Open Access Bioinformatics

Dovepress
open access to scientific and medical research

Open Access Full Text Article

REVIEW

Comparative studies of adipose triglyceride lipase genes and proteins: an ancient gene in vertebrate evolution

This article was published in the following Dove Press journal:
Open Access Bioinformatics
7 March 2012
Number of times this article has been viewed

Roger S Holmes

School of Biomolecular and Physical Sciences, Griffith University, Nathan, QLD, Australia

Abstract: At least eight families of mammalian patatin-like phospholipase domain-containing proteins (EC 3.1.1.3) catalyze the hydrolysis of triglycerides, including adipose triglyceride lipase (ATGL), which functions in triglyceride lipase metabolism in the body, especially in adipose tissue. Bioinformatic methods were used to predict the amino acid sequences, secondary and tertiary structures, and gene locations for *ATGL* genes and encoded proteins using data from several vertebrate genome projects. *ATGL* genes usually contained nine coding exons for each of the vertebrate genomes examined, whereas the invertebrate sea squirt (*Ciona intestinalis*) *ATGL* gene contained a single exon. Vertebrate ATGL subunits contained 473–504 residues, shared >46% sequence identities, and exhibited sequence alignments and identities for key amino acid residues and predicted motifs: an N-terminal lipid binding region (residues 7–29 for human ATGL); a patatin "motif" (residues 10–179); a putative active site oxyanion "hole" (Cys15, Gly16, and Leu18); and catalytic dyad active site residues (Ser47 and Asp166). Predicted tertiary structures for the ATGL patatin "motif" were similar to those reported for potato patatin, suggesting that this structure is strongly conserved during animal and plant evolution. Human *ATGL* contained a CpG131 island within the gene promoter; miR-124/506 and miR-108 binding sites within the mRNA 3′-noncoding region; several transcription factor binding sites, including PPARA and PPARG, which are key regulators of genes encoding enzymes of lipid metabolism; and exhibited wide tissue expression at a higher than average level (2.2×). Phylogenetic analyses of vertebrate *PNPLA*-like gene families suggest that *ATGL* is an ancient gene in vertebrate evolution which has been derived from an ancestral *ATGL* gene (encoding adipose triglyceride lipase) and undergone successive gene duplication events, forming ancestral genes for vertebrate *PNPLA1*, *ATGL*, *PNPLA3*, *PNPLA4*, and *PNPLA5* gene families.

Keywords: vertebrates, amino acid sequence, adipose triglyceride lipase, patatin-like phospholipase domain-containing proteins, gene duplication, evolution

Introduction

At least eight mammalian patatin-like phospholipase domain-containing (PNPLA-like) proteins (EC 3.1.1.3) and genes have been reported which encode patatin motif-containing lipases.[1] Human *ATGL* (also called *PNPLA2* or *PLPL2*) encodes adipose triglyceride lipase (ATGL, also known as desnutrin or PNPLA2) and is localized on chromosome 11 between the genes encoding ribosomal protein P2 (*RPLP2*) and EF-hand calcium-binding domain-containing protein 4A (*EFCAB4A*).[2] Other human *PNPLA*-like genes are separately localized on the human genome, including *PNPLA1* (or *PLPL1*, chromosome 6);[3] *PNPLA3* (*PLPL3*) and *PNPLA5* (*PLPL5*, chromosome 22);[4]

Correspondence: Roger S Holmes
School of Biomolecular and Physical Sciences, Griffith University, Nathan
Brisbane QLD 4111, Australia
Tel +617 3348 2834
Email r.holmes@griffith.edu.au

PNPLA4 (*PLPL4*, X chromosome);[5] *PNPLA6* (*PLPL6* or *NTE*) (neuropathy target esterase, chromosome 19);[6,7] *PNPLA7* (*PLPL7*, chromosome 9);[8] and *PNPLA8* (*PLPL8* or *IPLA2G*, calcium-independent phospholipase A2-gamma, chromosome 7).[9–11]

ATGL catalyzes the first step in the hydrolysis of triglycerides in adipocytes and lipid droplets in other tissues of the body, and also acts as an acylglycerol transacylase.[12,13] ATGL plays a key role in regulating the size and turnover of adipocytes and lipid droplets in nonadipose tissues in association with hormone-sensitive lipase.[13] Hormone-sensitive lipase serves as the major lipase for catecholamine-stimulated and natriuretic peptide-stimulated lipolysis, whereas ATGL mediates the hydrolysis of triglycerides during basal lipolysis.[14–16] *ATGL* gene expression has been shown to be upregulated during fasting,[17] downregulated by insulin,[18] and is associated with type 2 diabetes[2] and obesity.[19] Human *ATGL* also serves as a likely target for transcriptional activation by peroxisome proliferator-activated receptor gamma, a key adipogenic transcription factor,[18] and is upregulated in skeletal muscle by exercise training.[20] In addition, human *ATGL* gene mutations have been shown to cause a neutral lipid storage condition and severe myopathy.[21–23] Notari et al[24] have also identified ATGL in retinal and nerve cell plasma membranes as an extracellular pigment epithelium-derived factor which may contribute to ligand/receptor signaling function in these cells.

ATGL and other members of the PNPLA-like enzymes belong to the patatin family of acyl hydrolases, the proteins of which are characterized by conserved amino acid sequences for their active sites (human ATGL residues identified) involved in catalysis (Gly45-X46-Ser47-X48-Gly49) at their active sites[25] and a Ser-Asp catalytic dyad (Ser47/Asp166 for human ATGL) instead of the Ser-His-Asp/Glu triad reported for other lipases.[26] Although three-dimensional structural analyses have not been reported for mammalian ATGL, the crystal structure for human PNPLA8 (also IPLA2G or cytosolic phospholipase A2) has been described[27] which shows structural similarities to potato patatin.[25]

This paper reports the predicted gene structures and amino acid sequences for *ATGL* vertebrate genes and proteins, including primate (human [*Homo sapiens*] and chimpanzee [*Pan troglodytes*]), other eutherian mammals (mouse [*Mus musculus*], rat [*Rattus norvegicus*], guinea pig [*Cavia porcellus*], pig [*Sus scrofa*], cow [*Bos taurus*], dog [*Canis familiaris*]), panda (*Ailuropoda melanoleuca*), a marsupial mammal (opossum, *Monodelphis domestica*) and other vertebrates, including chicken (*Gallus gallus*), lizard (*Anolis carolensis*), frog (*Xenopus tropicalis*), and zebrafish (*Danio rerio*), and an invertebrate sea squirt (*Ciona intestinalis*). Predicted secondary and tertiary structures for vertebrate ATGL protein subunits are also described, as well as the structural, phylogenetic, and evolutionary relationships of these genes and enzymes with other *PNPLA*-like gene families.

Materials and methods

Vertebrate *ATGL* and other *PNPLA*-like gene and protein identification

BLAST (Basic Local Alignment Search Tool) studies were undertaken using web tools from the National Center for Biotechnology Information (http://blast.ncbi.nlm.nih.gov/Blast.cgi).[28] Protein BLAST analyses used human[29] and mouse[12] ATGL and PNPLA-like amino acid sequences deduced from reported sequences for these genes.[3–10] Nonredundant protein sequence databases for several mammalian and other vertebrate genomes were examined using the BLASTP algorithm with the default scoring parameters. The following genomes were examined: human (*H. sapiens*),[30] chimpanzee (*P. troglodytes*),[31] horse (*Equus caballus*),[32] cow (*B. Taurus*),[33] mouse (*M. musculus*),[34] rat (*R. norvegicus*),[35] dog (*C. familiaris*),[36] opossum (*M. domestica*),[37] chicken (*G. gallus*),[38] lizard (*A. carolensis*),[39] frog (*X. tropicalis*),[40] zebrafish (*D. rerio*),[41] and sea squirt (*C. intestinalis*).[42] 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 and predicted sequences with high similarity scores were generated (usually $>$ 800 with full sequence coverage). Patatin "motif" amino acid sequences were obtained from the following sources: potato POT1,[25] mouse ATGL,[12] mouse PNPLA1,[34] PNPLA3,[43] rat PNPLA4,[35] and mouse PNPLA5[34] using web tools from the National Center for Biotechnology Information (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

BLAST analyses were subsequently undertaken for each of the predicted ATGL and other PNPLA-like amino acid sequences using the UC Santa Cruz web browser (http://genome.ucsc.edu/cgi-bin/hgBlat)[44] with default settings to obtain the predicted locations for each of the vertebrate *PNPLA*-like genes, including predicted exon boundary locations and gene sizes (see Table 1). Structures for human *ATGL* isoforms were obtained using the AceView website to

Table 1 Vertebrate *ATGL* and other *PNPLA*-like lipase genes and proteins

Vertebrate	Species	PNPLA gene (other name)	Chromosome coordinates	Gene size	Exons strand	Subunit MW	Amino acids	pI	GenBank ID	UNIPROT ID	[1]NCBI reference ID [2]NCBI predicted ID
Human	*Homo sapiens*	*ATGL* (*PNPLA2*)	11:819,719–824,859	5,141	9 +ve	55,316	504	6.7	BC011958	Q96AD5	[1]NM_023376
Chimpanzee	*Pan troglodytes*	*ATGL* (*PNPLA2*)	11:879,389–884,735	5,347	9 +ve	55,429	504	6.9	na	na	[2]XP_003339482.1
Mouse	*Mus musculus*	*Atgl* (*Pnpla2*)	7:148,641,186–148,645,564	4,379	9 +ve	53,657	486	6.1	BC064747	Q8BJ56	[1]NR_028142
Rat	*Rattus norvegicus*	*Atgl* (*Pnpla2*)	1:201,642,058–201,646,343	4,286	9 +ve	52,567	478	6.2	AC109542	P0C548	[1]NM_001108509.2.
Guinea pig	*Cavia porcellus*	*ATGL* (*PNPLA2*)	[3]218:129,068–132,994	3,927	9 –ve	53,964	486	6.3	[5]ENSCPOT5922	na	na
Pig	*Sus scrofa*	*ATGL* (*PNPLA2*)	2:149,082–153,354	4,273	9 –ve	53,147	486	7.5	ABW06598	A5JT15	[1]NM_001098605.1
Cow	*Bos taurus*	*ATGL* (*PNPLA2*)	[4]Un.004.171:137,584–141,740	4,157	9 –ve	53,418	486	6.8	BC112803	Q2KI18	[1]NM_001046005.1
Panda	*Ailuropoda melanoleuca*	*ATGL* (*PNPLA2*)	[3]GL195194.1:5,234–9,734	4,501	9 –ve	52,834	482	8.1	[6]ENSAMET7602	na	na
Opossum	*Monodelphis domestica*	*ATGL* (*PNPLA2*)	[4]Un:45,368,040–45,372,831	4,792	9 +ve	53,547	490	6.8	na	na	[2]XP_001380646.2
Chick	*Gallus gallus*	*ATGL* (*PNPLA2*)	5:16,838,493–16,868,610	30,118	9 –ve	53,610	483	6.7	EU419874	A8WEN5	[1]NM_001113291.1
Lizard	*Anolis carolensis*	*ATGL* (*PNPLA2*)	1:73,687,462–73,733,079	45,618	9 –ve	53,795	483	6.4	na	na	[2]XP_003214822.1
Frog	*Xenopus tropicalis*	*ATGL* (*PNPLA2*)	[3]990:125,824–140,107	14,284	9 +ve	55,162	498	5.9	BC089233	Q5FWR9	[1]NM_001015693.1
Zebrafish	*Danio rerio*	*ATGL* (*PNPLA2*)	[3]3512:30,775–45,374	14,600	10 +ve	52,253	473	6.8	BC075928	na	[1]NM_001002338.1
Human	*Homo sapiens*	*PNPLA1* (*PLPL1*)	6:36,238,237–36,275,490	37,254	8 +ve	57,876	532	8.2	BC103906	Q8N8W4	[1]NM_001145716.1
Mouse	*Mus musculus*	*Pnpla1* (*Plpl1*)	17:28,995,812–29,023,893	28,082	9 +ve	65,171	592	8.6	AK132521	Q3V1D5	[1]NM_001034885.3.
Rat	*Rattus norvegicus*	*Pnpla3* (*Plpl3*)	7:122,152,145–122,171,991	19,847	9 +ve	45,908	414	6.8	EDM15609	na	na
Mouse	*Mus musculus*	*Pnpla3* (*Plpl3*)	15:83,998,304–84,016,512	18,209	9 +ve	45,772	413	6.6	BC028792	Q91WW7	[1]NM_054088.3
Human	*Homo sapiens*	*PNPLA4* (*PLPL4*)	X:7,866,804–7,895,475	29,493	6 –ve	27,980	253	9.0	BC020746	P41247	[1]NM_001142389.1
Rat	*Rattus norvegicus*	*Pnpla4* (*Plpl4*)	X:64,019,691–64,022,515	2,825	6 –ve	27,439	252	9.1	FQ216301	na	[2]XP_343791.1
Cow	*Bos taurus*	*PNPLA4* (*PLPL4*)	Un.004.9:49,923–92,708	42,786	6 –ve	28,117	253	9.7	BT021623	na	[2]XP_590366.2
Human	*Homo sapiens*	*PNPLA5* (*PLPL5*)	22:44,276,678–44,287,760	12,299	9 –ve	47,912	429	6.3	BC031820	Q7Z6Z6	[1]NM_138814
Rat	*Rattus norvegicus*	*Pnpla5* (*Plpl5*)	7:122,105,840–122,115,990	10,151	9 –ve	50,408	453	8.5	na	D3ZXU1	[1]NM_001130497.1
Mouse	*Mus musculus*	*Pnpla5* (*Plpl5*)	15:83,943,618–83,953,543	9,926	9 –ve	48,480	432	9.0	BC109360	Q32LZ8	[1]NM_029427.1
Human	*Homo sapiens*	*PNPLA6* (*NTE*)	19:7,600,675–7,626,445	25,771	32 +ve	149,996	1,366	7.7	BC050553	Q8IY17	[1]NM_006702
Rat	*Rattus norvegicus*	*Pnpla6* (*Nte*)	12:2,610,798–2,638,622	27,825	34 –ve	148,558	1,326	8.2	na	D3ZRF6	[2]XP_001057249.2
Mouse	*Mus musculus*	*Pnpla6* (*Nte*)	8:3,516,954–3,544,025	27,072	33 +ve	149,537	1,355	8.0	AF173829	Q3TRM4	[1]NM_001122818.1
Human	*Homo sapiens*	*PNPLA7* (*PLPL7*)	9:140,354,848–140,441,851	87,007	34 –ve	145,733	1,317	7.8	AK697623	Q6ZV29	[1]NM_001098537.1
Rat	*Rattus norvegicus*	*Pnpla7* (*Plpl7*)	3:5,125,219–5,200,872	75,654	35 +ve	150,156	1,349	6.7	BC083547	Q5BK26	[1]NM_144738.2
Mouse	*Mus musculus*	*Pnpla7* (*Plpl7*)	2:24,835,554–24,909,310	73,757	35 +ve	150,494	1,352	6.5	BC025621	A2AJ88	[1]NM_146251.4
Human	*Homo sapiens*	*PNPLA8* (*IPLA2G*)	7:108,112,848–108,155,935	43,088	9 –ve	88,477	782	9.3	BC032999	Q9NP80	[1]NM_015723.2
Rat	*Rattus norvegicus*	*Pnpla8* (*Ipla2g*)	6:63,671,683–63,704,337	32,655	9 +ve	87,961	776	9.2	FQ211866	na	[2]EDM03350.1
Mouse	*Mus musculus*	*Pnpla8* (*Ipla2g*)	12:45,383,654–45,412,583	28,930	9 +ve	87,294	775	9.3	BC019364	Q8K1N1	[1]NM_026164.2
Opossum	*Monodelphis domestica*	*PNPLA1* (*PLPL1*)	2:275,536,646–275,586,990	50,345	9 +ve	60,027	540	6.3	na	na	[2]XP_001378816.2
Sea squirt	*Ciona intestinalis*	*PNPLA2* (*ATGL*)	[3]127:40,380–41,927	1,548	1 +ve	57,387	516	8.1	AK112234	na	na
zebrafish	*Danio rerio*	*PNPLA7* (*PLPL7*)	5:26,715,402–26,738,943	23,542	34 –ve	147,166	1,293	6.5	BC092427	B0S7I0	[4]BX842571.5
Frog	*Xenopus tropicalis*	*PNPLA7* (*PLPL7*)	[3]886:206,377–247,043	40,667	32 +ve	148,479	1,323	7.9	BC129023	F7ASR4	[1]NM_001097275.1
Sea squirt	*Ciona intestinalis*	*PNPLA7* (*PLPL7*)	02q:1,109,384–1,121,040	11,657	26 +ve	123,881	1,097	6.4	AK116835	na	[2]XP_002127361.1

(*Continued*)

Table 1 (*Continued*)

Vertebrate	Species	PNPLA gene (other name)	Chromosome coordinates	Gene size	Exons strand	Subunit MW	Amino acids	pI	GenBank ID	UNIPROT ID	[1]NCBI reference ID [2]NCBI predicted ID
Zebrafish	*Danio rerio*	*PNPLA8 (PLPL8)*	4:12,692,490–12,707,208	14,719	9 +ve	76,937	696	9.0	na	na	[2]XP_001918731.2
Sea squirt	*Ciona intestinalis*	*PNPLA8 (PLPL8)*	04q:5,344,967–5,349,627	4,661	10 +ve	84,461	755	8.7	AK113484	na	[2]XP_002126910.1

Notes: [1]RefSeq, the reference amino acid sequence, [2]predicted ensemble amino acid sequence, [3]scaffold IDs are shown, [4]refers to 'unknown', and [5]refers to Ensembl ID. GenBank IDs are derived from National Center for Biotechnology Information sources (http://www.ncbi.nlm.nih.gov/genbank/). UNIPROT refers to UniprotKB/Swiss-Prot IDs for individual ATGL and other PNPLA-like lipase subunits (http://kr.expasy.org). Gene size refers to base pairs of nucleotide sequences.

Abbreviations: pI, theoretical isoelectric points (the number of coding exons are listed); na, data not available.

examine predicted gene and protein structures to interrogate this database of human mRNA sequences[45] (http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/).

Predicted structures and properties of vertebrate ATGL subunits

Alignments of predicted ATGL amino acid sequences were undertaken using a ClustalW method (http://www.ebi.ac.uk/Tools/msa/clustalw2/).[46] Predicted secondary and tertiary structures for vertebrate ATGL subunits were obtained using PSIPRED[47] and SWISS MODEL web tools, respectively.[48,49] The reported tertiary structure for potato patatin[25] served as the reference for the predicted ATGL tertiary structures, with a modeling range of residues 3–173 (pig ATGL), 4–176 (opossum ATGL), 9–176 (zebrafish ATGL), and 13–183 (sea squirt ATGL). Theoretical isoelectric points and molecular weights for vertebrate ATGL and PNPLA-like subunits were obtained using Expasy web tools (http://web.expasy.org/compute_pi/).[50] Predicted membrane binding helices for PNPLA-like sequences were obtained using web tools from the Center for Biological Sequence Analysis, Technical University of Denmark (http://www.cbs.dtu.dk/services/THMMH/).[51]

Human *ATGL* gene expression and predicted gene regulation sites

The human genome browser (http://genome.ucsc.edu)[44] was used to examine GNF Expression Atlas 2 data using various expression chips for the human and mouse *ATGL* genes (http://biogps.gnf.org).[52] Predicted CpG islands, microRNA (miRNA) binding sites, and transcription factor binding sites for human *ATGL* were obtained using the UC Santa Cruz genome browser (http://genome.ucsc.edu).[44]

Phylogenetic studies and sequence divergence

Alignments of ATGL and other PNPLA-like protein sequences were assembled using BioEdit v.5.0.1 and the default settings.[53] Alignment-ambiguous regions were excluded prior to phylogenetic analysis, yielding alignments of 241 residues for comparisons of vertebrate PNPLA-like sequences with the sea squirt PNPLA-like sequences (Table 1). Evolutionary distances were calculated using the Kimura option[54] in TREECON.[55] Phylogenetic trees were constructed from evolutionary distances using the neighbor-joining method[56] and "rooted" using the internode at the PNPLA6-PNPLA7-PNPLA8 branch. Tree topology was re-examined by the boot-strap method (100 bootstraps

were applied) of resampling and only values that were highly significant (≥95) are shown.[57]

Results and discussion

Alignments and biochemical features of vertebrate ATGL amino acid sequences

Amino acid sequence alignments for 13 previously unreported vertebrate and invertebrate ATGL amino acid sequences and an invertebrate (sea squirt, *C. intestinalis*) ATGL sequence are shown in Figure 1, together with the reported sequence for human[29] and mouse[12] ATGL. The vertebrate ATGL sequences exhibited > 46% identities, suggesting that these protein subunits are products of the same gene family, whereas the sequences for the predicted vertebrate PNPLA1, PNPLA3, PNPLA4, and PNPLA5 subunits were 23%–42% identical with the ATGL sequences, indicating that these are members of distinct but related *PNPLA*-like gene families (Table 2). The sequences for the vertebrate PNPLA6, PNPLA7, and PNPLA8 subunits examined were even more distantly related, having vertebrate ATGL sequences with identities of < 16% (Table 2). However , two of these sequences (PNPLA6 and PNPLA7) 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 473–504 (ATGL), 253 (PNPLA4), 429–532 (PNPLA1, PNPLA3, and PNPLA5), 782 (PNPLA8), and 1317–1366 (PNPLA6 or NTE and PNPLA7) residues (Table 1). Consequently, vertebrate PNPLA4 is the smallest among these PNPLA-like proteins with an average molecular weight of about 28,000, while others exhibited molecular weights which are about two times (ATGL, PNPLA1, PNPLA3, and PNPLA5), three times (PNPLA8), or five times larger (PNPLA6 and PNPLA7) than PNPLA4.

Structural studies for potato patatin[25,57,58] and site-directed mutagenesis studies of human PNPLA4[59,60] have enabled identification of key catalytic residues among those aligned for the vertebrate and invertebrate ATGL (Figure 1) and PNPLA-like (mouse PNPLA1, ATGL, PNPLA3, and PNPLA5; and rat PNPLA4) sequences examined (Figure 2). These included two proposed active site motifs for vertebrate ATGL. Active site motif 1 was identified as Cys-Gly-Xaa-Leu (mouse ATGL residues 15, 16, and 18) which may function in ATGL catalysis because it is aligned with the potato patatin motif (residues 37–40 Gly-Gly-Xaa-Lys) which functions as an oxyanion hole and stabilizes the oxyanion formed during triglyceride cleavage.[25] A second active ATGL motif was identified as Gly-Xaa-Ser-Yaa-Gly (mouse ATGL residues 45–49), together with active site residues Ser47 and Asp166 which serve as the catalytic dyad during catalysis.[25] ATGL Ser428 has been identified as a site subject to site-specific phosphorylation[61] which is conserved among all vertebrate ATGL sequences examined, with the exception of guinea pig, panda, and zebrafish ATGL. Neutral isoelectric points (pI) were observed for each of the vertebrate ATGL subunits examined (pI values 5.9–8.1), as well as the human, mouse, and rat PNPLA1, PNPLA3, PNPLA5, PNPLA6, and PNPLA7 sequences. This is in contrast with the human, rat, and cow PNPLA4 sequences and the human, mouse, and rat PNPLA8 subunits which exhibited high predicted pI values (> 9.0, Table 1).

Predicted secondary and tertiary structures for vertebrate ATGL subunits

Analyses of predicted secondary structures for ATGL sequences revealed similar α-helix and β-sheet structures for all of the vertebrate subunits examined, particularly near key residues or functional domains (Figure 1). Predicted secondary (Figure 1) and tertiary structures (Figure 3) were very similar to those reported for potato patatin[25] which have been retained for all of the vertebrate ATGL sequences examined. The predicted vertebrate and invertebrate ATGL tertiary structures (pig, opossum, zebrafish, and sea squirt, Figure 2) are based on partial sequences for this enzyme (residues 6–173, 4–176, 9–176, and 13–183, respectively) which reveal the relative positioning and predicted structures for each of 5 α-helices and 6 β-sheets. These included the N-terminus α-helix (designated as α1) which contained a predicted membrane bound helix (residues 7–29, see Supplementary Figure 1) and the proposed first active site motif Cys-Gly-Xaa-Leu (human ATGL residues 15, 16, and 18); a second α-helix (α2) and β-sheet (β2) which contains the second active site motif Gly-Xaa-Ser-Yaa-Gly (residues 45–49 for human ATGL); and a β-sheet (β6) which contains Asp166, the second member of the active site dyad of catalytic residues. These structures are proximally located within a putative active site cleft which supports the predicted three-dimensional structure for this enzyme. However, any firm conclusions must await further structural studies. Several conserved serine residues were also observed for the vertebrate and invertebrate ATGL sequences examined (human ATGL Ser117, 130, 145, and 170 residues, Figure 1) which may correspond to residues previously proposed for performing structural roles in potato papatin phospholipase A.[25–57]

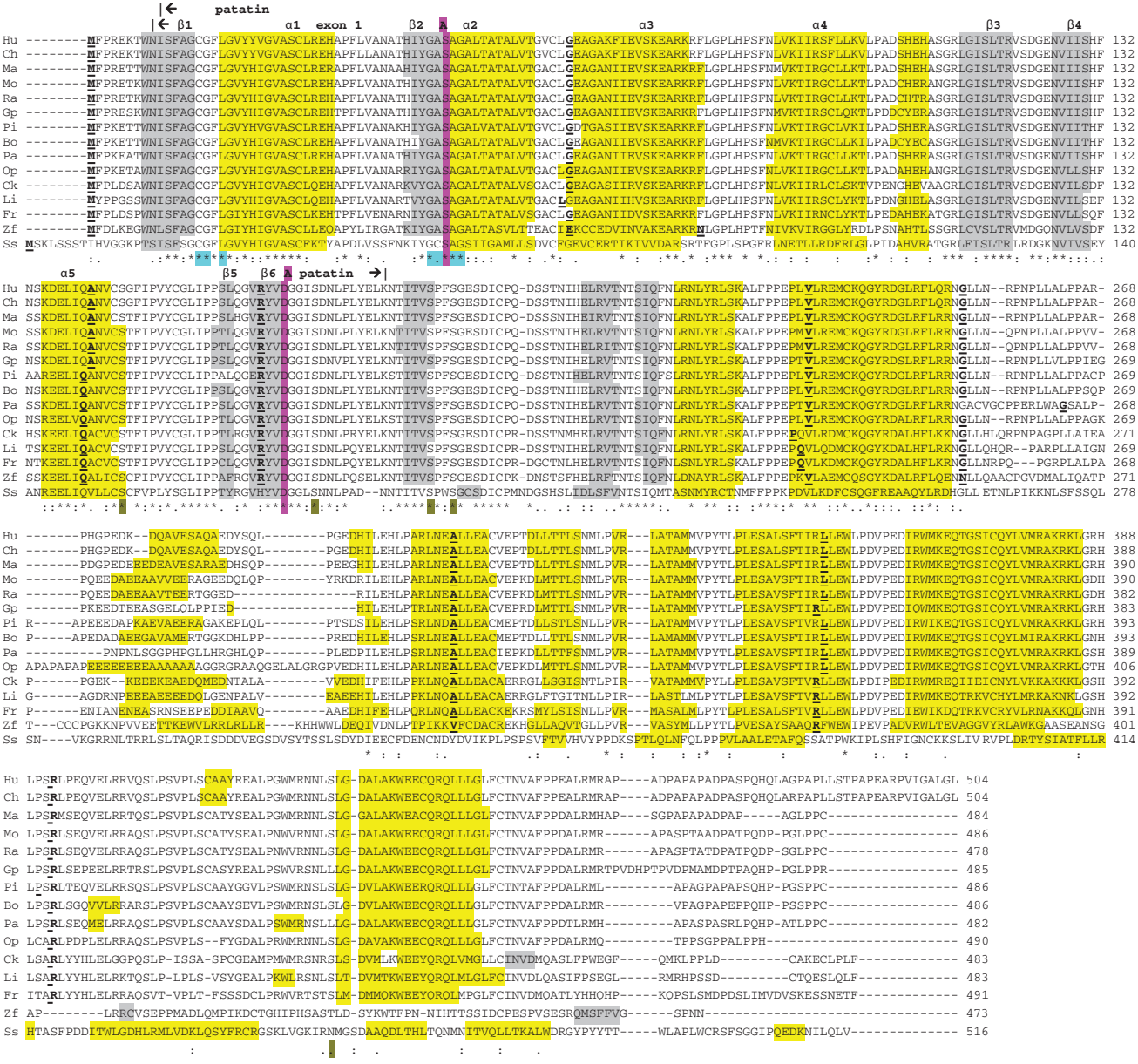

Figure 1 Amino acid sequence alignments for vertebrate ATGL sequences.

Notes: See Table 1 for sources of ATGL sequences. *Identical residues; 1 or 2 conservative substitutions; 1 or 2 nonconservative substitutions; patatin refers to predicted motif residues (6–173); predicted oxyanion hole motif (Cys-Gly-Xaa-Leu); active site motif (Gly-Xaa-Ser-Xaa-Gly); catalytic dyad residues Ser and Asp; predicted helix α1, α2; predicted sheet β1 β2; conserved serine residues; conserved Ser428 (P) for reported phosphorylation site; and bold underlined font shows predicted exon junctions.

Abbreviations: Hu, human; Ch, chimpanzee; Ma, marmoset; Mo, mouse; Ra, rat; Gp, guinea pig; Pi, pig; Bo, bovine; Pa, panda; Op, opossum; Ck, chicken; Li, lizard; Fr, frog; Zf, zebrafish; Ss, sea squirt.

Predicted gene locations, exonic structures, and expression for vertebrate *ATGL* genes

Table 1 summarizes the predicted locations and exonic structures for vertebrate *ATGL* genes based upon BLAST interrogations of several vertebrate genomes using the sequence for human[29] and mouse[12] ATGL and the predicted sequences for other vertebrate and invertebrate (sea squirt) ATGL subunits (Table 1) and the UC Santa Cruz web browser.[44] Vertebrate *ATGL* genes predominantly contained nine coding exons with the predicted exonic start sites in identical or similar positions (Figure 1). An exception was the zebrafish *ATGL* gene, which contained 10 coding exons, with the second exon for most of the other vertebrate ATGL genes being split into two coding exons. *ATGL* genes varied in size with all of the mammalian ATGL genes examined containing about 4–5 kilobase pairs, whereas other vertebrate *ATGL* genes were three

(frog [*X. tropicalis*] and zebrafish [*D. rerio*]) to nine (lizard [*A. carolensis*]) times larger (Table 1). This is mainly due to a major increase in the size of the first intron, particularly for chicken and lizard *ATGL* genes, for which intron 1 is more than10 times larger than for the other vertebrate *ATGL* genes (Supplementary Table 1).

Figure 4 illustrates the comparative predicted structures of premessenger RNA human *ATGL* gene transcripts (http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/).[45] There were nine introns and nine coding exons present for the premessenger mRNA reference (RefSeq) sequence (designated as transcript b [NM] in Figure 4) which contained a CpG131 island in the 5′-noncoding segment corresponding to the gene promoter and a second CpG108 island within the coding exon 6 region. Several transcription factor binding sites were also observed in this region: the aryl hydrocarbon receptor; the estrogen receptor, which regulates cellular responses to estrogens; early growth response 4 receptor, a transcription activator for genes involved in mitogenesis and differentiation; forkhead fox protein 01 (FOX01), a transcription factor which regulates response to oxidative stress; peroxisome proliferator-activated receptors alpha and gamma (PPARG/PPARA) which activate genes encoding enzymes of the peroxisomal β-oxidation pathway; chicken ovalbumin upstream promoter (COUP); homeobox protein 3 (HMX3), involved in the differentiation of nerve cells; and transcription regulator protein, BACH1 (Supplementary Table 2). These predicted gene regulation sites may contribute to the high level of gene expression (2.2 × the expression of the average human gene) and wide tissue expression observed for *ATGL*.[45] Elango and Yi[62] have recently reported that larger CpG islands are associated with gene promoters showing a broad range of gene expressions which contain more RNA polymerase II binding sites than other promoters. Consequently, the presence of CpG131 and the transcription factor binding sites observed within the *ATGL* gene may contribute significantly to the broad tissue expression observed for *ATGL* transcripts, as shown in Figure 5, which presents "heat maps" showing the comparative gene expression for various human and mouse tissues obtained from GNF Expression Atlas data using U133A, GNF1H, and U74A *ATGL* chips (http://genome.ucsc.edu; http://biogps.gnf.org).[52] Of particular interest are the very high levels of *ATGL* gene expression observed in human adipocytes, and in mouse brown fat, adipose tissue, mammary gland, trachea, and epidermis, which are consistent with the primary role for this enzyme in degrading triglycerides within fat cells and other tissues of the body.[12,13] The reference human *ATGL* transcript also contained an extended 3′-noncoding segment with two predicted miRNA binding sites (miR-124/506 and miR-377). miRNAs have been reported to function as post-transcriptional regulators that bind to complementary sequences on target messenger RNA transcripts (mRNAs), which result in translational repression or target degradation and gene silencing.[63]

Sequence identities and phylogeny of vertebrate ATGL and other PNPLA-like lipases

Supplementary Figure 2 shows a UCSC genome browser comparative genomics track depicting evolutionary conservation and alignment of nucleotide sequences for the human *ATGL* gene, including the intronic, exonic, and untranslated regions of this gene, with the corresponding sequences for eight vertebrate genomes. Extensive conservation was observed among these *ATGL* genomic sequences for the eutherian mammalian genomes, particularly for the primate species but also for the exonic regions for all eutherian genomes examined. In contrast with the mammalian genomes examined, the other vertebrate genomes examined lacked conserved sequences for exon 1 and the intron regions, which may reflect increased evolutionary divergence for these sequences.

A phylogenetic tree (Figure 6) was constructed from alignments of vertebrate PNPLA-like amino acid sequences with the predicted sea squirt (*C. intestinalis*) ATGL sequence, as well as sea squirt PNPLA7 and PNPLA8 sequences, serving to "root" the tree. The dendrogram showed clustering into three major groups of vertebrate PNPLA-like sequences: group 1, ATGL (the "apparent" ancestral sequence together with PNPLA1, PNPLA4, PNPLA3, and PNPLA5 sequences; group 2, PNPLA6 and PNPLA7; and group 3, PNPLA8. Group 1 is further divided according to the designation of *PNPLA*-like gene families, including the vertebrate ATGL sequences, which cluster with the sea squirt ATGL-like sequence. This suggests an ancestral relationship between invertebrate ATGL and vertebrate ATGL, as well as with other members of PNPLA-like group 1 sequences, ie, PNPLA1, PNPLA4, and PNPLA3/PNPLA5. Moreover, this is consistent with *ATGL* being an ancient gene present throughout vertebrate evolution, as well as within the invertebrate sea squirt genome.

Figure 6 also shows the number of times a clade (sequences common to a node or branch) occurred in the bootstrap analyses with replicate values of 90 or more (which are highly significant) for the 100 replicates

Table 2 Percentage identities for vertebrate ATGL and other PNPLA-like amino acid sequences

PNPLA-like Gene	Human	Mouse	Opossum	Chicken	Zebrafish	Sea squirt	Human	Mouse	Human	Mouse	Human	Rat	Human	Mouse	Human	Mouse	Human	Mouse	Human	Mouse
	ATGL	ATGL	ATGL	ATGL	ATGL	ATGL	PNPLA1	PNPLA1	PNPLA3	PNPLA3	PNPLA4	PNPLA4	PNPLA5	PNPLA5	PNPLA6	PNPLA6	PNPLA7	PNPLA7	PNPLA8	PNPLA8
Human ATGL	100	**86**	**81**	**67**	**46**	30	26	27	37	41	35	37	34	34	9	14	14	13	9	7
Mouse ATGL	**86**	100	**85**	**68**	**46**	33	25	27	36	40	35	37	34	33	7	9	10	10	9	10
Opossum ATGL	**81**	**85**	100	**70**	**47**	30	25	26	37	41	36	37	37	33	13	16	6	13	10	6
Chicken ATGL	**67**	**68**	**70**	100	**46**	30	25	26	35	42	34	34	33	34	11	12	7	12	8	8
Zebrafish ATGL	**46**	**46**	**47**	**46**	100	30	26	23	38	40	39	39	33	33	8	10	8	9	11	8
Sea squirt ATGL	30	33	30	30	30	100	23	21	23	28	35	33	27	26	8	8	8	10	5	4
Human PNPLA1	26	25	25	25	26	23	100	**63**	26	25	30	33	25	25	10	10	7	12	3	3
Mouse PNPLA1	27	27	26	26	23	21	**63**	100	25	26	30	32	26	24	8	6	5	6	5	5
Human PNPLA3	37	36	37	35	38	23	26	25	100	**65**	32	32	34	32	10	9	6	9	5	6
Mouse PNPLA3	41	40	41	42	40	28	25	26	**65**	100	28	28	27	12	7	11	12	12	8	5
Human PNPLA4	35	35	36	34	39	35	30	30	32	28	100	**61**	28	27	12	7	11	12	8	5
Rat PNPLA4	37	37	37	34	39	33	33	32	32	28	**61**	100	30	30	19	16	13	11	13	14
Human PNPLA5	34	34	37	33	33	27	25	26	34	27	28	30	100	**65**	9	9	9	7	8	6
Mouse PNPLA5	34	33	33	34	33	26	25	24	32	12	27	30	**65**	100	7	7	5	11	9	11
Human PNPLA6	9	7	13	11	8	8	10	8	10	7	12	19	9	7	100	**94**	**59**	**60**	9	8
Mouse PNPLA6	14	9	16	12	10	8	10	6	9	11	7	16	9	7	**94**	100	**61**	**58**	6	7
Human PNPLA7	14	10	6	7	8	8	7	5	6	12	11	13	9	5	**59**	**61**	100	**82**	11	7
Mouse PNPLA7	13	10	13	12	9	10	12	6	9	12	12	11	7	11	**60**	**58**	**82**	100	9	8
Human PNPLA8	9	9	10	8	11	5	3	5	5	8	8	13	8	9	9	6	11	9	100	**83**
Mouse PNPLA8	7	10	6	8	8	4	3	5	6	5	5	14	6	11	8	7	7	8	**83**	100

Notes: Numbers show the percentage of amino acid sequence identities. Numbers in bold show higher sequence identities for more closely related PNPLA-like family members.
Abbreviations: ATGL, adipose triglyceride lipase; PNPLA, patatin phospholipase.

undertaken in each case. Of particular interest are the nodes demonstrating highly significant separations for each of the vertebrate PNPLA-like group 1 gene family sequences (*ATGL*, *PNPLA4*, *PNPLA1*, *PNPLA3*, and *PNPLA5*) sequences during vertebrate evolution, which supports the distinct status for each of these genes. However, there are differences in the distribution of some gene families, with *PNPLA3* and *PNPLA5* being observed only among eutherian mammalian genomes, while *PNPLA1* was present in both eutherian and marsupial genomes, but apparently absent in other vertebrate genomes (Table 1, Figure 6). In contrast, both *ATGL* and *PNPLA4* genes were found among all vertebrate genomes examined, with *ATGL* also found in an invertebrate genome (sea squirt, *C. intestinalis*). This suggests that invertebrate *ATGL* may have served as a primordial gene for subsequent gene duplication events generating the five families of vertebrate *PNPLA*-like group 1 lipase genes. The likely sequence of these proposed group 1 *PNPLA*-like gene duplication events is summarized in Figure 6: an invertebrate *ATGL* gene undergoing duplication to form ancestral vertebrate *ATGL* and *PNPLA4* genes; vertebrate *PNPLA4* and *ATGL* genes are subsequently retained into mammalian genomes; vertebrate *ATGL* undergoes successive gene duplications to form primordial mammalian *PNPLA1*, *PNPLA3*, and *PNPLA5* genes, which are retained throughout subsequent eutherian mammalian evolution. The *PNPLA*-like gene phylogram for group 2 sequences is also consistent with a primordial invertebrate *PNPLA7* gene serving as an ancestor for vertebrate *PNPLA7*, followed by a gene duplication event generating the *PNPLA6* and *PNPLA7* sequences in eutherian mammalian genomes. The phylogram for vertebrate *PNPLA8* sequences suggests that this is an ancient gene in invertebrate and vertebrate evolution.

Previous phylogenetic studies of PNPLA-like proteins are predominantly consistent with these results, although these reports were restricted to vertebrate *PNPLA1–8*,[64] chicken *PNPLA1–8*,[65] and pig *PNPLA2–5*[66] genes and proteins, with each group inclusive of *ATGL* genes and proteins. In each case, evidence of separation into distinct PNPLA-like family groups was reported. However, this current study has also examined *PNPLA*-like genes and proteins from an invertebrate genome (sea squirt, *C. intestinalis*), and provides evidence for at least three ancestral genes (Figure 5): *ATGL* (proposed ancestral gene for vertebrate *PNPLA1*, *ATGL*, *PNPLA3*, *PNPLA4*, and *PNPLA5* genes and proteins); *PNPLA7* (proposed ancestral gene for vertebrate *PNPLA6* and *PNPLA7* genes and proteins); and *PNPLA8* (proposed ancestral gene for the vertebrate *PNPLA8* gene and protein).

Summary

The results of this study support those of previous studies for at least eight vertebrate *PNPLA*-like genes and encoded lipases, including five group 1 genes, namely *ATGL* (encoding adipose triglyceride lipase), *PNPLA1*, *PNPLA3*, and *PNPLA5*, and *PNPLA4* genes; two group 2 genes, *PNPLA6* (encoding neuropathy target esterase) and *PNPLA7*; and the group 3 gene, *PNPLA8* (encoding cytosolic phospholipase A2). Vertebrate ATGL sequences shared putative key conserved sequences, including two proposed active site motifs: Cys15-Gly16-Xaa-Leu18; and Gly45-Xaa-Ser47-Yaa-Gly49; active site residues (Ser47 and Asp166 for human ATGL); a phosphorylated Ser308 site; and several conserved serine residues. A predicted N-terminal membrane binding site proposed as a membrane anchor involved in the localization of the enzyme near adipocyte stores was also conserved within vertebrate ATGL sequences.

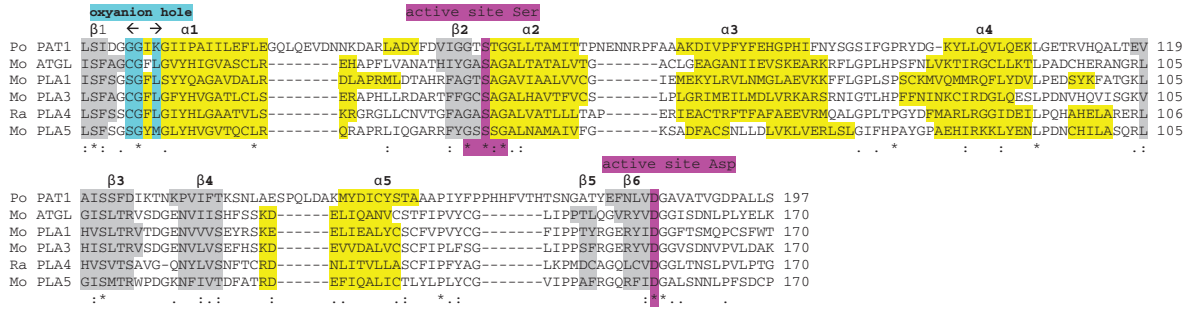

Figure 2 Amino acid sequence alignments for potato patatin, mouse ATGL, PNPLA1, PNPLA3, and PNPLA5, and rat PNPLA4 patatin "motif" sequences.
Notes: See Table 1 for sources of mouse and rat PNPLA-like sequences; the potato patatin "motif" sequence was obtained from.[25] *Identical residues; 1 or 2 conservative substitutions; 1 or 2 nonconservative substitutions; oxyanion hole motif (Gly-Gly-Xaa-Lys for potato patatin); active site motif (Gly-Xaa-Ser-Xaa-Gly); catalytic dyad residues Ser and Asp; predicted helix α1 α2; and predicted sheet β1 β2. Patatin "motif" sequences examined included Po PAT1 (potato patatin); Mo ATGL (mouse ATGL); Mo PLA1 (mouse PNPLA1); Mo PLA3 (mouse PNPLA3); Ra PLA4 (rat PNPLA4); and Mo PLA5 (mouse PNPLA5).

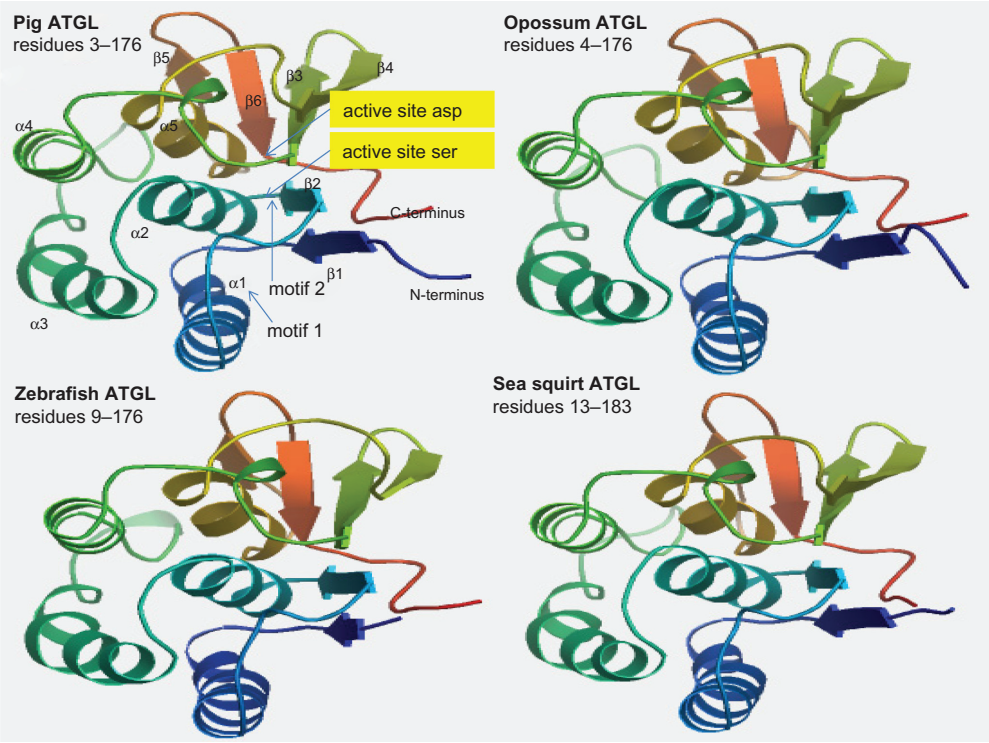

Figure 3 Predicted tertiary structures for vertebrate ATGL.
Notes: The predicted structures for pig, opossum, zebrafish, and sea squirt ATGL subunits are based on the reported structure for potato patatin[25] and obtained using the SWISS MODEL web site (http://swissmodel.expasy.org/workspace/). The rainbow color code describes the three-dimensional structures from the N-terminus (blue) to C-terminus (red color); predicted α-helices, β-sheets, proposed active site "motifs" 1 and 2; and active site residues (Ser47 and Asp166) are shown.
Abbreviation: ADPL, adipose triglyceride lipase.

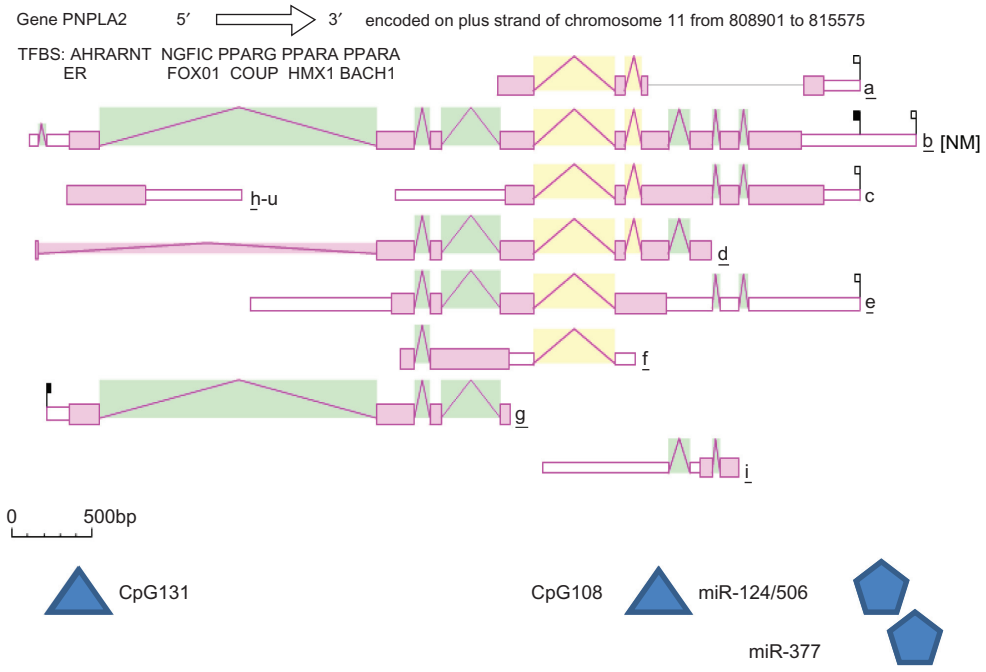

Figure 4 Gene structure and major isoforms for human ATGL.
Notes: From AceView website[46] http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/Mature isoform variants (designated as, eg, "a", "b") are shown for each *ATGL* transcript; capped 5′-ends and 3′-ends for the predicted mRNA sequences are identified; predicted CpG islands (CpG131 and CpG108), miRNA binding sites (miR-124/506 and miR-377); transcription factor binding sites (AHRARNT,[68] estrogen receptor,[69] early growth response 4 receptor,[70] FOX01,[71] PPARG,[72] COUP,[73] PPARA,[74] HMX1,[75] BACH1[76]); and a scale of base pairs of nucleotide sequences are shown.
Abbreviation: TFBS, transcription factor binding sites.

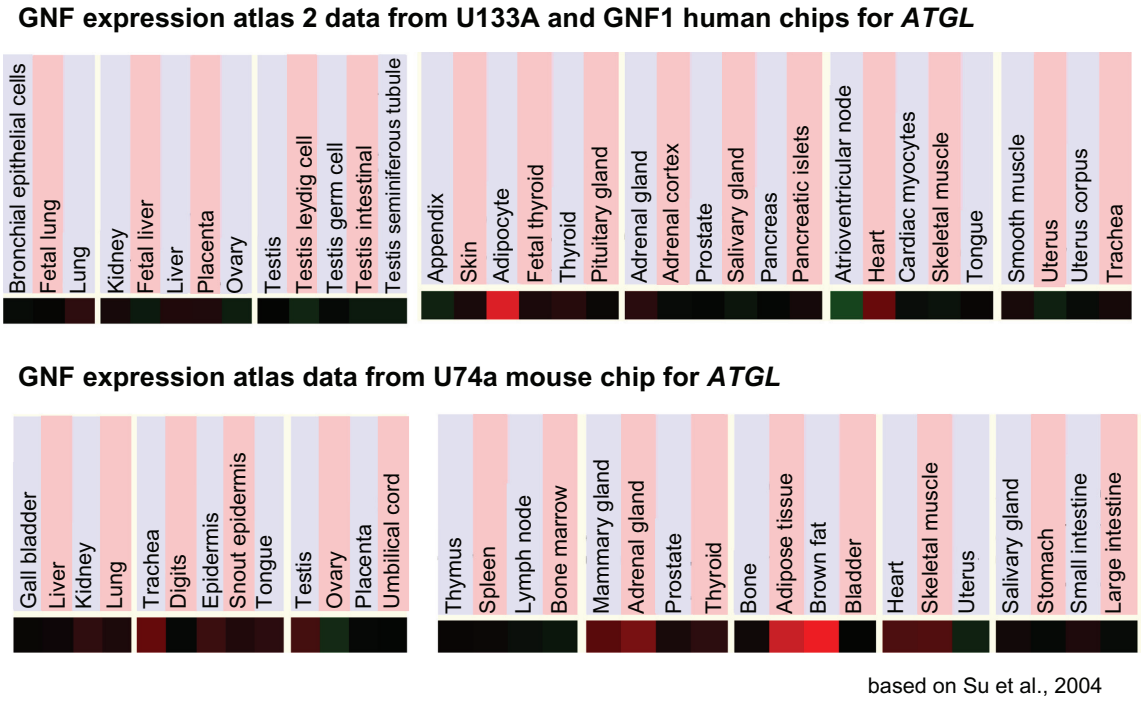

Figure 5 Human and mouse tissue gene expression "heat maps" for *ATGL*.
Notes: Taken from the human and mouse genome browsers (http://genome.ucsc.edu);[44] GNF Expression Atlas 2 data using expression chips for human and mouse ATGL (http://biogps.gnf.org);[52] comparative gene expression levels among human tissues: red (high); black, (intermediate); and green (low) expression levels.

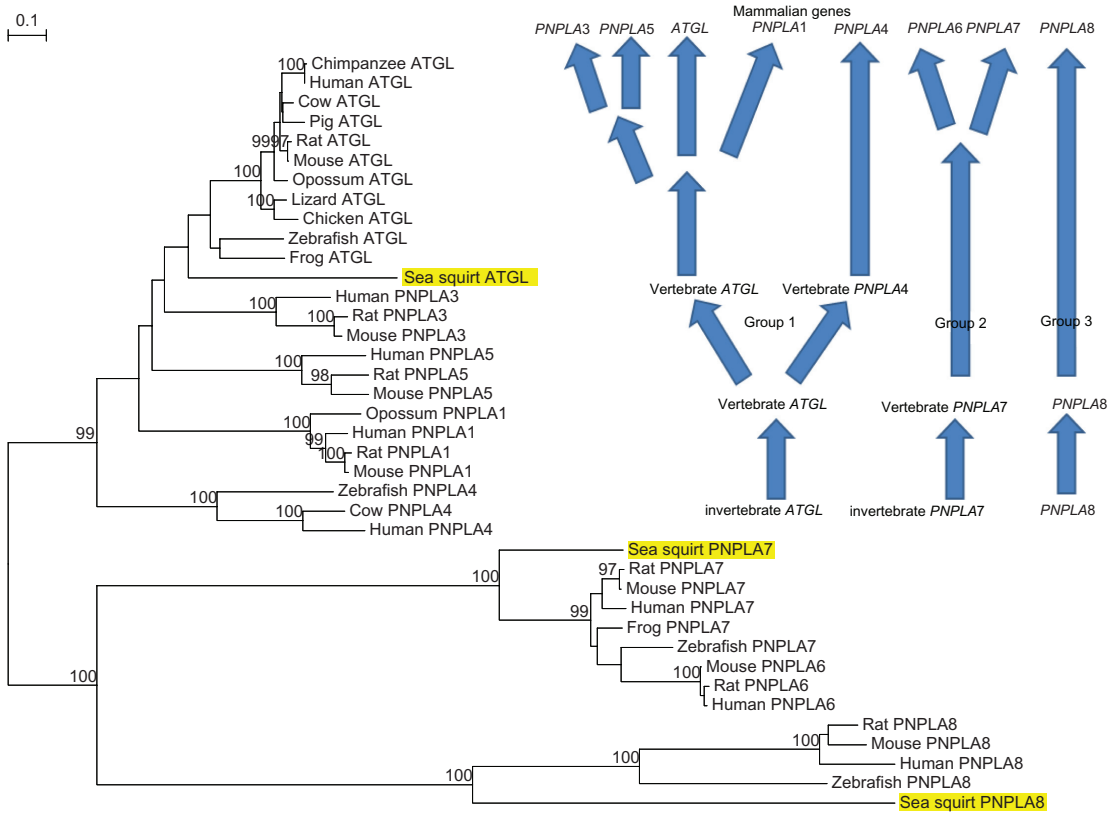

Figure 6 Phylogenetic tree of vertebrate ATGL and other PNPLA-like amino acid sequences with sea squirt (*Ciona intestinalis*) ATGL, PNPLA7, and PNPLA8 sequences.
Notes: The tree is labeled with the *ATGL* or other *PNPLA*-like gene name and the name of the vertebrate or invertebrate; note the major clusters into three groups of *PNPLA*-like gene families: Group 1 (*ATGL*, *PNPLA1*, *PNPLA3*, *PNPLA4*, and *PNPLA5*); Group 2 (*PNPLA6* and *PNPLA7*); and Group 3 (*PNPLA8*); proposed ancestral invertebrate *ATGL*, *PNPLA7*, and *PNPLA8* sequences which were used to "root" the tree are shown; gene duplication events generating these gene families are proposed to have occurred at different times: prior to the vertebrate common ancestor (*ATGL* and *PNPLA4*); prior to the appearance of marsupial mammals (*PNPLA1* and *PNPLA3*); and prior to the appearance of eutherian mammals (*PNPLA6*). A genetic distance scale is shown. The number of times a clade (sequences common to a node or branch) occurred in the bootstrap replicates is shown. Only replicate values of 90 or more which are highly significant are shown. 100 bootstrap replicates were performed in each case.

Gene expression data[52] showed that the human *ATGL* gene is broadly expressed at higher levels than those for the average gene.[45] A CpG island in the *ATGL* promoter region and several transcription factor binding sites of the *ATGL* gene may contribute to this high level of gene expression. Two miRNA binding sites localized in the extended 3′-noncoding region of the *ATGL* mRNA may contribute to the regulation of this gene during development. Phylogeny studies using several vertebrate and invertebrate ATGL subunits demonstrated that this is an ancient gene in vertebrate and invertebrate evolution. In addition, these studies of PNPLA-like subunits have suggested that there are at least three major groups based on their comparative structures and sequence phylogeny: group 1, *ATGL* (as the proposed ancestral gene), *PNPLA1*, *PNPLA3*, *PNPLA4*, and *PNPLA5*; group 2, *PNPLA6* and *PNPLA7*; and group 3, *PNPLA8*. Ancestral invertebrate group 1, group 2, and group 3 *PNPLA*-like genes and proteins were observed in the sea squirt (*C. intestinalis*) genome and used to "root" the phylogenetic tree for vertebrate *ATGL* and other *PNPLA*-like genes.

Acknowledgment

The expert assistance of Bharat Patel of Griffith University, Brisbane, Australia, with the phylogenetic analysis is gratefully acknowledged.

Disclosure

The author reports no conflict of interest in this work.

References

1. Wilson PA, Gardner SD, Lambie NM, Commans SA, Crowther DJ. Characterization of the human papatin-like phospholipase family. *J Lipid Res*. 2006;47(9):1940–1949.
2. Schoenborn V, Heid IM, Vollmert C, et al. The ATGL gene is associated with free fatty acids, triglycerides, and type 2 diabetes. *Diabetes*. 2006;55(5):1270–1275.
3. Mungall AJ, Palmer SA, Sims SK, et al. The DNA sequence and analysis of human chromosome 6. *Nature*. 2003;425(6960):805–811.
4. Dunham I, Shimizu N, Roe BA, et al. The DNA sequence of human chromosome 22. *Nature*. 1999;402(6761):489–495.
5. Lee WC, Salido E, Yen PH. Isolation of a new gene GS2 (DXS1283E) from a CpG island between STS and KAL1 on Xp22.3. *Genomics*. 1994;22(2):372–376.
6. Lush MJ, Li Y, Read DJ, Wills AC, Glynn P. Neuropathy target esterase and a homologous *Drosophila* neurodegeneration-associated mutant protein contain a novel domain conserved from bacteria to man. *Biochem J*. 1998:332(Pt 1):1–4.
7. Grimwood J, Gordon LA, Olsen A, et al. The DNA sequence of human chromosome 19. *Nature*. 2004;428(6982):528–535.
8. Humphray SJ, Oliver K, Hunt AR, et al. DNA sequence and analysis of human chromosome 9. *Nature*. 2009;429(6990):369–374.
9. Tanaka H, Takeya R, Sumimoto H. A novel intracellular membrane-bound calcium-independent phospholipase A(2). *Biochem Biophys Res Commun*. 2000;272(2):320–326.
10. Mancuso DJ, Jenkins CM, Gross RW. The genomic organisation, complete mRNA sequence, cloning, and expression of a novel membrane-associated calcium-independent phospholipase A(2). *J Biol Chem*. 2000;275(14):9937–9945.
11. Scherer SW, Cheung J, MacDonald JR, et al. Human chromosome 7: DNA sequence and biology. *Science*. 2005;300(5620):767–772.
12. Zimmermann R, Strauss JG, Haemmerle G, et al. Fat mobilization in adipose tissue is promoted by adipose triglyceride lipase. *Science*. 2004;306(5700):1383–1386.
13. Smirnova E, Goldberg EB, Makarova KS, Lin L, Brown WJ, Jackson CL. ATGL has a key role in lipid droplet/adiposome degradation in mammalian cells. *EMBO Rep*. 2006;7(1):106–113.
14. Langin D, Dicker A, Tavernier G, et al. Adipocyte lipases and defect of lipolysis in human obesity. *Diabetes*. 2005;54(11):3190–3197.
15. Zimmermann R, Lass A, Haemmerle G, Zechner R. Fate of fat: the role of adipose triglyceride lipase in lipolysis. *Biochim Biophys Acta*. 2009;1791(6):494–500.
16. Bezaire V, Mairal A, Ribet C, et al. Contribution of adipose triglyceride lipase and hormone-sensitive lipase to lipolysis in hMADS adipocytes. *J Biol Chem*. 2009;284(27):18282–18291.
17. Lake AC, Sun Y, Li JL, et al. Expression, regulation, and triglyceride hydrolase activity of adiponutrin family members. *J Lipid Res*. 2005;46(11):2477–2487.
18. Kim JY, Tillison K, Lee JH, Rearick DA, Smas CM. The adipose tissue triglyceride lipase ATGL/PNPLA2 is downregulated by insulin and TNF-alpha in 3T3-L1 adipocytes and is a target for transactivation by PPARgamma. *Am J Physiol Endocrinol Metab*. 2006;291(1):E115–127.
19. Jocken JW, Langin D, Smit E, et al. Adipose triglyceride lipase and hormone-sensitive lipase protein expression is decreased in the obese insulin-resistant state. *J Clin Endocrinol Metab*. 2007;92(6):2292–1199.
20. Alsted TJ, Nybo L, Schweiger M, et al. Adipose triglyceride lipase in human skeletal muscle is upregulated by exercise training. *Am J Physiol Endocrinol Metab*. 2009;296(3):E445–E453.
21. Akiyama M, Sakai K, Ogawa M, McMillan JR, Sawamura D, Shimizu H. Novel duplication mutation in the patatin domain of adipose triglyceride lipase (PNPLA2) in neutral lipid storage disease with severe myopathy. *Muscle Nerve*. 2007;36(6):856–859.
22. Campagna F, Nanni L, Quagliarini F, et al. Novel mutations in the adipose triglyceride lipase gene causing neutral lipid storage disease with myopathy. *Biochem Biophys Res Commun*. 2008;377(3):843–846.
23. Schweiger M, Lass A, Zimmermann R, Eichmann TO, Zechner R. Neutral lipid storage disease: genetic disorders caused by mutations in adipose triglyceride lipase/PNPLA2 or CGI-58/ABHD5. *Am J Physiol Endocrinol Metab*. 2009;297(2):E289–E296.
24. Notari L, Baladron V, Aroca-Aguilar JD, et al. Identification of a lipase-linked cell membrane receptor for pigment epithelium-derived factor. *J Biol Chem*. 2006;281(49):38022–38037.
25. Rydel TJ, Williams JM, Krieger E, et al. The crystal structure, mutagenesis, and activity studies reveal that patatin is a lipid acyl hydrolase with a Ser-Asp catalytic dyad. *Biochemistry*. 2003;42(22):6696–6708.
26. Cygler M, Schrag JD. Structure as basis for understanding interfacial properties of lipases. *Methods Enzymol*. 1997;284:3–27.
27. Dessen A, Tang A, Schmidt H, et al. Crystal structure of human phopholipase A(2) reveals a novel topology and catalytic mechanism. *Cell*. 1999;97(3):349–360.
28. Altschul F, Vyas V, Cornfield A, et al. Basic local alignment search tool. *J Mol Biol*. 1990;215(3):403–410.
29. Jenkins CM, Mancuso DJ, Yan W, Sims HF, Gibson B, Gross RW. 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*. 2004;279(47):48968–48975.
30. International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature*. 2001;409(6822):860–921.

31. Chimpanzee Sequencing and Analysis Consortium. Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature*. 2005;437(7055):69–87.
32. Wade CM, Giulotto E, Sigurdsson S, et al. Genome sequence, comparative analysis, and population genetics of the domestic horse. *Science*. 2009;326(5954):865–867.
33. Elsik CG, Tellam RL, Worley KC; The Bovine Genome Sequencing and Analysis Consortium. The genome sequence of taurine cattle: a window to ruminant biology and evolution. *Science*. 2009;324(5926): 522–528.
34. Mouse Genome Sequencing Consortium. Initial sequencing and comparative analysis of the mouse genome. *Nature*. 2002;420(6915):520–562.
35. Rat Genome Sequencing Project Consortium. Genome sequence of the brown Norway rat yields insights into mammalian evolution. *Nature*. 2004;428(6982):493–521.
36. Lindblad-Toh K, Wade CM, Mikkelsen TS, et al. Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature*. 2005;438(7069):803–819.
37. Mikkelsen TS, Wakefield MJ, Aken B, et al. Genome of the marsupial *Monodelphis domestica* reveals innovation in non-coding sequences. *Nature*. 2007;447(141):167–177.
38. International Chicken Genome Sequencing Consortium. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature*. 2004;432(701):695–716.
39. Alfoldi J, Di Palma F, Grabherr M, et al. The genome of the green anole lizard and a comparative analysis with birds and mammals. *Nature*. 2011;477(7366):587–591.
40. Hellsten U, Harland RM, Gilchrist MJ, et al. The genome of the western clawed frog *Xenopus tropicalis*. *Science*. 2010;328(5978):633–636.
41. Sprague J, Bayraktaroglu L, Bradford Y, et al. The zebrafish information network: the zebrafish model organism database. *Nucleic Acids Res*. 2005;34:D581–D585.
42. Dehal P, Satou Y, Campbell RK, et al. The draft genome of *Ciona intestinalis*: insights into chordate and vertebrate origins. *Science*. 2002;298(560):2157–2167.
43. Baulande S, Lasnier F, Lucas M, Pairalt J. Adiponutrin, a transmembrane protein corresponding to a novel dietary- and obesity-linked mRNA specifically expressed in the adipose lineage. *J Biol Chem*. 2001;276(36):33336–33344.
44. Kent WJ, Sugnet CW, Furey TS, et al. The human genome browser at UCSC. *Genome Res*. 2003;12(6):994–1006.
45. Thierry-Mieg D, Thierry-Mieg J. AceView: a comprehensive cDNA-supported gene and transcripts annotation. *Genome Biol*. 2006;7:S12.
46. Chenna R, Sugawara H, Koike T, et al. Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res*. 2003; 31(13):3497–3500.
47. McGuffin LJ, Bryson K, Jones DT. The PSIPRED protein structure prediction server. *Bioinformatics*. 2000;16(4):404–405.
48. Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modelling. *Electrophoresis*. 1997;18(15):2714–2723.
49. Kopp J, Schwede T. The SWISS-MODEL repository of annotated three-dimensional protein structure homology models. *Nucleic Acids Res*. 2004;32:D230–D234.
50. Gasteiger E, Hoogland C, Gattiker A, et al. Protein identification and analysis tools on the ExPASy server. In: Walker JM, editor. *The Proteomics Protocols Handbook*. Totowa, NJ: Humana Press; 2005.
51. Moller S, Croning MDR, Apweiler R. Evaluation of methods for the prediction of membrane spanning regions. *Bioinformatics*. 2001;17(7):646–653.
52. Su AI, Wiltshire T, Batalov S, et al. A gene atlas of the human and mouse protein encoding transcriptomes. *Proc Natl Acad Sci U S A*. 2004;101(16):6062–6067.
53. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp*. 1999;41:95–98.
54. Kimura M. *The Neutral Theory of Molecular Evolution*. Cambridge, MA: Cambridge University Press; 1983.
55. Van De Peer Y, de Wachter R. TreeCon for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. *Comput Appl Biosci*. 1994;10(5):569–570.
56. Saitou N, Nei N. The neighbour-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol*. 1987;4(4): 406–426.
57. Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*. 1985;39:783–791.
58. Hirschberg HJ, Simons JW, Dekker N, Egmond MR. Cloning, expression, purification and characterization of patatin, a novel phospholipase A. *Eur J Biochem*. 2001;268(19):5037–5044.
59. Gao J, Simon M. Identification of a novel keratinocyte retinyl ester hydrolase as a transacylase and lipase. *J Invest Dermatol*. 2005;124(6):1259–1266.
60. Gao JG, Shin A, Gruber R, Schmuth M, Simon M. GS2 as a retinol transacylase and as a catalytic dyad independent regulator of retinylester accretion. *Mol Genet Metabol*. 2009;96(4):253–260.
61. Dephoure N, Zhou C, Villén J, et al. A quantitative atlas of mitotic phosphorylation. *Proc Natl Acad Sci U S A*. 2008;105(31):10762–10767.
62. Elango N, Yi SV. Functional relevance of CpG island length for regulation of gene expression. *Genetics*. 2011;187(4):1077–1083.
63. Bartel DP. MicroRNAs: target recognition and regulatory functions. *Cell*. 2009;136(2):215–233.
64. Kienesberger PC, Oberer M, Lass A, Zechner R. Mammalian patatin-domain containing proteins: a family with diverse lipolytic activities involved in multiple biological functions. *J Lipid Res*. 2009; 50 Suppl:S63–S68.
65. Saarela J, Jung G, Hermann M, Nimpf J, Schneider WJ. The patatin-like lipase family in *Gallus gallus*. *BMC Genomics*. 2008;9:281.
66. Chen Z, Gao X, Lei T, et al. Molecular characterization, expression and chromosomal localization of porcine *PNPLA3* and *PNPLA4*. *Biotechnol Lett*. 2011;33(7):1327–1337.
67. Fujii-Kuriyama Y, Ema M, Mimura J, Sogawa K. Ah receptor: a novel ligand-activated transcription factor. *Exp Clin Immunogenet*. 1994;11(2–3):65–74.
68. Green S, Walter P, Kumar V, et al. Human oestrogen receptor cDNA: sequence, expression and homology to v-erb-A. *Nature*. 1986;320(6058): 134–139.
69. Decker EL, Nehmann N, Kampen E, Eibel H, Zipfel PF, Skerka C. Early growth response proteins (EGR) and nuclear factors of activated T cells (NFAT) form heterodimers and regulate proinflammatory cytokine gene expression. *Nucleic Acids Res*. 2003;31(3):911–921.
70. Biggs WV III, Cavenee WK, Arden KC. Identification and characterization of members of the FKHR (FOX O) subclass of winged-helix transcription factors in the mouse. *Mamm Genome*. 2001;12(6): 416–425.
71. Haemmerle G, Moustafa T, Woelkart G, et al. ATGL-mediated fat catabolism regulates cardiac mitochondrial function via PPAR-α and PGC−1. *Nat Med*. 2011;17(9):1076–1085.
72. Wang LH, Tsai SY, Cook RG, Beattie WG, Tsai MJ, O'Malley BW. COUP transcription factor is a member of the steroid receptor superfamily. *Nature*. 1989;340(6229):163–166.
73. Gearing KL, Crickmore A, Gustafsson JA. Structure of the mouse peroxisome proliferator activated receptor alpha gene. *Biochem Biophys Res Commun*. 1994;199(1):255–263.
74. Munroe RJ, Prabhu V, Acland GM, et al. Mouse H6 Homeobox 1 (Hmx1) mutations cause cranial abnormalities and reduced body mass. *BMC Dev Biol*. 2009:9;27.
75. Dohi Y, Ikura T, Hoshikawa Y, et al. Bach1 inhibits oxidative stress-induced cellular senescence by impeding p53 function on chromatin. *Nature Struct Mol Biol*. 2008;15(12):1246–1254.
76. Toki T, Itoh J, Kitazawa J, et al. Human small Maf proteins form heterodimers with CNC family transcription factors and recognize the NF-E2 motif. *Oncogene*. 1997;14(16):1901–1910.

Supplementary materials

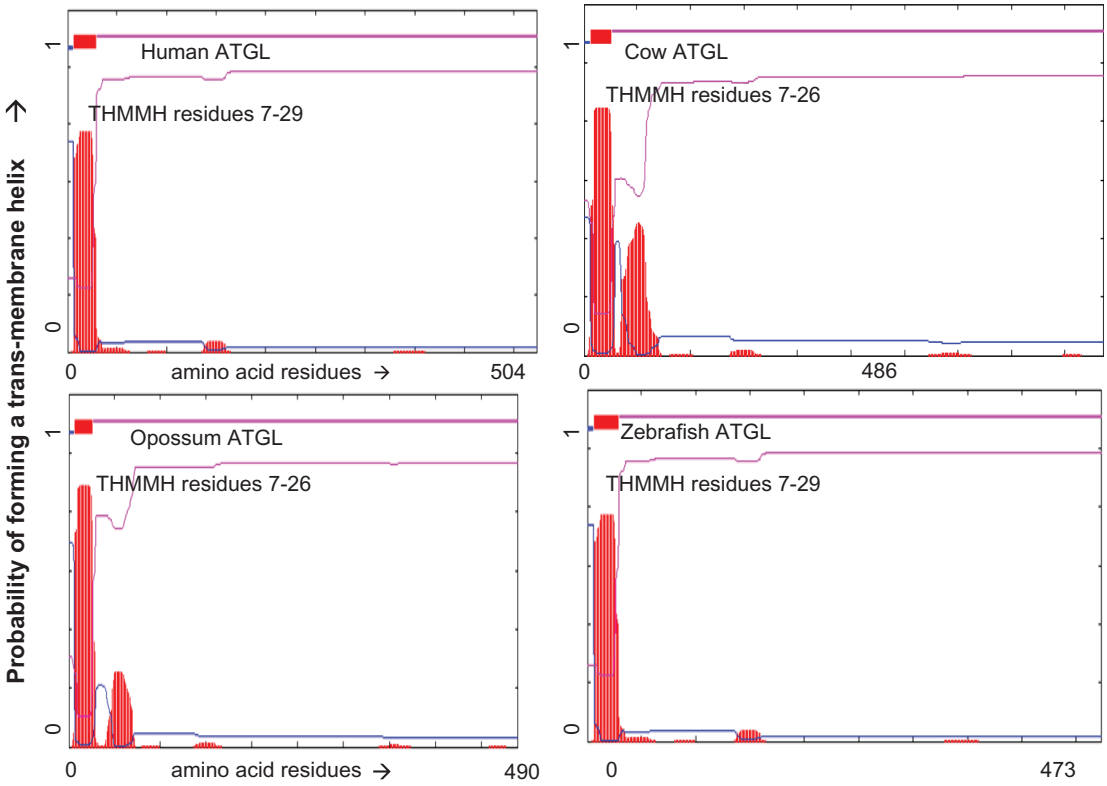

Figure S1 Predicted transmembrane structures (THMMH) for vertebrate ATGL sequences.
Notes: Derived from predictions using CBS web tools (Center for Biological Sequence Analysis, Technical University of Denmark (http://www.cbs.dtu.dk/services/THMMH/). Red peaks indicated positioning of predicted transmembrane structures (THMMH amino acid sequences are identified); pink-sequence outside membrane; blue sequence inside membrane; y axes show probabilities of forming THMMH (0→1); x axes show amino acids; individual ATGL sequences examined include human (*Homo sapiens*), cow (*Bos taurus*), opossum (*Monodelphis domestica*), and zebrafish (*Danio rerio*).

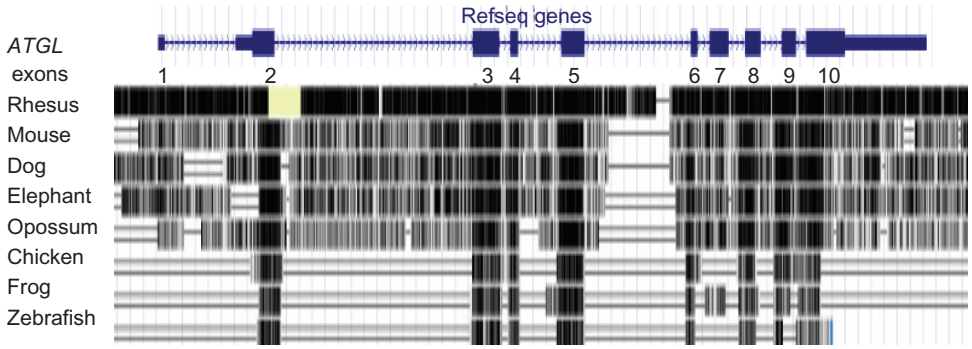

Figure S2 Comparative sequences for vertebrate ATGL sequences.
Notes: Derived from the UCSC genome browser using the comparative genomics track to examine alignments and evolutionary conservation of ATGL gene sequences; genomic sequences aligned for this study included primate (human and rhesus), other mammal (mouse, dog, elephant, and opossum), and other vertebrate species (chicken, frog, and zebrafish); conservation measures were based on conserved sequences across all of these species in the alignments; regions of sequence identity are shaded.

Table S1 Comparative intron sizes for vertebrate *ATGL* genes

Vertebrate	*ATGL* gene size (bps)	Intron 1 bps	Intron 2 bps	Intron 3 bps	Intron 4 bps	Intron 5 bps	Intron 6 bps	Intron 7 bps	Intron 8 bps
Human	5,141	1,722	97	373	920	106	142	183	86
Mouse	4,379	1,799	86	347	215	101	85	189	99
Rat	4,286	1,773	83	353	189	100	82	188	84
Pig	4,273	1,770	89	320	191	109	79	170	87
Cow	4,157	1,564	79	355	221	101	96	191	92
Opossum	4,792	1,543	146	184	783	106	314	177	69
Chicken	30,118	21,146	171	4,655	848	886	423	730	710
Lizard	45,618	29,925	239	6,135	1,526	2,008	652	1,576	2,048
Frog	14,284	6,276	1,521	1,610	278	672	116	1,165	1,055

Note: Numbers show the comparative sizes for vertebrate *ATGL* introns.
Abbreviation: bps, base pairs of nucleotides.

Table S2 Identification, location and comparative functions for human *ATGL* transcription factor binding sites

TFBS	Name	Chromosome coordinates	Chromosome location (strand)	Role
AHRARNT	Aryl hydrocarbon receptor[1]	chr11:819809–819824	Exon 2 (–)	Activates multiple phase I and II xenobiotic chemical metabolizing enzyme gene expression
ER	Estrogen receptor[2]	chr11:819883–819901	Exon 2 (–)	Involved in the hormonal regulation of eukaryotic gene expression
NGFIC	Early growth response 4[3]	chr11:820189–820200	Intron 2 (+)	Transcription activator of genes required for mitogenesis and differentiation.
FOX01	Forkhead fox protein 01[4]	chr11:820601–820614	Intron 2 (+)	Transcription factor which acts as a regulator of cell responses to oxidative stress
PPARG	Peroxisome proliferator-activated receptor gamma[5]	chr11:820704–820724	Intron 2 (–)	Activates genes encoding enzymes of the peroxisomal beta-oxidation pathway of fatty acids
COUP	Chicken ovalbumin upstream promoter[6]	chr11:820708–820721	Intron 2 (+)	Stimulates transcription initiation
PPARA	Peroxisome proliferator-activated receptor alpha[7]	chr11:820708–820727	Intron 2 (–)	Activates genes encoding enzymes of the peroxisomal beta-oxidation pathway of fatty acids
HMX1	Homeobox protein HMX1[8]	chr11:821301–821310	Intron 2 (+)	Transcription factor involved in the differentiation of neuronal cell types
PPARA	Peroxisome proliferator-activated receptor alpha[7]	chr11:821315–821334	Intron 2 (+)	Activates genes encoding enzymes of the peroxisomal beta-oxidation pathway of fatty acids
BACH1	Transcription regulator protein BACH1[9]	chr11:821379–821393	Intron 2 (–)	Coordinates gene transcription activation and repression by MAFK[10]
MYB	Transcriptional activator Myb[11]	chr11:821397–821406	Intron 3 (–)	Control of proliferation and differentiation of hematopoietic progenitor cells

Notes: Predicted transcription factor binding sites identified: [1]AHRARNT;[66,2] estrogen receptor;[67,3] early growth response 4 receptor;[68,4] FOX01;[69,5] PPARG;[70,6] COUP;[71,7] PPARA;[72,8] HMX1;[73,9] BACH1;[74,10] MAFK;[75,11] MYB[76].
Abbreviations: TFBS, transcription factor binding sites; UCSC, University of California Santa Cruz.