Journal of
Data Mining in Genomics & Proteomics

Holmes and Cox, J Data Mining in Genom Proteomics 2012, 3:1
DOI: 10.4172/2153-0602.1000112

Research Article Open Access

Bioinformatics and Evolution of Vertebrate Pancreatic Lipase and Related Proteins and Genes

Roger S Holmes[1,2,3]* and Laura A Cox[1,2]

[1]Department of Genetics, Griffith University, Nathan, QLD, Australia
[2]Southwest National Primate Research Center, Texas Biomedical Research Institute, San Antonio, TX, USA
[3]School of Biomolecular and Physical Sciences, Griffith University, Nathan, QLD, Australia

Abstract

Background: Pancreatic lipase (PTL) functions in the presence of colipase in the hydrolysis of emulsified fats in the small intestine following secretion from the pancreas. Pancreatic lipase related protein 1 (PLR1) is also found in pancreatic secretions and may perform a regulatory role in lipolysis; PLR2 catalyses pancreatic triglyceride and galactolipase reactions while PLR3 may serve a related but unknown lipase function. Comparative PTL, PLR1, PLR2 and PLR3 amino acid sequences and structures and gene locations and sequences were examined using data from several vertebrate genome projects.

Methods: Sequence alignments and conserved predicted secondary and tertiary structures were studied and key amino acid residues and domains identified based on previous reports on human and pig PTL. Comparative analyses of vertebrate *PTL*-like genes were conducted using the UC Santa Cruz Genome Browser. Phylogeny studies investigated the evolution of these vertebrate *PTL*-like genes.

Data: Human and mouse PTL sequences shared 78% identities but only 64-68% identities with human and mouse PLR1 and PLR2 sequences. Several vertebrate PTL and PTL-like protein domains were predicted using bioinformatics including an N-signal peptide; N-glycosylation site(s); an α/β hydrolase fold region containing a catalytic triad; a 'lid' region for the active site; a 'hinge' separating the lipase and PLAT regions; and a C-terminal PLAT region. Eutherian mammalian PLR1 sequences retained residues (196Val/198Ala) responsible for the loss of triacylglycerol lipase activity whereas lower vertebrate PLR1 sequences retained the 'active' lipase residues. In contrast, mammalian PTL, PLR2 and PLR3 sequences exhibited lipase 'active' residues (196Ala/198Pro); chicken and frog PLR1 sequences retained the lipase 'active' residues; and opossum and platypus PLR1 sequences exhibited 196Ala/Ser198 and 196Ser/198Pro residues. A phylogenetic tree analysis provided evidence for four distinct vertebrate *PTL*-like gene families.

Conclusions: Pancreatic lipase (*PTL*) and related genes and proteins (*PLR1* and *PLR2*) are present in all vertebrate genomes examined whereas *PLR3* is found only in primate genomes. The 'inactive' form of vertebrate PLR1 is restricted to eutherian mammals. Vertebrate *PTL*-like genes apparently originated in a vertebrate ancestor following gene duplication events of an ancestral *PTL*-like ancestral gene. Two separate lines of *PTL*-like gene evolution are proposed in lower vertebrates (*PTL/PLR1* and *PLR2),* with a further gene duplication event (*PLR2*/*PLR3*) for primate genomes.

Keywords: Vertebrate; Amino acid sequence; Pancreatic lipase; Pancreatic lipase related protein; Evolution; Bioinformatics

Introduction

Pancreatic lipase (PTL; EC 3.1.1.3; gene *LIPP, PTL* or *PNLIP*) is a major exocrine pancreas secretory enzyme which functions in the presence of a cofactor (colipase: CLPS; gene *CLPS*) in the hydrolysis of emulsified fats in the small intestine following secretion from the pancreas [1-5]. Two other PTL-related proteins, pancreatic lipase related-protein 1 (PLR1; gene *PLR1* or *PNLPR1*) and pancreatic lipase related protein 2 (PLR2; E.C.3.1.1.3; gene *PLR2* or *PNLPR2*) are also major pancreatic secretory proteins which are structurally similar to PTL [5-8]. A third PTL-related protein and gene, pancreatic lipase related protein 3 (PLR3; E.C.3.1.1.3; gene *PLR3* or *PNLPR3*), has also been reported in humans [9].

Biochemical studies of mammalian PTL lipases have shown that these enzymes are members of an esterase/lipase superfamily hydrolysing long-chain acyl-triglycerides at a water/lipid interface [4,5,10,11]. Structural and molecular modeling studies of human, pig and horse PTL [12-16] have shown that these enzymes contain several structural and functional domains, includng an N-terminal signal peptide; an N-glycosylation site [1, 4]; an α/β hydrolase fold containing a Ser/Asp/His active site catalytic triad [17]; a 'lid' which typically covers the active site but in the presence of substrate, colipase and bile salts, PTL undergoes a conformational change enabling access to the active site by the emulsified lipid substrate [14, 16-20]; a C-terminal β-sandwich (PLAT) domain which participates in lipid and colipase binding; and a 'hinge' region which separates the α/β hydrolase fold from the PLAT domain. Structural and biochemical studies of human PLR1 have shown similarities in sequence and domain regions with mammalian PTL [21], although no detectable *in vitro* catalytic activity with PTL-like substrates has been reported [6,8]. Human PLR1 has been 'reactivated' by site-directed mutagenesis at two sites corresponding to human PTL residues (Ala196Val and Pro198Ala)

***Corresponding author:** Roger S Holmes, School of Biomolecular and Physical Sciences, Griffith University, Nathan 4111 Qld Australia, Tel: (+) 61 7 3735 5077; Fax: (+) 61 7 3735 7773, E-mail: r.holmes@griffith.edu.au

Received November 03, 2011; **Accepted** November 21, 2011; **Published** January 14, 2012

Citation: Holmes RS, Cox LA (2012) Bioinformatics and Evolution of Vertebrate Pancreatic Lipase and Related Proteins and Genes. J Data Mining in Genom Proteomics 3:111. doi:10.4172/2153-0602.1000111

representing the active and inactive forms of PLR1, respectively [7,8]. No specific biochemical roles have as yet been reported for mammalian PLR1 although a possible role in modulating colipase binding to PTL has been suggested [5]. Ren and coworkers [22] have recently examined *PLR1* knock-out (KO) mice and reported mature onset obesity with increased fat mass, and have proposed that the enzyme may function as a metabolic inhibitor of PTL-colipase catalyzed triglyceride metabolism.

Structural and biochemical studies of mammalian PLR2 have reported strong similarities with PTL although differences were observed in kinetic and colipase binding properties and significant species differences were also reported [4-7,23-25]. Significantly, this enzyme has been shown to be predominantly responsible for pancreatic lipase-like neonatal fat digestion in rats and mice [26] and participates in the dietary fat digestion in human newborn children [27,28]. PLR2 also functions as a galactolipase (these substrates are the major lipids in plant cells) and participates in the hydrolysis of triglycerides, phospholipids and vitamin A esters [5,29-31]. There are no reports available concerning the functions of human PLR3 although a recent study [32] has demonstrated over-expression of this gene in hepatocellular carcinoma (HCC) indicating an association of *PLR3* with HCC.

Evolutionary studies of vertebrate *PTL*-like genes and proteins have been largely restricted to structural and genomic comparisons with other members of the neutral lipase gene superfamily, including lipoprotein lipases (LPL), hepatic lipases (HL), endothelial lipases (EL) and phospholipases A1 [5,10,33-37]. This paper reports predicted gene structures and amino acid sequences for several vertebrate pancreatic lipase (*PTL*) and related protein (*PLR1, PLR2* and *PLR3*) genes and enzymes not previously reported. Predicted secondary and tertiary structures for vertebrate pancreatic lipase (PTL) and related proteins (PLR1, PLR2 and PLR3) are also described as well as the structural and evolutionary relationships of these genes and enzymes. Comparative and phylogeny studies also showed that amino acid sequences for the PTL-like family are highly conserved during vertebrate evolution and suggested that the genes encoding these enzymes were generated from gene duplication events of an ancestral *PTL*-like ancestral gene resulting in two separate lines of mammalian gene evolution: *PTL/PLR1* and *PLR2/PLR3*, with the latter gene being observed only in primate genomes.

Methods

Vertebrate pancreatic lipase and related protein gene and protein identification

BLAST (Basic Local Alignment Search Tool) studies were undertaken using web tools from the National Center for Biotechnology Information (NCBI) (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Protein BLAST analyses used human PTL, PLR1, PLR2 and PLR3 amino acid sequences previously described (Table 1). Non-redundant protein sequence databases for several vertebrate genomes were examined using the blastp algorithm, including human (*Homo sapiens*) [38]; orangutan (*Pongo abelii*) (http://genome.wustl.edu) ; rhesus monkey (*Macaca mulatta*) [39]; marmoset (*Callithrix jacchus*) (http://genome.ucsc.edu/), cow (*Bos Taurus*) [40]; horse (*Equus caballus*) [41]; mouse (*Mus musculus*) [42]; rat (*Rattus norvegicus*) [43]; opossum (*Monodelphis domestica)* [45]; chicken (*Gallus gallus*) [45]; frog (*Xenopus tropicalis*) [46]; and zebrafish (*Danio rerio*) [47]. 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 mRNAs and encoded PTL-like proteins. The records were derived from annotated genomic sequences using the gene prediction method (GNOMON). Vertebrate genomes were chosen which showed high similarity scores for human PTL, PLR1, PLR2 and PLR3 (see Table 1 and Supplementary Table 1) and predicted PTL-like protein sequences were then subjected to analyses of predicted protein and gene structures.

BLAT (Blast Like Alignment Tool) analyses were subsequently undertaken for each of the predicted PTL-like amino acid sequences using the UC Santa Cruz genome browser [http://genome.ucsc.edu/cgi-bin/hgBlat] with the default settings to obtain the predicted locations for each of the vertebrate *PTL*-like genes, including predicted exon boundary locations and gene sizes for coding exons. Structures for human *PTL*, *PLR1*, *PLR2* and *PLR3* isoforms (splicing variants) were obtained using the AceView website to examine predicted gene and protein structures (http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/index.html?human) [48]. The UC Santa Cruz web browser genome browser (http://genome.ucsc.edu) was used to examine comparative structures for vertebrate *PTL*-like genes [49].

Predicted structures, properties and alignments of vertebrate PTL-like sequences

Predicted secondary and tertiary structures for human and other vertebrate PTL-like proteins were obtained using the PSIPRED v2.5 web site tools [http://bioinf.cs.ucl.ac.uk/psipred/psiform.html] and SWISS MODEL web tools [http://swissmodel.expasy.org/], respectively. The reported tertiary structure for human PTL [13] (PDB entry 1n8s) was used as the reference for the frog PTL tertiary structure with a modeling range of residues 17-465; and the reported structure for human PLR1 (unpublished results: PDB entry 2ppl) served as a reference for frog PLR1 with a modeling range of 19 to 467). Alignments of vertebrate PTL-like sequences were assembled using the ClustalW2 alignment program (http://www.ebi.ac.uk/Tools/clustalw2/index.html) [50].

Phylogenetic studies and sequence divergence

Alignments of vertebrate PTL-like sequences and human (*Homo sapiens*), mouse (*Mus musculus*), frog (*Xenopus tropicalis*) and zebrafish (*Danio rerio*) EL (endothelial lipase), LPL (lipoprotein lipase) and HL (hepatic lipase) sequences were assembled using BioEdit v.5.0.1 and the default settings [51]. Alignment ambiguous regions, including the amino and carboxyl termini, were excluded prior to phylogenetic analysis yielding alignments of 395 residues for comparisons of vertebrate PTL-like, EL, LPL and HL sequences with the fruit fly (*Drosophila melanogaster*) CG5966 lipase sequence (Table 1; Supplementary Table 1). Evolutionary distances were calculated using the Kimura option [52] in TREECON [53]. Phylogenetic trees were constructed from evolutionary distances using the neighbor-joining method [54] and rooted with the fruit fly CG5966 lipase sequence. Tree topology was reexamined by the boot-strap method (100 bootstraps were applied) of resampling and only values that were highly significant (≥90) are shown [55].

Results and Discussion

Alignments of vertebrate PTL amino acid sequences

The deduced amino acid sequences for rhesus monkey (*Macaca mulatta*), dog (*Canis familiaris*) and opossum (*Monodelphis domestica*) PTL, which were identified in the NCBI database (see Table 1) (http://www.ncbi.nlm.nih.gov/), are shown in Figure 1 together with previously reported PTL sequences for human (*Homo sapiens*) [12,56],

Pancreatic lipase-like Gene	Species	Gene ID	RefSeq ID [1]Ensembl	GenBank ID	UNIPROT ID	Amino acids	Chromosome location	Exons (strand)	Gene Size bps	pI	Subunit MW	Signal Peptide (Cleavage site)
Human *PTL*	*Homo sapiens*	*PTL* or *PNLIP* or *LIPP*	NM_000936	BC014319	P16233	465	10:118,295,595-118,317,297	12 (+)	21,703	6.3	51,157	1-16 [AG-KE]
Human *LIPR1*	*Homo sapiens*	*PLR1* or *PNPLRP1* or *PLRP1*	NM_006229	BC025784	P54315	467	10:118,341,942-118,358,615	12 (+)	16,674	5.5	51,848	1-17 [KG-KE]
Human *LIPR2*	*Homo sapiens*	*PLR2* or *PNPLRP2* or *PLRP2*	NM_005396	BC005989	P54317	469	10:118,370,808-118,394,595	12 (+)	23,788	5.4	51,988	1-17 [RG-KE]
Human *LIPR3*	*Homo sapiens*	*PLR3* or *PNPLRP3* or *PLRP3*	NM_001011709	BC117224	Q17RR3	467	10:118,177,515-118,226,652	12 (+)	49,138	8.6	52,254	1-17 [RG-KE]
Rhesus *PTL*	*Macaca mulatta*	*PTL* or *PNLIP* or *LIPP*	XP_001095070[1]	na	na	465	9:116,174,800-116,196,544	12 (+)	21,745	6.0	51,241	1-16 [AG-KE]
Rhesus *LIPR1*	*Macaca mulatta*	*PLR1* or *PNPLRP1* or *PLRP1*	XP_001094807[1]	na	na	467	9:116,220,484-116,239,377	12 (+)	18,894	5.4	51,776	1-17 [KG-KE]
Rhesus *LIPR2*	*Macaca mulatta*	*PLR2* or *PNPLRP2* or *PLRP2*	XP_001095293[1]	na	na	469	9:116,251,302-116,274,997	12 (+)	23,696	5.1	51,964	1-17 [RG-KE]
Rhesus *LIPR3*	*Macaca mulatta*	*PLR3* or *PNPLRP3* or *PLRP3*	XM_002805832[1]	na	na	474	9:116,058,416-116,107,264	13 (+)	48,849	8.6	52,916	1-17 [RG-KE]
Mouse *Ptl*	*Mus musculus*	*Ptl* or *Pnlip* or *Lipp*	NM_026925	BC061061	Q6P8U6	465	19:58,745,029-58,756,214	12 (+)	11,186	6.4	51,428	1-16 [AG-RE]
Mouse *LIPR1*	*Mus musculus*	*Plr1* or *Pnlrp1* or *Plrp1*	NM_017844	BC090985	Q5BKQ4	473	19:58,803,615-58,818,615	12 (+)	15,001	5.9	52,696	1-17 [QG-KE]
Mouse *LIPR2*	*Mus musculus*	*Plr2* or *Pnlrp2* or *Plrp2*	NM_011128	M30687	P17892	468	19:58,834,321-58,851,968	12 (+)	17,648	6.2	52,585	1-16 [GG-KE]
Opossum *PTL*	*Monodelphis domestica*	*PTL* or *PNLIP* or *LIPP*	XP_001377457[1]	na	na	465	1:93,547,211-93,565,111	12 (-)	17,901	5.8	51,583	1-16 [AG-KE]
Opossum *LIPR1*	*Monodelphis domestica*	*PLR1* or *PNPLRP1* or *PLRP1*	XP_001368186[1]	na	na	466	1:93,507,478-93,528,824	12 (-)	21,347	5.8	51,630	1-17 [RG-KE]
Opossum *LIPR2*	*Monodelphis domestica*	*PLR2* or *PNPLRP2* or *PLRP2*	XP_001377443[1]	na	na	469	1:93,430,659-93,463,464	12 (-)	32,806	7.9	52,796	1-17 [RG-KE]
Chicken *PTL*	*Gallus gallus*	*PTL* or *PNLIP* or *LIPP*	XP_42778[1]	na	na	467	6:30,327,112-30,337,503	12 (+)	10,392	6.1	51,281	1-17 [RG-SE]
Chicken *LIPR1*	*Gallus gallus*	*PLR1* or *PNPLRP1* or *PLRP1*	XP_001234657[1]	na	na	472	6:30,313,497-30,326,711	12 (+)	13,215	8.3	53,361	1-23 [LG-EE]
Chicken *LIPR2*	*Gallus gallus*	*PLR2* or *PNPLRP2* or *PLRP2*	XP_001234649[1]	na	na	446	6:30,285,522-30,297,237	11 (+)	11,716	9.1	49,496	na
Frog *PTL*	*Xenopus tropicalis*	*PTL* or *PNLIP* or *LIPP*	NM_204010	BC080957	Q28IT6	472	1076:4,143-10,564	12 (-)	6,422	6.2	51,908	1-17 [QG-EP]
Frog *LIPR1*	*Xenopus tropicalis*	*PLR1* or *PNPLRP1* or *PLRP1*	XP_002943432[1]	na	na	467	1394:21,973-29,042	12 [+]	7,070	7.1	51,417	1-17 [KG-GE]
Frog *LIPR2*	*Xenopus tropicalis*	*PLR2* or *PNPLRP2* or *PLRP2*	NP_001072431[1]	BC121677	na	468	1076:63,535-72,134	12 (-)	8,600	5.4	51,131	1-17 [QG-NE]
Seabream *PTL*	*Chrysophrys major*	*PTL* or *PNLIP* or *LIPP*	na	AB252856	Q0EDA5	452	na	na	na	6.5	50,520	1-23 [TA-QR]
Catfish *PTL*	*Ictalurus punctatus*	*PTL* or *PNLIP* or *LIPP*	na	GU589201	E3TFR7	471	na	na	na	5.8	52,732	1-19 [YG-AE]

Table 1: Vertebrate Pancreatic Lipase-like Genes and Proteins. RefSeq: the reference amino acid sequence; [1]predicted Ensembl amino acid sequence; scaffold IDs are identified for frog genome; na- not available; GenBank IDs are derived from the NCBI database http://www.ncbi.nlm.nih.gov/genbank/; Ensembl ID was derived from Ensembl genome database http://www.ensembl.org ; UNIPROT refers to UniprotKB/Swiss-Prot IDs for individual PTL-like enzymes (see http://kr.expasy.org); bps refers to base pairs of nucleotide sequences.

rat (*Rattus norvegicus*), horse (*Equus caballus*) [15], chicken (*Gallus gallus*) [58] and catfish (*Ictalurus punctatus*) [59]. These PTL sequences were chosen as representatives of primate (human and rhesus monkey), eutherian (rat, horse and dog), marsupial (opossum), bird (chicken) and fish (catfish) species. Alignments of human and other vertebrate PTL sequences showed between 55-93% identities, suggesting that these are products of the same family of genes, whereas comparisons of sequence identities of vertebrate PTL proteins with human and mouse hepatic lipase (HL), lipoprotein lipase (LPL) and endothelial lipase (EL) exhibited much lower levels of sequence identities (24-27%), indicating that these are members of distinct lipase families (Table 2).

The amino acid sequences for human, rhesus monkey, rat, horse, dog and opossum PTL contained 465 residues whereas chicken and catfish PTL contained 467 and 471 amino acids, respectively (Figure 1). Previous structural studies on horse and human PTL [13,15] have identified key residues for these vertebrate PTL proteins (sequence numbers refer to human PTL). These included the catalytic triad for the active site (Ser169; Asp193; His280); the hydrophobic N-terminus signal peptides (see also Table 1) which assist secretion of the enzyme from acinar pancreatic cells [5]; disulfide bond forming residues (Cys20/Cys26; Cys106/Cys118; Cys254/Cys278; Cys302/Cys313; Cys316/Cys321; Cys448/Cys465); the predicted 'lid' region (255-277) which covers the active site and participates in lipid substrate binding in analogous lipases [12,15]; and a 'hinge' region for vertebrate PTL which spatially separates the α/β N-terminal hydrolase region from the C-terminal PLAT β-sandwich region (residues 354-465) [12,15]. With the exception of the N-terminus signal peptides, nearly half of these residues (42%) are strictly conserved or underwent conservative substitutions (19%) indicating the essential nature of these residues in contributing to PTL structure and function. In contrast, the N-terminal region (residues 1-16) underwent major changes in the number and sequence of amino acid residues but retained a predicted signal peptide property in each case (Figure 1; Table 1). An N-glycosylation site, previously reported for human and pig PTL at 183Asn-184Gly-185Thr [12], was also predicted for the rhesus and dog PTL sequences (Figure 1). In contrast, this site was absent from the horse PTL sequence which may explain why this enzyme lacks a glycosyl group [60]. Rat, opossum and chicken PTL sequences also lacked this human PTL N-glycosylation site whereas catfish PTL contained three predicted N-glycosylation sites for this enzyme: 55NAS; 306NKT; 394NGT (Figure 1). The significance of these differences in N-glycosylation binding for vertebrate PTL remains to be investigated.

Alignments of vertebrate PTL-like amino acid sequences

Alignments of vertebrate PTL, PLR1 and PLR2, as well as primate PLR3 amino acid sequences examined, showed between 42-93% identities, suggesting that these are products of a conserved gene family during vertebrate evolution (Table 2). In addition, comparisons of mammalian PTL, PLR1, PLR2 and PLR3 sequences also showed higher levels of sequence identities within each of the PTL-like sub-groups of enzymes, suggesting that these are distinct members within the mammalian PTL gene family.

Amino acid sequence alignments for previously reported human (*Homo sapiens*) PTL [56], PLR1 [61], PLR2 [6] and PLR3 [9] and mouse PTL [42,62], PLR1 [61] and PLR2 [26,42] sequences are shown in Figure 2 (also see 1). The amino acid sequence for the human PLR1 subunit contained 467 residues while the mouse PLR1 sequence contained 473 amino acids, due to an extended C-terminus sequence. Human and mouse PLR2 sequences were similar in length to PTL and PLR1 with 469 and 468 amino acids respectively, while human PLR3 contained 467 amino acids. Structural studies of mammalian PLR1 and PLR2 sequences have enabled the identification of key residues for these enzymes identical to or comparable with those previously described for human PTL (see Figure 1). Key residues for each of the human and mouse PTL-like sequences included the following (human PTL sequence numbers were used) (Figure 2): the catalytic triad for the active site (Ser169; Asp193; His280); the predicted N-signal peptide regions (residues 1-16) with a glycine-↓-lysine/arginine peptide bond recognized for hydrolysis in each case; 12 disulfide bond forming residues (Cys20/Cys26; Cys106/Cys118; Cys254/Cys278; Cys302/Cys313; Cys316/Cys321; Cys448/Cys465); distinct predicted N-glycosylation sites, including human PTL at 183Asn-184Gly-185Thr [12], mouse PTL at 293Asn-294Pro-295Thr, mouse PLR1 at 159Asn-

Lipase	Human	Mouse	Frog	Human	Mouse	Frog	Human	Mouse	Frog	Human	Human	Mouse	Human	Mouse	Human	Mouse	Seasquirt
	PTL	PTL	PTL	PLR1	PLR1	PLR1	PLR2	PLR2	PLR2	PLR3	HL	HL	EL	EL	LPL	LPL	Lipase
Human PTL	100	**78**	55	67	67	64	64	68	49	47	25	26	27	27	24	24	35
Mouse PTL	**78**	100	54	67	66	62	62	69	47	45	25	27	26	27	22	24	34
Frog PTL	55	54	100	52	54	60	52	53	53	46	24	27	26	27	24	24	35
Human PLR1	67	67	52	100	**83**	60	62	63	47	48	29	30	29	26	26	25	36
Mouse PLR1	67	66	54	**83**	100	60	62	63	49	49	29	29	27	27	26	26	36
Frog PLR1	64	62	60	60	60	100	58	61	56	47	24	25	28	26	23	23	37
Human PLR2	64	62	52	62	62	58	100	**74**	47	47	27	26	25	25	24	24	36
Mouse PLR2	68	69	53	63	63	61	**74**	100	43	47	29	27	25	27	24	22	32
Frog PLR2	49	47	53	47	49	56	47	43	100	43	25	26	23	25	22	22	32
Human PLR3	47	45	46	48	49	47	47	47	43	100	29	27	22	27	24	22	32
Human HL	25	25	24	29	29	24	27	29	25	29	100	**74**	38	42	44	42	21
Mouse HL	26	27	27	30	29	25	26	27	26	27	**74**	100	42	40	44	43	25
Human EL	27	26	26	29	27	28	25	25	23	25	38	42	100	**80**	44	45	25
Mouse EL	27	27	27	26	27	26	25	27	23	27	42	40	**80**	100	45	46	25
Human LPL	24	22	24	26	26	23	24	24	22	24	44	44	44	45	100	**92**	25
Mouse LPL	24	24	24	25	26	23	24	22	23	22	42	43	45	46	**92**	100	25
Seasquirt Lipase	35	34	35	36	36	37	36	32	32	32	21	25	25	25	25	25	100

Table 2: Percentage Identities for Vertebrate and Sea Squirt Lipase Amino Acid Sequences. Numbers show the percentage of amino acid sequence identities. Numbers in **bold** show higher sequence identities for lipases from the same gene family: PTL; PLR1; PLR2; PLR3; HL, EL and LPL represent genes encoding hepatic lipase, endothelial lipase and lipoprotein lipase, respectively.

Citation: Holmes RS, Cox LA (2012) Bioinformatics and Evolution of Vertebrate Pancreatic Lipase and Related Proteins and Genes. J Data Mining in Genom Proteomics 3:111. doi:10.4172/2153-0602.1000111

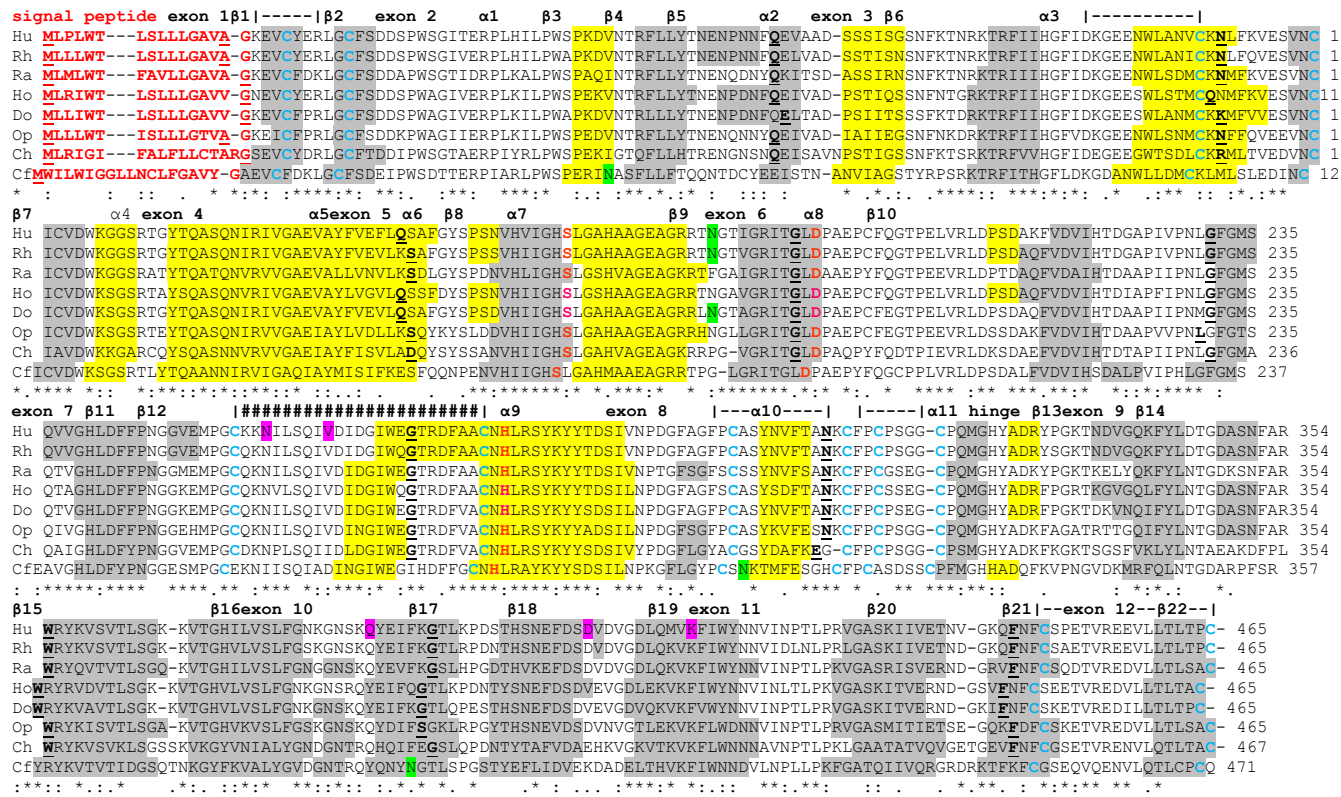

Figure 1: Amino Acid Sequence Alignments for Vertebrate Pancreatic Lipase (PTL) Sequences. See Table 1 for sources of vertebrate PTL sequences; * shows identical residues for proteins; : one similar alternate residue; . two similar alternate residues; key PTL active site residues: Ser169, Asp193 and His280 (human PTL); N-signal peptide residues are in **red**; N-glycosylation residues are in shaded green; five disulphide bonds are represented as |----| with Cys residues in shaded **blue**; β-sheets (β1-β22) are numbered according to human PTL sequences [12] in shaded grey; α-helices (α1-10) are also numbered according to the human PTL structure and are in shaded yellow; **|####|** represents 'lid' region covering active site; colipase binding residues are in purple; hinge refers to region connecting the 'lipase' and 'plat' domains; **bold** underlined font shows residues corresponding to known or predicted exon start sites; exon numbers refer to human *PTL* gene coding exons.

160Tyr-161Ser (none predicted for human PLR1), human PLR2 at 353Asn-354Phe-355Thr and four predicted N-glycosylation sites for human PLR3 at 74Asn-75Ser-76Ser, 125Asn-126Gly-127Ser, 338Asn-339Gly-340Ser and 412Asn-413Ile-414Thr; and residues previously recognized for forming salt bridges between human PTL and colipase [21], including 385Gln, 406Asp and 416Lys, which are retained for the human and mouse PTL and PLR2 sequences but replaced with lysine for the second of these residues for the human and mouse PLR1 sequences and with glutamate and glutamine residues respectively, for the first and third residues of the human PLR3 sequence.

Figure 3 presents amino acid sequence alignments for carboxyl ester lipase PTL, PLR1, PLR2 and primate PLR3 sequences for a region of the PTL active site previously shown to contribute significantly to lipase activity. Human PLR1 has been 'reactivated' by site-directed mutagenesis at two sites corresponding to human PTL residues (Ala196Val and Pro198Ala), representing the active and inactive forms of PLR1, respectively [7,8]. Comparisons of vertebrate PTL, PLR2 and primate PLR3 sequences in this region show that the 'active' Ala196/Pro198 residues were observed in each case, whereas 'inactive' Val196/Ala198 residues were retained for each of the 11 eutherian PLR1 sequences examined. In contrast, the predicted opossum and platypus PLR1 sequences contained Ala196/Ser198 and Ser196/Pro198 residues respectively, while the predicted chicken and frog PLR1 sequences retained the 'active' PTL-like residues, Ala196/Pro198. This suggests that the 'inactive-lipase' PLR1 property has been conserved for all eutherian PLR1 sequences examined, whereas PLR1 sequences from earlier mammalian and vertebrate species may be 'active-lipase' forms of this enzyme.

Supplementary Figure 1 shows amino acid alignments for the 'lid' regions for vertebrate PTL, PLR1, PLR2 and primate PLR3 sequences (residues 255-277 for human PTL) [12]. High levels of sequence identities were observed for each of the vertebrate PTL-like family sequences for this region, indicating that this is a highly conserved region for all vertebrate (PTL, PLR1 and PLR2) and primate (PLR3) sequences. In addition, the 'lid' sequences separated into 2 groups according to their comparative sequence identities: vertebrate PTL and PLR1; and vertebrate PLR and primate PLR3, which are very similar sequences in each case. The surface loop has been shown to cover the active site of human PTL and influences interfacial activation and lipid binding for this enzyme [18]. Different specificities for these properties are likely given the existence of two discrete PTL-like 'loop' sequences for the PTL/PLR1 and PLR2/PLR3 enzymes.

Secondary and tertiary structures for vertebrate PTL-like lipases

Predicted secondary structures for rhesus monkey, rat, dog,

Citation: Holmes RS, Cox LA (2012) Bioinformatics and Evolution of Vertebrate Pancreatic Lipase and Related Proteins and Genes. J Data Mining in Genom Proteomics 3:111. doi:10.4172/2153-0602.1000111

opossum, chicken and catfish PTL sequences were compared with those previously reported for human and horse PTL [12-16] (Figure 1). Alpha-helix and β-sheet structures for these sequences were numbered as for those described for horse PTL [15] which were predominantly retained for all vertebrate PTL sequences examined. Similar predicted secondary structure analyses were also undertaken for human and mouse PTL, PLR1, PLR2 and human PLR3 sequences (Figure 2) which demonstrated that the major structural features previously reported have been retained for each of these predicted sequences.

Figure 4 compares previously reported structures for human PTL [12] and PLR1 (submitted to PDB data bank: 2PPL) protein sequences with predicted tertiary structures for the frog (*Xenopus tropicalis*) PTL and PLR1 subunits. Major features for these proteins, based on previous reports [12-16], demonstrated separation into two major functional domains: a 'lipase' domain containing several parallel β-sheets, three large α-helices and an 'active-site' zone covered by a 'lid' and a 'plat' domain, which are separated by a 'hinge' region. These overall features are readily observed for the predicted frog PTL and PLR1 three-dimensional structures.

Gene locations, exonic structures and comparative tissue expression for vertebrate *PTL*-like genes

Table 1 and Supplementary Table 1 summarize the predicted locations for vertebrate *PTL*-like genes based upon BLAT interrogations of genomes using the reported sequences for human PTL, PLR1, PLR2 and PLR3 [6,9,56, 61] and the predicted sequences for other vertebrate PTL-like proteins and the UC Santa Cruz Genome Browser [49]. The mammalian *PTL*-like genes were predominantly transcribed on the positive strand and located in the following gene order: *PLR3-PTL-PLR1-PLR2* for human (see Supplementary Figure 2), orangutan and rhesus genomes; and *PTL-PLR1-PLR2* for other mammalian genomes, including mouse, rat, dog, pig, rabbit, opossum and platypus genomes (Table 1 and Supplementary Table 1). Horse *PTL*-like genes were however located on the negative strand but in the same gene order as for other eutherian mammalian genomes (Supplementary Table 1). In contrast, *PTL*-like genes were located on chicken chromosome 6 on the positive strand in the reverse order (*PLR2-PLR1-PTL*), which suggests that a chromosomal rearrangement has occurred in the region of the *PTL*-like genes, although in each case, the three genes are proximally localized on mammalian and bird genomes.

Figures 1 and 2 summarize the coding exon start sites for vertebrate *PTL* genes, human *PLR1, PLR2* and *PLR3* genes and mouse *PLR1* and *PLR2* genes showing 12 coding exons in identical or similar positions which are consistent with previous reports [9,63-64]. This is consistent with a common evolutionary origin for these genes. Figure 5 shows the predicted structures of mRNAs for human *PTL, PLR1, PLR2* and *PLR3* transcripts for the major isoform in each case [48]. These human mRNA transcripts varied in length from 18.8 kilobases to 50.1 kilobases for these *PTL*-like genes. The human *PLR2* and *PLR3* gene transcripts contained extended untranslated 3'-UTR regions, which in the case of the *PLR2* transcript, was encoded by 3 distinct exons. The levels of expression for these human *PTL*-like genes have been compared with the average level of gene expression observed in the human genome (see Table 1). Three of these genes exhibited higher levels of

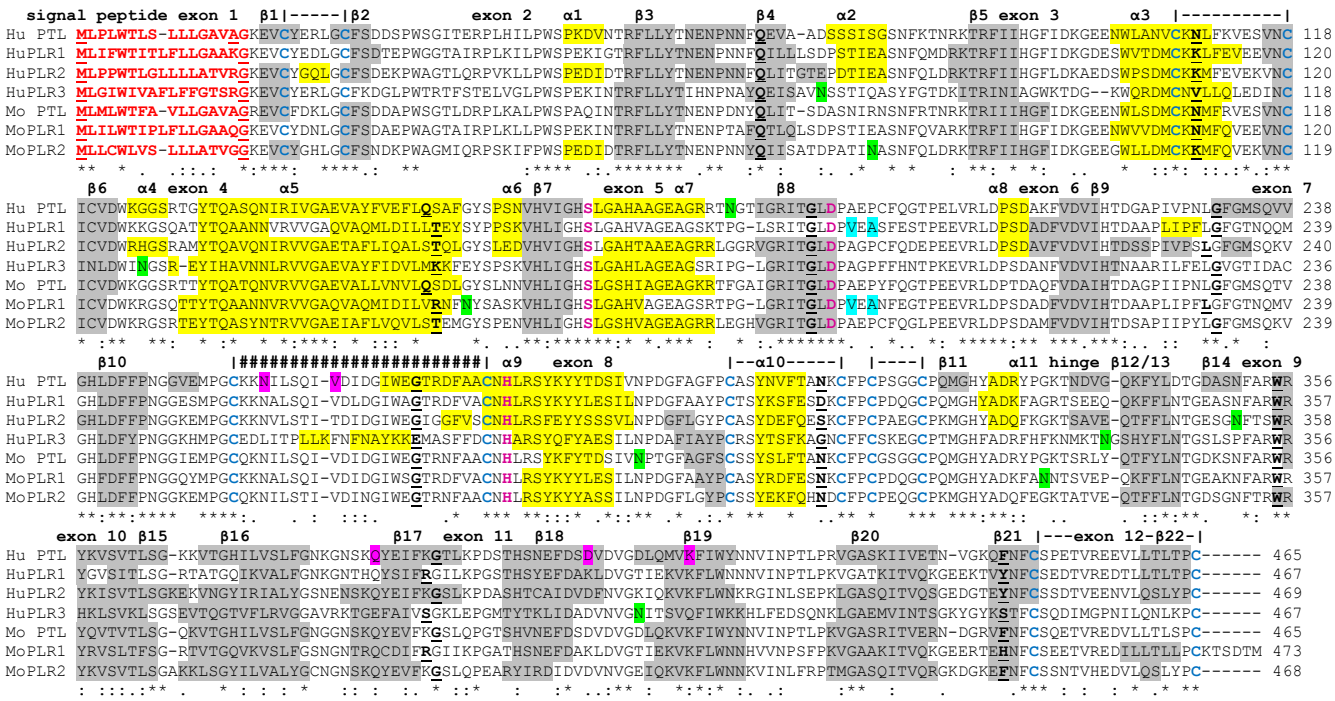

Figure 2: Amino Acid Sequence Alignments for Human and Mouse Pancreatic Lipase-like (PTL-like) Sequences. See Table 1 for sources of human and mouse PTL, PLR1, PLR2 and human PLR3 sequences; * shows identical residues for proteins; : one similar alternate residue; . two similar alternate residues; key PTL active site residues: Ser169, Asp193 and His280 (human PTL); N-signal peptide residues are in **red**; predicted or known N-glycosylation residues are in shaded green; five disulphide bonds are represented as |----| with Cys residues in shaded **blue**; β-sheets (β1-β22) are numbered according to human PTL sequences [12] in shaded grey; α-helices (α1-10) are also numbered according to the human PTL structure and are in shaded yellow; **|####|** represents 'lid' region covering active site; colipase binding residues are in pink; hinge refers to region connecting the 'lipase' and 'plat' domains; **bold** underlined font shows residues corresponding to known or predicted exon start sites; exon numbers refer to human *PTL* gene coding exons.

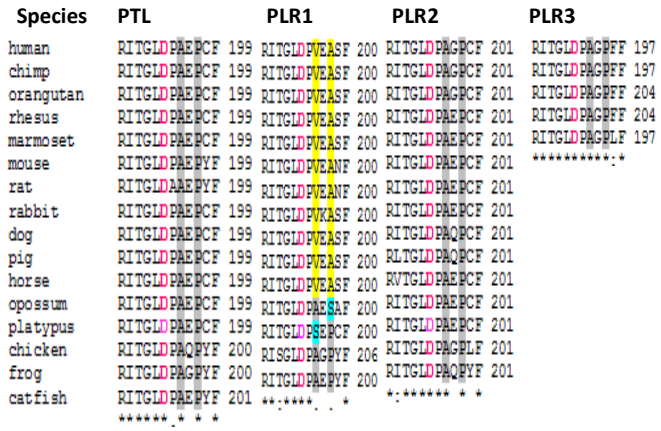

Figure 3: Amino Acid Sequence Alignments for Vertebrate Pancreatic Lipase-like (PTL-like) Sequences Containing the Active Aspartate Residue. See Table 1 and Supplementary Table 1 for sources of vertebrate PTL, PLR1, PLR2 and primate PLR3 sequences; * shows identical residues for proteins; : one similar alternate residue; . two similar alternate residues; key PTL active site residue (Asp193 for human PTL) is shown in pink ; shaded grey shows human PTL Ala195 and Pro197 residues for vertebrate PTL-like sequences; shaded yellow shows human PLR1 Val 196 and Ala198 mammalian PLR1-like sequences; shaded blue shows substituted Ser196 or Ser198 residues for opossum PLR1 and platypus PLR1, respectively.

expression [*PTL* (x1.8); *PLR1* (x1.5); *PLR2* (x1.5)], whereas human *PLR3* exhibited a lower than average level of expression (x0.2). Higher levels of gene expression were also observed for the mouse PTL-like genes: *Ptl* (x4.5 times); *Plr1* (x2.9); and *Plr2* (1.6) [48]. These high levels of gene expression are consistent with the major roles for these genes and encoded enzymes in lipid digestion in the body. Figure 6 presents 'heat maps' showing comparative gene expression for various human tissues obtained from GNF Expression Atlas Data using GNF1M chips [64] (http://genome.ucsc.edu; http://biogps.gnf.org). These data supported differential tissue expression for human *PTL*-like genes, with *PTL*, *PLR1* and *PLR2* showing highest levels in pancreatic islet tissue, which is consistent with previous reports for these genes (Figure 5). In contrast, the human *PLR3* gene was most highly expressed in bronchial epithelial cells with an as yet unknown function in the body.

Phylogeny and divergence of PTL-like sequences

A phylogenetic tree (Figure 7) was calculated by the progressive alignment of 55 vertebrate PTL-like amino acid sequences with twelve vertebrate hepatic lipase (HL), lipoprotein lipase (LPL) and endothelial lipase (EL) amino acid sequences. A sea squirt (*Ciona intestinalis)* lipase served as a 'root' for the vertebrate and invertebrate PTL-like sequences, and a fruit fly (*Drosophila melanogaster)* lipase sequence (CG5966) served as the 'root' for the tree overall (see Table 1 and Supplementary Table 1). The phylogram showed clustering of the lipase sequences into several groups, including mammalian PTL, PLR1 and PLR2 sequences; primate PLR3 sequences; chicken (*Gallus gallus*) and lizard (*Anolis carolinensis*) PTL sequences; frog (*Xenopus tropicalis*) PTL, PLR1 and PLR2 sequences; fish (seabream: *Chrysophrys major*; catfish: *Ictalurus punctatus*) PTL sequences; as well as distinct groups for the three neutral lipases: HL, LPL and EL. In addition, the phylogram suggested a sequence of gene duplication events for an ancestral *PTL*-like gene during vertebrate evolution (see Figure 7): (1) an ancestral *PTL* gene duplication generating the *PTL/PLR1* and *PLR2* precursor genes; (2) duplication of the vertebrate *PTL/PLR1* precursor gene to form the *PTL* and *PLR1* genes; and (3) duplication of the *PLR2* gene prior to primate evolution to form the *PLR2* and *PLR3* genes, currently observed in primate genomes. It is also apparent that vertebrate neutral lipase genes (*HL*; *EL*; and *LPL*) have been generated during vertebrate evolution from a distinct ancestral gene, which is consistent with previous studies on the enzymes and genes [33-37].

Conclusions

These results demonstrate that mammalian and other vertebrate pancreatic-like (*PTL*) genes and encoded enzymes comprise at least four distinct forms, designated as PTL, PLR1, PLR2 and PLR3, which is consistent with previous studies [5,9,31]. PTL functions in the presence of colipase in the hydrolysis of emulsified fats in the small intestine following secretion from the pancreas [1-5]; PLR1 is also found in pancreatic secretions but has an as yet unknown function and exhibits no detectable triacylglycerol, cholesterol ester or galactolipase activities [6,8]; PLR2 functions as a pancreatic secretion galactolipase but may also contribute to T cell cytotoxicity in the body [6,25,65]; while PLR3 serves a related, but as yet unknown, lipase function.

PTL, PLR1 and PLR2 were each encoded by single genes among most vertebrate genomes. These genes are highly expressed in human and mouse pancreatic islet cells, which is consistent with their roles in lipid digestion in the intestine. In contrast, *PLR3* was expressed at lower than average levels for human genes overall and was highly expressed in bronchial epithelial cells. Predicted secondary and tertiary structures for frog PTL and PLR1 proteins showed similarities with human and pig PTL [12-16] with two major functional domains: a 'lipase' domain containing several parallel β-sheets, three large α-helices and an 'active-site' zone covered by a 'lid' and a 'plat' domain separated by a 'hinge' region. Comparisons of PTL, PLR1, PLR2 and PLR3 amino acid sequences from vertebrates representative of mammals, birds, amphibians and bony fish, demonstrated that these are highly conserved proteins during evolution, not only for active site

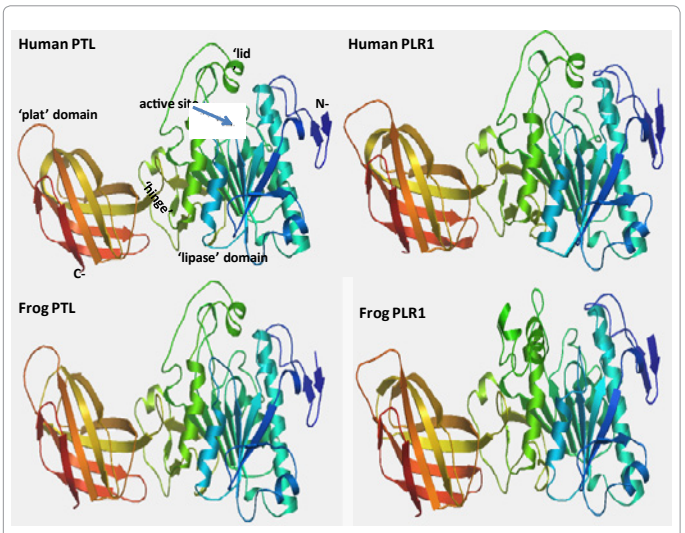

Figure 4: Tertiary Structures for Human PTL and PLR1 and Frog PTL and PLR1 Sequences. Tertiary structures for human PTL and PLR1 were obtained from Winkler et al. [12] and from a submitted report to the PDB data bank (2PPL) respectively, and using SWISS MODEL methods for the frog PTL and PLR1 sequences; the rainbow color code describes the known tertiary structures from the N- (blue) to C-termini (red color); arrows indicate directions for β-sheets; known or active site for human PTL; N-terminal and C-terminal regions are shown, as are predicted structures and locations for 'lipase' and 'plat' domains; the 'lid' covering the active site; and the 'hinge' separating the 2 domains.

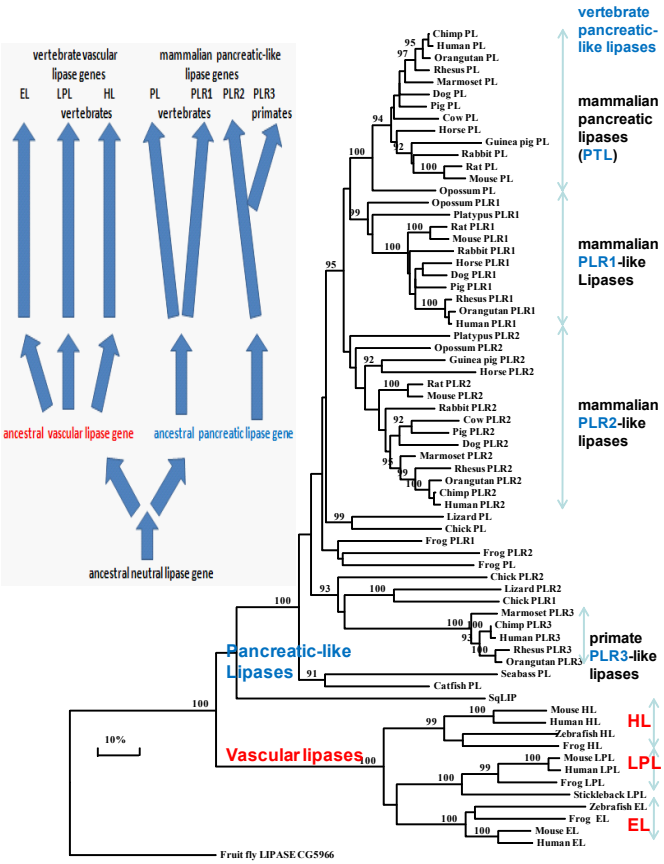

Figure 5: Gene Structures and Major Splicing Variant for Human *LIPP, LIPR1, LIPR2* and *LIPR3* Gene Transcripts. Derived from AceView [12]; mature isoform variants (a) are shown with capped 5'- and 3'- ends for the predicted mRNA sequences; NM refers to the NCBI reference sequence; exons are in shaded pink; untranslated 5'- and 3' sequences are in open pink; introns are represented as pink lines joining exons; the directions for transcription are shown as 5' → 3'; sizes of mRNA sequences are shown in kilobases (kb); comparative gene expression levels with the average human gene are shown.

residues but also for those involved in lipid binding ('plat' domain) and shielding the active site from the aqueous environment near the active site (the 'lid'). Vertebrate PTL, PLR1, PLR2 and PLR3 (the latter enzyme was restricted to primates) shared between 42-93% sequence identities but <30% with the vascular lipases, HL (hepatic lipase), EL (endothelial lipase) and LPL (lipoprotein lipase), demonstrating that the vertebrate PTL-like lipases are a distinct family of related genes and proteins. Sequence alignments, key amino acid residues and conserved predicted secondary and tertiary structures were examined, including active site residues; disulfide bonds; predicted N-glycosylation and signal peptide sites; 'plat' and 'lid' domains; and a 'hinge' region.

Phylogeny studies of vertebrate *PTL, PLR1, PLR2* and *PLR3* genes and enzymes suggested that they originated in a vertebrate ancestor from successive gene duplication events of an ancestral *PTL*-like gene, forming initially two precursor genes (*PTL/PLR1* and *PLR2*); followed by duplication of the former precursor gene to form the *PTL* and *PLR1* genes; and (3) duplication of the *PLR2* gene prior to primate evolution to form the *PLR2* and *PLR3* genes in primate genomes.

Overall, these results significantly contribute to our knowledge concerning the evolution of pancreatic lipase-like genes, enzymes and proteins, particularly with respect to comparative amino acid sequences, predicted secondary and tertiary structures for vertebrate PTL (pancreatic lipase), PLRI (pancreatic lipase-like protein 1), PLR2 (pancreatic lipase-like protein 2) and PLR3 (pancreatic lipase-like protein 3), comparative genomic structures for vertebrate *PTL*-like genes and the likely sequence of gene duplication events which have generated this family of *PTL*-like genes within vertebrate genomes. PTL and PLR2 amino acid sequences were predominantly conserved among the representative vertebrate species examined which is consistent

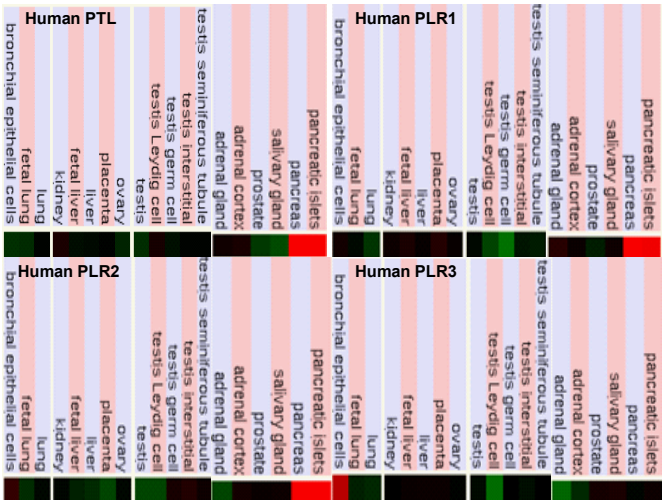

Figure 6: Comparative Tissue Expression for Human *PTL*-like Genes (*PTL, PLR1, PLR2* and *PLR3*). Expression 'heat maps' (GNF Expression Atlas 2 data) [64] were examined for comparative gene expression levels among human tissues for *PTL, PLR1, PLR2* and *PLR3* genes showing high (red); intermediate (black); and low (green) expression levels; derived from human genome browser.

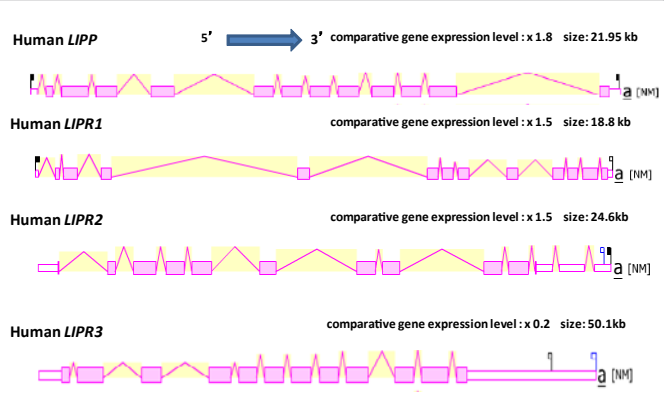

Figure 7: Phylogenetic Tree of Vertebrate PTL-like and Vascular Lipase Sequences and Invertebrate Lipase Amino Acid Sequences. The tree is labeled with the lipase name and the name of the animal and is 'rooted' with the fruit fly (*Drosophila melanogaster*) lipase sequence. Note the 4 major clusters of vertebrate PTL-like sequences corresponding to the *PTL, PLR1, PLR2* and *PLR3* gene families and the separate cluster of vertebrate vascular lipases, divided into 3 gene families: HL (hepatic lipase); EL (endothelial lipase); and LPL (lipoprotein lipase). A genetic distance scale is shown (% amino acid substitutions). The number of times a clade (sequences common to a node or branch) occurred in the bootstrap replicates are shown. Only replicate values of 90 or more which are highly significant are shown with 100 bootstrap replicates performed in each case. A proposed sequence of gene duplication events for neutral lipase genes during invertebrate and vertebrate evolution is shown.

with their major roles within pancreatic secretions in the digestion of emulsified fats in the duodenum, and in the case of PLR2, the digestion of plant lipids by serving as a galactolipase. In contrast, PLR3 genes and proteins were restricted to primates, serve an as yet unknown function(s) in the body and exhibit highest expression levels in human bronchial epithelial cells. Comparisons of vertebrate PTL, PLR2 and primate PLR3 sequences showed that the 'catalytically active' PL-like sequence (Ala196/Pro198) was observed for all species examined whereas a 'catalytically inactive' Val196/Ala198 PL-like sequence was observed for all eutherian mammalian PLR1 sequences examined. In contrast, chicken and frog PLR1 sequences retained 'active' PTL-like residues, Ala196/Pro198, whereas marsupial (opossum) and monotreme (platypus) PLR1 sequences contained Ala196/Ser198 and Ser196/Pro198 residues respectively. This suggests that the 'inactive-lipase' PLR1 form, which plays a specialized metabolic inhibitor role [22], may function only in eutherian mammals, whereas PLR1 sequences from earlier mammalian and vertebrate species are 'active-lipase' forms of this enzyme and may serve as additional forms of PTL-like lipases in these species. Further biochemical and genetic studies are needed to describe any specific role(s) for PLR1 in lower vertebrate species.

The bioinformatic methodologies used in this investigation of pancreatic lipase-like genes and proteins may be also readily applied to other pancreatic proteins as well as other gene families encoding enzymes and proteins, including neutral lipases [35-37], acid lipases [66], carboxylesterases [67], enolases [68] and other gene families [69-71].

Acknowledgement

This project was supported by NIH Grants P01 HL028972 and P51 RR013986. In addition, this investigation was conducted in facilities constructed with support from Research Facilities Improvement Program Grant Numbers 1 C06 RR13556, 1 C06 RR15456, 1 C06 RR017515. The expert assistance and advice of Dr Bharet Patel of Griffith University Australia is gratefully acknowledged.

Disclosure

The authors report no conflicts of interest in this work.

References

1. De Caro A, Figarella C, Amic J, Michel R, Guy O (1977) Human pancreatic lipase: A glycoprotein. Biochim Biophys Acta 490: 411–419.
2. Verger R (1984) Pancreatic lipase: In Lipases. Borgstrom B and Brockman H L, editors, Elsevier, Amsterdam. 84–150.
3. Payne RM, Sims HF, Jennens ML, Lowe ME (1994) Rat pancreatic lipase and two related proteins: enzymatic properties and mRNA expression during development. Am J Physiol 266: G914–G921.
4. Lowe ME (1997) Structure and function of pancreatic lipase and colipase. Ann Rev Nutr 17: 141-158.
5. Lowe ME (2002) The triglyceride lipases of the pancreas. J Lipid Res 43: 2007-2016.
6. Giller T, Buchwald P, Blum-Kaelin D, Hunziker W (1992) Two novel human pancreatic lipase related proteins, hPLRP1 and hPLRP2: differences in colipase dependency and in lipase activity. J Biol Chem 267: 16509–16516.
7. Roussel A, deCaro J, Bezzine S, Gastinel L, de Caro A, et al (1998) Reactivation of the totally inactive pancreatic lipase RP1 by structure-predicted point mutations. Proteins 32: 523–531.
8. Crenon I, Foglizzo E, Kerfelec B, Verine A, Pignol D, et al. (1998) Pancreatic lipase-related protein type I: a specialized lipase or an inactive enzyme. Protein Eng 11: 135–142.
9. Deloukas P, Earthrowl ME, Grafham DV, Rubenfield M, French L, et al. (2004) The DNA sequence and comparative analysis of human chromosome 10. Nature 429: 375–381.
10. Kirschgessner, TG, Chuat JC, Heinzmann C, Etienne J, Guilhot S, et al (1989) Organisation of the human lipoprotein lipase gene and evolution of the lipase gene family. Proc Natl Acad Sci USA 86: 9647-9651.
11. Thirstrup K, Verger R, Carriere F (1994) Evidence for a pancreatic lipase subfamily with new kinetic properties. Biochemistry 33: 2748–2756.
12. Winkler FK, D'Arcy A, Hunziker W (1990) Structure of human pancreatic lipase. Nature 343: 771–774.
13. van Tilbeurgh H, Sarda L, Verger R, Cambillau C (1992) Structure of the pancreatic lipase-procolipase complex. Nature 359: 159–162.
14. Van Tilbeurgh H, Egloff MP, Martinez C, Rugani N, Verger R, et al. (1993) Interfacial activation of the lipase-procolipase complex by mixed micelles revealed by x-ray crystallography. Nature 362: 814–820.
15. Bourne Y, Martinex C, Kerfelec B, Lombardo D, Chapus C, et al. (1994) Horse pancreatic lipase. The crystal structure refined at 2.3 A resolution. J Mol Biol 238: 709–732.
16. Egloff MP, Marguet F, Buono G, Verger R, Cambillau C, et al. (1995) The 2.46 A resolution of the pancreatic lipase-colipase complex inhibited by a C11 alkyl phosphonate. Biochemistry 24: 2751–2762.
17. Ollis DL, Cheah E, Cygler M, Dijkstra B, Frolow F, et al. (1992) The α/β hydrolase fold. Protein Eng 5: 197–211.
18. Jennens M L, Lowe ME (1994) A surface loop covering the active site of human pancreatic lipase influences interfacial activation and lipid binding. J Biol Chem 269: 25470–25474.
19. Freie AB, Ferrato F, Carrière F, Lowe ME (2006) Val-407 and Ile-408 in the beta5'-loop of pancreatic lipase mediate lipase-colipase interactions in the presence of bile salt micelles. J Biol Chem 281: 7793-800.
20. Bourbon-Freie A, Dub RE, Xiao X, Lowe ME (2009) Trp-107 and trp-253 account for the increased steady state fluorescence that accompanies the conformational change in human pancreatic triglyceride lipase induced by tetrahydrolipstatin and bile salt. J Biol Chem 284: 14157-14164.
21. Bezzine S, Ferrato F, Ivanova MG, Lopez V, Verger R, et al. (1999) Human pancreatic lipase: colipase dependence and interfacial binding of lid domain mutants. Biochemistry 38: 5499–5510.
22. Ren J, Chen Z, Zhang W, Li L, Sun R, et al. (2011) Increased fat mass and insulin resistance in mice lacking pancreatic lipase-related protein 1. J Nutr Biochem 22: 691-698.
23. Hjorth A, Carriere F, Cudrey C, Woldike H, Boel E, et al. (1993) A structural domain (the lid) found in pancreatic lipase is absent in the guinea pig (phospho) lipase. Biochemistry 32: 4702–4707.
24. Reboul E, Berton A, Moussa M, Kreuzer C, Crenon I, et al. (2006) Pancreatic lipase and pancreatic lipase-related protein 2, but not pancreatic lipase-related protein 1, hydrolyze retinyl palmitate in physiological conditions. Biochim Biophys Acta 1761: 4-10.
25. Eydoux C, De Caro J, Ferrato F, Boullanger P, Lafont D, et al. (2007) Further biochemical characterization of human pancreatic lipase-related protein 2 expressed in yeast cells. J Lipid Research 48: 1539-1549.
26. Li X, Lindquist S, Lowe M, Noppa L, Hernell O (2007) Bile salt-stimulated lipase and pancreatic lipase-related protein 2 are the dominating lipases in neonatal fat digestion in mice and rats. Pediatric Res 62: 537-541.
27. D'Agostino D, Lowe ME (2004) Pancreatic lipase-related protein 2 is the major colipase-dependent pancreatic lipase in suckling mice. J Nutr 134: 132-134.
28. Xiao X, Mukherjee A, Ross LE, Lowe ME (2011) Pancreatic lipase-related protein-2 (PLRP2) can contribute to dietary fat digestion in human newborns. J Biol Chem 286: 26353-26363.
29. Withers-Martinez C, Carrière F, Verger R, Bourgeois D, Cambillau C (1996) A pancreatic lipase with a phospholipase A1 activity: crystal structure of a chimeric pancreatic lipase-related protein 2 from guinea pig. Structure 4: 1363-1374.
30. Lowe ME (2000) Properties and function of pancreatic lipase related protein 2. Biochimie 82: 1–8.
31. Sias B, Ferrato F, Grandval P, Lafont D, Boullanger P, et al (2004) Human pancreatic lipase-related protein 2 is a galactolipase. Biochemistry 43: 10138–10148.
32. Saelee P, Wongkham S, Puapairoj A, Khuntikeo N, Petmitr S, et al. (2009)

Novel PNLIPRP3 and DOCK8 gene expression and prognostic implications of DNA loss on chromosome 10q25.3 in hepatocellular carcinoma. Asian Pac J Cancer Prev. 10: 501-506.

33. Hide WA, Chan L, Li W-H. (1992) Structure and evolution of the lipase superfamily. J Lipid Res 33: 167-178.

34. Carriere F, Thirstrup K, Hjorth S, Ferrato F, Nielsen PF, et al. (1997) Pancreatic lipase structure-function relationships by domain exchange. Biochemistry 36: 239–248.

35. Holmes, RS, VandeBerg JL, Cox LA (2011) Vertebrate endothelial lipase: comparative studies of an ancient gene and protein in vertebrate evolution. Genetica 139: 291-304.

36. Holmes RS, VandeBerg JL, Cox LA (2011) Comparative studies of vertebrate lipoprotein lipase: a key enzyme of very low density lipoprotein metabolism. Comp Biochem Physiol D Genomics Proteomics 6: 224-234.

37. Holmes RS, VandeBerg JL, Cox LA (2011) Vertebrate hepatic lipase genes and proteins: a review supported by bioinformatic studies. Open Access Bioinformatics 3: 85-95.

38. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC et al. (2001) Initial sequencing and analysis of the human genome. Nature 409: 860-921.

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

40. Elsik CG, Tellam RL, Worley KC, Gibbs RA, Muzny DM et al. (2009) The genome sequence of taurine cattle: A window to ruminant biology and evolution. Science 324: 522-528.

41. Wade CM, Giulotto, E, Sigurdsson S, Zoli M, Gnerre S et al. (2009) Genome sequence, comparative analysis and population genetics of the domestic horse. Science 326: 865-867.

42. Mouse Genome Sequencing Consortium (2002) Initial sequencing and comparative analysis of the mouse genome. Nature 420: 520-562.

43. Rat Genome Sequencing Project Consortium (2004) Genome sequence of the Brown Norway rat yields insights into mammalian evolution. Nature 428: 493-521.

44. Mikkelsen TS, Wakefield MJ, Aken B, et al. (2007) Genome of the marsupial *Monodelphis domestica* reveals innovation in noncoding sequences. Nature 447: 167-175.

45. International Chicken Genome Sequencing Consortium (2004) Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature 432: 695-716.

46. Hellsten U, Harland RM, Gilchrist LJ, Hendrix D, Jurka J, et al. (2010) The genome of the western clawed frog Xenopus tropicalis. Science 328: 633-636.

47. Sprague J, Bayraktaroglu L, Clements D, Sprague J, Bayraktaroglu L, et al. (2005) The zebrafish information network: the zebrafish model organism database. Nucleic Acids Res. 34: D581-D585.

48. Thierry-Mieg D, Thierry-Mieg J (2006) AceView: A comprehensive cDNA-supported gene and transcripts annotation. Genome Biol 7: S12.

49. Kent WJ, Sugnet CW, Furey TS (2002) The human genome browser at UCSC. Genome Res. 12: 994-1006.

50. Larkin MA, Blackshields G, Brown NP, et al. (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23: 2947-2948.

51. Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.

52. Kimura M. (1983) The Neutral Theory of Molecular Evolution. Cambridge University Press.

53. Van De Peer Y, de Wachter R (1994) TreeCon for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. Comput Appl Sci 10: 569–570.

54. Saitou N, Nei M (1987) The neighbour-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4: 406–425.

55. Felsenstein J. (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39: 783–791.

56. Lowe ME, Rosenblum JL, Strauss AW (1989) Cloning and characterization of human pancreatic lipase cDNA. J Biol Chem 264: 20042-20048.

57. Fendri A, Frikha F, Mosbah H, Miled N, Zouari N, et al. (2006) Biochemical characterization, cloning and molecular modeling of chicken pancreatic lipase. Arch Biochem Biophys 451: 149–159.

58. Chen F, Lee Y, Jiang Y, Wang S, Peatman E, et al. (2010) Identification and characterization of full-length cDNAs in channel catfish (*Ictalurus punctatus*) and blue catfish (*Ictalurus furcatus*). PLoS One 5: e11546.

59. Bianchetta JD, Bidaud J, Guidoni AA, Bonicel JJ, Rovery M (1979) Porcine pancreatic lipase. Sequence of the first 234 amino acids of the peptide chain. Eur J Biochem 97: 395-405.

60. De Caro J, Carrière F, Barboni P, Giller T, Verger R, et al. (1998) Pancreatic lipase-related protein 1 (PLRP1) is present in the pancreatic juice of several species. Biochim Biophys Acta 1387: 331-341.

61. van Bennekum AM, Fisher EA, Blaner WS, Harrison EH (2000) Hydrolysis of retinyl esters by pancreatic triglyceride lipase. Biochemistry 39: 4900-4906.

62. Sims HF, Jennens ML, Lowe ME (1993) The human pancreatic lipase-encoding gene: structure and conservation of an Alu sequence in the lipase gene family. Gene 131: 281-285.

63. MGC Project Team (2004) The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC). Genome Res 14: 2121-2127.

64. Su AI, Wiltshire T, Batalov S, Lapp H, Ching KA, et al. (2004) A gene atlas of the mouse and human protein encoding transcriptomes. Proc Natl Acad Sci USA 101: 6062-6067.

65. Alves BN, Leong J, Tamang DL, Elliott V, Edelnant J, et al. (2009) Pancreatic lipase-related protein 2 (PLRP2) induction by IL-4 in cytotoxic T lymphocytes (CTLs) and reevaluation of the negative effects of its gene ablation on cytotoxicity. J Leuk Biol 86: 701-712.

66. Holmes RS, Cox LA, VandeBerg JL (2010) Comparative studies of mammalian acid lipases: evidence for a new gene family in mouse and rat. Comp Biochem Physiol D 5:217-226.

67. Holmes RS, Wright MW, Laulederkind SJF, Cox LA, Hosokawa M, et al. (2010) Recommended nomenclature for five mammalian carboxylesterase gene families: human, mouse and rat genes and proteins. Mammal Genome 21:427–441.

68. Holmes RS (2011) Bioinformatic studies of vertebrate enolases: multifunctional genes and proteins. Open Access Bioinform 3:43–59.

69. Holmes RS, Goldberg E (2009) Computational analyses of mammalian lactate dehydrogenases: human, mouse, opossum and platypus LDHs. Comp Biol Chem 33:379-385.

70. Holmes RS (2009) Opossum aldehyde dehydrogenases. Evidence for four *ALDH1A1*-like genes on chromosome 6 and *ALDH1A3* genes on chromosome 1. Biochem Genet 47: 609-624.

71. Holmes RS, Rout UK (2011) Comparative studies of vertebrate beta integrin genes and proteins: Ancient genes in vertebrate evolution. Biomolecules 1:3-31.