

Genomics and proteomics of vertebrate cholesterol ester lipase (LIPA) and cholesterol 25-hydroxylase (CH25H)

Author

Holmes, Roger S, VandeBerg, John L, Cox, Laura A

Published

2011

Journal Title

3 Biotech

DOI

[10.1007/s13205-011-0013-9](https://doi.org/10.1007/s13205-011-0013-9)

Rights statement

© 2011 Springer Berlin / Heidelberg. This is an electronic version of an article published in 3 Biotech, Vol. 1(2), pp. 99-109, 2011. 3 Biotech is available online at: <http://www.springerlink.com/> with the open URL of your article.

Downloaded from

<http://hdl.handle.net/10072/42456>

Griffith Research Online

<https://research-repository.griffith.edu.au>

Genomics and Proteomics of Closely Linked Vertebrate Genes and Enzymes of Cholesterol Metabolism: Cholesterol Ester Lipase (*LIPA*) and Cholesterol 25-Hydroxylase (*CH25H*)

Roger S Holmes^{1,4}, Laura A Cox^{1,2} and John L VandeBerg^{1,2}

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

⁴Corresponding Author:

Roger S Holmes, D.Sc.

Department of Genetics

Southwest National Primate Research Center

Southwest Foundation for Biomedical Research

San Antonio, TX, USA 78227

Email: rholmes@sfbgenetics.org

Phone: 210-258-9687

Fax: 210-258-9600

Keywords: Vertebrates; lipase A; Cholesterol 25 hydroxylase; cholesterol metabolism

Running Head: Vertebrate *LIPA* and *CH25H* Genes and Proteins

Summary

Objective. To examine the comparative genomics and proteomics of vertebrate cholesterol ester lipase (*LIPA*) and cholesterol 25-hydroxylase (*CH25H*) genes and proteins. *Background.* *LIPA* (E.C.3.1.1.13) and *CH25H* (E.C.1.14.99.48) play essential roles in cholesterol metabolism in the body: *LIPA* hydrolyses cholesteryl esters and triglycerides within lysosomes and *CH25H* catalyses the formation of 25-hydroxycholesterol from cholesterol which acts to repress cholesterol biosynthesis. *Methods.* *In silico* techniques were used to predict the amino acid sequences, structures and genomic features of several vertebrate *LIPA* and *CH25H* genes and proteins and to examine the phylogeny of vertebrate *LIPA*. *Results.* Amino acid sequence alignments and predicted subunit structures enabled the identification of key sequences previously reported for human *LIPA* and *CH25H* and transmembrane structures for vertebrate *CH25H* sequences. Vertebrate *LIPA* and *CH25H* genes were located in tandem on all vertebrate genomes examined and showed several predicted transcription factor

binding sites and CpG islands located within the 5' regions of the human genes. Vertebrate *LIPA* genes contained coding 9 exons while all vertebrate *CH25H* genes were without introns. A phylogenetic analysis demonstrated the distinct nature of the vertebrate *LIPA* gene and protein family in comparison with other vertebrate acid lipases. *Conclusions.* *LIPA* and *CH25H* genes are located in tandem on vertebrate genomes examined and encode proteins that are conserved throughout vertebrate evolution and perform essential roles in cholesterol metabolism. The *LIPA* family of acid lipases is distinct from other acid lipases and has apparently evolved from an ancestral *LIPA* gene which predated the appearance of vertebrates.

1. Background/Review of the Literature

Lysosomal acid lipase or cholesteryl ester hydrolase (also called lipase A or *LIPA*) (E.C.3.1.1.13) catalyses the hydrolysis of cholesterol esters or triglycerides and has been localized within lysosomes following a receptor-mediated endocytosis of low-density lipoprotein (LDL) particles [1-3]. Inborn errors of metabolism for the human gene encoding this enzyme (*LIPA*) have been described, including Wolman disease (WOD), resulting from a major defect of the gene which leads to a cholesteryl ester storage disease and loss of life, usually within one year of age while a second defect of the human *LIPA* gene generates a milder late-onset cholesteryl ester storage disease (CESD) [4-6].

LIPA is localized on chromosome 10 of the human genome and highly expressed throughout the body, and contains 9 coding exons [7-9]. Several other acid lipase genes, including *LIPF* (encoding gastric triacylglycerol lipase); *LIPJ* (encoding lipase J); *LIPK*; *LIPM* and *LIPN* (encoding epidermis acid lipases K, M and N), are also located within an acid lipase gene cluster on human chromosome 10 [10-12]. This acid lipase gene cluster encode enzymes with similar sequences which are distinct from the 'neutral lipases', including endothelial lipase (LIPE), lipoprotein lipase (LPL) and hepatic lipase (HL), which perform specific roles in high-density lipoprotein (HDL), LDL and hepatic lipid metabolism, respectively [13-18].

Cholesterol 25-hydroxylase (*CH25H* or cholesterol 25- monooxygenase) (EC 1.14.99.38) catalyses 25-hydroxycholesterol biosynthesis from cholesterol, which may serve as a corepressor of cholesterol biosynthetic enzymes by blocking sterol regulatory element binding protein processing [19]. 25-Hydroxysterol is also an activator of gene signalling pathways and an immunoregulatory lipid produced by macrophages to negatively regulate the adaptive immune response in mice [20, 21]. *CH25H* is a member of an enzyme family that utilizes diiron cofactors to catalyse the hydroxylation of sterol substrates, is encoded by an intronless gene (*CH25H*)

located proximally to *LIPA* on human chromosome 10 and is an integral membrane protein located in the endoplasmic reticulum of liver and many other tissues of the body [11, 19].

This study describes the predicted sequences, structures and phylogeny of several mammalian and other vertebrate *LIPA* and *CH25H* genes and compares these results for those previously reported for human (*Homo sapiens*) and mouse (*Mus musculus*) *LIPA* and *CH25H*. *In silico* methods were used to predict the sequences and structures for vertebrate *LIPA* and *CH25H* and gene locations for these genes, using data from the respective genome sequences. Phylogenetic analyses also describe the relationships and potential origins of vertebrate *LIPA* genes during mammalian and vertebrate evolution in comparison with other acid lipase genes.

2. Materials and Methods

2.1 *In silico* vertebrate lipase and cholesterol 25-hydroxylase 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>) [22]. Non-redundant protein sequence databases for several vertebrate genomes were examined using the blastp algorithm, including the chimpanzee (*Pan troglodytes*) [23], macaque monkey (*Mucaca mulatta*) [24], horse (*Equus caballus*) [25], cow (*Bos Taurus*) [26]; mouse (*Mus musculus*) [27]; rat (*Rattus norvegicus*) [28]; guinea pig (*Cavia porcellus*) [29]; dog (*Canis familiaris*) [30]; chicken (*Gallus gallus*) [31]; and frog (*Xenopus tropicalis*) [32]. 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 *LIPA*- and *CH25H*-like proteins. These were derived from annotated genomic sequences using the gene prediction method: GNOMON and predicted sequences with high similarity scores for many of the vertebrate *LIPA* and *CH25H* genes and proteins examined (see Table 1). The orangutan (*Pongo abelii*) and marmoset (*Callithrix jacchus*) genomes was subjected to BLAT (BLAST-Like Alignment Tool) *in silico* analysis using the human *LIPA* protein sequence and the UC Santa Cruz genome browser [<http://genome.ucsc.edu/cgi-bin/hgBlat>] with the default settings to obtain an Ensembl generated protein sequence using the methods of Hubbard and coworkers [33] (<http://www.ensembl.org/index.html>).

BLAT analyses were then undertaken for each of the predicted *LIPA* and *CH25H* amino acid sequences using the UC Santa Cruz web browser [<http://genome.ucsc.edu/cgi-bin/hgBlat>] [34] with the default settings to obtain the predicted locations for each of the vertebrate *LIPA* and *CH25H* genes, including predicted

exon boundary locations and gene sizes. BLAT analyses were also performed of human *LIPF*, *LIPJ*, *LIPK*, *LIPM* and *LIPN* genes and the mouse *Lipo1-like* gene using previously reported sequences for encoded subunits in each case (see Table 1). Structures for the major human *LIPA* and *CH25H* isoforms (gene splicing variants) were obtained using the AceView website to examine predicted gene structures using the human *LIPA* and *CH25H* genes to interrogate the database of human mRNA sequences [35] (<http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/index.html?human>). Predicted transcription factor binding sites (TFBS) and CpG islands for human *LIPA* and *CH25H* genes were identified using the UC Santa Cruz web browser [<http://genome.ucsc.edu/cgi-bin/hgBlat>] [34].

2.2 Predicted Structures and Properties for Vertebrate LIPA Subunits.

Predicted structures for vertebrate LIPA subunits were obtained using the SWISS MODEL web tools [<http://swissmodel.expasy.org/>] [36]. The reported tertiary structure for human LIPF and predicted tertiary structure for human LIPA [50] served as the reference for the predicted vertebrate LIPA tertiary structures, with a modeling range of residues 24-395. Theoretical isoelectric points and molecular weights for vertebrate LIPA and CH25H subunits were obtained using Expasy web tools (http://au.expasy.org/tools/pi_tool.html). SignalP 3.0 web tools were used to predict the presence and location of signal peptide cleavage sites (<http://www.cbs.dtu.dk/services/SignalP/>) for each of the predicted vertebrate LIPA sequences [38]. The NetNGlyc 1.0 Server was used to predict potential N-glycosylation sites for vertebrate LIPA subunits (<http://www.cbs.dtu.dk/services/NetNGlyc/>).

2.3 Predicted transmembrane structures for vertebrate CH25H subunits

Predicted transmembrane structures for vertebrate CH25H subunits were obtained using the server (<http://www.cbs.dtu.dk/services/TMHMM-2.0>) provided by the Center for Biological Sequence Analysis of the Technical University of Denmark [40].

2.4 Phylogenetic Studies and Sequence Divergence

Alignments of acid lipase protein sequences were assembled using BioEdit v.5.0.1 and the default settings [40]. Alignment ambiguous regions, including the amino and carboxyl termini, were excluded prior to phylogenetic analysis yielding alignments of 365 residues for comparisons of vertebrate LIPA, human LIPF, LIPJ, LIPK, LIPM and LIPN, mouse LIP01 and *Drosophila melanogaster* LIP3 sequences (Table 1). Evolutionary distances were calculated using the Kimura option [41] in TREECON [42]. Phylogenetic trees were constructed from evolutionary distances using the neighbor-joining method [43] and were rooted using the

Drosophila melanogaster LIP3 sequence. Tree topology was reexamined by the boot-strap method (100 bootstraps were applied) of resampling [44].

3. Results and Discussion

3.1 Alignments of vertebrate LIPA amino acid sequences.

The amino acid sequences of *in silico* derived LIPA subunits are shown in Figure 1 together with previously reported sequences for human and mouse LIPA [8, 9, 45]. Alignments of human LIPA with other predicted vertebrate LIPA sequences showed 64-98% identities, whereas lower levels of identities were observed with human LIPF, LIPJ, LIPK, LIPM and LIPN and with mouse LIPO1 sequences (49-63% identities), and with the *Drosophila melanogaster* LIP3 sequence (38% identity) (alignments of vertebrate LIPA sequences with human and mouse acid lipase gene families are not shown) (Table 2). This comparison suggests that the vertebrate subunits identified are all products of a single gene family (*LIPA*) which is distinct from those previously described for mammalian *LIPF*, *LIPJ*, *LIPK*, *LIPM* and *LIPN* gene families [10-18], and from a new rodent acid lipase gene family, designated as *Lipo* [unpublished results].

The predicted amino acid sequences for these vertebrate LIPA subunits were all of similar length (397-404 residues) and shared many (~ 34%) of identically aligned residues (Figure 1; Table 1). In addition, key residues previously described for human gastric acid lipase (LIPF) and human LIPA [50] involved in catalysis and maintaining enzyme structure have been conserved. Those retained for catalytic function included the active site residues involved with the charge relay system (human LIPA residue numbers used) (Ser174; Asp345; His374); the active site motif (Gly-Xaa-Ser-Yaa-Gly) (residues 172-176); residues Leu89 and Gln175 (replaced with 175Glu for chicken LIPA) which stabilize the 'oxyanion' transition state; and cysteine residues which form a disulfide bond (Cys248/Cys257 [37] to support the enzyme's structure.

The hydrophobic N-terminus signal peptide function (residues 1-18 for human LIPA) and the mannose-6-phosphate containing N-glycosylation site (residues 161-3: Asn-Lys-Thr) have been retained for all vertebrate LIPA sequences examined with the exception of frog LIPA, which has lost this N-glycosylation site with an Asn161 → Lys substitution. A potential basic amino acid 'patch' within the C-terminal sequence (386Arg..390Lys..396Arg-397Lys for human LIPA) [46], has also been retained or have undergone conservative substitution(s) for the vertebrate LIPA sequences examined (Figure 1). Two of the other high probability N-glycosylation sites for human LIPA (Asn36-Val37-Ser38; and Asn273-274Met-275Ser) were also conserved for the vertebrate LIPA sequences examined, while

another was retained for some vertebrate LIPA sequences (Asn72-His73-Ser74 for human, rhesus, horse and cow LIPA) (Figure 1; Table 3). There were species differences observed for the theoretical isoelectric points (pI) of the vertebrate LIPA subunits, with predicted higher values (pI values > 8) found for mouse and chicken LIPA as compared with the other LIPA subunits examined (Table 1).

3.2 Alignments of vertebrate CH25H amino acid sequences

Amino acid sequence alignments of *in silico* derived CH25H subunits are shown in Figure 2 together with previously reported sequences for human and mouse CH25H [19]. Most of the vertebrate CH25H sequences were 270-274 amino acid residues in length, with the exception of mouse and rat CH25H which exhibited extended C-termini, and contained 298 residues. Three histidine boxes reported for human CH25H [19] have been conserved for all vertebrate CH25H sequences examined, including box 1 (Trp-His-Leu/Val-Leu-Val-His-His) for residues 142-148; box 2 (Phe/Ile-His-Lys-Val/Met/Leu-His-His) for residues 157-162; and box 3 (His-His-Asp-Leu/Met-His-His) for residues 238-244 (Figure 2). These have been previously shown to be essential for CH25H catalytic activity and bind the iron atoms which assist in the hydroxylation reaction [47]. Predicted transmembrane structures for vertebrate CH25H are also shown (Figures 2), for which three such regions are predominantly retained for the sequences examined, with the exception of dog CH25H, with only the first of the predicted transmembranes. Figure 4 examines in more detail the predicted positioning of the three transmembranes within the human CH25H sequence and suggests that the N-terminus commences outside the endoplasmic reticulum membrane, and that the three active site histidine boxes are localized inside the membrane, where CH25H catalytic activity is likely to take place.

3.3 Comparative Vertebrate LIPA and CH25H Genomics

The AceView web browser defines the human *LIPA* gene by 1443 GenBank accessions from cDNA clones derived from spleen, brain, liver and many other tissues and reports a high expression level (~4.9 times the average human gene) (<http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/>) [35]. Similar high expression levels and wide tissue expression have also been reported for mouse *Lipa* (~3.2 times the average mouse gene) [35]. This high level expression is consistent with LIPA being responsible for an essential intracellular lipid catabolic process within lysosomes of many tissues and cells of the body [1-6].

Human *LIPA* transcripts included 22 alternatively spliced variants, which differed by truncations of the 5' or 3' ends, presence or absence of 10 cassette exons, or having overlapping exons with different boundaries. Of these, 5 were complete proteins, including isoform *LIPAb* (RefSeq NM_00235) which is shown in Figure 4. The predicted 38.47kb sequence contained 10 pre-messenger exons and 9 coding exons, as well as several transcription

factor binding sites (TFBS) and a CpG island (designated as CpG45) within the 5'-untranslated region for the human *LIPA* gene (Figure 4). Figure 1 compares the locations of the intron-exon boundaries for the vertebrate *LIPA* gene products examined. Exon 1 corresponded to the encoded signal peptide in each case, and exon 4 encoded the lysosomal targeting sequence (for human *LIPA* residues 161-3 Asn-Lys-Thr) [46]. There is identity or near identity for the intron-exon boundaries for each of the vertebrate *LIPA* genes suggesting conservation of these exons during vertebrate evolution.

In contrast to human *LIPA*, the human *CH25H* gene is defined by only 29 GenBank accessions for the AceView web browser from cDNA clones derived from 14 tissues including pancreas, brain and lung and shows a lower expression level (~25% of the average human gene) (<http://www.ncbi.nlm.nih.gov/IEB/Research/Aceembly/>) [35]. Moreover, a single human *CH25H* transcript was recorded covering 1.7kb of sequence which was intronless and contained a large 5' untranslated sequence proximally located near the 3' region of the *LIPA* gene (Figure 4), which is consistent with results previously reported [19]. The human *CH25H* genome sequence contained several predicted TFBS sites and a CpG island (CpG33) which was located in the intragenic region (~7.5kb) separating the human *CH25H* and *LIPA* genes on chromosome 10. Of particular significance are the CREB (cyclic-AMP response element-binding) binding sites, which may play a role in driving expression from the *CH25H* promoter [48]. Close proximal location for these genes was also observed for all other mammalian genomes examined (<20kb) (Table 1), while chicken (*Gallus gallus*) and frog (*Xenopus tropicalis*) *LIPA* and *CH25H* genes were more distantly located located in tandem (~160kb). CpG islands were observed in the human *LIPA-CH25H* intragenic region and in the 5'-untranslated *LIPA* region which may reflect roles for these CpG islands in up-regulating gene expression [49], given their collocation with the *LIPA* and *CH25H* gene promoters.

3.4 Secondary and Tertiary Structures for Vertebrate *LIPA* Sequences

Figure 1 shows the secondary structures predicted for vertebrate *LIPA* sequences. Similar α -helix β -sheet structures were observed for all of the vertebrate *LIPA* subunits examined, particularly near key residues or functional domains, including the α -helix within the N-terminal signal peptide, the β -sheet and α -helix structures surrounding the active site Ser174 (for human *LIPA*) and the α -helix enclosing the lysosomal targeting signal residues (Asn-Lys-Thr residues 161-3 for human *LIPA*) [46]. The pattern of secondary structures were very similar to those reported for human *LIPF* and predicted for human *LIPA* and are numbered accordingly [50]. Roussel and coworkers [50] have previously described these globular enzymes as being α/β hydrolase-like, with a core domain between residues 28-202 and 328-398 (Figure 1) which contains a central β -sheet composed of 8 strands, designated

as $\beta 1 - \beta 8$, and 6 α -helices, designated as $\alpha 1$, αA , $\alpha B1/B2$, $\alpha C1/C2$, αE and αF , with 3 helices on each side of the central β -sheet. In addition, a 'Cap' domain is described for human LIPF and LIPA with 8 helices (designated as $\alpha e1-\alpha e8$) within residues 203-329 [50]. This domain apparently acts as a 'lid' for the active site Ser174 which restricts access by the aqueous environment and but allows for cholesteryl ester and other substrate entry when the 'lid' opens. All of these secondary structures have been retained for the vertebrate LIPA subunits examined however these predictions of LIPA secondary structures may not fully reflect structures *in vivo* and serve only as a guide to the comparative structures for vertebrate LIPA subunits. The predicted tertiary structures for human, mouse, cow and chicken LIPA were sufficiently similar to the previously reported dog and human LIPF (gastric acid lipase) and the predicted human LIPA structures [50] (Figure 5) to enable predictions of these tertiary structures. These were based however, on incomplete sequences for these enzymes (residues 24-395 for human LIPA). The predictions observed suggest that the major structural features for human LIPA recently reported [50] resemble those for other vertebrate LIPA proteins, as well as for human gastric LIPF.

3.5 Phylogeny of Vertebrate LIPA and other Human Acid Lipase Genes and Proteins

Phylogenetic trees (Figure 6) were constructed from alignments of vertebrate LIPA-like amino acid sequences with human LIPF, LIPJ, LIPK, LIPM and LIPN and mouse LIPO1 sequences. The dendrogram was rooted using a *Drosophila melanogaster* LIP3 sequence [51] and showed clustering of all of the LIPA-like sequences which were distinct from the other human and mouse acid lipase gene families. These results are consistent with the vertebrate LIPA gene family being present throughout vertebrate evolution, and of apparent ancient origin of more than 500 million years ago [52]. 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 undertaken in each case. Of particular interest is the node demonstrating a highly significant separation of the vertebrate LIPA sequences from other acid lipase gene family (LIPF, LIPJ, LIPK, LIPM, LIPN and LIPO) sequences during vertebrate evolution, which supports the separate family status for vertebrate LIPA. Table 2 summarizes the percentages of identity for these enzymes and shows that vertebrate LIPA sequences are $\geq 64\%$ identical which is in comparison with the 44-63% identities observed comparing sequence identities between acid lipase families. In addition, more closely related species showed higher levels of sequence identity for LIPA, such as the primate species (human and rhesus monkey) which were 98% identical, as compared with the bird (chicken) and human LIPA sequences, with 72% identical sequences.

4. Conclusions

This study has compared the predicted genomic and protein structures of lysosomal lipases (LIPA) from several mammalian and other vertebrate genome sources. All of these vertebrate LIPA sequences shared key conserved sequences and predicted secondary and tertiary structures that have been reported for human LIPJ (gastric acid lipase) and LIPA [37], including active site and catalytic transition state supporting residues and disulfide bond forming cysteine residues. A specific N-glycosylation site, previously identified as a lysosomal targeting signal [46], was also conserved for mammalian and chicken LIPA sequences examined, but was missing from the frog (*Xenopus tropicalis*) LIPA sequence. Comparative gene expression data showed that human and mouse *LIPA* genes are expressed at higher levels than those for the average gene which is consistent with the essential metabolic roles for these enzymes in lysosomal cholesterol ester and triglyceride metabolism. Phylogeny studies suggest that the *LIPA* family of acid lipases is distinct from other acid lipases and has apparently evolved from an ancestral *LIPA* gene during or prior to vertebrate evolution.

5. Acknowledgements

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. We gratefully acknowledge the expert assistance of Dr Bharet Patel of Griffith University with the phylogenetic studies.

REFERENCES

1. J. L. Goldstein, S. E. Dana, J. R. Faust, A. L. Beaudet, and M. S. Brown, "Role of lysosomal acid lipase in the metabolism of plasma low density lipoprotein. Observations in cultured fibroblasts from a patient with cholesteryl ester storage disease," *Journal of Biological Chemistry*, vol. 250, pp. 8487–8495, 1975.
2. R. A. Anderson, R. S. Byrum, P. M. Coates, and G. N. Sando, "Mutations at the lysosomal acid cholesteryl ester hydrolase gene locus in Wolman disease," *Proceedings of the National Academy of Sciences USA*, vol. 91, pp. 2718-2722, 1994.
3. F. Wang, W. Wang, K. Wahala, H. Adlercreutz, E. Ikonen, and M. J. Tikkanen, "Role of lysosomal acid lipase in the intracellular metabolism of LDL-transported dehydroepiandrosterone-fatty acyl esters", *American Journal of Physiology Endocrinology and Metabolism*, vol. 295, pp. E1455-61, 2008.
4. A. L. Beaudet, G. D. Ferry, B. L. Nichols Jr., and H. S. Rosenberg, "Cholesterol ester storage disease: clinical, biochemical, and pathological studies," *Journal of Pediatrics*, vol. 90, pp. 910-914, 1977.
5. B. K. Burton, and S. P. Reed, "Acid lipase cross-reacting material in Wolman disease and cholesterol ester storage disease," *American Journal of Human Genetics*, vol. 33, pp. 203-208, 1981.
6. J. M. Hoeg, S. J. Demosky Jr., O. H. Pescovitz, and H. B. Brewer Jr., "Cholesteryl ester storage disease and Wolman disease: phenotypic variants of lysosomal acid cholesteryl ester hydrolase deficiency," *American Journal of Human Genetics*, vol. 36, pp. 1190-1203, 1984.
7. G. Koch, P. A. Lalley, M. McAvoy, and T.B. Shows, "Assignment of LIPA, associated with human acid lipase deficiency to human chromosome 10 and comparative assignment to mouse chromosome 19," *Somatic Cell Genetics*, vol. 7, pp. 345-358, 1981.

8. R. A. Anderson, and G. N. Sando, "Cloning and expression of cDNA encoding human lysosomal acid lipase/cholesterol ester hydrolase: similarities to gastric and lingual lipases," *Journal of Biological Chemistry*, vol. 266, pp. 22479-22484, 1991.
9. D. Ameis, M. Merkel, C. Eckerskorn, and H. Greten, "Purification, characterization and molecular cloning of human hepatic lysosomal acid lipase," *European Journal of Biochemistry*, vol. 219, pp. 905-914, 1994.
10. M. W. Bodmer, S. Angal, *et al.*, "Molecular cloning of a human gastric lipase and expression of the enzyme in yeast," *Biochimica et Biophysica Acta*, vol. 909, pp. 237-244, 1987.
11. P. Deloukas, M. E. Earthrowl, *et al.*, "The DNA sequence and comparative analysis of human chromosome 10," *Nature*, vol. 429, pp. 375-381, 2004.
12. E. Toulza, Mattiuzzo, *et al.*, "Large-scale identification of human genes implicated in epidermal barrier function," *Genome Biology*, vol. 8, pp. R107.1-R107.23, 2007.
13. K. Hirata, H. L. Dichek, *et al.*, "Cloning of a unique lipase from endothelial cells extends the lipase gene family," *Journal of Biological Chemistry*, vol. 274, pp. 14170-14175, 1999.
14. M. Jaye, K. J. Lynch, *et al.*, "A novel endothelial-derived lipase that modulates HDL metabolism," *Nature Genetics*, vol. 21, pp. 424-428, 1999.
15. K. L. Wion, T. G. Kirchgessner, *et al.*, "Human lipoprotein lipase complementary DNA sequence," *Science*, vol. 235, pp. 1638-1641, 1987.
16. K. Ishimura-Oka, F. Faustinella, *et al.*, "A missense mutation (Trp86-->Arg) in exon 3 of the lipoprotein lipase gene: a cause of familial chylomicronemia," *American Journal of Human Genetics*, vol. 50, pp. 1275-1280, 1992.
17. G. A. Martin, S. J. Busch, *et al.*, "Isolation and cDNA sequence of human postheparin plasma hepatic triglyceride lipase," *Journal of Biological Chemistry*, vol. 263, pp. 10907-10914, 1988.
18. S. J. Cai, D. M. Wong, *et al.*, "Structure of the human hepatic triglyceride lipase gene," *Biochemistry*, vol. 28, pp. 8966-8971, 1989.
19. E. G. Lund, T. A. Kerr, *et al.*, "cDNA cloning of mouse and human cholesterol 25-hydroxylases, polytopic membrane proteins that synthesize a potent oxysterol regulator of lipid metabolism," *Journal of Biological Chemistry*, vol. 273, pp. 34316-34327, 1998.
20. D. R. Baumann, A. D. Bitmansour, *et al.*, "25-Hydroxycholesterol secreted by macrophages in response to Toll-like receptor activation suppresses immunoglobulin A production," *Proceedings of the National Academy of Sciences USA*, vol. 106, pp. 16764-16869, 2009.
21. J. R. Dwyer, N. Sever, *et al.*, "Oxysterols are novel activators of the hedgehog signaling pathway in pluripotent mesenchymal cells," *Journal of Biological Chemistry*, vol. 282, pp. 8959-68, 2007.
22. S. F. Altschul, W. Gish, *et al.*, "Basic local alignment search tool," *Journal of Molecular Biology*, vol. 215, pp. 403-410, 1990.
23. The Chimpanzee Sequencing and Analysis Consortium, "Initial sequence of the chimpanzee genome and comparison with the human genome," *Nature*, vol. 437, pp. 69-87, 2005.
24. R. A. Gibbs, J. Rogers *et al.*, "Evolutionary and biomedical insights from the rhesus macaque genome," *Science* vol. 316 pp. 222-34, 2007.
25. Horse Genome Sequencing Consortium, <http://www.broadinstitute.org/mammals/horse>, 2007.
26. Bovine Genome Project, <http://www.hgsc.bcm.tmc.edu/projects/bovine/>, 2008.
27. Mouse Genome Sequencing Consortium, "Initial sequencing and comparative analysis of the mouse genome," *Nature*, vol. 420, pp. 520-562, 2002.
28. Rat Genome Sequencing Project Consortium. Genome sequence of the Brown Norway rat yields insights into mammalian evolution. *Nature* **428**: 493-521 (2004).
29. Guinea Pig Genome Sequencing Project, <http://www.broadinstitute.org/science/projects/mammals-models/guinea-pig/guinea-pig>, 2008.
30. Dog Genome Project, <http://www.broadinstitute.org/mammals/dog>, 2005.
31. International Chicken Genome Sequencing Consortium, "Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution," *Nature*, vol. 432, pp. 695-716, 2004.
32. *Xenopus tropicalis* Genome Project, <http://genome.jgi-psf.org/Xentr4/Xentr4.home.html>, 2005.
33. T. J. P. Hubbard, B. L. Aken, *et al.*, "Ensembl," *Nucleic Acids Research*, vol. 35, pp. D610-D617, 2007.
34. W. J. Kent, C. W. Sugnet *et al.*, "The human genome browser at UCSC," *Genome Research*, vol. 12, pp. 994-1006, 2002.
35. D. Thierry-Mieg, and J. Thierry-Mieg, "AceView: A comprehensive cDNA-supported gene and transcripts annotation," *Genome Biology* vol. 7, p. S12, 2006.
36. J. Kopp and T. Schwede, "The SWISS-MODEL repository of three dimensional protein structure homology models," *Nucleic Acids Research*, vol. 32, pp. D230-D234, 2004.

37. A. Roussel, S. Canaan, *et al.*, “Crystal structure of human gastric lipase and model of lysosomal acid lipase, two lipolytic enzymes of medical interest,” *Journal of Biological Chemistry*, vol. 274, pp. 16995-17002, 1999.
38. O. Emanuelsson, S. Brunak, *et al.*, “Locating proteins in the cell using TargetP, SignalP and related tools,” *Nature Protocols* vol. 2, pp. 953-971, 2007.
39. A. Krogh, B. Larsson, *et al.*, “Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes,” *Journal of Molecular Biology*, vol. 305, pp. 567-580, 2001.
40. T. A. Hall, “BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT,” *Nucleic Acids Symposium Series*, vol. 41, pp. 95-98, 1999.
41. M. Kimura “*The Neutral Theory of Molecular Evolution*,” Cambridge University Press, Cambridge Mass. , 1983
42. Y. Van De Peer, and R. de Wachter, “TreeCon for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment,” *Computational and Applied Science*, vol. 10, pp. 569-570, 1994.
43. N. Saitou, and M. Nei, “The neighbour-joining method: a new method for reconstructing phylogenetic trees,” *Molecular Biology and Evolution*, vol. 4, pp. 406-426, 1987.
44. J. Felsenstein, “Confidence limits on phylogenies: an approach using the bootstrap,” *Evolution*, vol. 39, pp. 783-791, 1985.
45. H. Du, D. P. Witte, and G. A. Grablowski, “Tissue and cellular specific expression of murine lysosomal acid lipase mRNA and protein,” *Journal of Lipid Research*, vol. 37, pp. 937-949, 1996.
46. D. E. Sleat, H. Zheng, M. Qian M, and P. Lobel, “Identification of sites of mannose 6-phosphorylation on lysosomal proteins,” *Molecular and Cell Proteomics*, vol. 5, pp. 686-701, 2006.
47. B. G. Fox, J. Shanklin, J. Ai, T.M. Loehr, and J. Sanders-Loehr J, “Resonance Raman evidence for an Fe-O-Fe center in stearyl-ACP desaturase. Primary sequence identity with other diiron-oxo proteins,” *Biochemistry* vol. 33, pp.12776-86, 1994.
48. D. J. Watters, and J. Nourse, “Identification and characterization of the cholesterol 25-hydroxylase gene promoter,” *FASEB Journal*, Abstract 689.9, 2009.
49. S. Saxonov, P. Berg, and D.L. Brutlag, “A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters,” *Proceedings of the National Academy of Sciences USA*, vol. 103, pp. 1412–1417, 2006.
50. A. Roussel, M. Miled, *et al.*, “Crystal Structure of the Open Form of Dog Gastric Lipase in Complex with a Phosphonate Inhibitor,” *Journal of Biological Chemistry*, vol. 277, pp. 2266-2274, 2002.
51. D. Pistillo, A. Manzi, *et al.*, “The *Drosophila melanogaster* lipase homologs: a gene family with tissue and developmental specific expression,” *Journal of Molecular Biology*, vol. 276, pp. 877-885, 1998.
52. P. C. J. Donoghue, and M. J. Benton, “Rocks and clocks: calibrating the tree of life using fossils and molecules,” *Trends in Genetics*, vol. 22, pp. 424-431, 2007.

Figure 1: Amino acid sequence alignments for vertebrate LIPA sequences.

HuA: human LIPA; RhA: rhesus LIPA; HoA: horse LIPA; MoA: mouse LIPA; RaA: rat LIPA; CoA: cow LIPA; DoA: dog LIPA; ChA: chicken LIPA; XeA: frog LIPA. See Table 1 for sources of LIPA sequences; * identical residues; : 1 or 2 conservative substitutions; . 1 or 2 non-conservative substitutions; residues involved in processing at N-terminus (signal peptide); potential N-glycosylation sites including residues NKT (161-163) which serves as a lysosomal targeting sequence ^LTS; active site residues Ser174; Asp345; and His374; disulfide bond C residues Cys248 and Cys257 for human LIPA; helix (human LIPA) or predicted helix; Sheet (human LIPA) or predicted sheet; numbered according to Roussel *et al* [37]; potential basic amino acid ‘patches’ for lysosomal targeting #; residues Leu89 and Gln175 contributing to the oxyanion ‘hole’ near active site [37]; and bold underlined font shows known or predicted exon junctions.

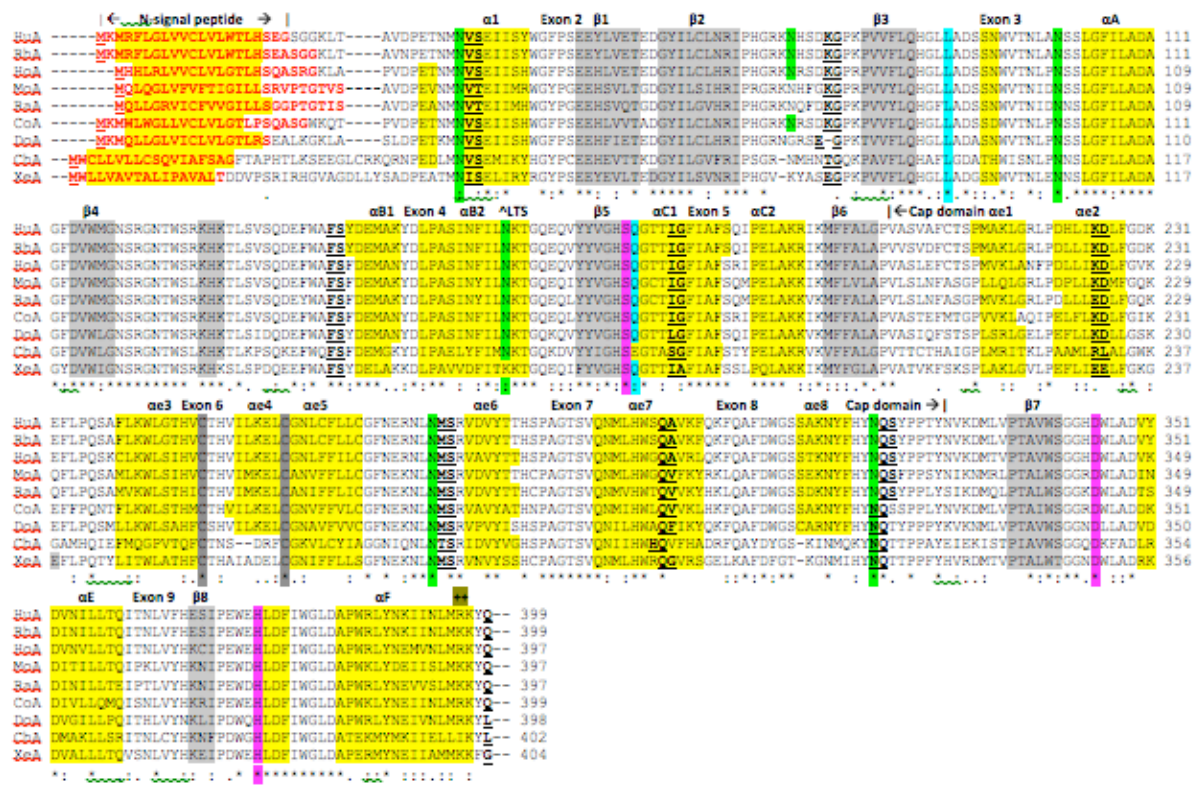
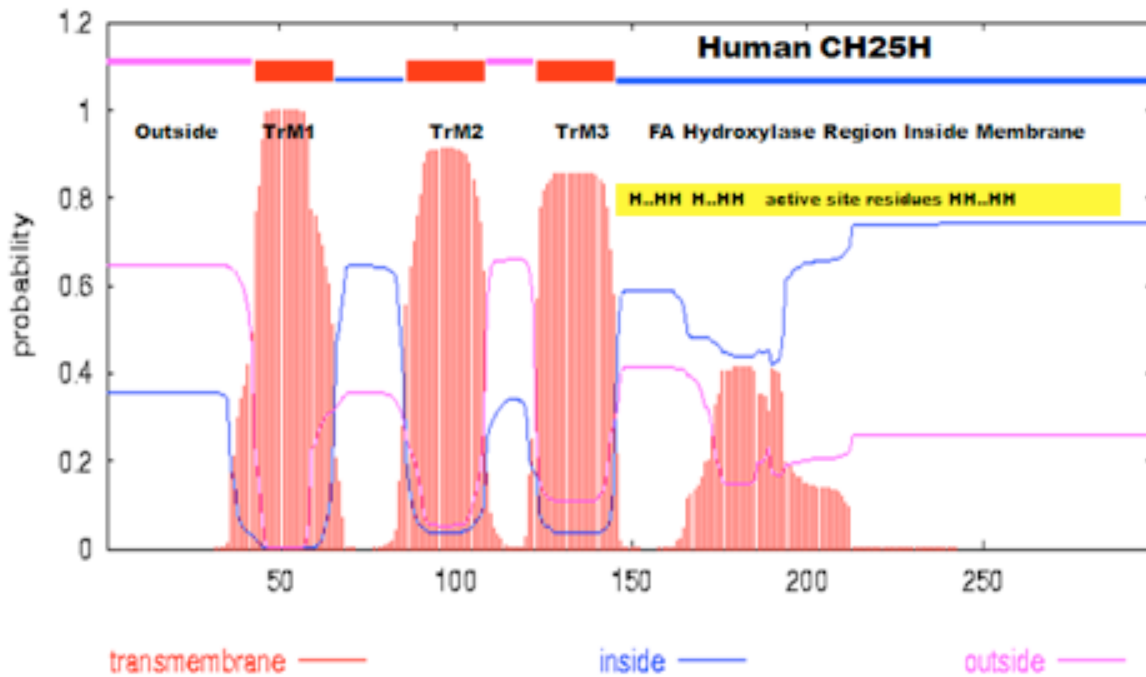


Figure 2: Amino acid sequence alignments for vertebrate CH25H sequences.



HuCH25H: human CH25H; RhCH25H: rhesus CH25H; MoCH25H: mouse CH25H; RaCH25H: rat CH25H; CoCH25H: cow CH25H; HoCH25H: horse CH25H; DoCH25H: dog Ch25H; ChCH25H: chicken CH25H; XeCH25H: frog CH25H. See Table 1 for sources of CH25H sequences; * identical residues; : 1 or 2 conservative substitutions; . 1 or 2 non-conservative substitutions; histidine residues active site boxes 1, 2 and 3; predicted helix; predicted sheet; predicted transmembrane regions; bold underlined font shows known or predicted exon junctions (single exon CH25H genes observed in each case).

Figure 3: Predicted locations for transmembrane regions for human CH25H.



The graph shows probability (0-1 on y-axis) of transmembrane regions (TrM1, TrM2 and TrM3 shown in red) for the human CH25H amino acid sequence (on x-axis). Predicted outside membrane CH25H residues are shown in red; predicted inside membrane CH25H residues are shown in blue; predicted positioning of the three histidine active site boxes are shown as H..HH or HH..HH and are localized inside the membrane.

Figure 4: Gene structures and tandem locations for the human *CH25H* and *LIPA* genes on chromosome 10.



Derived from the AceView website <http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/> [35]; isoform variant LIPAb and CH25H mRNAs are shown with capped 5'- and validated 3'-ends for the predicted sequences; predicted exon regions are shaded; note that CH25H is predicted as a single exon gene; 5'UTR and 3'UTR refer to untranslated 5' and 3' regions respectively; predicted transcription factor binding sites are shown: NKX25: homeobox protein 2.5 ; RP58: transcriptional repressor RP58; ROAZ: zinc finger protein 423; TAXCREB, CREBP1 and CREBP1C: cyclic-AMP responsive element-binding proteins; PPARG: peroxisome proliferator-activated receptor gamma; HNF4: hepatocyte nuclear factor 4-alpha ; COMP1: muscle specific transcription enhancer ; HNF3B: hepatocyte nuclear factor 3-beta ; GF11: zinc finger protein GF11; RORA2: alpha orphan nuclear receptor; EVI1: zinc finger protein EVI1; FREAC4: forkhead box protein; STAT3: identified in the promoters of acute-phase genes; HEN1: helix-loop-helix protein 1; and OCT1: transcription factor that binds to the octomer motif; predicted locations for CpG islands (CPG45; CPG33) are shown by shaded triangles.

Figure 5: Predicted three dimensional structures for vertebrate LIPA subunits.

Predicted 3-D structures were obtained using the SWISS MODEL (<http://swissmodel.expasy.org/workspace/index.php>) web site and the predicted amino acid sequences for vertebrate LIPA subunits (see Table 1). The rainbow color code describes the 3-D structures from the N- (blue) to C-termini (red color). The structures are based on the known 3-D structures for human LIPF [50] (with a modeling range of residues 24 to 395).

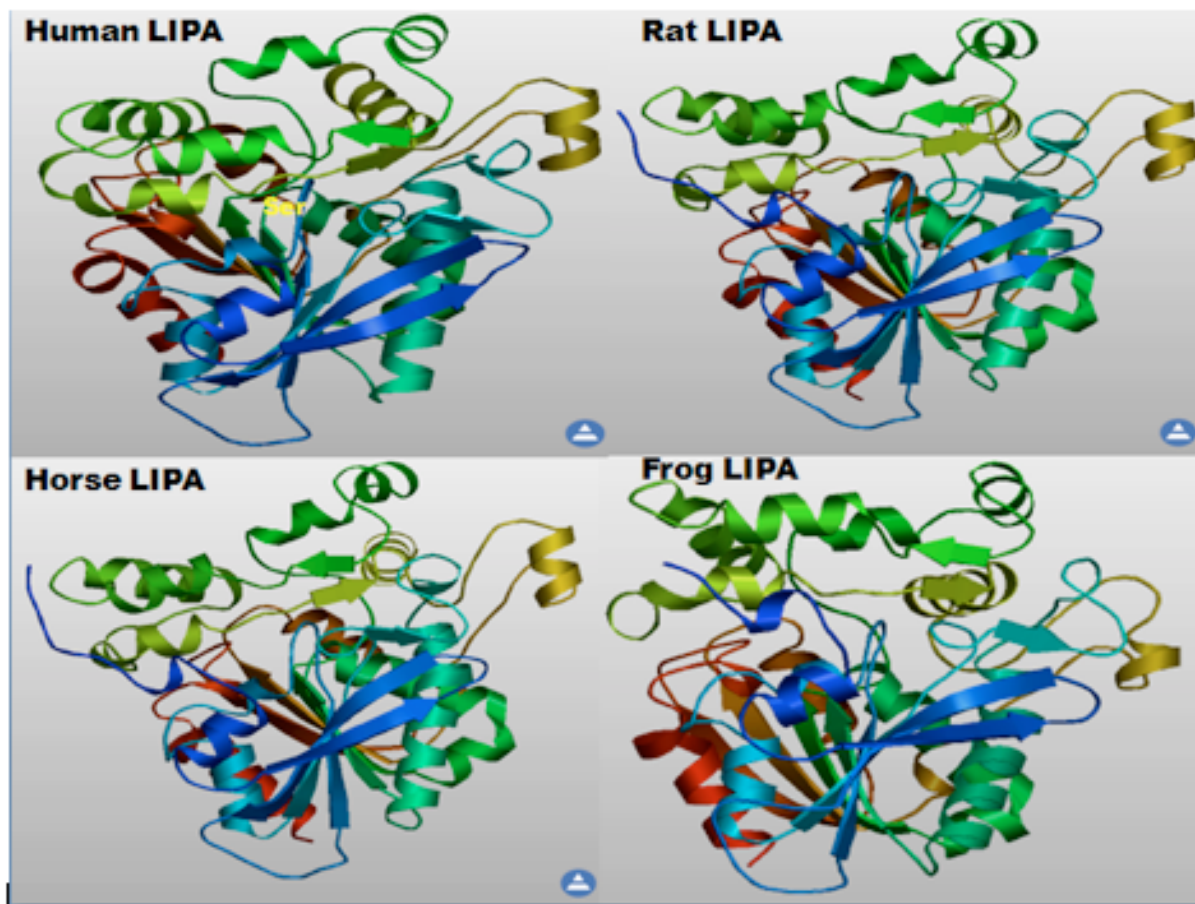


Figure 6: Phylogenetic tree of vertebrate LIPA, human LIPJ, LIPF, LIPK, LIPM and LIPN, mouse LIPO and *Drosophila melanogaster* LIP3 sequences.

The tree is labeled with the lipase gene family number and the species name. Note the separation of the human LIPF, LIPJ, LIPK, LIPM, LIPN and mouse LIPO sequences from the vertebrate LIPA family cluster. The *Drosophila melanogaster* LIP3 sequence was used to root the tree. A genetic distance scale is shown. 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. 100 bootstrap replicates were performed in each case. Of particular interest is the node marked with an asterisk* supporting the significant separation of vertebrate LIPA from other acid lipase sequences examined.



Table 1: Vertebrate Lipase A (LIPA), Human and Mouse Acid Lipases and Cholesterol 25-hydroxylase (CH24H) Genes and Enzymes Examined

RefSeq: the reference mRNA sequence;¹ predicted Ensembl mRNA sequence; and GenBank mRNA (or cDNA) IDs are shown (see <http://www.ncbi.nlm.nih.gov>); ²result not available; ³cleavage site predicted for signal peptide at N-termini; ⁴Ensembl gene prediction; ⁵N-scan gene prediction using the software from the Computational Genomics Lab at Washington University in St. Louis, MO, USA (see <http://genome.ucsc.edu>); ⁶Contig ID given; UNIPROT refers to UniprotKB/Swiss-Prot IDs for individual LIPA, other acid lipase or CH25H subunits (see <http://kr.expasy.org>); bps refers to base pairs of nucleotide sequences; pI refers to theoretical isoelectric points; the number of coding exons are listed. Sources for LIPA, other human and mouse acid lipases and CH25H sequences were provided by the above sources

(see Table 1).

Lipase Gene	Species	RefSeq	GenBank ID	UNIPROT	Amino acids	Chromosome location	Exons (strand)	Gene Size (bps)	pl	Subunit MW	Signal Peptide (Cleavage site)
Human LIPA	<i>Homo sapiens</i>	NM_001127605	BC012287	P38571	399	10:90,964,568-90,997,385	9 (-ve)	32,818	6.42	45,419	1-21 (EG-SG) ¹
Chimp LIPA	<i>Pan troglodytes</i>	XP_521552	BC012287	P38571	399	10:89,482,403-89,515,834	9 (-ve)	33,432	6.42	45,419	1-21 (EG-SG) ¹
Orangutan LIPA	<i>Pongo abelii</i>	*ENSPPYT0000002953	2	2	399	10:87,901,737-87,934,555	9 (-ve)	32,819	6.42	45,452	1-21 (EG-SG) ¹
Rhesus LIPA	<i>Macaca mulatta</i>	XP_001085160	2	2	399	9:88,805,255-88,839,453	9 (-ve)	34,199	6.39	45,480	1-24 (GG-KL) ¹
Marmoset LIPA	<i>Callithrix jacchus</i>	¹⁹ 4107.004.a	2	2	399	⁴ 4107:158,060-204,094	9 (+ve)	46,035	6.34	45,424	2
Mouse LIPA	<i>Mus musculus</i>	NM_021460	BC058064	Q920M5	397	19:34,568,473-34,599,332	9 (-ve)	30,860	8.15	45,325	1-25 (VS-AV) ¹
Rat LIPA	<i>Rattus norvegicus</i>	NM_012732	BC072532	Q64194	397	1:238,468,218-238,497,746	9 (-ve)	29,529	6.30	45,079	1-25 (IS-AV) ¹
Guinea Pig LIPA	<i>Cavia porcellus</i>	XP_001503012	2	2	397	1:39,069,159-39,103,381	9 (+ve)	34,223	7.29	46,327	1-22 (RG-KL) ¹
Horse LIPA	<i>Equus caballus</i>	XP_001503012	2	2	397	1:39,069,159-39,103,381	9 (+ve)	34,223	7.29	46,327	1-22 (RG-KL) ¹
Cow LIPA	<i>Bos taurus</i>	NP_001096793	BC146075	2	399	26:11,349,737-11,387,245	9 (-ve)	37,509	7.23	45,671	1-23 (SG-WK) ¹
Pig LIPA	<i>Sus scrofa</i>	NP_001116606	2	2	399	2	2	7.75	45,347	1-19 (HS-EA) ¹	
Dog LIPA	<i>Canis familiaris</i>	XP_853280	2	2	398	26:41,958,963-41,981,595	9 (-ve)	22,633	6.70	45,063	1-19 (RS-EA) ¹
Chicken LIPA	<i>Gallus gallus</i>	XP_426515	2	2	402	6:20,252,280-20,262,074	9 (+ve)	9,795	8.44	45,610	1-18 (AG-FT) ¹
Frog LIPA	<i>Xenopus tropicalis</i>	NM_001015847	BC090136	2	404	sc ¹ 150:1,826,750-1,838,449	9 (+ve)	11,700	5.81	45,454	1-17 (LT-DD) ¹
Fruit Fly Lip3	<i>Drosophila melanogaster</i>	NM_057983	BT023252	O46108	394	3R:9,195,960-9,197,626	3 (-ve)	1,667	5.40	44,901	1-20 (LA-GS) ¹
Human LIPF	<i>Homo sapiens</i>	NM_004190	AK208334	P07098	398	10:90,424,214-90,438,571	9 +ve	14,357	6.8	45,238	1-19 (HG-LF) ¹
Human LIPJ	<i>Homo sapiens</i>	NM_001010939	BC031219	Q5W064	366	10:90,346,519-90,366,732	9 +ve	20,213	6.1	42,388	2
Human LIPK	<i>Homo sapiens</i>	NM_001080518	EF426482	Q5VXJ0	399	10:90,474,281-90,502,490	9 +ve	28,210	8.4	45,563	1-19 (HG-YO) ¹
Human LIPM	<i>Homo sapiens</i>	NM_001128215	EF426484	Q5VYY2	423	10:90,552,634-90,570,235	9 +ve	17,602	6.6	48,233	1-33 (NS-VH) ¹
Human LIPN	<i>Homo sapiens</i>	NM_001102469	EF426483	Q5VXJ9	399	10:90,511,143-90,527,979	9 +ve	16,834	6.4	45,534	1-18 (NA-GG) ¹
Mouse LIPO1	<i>Mus musculus</i>	NP_001013792	A1747699	2	399	19:33,851,027-33,861,942	9 (-ve)	10,916	6.1	44,668	1-19 (FC-LI) ¹
CH25H Gene											Intragenic (bps)
Human	<i>Homo sapiens</i>	NM_003956	BC072430	O09592	272	10:90,956,214-90,957,029	1 (-ve)	816	6.77	31,745	7,539
Rhesus	<i>Macaca mulatta</i>	XP_001083208	2	2	272	9:88,797,797-88,798,612	1 (-ve)	816	6.75	31,850	6,643
Mouse	<i>Mus musculus</i>	NP_034020	BC039919	Q920F4	298	19:34,548,723-34,549,616	1 (-ve)	894	7.67	34,672	18,857
Rat	<i>Rattus norvegicus</i>	NP_001020586	BC097064	2	298	1:238,457,437-238,458,330	1 (-ve)	894	7.07	34,414	9,888
Horse	<i>Equus caballus</i>	XP_001503057	2	2	270	1:39,111,964-39,112,773	1 (+ve)	810	6.73	31,464	8,613
Cow	<i>Bos taurus</i>	NP_001068711	BC120312	2	270	26:11,336,966-11,337,775	1 (-ve)	810	6.88	31,326	11,972
Dog	<i>Canis familiaris</i>	XP_543596	2	2	270	26:41,949,857-41,950,666	1 (-ve)	810	8.88	30,426	8,297
Chicken	<i>Gallus gallus</i>	XP_421660	2	2	274	6:20,415,141-20,415,962	1 (-ve)	822	8.15	32,456	163,682
Frog	<i>Xenopus tropicalis</i>	¹⁹ sc.150.119	2	2	272	sc ¹ 150:1,998,638-1,999,453	1 (-ve)	816	8.61	31,405	158,189

Table 2: Percentage Identities for Vertebrate LIPA, Human LIPF, LIPJ, LIPK, LIPM and LIPN, Mouse LIPO1 and Fruit Fly (*Drosophila melanogaster*) LIP3 Amino Acid Sequences

Acid Lipase Gene	Human LIPA	Rhesus LIPA	Mouse LIPA	Chicken LIPA	Frog LIPA	Human LIPF	Human LIPJ	Human LIPK	Human LIPM	Human LIPN	Mouse LIPO1	Fruit Fly LIP3
Human LIPA	100	98	77	72	69	61	53	59	63	55	49	38
Rhesus LIPA	98	100	78	71	69	60	53	59	63	55	49	37
Mouse LIPA	77	78	100	65	64	55	49	55	58	52	47	38
Chicken LIPA	72	71	65	100	75	62	54	60	63	55	48	38
Frog LIPA	69	69	64	75	100	59	52	58	64	53	50	37
Human LIPF	61	60	55	62	59	100	55	66	55	53	50	37
Human LIPJ	53	53	49	54	52	55	100	57	51	48	46	33
Human LIPK	59	59	55	60	58	66	57	100	57	52	51	32
Human LIPM	63	63	58	63	64	55	51	57	100	54	49	35
Human LIPN	55	55	52	55	53	53	48	52	54	100	44	32
Mouse LIPO1	49	49	47	48	50	50	46	51	49	44	100	35
Fruit Fly LIP3	38	37	38	38	37	37	33	32	35	32	35	100

Numbers show the percentage of amino acid sequence identities. Numbers in **bold** show higher sequence identities for vertebrate LIPA sequences.

Table 3: Predicted N-Glycosylation Sites for Vertebrate LIPA Subunits.

Numbers refer to amino acids in the LIPA sequences, including N-asparagine; K-lysine; I-isoleucine; M-methionine; H-histidine; S-serine; R-arginine; T-threonine; Q-glutamine; and V-valine. Note that there are 6 potential sites identified. High (**in bold**) and lower probability N-glycosylation sites were identified using the NetNGlyc 1.0 web server (<http://www.cbs.dtu.dk/services/NetNGlyc/>). **Highlighted** column refers to N-glycosylation site 4 which has been shown to contribute to lysosomal targeting for LIPA [46]

Vertebrate	Species	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6	Potential N-glycosylation sites	High probability sites (>0.75)	Lower Probability Sites (0.5-0.74)
LIPA										
Protein										
Human	<i>Homo sapiens</i>	36NVS	72NHS	101NSS	161NKT	273NMS	321NQS	6	4	2
Rhesus	<i>Macaca mulatta</i>	36NVS	72NHS	101NSS	161NKT	273NMS	321NQS	6	4	2
Mouse	<i>Mus musculus</i>	34NVT		99NSS	159NKT	271NMS	319NQS	5	3	2
Rat	<i>Rattus norvegicus</i>	34NVT		99NSS	159NKT	271NMS	319NQS	5	3	2
Horse	<i>Equus caballus</i>	34NVS		99NSS	159NKT	271NMS	319NQS	5	3	2
Cow	<i>Bos taurus</i>	36NVS	72NRS	101NSS	161NKT	273NMS	321NQS	6	4	2
Dog	<i>Canis familiaris</i>	36NVS		100NSS	160NKT	272NMS	320NQT	5	2	3
Chicken	<i>Gallus gallus</i>	43NVS		107NNS	167NKT	277NMS	324NQT	5	1	4
Frog	<i>Xenopus tropicalis</i>	43NIS		107NNS		279NMS	326NQT	4	1	3
Fish	<i>Gasterosteus aculeatus</i>	39NIS				277NMT	324NQS	3	2	0