Vertebrate patatin-like phospholipase domain-containing protein 4 (PNPLA4) genes and proteins: a gene with a role in retinol metabolism

Roger S. Holmes

Abstract At least eight families of mammalian patatin-like phospholipase domain-containing proteins (PNPLA) (E.C. 3.1.1.3) catalyse the hydrolysis of triglycerides, including PNPLA4 (alternatively PLPL4 or GS2), which also acts as a retinol transacylase and participates in retinol ester metabolism in the body. Bioinformatic methods were used to predict the amino acid sequences, secondary and tertiary structures and gene locations for PNPLA4 genes and encoded proteins using data from several vertebrate genome projects. PNPLA4 genes were located on the X-chromosome for the eutherian mammalian genomes examined. Opossum (marsupial), chicken, anole lizard, clawed toad, zebrafish and lancelet PNPLA4 genes were also identified. Most vertebrate PNPLA4 genes typically contained six coding exons whereas the lancelet PNPLA4 gene contained five coding exons. PNPLA4 subunits were the smallest among the PNPLA-like proteins examined containing 252–255 residues, shared >64% sequence identities and key amino acid residues and predicted motifs, including ‘patatin’ (residues 6–176); putative catalytic dyad active site residues, Ser43 and Asp163; oxy-anion ‘hole’ residues (10–15); and conserved serine residues, which may perform structural roles for this enzyme. Predicted tertiary structures for PNPLA4 ‘patatin’ were similar to those reported for potato ‘patatin’, suggesting that it is strongly conserved during evolution. Human PNPLA4 contained a CpG49 island within the gene promoter, a miRNA-186 binding site within the mRNA 3’-noncoding region for the PNPLA4b isoform and exhibited wide tissue expression at a higher than average level. These and previous studies of vertebrate PNPLA-like gene families have suggested that PNPLA4 is an ancient gene in evolution which has resulted from a duplication of an ancestral invertebrate ATGL-like gene (encoding adipose triglyceride lipase).

Keywords Patatin-like phospholipase domain containing proteins · Vertebrate PNPLA4 · PNPLA4 · X-chromosome · Comparative studies

Introduction

At least eight mammalian patatin-like phospholipase domain-containing proteins (PNPLA-like) (E.C. 3.1.1.3) and genes have been reported which encode patatin-motif containing lipases (Wilson et al. 2006; Kienesberger et al. 2009; Holmes 2012). Human PNPLA4 (also designated as PLPL4 or GS2) is localized on the X-chromosome at Xp22.3, between the genes for steroid sulfatase (STS) and Kallman syndrome (KAL) (Lee et al. 1994). Other human PNPLA-like genes are separately localized on the human genome, including PNPLA1 (chromosome 6) (Mungall et al. 2003), ATGL (also called PNPLA2 or adipose triglyceride lipase) (chromosome 11), PNPLA3 (PLPL3) and PNPLA5 (PLPL5) (chromosome 22) (Schoenborn et al. 2006), PNPLA6 (PLPL6 or NTE) (neuropathy target esterase) (chromosome 19) (Lush et al. 1998; Grimwood et al. 2004), PNPLA7 (PLPL7) (chromosome 9) (Humphray et al. 2009) and PNPLA8 (PLPL8 or IPLA2G) (calcium-independent phospholipase A2-gamma) (chromosome 7) (Tanaka et al. 2000; Mancuso et al. 2000; Scherer et al. 2005).
PNPLA4 catalyses the hydrolysis of triglycerides and participates in retinol-ester metabolism in the body, with a specific role reported for this enzyme in the epidermis in regulating access to retinol from retinol-ester storage depots (Kienesberger et al. 2009; Gao and Simon 2005; Gao et al. 2009). Retinol and related retinoid compounds play key roles in the body including supporting vision (Palczewski 2011), regulating epithelial cell growth and differentiation (Long et al. 2010), contributing to the growth of bone tissue (Oki et al. 2008), immune function (Pino-Lagos et al. 2010) and the activation of tumor suppressor genes (Ye et al. 2009). This retinol-ester metabolic role is in contrast to functions reported for other PNPLA-like enzymes including ATGL (or adipose triglyceride lipase) in triglyceride hydrolysis in adipocyte and non-adipocyte lipid droplets (Zimmermann et al. 2004; Haemmerle et al. 2011); PNPLA3 in contributing to hepatic fat metabolism and non-alcoholic fatty liver disease (Romeo et al. 2008); PNPLA6 (or neuropathy target esterase) which contributes to membrane lipid homeostasis and assists in maintaining axonal integrity (Zaccheo et al. 2004; Rainier et al. 2008); and PNPLA8 which serves as a calcium-independent phospholipase A2 and catalyzes the hydrolysis of membrane phospholipids (Tanaka et al. 2000; Mancuso et al. 2000).

PNPLA4 and other members of the PNPLA-like enzymes belong to the patatin family of acyl hydrolases whose proteins are characterized by a conserved amino acid sequence of Gly-X-Ser-Gly at their active sites, a Ser-Asp catalytic dyad (Ser43/Asp163 for human PNPLA4) (Rydel et al. 2003; Holmes 2012) instead of the Ser-His-Asp/Glu triad reported for other lipases (Cygler and Schrag 1997) and an oxy-anion ‘hole’ providing access to the active site (Rydel et al. 2003). Although three-dimensional structural analyses have not been reported for mammalian PNPLA4, the crystal structure for human PNPLA8 (also IPLA2G or cytosolic phospholipase A2) has been described (Dessen et al. 1999) showing structural similarity to potato patatin (Rydel et al. 2003).

This paper reports the predicted gene structures and amino acid sequences for PNPLA4 genes and proteins, including primate (human [Homo sapiens], chimpanzee [Pan troglodytes], orang-utan [Pongo abelii], rhesus monkey [Rhesus mulatta], marmoset [Callithrix jacchus]), other eutherian mammals (rat [Rattus norvegicus], horse [Equus caballus], cow [Bos taurus], dog [Canis familiaris]), a marsupial mammal (opossum) [Monodelphis domestica] and other vertebrates, including chicken [Gallus gallus], lizard [Anolis carolinensis], frog [Xenopus tropicalis], zebrafish [Danio rerio] and lancelet [Branchiostoma floridae]. Predicted secondary and tertiary structures for PNPLA4 protein subunits are also described, as well as the structural relationships of these genes and enzymes with other PNPLA-like gene families.

Methods

PNPLA4 and other PNPLA-like gene and protein identification

Basic Local Alignment Search Tool (BLAST) studies were undertaken using web tools from the National Center for Biotechnology Information (NCBI) (http://blast.ncbi.nlm.nih.gov/Blast.cgi) (Altschul et al. 1990). Protein BLAST analyses used the human PNPLA4 (Gao and Simon 2005) and PNPLA-like amino acid sequences deduced from reported sequences for these genes (Schoenborn et al. 2006; Dunham et al. 1999; Lush et al. 1998; Grimwood et al. 2004; Humphrey et al. 2009; Tanaka et al. 2000; Mancuso et al. 2000). Non-redundant protein sequence databases for several mammalian and other vertebrate genomes were examined using the blastp algorithm, including human (Homo sapiens) (International Genome Sequencing Consortium 2001); chimpanzee (Pan troglodytes) (Chimpanzee Sequencing and Analysis Consortium 2005); orang-utan (Pongo abelii) (Locke et al. 2011); rhesus monkey (Macaca mulatta) (Gibbs et al. 2007); marmoset (Callithrix jacchus) (http://genome.ucsc.edu/cgi-bin/hgGateway?db=caJac1); horse (Equus caballus) (Wade et al. 2009); cow (Bos taurus) (The Bovine Genome Sequencing and Analysis Consortium et al. 2009); mouse (Mus musculus) (Mouse Genome Sequencing Consortium 2002); rat (Rattus norvegicus) (Rat Genome Sequencing Project Consortium 2004); dog (Canis familiaris) (Lindblad-Toh et al. 2005); opossum (Monodelphis domestica) (Mikkelsen et al. 2007); chicken (Gallus gallus) (International Chicken Genome Sequencing Consortium 2004); lizard (Anolis carolinensis) (Alfoldi et al. 2011); frog (Xenopus tropicalis) (Hellsten et al. 2010); zebrafish (Danio rerio) (Sprague et al. 2005); sea squirt (Ciona intestinalis) (Dehal et al. 2002); and lancelet (Branchiostoma floridae) (Putnam et al. 2008). This procedure produced multiple BLAST ‘hits’ for each of the protein databases which were individually examined and retained in FASTA format, and a record kept of the sequences for predicted encoded PNPLA-like proteins. These records were derived from annotated genomic sequences using the gene prediction method: GNOMON (http://www.ncbi.nlm.nih.gov/genome/guide/gnomon.shtml) and predicted sequences with high similarity scores generated.

BLAT analyses were subsequently undertaken for each of the predicted PNPLA4 and other PNPLA-like amino acid sequences using the UC Santa Cruz web browser (Kent et al. 2003) with the default settings to obtain the predicted locations for each of the vertebrate PNPLA-like genes, including predicted exon boundary locations and gene sizes (Table 1; Supplementary Table 1). Structures for human PNPLA4 isoforms were obtained using the AceView website to examine predicted gene and protein
| Animal Species                  | Gene    | RefSeq: the reference amino acid sequence; predicted Ensembl amino acid sequence; and NCBI reference ID |
|--------------------------------|---------|------------------------------------------------------------------------------------------------------|
| Human                          |         |                                                                                                      |
| Chimpanzee                     |         |                                                                                                      |
| Orangutan                      |         |                                                                                                      |
| Rhesus monkey                  |         |                                                                                                      |
| Guinea pig                     |         |                                                                                                      |
| Horse                          |         |                                                                                                      |
| Dog                            |         |                                                                                                      |
| Rat                            |         |                                                                                                      |
| Cow                            |         |                                                                                                      |
| Panda                          |         |                                                                                                      |
| Elephant                       |         |                                                                                                      |
| Pig                            |         |                                                                                                      |
| Cow                            |         |                                                                                                      |
| Zebrafish                      |         |                                                                                                      |
| Sea squirt                     |         |                                                                                                      |
| Chicken                        |         |                                                                                                      |
| Mouse                          |         |                                                                                                      |
| Rat                            |         |                                                                                                      |
| Cow                            |         |                                                                                                      |
| Human                          |         |                                                                                                      |
| Chimpanzee                     |         |                                                                                                      |
| Orangutan                      |         |                                                                                                      |
| Rhesus monkey                  |         |                                                                                                      |
| Guinea pig                     |         |                                                                                                      |
| Horse                          |         |                                                                                                      |
| Dog                            |         |                                                                                                      |
| Rat                            |         |                                                                                                      |
| Cow                            |         |                                                                                                      |
| Zebrafish                      |         |                                                                                                      |
| Sea squirt                     |         |                                                                                                      |
| Chicken                        |         |                                                                                                      |
| Mouse                          |         |                                                                                                      |
| Rat                            |         |                                                                                                      |
| Cow                            |         |                                                                                                      |
| Human                          |         |                                                                                                      |
| Chimpanzee                     |         |                                                                                                      |
| Orangutan                      |         |                                                                                                      |
| Rhesus monkey                  |         |                                                                                                      |
| Guinea pig                     |         |                                                                                                      |
| Horse                          |         |                                                                                                      |
| Dog                            |         |                                                                                                      |
| Rat                            |         |                                                                                                      |
| Cow                            |         |                                                                                                      |
| Zebrafish                      |         |                                                                                                      |
| Sea squirt                     |         |                                                                                                      |
| Chicken                        |         |                                                                                                      |
| Mouse                          |         |                                                                                                      |
| Rat                            |         |                                                                                                      |
| Cow                            |         |                                                                                                      |
| Human                          |         |                                                                                                      |
| Chimpanzee                     |         |                                                                                                      |
| Orangutan                      |         |                                                                                                      |
| Rhesus monkey                  |         |                                                                                                      |
| Guinea pig                     |         |                                                                                                      |
| Horse                          |         |                                                                                                      |
| Dog                            |         |                                                                                                      |
| Rat                            |         |                                                                                                      |
| Cow                            |         |                                                                                                      |
| Zebrafish                      |         |                                                                                                      |
| Sea squirt                     |         |                                                                                                      |
| Chicken                        |         |                                                                                                      |
| Mouse                          |         |                                                                                                      |
| Rat                            |         |                                                                                                      |
| Cow                            |         |                                                                                                      |

**Table 1**: PNPLA and other PNPLA-like lipase genes and proteins. GenBank IDs are derived from NCBI sources (http://www.ncbi.nlm.nih.gov/Genbank/). UniprotKB/Swiss-Prot IDs available in the Ensembl database are provided for individual PNPLA4, and other PNPLA-like lipase subunits (see http://www.expasy.org). Gene size refers to base pairs of nucleotide sequences; pI refers to theoretical isoelectric points; the number of coding exons are listed; 'na' means data not available.
structures to interrogate this database of human mRNA sequences (Thierry-Mieg and Thierry-Mieg 2006).

Predicted structures and properties of vertebrate PNPLA4 subunits

Alignments of predicted PNPLA4 amino acid sequences were undertaken using a ClustalW method (http://www.ebi.ac.uk/Tools/msa/clustalw2/) (Chenna et al. 2003). Predicted secondary and tertiary structures for vertebrate PNPLA4 subunits were obtained using PSIPRED (McGuffin et al. 2000) and SWISS MODEL web tools, respectively (Guex and Peitsch 1997; Kopp and Schwede 2004). The reported tertiary structure for potato patatin (Rydel et al. 2003) served as the reference for the predicted PNPLA4 tertiary structures, with a modeling range of residues 6–173. Theoretical isoelectric points and molecular weights for vertebrate PNPL4 and PNPLA-like subunits were obtained using Expasy web tools (http://web.expasy.org/compute_pi/) (Gasteiger et al. 2005). Predicted trans-membrane helices for PNPLA-like sequences were obtained using CBS web tools (Center for Biological Sequence Analysis, Technical University of Denmark) (http://www.cbs.dtu.dk/services/TMHMM/) (Moller et al. 2001). Patatin-motifs were identified for PNPLA-like sequences using web tools from the National Center for Biotechnology Information (NCBI) (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

Human PNPLA4 gene expression and predicted gene regulation sites

The human genome browser (http://genome.ucsc.edu) (Kent et al. 2003) was used to examine GNF Expression Atlas 2 data using various expression chips for the human PNPLA4 gene (http://biogps.gnf.org) (Su et al. 2004). Predicted CpG islands and microRNA (miRNA) binding sites for human PNPLA4 were obtained using the UC Santa Cruz Genome Browser (http://genome.ucsc.edu).

Results and discussion

Alignments and biochemical features of PNPLA4 amino acid sequences

Amino acid sequence alignments for 14 previously unreported vertebrate PNPLA4 amino acid sequences are shown in Fig. 1, together with the reported sequence for human PNPLA4 (Gao and Simon 2005; Gao et al. 2009). The PNPLA4 sequences exhibited >60 % identities, suggesting that these protein subunits are products of the same gene family, whereas the sequences for the predicted vertebrate PNPLA1, ATGL, PNPLA3 and PNPLA5
The subunits were 27–37% identical with the PNPLA4 sequences, indicating that these are members of distinct, but related PNPLA-like gene families (Supplementary Table 2). The sequences for the vertebrate PNPLA6, PNPLA7 and PNPLA8 subunits examined were even more distantly related with vertebrate PNPLA4 sequences with identities of 16% (Supplementary Table 2). Two of these sequences (PNPLA6 and PNPLA7), however, showed comparatively high sequence identities (58–61%), suggesting that these are closely related gene families. Amino acid sequences for the eight human PNPLA-like proteins examined contained 253 (PNPLA4), 429–532 (PNPLA1, PNPLA2, PNPLA3 and PNPLA5), 782 (PNPLA8) and 1,317–1,366 (PNPLA6 or NTE and PNPLA7) residues (Table 1; Supplementary Table 1). Consequently, vertebrate PNPLA4 is the smallest among these PNPLA-like proteins with an average molecular weight of ~28,000, while others exhibited MWs which are ~2 (PNPLA1, PNPLA2, PNPLA3 and PNPLA5), ~3 (PNPLA8) or ~5 times larger (PNPLA6 and PNPLA7) than PNPLA4.

Site-directed mutagenesis studies for human PNPLA4 (Gao and Simon 2005; Gao et al. 2009) and potato patatin (Hirschberg et al. 2001) have enabled the identification of key catalytic residues among those aligned for the vertebrate PNPLA4 sequences examined (Fig. 1). These included an active site motif (Gly-Xaa-Ser-Yaa-Gly designated as motif 2) (human PNPLA4 residues 41–45); active site residues Ser43 and Asp163 which serve as the catalytic dyad during catalysis; and a putative oxy-anion hole with a consensus sequence for this motif (Cys-Gly-Phe-Leu-Gly for residues 11–15 designated as motif 1). These residues are conserved among all of the vertebrate PNPLA4 sequences examined (with the exception of a Ala10?Ser10 substitution for opossum PNPLA4), in addition to Thr115 and serine residues; and bold underlined font shows predicted exon junctions.
site-specific phosphorylation (Daub et al. 2008). High theoretical isoelectric points (pI) were observed for each of the vertebrate PNPLA4 subunits examined (pI values range from 8.2–9.3), as compared with the other PNPLA-like subunits examined which exhibited lower predicted pI values, with the exception of the vertebrate PNPLA8 subunits (pI values of 9.2–9.3) (Table 1; Supplementary Table 1).

Predicted secondary and tertiary structures for vertebrate PNPLA4 subunits

Analyses of predicted secondary structures for PNPLA4 sequences revealed similar α-helix and β-sheet structures for all of the vertebrate subunits examined, particularly near key residues or functional domains (Fig. 1). Predicted secondary (Fig. 1) and tertiary structures (Fig. 2) were very similar to those reported for potato patatin (Rydel et al. 2003), which have been retained for all of the vertebrate PNPLA4 sequences examined. The predicted PNPLA4 tertiary structure (Fig. 2) is based on a partial sequence for this enzyme (residues 6–173) revealing the relative positioning and predicted structures for each of 5α-helices and 5β-sheets. These included the N-terminus α-helix (designated as α1), which may serve as a membrane anchor for PNPLA4 (no predicted trans-membrane properties were, however, observed for the α1 helix); an oxy-anion hole proposed for the motif previously reported (Cys-Gly-Phe-Leu-Gly for residues 11–15 designated as motif 1) located near the active site cleft (Fig. 2) which is similar to the oxy-anion hole reported for potato patatin (Rydel et al. 2003) and human PNPLA8 (encoding cytosolic phospholipase A2) (Dessen et al. 1999); a second α-helix (α2) and β-sheet (β2) which contain the active site motif Gly-Xaa-Ser/Yaa-Gly (residues 41–45 for human PNPLA4 designated as motif 2); and a β-sheet (β5) which contains Asp163, the second member of the active site dyad of catalytic residues.

These structures are proximally located within a putative active site cleft supported by the predicted three-dimensional structure for this enzyme, however, any firm conclusions must await further studies. Several conserved serine residues were also observed for the vertebrate PNPLA4 sequences which may correspond to residues previously proposed for performing structural roles in potato patatin phospholipase A (Hirschberg et al. 2001; Rydel et al. 2003).

Predicted gene locations, exonic structures and expression for vertebrate PNPLA4 genes

Table 1 summarizes the predicted locations for vertebrate PNPLA4 genes based on BLAT interrogations of several vertebrate genomes using the sequence for human PNPLA4 (Gao and Simon 2005; Gao et al. 2009) and the predicted sequences for other vertebrate PNPLA4 enzymes and the UC Santa Cruz Web Browser (Kent et al. 2003). Eutherian mammalian PNPLA4 genes were located on the X-chromosome in each case, however, the marsupial PNPLA4 gene (opossum; Monodelphis domestica) was located on an autosome (chromosome 7), suggesting that the X-chromosome location for PNPLA4 is restricted to eutherian mammalian genomes. Table 1 also provides data for other vertebrate PNPLA4 genes, including the previously reported chicken PNPLA4 sequence (Saarela et al. 2008), and those predicted for lizard (Anolis carolensis), frog (Xenopus tropicalis), zebrafish (Danio rerio) and lancelet (Branchiostoma floridae) genomes, which have distinct locations to those reported here for the other vertebrate PNPLA-like genes. Figure 1 summarizes the predicted exonic start sites for several vertebrate PNPLA4 genes with each having six coding exons in identical or similar positions. In contrast, lancelet PNPLA4 contained 5 coding exons, with exon 5 corresponding to exons 5 and 6 for the vertebrate PNPLA4 genes.

Figure 3 examined the predicted location of the human PNPLA4 gene on the human X-chromosome as well as comparative sequence identities for vertebrate PNPLA4 sequences. The absence of a mouse PNPLA4 gene was readily apparent from this study. Moreover, a major decrease in sequence identities for vertebrate PNPLA4 genes with the human PNPLA4 gene was observed for the more distantly related species examined, especially for the intronic sequences and for exons 5 and 6 of chicken, frog and zebrafish PNPLA4 genes. It is suggested that this may reflect a higher level of conservation for the ‘patatin’
encoding regions for the vertebrate PNPLA4 sequences, which are encoded by exons 1–4 of the vertebrate PNPLA4 genes examined (Fig. 1).

Supplementary Table 3 examined the comparative sizes for several vertebrate PNPLA4 genes and intronic sequences (introns 1–5 for vertebrate PNPLA4 genes and introns 1–4 for the lancelet PNPLA4 gene examined). The rat PNPLA4 gene was much smaller than other PNPLA4 genes examined, being >10 times smaller than the human gene, which is reflected in the smaller sizes observed for introns 1, 3, 4 and 5. Moreover, a mouse PNPLA4 gene was not detected in this and previous studies and further investigations are required to demonstrate whether this gene is absent from the mouse genome or has escaped detection at this stage. The guinea pig (Cavia porcellus) PNPLA4 gene, however, resembled other mammalian PNPLA4 genes in the comparative sizes of introns, which suggested that the small size for the rat PNPLA4 gene was not a common feature for other rodent PNPLA4 genes. Comparisons of intron sizes for vertebrate and invertebrate PNPLA4 genes also showed that intron 2 was much smaller for all mammalian (also chicken and lizard) PNPLA4 genes examined than other introns, although intron 2 sequences for frog (Xenopus tropicalis), zebrafish (Danio rerio) and lancelet (Branchiostoma floridae) PNPLA4 genes were much larger than for the mammalian PNPLA4 genes.

Figure 4 illustrates the comparative predicted structures of pre-messenger RNA human PNPLA4 gene transcripts (http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/) (Thierry-Mieg and Thierry-Mieg 2006). There were 6 introns present for the pre-messenger mRNA PNPLA4a and PNPLA4b transcripts, with the latter containing a CpG49 island in the 5’-noncoding segment corresponding to the promoter for this gene. In addition, the PNPLA4b transcript contained an extended 3’-noncoding segment with a predicted miRNA-186 binding site. These predicted gene regulation sites may contribute to the high level of gene expression (×1.5 times the expression of the average gene) and wide tissue expression observed for PNPLA4. Elango and Yi (2011) have previously reported that larger CpG islands are associated with gene promoters of housekeeping genes showing a broad range of gene expression and containing more RNA polymerase II binding sites than other promoters. Moreover, miRNAs are post-transcriptional regulators that bind to complementary sequences on target messenger RNA transcripts (mRNAs), usually resulting in translational repression or target degradation and gene silencing (Bartel 2009). Consequently, the presence of CpG49 and miRNA-186 within the PNPLA4 gene may contribute significantly to the broad tissue expression observed for PNPLA4 transcripts. Figure 5 presents ‘heat maps’ showing the comparative
gene expression for various human tissues obtained from GNF Expression Atlas Data using U133A and GNF1H PNPLA4 chips (Su et al. 2004) with higher levels being observed in bronchial epithelial cells and heart as well as significant expression in the other tissues examined.

Fig. 4 Gene structure and major isoforms for human PNPLA4. From AceView website (Thierry-Mieg and Thierry-Mieg 2006) http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/ mature isoform variants (designated as ‘a’, ‘b’ etc.) are shown for each PNPLA4 transcript; capped 5’- and 3’- ends for the predicted mRNA sequences are identified; a predicted CpG49 island, a miRNA binding site (miR-186) and a scale of base pairs of nucleotide sequences are shown.

Fig. 5 Human tissue gene expression ‘heat maps’ for PNPLA4 taken from the human genome browser (Kent et al. 2003) (http://genome.ucsc.edu); GNF Expression Atlas 2 data using expression chips for human PNPLA4 (http://biogps.gnf.org) (Su et al. 2004); comparative gene expression levels among human tissues: red (high) and black (intermediate) expression levels.

Phylogeny of vertebrate PNPLA4 and other PNPLA-like lipases

A phylogenetic tree has been previously described from alignments of vertebrate ATGL-like amino acid sequences (PNPLA1, ATGL, PNPLA3, PNPLA4 and PNPLA5) with the predicted fruit fly (Drosophila melanogaster) ATGL sequence serving to ‘root’ the tree (Holmes 2012). Clustering was reported for five major groups of vertebrate ATGL-like sequences: PNPLA1; ATGL (or PNPLA2); PNPLA3; PNPLA4; and PNPLA5. Clustering into sub-groupings was also described, including PNPLA3 and PNPLA5, with ATGL; and PNPLA4 with PNPLA1. These results were consistent with the presence of ATGL-like and PNPLA4-like genes within primitive vertebrate genomes examined, and were suggestive of an initial gene duplication event for ATGL generating both of these genes, during the evolutionary appearance of vertebrates. This is consistent with PNPLA4 being an ancient gene, appearing in some primitive vertebrate genomes and being present throughout vertebrate evolution over a period of evolution of >500 million years, which is reported for the timing of the appearance of vertebrates during evolution (Donoghue and Benton 2007).

These phylogenetic studies were also extended to include other PNPLA-like genes and proteins, namely PNPLA6, PNPLA7 and PNPLA8 sequences (Holmes 2012). The
results were indicative of at least three major PNPLA-like sequence groups, including the ATGL-like sequences (PNPLA1, ATGL (PNPLA2), PNPLA3, PNPLA4 and PNPLA5 (Group 1); the PNPLA6 and PNPLA7 sequences (Group 2); and the PNPLA8 sequences (Group 3). Group 1 sequences were further divided according to the designation of ATGL-like gene families, which clustered with the sea squirt ATGL-like sequence, and were suggestive of an ancestral relationship between early vertebrate ATGL and PNPLA4 genes, with other members of PNPLA-like group 1 sequences, which appeared later during vertebrate evolution: PNPLA1 and PNPLA3/PNPLA5. This report (Holmes 2012) also suggested that vertebrate PNPLA6 and PNPLA7 sequences shared a common evolutionary origin distinct to the ATGL-like and PNPLA8 sequences, which were ‘rooted’ with the sea squirt (Ciona intestinalis) PNPLA7 sequence, whereas the vertebrate PNPLA8 sequences were also distinct and separately ‘rooted’ with the sea squirt (Ciona intestinalis) PNPLA8 sequence.

Overall, it is likely that the three PNPLA-like groups of genes and proteins have arisen from distinct ancestral genes during vertebrate evolution, namely ATGL (the proposed ancestral gene for group 1 PNPLA-like genes [PNPLA1; ATGL; PNPLA3/PNPLA5]; and PNPLA4); PNPLA6/PNPLA7 for group 2 PNPLA-like genes; and PNPLA8 for group 3 PNPLA-like genes (Holmes 2012).

**Summary**

The results of this study support previous studies (Wilson et al. 2006; Kienesberger et al. 2009; Saarela et al. 2008; Holmes 2012) for at least eight vertebrate PNPLA-like genes and encoded lipases, including five ATGL-like genes, namely PNPLA4 (encoding PNPLA4) and PNPLA1, ATGL (encoding adipose triglyceride lipase), PNPLA3 and PNPLA5 genes; two PNPLA6-like genes, PNPLA6 (encoding neuropathy target esterase) and PNPLA7; and PNPLA8 (encoding cytosolic phospholipase A2). Vertebrate PNPLA4 sequences shared key conserved sequences reported for human PNPLA4 (Gao and Simon 2005; Wilson et al. 2006; Gao et al. 2009), including active site residues, an oxy-anion ‘hole’ sequence, a phosphorylated Thr site and several conserved serine residues. Gene expression data showed that the human PNPLA4 gene is broadly expressed at higher levels than those for the average gene, for which a CpG island localized in the PNPLA4 promoter and a miRNA binding site localized in the extended 3’noncoding region of PNPLA4b mRNA isoform may contribute to these high expression levels. A recent phylogenetic study (Holmes 2012) has suggested that PNPLA4 is an ancient gene in vertebrate evolution derived from a duplication of an ancestral ATGL-like gene within a primitive vertebrate genome.

**Acknowledgments** The advice of Dr Laura Cox of the Texas Biomedical Research Institute is gratefully acknowledged.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

**References**

Alfoldi J, Di Palma F, Grabherr M, Williams C, Kong L, Maucci E et al (2011) The genome of the green anole lizard and a comparative analysis with birds and mammals. Nature 477:587–591

Altshul SF, Vyas V, Cornfield A, Goodin S, Ravikumar TS, Rubin EH, Gupta E (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Bartel DP (2009) MicroRNAs: target recognition and regulatory functions. Cell 136:215–233

Chenna R, Sugawara R, Koike T, Lopez R, Gibson TJ, Higgins DJ, Thompson JD et al (2003) Multiple sequence alignment with the clustal series of programs. Nucleic Acids Res 31:3497–3500

Chimpanzee Sequencing and Analysis Consortium (2005) Initial sequence of the chimpanzee genome and comparison with the human genome. Nature 437:69–87

Cygler M, Schrag JD (1997) Structure as basis for understanding interfacial properties of lipases. Methods Enzymol 284:3–27

Daub H, Olsen JV, Barlein M, Gnad F, Oppermann FS, Korner R et al (2008) Kinase-selective enrichment enables phosphoproteomics of the kinome across the cell cycle. Mol Cell 31:438–448

Dehal P, Satou Y, Campbell RK, Chapman J, Degnan B, De Tomaso A et al (2002) The draft genome of Ciona intestinalis: insights into chordate and vertebrate origins. Science 298:2157–2167

Dessen A, Tang A, Schmidt H, Stahl H, Clark JD, Seehra J, Somers WS (1999) Crystal structure of human phosphohydrolase A2 reveals a novel topology and catalytic mechanism. Cell 97:349–360

Donoghue PCJ, Benton MJ (2007) Rocks and clocks: calibrating the tree of life using fossils and molecules. Trends Genet 22:424–431

Dunham I, Shimizu N, Roe BA, Chissoe S, Hunt AR, Collins JE et al (1999) The DNA sequence of human chromosome 22. Nature 402:489–495

Elango N, Yi SV (2011) Functional relevance of Cpg island length for regulation of gene expression. Genetics 187:1077–1083

Gao J, Simon M (2005) Identification of a novel keratinocyte retinol ester hydrolase as a transacylase and lipase. J Invest Dermatol 124:1259–1266

Gao JG, Shin A, Gruber R, Schmuth M, Simon M (2009) GS2 as a retinol transacylase and as a catalytic dyad independent regulator of retinylester accretion. Mol Genet Metabol 96:253–260

Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR et al (2005) Protein identification and analysis tools on the ExPASy Server. In: Walker JM (ed) The proteomics protocols handbook. Humana Press, USA

Gibbs RA, Rogers J, Katze MG, Bumgarner R, Weinstock GM, Mardis ER et al (2007) Evolutionary and biomedical insights from the rhesus macaque genome. Science 316:222–234

Grimwood J, Gordon LA, Olsen A, Terry A, Schmutz J, Lamerdin J et al (2004) The DNA sequence of human chromosome 19. Nature 428:528–535

Guex N, Peitsch MC (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modelling. Electrophoresis 18:2714–2723

Haemmerle G, Moustafa T, Wielkart G, Büttrn S, Schmidt A, van de Weijer T et al (2011) ATGL-mediated fat catabolism regulates
