Characterization of the human patatin-like phospholipase family[S]

Paul A. Wilson,[1,*] Scott D. Gardner,[†] Natalie M. Lambie,[§] Stephane A. Commans,[**] and Daniel J. Crowther[*]

Bioinformatics Discovery and Analysis* and Disease and Biomarker Transcriptomics,[§] GlaxoSmithKline Research and Development, Stevenage, England SG1 2NY; Quantitative Expression,[†] GlaxoSmithKline Research and Development, Upper Providence, Collegeville, PA 19426; and Laboratoires GlaxoSmithKline,** Centre de Recherches, 91951 Les Ulis, France

Abstract Several publications have described biological roles for human patatin-like phospholipases (PNPLAs) in the regulation of adipocyte differentiation. Here, we report on the characterization and expression profiling of 10 human PNPLAs. A variety of bioinformatics approaches were used to identify and characterize all PNPLAs encoded by the human genome. The genes described represent a divergent family, most with a highly conserved ortholog in several mammalian species. In silico characterization predicts that two of the genes function as integral membrane proteins and are regulated by cAMP/cGMP. A structurally guided protein alignment of the patatin-like domain identifies a number of conserved residues in all family members. Quantitative PCR was used to determine the expression profile of each family member. Affymetrix-based profiling of a human preadipocyte cell line identified several members that are differentially regulated during cell differentiation. Cumulative data suggest that patatin-like genes normally expressed at very low levels are induced in response to environmental signals. Given the observed conservation of the patatin fold and lipase motif in all human PNPLAs, a single nomenclature to describe the PNPLA family is proposed.—Wilson, P. A., S. D. Gardner, N. M. Lambie, S. A. Commans, and D. J. Crowther. **Characterization of the human patatin-like phospholipase family. *J. Lipid Res.* 2006.** 47: **1940–1949.**

Supplementary key words patatin-like phospholipase • adiponutrin • desnutrin • neuropathy target esterase • phospholipase A2, group VI • Intracellular membrane-associated calcium-independent phospholipase A2γ • adipocyte • TaqMan • Affymetrix

The patatin glycoprotein is a nonspecific lipid acyl hydrolase that is found in high concentrations in mature potato tubers (1). Patatin is reported to play a role in plant signaling (2), to cleave fatty acids from membrane lipids (2), and to act as defense against plant parasites (3). Proteins encoding a patatin-like domain are ubiquitously distributed across all life forms, including eukaryotes and prokaryotes [see the species distribution tree of the PFAM (protein family) database (4), accession number PF01734 (http://www.sanger.ac.uk//cgi-bin/Pfam/getacc?PF01734), for further details], and are observed to participate in a miscellany of biological roles, including sepsis induction (5), host colonization (6), triglyceride metabolism (7), and membrane trafficking (8). Of interest is the observation that prokaryotic patatin-like proteins appear more similar to eukaryotic paralogues than any other bacterial lipases (9). This suggests that distant homologs may have arisen from a common ancestor. Together, these observations indicate that the patatin-like domain represents a unique superfamily quite distinct from all other lipolytic families.

Recent studies investigating the role of the hormone-sensitive lipase in mammalian adipose tissue identified a novel patatin-like lipase (TTS-2.2) that was significantly upregulated in differentiating murine adipocytes and appeared to act coordinately with hormone-sensitive lipase in the catabolism of triglycerides (10). The rat ortholog of this gene was also found to be upregulated in differentiating adipocytes and transiently induced during fasting (11). The yeast protein Tgl4 has since been reported to be a functional ortholog of the mammalian TTS-2.2 gene (7).

Jenkins et al. (12) reported details of three murine patatin-like proteins [adiponutrin (ADPN), TTS-2.2, and GS2] with abundant triacylglyerol (catabolic) and transacylation (anabolic) activities. All three were observed to associate with the cell membrane. The ADPN gene transcript was found to be virtually absent in a fasted state and dramatically upregulated in response to feeding. A more recent report (13) characterized four human ADPN-like lipases (ADPN, TTS-2.2, GS2, and GS2-like) and detailed changes in gene expression during murine adipocyte differentiation. Both TTS-2.2 and GS2 were highly expressed in metabolically active tissues. It was also reported that,

Manuscript received 26 April 2006 and in revised form 9 June 2006.

Published, JLR Papers in Press, June 25, 2006.
DOI 10.1194/jlr.M600185-JLR200

[1] To whom correspondence should be addressed.
e-mail: paw84313@gsk.com

[S] The online version of this article (available at http://www.jlr.org) contains supplemental data.

except for ADPN, overexpression of the ADPN-like proteins resulted in decreased intracellular triglyceride levels (13).

Given the significant association of the patatin-like proteins in adipocyte differentiation and their response to metabolic stimuli, the purpose of this investigation was to use a variety of bioinformatics approaches to identify and characterize all human proteins encoding a patatin-like domain. Combined with gene expression data, the collected observations will be used to further assess and characterize the role of these phospholipases in mammalian lipid metabolism.

MATERIALS AND METHODS

Bioinformatics analysis

A redundant human sequence set was identified from the PFAM (4) patatin family alignment (accession number PF01734). Each of these sequences was used as input to a Position-Specific Iterated Basic Local Alignment Search Tool (PSI-BLAST) (14) search to query a nonredundant (>94% identity between sequences) database composed of all available human gene and gene prediction sequences (included were the TREMBL, Swiss-Protein, GenPept, ENSEMBL, GeneSeqP, PIR, and PDB databases).

Equivalent full-length transcript sequences of each patatin-like gene were used to query a redundant human expressed sequence tag (EST) database (GenBank release 151, December 2005) in an effort to verify the expression of the transcripts. Gene expression was further verified using the GeneLogic Genesis Enterprise database (http://www.genelogic.com/genomics/genesis.cfm).

The patatin domain encoded in each protein was delineated using the SMART database (15). Secondary structure was predicted using JNET (16), and fold prediction was completed using the FUGUE application (17). Low-resolution homology models were generated using the Swiss-PDB Viewer application (18).

Sequence alignments were completed using ClustalW (19). The resulting alignment was used as input for PhyloWin (20) to determine putative evolutionary relationships between members of the human patatin-like gene family. Full-length protein sequences were used as input to TMHMM (21) and SignalP (22) to predict transmembrane domains and signal peptide sequences, respectively.

Putative orthologs in mouse and rat were identified by reverse BLASTing a nonredundant nucleotide database derived from the EMBL nucleotide database (23).

TaqMan real-time quantitative reverse transcription-PCR analysis

Tissue processing and sample preparation were described previously (24). Briefly, cDNA prepared from 0.4 ng of poly(A+) RNA extracted from each tissue sample was loaded into each well on a 384-well optical microplate. Gene-specific primers were designed using Primer Express software (Applied Biosystems) and confirmed by BLAST searches of public and propriety sequence databases. Details of each primer are as follows (gene name/forward/reverse/probe): patatin-like phospholipase 3 (PNPLA3)/ADPN/AAATGCCAGTGAGCAGCCAA/TCTCTGCTGGACAGCCCTTG/FAM-CTCCCCATGCACACCTGAGCAGGACTTAMRA; PNPLA7/FLJ43070/GCCTCTGTACCTGCCCTGCT/CTGTATGCAGGGCTGCTGGT/FAM-CCCAGAGAACCCTAACACAGCCTGGGG-TAMRA; PNPLA6/NTE/AGCCACAGATGCCTGAGGAC/GGGCAGGTCAGTCCAGTGTG/FAM-CTCACTCCCCCTCCTGCTGCTATGCCT-TAMRA; PNPLA9/PLA2G6/ACCGCGAGGAGTTCCAGAAG/AATGGACGAGGTCAGCTGGG/FAM-TCATCCACCTGCTGCTCTCACCCTGAGTAMRA; PNPLA8/iPLA2γ/GCTAGACCCTGTTGCCCAGA/GGTTCTTGCAGCAAAGGCAG/FAM-TTGAACCACATCTCACAGCCTCTGTGA-TAMRA; PNPLA1/GCAGTGGGAGATTGGGCTTT/TGTCCCAACCATGATCCCTG/FAM-AAAATTCCTGCTCTGCCACAGCTCCAC-TAMRA; PNPLA4/GS2/AACGTGTGGCAATTGTGGGA/CCCTAACTGCACTGGGAGCA/FAM-TTGGCTGTGTCCCCACCCAAATCTCATTAMRA; PNPLA5/GS2-like/GGCGGACTTGTGGTGGATG/GATGGGCCCAAGGAGCTG/FAM-AAACCTCGAGGGCCATGTTCCTCAGC-TAMRA. The concentrations of the forward and reverse primer and probe for each assay were 900, 900, and 100 nM, respectively. Quantitative PCR was carried out using the 7900HT Sequence Detector System (Applied Biosystems) in a 10 μl reaction volume. TaqMan Universal PCR Master Mix 2X (Applied Biosystems) and universal PCR conditions recommended by the vendor (Applied Biosystems) were used. Abundance values for each tissue sample were calculated using a genomic DNA standard curve to convert TaqMan fluorescent output to a known number of sequence copies.

Microarray analysis

Differentiation of multipotent adipose-derived stem cells, isolated from human adipose tissue, was completed as described previously (25). Total RNA was extracted using Trizol (Invitrogen, Carlsbad, CA). RNA purity and integrity were evaluated using the Agilent Bioanalyzer and Optical Density spectrophotometer. All samples were prepared according to the Affymetrix (Palo Alto, CA) protocol (http://www.affymetrix.com/support/). Fifteen micrograms of labeled copy RNA samples was hybridized onto U133plus2 microarrays (Affymetrix) using standard Affymetrix protocols. Quality-control parameters were as follows: noise (RawQ) < 3, consistent scale factors, background signal < 50, consistent number of genes called as "present" across arrays, and β-actin and GAPDH 5′/3′ signal ratios < 3. Signal intensities were collated and analyzed using Rosetta's Resolver Biosoftware (http://www.rosettabio.com/products/resolver/default.htm).

RESULTS

Extracting a nonredundant set of human sequences from the PFAM PNPLA alignment (accession number PF01734) generated eight unique patatin-like sequences (ADPN, TTS-2.2, Neuropathy Target Esterase (NTE), FLJ43070, IPA2γ, PNPLA1, PNPLA4, and PNPLA5). Each sequence was used to search a nonredundant human transcript database using a PSI-BLAST-based algorithm (14). This identified two additional patatin-like sequences, XM_066899 and Phospholipase A2, group VI (PLA2G6), not included in the PFAM alignment. BLAST alignments indicated that both sequences are most homologous to intracellular membrane-associated calcium-independent phospholipase A2γ (iPLA2γ) (data not shown).

All 10 human sequences that encoded a significant patatin-like domain were compared using the PFAM motif search tool and the FUGUE profile search tool. Both the PFAM and FUGUE similarity scores for each of the human patatin-like domains are detailed in **Table 1**. Both techniques were in broad agreement and indicated that the NTE, NTE-like, and iPLA2γ sequences had higher homology to the patatin domain than did the ADPN-like subfamily.

TABLE 1. Summary of predicted key characteristics of the human PNPLA family

Gene Name	Peptide Length	PFAM E Value (PF01734)	FUGUE Similarity Score (hs1oxwa)	Predicted Signal Peptide	Predicted Transmembrane Domain	Other Predicted Motifs
PNPLA1	437	2.7e-9 (5–81)	6.91			Tyrosine kinase phosphorylation site (29–36 and 169–177)
PNPLA8	782	2e-52 (445–640)	19.56			Tyrosine kinase phosphorylation site (217–223)
PNPLA7	1,317	4.5e-56 (928–1,094)	16.87		13–32	Cyclic nucleotide binding sites (145–694) Glycosaminoglycan attachment site (1,027–1,030) Tyrosine kinase phosphorylation site (1,013–1,020, 1,127–1,135, and 1,181–1,188) Cell attachment sequence (658–660)
PNPLA2	504	1.8e-25 (10–179)	13.36		7–26	Prokaryotic membrane lipoprotein lipid attachment site (36–46)
PNPLA6	1,327	2.9e-63 (933–1,099)	17.53		9–31	Cyclic nucleotide binding sites (147–701) Glycosaminoglycan attachment site (735–738) Tyrosine kinase phosphorylation site (403–410) Cell attachment sequence (663–665) Uncharacterized protein family UPF0028 signature (954–986)
PNPLA9	806	4.1e-55 (481–665)	13.34			Ankyrin repeats (151–412)
PNPLA5	429	5.7e-13 (12–181)	16.19			Glycosaminoglycan site (33–36 and 353–356)
PNPLA3	481	3.7e-25 (10–179)	14.90			
PNPLA4	253	9.8e-41 (6–176)	15.73	Yes		Tyrosine kinase phosphorylation site (246–253)

PNPLA, patatin-like phospholipase. In silico characterization of the human patatin-like family of phospholipases. Summary of in silico predictions: column 2 indicates the length of the longest representative human peptide sequence in the public domain. Columns 3 and 4 list the scoring outputs for each of the predicted patatin-like domains from the PFAM and FUGUE algorithms, respectively. Columns 5, 6, and 7 summarize predicted signal peptides, transmembrane domains, and significant functional motifs as determined by the SignalP, TMHMM, and SMART applications. The relative positions of the predicted motifs shown in parentheses refer to the full-length peptide sequence used as input.

Attempts to align the human protein sequences resulted in low-quality alignments as a result of minimal homology between input sequences. The residues that determined homology to the patatin domain were identified using the PFAM protein query tool (accession number PF01734/patatin domain), and the alignment procedure was repeated using the restricted sequence as input. Although this approach demonstrated an improvement relative to the full-length protein alignment, much of the sequence remained poorly aligned. Information derived from secondary structure prediction (JNET) and fold prediction (FUGUE) was used to generate a manual alignment of the patatin-like sequences. The resulting alignment highlighted several regions of high conservation, including the characteristic glycine-rich region, the nucleophilic elbow, and the aspartate-glycine residues of the catalytic center. When shaded by physiochemical characteristics (**Fig. 1**), the protein alignment clearly illustrated the conservation of similar residues at several defined positions dispersed throughout the patatin-like domain. Such residues are obvious candidates for future mutagenesis studies. The alignment also highlighted blocks of physiochemical conservation across those regions predicted to fold as α-helices and β-sheets interspersed by highly divergent loop regions.

The alignment also highlights the fact that the PNPLA1 sequence does not include the N-terminal region of the patatin domain. In particular, it is missing both the glycine-rich region and the G-X-S-X-G motif, which includes the nucleophilic serine, an essential component of the catalytic dyad. Without these residues, it is highly unlikely that PNPLA1 could function as a lipolytic enzyme. Extensive searches of the EMBL nucleotide sequence database (release 85, December 2005) did not identify additional human sequences that could be used to extend the N-terminal coding sequence of PNPLA1. However, predicted ortholog sequences from rat (XP_342106), mouse (XP_484618), dog (XP_538884), and chimpanzee (XP_527370) all demonstrate a highly conserved N-terminal sequence that encodes the lipase signature motifs. Use of the predicted PNPLA1 sequences to query the human genome sequence with BLASTP (data not shown) identified what appears to be an additional, upstream exon that would encode both the glycine-rich region and the G-X-S-X-G motif, immediately upstream of the human PNPLA1 locus. The predicted full-length sequence of the human PNPLA1 gene has been submitted to GenBank (accession number AM182887).

PNPLA4 includes an alanine substitution for the first conserved glycine in the G-X-G-X-X-G motif. The alanine substitution is seen in all available human transcripts and is also the equivalent residue in the dog, mouse, cow, and frog ortholog sequences (alignment not shown). A serine residue occurs in the equivalent position in the rat ortholog sequence. The functional implications of this substitution are currently unclear, but it does not negate lipolytic activity, as PNPLA4 has previously demonstrated both lipase (13) and transacylation (12) activities.

Sequence comparisons indicated that the putative human patatin-like family could be divided into four subfamilies: an ADPN-like family, as proposed by Lake et al. (13), an NTE-like subfamily, and two unique members represented by PNPLA9 (PLA2G6) and PNPLA8 (iPLA2γ). This observation was reinforced by a phylogenetic comparison of the patatin-like domain sequences (**Fig. 2**). PNPLA1 was found to cluster with the ADPN-like subfamily (data not shown). Because the method incorporated a global gap-removal function, the initial tree was

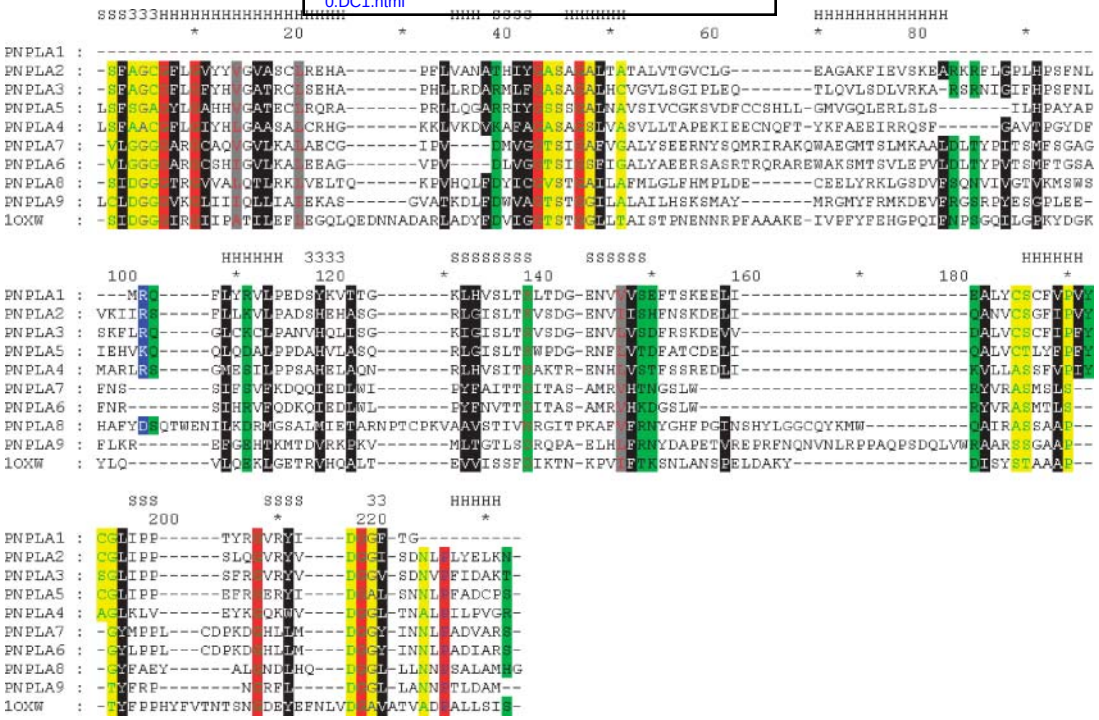

Fig. 1. ClustalW family alignment of the human patatin-like domains. Low-resolution homology models and secondary structure predictions were combined to generate this structurally "guided" protein alignment of human patatin-like domains and the plant patatin isozyme, plant patatin (MSD entry 1OXW). The letter H above the alignment indicates regions folded as α-helix, the letter S indicates regions folded as β-sheet, and the numeral 3 indicates regions folded as 3_{10}-helix. The alignment is shaded to represent the physiochemical properties of the predicted conserved residues as follows: black indicates a hydrophobic residue; yellow indicates a small residue; red with green letters indicates a glycine; red with blue letters indicates a proline; gray indicates an aliphatic residue; green indicates a polar residue; and blue indicates a charged residue. Residues that form the predicted catalytic dyad are positioned in columns 45 and 217. PNPLA, patatin-like phospholipase.

derived from only 72 residue positions. To increase confidence in the analysis, the partial domain sequence represented by PNPLA1 was removed. The resulting phylogenetic tree was constructed from 144 residue positions and did not alter the previous relationship (i.e., when PNPLA1 was included in the analysis) of the remaining sequences.

Low-resolution homology models of each of the human PNPLAs were generated using Macromolecular Structure Database (MSD) entry 1OXW (i.e., a crystal structure derived from a plant patatin isozyme) as a template. These were used to give an approximation of the three-dimensional organization of conserved residues. A representative model of the human PNPLA3 (ADPN) is shown in **Fig. 3** and illustrates both the conserved α/β, patatin-like fold and the spatial orientation of the residues that compose the catalytic dyad (model coordinates are available as supplementary data). When each model was examined in conjunction with the sequence alignment, it was apparent that the catalytic residues were "framed" by several conserved hydrophobic residues (e.g., residues 8, 49, 109, 189, and 195 of the alignment in Fig. 2). When structural data were combined with the observed evolutionary relationships (Fig. 2), other positions were predicted to determine subfamily specificity (e.g., a conserved leucine residue at position 9 is specific to the ADPN subfamily, whereas an isoleucine residue at position 46 appears to be a characteristic of the NTE-like subfamily). Such observations will be the input to future mutagenesis studies that will fully characterize both enzyme function and substrate recognition.

The collective specificity to the unique patatin enzyme suggested that these proteins represented a human patatin-like family. Given the numerous aliases used to describe the various family members, we derived the following consistent nomenclature: PNPLA1, PNPLA2 (TTS2.2), PNPLA3 (ADPN), PNPLA4 (GS2), PNPLA5 (GS2-like), PNPLA6 (NTE), PNPLA7 (NTE-like), PNPLA8 (iPLA2γ), PNPLA9 (PLA2G6), and PNPLA10P (pseudogene). This proposed nomenclature was submitted to the Human Genome Organization (HUGO) Gene Nomenclature Committee for approval.

All protein sequences were further characterized using a variety of in silico prediction algorithms (Table 1). Briefly, PNPLA6 (NTE), PNPLA7 (NTE-like), and PNPLA2 (TTS-2.2) are all predicted to encode single transmembrane domains. PNPLA4 is predicted to be a secreted protein, which along with the PNPLA1, PNPLA8 (iPLA2γ), PNPLA6, and PNPLA7 protein sequences all encode one or more tyrosine kinase phosphorylation sites. Both the PNPLA6 and PNPLA7 sequences encode regions of cyclic nucleotide binding sites, whereas PNPLA9 is characterized by a region of ankyrin repeats.

The expression profile of each gene was verified using TaqMan analysis of a representative sample set of human tissues (see supplementary data for all expression data).

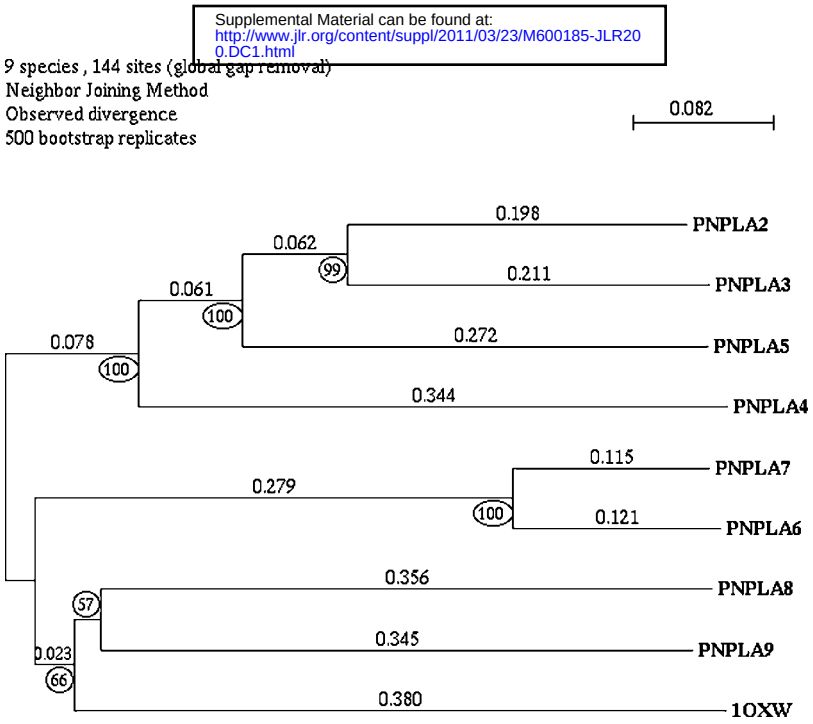

Fig. 2. Phylogenetic relationships of the human patatin-like domains. Residues encoding the patatin-like domains of each of the human phospholipases and the plant patatin isozyme Pat17 were determined using the PFAM prediction tool, and the relevant regions of sequence were aligned as described. The neighbor-joining method was used to predict putative evolutionary relationships using both bootstrap values (circled) and genetic distances.

This indicated that PNPLA3 (ADPN) was expressed at very low levels in a number of tissues with marginally higher copy numbers detected in liver, bone, and macrophages (**Fig. 4**). Equivalent data extracted from the GeneLogic Genesis Enterprise database (probe set 233030_at) supported the observation of increased PNPLA3 expression in liver tissue (data not shown) and contradict previous descriptions of this lipase as an adipocyte-specific gene (11), although this discrepancy may reflect variation between the rat and human species.

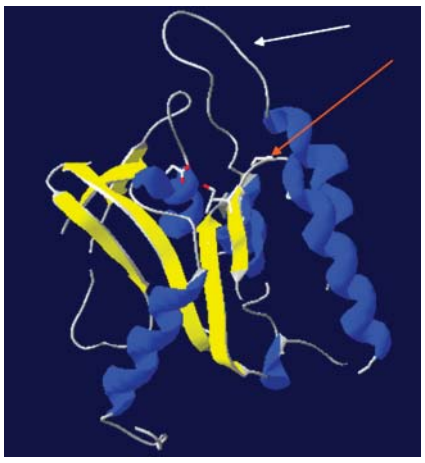

Fig. 3. Low-resolution homology model of the human PNPLA3 patatin-like domain. Partial low-resolution homology models of each of the human patatin-like domains were generated using the DeepView/Swiss-PDB Viewer (see supplementary data for model coordinates). Regions predicted to fold as β-sheets and α-helices are shaded yellow and blue, respectively. Side chains of the catalytic aspartate and serine residues are rendered in stick format and the atoms colored using standard Corey, Pauling & Kultun (CPK). The lycine-rich region is indicated by the red arrow and the putative "lid" region by the white arrow.

PNPLA1 was also observed to be expressed at very low levels in a number of tissues, notably in the digestive system (see supplementary data). There are no representative Affymetrix probe sets for this gene, so the observed TaqMan data could not be verified using the GeneLogic databases in this instance.

Moderate copy numbers of PNPLA5 were detected predominantly in brain tissue, although very low copy numbers were detected in liver tissue (Fig. 4). PNPLA5 was not observed (data not shown) to be upregulated during adipocyte differentiation; however, probe set 233597_at is of low quality (data not shown). Lake et al. (13) observed that PNPLA5 was strongly induced in the liver of *ob/ob* mice and proposed a role in hepatic lipogenesis. The significance of this observation is not clear, as PNPLA5 was not expressed in the liver tissue of C57BI/6J mice. Of the 23 human ESTs representing the human PNPLA5 transcript in the EMBL (23) nucleotide database (release 84), 7 had been isolated from brain tissue, 7 from assorted skin tissue samples, and the remaining constructs from miscellaneous tissues.

TaqMan data indicated that the sequence XM_066899 was not expressed in any of the tissues tested. Combined with the absence of supporting EST data and the fact that the predicted transcript does not encode a catalytic aspartate residue (alignment not shown), it was concluded

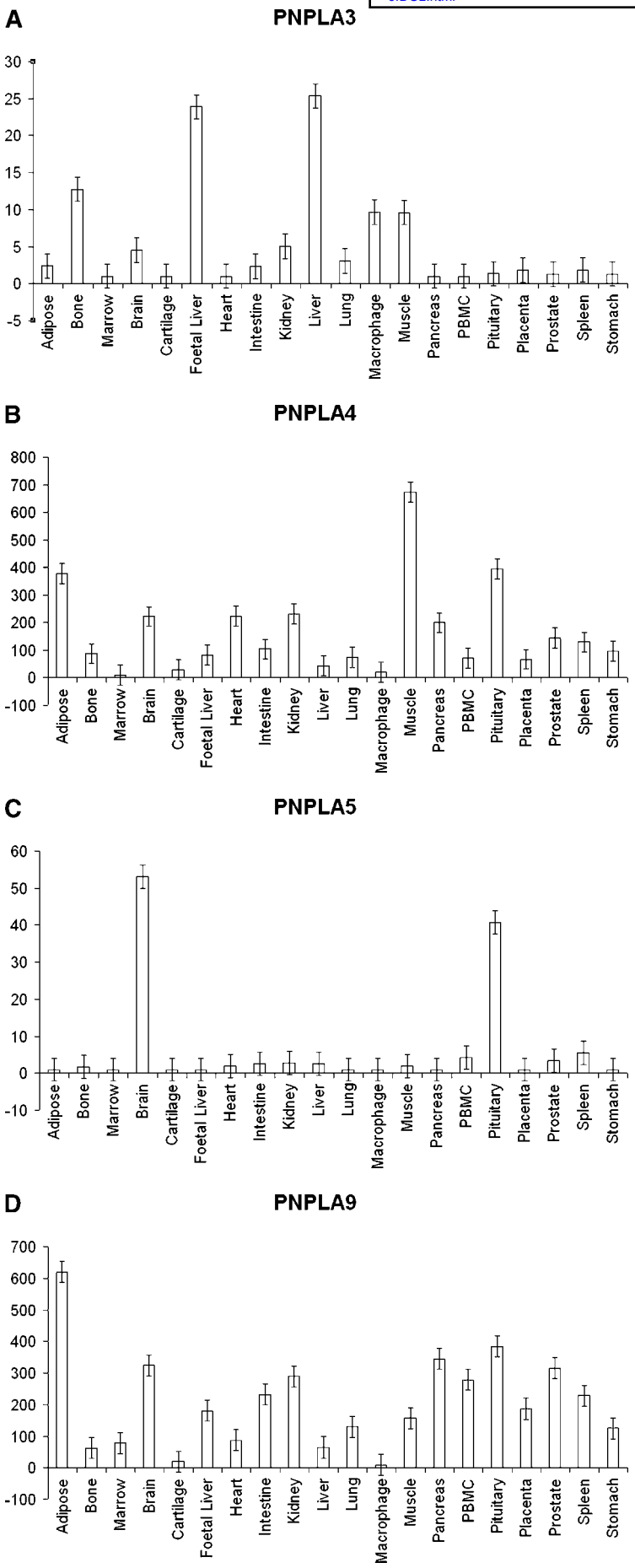

Fig. 4. Expression profiles of the human patatin-like family. Gene-specific TaqMan primers were used to determine the expression profile of each of the human PNPLAs. Equivalent RNA was extracted from the tissues of four unrelated donors as described in Materials and Methods. Abundance values for each tissue sample were calculated using a genomic DNA standard curve to convert TaqMan fluorescent output to the number of sequence copies. The average normalized copy number [i.e., copies of mRNA detected/ng poly(A+) RNA] is represented by the y axis of each graph. Error bars represent the observed SD. All expression data values and individual graphs can be obtained from the supplementary data.

that this sequence likely represents a pseudogene. The predicted peptide sequence is most similar to the PNPLA8 (iPLA2γ) sequence (data not shown). The remaining genes were expressed in most tissues tested. In particular, PNPLA9 (PLA2G6) was predominantly expressed in adipose and brain tissue, PNPLA4 was predominantly expressed in muscle and adipose (Fig. 4), PNPLA7 (NTE-like) was predominantly expressed in all prostrate, pancreas, and adipose tissues tested, and PNPLA8 (iPLA2γ) was moderately expressed in all tissues tested. All experimental values are available as supplementary data.

In an effort to determine a possible role in adipocyte metabolism, the expression profile of each patatin-like gene was monitored in differentiating human adipocyte tissue. Available data concurred with previous reports (12, 13) of significant upregulation of both PNPLA3 (ADPN) and PNPLA2 (TTS-2.2) during the differentiation process. The upregulation for both genes was observed from day 2 of the differentiation process and continued to increase until day 14, when adipocyte differentiation was complete (25). PNPLA8 (iPLA2γ) was observed to be upregulated on day 2 (P = 7.4E-07) and then downregulated by day 7 (**Fig. 5**). PNPLA6 (NTE) was marginally downregulated from day 2 (P = 6.0E-03), whereas PNPLA9 (PLA2G6) was upregulated on day 14 (P = 5.9E-03) of the differentiation process (Fig. 5). No differential expression was observed for PNPLA5 (GS2-like), as reported previously (13). This may reflect species differences, although the authors describe the use of very high RT-PCR cycle numbers to detect gene expression (which could not be detected by Northern blotting). All experimental values are available as supplementary data.

DISCUSSION

Given previous reports detailing a critical role for the PNPLAs in energy mobilization and adipose storage, it was considered important to identify and further characterize additional family members to further understand the physiological role of this lipase family.

The combined analyses identified nine genes and one likely pseudogene that encode a patatin-like domain in the human genome. However, using conventional sequence alignment algorithms, the family demonstrates negligible sequence homology outside of the patatin-like domain. That the nine expressed genes all belong to a single superfamily is supported by the P values reported by the PFAM search tool combined with the predicted z scores obtained from the FUGUE fold prediction application (Table 1). The latter predicts that all nine of the expressed genes should be categorized as a FabD/lysophospholipase-like fold (SCOP database entry 52150; http://scop.mrc-lmb.cam.ac.uk/scop/). This protein fold incorporates both the dual glycine-rich and G-X-S-X-G characteristic lipase motifs and a catalytic aspartate, all within close proximity in the three-dimensional structure. Homology modeling predicts that all of these features are spatially conserved in all of the human PNPLAs.

Sequences were manually aligned using structural prediction data. Comparison of aligned residues indicated that several residues, in addition to those known to be critical for enzyme activity, are conserved in all of the proposed family members (Fig. 1). For example, the serine/threonine residue at position 186 in the structurally guided alignment (Fig. 1) is situated close to (4–6 Å) the catalytic serine (measurement derived from MSD entry 1OXW). Given the proximity and conservation of this position, it is likely that this residue has a role in determining the activity of the PNPLAs.

Phylogenetic analysis of the patatin-like (domain) sequence indicates that five of the genes cluster together in what has been described as the ADPN family (13). Of the remaining four sequences, PNPLA6 (NTE) and PNPLA7 (FLJ43070) cluster together and PNPLA8 (iPLA2γ) and PNPLA9 (PLA2G6) separate out as two unique family members. Given the comparative levels of sequence conservation, it seems probable that the ADPN and NTE subfamilies are the result of distinct gene duplication events. In spite of the conservation of critical catalytic residues, the observed sequence divergence suggests that family members will not display a similar substrate specificity profile.

Expression data indicate that several of these genes are expressed at relatively low levels in a wide range of human tissues. However, the significant upregulation of PNPLA3 (ADPN), PNPLA2 (TTS-2.2), and PNPLA8 (iPLA2γ) observed during adipocyte differentiation (Fig. 5), combined with reports of upregulation in response to fasting (11) and feeding (13), suggests that these three genes respond rapidly to a variety of physiological stimuli that are indicative of their role in facilitating both energy mobilization and lipid storage in adipocyte tissue (12). In contrast to a previous study (10), it was observed that human PNPLA2 gene expression continued to increase until day 14 rather than day 6, as observed previously. If validated, this would highlight a functional difference between species that may have implications for selecting suitable animal models to further evaluate the physiological role of this enzyme in lipid metabolism.

It is tempting to speculate that an increase in transcription similar to that observed for other members of the ADPN subfamily will also be observed for the PNPLA1 transcript when the functional role of this protein is elucidated and an appropriate stimulus is applied. The various analyses completed indicate that the human PNPLA1 protein shares most homology with the ADPN-like subfamily, but available sequences do not encode a G-X-S-X-G motif and as such are unlikely to act as a hydrolase. Predictions based on the human genomic sequence indicate that the public domain transcripts of the human PNPLA1 gene represent a 5′ truncated version of the gene and that the full-length transcript would encode the characteristic lipase motifs. The predicted full-length human transcript can only be validated by gene cloning. The C-terminal sequence is proline-rich but encodes no additional significant motifs. What functions this protein confers are currently unknown.

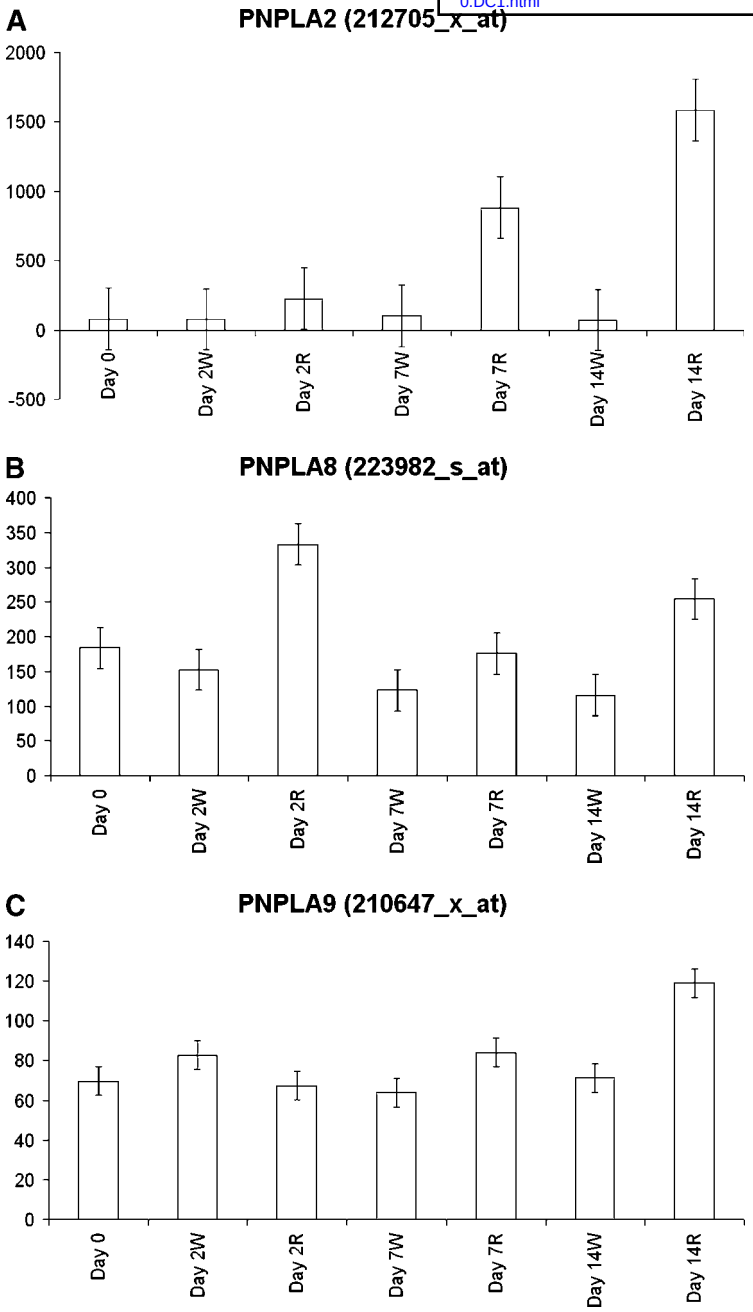

Fig. 5. Expression of PNPLAs in a differentiating human adipocyte cell line. Triplicate RNA samples were purified from both control and differentiating human adipocyte tissues on days 0, 2, 7, and 14. Equivalent samples were harvested from both treated [rosiglitazone (R)] and untreated [water (W)] cultures (see Materials and Methods for further details). Expression data were determined using Affymetrix U133 chips and analyzed using the Resolver™ ANOVA algorithm with standard methodology. The average intensity value ($P \leq 0.05$) at each time point is represented by the y axis. Error bars represent the observed SD. All expression data values and individual graphs used to generate the graphic summaries can be obtained from the supplementary data.

The observed upregulation of the PNPLA8 (iPLA2γ) gene in differentiating adipocytes is a "novel" observation that will require further investigation. A discrete isoform of this membrane-associated phospholipase has been observed in the mitochondrial compartment of the rat myocardium and appeared to facilitate integrating respiration with mitochondrial thermogenesis (26).

In contrast, the PNPLA7 (NTE-like), PNPLA9 (PLA2G6), and PNPLA4 transcripts are relatively highly expressed in a number of human tissues. PNPLA4 is predominantly expressed in both muscle and adipose tissue (Fig. 4), and overexpression of the gene has been associated with the decrease of intracellular triglyceride levels (12). Knockdown assays will be required to confirm a precise role in cellular/lipid metabolism. Note, however, that a murine PNPLA4 ortholog has not yet been confirmed.

Previous publications report that PNPLA6 is expressed in blood lymphocytes (27) and avidly hydrolyzes a number of lysophospholipids, indicating a role in intracellular membrane trafficking (28, 29). The human PNPLA6 and PNPLA7 proteins share ~60% sequence identity and are highly conserved in mouse, rat, and dog (>90% and >80% sequence identity, respectively; alignments not shown). This high level of conservation is often indicative of a conserved function. In addition, both are integral membrane proteins and encode cyclic nucleotide binding sites, indicating that they may be regulated by either cAMP or cGMP.

PNPLA9 (PLA2G6) is a relatively well-characterized protein that is activated during apoptosis (30, 31) and reported to hydrolyze platelet-activating factor at the *sn*-2 fatty acid position (32). These observations indicate a key

role in mediating aspects of the inflammatory response. Protein function is further complicated by extensive alternative splicing (33) and the presence of ankyrin repeats (indicative of mediating protein-protein interactions), which are required for enzyme activity (32).

Partial, low-resolution models of each of the human PNPLAs predict that the central core of each protein is composed of a β-sheet sandwiched by α-helices (α/β/α) in approximately three layers and that the catalytic serine is situated in a nucleophilic elbow after a β-sheet and preceding an α-helix. There appears to be spatial conservation of several hydrophobic residues surrounding the conserved serine nucleophile that may aid substrate binding, although precise coordinates are unavailable given the minimal sequence conservation between the target and template sequences. The significant divergence across the predicted loop regions is readily apparent, and it is tempting to speculate that it is these regions that play a key role in determining the unique characteristics of each family member.

The plant patatin phospholipase is reported not to exhibit interfacial activation (1); however, it is apparent from the patatin (MSD entry 1OXW; http://www.ebi.ac.uk/msd/) structure that a loop (residues represented by positions 82–95 in the alignment shown in Fig. 1) appears to stand "proud" of the main protein structure (Fig. 3). The predicted topology indicates that this loop could theoretically reposition over the catalytic residues (i.e., as a lid), although the reported B-factor values derived from the crystal structure are relatively high for this stretch of residues, indicating minimal flexibility. The authors report that the crystal structure has three molecules in the asymmetric unit. When a number of asymmetric units are assembled according to the symmetry of the crystal, the loops are observed to form crystal contacts with symmetry-related molecules. It is possible that these may provide some stability to this region when packed in the crystal structure. Interfacial activation studies have not yet been reported for any of the human PNPLAs, but PNPLA9 (PLA2G6) has been observed to specify a preference for 1,2-dipalmitoyl phosphatidic acid when the substrate is presented in vesicles (32). PNPLA9 has been observed in a large multimeric complex, so it is not clear whether the observed specificity is attributable to interfacial activation or another mechanism involving one or more of the multimer components. It will be of interest to further investigate the residues in this loop region to determine either a role in substrate recognition or the regulation of enzyme activity.

As with most preliminary efforts to characterize novel protein families, a variety of aliases are used to describe the same entity. This can result in confusion and hinder the detection of shared functional characteristics. Given the conservation of a unique patatin-like lipase motif and the predicted characteristic patatin-like fold found in each of the described proteins, we recommend that the family be henceforth described using the prefix PNPLA, as described above. This proposal has been forwarded to the HUGO Gene Nomenclature Committee, and the proposed names PNPLA3, PNPLA7, PNPLA8, and PNPLA10P have been accepted by the committee (PNPLA1, PNPLA2, PNPLA4, and PNPLA5 had previously been agreed to). Both the PNPLA6 and PNPLA9 proposals are under discussion.

In summary, described are pertinent details and selected expression data that partially characterize the human patatin-like family of phospholipases. The proteins described appear to represent a unique lipolytic family. We recommend further study to determine both the precise, normal physiological role and the putative disease association of each of the described proteins.

The authors thank Pam Thomas for useful comments regarding the implications of crystal contacts within a protein structure and Michael Jaye, Laurie Scott, and Paul Murdoch for their contributions in preparing the manuscript.

REFERENCES

1. Rydel, T. J., J. M. Williams, E. Krieger, F. Moshiri, W. C. Stallings, S. M. Brown, J. C. Pershing, J. P. Purcell, and M. F. Alibhai. 2003. The crystal structure, mutagenesis, and activity studies reveal that patatin is a lipid acyl hydrolase with a Ser-Asp catalytic dyad. *Biochemistry.* **42:** 6696–6708.
2. Holk, A., S. Rietz, M. Zahn, H. Quader, and G. F. E. Scherer. 2002. Molecular identification of cytosolic, patatin-related phospholipases A from Arabidopsis with potential functions in plant signal transduction. *Plant Physiol.* **130:** 90–101.
3. Strickland, J. A., G. L. Orr, and T. A. Walsh. 1995. Inhibition of Diabrotica larval growth by patatin, the lipid acyl hydrolase from potato tubers. *Plant Physiol.* **109:** 667–674.
4. Bateman, A., L. Coin, R. Durbin, R. D. Finn, V. Hollich, S. Griffiths-Jones, A. Khanna, M. Marshall, S. Moxon, E. L. L. Sonnhammer, et al. 2004. The PFAM protein families database. *Nucleic Acids Res.* **32:** D138–D141.
5. Allewelt, M., F. T. Coleman, M. Grout, G. P. Priebe, and G. B. Pier. 2000. Acquisition of expression of the *Pseudomonas aeruginosa* ExoU cytotoxin leads to increased bacterial virulence in a murine model of acute pneumonia and systemic spread. *Infect. Immun.* **68:** 3998–4004.
6. La Camera, S., P. Geoffroy, H. Samaha, A. Ndiaye, G. Rahim, M. Legrand, and T. Heitz. 2005. A pathogen-inducible patatin-like lipid acyl hydrolase facilitates fungal and bacterial host colonization in Arabidopsis. *Plant J.* **44:** 810–825.
7. Kurat, C. F., K. Natter, J. Petschnigg, H. Wolinski, K. Scheuringer, H. Scholz, R. Zimmermann, R. Leber, R. Zechner, and S. D. Kohlwein. 2006. Obese yeast: triglyceride lipolysis is functionally conserved from mammals to yeast. *J. Biol. Chem.* **281:** 491–500.
8. Shohdy, N., J. A. Efe, S. D. Emr, and H. A. Shuman. 2005. Pathogen effector protein screening in yeast identifies Legionella factors that interfere with membrane trafficking. *Proc. Natl. Acad. Sci. USA.* **102:** 4866–4871.
9. Banerji, S., and A. Flieger. 2004. Patatin-like proteins: a new family of lipolytic enzymes present in bacteria? *Microbiology.* **150:** 522–525.
10. Zimmermann, R., J. G. Strauss, G. Haemmerle, G. Schoiswohl, R. Birner-Gruenberger, M. Riederer, A. Lass, G. Neuberger, F. Eisenhaber, A. Hermetter, et al. 2004. Fat mobilization in adipose tissue is promoted by adipose triglyceride lipase. *Science.* **306:** 1383–1386.
11. Villena, J. A., S. Roy, E. Sarkadi-Nagy, K. H. Kim, and H. S. Sul. 2004. Desnutrin, an adipocyte gene encoding a novel patatin domain-containing protein, is induced by fasting and glucocorticoids: ectopic expression of desnutrin increases triglyceride hydrolysis. *J. Biol. Chem.* **279:** 47066–47075.
12. Jenkins, C. M., D. J. Mancuso, W. Yan, H. F. Sims, B. Gibson, and R. W. Gross. 2004. Identification, cloning, expression, and purification of three novel human calcium-independent phospholipase A2 family members possessing triacylglycerol lipase and acylglycerol transacylase activities. *J. Biol. Chem.* **279:** 48968–48975.
13. Lake, A. C., Y. Sun, J. Li, J. Kim, J. W. Johnson, D. Li, T. Revett, H. H. Shih, W. Liu, J. E. Paulsen, et al. 2005. Expression, regulation, and triglyceride hydrolase activity of adiponutrin family members. *J. Lipid Res.* **46:** 2477–2487.

14. Koretke, K. K., R. B. Russell, and A. N. Lupas. 2002. Folds without a fold. *Protein Sci.* **11:** 1575–1579.
15. Schultz, J., F. Milpetz, P. Bork, and C. P. Ponting. 1998. SMART, a simple modular architecture research tool: identification of signalling domains. *Proc. Natl. Acad. Sci. USA.* **95:** 5857–5864.
16. Cuff, J. A., and G. J. Barton. 1999. Application of enhanced multiple sequence alignment profiles to improve protein secondary structure prediction. *Proteins.* **40:** 502–511.
17. Shi, J., T. L. Blundell, and K. Mizuguchi. 2001. FUGUE: sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. *J. Mol. Biol.* **310:** 243–257.
18. Guex, N., and M. C. Peitsch. 1997. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis.* **18:** 2714–2723.
19. Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22:** 4673–4680.
20. Galtier, N., M. Gouy, and C. Gautier. 1996. SeaView and Phylo_win, two graphic tools for sequence alignment and molecular phylogeny. *Comput. Appl. Biosci.* **12:** 543–548.
21. Sonnhammer, E. L., G. von Heijne, and A. Krogh. 1998. A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **6:** 175–182.
22. Bendtsen, J. D., H. Nielsen, G. von Heijne, and S. Brunak. 2004. Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* **340:** 783–795.
23. Kanz, C., P. Aldebert, N. Althorpe, W. Baker, A. Baldwin, K. Bates, P. Browne, A. van den Broek, M. Castro, G. Cochrane, et al. 2005. The EMBL nucleotide sequence database. *Nucleic Acids Res.* **33:** D29–D33.
24. Sarau, H. M., R. S. Ames, J. Chambers, C. Ellis, N. Elshourbagy, J. J. Foley, D. B. Schmidt, R. M. Muccitelli, O. Jenkins, P. R. Murdock, et al. 1999. Identification, molecular cloning, expression, and characterization of a cysteinyl leukotriene receptor. *Mol. Pharmacol.* **56:** 657–663.
25. Rodriguez, A., C. Elabd, F. Delteil, J. Astier, C. Vernochet, P. Saint-Marc, J. Guesnet, A. Guezennec, E. Amri, C. Dani, et al. 2004. Adipocyte differentiation of multipotent cells established from human adipose tissue. *Biochem. Biophys. Res. Commun.* **315:** 255–263.
26. Mancuso, D. J., C. M. Jenkins, H. F. Sims, J. M. Cohen, J. Yang, and R. W. Gross. 2004. Complex transcriptional and translational regulation of iPLA2γ resulting in multiple gene products containing dual competing sites for mitochondrial or peroxisomal localization. *Eur. J. Biochem.* **271:** 4709–4724.
27. Bertoncin, D., A. Russolo, S. Caroldi, and M. Lotti. 1985. Neuropathy target esterase in human lymphocytes. *Arch. Environ. Health.* **40:** 139–144.
28. van Tienhoven, M., J. Atkins, Y. Li, and P. Glynn. 2002. Human neuropathy target esterase catalyzes hydrolysis of membrane lipids. *J. Biol. Chem.* **277:** 20942–20948.
29. Zaccheo, O., D. Dinsdale, P. A. Meacock, and P. Glynn. 2004. Neuropathy target esterase and its yeast homologue degrade phosphatidylcholine to glycerophosphocholine in living cells. *J. Biol. Chem.* **279:** 24024–24033.
30. Kim, S. J., D. Gershov, X. Ma, N. Brot, and K. B. Elkon. 2002. I-PLA(2) activation during apoptosis promotes the exposure of membrane lysophosphatidylcholine leading to binding by natural immunoglobulin M antibodies and complement activation. *J. Exp. Med.* **196:** 655–665.
31. Perez, R., R. Melero, M. A. Balboa, and J. Balsinde. 2004. Role of group VIA calcium-independent phospholipase A2 in arachidonic acid release, phospholipid fatty acid incorporation, and apoptosis in U937 cells responding to hydrogen peroxide. *J. Biol. Chem.* **279:** 40385–40391.
32. Tang, J., R. W. Kriz, N. Wolfman, M. Shaffer, J. Seehra, and S. S. Jones. 1997. A novel cytosolic calcium-independent phospholipase A2 contains eight ankyrin motifs. *J. Biol. Chem.* **272:** 8567–8575.
33. Larsson, P. K., H. E. Claesson, and B. P. Kennedy. 1998. Multiple splice variants of the human calcium-independent phospholipase A2 and their effect on enzyme activity. *J. Biol. Chem.* **273:** 207–214.