Evolution of the acyl-CoA binding protein (ACBP)

Mark BURTON*, Timothy M. ROSE†, Nils J. FÆRGEMAN* and Jens KNUDSEN*[1]

*Department of Biochemistry and Molecular Biology, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark, and †Department of Pathobiology, School of Public Health and Community Medicine, University of Washington, Seattle, WA, 98195, U.S.A.

Acyl-CoA-binding protein (ACBP) is a 10 kDa protein that binds C_{12}–C_{22} acyl-CoA esters with high affinity. *In vitro* and *in vivo* experiments suggest that it is involved in multiple cellular tasks including modulation of fatty acid biosynthesis, enzyme regulation, regulation of the intracellular acyl-CoA pool size, donation of acyl-CoA esters for β-oxidation, vesicular trafficking, complex lipid synthesis and gene regulation. In the present study, we delineate the evolutionary history of ACBP to get a complete picture of its evolution and distribution among species. ACBP homologues were identified in all four eukaryotic kingdoms, Animalia, Plantae, Fungi and Protista, and eleven eubacterial species. ACBP homologues were not detected in any other known bacterial species, or in archaea. Nearly all of the ACBP-containing bacteria are pathogenic to plants or animals, suggesting that an *ACBP* gene could have been acquired from a eukaryotic host by horizontal gene transfer. Many bacterial, fungal and higher eukaryotic species only harbour a single ACBP homologue. However, a number of species, ranging from protozoa to vertebrates, have evolved two to six lineage-specific paralogues through gene duplication and/or retrotransposition events. The ACBP protein is highly conserved across phylums, and the majority of *ACBP* genes are subjected to strong purifying selection. Experimental evidence indicates that the function of ACBP has been conserved from yeast to humans and that the multiple lineage-specific paralogues have evolved altered functions. The appearance of ACBP very early on in evolution points towards a fundamental role of ACBP in acyl-CoA metabolism, including ceramide synthesis and in signalling.

Key words: acyl-CoA ester, evolution, lipid transport, phylogeny, selection pressure.

INTRODUCTION

Acyl-CoA binding protein (ACBP) is a highly conserved, approximately 10 kDa cytosolic protein. It was first identified in mammals as a neuropeptide which inhibited diazepam binding to the GABA (γ-aminobutyric acid) receptor and was known as diazepam binding inhibitor (DBI)/endozepine (EP) [1]. Subsequent studies have shown that ACBP binds LCACoA (long-chain acyl-CoA) esters (C_{12}–C_{22}) with high specificity and affinity (K_d, 1–10 nM) [2]. A large body of *in vitro* evidence indicates that ACBP acts as an intracellular acyl-CoA transporter and pool former [2–4]. An ACBP homologue has been identified in yeast, where it is required for fatty acid chain elongation, sphingolipid synthesis, protein sorting and vesicular trafficking [3,5]. These and other studies demonstrate a major functional role for ACBP in binding and transporting acyl-CoA, although its ability to function as a neuropeptide has not been verified.

The structure of the bovine ACBP has been determined by NMR to be a four α-helix bundle protein [6,7]. Sequence analysis of chordate ACBP proteins revealed a large number of conserved residues that are either hydrophobic or involved in binding of the acyl-CoA ester in the binding site [8,9]. The majority of these conserved hydrophobic residues participate in forming three mini-cores within the ACBP structure, providing a network of interactions between the ACBP helixes [9].

The presence of multiple independent genes within a single species encoding different functional paralogues of ACBP has been noted. In mammals, in addition to the protein initially identified as ACBP/DBI/EP, first characterized in liver (L-ACBP), homologous distinct ACBP-like genes have been identified in testes [10,11] (T-ACBP) and brain [11,12] (B-ACBP). Similarly, two distinct ACBP-like genes have been identified in the plant, *Digitalis lanata* [12]. In addition, multiple inactive *ACBP* pseudogenes have been identified in different mammalian genomes, including rat [13] and human [14,15].

Although ACBP occurs as a completely independent protein, intact ACBP-like domains have been identified in a number of large, multifunctional proteins in a variety of eukaryotic species. These include large membrane-associated proteins with N-terminal ACBP domains [16–18], multifunctional enzymes with both ACBP and peroxisomal enoyl-CoA Δ^3,Δ^2-enoyl-CoA isomerase domains [19,20], and proteins with both an ACBP domain and ankyrin repeats [18].

The aim of the present study was to characterize the distribution and conservation of ACBP among species during evolution. We have searched the existing nucleotide, protein and genome databases for ACBP-related sequences. Large multi-domain proteins containing an ACBP domain have been ignored in this study. We identified ACBP homologues in all four eukaryotic kingdoms, Animalia, Plantae, Fungi and Protista, and in eleven eubacterial species. ACBP homologues were not detected in any other known bacterial species or in archaea. We show that the *ACBP* gene family is highly conserved and that multiple independent lineage-specific duplication events have generated variable numbers of related ACBP paralogues in different species.

MATERIALS AND METHODS

Identification of ACBP sequences and listing of nucleotide accession numbers

Identification of ACBP-related sequences was performed using BLAST similarity searches, using known ACBP protein

Abbreviations used: aa, amino acid; ACBP, acyl-CoA binding protein (identified in B-ACBP, brain; L-ACBP, liver and T-ACBP, testes); A.tha, etc., *Arabidopsis thaliana* ACBP etc.; EST, expressed sequence tag; L_a, average non-synonymous nucleotide substitution rate; L_s, average synonymous nucleotide substitution rate; LCACoA, long-chain acyl-CoA esters; Mya, million years ago.

[1] To whom correspondence should be addressed (email jkk@bmb.sdu.dk).

sequences as probes. Targeted databases included the non-redundant nucleotide and protein databases, EST (expressed sequence tag) databases, whole genome sequence databases, contig databases and patent databases from the following sites: NCBI (National Center for Biotechnology Information, http://www.ncbi.nlm.nih.gov/BLAST/), TIGR (the Institute for Genomic Research, http://tigrblast.tigr.org/tgi/), the Sanger Institute (http://www.sanger.ac.uk/DataSearch/), COGEME (Consortium for the Functional Genomics of Microbial Eukaryotes, http://cogeme.ex.ac.uk/index.html), the DOE Joint Genome Institute (http://www.jgi.doe.gov/) and Wormbase (http://www.wormbase.org/). The accession numbers of identified ACBP sequences are listed in Tables S1–S7 (http://www.BiochemJ.org/bj/392/bj3920299add.htm) and, unless otherwise indicated, refer to the NCBI GenBank accession. For reference purposes within this manuscript, we have provided a unique label for each gene and sequence characterized in this study which is based on a four letter code for the genus and species of the organism from which it is derived. An additional number is included to indicate paralogous ACBP subfamilies (see Tables S1–S7).

Alignments

The sequences used for phylogenetic analysis were aligned using CLUSTALX v.1.82 [21]. The protein alignments were conducted with the Gonnet matrix and the DNA alignments, needed for calculations of the L_a(average non-synonymous nucleotide-substitution rate)/L_s(average synonymous nucleotide-substitution rate) ratios, were aligned using the IUB DNA matrix with default settings followed by manual editing.

Construction of the phylogenetic trees

The maximum likelihood trees were constructed using the PROTML program implemented in the PHYLIP v.3.63 software package [22], with the following changes to the default settings: (1) global arrangements were set to 'on'; (2) quick and speedy processing was turned off and thorough processing was used instead. Bootstrap values were based on 100 pseudo-replications produced by the SEQBOOT program. The bootstrap values were calculated using the majority-rule implemented in the CONSENSE program.

Calculation of the L_a and L_s ratios

The sequences in the respective clusters were aligned using the ClustalX program [21]. The Kumar method was used for calculation of the non-synonymous (d_a) and synonymous (d_s) nucleotide substitution distances between ACBP orthologue sequences and an ancestor or ancestral group within or between specific phyla. All calculations were carried out using the MEGA 2.1 software package [23]. The L_s and the L_a were calculated by dividing the synonymous nucleotide (d_s) or non-synonymous (d_a) nucleotide substitution distance with two times the divergence time (T) to their last shared common ancestor(s) as follows: $L_s = d_s/2 \times T$ and $L_a = d_a/2 \times T$.

The common ancestors and divergence times used for calculating the respective L_a/L_s values were: B-ACBP and L-ACBP [fish versus amphibia, birds and mammals, 450 Mya (million years ago)] [24]; D.mel1 homologues (crustacean, 437 Mya) [25]; plant ACBPs (eudicot versus monocot plants, 127 Mya); plant ACBPs (gymnosperms versus monocots and eudicots, 450 Mya); plant ACBPs (moss versus gymnosperms, monocots and eudicots, 700 Mya) [26]; Apicomplexa ACBPs (Coccidia versus piroplamida, 502.5 Mya) [27]; filamentous fungi (Pyrenomycetes versus plectomycetes, 670 Mya) [28]; and yeast (*Schizosaccharomyces pombe* versus *Saccharomyces* genus, 1140 Mya) [28].

RESULTS

The chordate ACBP gene family

A BLAST analysis, using the amino acid sequences of L-H.sap (human L-ACBP), B-A.pla (duck B-ACBP) and T-M.mus (mouse T-ACBP) as probes, identified single or multiple homologous sequences in all chordate species (Table S1, http://www.BiochemJ.org/bj/392/bj3920299add.htm). Only a single homologous sequence was identified in the completed genome of the urochordate, *Ciona intestinalis*, a sea squirt.

Phylogenetic analysis was performed on the multiply aligned chordate ACBP sequences (Figure 1), by protein-maximum likelihood, using the *C. intestinalis* (C.int) ACBP as an outgroup (Figure 2). This analysis revealed the presence of three major clusters corresponding to separate ACBP gene subfamilies. One subfamily contained orthologues of the duck B-ACBP which were identified in all vertebrate species ranging from fish to higher primates (Table S1, http://www.BiochemJ.org/bj/392/bj3920299add.htm). In humans and chimpanzees, two closely related B-ACBP paralogues were identified (B-H.sap1, B-P.tro1 and B-H.sap2, B-P.tro2) (Table S1 and Figure 2). The absence of a second B-ACBP paralogue in other mammalian species suggests that this gene evolved after the divergence of higher primates. A second ACBP subfamily contained orthologues of human L-ACBP which were also identified in all vertebrate species and the cephalochordate *Branchiostoma floridae* (Table S1). In mouse and fish, two L-ACBP paralogues were identified (L-M.mus1, L-M.mus2 and L-O.myk1, L-O.myk2). The separate clustering of the murine and fish paralogues suggested that these genes derived from independent lineage-specific duplication events, which have not been detected in other species. A third ACBP subfamily contained orthologues of murine T-ACBP and were identified in only a limited number of mammalian species. The T-ACBP subfamily branched from within the L-ACBP cluster, but clustered separately from the L-ACBP sequences. Although previous studies have identified T-ACBP-related pseudogenes in various primate species [15], we obtained no evidence for orthologous sequences in lower organisms, including fish, amphibians or birds.

Our finding of only single ACBP gene homologues in the genomes of several invertebrates, including the urochordate sea squirt, *C. intestinalis*, suggests that the genetic events generating the paralogous chordate ACBP subfamilies occurred after the divergence of the vertebrate and invertebrate lineages. Members of the L-ACBP and B-ACBP subfamilies exhibit conserved exon/intron gene structures ([35], and results not shown), suggesting that they have derived from a gene duplication event, which must have occurred prior to the divergence of fish and higher vertebrates 450 Mya [28]. The presence of a putative L-ACBP orthologue in the incomplete genome of the cephalochordate *B. floridae* suggests that the gene duplication event may have occurred after divergence of the urochordates, but prior to the divergence of cepholochordates and higher animals. All known members of the T-ACBP gene subfamily lack the conserved exon/intron structure found in the L-ACBP and B-ACBP subfamilies. The lack of conserved introns and the sequence similarity to the L-ACBP subfamily suggests an origin from a retrotransposition of an ancestral L-ACBP mRNA transcript [15]. The inability to identify a T-ACBP orthologue in birds and fish in our study suggests that the retrotransposition event occurred

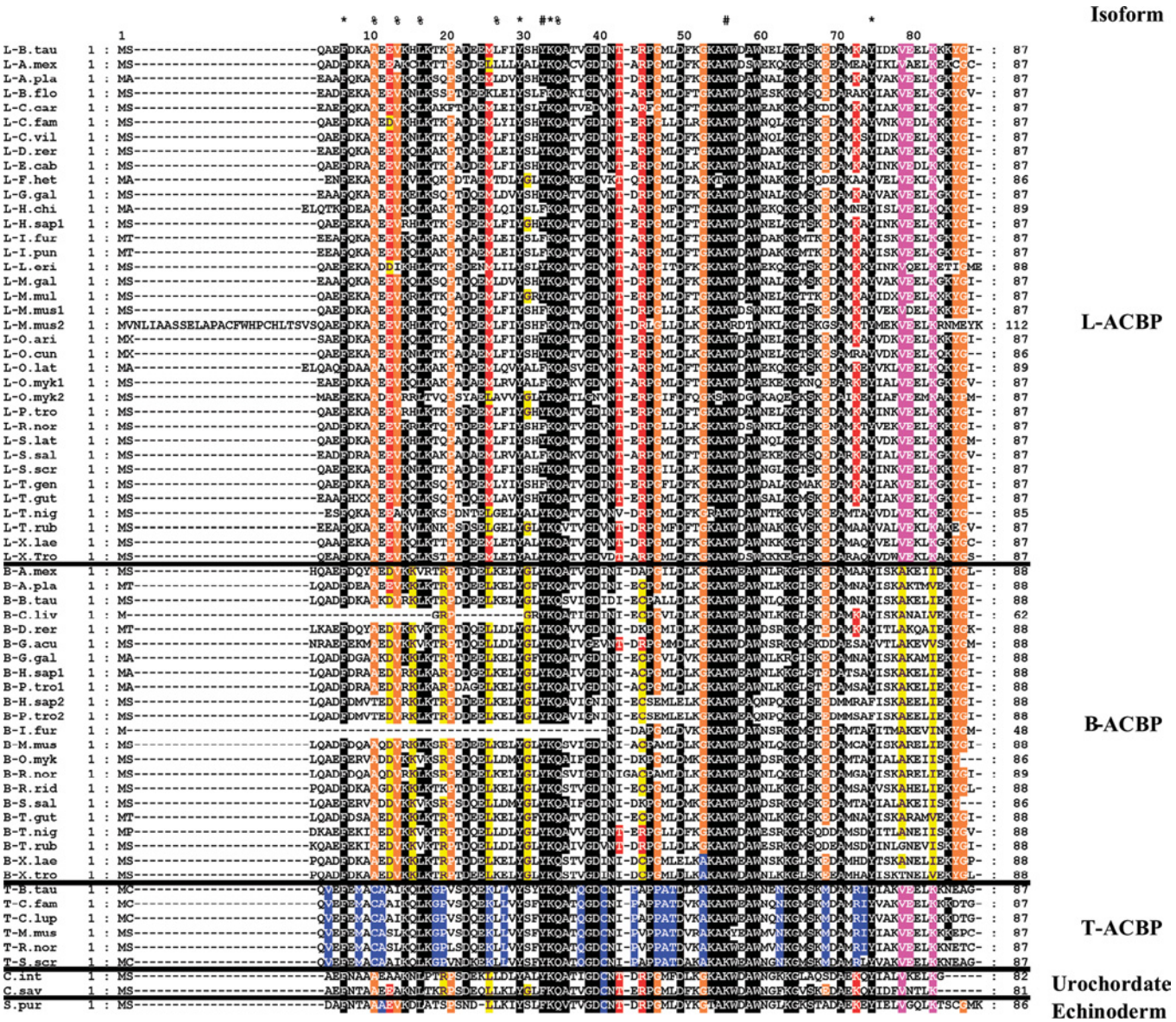

Figure 1 Sequence comparison of the vertebrate, urochordate and echinoderm *ACBP* gene family

Shading illustrates conserved residues: black, all sequences.; orange, L-ACBP and B-ACBP paralogues; purple, L-ACBP and T-ACBP; red, L-ACBP-specific; yellow, B-ACBP-specific; blue, T-ACBP-specific. Residues important for ACBP stability, ligand binding or both are marked by %, # and * respectively (see text).

after divergence of the avian and reptile lineages (≈310 Mya), but prior to the radiation of mammals. Similarly, the second paralogues, identified within the L-ACBP subfamily in mouse and the B-ACBP subfamily in higher primates, lacked the conserved intron/exon structure suggesting similar origins from lineage-specific retrotransposition events.

The multiple sequence alignments of the members of the chordate L-, B- and T-ACBP gene subfamilies revealed a strong conservation of 26 amino acid positions (shaded in black, Figure 1). Many of these residues have been shown to be important for proper ligand binding (marked # in Figure 1) and protein stability (marked %) or both binding and stability (marked *) [9,29]. In addition to the residues conserved within all three chordate ACBP subfamilies, conservation of subfamily-specific residues was also apparent. Five residues were highly conserved within the L-ACBP subfamily (Table 1; shaded in red, Figure 1). Eight residues were highly conserved within the B-ACBP subfamily (Table 1; shaded in yellow, Figure 1), and nineteen residues were highly conserved within the T-ACBP subfamily (Table 1; shaded in blue, Figure 1). The L- and B-ACBP subfamilies shared eight conserved residues which were not present in the T-ACBP subfamily (shaded in orange, Figure 1). Similarly, the L-ACBP and T-ACBP subfamilies shared three conserved residues Val-78, Glu-79 and Lys-82 that were not found in the B-ACBP subfamily (shaded purple, Figure 1). The B-ACBP and T-ACBP subfamilies shared only one residue, Ile-42, which was not shared with the L-ACBPs. We have shown that the L-ACBP, B-ACBP and T-ACBP paralogues in mice bind acyl-CoA but have different abilities to complement growth of ACBP yeast knockouts (J. Knudsen and N. J. Færgeman, unpublished work). This suggests that the sequence differences in each ACBP vertebrate subfamily are important for specific interactions with

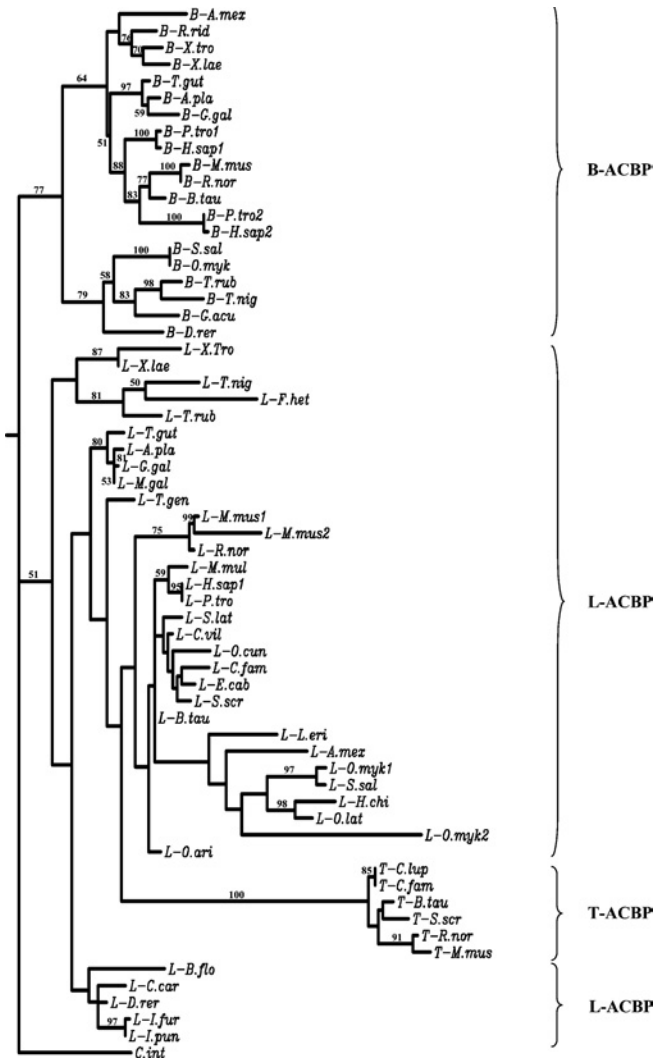

Figure 2 Phylogeny of the vertebrate ACBP gene family

C. intestinalis (C.int) was used as outgroup. Bootstrap support of 50 % or greater is indicated. Ln likelihood = − 3529.28254).

Table 1 Amino acid residues conserved in the vertebrate ACBP paralogues

Paralogue	Paralogue specific sites
B-ACBP	Asp-12, Lys-15, Arg-19, Leu-25, Gly-30, Cys-44, Ala-78, Ile-82/Val-82
L-ACBP	Glu-12, Met-25, Thr-42, Arg-44, Lys-72
T-ACBP	Val-4, Met-8, Cys-10, Ala-11, Gly-18, Pro-19, Lys-25, Leu-27, Gln-37, Cys-40, Pro-43, Pro-46, Ala-47, Thr-48, Ala-52, Asn-62, Met-68, Arg-72, Ile-73
L- + B-ACBP	Ala-10, Val-13, Pro-20, Gly-46, Gly-52, Glu-68, Tyr-85, Gly-86
L- + T-ACBP	Val-78, Glu-79, Lys-82

different acyl-CoA-utilizing proteins and thus provide separate functionalities.

The single ACBP sequences identified in the urochordates, *C. intestinalis* (C.int) and *Ciona savignyi* (C.sav), and the echinoderm, *Strongylocentrotus pupuratus* (S.pup), revealed a mixture of B-, L- and T-ACBP-specific residues with no clear orthologous relationship to a specific vertebrate ACBP subfamily (Figure 1).

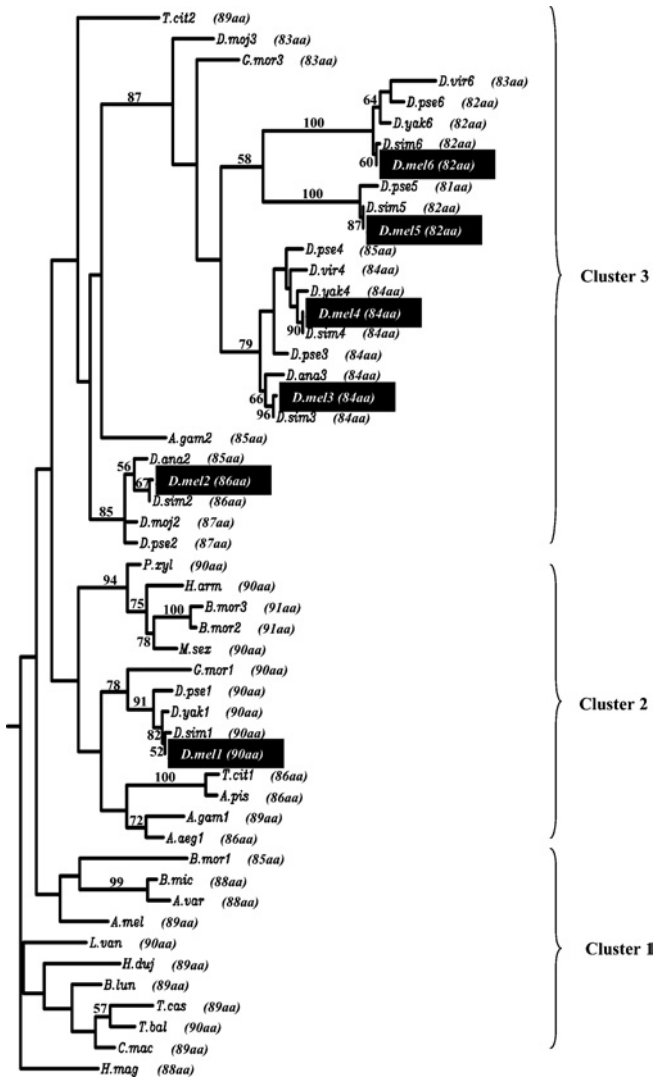

Figure 3 Phylogeny of the arthropod ACBP gene family

Bootstrap support of 50 % or greater is indicated. The hydra *H. magnipapillata* (H.mag) was used as outgroup (ln likelihood = − 3726.66918). The *D. melanogaster* paralogues are indicated with shaded boxes.

The arthropod ACBP gene family

A search of the completed genome of the fruit fly, *Drosophila melanogaster*, revealed six distinct ACBP paralogues, herein designated as D.mel1, D.mel2, D.mel3, D.mel4, D.mel5 and D.mel6 (Table S2, http://www.BiochemJ.org/bj/392/bj3920299add.htm). Similarly, six ACBP paralogues were also identified in the incomplete *Drosophila pseudoobscura* genome (Table S2). However, only two ACBP paralogues were identified in the complete genome of the distantly related malaria mosquito, *Anopheles gambiae*. Additional arthropod genomes with incomplete sequence coverage yielded from one to three ACBP sequences (see Table S2).

Phylogenetic analysis using the hydra *Hydra magnipapillata* (H.mag) as outgroup showed that the arthropod ACBP sequences grouped into three definable clusters (Figure 3). Cluster 1, which was most related to the outgroup, contained the single ACBP sequences identified in the lower arthropods, including shrimp and tardigrades. The closely related cluster 2 contained

D.mel1 and other closely related ACBP sequences that were found in all winged insects (Pterygota). As in the sequences in cluster 1, the majority of the sequences in cluster 2 were 88–91 aa (amino acids) in length. Two silkworm sequences, B.mor2 and B.mor3, were identified which clustered closely together within cluster 2, suggesting that these paralogues were generated by a lineage-specific duplication event.

Cluster 3 is a supercluster containing additional paralogues in numerous Pterygota species. Some species had a single paralogue in cluster 3, including the mosquito (A.gam2) and tetse fly (G.mor3), whereas other species had multiple paralogues. The presence of paralogues in both cluster 2 and cluster 3 in both Diptera and Hemiptera species suggests that an ACBP gene duplication event occurred prior to the divergence of these two arthropod lineages. Five *Drosophila* ACBP subfamilies were identified within cluster 3, consisting of the 81–86 aa residue paralogues D.mel2, D.mel3, D.mel4, D.mel5 and D.mel6 of *D. melanogaster* and orthologous sequences in other *Drosophila* species. Within cluster 3, the *Drosophila* subfamilies showed distinct clustering hierarchies. The *Drosophila* subfamilies, ACBP3 and ACBP4, containing D.mel3 and D.mel4 respectively, clustered together whereas the ACBP5 and ACBP6 subfamilies, containing D.mel5 and D.mel6 respectively, clustered together, showing evidence for evolutionary relatedness. Interestingly, the D.mel4 ACBP gene is intronless, suggesting an origin from a retrotransposition of an mRNA transcript of an ancestor of the closely related ACBP3 subfamily (results not shown). These results indicate that multiple ACBP duplication events have occurred prior to the divergence of the different *Drosophila* lineages.

Alignment of the *Drosophila* ACBP paralogues with the L-ACBP and B-ACBP paralogues from humans revealed a strong conservation of the residues shown to be important for ACBP stability and acyl-CoA binding [29]. The *Drosophila* ACBP1 subfamily, containing D.mel1, showed the most similarity to the vertebrate ACBP homologues (Figure S1, http://www.BiochemJ.org/bj/392/bj3920299add.htm). However, no direct orthologous relationship was evident between the *Drosophila* and vertebrate ACBP subfamilies suggesting that these ACBP lineages have evolved independently after divergence of invertebrates and vertebrates. We have observed that the *D. melanogaster* ACBP paralogues exhibit very different abilities to complement growth, vacuole structure and ceramide synthesis in ACBP-depleted yeast, (J. Knudsen and N. J. Færgeman, unpublished work) suggesting that they have evolved different biological functions. This is supported by the observations that D.mel3 and D.mel6 are differentially upregulated, in contrast with D.mel1, D.mel2, D.mel4 and D.mel5, following bacterial infection [30], and that the D.mel2 level is increased in DDT [1,1,1-trichloro-2,2-bis-(*p*-chlorophenyl)ethane]-resistant *D. melanogaster* [31].

ACBP gene family in nematodes and other lower metazoans

A BLASTP search using the *Caenorhabditis elegans* ACBP homologue (C.ele1) as probe identified three additional ACBP paralogues in this organism: C.ele2, C.ele3 and C.ele4, consisting of 116, 146 and 115 aa respectively (Table S3, http://www.BiochemJ.org/bj/392/bj3920299add.htm). A BLAST search using these four sequences as probes identified four orthologous sequences in the completed *Caenorhabditis briggsae* genome and one or two sequences in a number of other nematode species (Table S3). Single ACBP homologues were identified in the platyhelminthes *Schistosoma japonicum* (S.jap), *Echinococcus granulosus* (E.gra), and in the cnidaria *H. magnipapillata* (H.mag) and *Hydra oligactis* (H.oli) (Table S3).

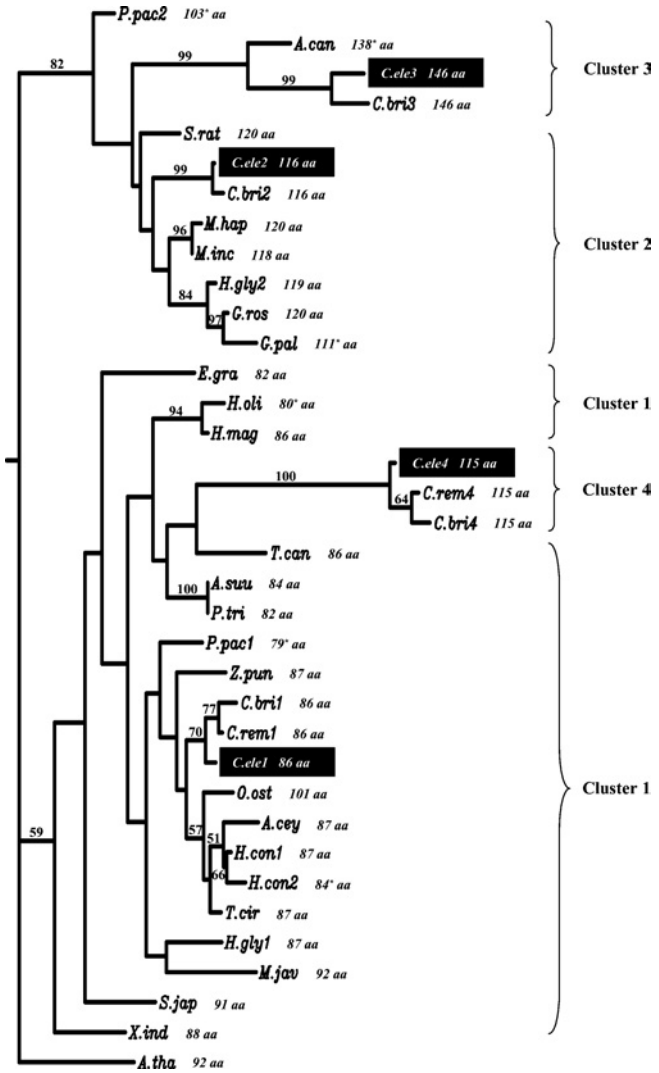

Figure 4 Phylogeny of the hydra, trematode and nematode ACBP gene family

A. thaliana was used as outgroup (ln likelihood = −4456.87077). Bootstrap support of 50% or greater is indicated. For each taxa the length of the amino acid sequence is shown; * indicates a partial sequence.

Phylogenetic analysis using the plant *Arabidopsis thaliana* ACBP (A.tha) as outgroup revealed four major clusters of lower metazoan ACBP sequences (Figure 4). Cluster 1 contained the *Caenorhabditis* orthologues, C.ele1 and C.br1 (86 aa) and similar sized ACBP sequences (82–99 aa) from other nematodes, platyhelminth and cnidaria species. Cluster 2 contained a number of longer nematode ACBP sequences (116–120 aa), closely related to the C.ele2 and C.bri2 orthologues (116 aa). Cluster 3, a distantly related outgroup from cluster 2, contained the 146 aa C.ele3 and C.bri3 orthologues and the incomplete (> 138 aa) dog hookworm sequence (A.can). Cluster 4, a distantly related outgroup from cluster 1, included the 115 aa C.ele4, C.bri4 and C.rem4 orthologues. Cluster 2, 3 and 4 sequences were detected in a number of nematode species, but were absent from other lower metazoan species. This suggests that the four ACBP subfamilies detected in *C. elegans* and other nematodes have evolved through nematode-specific duplication events after the divergence of nematodes, platyhelminthes and cnidaria. Unlike the vertebrate or arthropod ACBP subfamilies, the nematode ACBP

subfamilies have undergone significant changes in protein size and sequence similarity (Figure 4). As the complete genome sequences have not been determined for any of the platyhelminth or cnidaria species, it is not clear at this time whether these phyla have undergone similar lineage-specific duplications of the ACBP gene.

Alignment of the *Caenorhabditis* members of the four nematode ACBP subfamilies with the human members of the vertebrate L-ACBP and B-ACBP subfamilies revealed a strong conservation of residues important for ACBP stability and acyl-CoA binding (Figure S2, http://www.BiochemJ.org/bj/392/bj3920299add.htm). The *C. elegans* member of the nematode ACBP1 subfamily, C.ele1, was nearly identical in length (86 versus 87 aa) and was equally similar in sequence to both human L- and B-ACBP subfamily members. We have shown that C.ele1 can functionally complement the yeast ACBP (J. Knudsen and N. J. Færgeman, unpublished work). Furthermore, others have shown that RNAi (RNA interference) knockdown of C.ele1 causes a reduction of whole body fat content [32]. These findings are consistent with the hypothesis that C.ele1 has maintained the basic functional properties of a prototypical ACBP. The C.ele2 ACBP paralogue is highly expressed in the oocyte and during embryonic and L1 larval stages [33], while the C.ele3 ACBP paralogue is only expressed during the L3 and L4 stages and the reproductive stages of the adult hermaphrodite [33]. The significant difference in amino acid sequence and the differential expression of these two ACBP paralogues during development suggest that they perform different functions related to acyl-CoA binding. The members of the nematode ACBP2 and ACBP3 subfamilies all contained a single proline insertion after Leu-16 and a three-amino-acid insertion after Pro-19 (see Figure S2). The members of the nematode ACBP4 subfamily contain a unique insertion of three additional amino acids further downstream between amino acids 44 and 45 (human L-ACBP numbering). These changes in primary structure may significantly alter the ability of these paralogues to bind acyl-CoA in comparison with the C.ele1 ACBP prototype. In this connection, it is interesting that a two-residue extension of loop one in *Plasmodium falciparum* ACBP has altered the size of the hydrophobic binding pocket and broadened the ligand specificity range for both shorter and longer acyl-CoA esters compared with bovine L-ACBP [34].

Non-metazoan ACBP gene families

A BLAST search against the completed fungal genomes of *Neurospora crassa* (N.cra), *Magnaphorthe grisea* (M.gri), *Gibberella zeae* (G.zea), *Ustilago maydis* (U.may) and *S. pombe* (S.pom) using the known *S. cerevisiae* ACBP [35–37] as probe identified only a single ACBP sequence in each of these species (Table S4, http://www.BiochemJ.org/bj/392/bj3920299add.htm). Similarly, only single ACBP sequences were identified in the incomplete genomes of twenty fungal species, three blight species *Phytophthora infestans* (P.inf), *Phytophthora nicotianae* (P.nic) and *Phytophthora sojae* (P.soj) and the cellular slime mould, *Dictyostelium discoideum* (Table S4). The ACBP homologues from the different yeast species were similar in length to the mammalian ACBP homologues (86–87 aa residues), whereas the ACBP homologues from the basidimycota (P.inv and U.may) and filamentous fungi (B.gra, B.fuc, C.bas, F.spo, G.zea, M.gri and N.cra) were longer proteins containing 100–112 aa due to a 2–3-amino-acid N-terminal extension and a 12–17 residue C-terminal extension (Figure S3, http://www.BiochemJ.org/bj/392/bj3920299add.htm). The basidiomycotic ACBP homologues (P.inv and U.may) were further characterized by the presence of a four-residue insert in loop 1 between Thr-18 and Lys-19 (numbering from bovine L-ACBP). In contrast, the ACBP homologues of the filamentous fungi had a one-residue insert in loop 2 between Thr-42 and Glu-43 and a 2–12-residue insert in loop 3 between Lys-63 and Gly-64.

In plants, ACBP homologues have been previously identified in *A. thaliana* (A.tha, 92 aa) [38], and *D. lanata* (D.lan1, 90 aa; D.lan2, 92 aa) [12]. A BLAST search of the completed *A. thaliana* and rice genomes using the *A. thaliana* ACBP sequence as probe failed to identify any additional ACBP paralogues in *A. thaliana* but revealed the presence of three ACBP paralogues in rice (Table S5, http://www.BiochemJ.org/bj/392/bj3920299add.htm). A search of the incomplete algae and plant genomes and EST databases identified single ACBP homologues in a number of species (Table S5). Although the majority of the plant ACBP homologues, including A.tha, D.lan1, D.lan2 and O.sat1, were structurally similar to the mammalian ACBPs with protein lengths of 87–93 aa, the two additional ACBP paralogues in rice, O.sat2 (115 aa) and O.sat3 (155 aa), had a 24-residue N-terminal and 64-residue C-terminal extension respectively (Figure S4, http://www.BiochemJ.org/bj/392/bj3920299add.htm). Phylogenetic analysis of the multiple paralogues identified in *D. lanata* and *O. sativa* failed to reveal a clear clustering across species, suggesting that the multiple paralogues derived from lineage-specific duplication events after the divergence of the different plant species (results not shown).

A BLAST search using the known ACBP homologues of *Trypanosoma brucei* [39] (T.bru) and the African malaria mosquito parasite, *P. falciparum* [34] (P.fal1), revealed four distinct 90 aa ACBP paralogues in *P. falciparum 3D7* (P.fal1, P.fal2, P.fal3 and P.fal4) and a single 89 aa ACBP homologue in *Plasmodium yoelii yoelii* (Table S6, http://www.BiochemJ.org/bj/392/bj3920299add.htm). A search of the incomplete protozoan genomes and EST sequence databases identified single ACBP sequences of similar size in eight other species (Table S6). Phylogenetic analysis failed to reveal a clustering between the *P. falciparum* ACBP paralogues and other ACBP lineage-specific metazoan and non-metazoan ACBP subfamilies, suggesting that these genes have evolved through separate lineage-specific duplication events.

BLAST searches identified ACBP sequences in 11 soil eubacterial species (Table S7, http://www.BiochemJ.org/bj/392/bj3920299add.htm). Eight of the eubacterial species belong to the β-proteobacteria lineage, whereas the other bacterial species included the myxobacterium *Stigmatella aurantiaca* (S.aur), the flavobacterium *Cytophaga hutchinsonii* (C.hut) and the radiation-resistant bacterium *Deinococcus radiodurans* (D.rad). The eubacterial ACBP homologues were similar in size to the mammalian ACBP homologues, with sizes ranging from 84 to 89 aa. No ACBP homologues were detected in any of the other known bacterial species, or in any of the archaea. It is of interest that all of the ACBP-containing bacteria, with the exception of *Ralstonia eutropha* and *Ralstonia metallidurans*, are known to be pathogenic to plants or animals, suggesting that an *ACBP* gene could have been acquired from a eukaryotic host by horizontal gene transfer. The lack of an ACBP homologue in most bacterial species and the phylogenetic branching of the bacterial ACBP sequences from within the eukaryotic ACBP cluster would support this hypothesis.

We were unable to identify an ACBP homologue in the diplomonodida (i.e. *Giardia intestinalis*) or other amitochondrial eukaryotes, which suggests that eukaryotes acquired the ACBP gene after acquisition of the mitochondria. This suggests that ACBP is important for basal mitochondrial functions including donation of acyl-CoA esters for fatty acid β-oxidation. This hypothesis is supported by *in vitro* experimental observations that

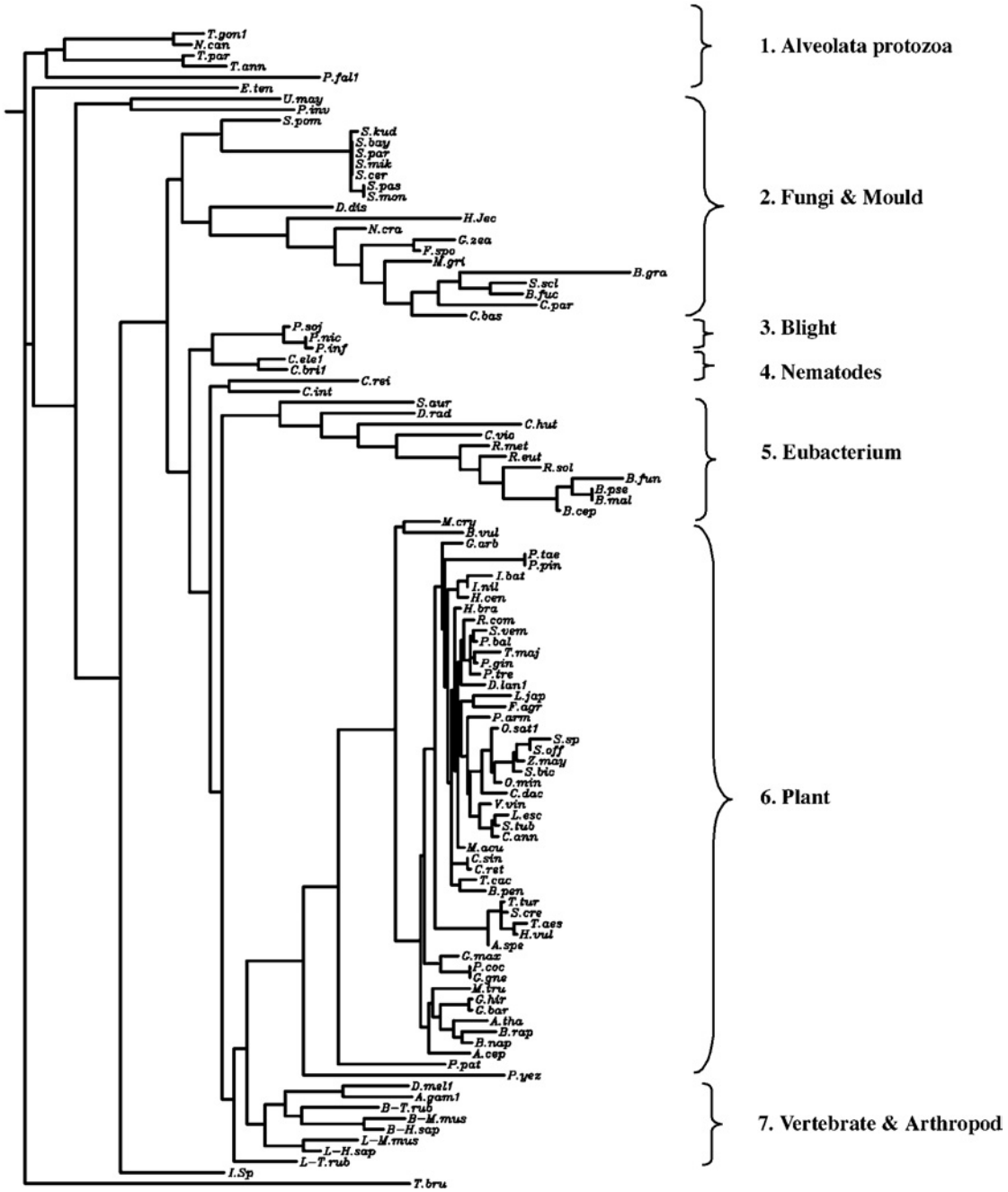

Figure 5 Phylogeny of the metazoan and non-metazoan ACBP gene family

The protozoan parasite *T. brucei* was chosen as outgroup (ln likelihood = – 8686.75013). See Tables S1 and S4–7.

CPT1 (carnitine palmitoyltransferase 1) prefers the ACBP–acyl-CoA complex as substrate over free acyl-CoA [40–42].

An alignment of the fungal, plant, protozoan and eubacterial ACBP sequences with the vertebrate ACBP sequences revealed a strong conservation of amino acids at the majority of sites determined to be important for ACBP structure and function (Figure S4, http://www.BiochemJ.org/bj/392/bj3920299add.htm).

Overall phylogeny of the *ACBP* gene family

To determine the overall evolutionary relationship between the eukaryotic and prokaryotic ACBP homologues, we aligned a single prototype ACBP sequence from the different metazoan and non-metazoan ACBP subfamilies and performed a phylogenetic analysis. The selected animal, fungi, plant, protozoa and eubacterial ACBP homologues grouped into seven clusters according to phylum or lineage (Figure 5). Cluster 1 contained the apicomplexian protozoan ACBP homologues to which *Eimeria tenella* ACBP (E.ten) associated as a distant outgroup. Cluster 2 contained the ACBP homologues from yeast, filamentous fungi and the cellular slime mould *D. discoideum* (D.dis). The basidomycotic ACBP homologues (U.may & P.inv) associated with this cluster as an outgroup, with a long branch agreeing with the early divergence of this lineage in the evolution of fungi.

Table 2 Synonymous (L_s) and non-synonymous (L_a) nucleotide substitution rates in the *ACBP* gene family

T is the divergence time to the last compared common ancestor(s) within each cluster, in Mya (see Materials and methods section) and N is the number of sequence comparisons.

	T	N	L_a	L_s	L_a/L_s
L-ACBP	450 (Fish)	86	0.24	1.76	0.14 ± 0.043
B-ACBP	450 (Fish)	30	0.28	1.81	0.16 ± 0.065
Arthropod-1	268 (Lepidoptera)	12	0.72	2.91	0.25 ± 0.080
Plant	127 (Monocot)	414	0.52	4.57	0.11 ± 0.056
Plant	246 (Gymnosperm)	87	0.41	2.78	0.15 ± 0.050
Apicomplexa	502.5	2	0.64	1.25	0.51 ± 0.171
Filamentous fungi	670 (Pyrenomycetes- Plectomycetes)	9	0.34	1.42	0.24 ± 0.089
Plant	700 (Moss)	43	0.27	1.02	0.27 ± 0.070
Yeast	1140 (Taohrinomycotina)	3	0.15	0.91	0.17 ± 0.025

Cluster 3 contained ACBP homologues from the three blight species (P.soj, P.inf and P.nic). Cluster 4 contained the nematode ACBP homologues from *C. elegans* (C.ele1) and *C. briggsae* (C.bri1); cluster 5 contained the 11 prokaryotic ACBP sequences in which all the β-proteobacteria ACBP homologues were grouped together. Cluster 6 contained the ACBP homologues from plants to which the moss (P.pat) and the seaweed (P.yez) were associated, with long branch length agreeing with the expected evolutionary pattern of species. Cluster 7 contained the vertebrate L- and B-ACBP homologues and the arthropod ACBP orthologues (D.mel1 and A.gam1). Other ACBP sequences appeared as orphans with no significant association with any of the seven clusters, such as the ACBP from *Isotricha sp.* (I.sp), which is distantly related to the apicomplexa protozoan ACBP cluster (cluster 1).

Examination of the phylogenetic tree obtained from analysing the aligned sequences of multiple ACBP paralogues from different species failed to reveal any clear association of ACBP paralogues across distantly related kingdoms or phyla (results not shown). This supports the notion that evolution of ACBP paralogues has occurred through lineage-specific duplication events.

The *ACBP* gene experienced a purifying selection during evolution

To determine the constraints on the *ACBP* gene evolution, we calculated the L_s and L_a from the nucleotide substitution distances to the last common ancestor during comparable evolutionary periods, as described in the Materials and methods section. The vertebrate *L-ACBP* and *B-ACBP* and the arthropod *ACBP1* genes had an average L_a/L_s-ratio of 0.14, 0.16 and 0.25, demonstrating that these animal *ACBP* genes were subjected to a strong purifying selection (Table 2). Similarly low average L_a/L_s-ratios were found among plants when comparing monocots with eudicots and when comparing gymnosperms with ACBPs from higher plants (monocots and eudicots), which had L_a/L_s-ratios of 0.11 and 0.15 respectively (Table 2). However, a comparison of moss ACBP with higher ACBPs showed a slightly increased average L_a/L_s-ratio of 0.27 (Table 2). The ACBP genes from fungi were also found to be subjected to a purifying selection with average L_a/L_s-ratios of 0.17 and 0.24 when *S. pombe* was compared with orthologue ACBPs from the *Saccharomyces* genus and for filamentous fungi when pyrenomyces and plectomycetes were compared respectively (Table 2). These results show that these *ACBP* genes are highly conserved and that non-synonymous substitutions are deleterious and are selected against. The low L_a/L_s-ratio obtained for the L-ACBP cluster supports primary structural information in which bovine L-ACBP possesses 26 highly conserved sequence positions that are essential for protein stability, folding and for binding of acyl-CoA esters [9].

The low L_a/L_s-ratio of 0.14–0.16 for the vertebrate ACBP paralogues, across species since divergence from fish at 450 Mya, were consistent with the selection pressure found for other genes involved in fatty-acid transport, $L_a/L_s = 0.12$–0.27 [43], and in fatty acid and triacylglycerol metabolism between humans and mice ($L_a/L_s = 0.199$) [44].

In contrast with the low L_a/L_s-ratio for other ACBP genes, the apicomplexian ACBPs were found to have a significantly higher L_a/L_s-ratio of 0.51 (Table 2), when comparing the *ACBP* genes from the *Plasmodium* genus with orthologous genes from other apicomplexian species (see the Materials and methods section). This suggests that the *ACBP* genes within this cluster are subjected to relaxed evolutionary constraints, allowing changes in the primary structure (Table 2).

What then would be the advantage to the few bacterial species that have acquired an *ACBP* gene? ACBP has been shown to donate acyl-CoA during phospholipid synthesis [40] and to be required for membrane assembly, fatty acid chain elongation and ceramide synthesis in yeast [3,5]. In *S. aurantiaca* the ACBP open reading frame flanks an approximately 65 kbp *sti* gene cluster encoding a bacterial modular type I polyketide synthase [45]. This enzyme synthesizes aromatic polyketides from activated CoA ester primers (i.e. isobutyryl-CoA or 2-methylbutyryl-CoA) [45]. Under anaerobic conditions or during exposure to an excessive carbon source, *R. metallidurans* and *R. eutropha* are able to synthesize polyhydroxyalkanoates by condensation of two acetyl-CoA molecules, leading to formation of 3-hydroxybutyryl-CoA. Acquisition of ACBP may have been beneficial for synthesis of the lipid components needed for the above-mentioned bacterial processes and thereby induce a selective pressure within these species.

REFERENCES

1 Guidotti, A., Forchetti, C. M., Corda, M. G., Konkel, D., Bennett, C. D. and Costa, E. (1983) Isolation, characterization, and purification to homogeneity of an endogenous polypeptide with agonistic action on benzodiazepine receptors. Proc. Natl. Acad. Sci. U.S.A. **80**, 3531–3535

2 Knudsen, J., Burton, J. and Færgeman, N. J. (2004) Long chain acyl-CoA esters and acyl-CoA binding protein (ACBP) in cell function. In Lipobiology (van der Vusse, G., ed.), pp. 123–153, Elsevier B.V., Amsterdam

3 Faergeman, N. J., Feddersen, S., Christiansen, J. K., Larsen, M. K., Schneiter, R., Ungermann, C., Mutenda, K., Roepstorff, P. and Knudsen, J. (2004) Acyl-CoA-binding protein, Acb1p, is required for normal vacuole function and ceramide synthesis in Saccharomyces cerevisiae. Biochem. J. **380**, 907–918

4 Mandrup, S., Færgeman, N. J. and Knudsen, J. (2004) Structure, function and phylogeny of acyl-CoA binding protein. In Cellular Proteins and Their Fatty Acids in Health and Disease (Duttaroy, A. K. and Spener, F., eds.), pp. 151–172, Wiley-VCH gmbH & Co. Kaga, Weinheim

5 Gaigg, B., Neergaard, T. B., Schneiter, R., Hansen, J. K., Faergeman, N. J., Jensen, N. A., Andersen, J. R., Friis, J., Sandhoff, R., Schroder, H. D. et al. (2001) Depletion of acyl-coenzyme A-binding protein affects sphingolipid synthesis and causes vesicle accumulation and membrane defects in Saccharomyces cerevisiae. Mol. Biol. Cell **12**, 1147–1160

6 Andersen, K. V., Mandrup, S., Knudsen, J. and Poulsen, F. M. (1991) The secondary structure in solution of acyl-coenzyme A binding protein from bovine liver using 1H nuclear magnetic resonance spectroscopy. Biochemistry **30**, 10654–10663

7 Andersen, K. V. and Poulsen, F. M. (1992) Three-dimensional structure in solution of acyl-coenzyme A binding protein from bovine liver. J. Mol. Biol. **226**, 1131–1141

8 Kragelund, B. B., Andersen, K. V., Madsen, J. C., Knudsen, J. and Poulsen, F. M. (1993) Three-dimensional structure of the complex between acyl-coenzyme A binding protein and palmitoyl-coenzyme A. J. Mol. Biol. **230**, 1260–1277

9 Kragelund, B. B., Hojrup, P., Jensen, M. S., Schjerling, C. K., Juul, E., Knudsen, J. and Poulsen, F. M. (1996) Fast and one-step folding of closely and distantly related homologous proteins of a four-helix bundle family. J. Mol. Biol. **256**, 187–200

10 Mogensen, I. B., Schulenberg, H., Hansen, H. O., Spener, F. and Knudsen, J. (1987) A novel acyl-CoA-binding protein from bovine liver. Effect on fatty acid synthesis. Biochem. J. **241**, 189–192

11 Owens, G. P., Sinha, A. K., Sikela, J. M. and Hahn, W. E. (1989) Sequence and expression of the murine diazepam binding inhibitor. Brain Res. Mol. Brain Res. **6**, 101–108

12 Metzner, M., Ruecknagel, K. P., Knudsen, J., Kuellertz, G., Mueller-Uri, F. and Diettrich, B. (2000) Isolation and characterization of two acyl-CoA-binding proteins from proembryogenic masses of Digitalis lanata Ehrh. Planta **210**, 683–685

13 Mandrup, S., Hummel, R., Ravn, S., Jensen, G., Andreasen, P. H., Gregersen, N., Knudsen, J. and Kristiansen, K. (1992) Acyl-CoA-binding protein/diazepam-binding inhibitor gene and pseudogenes. A typical housekeeping gene family. J. Mol. Biol. **228**, 1011–1022

14 Gersuk, V. H., Rose, T. M. and Todaro, G. J. (1995) Molecular cloning and chromosomal localization of a pseudogene related to the human acyl-CoA binding protein/diazepam binding inhibitor. Genomics **25**, 469–476

15 Ivell, R., Pusch, W., Balvers, M., Valentin, M., Walther, N. and Weinbauer, G. (2000) Progressive inactivation of the haploid expressed gene for the sperm-specific endozepine-like peptide (ELP) through primate evolution. Gene **255**, 335–345

16 Chye, M. L. (1998) Arabidopsis cDNA encoding a membrane-associated protein with an acyl-CoA binding domain. Plant Mol. Biol. **38**, 827–838

17 Chye, M. L., Huang, B. Q. and Zee, S. Y. (1999) Isolation of a gene encoding Arabidopsis membrane-associated acyl-CoA binding protein and immunolocalization of its gene product. Plant J. **18**, 205–214

18 Chye, M. L., Li, H. Y. and Yung, M. H. (2000) Single amino acid substitutions at the acyl-CoA-binding domain interrupt 14[C]palmitoyl-CoA binding of ACBP2, an Arabidopsis acyl-CoA-binding protein with ankyrin repeats. Plant Mol. Biol. **44**, 711–721

19 Geisbrecht, B. V., Zhang, D., Schulz, H. and Gould, S. J. (1999) Characterization of PECI, a novel monofunctional Delta(3), Delta(2)-enoyl-CoA isomerase of mammalian peroxisomes. J. Biol. Chem. **274**, 21797–21803

20 Suk, K., Kim, Y. H., Hwang, D. Y., Ihm, S. H., Yoo, H. J. and Lee, M. S. (1999) Molecular cloning and expression of a novel human cDNA related to the diazepam binding inhibitor. Biochim. Biophys. Acta **1454**, 126–131

21 Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. and Higgins, D. G. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **25**, 4876–4882

22 Felsenstein, J. (1989) PHYLIP – Phylogeny Inference Package (Version 3.2). Cladistics **5**, 164–166

23 Kumar, S., Tamura, K., Jakobsen, I. B. and Nei, M. (2001) MEGA2: molecular evolutionary genetics analysis software. Bioinformatics **17**, 1244–1245

24 Kumar, S. and Hedges, S. B. (1998) A molecular timescale for vertebrate evolution. Nature (London) **392**, 917–920

25 Burmester, T. (2001) Molecular evolution of the arthropod hemocyanin superfamily. Mol. Biol. Evol. **18**, 184–195

26 Heckman, D. S., Geiser, D. M., Eidell, B. R., Stauffer, R. L., Kardos, N. L. and Hedges, S. B. (2001) Molecular evidence for the early colonization of land by fungi and plants. Science (Washington, D.C.) **293**, 1129–1133

27 Escalante, A. A. and Ayala, F. J. (1994) Phylogeny of the malarial genus Plasmodium, derived from rRNA gene sequences. Proc. Natl. Acad. Sci. U.S.A. **91**, 11373–11377

28 Hedges, S. B. (2002) The origin and evolution of model organisms. Nat. Rev. Genet. **3**, 838–849

29 Kragelund, B. B., Poulsen, K., Andersen, K. V., Baldursson, T., Kroll, J. B., Neergard, T. B., Jepsen, J., Roepstorff, P., Kristiansen, K., Poulsen, F. M. et al. (1999) Conserved residues and their role in the structure, function, and stability of acyl-coenzyme A binding protein. Biochemistry **38**, 2386–2394

30 De Gregorio, E., Spellman, P. T., Rubin, G. M. and Lemaitre, B. (2001) Genome-wide analysis of the Drosophila immune response by using oligonucleotide microarrays. Proc. Natl. Acad. Sci. U.S.A. **98**, 12590–12595

31 Pedra, J. H., McIntyre, L. M., Scharf, M. E. and Pittendrigh, B. R. (2004) Genome-wide transcription profile of field- and laboratory-selected dichlorodiphenyltrichloroethane (DDT)-resistant Drosophila. Proc. Natl. Acad. Sci. U.S.A. **101**, 7034–7039

32 Ashrafi, K., Chang, F. Y., Watts, J. L., Fraser, A. G., Kamath, R. S., Ahringer, J. and Ruvkun, G. (2003) Genome-wide RNAi analysis of Caenorhabditis elegans fat regulatory genes. Nature (London) **421**, 268–272

33 Hill, A. A., Hunter, C. P., Tsung, B. T., Tucker-Kellogg, G. and Brown, E. L. (2000) Genomic analysis of gene expression in C. elegans. Science (Washington D.C.) **290**, 809–812

34 van Aalten, D. M., Milne, K. G., Zou, J. Y., Kleywegt, G. J., Bergfors, T., Ferguson, M. A., Knudsen, J. and Jones, T. A. (2001) Binding site differences revealed by crystal structures of Plasmodium falciparum and bovine acyl-CoA binding protein. J. Mol. Biol. **309**, 181–192

35 Rose, T. M., Schultz, E. R. and Todaro, G. J. (1992) Molecular cloning of the gene for the yeast homologue (ACB) of diazepam binding inhibitor/endozepine/acyl-CoA-binding protein. Proc. Natl. Acad. Sci. U.S.A. **89**, 11287–11291

36 Knudsen, J., Faergeman, N. J., Skott, H., Hummel, R., Borsting, C., Rose, T. M., Andersen, J. S., Hojrup, P., Roepstorff, P. and Kristiansen, K. (1994) Yeast acyl-CoA-binding protein: acyl-CoA-binding affinity and effect on intracellular acyl-CoA pool size. Biochem. J. **302**, 479–485

37 Borsting, C., Hummel, R., Schultz, E. R., Rose, T. M., Pedersen, M. B., Knudsen, J. and Kristiansen, K. (1997) Saccharomyces carlsbergensis contains two functional genes encoding the acyl-CoA binding protein, one similar to the ACB1 gene from S. cerevisiae and one identical to the ACB1 gene from S. monacensis. Yeast **13**, 1409–1421

38 Engeseth, N. J., Pacovsky, R. S., Newman, T. and Ohlrogge, J. B. (1996) Characterization of an acyl-CoA-binding protein from Arabidopsis thaliana. Arch. Biochem. Biophys. **331**, 55–62

39 Milne, K. G. and Ferguson, M. A. (2000) Cloning, expression, and characterization of the acyl-CoA-binding protein in African trypanosomes. J. Biol. Chem. **275**, 12503–12508

40 Rasmussen, J. T., Faergeman, N. J., Kristiansen, K. and Knudsen, J. (1994) Acyl-CoA-binding protein (ACBP) can mediate intermembrane acyl-CoA transport and donate acyl-CoA for beta-oxidation and glycerolipid synthesis. Biochem. J. **299**, 165–170

41 Bhuiyan, A. K., Murthy, M. S. and Pande, S. V. (1994) Some properties of the malonyl-CoA sensitive carnitine long/medium chain acyltransferase activities of peroxisomes and microsomes of rat liver. Biochem. Mol. Biol. Int. **34**, 493–503

42 Abo-Hashema, K. A., Cake, M. H., Lukas, M. A. and Knudsen, J. (2001) The interaction of acyl-CoA with acyl-CoA binding protein and carnitine palmitoyltransferase I. Int. J. Biochem. Cell Biol. **33**, 807–815

43 Schaap, F. G., van der Vusse, G. J. and Glatz, J. F. (2002) Evolution of the family of intracellular lipid binding proteins in vertebrates. Mol. Cell. Biochem. **239**, 69–77

44 Kitami, T. and Nadeau, J. H. (2002) Biochemical networking contributes more to genetic buffering in human and mouse metabolic pathways than does gene duplication. Nat. Genet. **32**, 191–194

45 Gaitatzis, N., Silakowski, B., Kunze, B., Nordsiek, G., Blocker, H., Hofle, G. and Muller, R. (2002) The biosynthesis of the aromatic myxobacterial electron transport inhibitor stigmatellin is directed by a novel type of modular polyketide synthase. J. Biol. Chem. **277**, 13082–13090

Received 25 April 2005/30 June 2005; accepted 14 July 2005
Published as BJ Immediate Publication 14 July 2005, doi:10.1042/BJ20050664