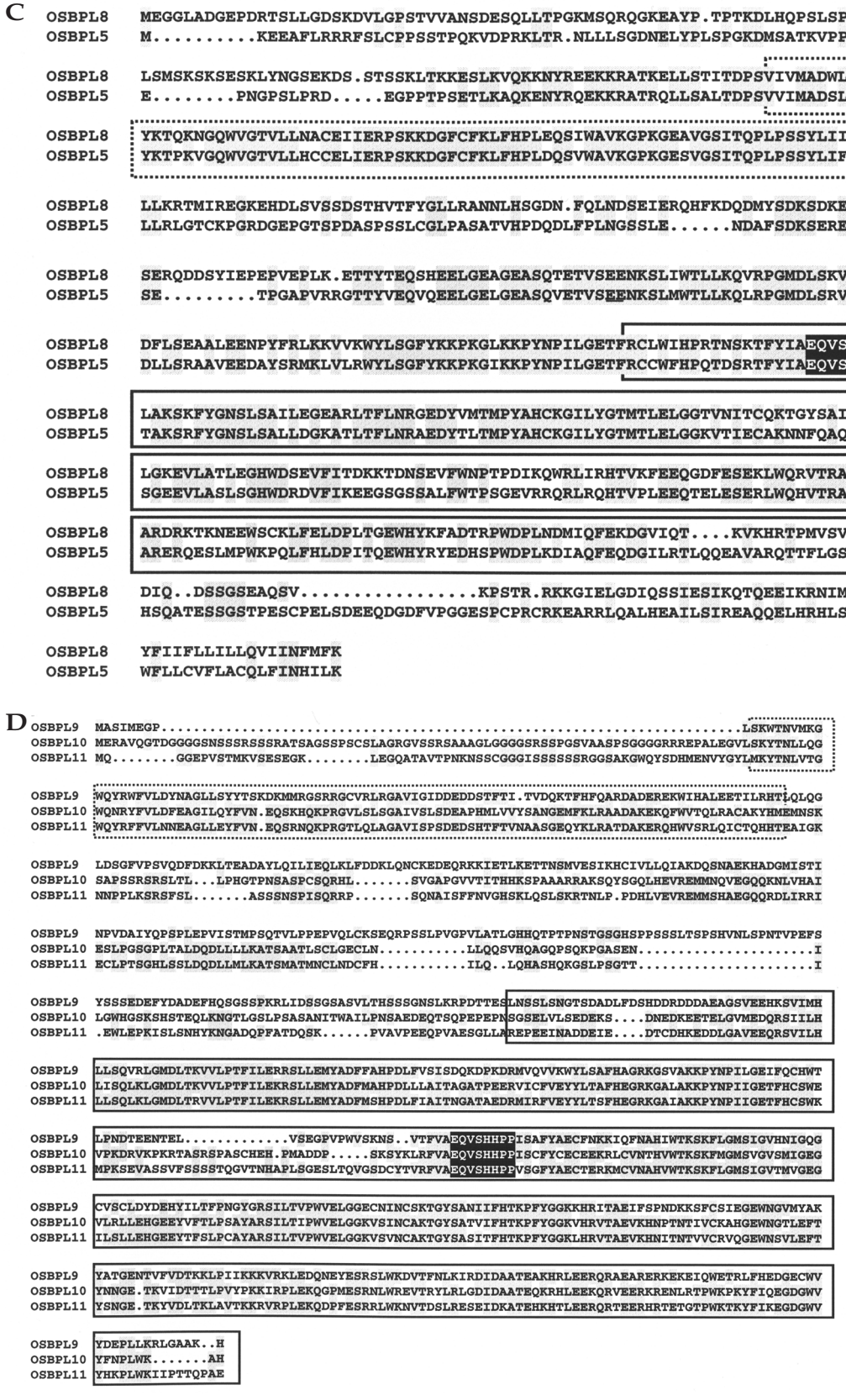

FIG. 1. Peptide sequences of the human OSBPs. Human OSBP family members are divided into groups according to sequence similarity. Global alignments of each group: gray, identical amino acid residues; boxes with solid black lines, oxysterol-binding domains as defined by Pfam; boxes with dotted lines, PH domains; white letters with black background, OSBP signature motif. (A) (Opposite page) Sub-group A2: OSBPL3, OSBPL7, and OSBPL6. (B) (Opposite page) Sub-group A3: OSBPL2 and OSBPL1. (C) (Opposite page) Sub-group B1: OSBPL8 and OSBPL5. (D) (Opposite page) Sub-group B2: OSBPL9, OSBPL10, and OSBPL11. Alignments of OSBP and OSBP2 (sub-group A1) have been published before [17].

doi:10.1006/geno.2001.6663, available online at http://www.idealibrary.com on IDEAL

TABLE 1: Accession numbers and other database reference numbers

	cDNA	Protein	Gene acc. no.	Other designations[a]	Unigene	Locus ID
OSBP	AF185696 NM_002556 M86917 J04757 XM_012050	M86917 NP_002547 P22059 A34581 AAG28373 AAG17011	AP000442 AF185697 to AF185705		Hs.24734	5007
OSBP2	AF288741 AB051451 AF323731	BAB33334 AAG53406	AF288742 AC004542 AL022336 AL079299	ORP-4 KIAA1664	Hs.163427	23762
OSBPL1A	**AF392449** AK001079	BAA91496.1	AC016027 AC016186 AP001177	ORP-1	Hs.252716	
OSBPL1B	**AF392450**		AC023989			
OSBPL2	**AF392447** AB018315 BC004455 BC000296 XM_009562 AY028168 NM_014835 AK000230	CAC22306 AAH00296 AAH04455 AAK18044 BAA34492	AL354836 AL078633	ORP-2 KIAA0772 FLJ20223	Hs.15519	9885
OSBPL3	**AF392444** AB014604 AY008372 AF323727	BAA31679 AAG23400 AAG53408.1 AAB83939	AC003093 AC004016 AC004008	ORP-3 KIAA0704	Hs.197955	26031
OSBPL5	**AF392453** AL136918 AB040967	BAA96058 CAB66852	AC016765	ORP-5 Kiaa1534 OBHP-1	Hs.112034	57656
OSBPL6	**AF392448** AF323728	AAG53409 346291	AC011743	ORP-6	Hs. 318775	
OSBPL7	**AF392446** AK000267 AF323729 XM_008550	AAG53410 BAA91043 NP_060201	AC003665	ORP-7 FLJ20260	Hs.274370	54871
OSBPL8	**AF392452** AL049923	BAA95975	AC017108	ORP-8 Kiaa1451	Hs.109694	57601
OSBPL9	**AF392445** AK022554 XM_016101	BAB14096 NP_078862 XP_016101	AL050343 AL359372	FLJ12492	Hs.21938	79638
OSBPL10	**AF392451** AK000370 BC003168 AF346291	BAA91118 AAH03168 NP_060254	AC020625	ORP-10 FLJ20363	Hs.321622	54907
OSBPL11	**AF392454** AK023074 AK023226 XM_015926 AF346292	BAB14391 BAB14477 XP_015926 NP_073613 AAK31140	AF238377 AC016959 AL357146	ORP-11 FLJ13012 FLJ13164	Hs.61260	64791

GenBank accession numbers for sequences that correspond to cDNA, peptide, and genomic sequences of the 12 human OSBPs. The sequences listed have varying degrees of completeness. Unigene and Locuslink reference numbers from the NCBI databases are included if available. Our own sequences are in bold.

[a]Clone identifiers from sequencing projects or other surveys of OSBP family members as listed in GenBank.

TABLE 2: Characteristics of the human OSBP family members

Name	Chromosomal localization	Peptide length	cDNA	PH domain	Exon number	3′-UTR	Estimated molecular mass
OSBP1	11q12-q13	807	5083	+	14	2180	89.4
OSBP2	22q12	878/916	2791; 4238	+	14		101.3
OSBPL1A	18q11-12	438;	2930;		13;		50.4
OSBPL1B		950	4158	-/+	27	1130	108.5
OSBPL2	20qt3l	480	3970	-	15	2325	55.2
OSBPL3	7p15	887	6689	+	23	3640	101.2
OSBPL5	11p15.4	879	3798	+	22	1087	98.6
OSBPL6	2q32.1	934	3344	+	25	161	106.3
OSBPL7	17q21.2 / 17pter-p13.1	842	3349	+	24	564	95.5
OSBPL8	12p	889 (847)	7240	+	23	4089	101.2
OSBPL9	1p32.2-34.2	737	2911	+	22 (+)	689	83.2
OSBPL10	3p22/3p25.3-3p24.1	764	3933	+	?	1257	84.0
OSBPL11	8	747	4208	+	13	1656	83.6

large and feature N-terminal pleckstrin homology (PH) domains. Similar diversity in the structure of OSBPs is found in the genomes of *Drosophila melanogaster*, *Caenorhabditis elegans*, and *Arabidopsis thaliana* (unpublished data). Our laboratory has detected four OSBP genes in *Drosophila*: three have PH domains and one encodes only the OSBP domain (GenBank acc. nos. AAF47130, AAF58878, AAG22160, and CAA74289; unpublished data). *C. elegans* and *A. thaliana* have 5 and 12 OSBP genes, respectively, and both species have forms with and without PH domains. *A. thaliana* has two variants that have perfectly matching signature sequences (unpublished data).

The existence of multiple OSBP family members with two different structural designs in these species led to the speculation that other OSBP family members might also be present in the human genome. This idea is supported by the fact that cursory surveys of GenBank peptide and nucleotide data have demonstrated the presence of previously unknown sequences within the human genome that are recognizably similar yet obviously distinct from human OSBP and OSBP2 (unpublished data). Similar observations have been made before [18]. Expressed sequence tag (EST) evidence has been demonstrated for other human OSBPs, which were called OSBP-related proteins (ORPs) [19]. Two of these have been cloned and the peptides characterized in detail [20].

The current avalanche of genomic and expressed sequence data provides a resource for discovering and assembling cDNA for previously unknown homologues of known genes and peptides. Consequently, we have used this resource as a basis for examining the OSBP family in the human genome. We have found 10 genes, in addition to *OSBP* and the recently reported *OSBP2*, that share the defining features of the oxysterol-binding domain. Here we describe the cDNA and deduced amino acid sequences, examine differences in protein features and mRNA expression patterns, and present gene structures and chromosomal localization for these previously unknown genes.

RESULTS

Assembly of cDNA and Deduced Peptide Sequences

We began to assemble and compile sequences related to human OSBP by using its peptide sequence as a query with BLASTP and TBLASTN searches of peptide and translated nucleotide databases. The large number of positive matches necessitated a refinement in search techniques. First, we excluded sequences that had similarity to other regions of the OSBP peptide but did not contain the OSBP domain. We found an example of this in genomic DNA on chromosome 19 (GenBank acc. no. AC005795), in which TBLASTN defined exons encoding a PH domain very similar to those of the human OSBPs. However, construction of the remainder of the cDNA showed a deduced peptide lacking the OSBP domain, identified as PLEKHA3, a member of a phosphoinositide binding-specific pleckstrin homology domain-containing family.

Next, it became apparent that the search output derived from a multiplicity of new family members. To sort the new sequences by family member, we needed a means of rapidly differentiating and classifying the sequences. We designed peptide query statements for BLAST searches using the OSBP signature motif and extending it to include the variable segment of 30 amino acids immediately downstream, thus limiting each search result to a gene or gene product containing the OSBP domain. The additional segment allowed comparisons, making it possible to make distinctions among products of different OSBP genes.

In all, we obtained 12 variants of the extended signature from the output, each representing a distinct gene that encodes an OSBP domain. We assembled the corresponding cDNA by incorporating overlapping ESTs and cDNA sequences and complementing them with genomic sequences. We extended the constructed cDNA with 5′ and 3′ rapid amplification of cDNA ends (RACE) procedures [21]. DNA sequencing of RT-PCR and RACE products confirmed and completed the constructed sequences. The 12 genes give rise to 13 mRNAs, as there are two forms of *OSBPL1*.

doi:10.1006/geno.2001.6663, available online at http://www.idealibrary.com on IDEAL

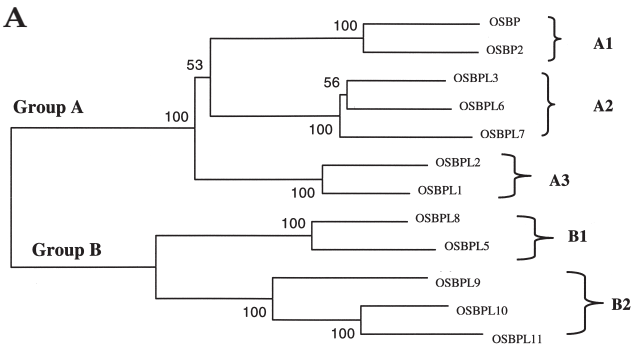

FIG. 2. Sequence comparison of the oxysterol-binding domains from the 12 human OSBPs. (A) Cladogram of amino acid sequences of the human OSBPs. A neighbor-joining tree was produced from the sequence distance data of the aligned oxysterol-binding domains. The branching pattern shows the relationships among the 12 peptides based on sequence similarities. The human OSBPs separate into two clusters: group A and group B. Each group divides further, forming a total of five distinct classes. Numbers represent results of 500 bootstrap replications. (B) Alignment of OSBP domains from the 12 human family members. Identical amino acids are indicated by shading at positions where seven or more match. The location of the aligned signature motif is indicated by double lines above and below the alignment.

B

```
OSBP1   ISG.ASSDISLDEQYKHQLEETKKEKRTRIPYK..PNYSLNLWSIMKNCIGKELSKIPMPVNFNEPLSMLQRLTEDLEYHELLDRAAKCENSLEQLCYVAAFTVSSYSTTVF.RTS.KPF
OSBP2   SVDWSSADNVLDGASLVPKGSSKVKRRVRIPNK..PNYSLNLWSIMKNCIGRELSRIPMPVNFNEPLSMLQRLTEDLEYHHLLDKAVHCTSSVEQMCLVAAFSVSSYSTTVH.RIA.KPF
OSBPL3  SNDLDNERQTLG.PVLDSGREAKSRRRTCLPAPCPSSSNISLWNILRNNIGKDLSKVAMPVELNEPLNTLQRLCEELEYSELLDKAAQIPSPLERMVYVAAFAISAYASSYY.RAGSKPF
OSBPL9  SLNSSLSNGTSDADLFDSHDDRDDDAEAGSV.EEHKSVIMHLLSQVR..LGMDLTKVVLPTFILERRSLLEMYADFFAHPDLFVSISDQKDPKDRMVQVVKWYLSAFHAGRKGSVAKKPY
OSBPL7  MLDLRGAERCQK.GGCVPGRPMGPPRRRCLPAASGPGADVSLWNILRNNIGKDLSKVSMPVQLNEPLNTLQRLCEELEYSSLLDQASRIADPCERMVYIAAFAVSAYSSTYH.RAGCKPF
OSBPL2  DLDTSKNNRIGKTGERPSQENGIQKHRTSLPAPMFSRSDFSVWTILKKCVGLELSKITMPIAFNEPLSFLQRITEYMEHVYLIHRASCQPQPLERMQSVAAFAVSAVASQWE.RTG.KPF
OSBPL6  SVADNISRQILN.GELTGG.AFRNGRRACLPAPCPDTSNINLWNILRNNIGKDLSKVSMPVELNEPLNTLQHLCEEMEYSELLDKASETDDPYERMVLVAAFAVSGYCSTYF.RAGSKPF
OSBPL1  .MSEEKD.....CGGGDALSNGIKKHRTSLPSPMFSRNDFSIWSILRKCIGMELSKITMPVIFNEPLSFLQRLTEYMEHTYLIHKASSLSDPVERMQCVAAFAVSAVASQWE.RTG.KPF
OSBPL10 N...SGSELVLSEDEKSDNEDKEET.ELGVM.EDQRSIILHLISQLK..LGMDLTKVVLPTFILEKRSLLEMYADFMAHPDLLLAITAGATPEERVICFVEYYLTAFHEGRKGALAKKPY
OSBPL8  K......ETTYTEQSHEELGE...ASQTETVSEENKSLIWTLLKQVR..PGMDLSKVVLPTFILEPRSFLDKLSDYYYHADFLSEAALEENPYFRLKKVVKWYLSGFYKKPKGL..KKPY
OSBPL5  R.....RGTTYVEQVQEELGELGEASQVETVSEENKSLMWTLLKQLR..PGMDLSRVVLPTFVLEPRSFLNKLSDYYYHADLLSRAAVEEDAYSRMKLVLRWYLSGFYKKPKGI..KKPY
OSBPL11 GLLAREPEEINADDEIEDTCDHKED.DLGAV.EEQRSVILHLLSQLK..LGMDLTRVVLPTFILEKRSLLEMYADFMSHPDLFIAITNGATAEDRMIRFVEYYLTSFHEGRKGAIAKKPY

OSBP1   NPLLGETFELDRLEENG.....................................YRSLCEQVSHHPPAAAHHAES.KNGWTLRQEIKITSKFRGKYLSIMPLGTIHCIFHATGHHYTWKKVT
OSBP2   NPMLGETFELDRLDDMG.....................................LRSLCEQVSHHPPSAAHYVFS.KHGWSLWQEITISSKFRGKYISIMPLGAIHLEFQASGNHYVWRKST
OSBPL3  NPVLGETYECIR.EDKG.....................................FQFFSEQVSHHPPISACHAES..RNFVFWQDVRWKNKFWGKSMEIVPIGTTHVTLPVFGDHFEWNKVT
OSBPL9  NPILGEIFQCHWTLPNDTEENTEL..............VSEGPVPWVSKNS.VTFVAEQVSHHPPISAFYAECFNKKIQFNAHIWTKSKFLGMSIGVHNIGQGCVSCLDYDEHYILTFPN
OSBPL7  NPVLGETYECER.PDRG.....................................FRFISEQVSHHPPISACHAES..ENFAFWQDMKWKNKFWGKSLEIVPVGTVNVSLPRFGDHFEWNKVT
OSBPL2  NPLLGETYELIR.EDLG.....................................FRFISEQVSHHPPISAFHSEGLNHDFLFHGSIYPKLKFWGKSVEAEPRGTITLELLKHNEAYTWTNPT
OSBPL6  NPVLGETYECIR.EDKG.....................................FRFFSEQVSHHPPISACHCES..KNFVFWQDIRWKNKFWGKSMEILPVGTLNVMLPKYGDYYVWNKVT
OSBPL1  NPLLGETYELVR.DDLG.....................................FRLISEQVSHHPPISAFHAEGLNNDFIFHGSIYPKLKFWGKSVEAEPKGTITLELLEHNEAYTWTNPT
OSBPL10 NPIIGETFHCSWEVPKDRVKPKRTASRSPASCHEH.PMADDP.....SKSYKLRFVAEQVSHHPPISCFYCECEEKRLCVNTHVWTKSKFMGMSVGVSMIGEGVLRLLEHGEEYVFTLPS
OSBPL8  NPILGETFRCLWIHPRTN.....................................SKTF...YIAEQVSHHPPISAFYVSNRKDGFCLSGSILAKSKFYGNSLSAILEGEARLTFLNRGEDYVMTMPY
OSBPL5  NPILGETFRCCWFHPQTD.....................................SRTF...YIAEQVSHHPPVSAFHVSNRKDGFCISGSITAKSRFYGNSLSALLDGKATLTFLNRAEDYTLTMPY
OSBPL11 NPIIGETFHCSWKMPKSEVASSVFSSSSTQGVTNHAPLSGESLTQVGSDCYTVRFVAEQVSHHPPVSGFYAECTERKMCVNAHVWTKSKFLGMSIGVTMVGEGILSLLEHGEEYTFSLPC

OSBP1   TTVHNIIVGKLWIDQSGEIDIVNHKTGD.KCNLKFVPYSYFSRDVARKVTGEVTDPSGKVHFALLGTWDEKM........ECFKV.QPVIGENG....GDARQRGHEAEE..........
OSBP2   STVHNIIVGKLWIDQSGDIEIVNHKTND.RCQLKFLPYSYFSKEAARKVTGVVSDSQGKAHYVLSGSWDEQM........ECSKVMHSSPSSPS....SDGKQKTVYQTL..........
OSBPL3  SCIHNILSGQRWIEHYGEIVIKNLHDDSCYCKVNFIKAKYWSTNA.HEIEGTVFDRSGKAVHRLFGKWHESIYCGGG.....................................
OSBPL9  GYGRSIL.TVPWVELGGECNINCSKTG.YSANIIFHTKPFYGGKKHRITAEIFSPNDKKSFCSIEGEWNGVMY......AKYATGENTVFVDTKKLPIIKKKVR....KLEDQNEYESRS
OSBPL7  SCIHNVLSGQRWIEHYGEVLIRNTQDSSCHCKITFCKAKYWSSNV.HEVQGAVLSRSGRVLHRLFGKWHEGLYRGPT.....................................
OSBPL2  CCVHNVIIGKLWIEQYGTVEILNHRTGH.KCVLHFKPCGLFGKEL.HKVEGHIQDKNKKKLFMIYGKWTECLWGIDPVSYESFKKQERRGDHLRKAKLDEDSGKADSDVADDVPVAQ.ET
OSBPL6  TCIHNILSGRRWIEHYGEVTIRNTKSSVCICKLTFVKVNYWNSNM.NEVQGVVIDQEGKAVYRLFGKWHEGLYCGVA.....................................
OSBPL1  CCVHNIIVGKLWIEQYGNVEIINHKTGD.KCVLNFKPCGLFGKEL.HKVEGYIQDKSKKKLCALYGKWTECLYSVDPATFDAYKKNDKKNTEEKK....NSKQMSTSEELDEMPVPDSES
OSBPL10 AYARSIL.TIPWVELGGKVSINCAKTG.YSATVIFHTKPFYGGKVHRVTAEVKHNPTNTIVCKAHGEWNGTLE......FTYNNGE.TKVIDTTTLPVYPKKIR....PLEKQGPMESRN
OSBPL8  AHCKGILYGTMTLELGGTVNITCQKTG.YSAILEFKLKPFLGSSDCVNQISGKLKLGKEVLATLEGHWDSEVF......ITDKKTDNSEVFWNPTPDIKQWRLIRHTVKFEEQGDFESEK
OSBPL5  AHCKGILYGTMTLELGGKVTIECAKNN.FQAQLEFKLKPFFGGSTSINQISGKITSGEEVLASLSGHWDRDVF......IKEEGSGSSALFWTPSGEVRRQRLRQHTVPLEEQTELESER
OSBPL11 AYARSIL.TVPWVELGGKVSVNCAKTG.YSASITFHTKPFYGGKLHRVTAEVKHNITNTVVCRVQGEWNSVLE......FTYSNGE.TKYVDLTKLAVTKKRVR....PLEKQDPFESRR

OSBP1   .....SRVMLWKRNPLPKNAENMYYFSELALTLNAWESGT....APTDSRLRPDQRLMENGRWDEANAEKQRLEEKQRLSRKKR....EAEAMKATEDGTPYDPYKALWFERKKDPVTKE
OSBP2   .....SAKLLWKKYPLPENAENMYYFSELALTLNEHEEGV....APTDSRLRPDQRLMEKGRWDEANTEKQRLEEKQRLSRRRRLEACGPGSSCSSEEEKEADAYTPLWFEKRLDPLTGE
OSBPL3  ....SSSACVWRANPMPKGYEQYYSFTQFALELNEMDPSSKSLLPPTDTRFRPDQRFLEEGNLEEAEIQKQRIEQLQRERRRVL.............EENHVEHQPRFFRK.....SDD
OSBPL9  .........LWKDVTFNLKIRDIDAATEAKHRLEERQRAEARERKEKEIQWET..RLFHED..GECWVYD........EPLLKRL.....................GAAK....
OSBPL7  ....PGGQCIWKPNSMPPDHERNFGFTQFALELNELTAELKRSLPSTDTRLRPDQRYLEEGNIQAAEAQKRRIEQLQRDRRKVM.............EENNIVHQARFFRRQTDS.SGK
OSBPL2  VQVIPGSKLLWRINTRPPNSAQMYNFTSFTVSLNELETGMEKTLPPTDCRLRPDIRGMENGNMDLASQEKERLEEKQREARRER.............AKEEAEWQTRWFYPGNNPYTGT
OSBPL6  ....PSAKCIWRPGSMPTNYELYYGFTRFAIELNELDPVLKDLLPPTDARFRPDQRFLEEGNLEAAASEKQRVEELQRSRRRYM.............EENNLEHIPKFFKKVIDA.NQR
OSBPL1  VFIIPGSVLLWRIAPRPSNSGQMYNFTSFAMVLNEVDKDMESVIPKTDCRLRPDIRAMENGEIDQASEEKKRLEEKQRAARKNR.............SKSEEDWKTRWFHQGPNPYNGA
OSBPL10 .........LWREVTRYLRLGDIDAATEQKRHLEEKQRVEERKRENLRTPWKP..KYFIQE..GDGWVYF........NPLWK.....................................
OSBPL8  .........LWQRVTRAINAKDQTEATQEKYVLEEAQRQAARDRKTKNEEWSC..KLFELDPLTGEWHYKFA.DTRPWDPLNDMIQFEKDGVIQT....K..VKHRTPMVSVPKMK....
OSBPL5  .........LWQHVTRAISKGDQHRATQEKFALEEAQRQRARERQESLMPWKP..QLFHLDPITQEWHYRYE.DHSPWDPLKDIAQFEQDGILRTLQQEA..VARQTTFLGSPGPR....
OSBPL11 .........LWKNVTDSLRESEIDKATEHKHTLEERQRTEERHRTETGTPWKT..KYFIKE..GDGWVYH........KPLWKII.....................................PTTQ....
```

The cDNA sequences are deposited in the GenBank library (Table 1, accession numbers). In Fig. 1, the deduced amino acid sequences are shown in alignments in which each is shown with its most similar family members. Characteristics of the OSBP peptides are described in Table 2. The peptides range in size from 438 to 950 amino acids, with predicted molecular weights ranging from 50 to 108 kDa for the unmodified peptides. The deduced peptide sequence of each human OSBP contains the signature sequence diagnostic of OSBPs (EQVSHHPP) embedded within a highly conserved OSBP domain of approximately 400 amino acids. PH domains are found near the N terminus of each of the human OSBPs, except for OSBPL2 and the short form of OSBPL1. Targets for phosphorylation (Prosite PS0005), as well as regions resembling rhodopsin–G-protein-

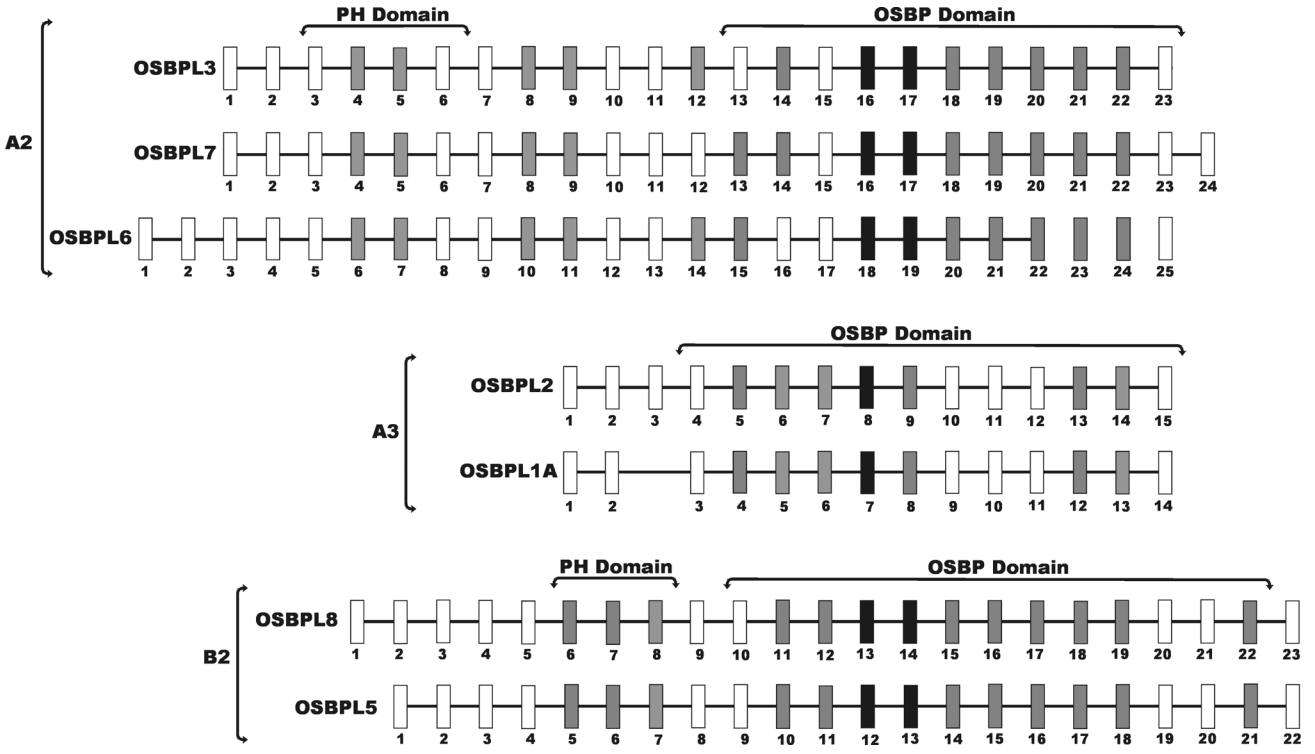

FIG. 3. Gene structures of previously unknown OSBPs. The OSBP genes are presented in the groupings defined in Fig. 2A. Boxes indicate exons. Gray shading indicates that the exons at that position have identical sizes. Boxes shaded in black, also identical, encode the OSBP signature motif. Regions encoding the PH and OSBP domains are indicated above diagrams. Structures are not drawn to scale.

coupled receptor sequences (GPCRRHODOPSN, from the PRINTS database, available at http://www.motif.genome.ad.jp/), are present in multiple copies in all the human OSBP peptides. Other peptide sequence features are found only in selected family members. OSBP and OSBP2 contain leucine zipper motifs. The long form of OSBPL1 has three ankyrin repeats amino-terminal to its PH domain at positions 47–76, 80–109, and 175–204. Typically, ankyrin repeats appear in a cluster of four consecutive repeats.

Family Structure of Human OSBPs

Although the OSBP domains and PH domains are highly conserved, other portions of the human OSBP peptides are not very similar. Consequently, sequence alignments were restricted to a comparison of the oxysterol-binding domains. The 12 domains found in the human family members, as defined by Pfam, are presented in alignment in Fig. 2B. Amino acids are shaded to indicate that at least 7 of the 12 residues in a given position are identical. Many blocks of sequence are conserved among all 12 human OSBPs. Most notable are the signature sequence (EQVSHHPP), bracketed by double lines above and below, and a region in close proximity upstream (KPF/YNXLGETF/Y).

Sequence distances based on this alignment were used to generate a phylogenetic tree with the neighbor-joining algorithm in MEGA2 [22]. The topology of the tree (Fig. 2A) demonstrates the relationships among the human OSBPs and serves to divide the family members into groups on the basis of sequence similarity. Bootstrap values (Fig. 2A, branch lines) serve as indicators of the reliability of each inferred grouping. The domains from the 12 OSBP peptides fall neatly into two clusters, both of which are further divided into sub-groups.

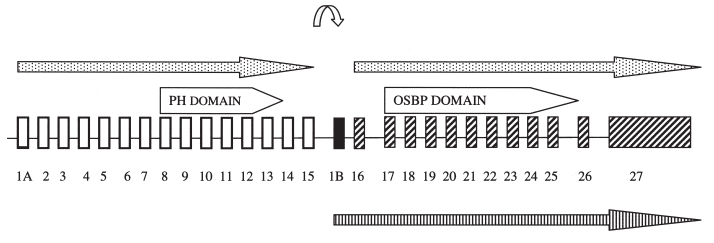

FIG. 4. Gene structure of *OSBPL1*. The origins of the two transcripts arising from *OSBPL1* are shown. Open boxes, exons unique to *OSBPL1B*; filled box, exon 1B, unique to *OSBPL1A*; striped boxes, exons common to both. The shorter transcript is made up of 13 exons, producing a peptide, like OSBPL2, that consists essentially of a free-standing oxysterol-binding domain. The first exon of this transcript is designated 1B; translation begins in the next exon, exon 16. The longer transcript (*OSBPL1B*) contains 15 additional exons upstream, skips exon 1B, then continues onward as for *OSBPL1A*. Open arrows, exons that encode the PH and OSBP domain.

doi:10.1006/geno.2001.6663, available online at http://www.idealibrary.com on IDEAL

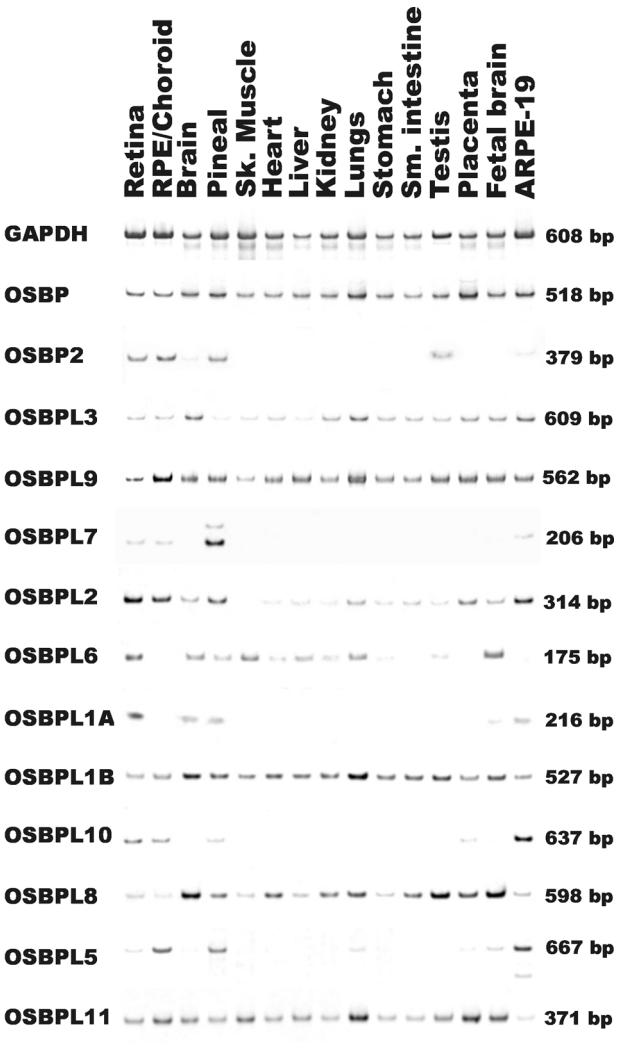

FIG. 5. Survey of OSBP expression. For RT-PCR we used solid-phase cDNA generated from different human tissues (above gel image). Primer pairs are specific for each *OSBP*, both forms of *OSBPL1*, and a *GAPDH* control. All amplifications proceeded for 20 cycles, except for *OSBPL5* analysis, which used 30 cycles. The amplified products were separated by electrophoresis through 10% polyacrylamide gels, stained with SYBR Green, and scanned in a Storm phosphorimager.

The top cluster in the cladogram, called group A, has a bootstrap value of 100, indicating that the members of this group remain together under the most stringent statistical test.

Within group A are three subgroups (two pairs and one set of three) and each subgroup is also supported by a bootstrap value of 100. Subgroup A1 consists of OSBP and OSBP2. OSBP and OSBP2 are highly similar in protein design, sequence, and gene structure, as described [17]. In the context of the other OSBP peptide sequences, this similarity is reflected in the highly significant score for clustering seen here. OSBPL3, OSBPL7, and OSBPL6 make up subgroup A2. The peptide sequences of their oxysterol-binding domains range in similarity from 68% to 74%, and are 64–68% identical. Common characteristics of their gene structures and expression are discussed below. The remaining subgroup, A3, consists of OSBPL2 and OSBPL1. At the peptide level, they are 68% identical and 78% similar, overall. They also share broader features in protein design, representing the only two human OSBP genes that encode peptides lacking a PH domain.

Subgroups A1 and A2 are grouped together with a bootstrap value of 53, which is considerably less robust than the clustering that defines the subgroups. This is a reflection of many examples of sequence similarities shared by subgroup A3 with only one of the other two subgroups.

The remaining sequences (OSBPL5, OSBPL8, OSBPL9, OSBPL10, and OSBPL11) are excluded from group A, defining group B by default. The group B members also segregate into subgroups: OSBPL5 and OSBPL8 form subgroup B1, and subgroup B2 contains OSBPL9 and the OSBPL10 and OSBPL11 pair. Like the subgroups of group A, both of these clusters are characterized by bootstrap values of 100.

With the exception of OSBPL9, each of the human OSBPs has at least one highly similar partner, with an overall sequence identity in the range of 33.2–70.7% (OSBPL9 and OSBPL11, and OSBPL2 and OSBPL1A, respectively). These partners form the basis for the five distinct classes delineated by the cladogram.

Thus, the human OSBPs seem to belong to five distinct classes based on sequence similarities, or shared derived amino acid substitutions, within the oxysterol-binding domains. Trees drawn from alignments of the PH domains, also highly conserved, segregated the OSBPs according to the same pattern (data not shown). Trees drawn with the unweighted pair group method using arithmetic averages and maximum parsimony methods gave similar results. We arrived at essentially the same topologies, with minor differences in bootstrap values. The extensive homology evident in the global alignments of individual groups in Fig. 1 further supports this classification scheme.

Gene Structure

We determined the gene structure for each new OSBP by comparing cDNA sequences with genomic sequences, available from GenBank/National Center for Biotechnology Information (NCBI) as completed bacterial or P1-derived artificial chromosome clones or from high-throughput genome sequences, which are preliminary reports of unordered, noncontiguous sequence (Table 1, accession numbers for genomic sequences). Structures have not been completed for *OSBPL9*, *OSBPL10*, and *OSBPL11* because the finished genomic sequences are not yet available.

The genes encoding OSBP and OSBP2 each have 14 exons, whose sizes and boundaries are nearly identical [17]. The genes in groups A2, A3, and B1 range from 22 to 27 exons in size. Similarity in gene structures is also evident for these newly defined groups (Fig. 3). The gene components are not drawn to scale in Fig. 3, but instead are designed to show the correspondence in exon sizes. Exon sizes, for the most part, are well conserved in the regions encoding the PH and oxysterol-binding domains, in contrast to the peripheral areas. The

genes encoding OSBPL3, OSBPL6, and OSBPL7 also have exons of identical size between the domains. *OSBPL9* has at least 22 exons defined from the current data; we have identified 9 and 13 exons, respectively, for *OSBPL10* and *OSBPL11*.

Many of the nonidentical exons are nevertheless very similar. The PH domain regions of group B1 are flanked on each side by exons that differ by a single codon. In another example, the fifteenth exon of *OSBPL3* and *OSBPL7*, along with the seventeenth exon of *OSBPL6*, are within one codon of each other, with sizes of 251, 248, and 254 bp, respectively. The similarity of gene structure as well as sequence similarity among the members of each class is consistent with a common evolutionary origin.

Two Initiation Sites for *OSBPL1*

The gene encoding OSBPL1 has a complicated means of expression, producing two very different transcripts. *OSBPL1A* is made up of 13 exons. It encodes a protein of 438 amino acids and, like OSBPL2, contains only an OSBP domain. An alternative transcript, *OSBPL1B*, is produced from an upstream starting point. It includes 15 additional exons at the 5′ end, skipping the initial exon (1B) of *OSBPL1A*, and gives rise to a 950-amino-acid peptide containing both the OSBP domain and a PH domain. The relationship between the two transcripts is shown in Fig. 4. Presumptive promoters are upstream of exons 1A and 1B.

Expression Survey

For RT-PCR, we used single-stranded cDNA, synthesized on magnetic beads, from 15 different human tissues, including ARPE-19 cells, a cultured human retinal epithelium (RPE) cell line. We used primer pairs specific for each OSBP and glyceraldehyde phosphodehydrogenase (GAPDH). The amplifications proceeded for 20 cycles, remaining within the exponential range under these conditions. The results (Fig. 5) were intended to provide a broad comparison of the tissue distribution of the various OSBP transcripts. However, because of the great variability in human postmortem tissues, we were not able to quantify their expression accurately. *GAPDH*, our positive control, was readily amplified from each tissue.

Several of the OSBP genes are widely expressed. *OSBP1* and *OSBPL1B*, *OSBPL2*, *OSBPL6*, *OSBPL8*, *OSBPL9*, and *OSBPL11* were present in all tissues, but *OSBPL11* was only barely detectable in ARPE-19 cells. *OSBPL3* may also fit this category, but the overall level of expression seems to be much lower. The remaining OSBPs are far more restricted in distribution. The RT-PCR profiles of *OSBP* and *OSBP2* corroborate previous northern blot analysis evidence [17], showing that unlike *OSBP*, *OSBP2* is found mainly in retina, RPE, pineal, and testis tissue. In these conditions, *OSBPL7* was found in pineal tissue, with faint bands present in retina and RPE/choroid tissue. *OSBPL1A*, in contrast to the longer product of the same gene, was limited to brain, retina, and RPE/choroid tissue, and the ARPE-19 cells. The contrasting expression patterns of *OSBPL1A* and *OSBPL1B* indicate that the upstream and downstream promoters are under very different regulatory control. *OSBPL5* bands could not be found with the conditions used for RT-PCR of the other OSBPs; however, increasing the number of PCR cycles from 20 to 30 yielded products in RPE/choroid and pineal tissues and ARPE-19 cultured cells.

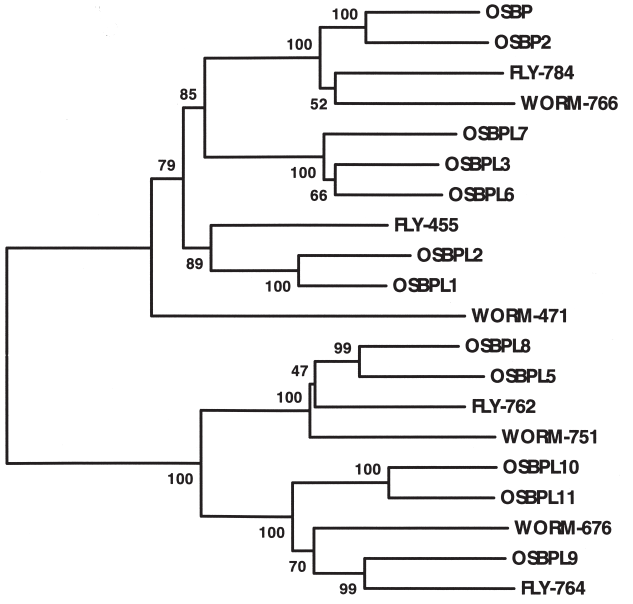

FIG. 6. Cladogram of OSBPs from multicellular animals. A neighbor-joining tree was drawn from the amino acid sequences of the oxysterol-binding domains from human, fruit fly, and nematode OSBPs. Non-human peptides are designated by species name combined with length of the entire peptide. Corresponding accession numbers are: fly-784, CAA74289; fly-455, AAF47130; fly 762, AAG22160; fly-764, AAF58878; worm-766, CAB57894; worm-471, CAB02934; worm 751, CAA94223; and worm 676, AAC24270.

It is salient that for each subgroup of highly similar peptides, the expression patterns are highly divergent. The sequence pairs of groups A1 and B1 each have a widely expressed member coupled with a more selectively expressed counterpart. Members of group A2 are similar in that they cannot be detected by northern blot analysis (data not shown); *OSBPL7* differs in its "preferred" expression in retinal and pineal tissue. Each subgroup may therefore represent gene duplication products that assumed specialized patterns of expression.

DISCUSSION

OSBPs are known to be part of multi-gene families in many eukaryotes. Evidence from sequence databases indicates that this is also true for mammals. EST and cDNA fragments that encode peptides similar but not identical to OSBP have been noted in various contexts [18,19]. The gene and gene products of a second human oxysterol-binding protein, OSBP2, has been described; its expression had been found earlier as a marker of metastasis in lung tissue [17]. Two additional genes containing oxysterol-binding domains, called ORP-1 and ORP-2, have been reported [20].

doi:10.1006/geno.2001.6663, available online at http://www.idealibrary.com on IDEAL

We endeavored to document the remainder of the human OSBPs to provide a more complete basis for experimental designs and a better foundation for understanding functional aspects of the OSBP family of proteins. The nomenclature used here reflects the Pfam definition of the oxysterol-binding domain (PF01237), which is based on sequence similarity and the presence of a stringently conserved signature motif. Examples of proteins containing the OSBP domain are drawn from sequence databases, and notably include entries lacking proven function.

Mining the sequence databases has served as a starting point for assembling other human homologues of OSBPs. Combining bioinformatics with confirmatory cloning and sequencing, we found a total of 12 genes that belong to this family of proteins. With one exception, the genes encode peptides of the same design, combining two well-characterized protein domains. OSBPL2 alone consists only of an oxysterol-binding domain, whereas each of the other genes also encodes an N-terminal PH domain. The gene for OSBPL1 generates two transcripts, one of which produces a peptide lacking a PH domain and is therefore similar to OSBPL2. A second, upstream promoter produces a longer transcript that includes sequences for the PH domain.

In addition to the use of alternative promoters, the human OSBP genes show many examples of alternative splicing. A variant *OSBPL2* transcript omits exon 3, which has no counterpart in the structurally similar gene *OSBPL1*. This variant maintains the same reading frame, and has no effect on the oxysterol-binding domain, differing only by the lack of 12 amino acids N-terminal to the OSBP domain. Similarly, in-frame exon skipping is found for exon 4 of *OSBPL3*. *OSBP2* has a variant lacking exon 12, which alters the amino acid sequence at the carboxy terminus of the peptide, but does not substantially change the oxysterol-binding domain. Several alternatively spliced transcripts of the subgroup A2 (*OSBPL3*, *OSBPL6*, and *OSBPL7*), however, seem to produce salient consequences. These variants are now being investigated in our laboratory.

The existence of two OSBP designs whereby some peptides lack the PH domain is also noted in yeast (*S. cerevisiae*) [13,23–25], nematode (*C. elegans*), and fruit fly (*D. melanogaster*; unpublished data). More extensive similarities in family structure are demonstrated by the cladogram in Fig. 6, based on an alignment of OSBP domains from human, nematode, and fruit fly. The division into group A and group B sequences is conserved in the other animal species. The nematode and fruit fly OSBPs are segregated from same-species family members; most are placed within the disparate clades of the human tree. For example, each fruit fly OSBP is more similar in sequence to specific groups of the human peptides than to other fruit fly OSBP peptides. The short form of fruit fly OSBP is most similar in sequence to OSBPL2 and OSBPL1, which also lack PH domains. The conservation of family structure is consistent with the idea that a divergent family of specialized OSBP genes was present at very early stages of evolutionary history.

Oxysterols have a variety of biochemical and physiological functions. These include regulation of cholesterol synthesis [26–28], modulation of vesicular movement [24], induction of differentiation [6], and involvement with cell cycle regulation and apoptosis [29]. Presumably, these functions are mediated by proteins that bind oxysterols. In addition to the OSBP family of proteins, there are sterol regulatory element-binding proteins (SREBP) [30] and liver X receptor (LXR) a-receptors [31–34], which have also been shown to bind oxysterols. The SREBPs are basic helix–loop–helix transcription factors, and the LXR receptors belong to the class of heterodimeric nuclear receptors. As noted before [13], the OSBP domain has no obvious sequence homology to either the SREBPs or the LXR families. Both of these families have been implicated in transcriptional regulation of sterol biosynthetic pathways [30,35], leaving the OSBP family of proteins as obvious candidates for the other biological functions of oxysterols.

A function for the OSBP family of proteins different from the control of sterol metabolic pathways is further indicated by the fact that the fruit fly does not use the mevalonate pathway, relying instead on dietary sources for the sterol structure [15]. Nevertheless, the fruit fly has four distinct OSBP genes (unpublished data). Studies of the yeast OSBP family members, using complementation assays, showed two family members to be associated with cell cycle control by their involvement with the actions of the Wee1 protein [13]. The high degree of conservation among the phylogenetically disparate OSBPs indicates that the human OSBPs may have functions similar to their counterparts in other species.

The compilation of data here describing the human OSBP family enables us to characterize and compare functional details of the individual proteins. Experiments are now underway in our laboratory to examine subcellular targeting by the different PH domains. Another important project now possible is to express the proteins and study the ligand binding properties of each. OSBP and OSBP2 have different preferences for oxysterols [17]. OSBP2 binds 7-ketocholesterol but not 25-hydroxycholesterol. Binding characteristics of ORP-1 and ORP-2, which correspond to OSBPL1A and OSBPL2, respectively, have been studied [20]. Like OSBP2, these peptides do not bind 25-hydroxycholesterol, but instead bind phospholipids. Elucidation of the targeting, binding specificities, and expression should lead to further understanding of the biological functions of oxysterols and their binding proteins.

Materials and Methods

Search paradigms. We initially searched GenBank at NCBI for OSBP homologues with human OSBP as a peptide query sequence, using BLASTP [36] for the non-redundant protein database and using TBLASTN to search ESTs, non-redundant nucleotide sequences, and high-throughput genomic databases. Default parameters were used.

To simplify subsequent searches, we sought to generate a preliminary classification and organization of the output by clustering "hits" produced by the same gene or peptide. Therefore, we restricted the query sequence(s) to the OSBP signature (EGVSHHPP) plus approximately 20 amino acids immediately following. This ensured that each BLAST hit represented a protein containing

the sequence diagnostic for OSBPs, and also allowed a rapid means to distinguish among the various proteins as they were being discovered.

Construction of cDNA for mined family members. The nucleotide sequence for each new OSBP family member was extended by searching for overlapping EST and non-redundant cDNA sequences and assembling a contig using the SeqMan program from Lasergene (DNAstar, Inc., Madison, WI). Further extensions of the 5′ and 3′ ends were achieved by RACE, and the consensus sequences were verified by DNA sequencing of RT-PCR products. For the 5′-RACE procedure we used human retina cDNA synthesized on magnetic Dynabeads (Dynal Inc., Oslo, Norway) as described [21]. We used the same solid-phase cDNA for 3′-RACE with an oligo-dT anchor primer, and for standard RT-PCR reactions with pairs of specific primers.

RT-PCR, cloning, and sequencing. For PCR we used the solid-phase cDNA templates and AmpliTaq (Perkin-Elmer Cetus Applied Biosystems, Foster City, CA) according to the manufacturer's instructions. Primers were used at final concentrations of 0.8 μM. The PCR profile consisted of an initial 2 min at 94°C, followed by 30 cycles of 10 s at 94°C, 30 s at 60°C, and 1 min at 72°C, with a final incubation of 10 min at 72°C. Amplified products were either sequenced directly or were cloned using the TOPO-TA cloning kits (Invitrogen, San Diego, CA). DNA was sequenced using an ABI Prism 377 automated DNA sequencer (Applied Biosystems, Foster City, CA) or a Beckman CEQ-2000 capillary fluorescent sequencer (Beckman Instruments, Fullerton, CA). Sequencing reactions were done according to the manufacturer's specifications.

Analysis of expression. RT-PCR was done as described above on samples from various tissues, using primer pairs specific for each OSBP cDNA and for GAPDH as a control. We used equal amounts of cDNA saturated solid-phase template representing the following tissues: retina, RPE/choroid, brain, pineal, skeletal muscle, heart, liver, kidney, lung, stomach, small intestine, testis, placenta, fetal brain, and ARPE-19 cultured human retinal pigment epithelial cells [37]. All amplifications proceeded for 20 cycles, except for those for OSBPL5, which used 30 cycles. The amplified products were separated by electrophoresis through 10% polyacrylamide TBE gels (InVitrogen, Carlsbad, CA) and were visualized on a STORM 860 imager (Molecular Dynamics, Inc., Sunnyvale, CA) after being stained with SYBR Green (Molecular Probes, Eugene, OR).

Sequence analysis. Gene structure was determined by using cDNA sequences to query genomic sequences in the database using BLASTN. The splitting of the cDNA sequence to matching regions of the genomic sequence defined exon and intron structures and further served to confirm the cDNA sequence itself. We searched for protein sequence motifs at http://www.motif.genome.ad.jp/ and used the Pfam and prosite entries to define the OSBP signature sequence, OSBP domains, and the PH domains within the OSBP peptides. Nucleotide and peptide sequence alignments were produced with the MegAlign program of the Lasergene software suite (DNAstar, Inc., Madison, WI), using default parameters.

Cladistic analysis. Cladograms of the OSBP peptide sequences were produced by the MEGA2 program [22] using the neighbor-joining algorithm. We calculated distances from simple sequence differences; gaps were handled by pairwise exclusion. Phylogeny was tested with 500 bootstrap replications.

Chromosomal localization. Chromosome location was determined, where possible, through the use of the NCBI's Unigene clusters by querying the database with EST file names for each OSBP. For those not available through Unigene, assignment was made using the Standford G3 panel purchased from Research Genetics (Huntsville, AL).

Oligonucleotide primers. Primers were synthesized by GeneProbe Technology (Gaithersburg, MD). Primers and product sizes for the RT-PCR expression study are as follows (F, forward; R, reverse): OSBP (518 bp): F, 5′-TCACAAGACAGGAGACAAGTG-3′, and R, 5′-TCCGCTCAAACCACAGTGCCTTAT-3′; OSBP2 (379 bp): F, 5′-GAAATGGCCCACACGTGCCGTGG-3′, and R, 5′-CAGATGGGATCTTGAGGCCGTCCAGC-3′; OSBPL3 (609 bp): F, 5′-AGATGCCAAGATGAACACAATG-3′, and R, 5′-AAAGGGACAGAAGGAGAAAAGAAG-3′; OSBPL9 (562 bp): F, 5′-CAGAACGAGTATGANTCCCGCAGCCTTTGG-3′, and R, 5′-CTTAACAGCAGTACTAGAGCGCAGAGT-3′; OSBPL7 (206 bp): F, 5′-GCCGCTCAACACTCTGCAGCGGCT-3′, and R, 5′-GGTCAGGCCGCTCACACTCGTAGGT-3′; OSBPL2 (314 bp): F, 5′-GAGTGGCAGACGAGGTGGTTCTAC-3′, and R, 5′-AAGAGACGGCACAGCAGAGTGAA-3′; OSBPL6 (174 bp): F, 5′-CAGAGATCTCGGAGACGATATATGGAAG-3′, and R, 5′-CAGTCTACCAAAGAACAGGGCTGTCTACT-3′; OSBPL1A (216 bp): F, 5′-CTGGCGGGCAACGCCTCTGCCCGACC-3′, and R, 5′-TCCACCCTGCGCTCCTCGCAAGCTC-3′; OSBPL1B (527 bp): F, 5′-TTCTGGGTAGTGTTAGAGCATGGAG-3′, and R, 5′-CTCACCCCTTCTTGTTTGGTGAAGAG-3′; OSBPL10 (637 bp) F: 5′-GTGGAAGGGCAGCAGAAGGACCTTG-3′, and R: 5(-CGAAGCAAATGACTCTCTCCTCTGG-3′; OSBPL8 (598 bp): F, 5′-GGCTTACTACGTGCTAACAATCTCC-3′, and R, 5′-CAACGGAAAGTCTCGCCAAGTATAGG-3′; OSBPL5 (667 bp): F, 5′-GCCAGGGTCACCAAGAAGGAGACT-3′, and R, 5′-CTGACTCCTCAGGGTTCTCTCTC-3′; OSBPL11 (371 bp): F, 5′-AGAGTGAGACCTCTGGAGAAGCAGG-3′, and R, 5′-TCAGTCTGCGCAATCAGGAAGCAG-3′; GAPDH (608 p): F, 5′-CCACCCATGGCAAATTCCATGG-3′, and R, 5′-TCTAGACGGCAGGTCAGGTCCA-3′.

RECEIVED FOR PUBLICATION JUNE 25;
ACCEPTED SEPTEMBER 28, 2001.

REFERENCES

1. Wolf, G. (1999). The role of oxysterols in cholesterol homeostasis. *Nutr. Rev.* **57:** 196–198.
2. Thompson, E. B., and Ayala-Torres, S. (1999). Oxysterols and apoptosis: evidence for gene regulation outside the cholesterol pathway. *Crit. Rev. Biochem. Mol. Biol.* **34:** 25–32.
3. Rusinol, A. E., *et al.* (2000). Isolation of a somatic cell mutant resistant to the induction of apoptosis by oxidized low density lipoprotein. *J. Biol. Chem.* **275:** 7296–7303.
4. Bakos, J. T., Johnson, B. H., and Thompson, E. B. (1993). Oxysterol-induced cell death in human leukemic T-cells correlates with oxysterol binding protein occupancy and is independent of glucocorticoid-induced apoptosis. *J. Steroid Biochem. Mol. Biol.* **46:** 415–426.
5. Kolsch, H., Lutjohann, D., Tulke, A., Bjorkhem, I., and Rao, M. L. (1999). The neurotoxic effect of 24-hydroxycholesterol on SH-SY5Y human neuroblastoma cells. *Brain Res.* **818:** 171–175.
6. Hanley, K., *et al.* (2000). Oxysterols induce differentiation in human keratinocytes and increase Ap-1-dependent involucrin transcription. *J. Invest. Dermatol.* **114:** 545-553.
7. Peng, S. K., and Morin, R. J. (1987). Effects on membrane function by cholesterol oxidation derivatives in cultured aortic smooth muscle cells. *Artery* **14:** 85–99.
8. Morin, R. J., and Peng, S. K. (1989). The role of cholesterol oxidation products in the pathogenesis of atherosclerosis. *Ann. Clin. Lab. Sci.* **19:** 225–237.
9. Levanon, D., *et al.* (1990). cDNA cloning of human oxysterol-binding protein and localization of the gene to human chromosome 11 and mouse chromosome 19. *Genomics* **7:** 65–74.
10. Dawson, P. A., Ridgway, N. D., Slaughter, C. A., Brown, M. S., and Goldstein, J. L. (1989). cDNA cloning and expression of oxysterol-binding protein, an oligomer with a potential leucine zipper. *J. Biol. Chem.* **264:** 16798-16803.
11. Patel, N. T., and Thompson, E. B. (1990). Human oxysterol-binding protein. I. Identification and characterization in liver. *J. Clin. Endocrinol. Metab.* **71:** 1637–1645.
12. Ridgway, N. D., Dawson, P. A., Ho, Y. K., Brown, M. S., and Goldstein, J. L. (1992). Translocation of oxysterol binding protein to Golgi apparatus triggered by ligand binding. *J. Cell Biol.* **116:** 307–319.
13. Beh, C. T., Cool, L., Phillips, J., and Rine, J. (2001). Overlapping functions of the yeast oxysterol-binding protein homologues. *Genetics* **158:** 1387A–1387.
14. The C. elegans Sequencing Consortium. (1998). Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* **282:** 2012–2018.
15. Alphey, L., Jimenez, J., and Glover, D. (1998). A Drosophila homologue of oxysterol binding protein (OSBP)—implications for the role of OSBP. *Biochim. Biophys. Acta* **1395:** 159–164.
16. Lin, X., *et al.* (1999). Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*. *Nature* **402:** 761–768.
17. Moreira, E. F., Jaworski, C., Li, A., and Rodriguez, I. R. (2001). Molecular and biochemical characterization of a novel oxysterol-binding protein (OSBP2) highly expressed in retina. *J. Biol. Chem.* **276:** 18570–18578.
18. Fournier, M. V., *et al.* (1999). Identification of a gene encoding a human oxysterol-binding protein- homologue: a potential general molecular marker for blood dissemination of solid tumors. *Cancer Res.* **59:** 3748–3753.
19. Laitinen, S., Olkkonen, V. M., Ehnholm, C., and Ikonen, E. (1999). Family of human oxysterol binding protein (OSBP) homologues. A novel member implicated in brain sterol metabolism. *J. Lipid Res.* **40:** 2204–2211.
20. Xu, Y., Liu, Y., Ridgway, N. D., and McMaster, C. R. (2001). Novel members of the human oxysterol-binding protein family bind phospholipids and regulate vesicle transport. *J. Biol. Chem.* **276:** 18407–18414.
21. Rodriguez, I. R., Mazuruk, K., Schoen, T. J., and Chader, G. J. (1994). Structural analysis of the human hydroxyindole-O-methyltransferase gene. Presence of two distinct promoters. *J. Biol. Chem.* **269:** 31969–31977.
22. Kumar, S., Tamura, K., Jakobsen, I. B., and Nei, M. (2001). Molecular Evolutionary Genetics Analysis software. Arizona State University, Tempe, AZ.
23. Schmalix, W. A., and Bandlow, W. (1994). SWH1 from yeast encodes a candidate nuclear

factor containing ankyrin repeats and showing homology to mammalian oxysterol-binding protein. *Biochim. Biophys. Acta* **1219:** 205–210.

24. Fang, M., *et al.* (1996). Kes1p shares homology with human oxysterol binding protein and participates in a novel regulatory pathway for yeast Golgi-derived transport vesicle biogenesis. *EMBO J.* **15:** 6447–6459.

25. Jiang, B., Brown, J. L., Sheraton, J., Fortin, N., and Bussey, H. (1994). A new family of yeast genes implicated in ergosterol synthesis is related to the human oxysterol binding protein. *Yeast* **10:** 341–353.

26. Kandutsch, A. A., and Chen, H. W. (1978). Inhibition of cholesterol synthesis by oxygenated sterols. *Lipids* **13:** 704–707.

27. Kandutsch, A. A., Chen, H. W., and Heiniger, H. J. (1978). Biological activity of some oxygenated sterols. *Science* **201:** 498–501.

28. Chen, H. W., Kandutsch, A. A., Heiniger, H. J., and Meier, H. (1973). Elevated sterol synthesis in lymphocytic leukemia cells from two inbred strains of mice. *Cancer Res.* **33:** 2774–2778.

29. Schroepfer, G. J., Jr. (2000). Oxysterols: modulators of cholesterol metabolism and other processes. *Physiol. Rev.* **80:** 361–554.

30. Brown, M. S., and Goldstein, J. L. (1997). The SREBP pathway: regulation of cholesterol metabolism by proteolysis of a membrane-bound transcription factor. *Cell* **89:** 331–340.

31. Janowski, B. A., Willy, P. J., Devi, T. R., Falck, J. R., and Mangelsdorf, D. J. (1996). An oxysterol signalling pathway mediated by the nuclear receptor LXR α. *Nature* **383:** 728–731.

32. Taylor, F. R., Kandutsch, A. A., Anzalone, L., Phirwa, S., and Spencer, T. A. (1988). Photoaffinity labeling of the oxysterol receptor. *J. Biol. Chem.* **263:** 2264–2269.

33. Saucier, S. E., Kandutsch, A. A., Gayen, A. K., Swahn, D. K., and Spencer, T. A. (1989). Oxysterol regulators of 3-hydroxy-3-methylglutaryl-CoA reductase in liver. Effect of dietary cholesterol. *J. Biol. Chem.* **264:** 6863–6869.

34. Spencer, T. A., *et al.* (2001). Pharmacophore analysis of the nuclear oxysterol receptor LXRalpha. *J. Med. Chem.* **44:** 886–897.

35. Brown, M. S., and Goldstein, J. L. (1999). A proteolytic pathway that controls the cholesterol content of membranes, cells, and blood. *Proc. Natl. Acad. Sci. USA* **96:** 11041–11048.

36. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215:** 403–410.

37. Dunn, K. C., Aotaki-Keen, A. E., Putkey, F. R., and Hjelmeland, L. M. (1996). ARPE-19, a human retinal pigment epithelial cell line with differentiated properties. *Exp. Eye Res.* **62:** 155–169.

Sequence data from this article have been deposited with the DDBJ/EMBL/GenBank Data Libraries under accession numbers AF392449 (OSBPL1A), AF392450 (OSBPL1B), AF392447 (OSBPL2), AF392444 (OSBPL3), AF392453 (OSBPL5), AF392448 (OSBPL6), AF392446 (OSBPL7), AF392452 (OSBPL8), AF392445 (OSBPL9), AF392451 (OSBPL10), and AF392454 (OSBPL11).