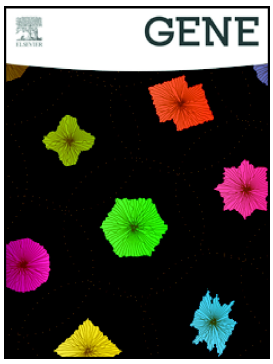
Damsteegt E, Davie A & Lokman PM (2018) The evolution of apolipoprotein B and its mRNA editing complex. Does the lack of editing contribute to hypertriglyceridemia?, *Gene*, 641, pp. 46-54.

DOI: 10.1016/j.gene.2017.10.024



The evolution of apolipoprotein B and its mRNA editing complex. Does the lack of editing contribute to hypertriglyceridemia?

L. Damsteegt Erin, Andrew Davie, P. Mark Lokman

PII: S0378-1119(17)30849-1
DOI: doi:10.1016/j.gene.2017.10.024
Reference: GENE 42243

To appear in: *Gene*

Received date: 30 May 2017
Revised date: 18 September 2017
Accepted date: 10 October 2017

Please cite this article as: L. Damsteegt Erin, Andrew Davie, P. Mark Lokman , The evolution of apolipoprotein B and its mRNA editing complex. Does the lack of editing contribute to hypertriglyceridemia?. The address for the corresponding author was captured as affiliation for all authors. Please check if appropriate. Gene(2017), doi:10.1016/j.gene.2017.10.024

The evolution of apolipoprotein B and its mRNA editing complex. Does the lack of editing contribute to hypertriglyceridemia?

Erin, L. Damsteegt[*1], Andrew Davie[2] and P. Mark Lokman[1].

(1) Department of Zoology, University of Otago, 340 Great King Street, PO Box 56, Dunedin 9054, New Zealand

(2) Institute of Aquaculture, School of Natural Sciences, University of Stirling, Stirling, FK9 4LA, United Kingdom

*Author for correspondence: Tel: +64 3 479 7986; Fax: +64 3 479 7584;

email: erin.damsteegt@otago.ac.nz

Abbreviations: activation-induced deaminase (Aid); apolipoprotein B (Apob/*apob*); apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like complex-one (Apobec1); Apobec complementation factor (Acf); apolipoprotein E (Apoe).

ABSTRACT

The evolution of apolipoprotein B (Apob) has been intensely researched due to its importance during lipid transport. Mammalian full-length *apob100* can be post-transcriptionally edited by the enzyme apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like complex-one (Apobec1) resulting in a truncated Apob, known as Apob48. Whilst both full-length and truncated forms of Apob are important for normal lipid homeostasis in mammals, there is no evidence for the presence of *apob* mRNA editing prior to the divergence of the mammals, yet, non-mammalian vertebrates appear to function normally with only Apob100. To date, the majority of the research carried out in non-mammalian vertebrates has focused on chickens with only a very limited number examining *apob* mRNA editing in fish. This study focused on the molecular evolution of Apobec1 and Apob in order to ascertain if *apob* mRNA editing occurs in eels, a basal teleost which represents an evolutionarily important animal group. No evidence for the presence of Apobec1 or the ability for eel *apob* to be edited was found. However, an important link between mutant mice and the evident hypertriglyceridemia in the plasma of non-mammalian vertebrates was made. This study has provided imperative evidence to help bridge the evolutionary gap between fish and mammals and provides further support for the lack of *apob* mRNA editing in non-mammalian vertebrates.

Keywords: apolipoprotein B, mRNA editing, Apobec1, hypertriglyceridemia

1. INTRODUCTION

The activation-induced deaminase/apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like (Aid/Apobec) family is a group of proteins belonging to the cytidine deaminase superfamily. To date, one Aid and five Apobecs have been identified in mammals (Conticello, 2008) and only one Apobec has been found in non-mammalian species (with a few notable exceptions). As their name suggests, members of this family have the ability to edit RNA. Functions other than editing have also been reported for certain members of this family, e.g., Apobec2 has been linked to tissue regeneration (Powell *et al*., 2012), Apobec3 has antiviral properties (Harris and Liddament, 2004; Pham *et al*., 2005; Cullen, 2006) and Apobec4 is possibly involved in regulating spermatogenesis (Rogozin *et al*., 2005).

Editing of apob mRNA results in the formation of a modified transcript in mammals, a nuclear event that occurs post-transcriptionally, yielding a truncated protein, Apob48 (Lau *et al*., 1991). During editing, Apobec replaces a cytidine (C) at the start of the 2153rd codon (CAA) with a uridine (U) by hydrolytic deamination to produce a stop codon (UAA) resulting in translational termination (Navaratnam *et al*., 1993; Teng *et al*., 1993). This process is very precise, targeting a single cytidine hidden amongst the approximately 14,000 nucleotides that comprise the *apob* transcript. Research has shown that RNA editing has stringent cis-acting requirements that include sequence-specific elements located both upstream (the 5' efficiency sequence, the A-rich region and the efficiency sequence) (Shah *et al*., 1991; Hersberger and Innerarity, 1998; Hersberger *et al*., 1999) and downstream (the mooring sequence and the 3' efficiency sequence) (Backus and Smith, 1992; Richardson *et al*., 1998) of the editing site.

Editing of the *apob* transcript is mediated by a multi-component enzyme complex consisting of an Apobec1 homodimer and an auxiliary factor called Apobec complementation factor (Acf) (see: Blanc and Davidson, 2010). The Acf recognises and attaches to the mooring sequence, positioning

the Apobec1 above the cytidine to be edited (Wedekind *et al*., 2003). The Apobec1 contains a cluster of residues that bind zinc which can activate a zinc-bound water molecule and initiate proton transfer (i.e. a free hydroxide ion). This in turn removes ammonia (deaminates) from the cytidine, converting it into uridine (Carter, 1995). Mutations in the mooring sequence can greatly reduce or completely abolish RNA editing (see: Adler and Hajduk, 1994; Chan *et al*., 1997; Davidson and Shelness, 2000 for in-depth reviews on the mechanisms and regulations of RNA editing). Consequently, the editing site and mooring region are highly conserved in mammals (Richardson *et al*., 1998).

It was originally thought that *apob* editing was not only tissue specific (i.e. editing was thought to only occur in the intestine) but also specific to placental mammals. However, recent investigations have quashed both of these theories; *apob* editing has been discovered in the liver of rats and mice (Srivastava *et al*., 1992) and more recently, the detection of Apobec1 editing activity in non-placental mammalian species (i.e. marsupials and monotremes (Fujino *et al*., 1999)) has seen the editing paradigm extended accordingly. Several other studies have examined the presence of Apobec1 and *apob* mRNA editing in lower vertebrates but, to date, there is no evidence indicating its presence prior to the divergence of the mammals (Blue *et al*., 1980; Tarugi *et al*., 1990; Teng and Davidson, 1992; Anant *et al*., 1998; Fujino *et al*., 1999; Dance *et al*., 2001; Conticello *et al*., 2005). The presence of two forms of Apob in mammals coupled with the seemingly normal functioning of non-mammalian vertebrates with only Apob100 begs an interesting question; what is the physiological relevance and benefit of having two forms of Apob? Several studies have genetically modified mice so that they exclusively synthesise either only Apob48 or only Apob100 (Farese *et al*., 1996; Hirano *et al*., 1996; Morrison *et al*., 1996; Nakamuta *et al*., 1996). These mice appear to develop and grow normally and are fertile. However, closer examination of serum lipid levels showed that the Apob100-only mice had elevated levels of triglycerides (hypertriglyceridemia), an

important issue especially for human health. This highlights a major functional difference between mammalian Apob100 and Apob48.

The importance of normal lipid homeostasis and its evident link with the two mammalian forms of Apob has led to a growing amount of research investigating the benefits of each ApoB and the mechanisms of *apob* editing in mammals (Farese *et al*., 1996; Hirano *et al*., 1996; Morrison *et al*., 1996; Nakamuta *et al*., 1996). Whilst a considerable amount of studies have been carried out on the chicken (Blue *et al*., 1980; Tarugi *et al*., 1990; Teng and Davidson, 1992), there has been minimal research into the presence of *apob* editing in any fish species and the consequences that a lack of editing may have on fish lipid homeostasis. This study focused on the molecular evolution of Apobec1 and Apob in order to ascertain whether *apob* mRNA editing occurs in anguillid eels, basal teleosts which represent an evolutionarily important animal group. Nucleotide, amino acid and partial genome alignments were complemented by phylogenic tree constructions and a synteny analysis. We conclude that having two forms of Apob is beneficial in terms of avoiding hypertriglyceridemia; as apparent from differences in lipid homeostasis between mammals and non-mammalian vertebrates.

2 METHODS

2.1. Phylogenetic trees and alignments

The amino acid sequences for Apob and Apobec1-4 from several mammalian and oviparous vertebrate species were obtained after a search of the National Center for Biotechnology Information (NCBI) database (http://www.ncbi.nlm.nih.gov/protein) (for accession numbers and species names see supplementary Table 1). Apobec5 is a very recent discovery (Severi et al., 2011) which has not yet been fully characterized or annotated in any species and was therefore omitted from this study. Searches were carried out in both the nucleotide and amino acid databases as well

as the whole genome databases. Failed Apobec1 searches were followed up by a BLAT (BLAST-like alignment tool) search of the Genome Bioinformatics database (http://genome.ucsc.edu/index.html) which yielded possible remnant sequences in the zebrafish genome. The sequences for *Anguilla australis* were taken from a liver transcriptome (unpublished data). Where insufficient sequence data were available from the transcriptome, the *A. anguilla* genome was searched (http://www.zfgenomics.org/sub/eel). Whilst the full-length sequences were used for the analysis of the Apobecs, only the exon 26 (approximately 2500 amino acids), which includes the editing site, was used for the analysis of Apob due to its large size. The Neighbor-Joining method (Saitou and Nei, 1987) was used to infer the evolutionary history of both the Apob sequences and the Apobec sequences. A bootstrap test (10,000 replicates) was carried out (Felsenstein, 1985). The evolutionary distances were computed using the p-distance method (Nei and Kumar, 2000). Phylogenetic tree construction was done using MEGA5 (Tamura *et al*., 2011). The amino acid sequences of Apob which span the Apobec1 editing site were aligned using CLC Sequence Viewer Version 6.3 to enable nucleotide differences in the efficiency sequence, editing site and mooring region to be observed.

2.2. Synteny analysis & genome alignment.

The synteny analysis of the human APOBEC1 genomic region in relation to chimpanzee (*Pan troglodytes*), mouse (*Mus musculus*), sheep (*Ovis aires*), dog (*Canis familiaris*), elephant (*Loxodonta Africana*), opossum (*Monodelphis domestica*), chinese turtle (*Pelodiscus sinensis*), zebra finch (*Taeniopygia guttata*), flycatcher (*Ficedula albicollis*), duck (*Anas platyrhynchos*), chicken (*Gallus gallus*), turkey (*Meleagris gallopavo*), lizard (*Anolis carolinensis*), frog (*Xenopus tropicalis*), spotted gar (*Lepisosteus oculatus*), stickleback (*Gasterosteus aculeatus*), tetradon (*Tetraodon nigroviridis*), fugu (*Takifugu rubripes*), tilapia (*Oreochromis niloticus*), medaka (*Oryzias latipes*), amazon molly (*Poecilia Formosa*), platyfish (*Xiphophorus maculatus*), atlantic cod (*Gadus morhua*) and zebrafish (*Danio rerio*) was completed using Genomicus v86.01 web site

(Louis et al 2014). The synteny map was constructed with AlignView using human APOBEC1 (ENSF00000111701) as the reference gene and the root species was expanded to Euteleostomi (~420 million years) to allow the inclusion of teleost species facilitate alignment, the synteny analysis was expanded to include two synteny blocks (FOXJ2 – SLC2A14 and PEX5 – C1R) to root all species alignments. Based on this synteny analysis, a physical alignment of genomic regions from human chromosome 12: 7,080,001 – 8,080,000, stickleback linkage group XX: 11,580,000 - 12,100,000 and zebrafish chromosome 16: 13,800,000 – 14,540,000 was manually created using GNU Image Manipulation Program (v2.8 https://www.gimp.org). Sequence information and annotation were from Ensemble genomes Human (GRCh38), stickleback (BROADS1) and zebrafish (Zv9).

3. RESULTS

3.1. Apobec Phylogeny

Searches of both the *A. australis* transcriptome and *A. anguilla* genome resulted in sequence information for only one of the four Apobecs, which showed greatest sequence identity with Apobec2 (Figure 1). A protein search of the NCBI database resulted in full sequence information for the Apobec genes from a variety of mammalian (placentals, monotremes and marsupials) and oviparous (birds, lizards, amphibians and fish) vertebrates. For most mammalian species, all four Apobec sequences were retrieved (except Apobec3 in marsupials) (Figure 1). Apobec1 was near-universally absent in non-mammalian vertebrates except for Apobec1 in lizard and Apobec1-like in reptiles and some birds (Figure 1). Each member of the family grouped more closely with its orthologs found in other species than to the other enzymes found in the same species (except turtle Apobec3-like which clustered with the Apobec1s), i.e. all Apobec1 sequences across the species formed a cluster rather than all human Apobecs forming a cluster (Figure 1).

3.2. Apob Phylogeny

The *apob* gene is highly conserved amongst the different vertebrate clades and their sequences therefore cluster together within the tree (Figure 2). Indeed, alignments of the amino acid sequences for exon 26 showed similarities greater than 75 % within the clusters of the tree (i.e., primates (96 % homology between humans and gibbons); carnivores (85 % homology between cats and dogs); ungulates (82 % homology between bull and elephant); rodents (88 % homology between hamsters and rats); birds (85 % homology between chickens and flycatchers)). Even amongst the placental mammals, the two species separated by the greatest distance on the tree (human and rat) had 79 % homology. Although still high, once outside of the placental mammals the homology decreased quickly; there is 68 % identity between platypus and humans and 60 % homology between humans and turkeys. However, there is much lower homology between the fishes and the primates with 34 % homology shared between zebrafish and humans across the approximately 2500 amino acids encoded by exon 26 of Apob.

The single cytidine residue to which Apobec1 mRNA editing is targeted, also known as the editing site, is highly conserved amongst *apob* in all mammals, placental or otherwise (Figure 3). This cytidine forms the start of a glutamine codon. This glutamine codon is also present, in-frame, in the chicken, turkey, duck, turtle, alligator, cichlid (*Haplochromis burtoni*) and tilapia; however, there are between two and nine nucleotides that differ in the *apob* mooring region of these species compared to the highly homologous mooring region of mammals and between seven and thirteen differences in the efficiency sequences (Figure 3). The CAA sequence which encodes the glutamine appears to be out of frame in the lizard with an asparagine codon present at the editing site instead (Figure 3). Although the amino acid sequences from the lower vertebrates align well with the mammalian amino acid sequence both up- and down-stream of the editing region (providing credibility for the alignment), there is very little homology (ranging from approximately 8 – 66 % of amino acids) inside the area of focus. However, as editing occurs at the nucleotide level, the amino acid alignment is of little relevance.

3.3. Synteny analysis and genomic alignment.

Among 17 protein-encoding mammalian genes, present in a block consisting of apobec1 and eight genes either side of apobec1, two synteny groups (FOXJ2 – SLC2A14 and PEX5 – C1R) were conserved within teleosts (Figure 4). Significantly, a nine-gene group that is linked and centers on apobec1 (nanog – acsm4) is not present in fish. Despite the presence of apobec1 in a number of non-mammalian species (turtle, zebra finch, flycatcher, duck and lizard), it appears that the synteny within which apobec1 is found is a mammal-specific grouping. To add further credence to the lack of apobec1 within the teleost lineage at this location, a specific genomic comparison between human and data-rich teleost genomes of stickleback and zebrafish shows that the common syntenic groups are separated by a 0.59Mb, 0.42Mb and 0.38Mb genomic region for each species respectively (Figure 5). Closer examination of the two teleost regions reveals a teleost-specific synteny grouping including seven common genes that occupies this space.

4. DISCUSSION

The evolutionary history of Apob and Apobec1-4 was inferred using sequences retrieved from the NCBI and Genome Bioinformatics databases as well as the *A. australis* transcriptome and *A. anguilla* genome. These inferences revealed a lack of *apob* mRNA editing in lower vertebrates.

Surprisingly, recent research has revealed a homologue of the mammalian Apobec1 in reptiles (lizard: Severi *et al*., 2011; NB: the turtle, alligator, duck, zebrafinch and flycatcher sequences are predicted by automated computational analysis). Despite the presence of this homologue, there has been no evidence for *apob* mRNA editing outside of mammals suggesting that this gene arose by duplication independently well after the divergence of mammals and reptiles. Indeed, closer examination of the sequence revealed that the Apobec1 homologue was missing the C-terminal region that is required for *apob* mRNA editing (Severi *et al*., 2011). This study searched for eel

Apobec1 within transcriptome information gained from sequencing of the liver, as a gut transcriptome is currently unavailable; this may have lead to Apobec1 being overlooked due to the wrong tissue being examined. However, the genomes of all species of fish used in this study were also searched to no avail. Further support for the lack of Apobec1 comes from the results of synteny analysis. A whole block of nine mammalian genes which center on Apobec1 (nanog – acsm4) is missing in this linkage group of non-mammalian vertebrates. Interestingly, Acf was found in the transcriptome/genome of all species indicating that some of the editing complex is present; however, as with mammals, it appears Apobec1 is absolutely required for *apob* mRNA editing.

Most studies agree that the single-domain cytidine deaminase, Apobec2, is the ancestral form from which all other Apobecs arose (Conticello *et al*., 2005; Rogozin *et al*., 2007; Conticello, 2008). Accordingly, Apobec2 was found in all species included in this study, and although not shown in this tree, in the case of most of the bony fishes (eels included), two distinct paralogues were present. These two paralogues more than likely reflect the genome duplication which is thought to have occurred around the origin of the ray-finned fishes (Taylor *et al*. 2003; Christoffels *et al*., 2004).

No non-mammalian species were found to have a true Apobec3 (only an Apobec3-like gene). In fact, Apobec3 was not found in the possum, Tasmanian devil or platypus either, supporting the finding by LaRue and colleagues (2008) that Apobec3 arose from Apobec2 after the divergence of placental and non-placental mammals. Identified as the most recently diverged family, the Apobec3s appear to be still expanding at a rapid rate (Rogozin *et al*., 2005). The genomes of primates encode up to seven Apobec3 proteins on a single chromosome (Jarmuz *et al*., 2002; Zhang and Webb, 2004), compared to other placental mammals which average between two and four copies. This indicates a recent expansion in primates which is thought to be driven by selective pressure from retroviruses as primate Apobecs have been implicated in immune function (Conticello *et al*., 2005; Koito and Ikeda, 2012).

The discovery of Apobec4 occurred only recently (Rogozin *et al*., 2005) and little is known about its function (other than its transcript in the testis indicating a potential role in spermatogenesis). Despite its recent identification, the identification of Apobec4 in reptiles and birds, i.e. non-mammalian species, indicates that Apobec4 diverged earlier than Apobec1 and 3. However, searches of the (nearly) complete genomes of several fish species failed to identify Apobec4, indicating that Apobec4 may have arisen prior to the divergence of the amniotes, possibly as a tetrapod-specific duplication.

The lack of Apobec1 coupled with the low homology between mammalian Apob and non-mammalian Apob within the sequence required for *apob* mRNA editing, supports the notion that *apob* mRNA editing does not occur in lower vertebrates (Blue *et al*., 1980; Tarugi *et al*., 1990; Teng and Davidson, 1992; Anant *et al*., 1998; Fujino *et al*., 1999; Dance *et al*., 2001; Conticello *et al*., 2005). The chicken, turkey and duck provide interesting models; although the Apobec1 has not been identified in any of these species, all of their Apob sequences contain the editing site. However, all three species have several nucleotides different from the mammalian mooring sequence and these have been shown to abolish mammalian Apobec1 editing abilities *in vitro* (Teng and Davidson, 1992). Another interesting case arises in the lizard. As mentioned above, an Apobec1 orthologue has been recently discovered (Severi *et al*., 2011) but the *apob* mRNA editing site appears out-of-frame in the lizard *apob* gene. This, coupled with the lack of the C-terminus of the Apobec1 orthologue removes its ability to edit (Severi *et al*., 2011). Further, nucleotide differences in the mooring sequence of the lizard *apob* indicates that editing would likely not occur even in the presence of a functional Apobec1 because previous *in vitro* studies have shown that nearly all discrepancies in this eleven nucleotide sequence abolish *apob* mRNA editing (Shah *et al*., 1991; Teng and Davidson, 1992; Driscoll *et al*., 1993; Backus *et al*., 1994). The cichlid and tilapia also have the *apob* mRNA editing site (in-frame) present in their *apob* sequences. However, both species

have nine nucleotides different in the mooring region, so like the chicken, turkey and the lizard, it is very unlikely that it could be edited in the presence of a functional Apobec1.

The benefit of having two forms of Apob has been a focus of mammalian lipid research over the last two decades (for in depth reviews see: Kim and Young, 1998; Veniant *et al*., 1999; Marschang and Herz, 2003). However, the results from the present study may provide an important link between mammals and non-mammalian vertebrates that has been previously unreported. Targeted mutagenesis of the *apob* gene has produced mice that exclusively synthesise either only *apob48* or only *apob100* (Farese *et al*., 1996). As mentioned above, both of these groups of mice appeared to develop and grow normally and were fertile. However, upon examination of their serum it was found that Apob48-only mice had significantly lower triglyceride levels than wild-type mice, while Apob100-only mice had significantly higher levels (Farese *et al*., 1996).

It has been suggested that *apob* mRNA editing activity arose in order to ensure dietary triglycerides were sent specifically to the liver for further processing and were cleared quickly from circulation (Hodges and Scott, 1992). Indeed, Apob48 lacks the low density lipoprotein receptor binding site of Apob100 (Mahley, 1988; Segrest *et al*., 2001; Hinsdale *et al*., 2002) and therefore relies on apolipoprotein E (Apoe) and the highly expressed hepatic Apoe receptors for uptake (Mahley *et al*., 1981; Linton *et al*., 1993). During mammalian intestinal chylomicron formation there is a convenient trade-off by which the reduced size of the Apob48 allows more Apoe to be associated with the lipoprotein. It is possible that non-mammalian vertebrates may not produce chylomicrons. Chickens produce molecules called portomicrons which have similar diameters and lipid compositions to mammalian chylomicrons but have Apob100 associated with them (Fraser et al., 1986). In reptiles and fish, the nomenclature surrounding the lowest density lipoproteins is confusing and inconsistent with the terms chylomicron, portomicron and very-low density lipoprotein used interchangeably. Despite the erratic naming of these molecules in non-mammalian

vertebrates there has been no evidence to suggest that any Apob other than Apob100 is associated with these molecules. As mentioned above, Apob100 can be associated with chylomicrons (Farese *et al*., 1996; Hirano *et al*., 1996) but this results in elevated plasma triglyceride levels (Farese *et al*., 1996). This increase in triglycerides is likely because the large size of the Apob100 displaces the amount of Apoe that can be incorporated into the lipoprotein, reducing the number of ligands available for binding and consequently increasing the time each lipoprotein is in circulation before being cleared. Also, the presence of Apob100 may reduce the accessibility of the triglycerides to lipase and thereby reduce lipolytic processing. These potential increases in Apob100 associated lipoprotein clearance times highlights an important physiological difference between mammalian Apob100 and Apob48.

This apparent excess of triglycerides in the serum of Apob100-only mice may help to explain the evident hypertriglyceridemia seen in non-mammalian vertebrates (i.e. chickens: Kudzma *et al*., 1975; Cho *et al*., 1984; Smith *et al*., 1985, and lizards: Knotkova *et al*., 2005; Lawton, 2005) in which only Apob100 has been confirmed. High levels of triglycerides have also been reported in anguillids (Larsson and Fange, 1969; Dave *et al*., 1975; Damsteegt *et al*., 2015) and evidence from this study suggests only Apob100 is present in eels. It is possible that in lower vertebrates the larger Apob100 is also affecting triglyceride transport from the gut resulting in hypertriglyceridemia. Further evidence that only Apob100 is present in eels comes from histological examination of the intestine. Feeding eels have a large amount of cytosolic fat in the enterocytes of their intestine (compared to non-feeding eels; see Figure 1 supplementary material), a pathological finding similar to that of mice with no Apob48 (Young *et al*., 1995) and humans with low levels of Apob (hypobetalipoproteinemia) or those that have normal levels of Apob but because of a mutation in microsomal triglyceride transfer protein can not synthesise Apob-containing lipoproteins (abetalipoproteinemia) (Gregg and Wetterau, 1994).

It appears that three components are necessary for successful editing of Apob – the presence of the editing site, a conserved mooring region and a functional Apobec1. However, using the current genomic data available we are still not able to ascertain which came first – Apobec1 or an editable Apob transcript. Presumably Apobec1 arose following a duplication event in early tetrapod evolution, preceding the mammalian-reptilian-avian split. It seems likely that the mooring region and Apobec1 have co-evolved and that Apobec1 neo-functionalised over time, but what were the implications of this? Did the ability to edit Apob give early mammals a fitness advantage which then lead to the fine-tuning of editing over time? As mentioned above, mammals that cannot edit Apob are disadvantaged, presenting with hypertriglyceridemia and the associated risks of this condition with heart disease. However, these detrimental effects do not appear to be problematic in non-mammalian vertebrates.

This study has added to the almost overwhelming support that *apob* mRNA editing does not occur outside of mammals. It was concluded that only one of the Apobecs (Apobec2) is present in the shortfinned eel. The lack of Apobec1, the low homology in the mooring region, the dispersal of peptides throughout the Apob protein (after amino acid sequencing), the evident hypertriglyceridemia and the build-up of fat in the enterocytes of the intestine provide highly compelling evidence for the presence of only one form of Apob is in eels (not including those Apob forms resulting from post-translational cleavage) and that it likely corresponds to mammalian Apob100. This is also one of the first studies to provide the link between hypertriglyceridemia in lower vertebrates and the presence of only mammalian Apob100 homologues, a hugely important finding that may have future benefits for human health.

5. ACKNOWLEDGEMENTS

The authors would like to thank Professor Graham Wallis (Department of Zoology, University of Otago) for helping with phylogenetic analyses. We would also like to gratefully acknowledge the constructive feedback to a prior version of this manuscript by the late Dr Elspeth Gold (Department of Anatomy, University of Otago), Dr Andrew Munkacsi (School of Biological Sciences, Victoria University of Wellington) and Professor Patrick Babin (Rare Diseases: Genetics and Metabolism, University of Bordeaux). This work was made possible through a University of Otago Doctoral Scholarship (ELD) and was funded in part by a Research Enhancement Grant (PBRF2012-63 and PBRF 2017-63) to PML.

6. REFERENCES

Adler, B.K. and Hajduk, S.L. (1994) Mechanisms and origins of RNA editing. *Current Opinion in Genetics and Development*, 4: 316- 22.

Anant, S., Yu, H. and Davidson, N.O. (1998) Evolutionary origins of the mammalian apolipoproteinB RNA editing enzyme, apobec-1: structural homology inferred from analysis of a cloned chicken small intestinal cytidine deaminase. *Biological Chemistry*, 379: 1075-1081.

Backus, J.W. and Smith, H.C. (1992) Three distinct RNA sequence elements are required for efficient apolipoprotein B (apoB) RNA editing *in vitro*. *Nucleic Acids Research*, 20: 6007-6014.

Blanc, V. and Davidson, N.O. (2010) APOBEC-1-mediated RNA editing. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 2: 594-602.

Blue, M.L., Protter, A.A., and Williams, D.L. (1980) Biosynthesis of apolipoprotein B in rooster kidney, intestine, and liver. *Journal of Biological Chemistry*, 255: 10048-10051.

Carter, C.W. (1995) The nucleoside deaminases for cytidine and adenosine: structure, transition state stabilization, mechanism, and evolution. *Biochimie*, 77: 92-98.

Chan, L., Chang, B.H., Nakamuta, M., Li, W.H. and Smith, L.C. (1997) Apobec-1 and apolipoprotein B mRNA editing. *Biochimica et Biophysica Acta*, 1345: 11-26.

Cho, B.H.S., Lawson, L.D., Toda, T. and Kummerow, F.A. (1984) Oxidation of fatty acid by heart mitochondria of chickens with endogenous hyperlipidemia. *Biochemical Medicine*, 31: 347-351.

Christoffels, A., Koh, E.G., Chia, J.M., Brenner, S., Aparicio, S., and Venkatesh, B. (2004) Fugu genome analysis provides evidence for a whole-genome duplication early during the evolution of ray-finned fishes. *Molecular Biology and Evolution*, 21: 1146-1151.

Conticello, S.G. (2008) The AID/APOBEC family of nucleic acid mutators. *Genome Biology*, 9: 229-239.

Conticello, S.G., Thomas, C.J.F., Petersen-Mahrt, S.K. and Neuberger, M.S. (2005) Evolution of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases. *Molecular Biology and Evolution*, 22: 367-377.

Cullen, B.R. (2006) Role and mechanism of action of the apoBEC3 family of antiretroviral resistance factors. *Journal of Virology*, 80: 1067-1076.

Damsteegt, E., Mizuta, H., Ozaki, Y., Hiramatsu, N., Todo, T., Hara, A., Ijiri, S., Adachi, S. and Lokman, P.M. (2014) Development and partial characterisation of an antiserum against apolipoprotein B of the short-finned eel, *Anguilla australis*. *Journal of Comparative Physiology B*, 184: 589-599.

Damsteegt, E., Falahatimarvast, A., McCormick, S.P.A. and Lokman, P.M. (2015) Triacylglyceride physiology in the short-finned eel, *Anguilla australis* - changes throughout early oogenesis. *American Journal of Physiology – Regulatory, Integrative and Comparative Physiology*.

Dance, G.S., Beemiller, P., Yang, Y., Mater, D.V., Mian, I.S. and Smith, H.C. (2001) Identification of the yeast cytidine deaminase CDD1 as an orphan C / U RNA editase. *Nucleic Acids Research*, 29: 1772-1780.

Dave, G., Johansson-Sjobeck, M-L., Larsson, A., Lewander, K. and Lidman, U. (1975) Metabolic and hematological effects of starvation in the European eel, *Anguilla anguilla* L. - I. Carbohydrate, lipid, protein and inorganic ion metabolism. *Comparative Biochemistry and Physiology*, 52: 423-430.

Davidson, N.O. and Shelness, G.S. (2000) Apolipoprotein B: mRNA Editing, Lipoprotein Assembly, and Presecretory Degradation. *Annual Reviews in Nutrition*, 20: 169-193.

Driscoll, D.M., Lakhe-Reddy, S., Oleska, L.M. and Martinez, D. (1993) Induction of RNA editing at heterologous sites by sequences in apolipoprotein B mRNA. *Molecular and Cellular Biology*, 13: 7288-7294.

Farese, R.V., Veniant, M.M., Cham, C.M., Flynn, L.M., Pierotti, V., Loring, J.F., Traber, M., Ruland, S., Stokowski, R.S., Huszar, D. and Young, S.G. (1996) Phenotypic analysis of mice expressing exclusively apolipoprotein B48 or apolipoprotein B100. *Proceedings of the National Academy of Sciences*, 93: 6393-6398.

Felsenstein, J. (1985) Confidence limits on phylogenies: An approach using the bootstrap. *Evolution*, 39: 783-791.

Fraser, R., Heslop, V.R., Murray, F.E.M. and Day, W.A. (1986). Ultrastructuralstudies of the portal transport of fat in chickens. *British Journal of Experimental Pathology*, 67: 783-791.

Fujino, T., Navaratnam, N., Jarmuz, A., Von Haeseler, A. and Scott, J. (1999) C→U editing of apolipoprotein B mRNA in marsupials: identification and characterization of APOBEC-1 from the American opossum *Monodelphus domestica*. *Nucleic Acids Research*, 27: 2662-2671.

Gregg, R.E. and Wetterau, J.R. (1994) The molecular basis of abetalipoproteinemia. *Current Opinion in Lipidology*, 5: 81-86.

Harris, R.S. and Liddament, M.T. (2004) Retroviral restriction by APOBEC proteins. *Nature Reviews Immunology*, 4: 868-877.

Hersberger, M. and Innerarity, T.L. (1998) Two efficiency elements flanking the editing site of cytidine 6666 in the apolipoprotein B mRNA support mooring-dependent editing. *Journal of Biological Chemistry*, 273: 9435-9442.

Hergberger, M., Patarroyo-White, S., Arnold, K.S. and Innerarity, T.L. (1999) Phylogenetic analysis of the apolipoprotein B mRNA-editing region. Evidence for a secondary structure between the mooring sequence and the 3' efficiency element. *The Journal of Biological Chemistry*, 274: 34590-34597.

Hinsdale, M.E., Sullivan, P.M., Mezdour, H. and Maede, N. (2002) ApoB-48 and apoB-100 differentially influence the expression of type-III hyperlipoproteinemia in APOE*2 mice. *The Journal of Lipid Research*, 43: 1520-1528.

Hirano, K-I., Young, S.G., Farese, R.V., Ng, J., Sande, E., Warburton, C., Powell-Braxton, L.M. and Davidson, N.O. (1996) Targeted disruption of the mouse apobec-1 gene abolishes

apolipoprotein B mRNA editing and eliminates apolipoprotein B48. *The Journal of Biological Chemistry*, 271: 9887-9890.

Hodges, P. and Scott, J. (1992) Apolipoprotein B mRNA editing: a new tier for the control of gene expression. *Trends in Biochemical Sciences*, 17: 77-81.

Jarmuz, A., Chester, A., Bayliss, J., Gisbourne, J., Dunham, I., Scott, J. and Navaratnam, N. (2002) An anthropoid-specific locus of orphan C to U RNA-editing enzymes on chromosome 22. *Genomics*, 79: 285-296.

Kim, E. and Young, S.G. (1998) Genetically modified mice for the study of apolipoprotein B. *The Journal of Lipid Research*, 39: 703-723.

Knotkova, Z., Pejrilova, S., Trnkova, S., Matouskova, O. and Knotek, Z. (2005) Influence of reproductive season upon plasma biochemistry values in green iguanas. *Acta Veterinaria Brno*, 74: 515-520.

Koito, A. and Ikeda, T. (2012) Apolipoprotein B mRNA-editing, catalytic polypeptide cytidine deaminases and retroviral restriction. *Wiley Interdisciplinary Reviews, RNA*, 3: 529-541.

Kudzma, D.J., St Claire, F., DeLallo, L. and Friedberg, S.J. (1975) Mechanism of avian estrogen-induced hypertriglyceridemia: evidence for overproduction of triglyceride. *The Journal of Lipid Research*, 16: 123-133.

Larsson, A. and Fange, R (1969) Chemical differences in the blood of yellow and silver phases of the European eel (*Anguilla anguilla* L.). *Archives internationals de Physiologie et de Biochemie*, 77: 701-709.

LaRue, R.S., Jonsson, S.R., Silverstein, K.A.T., Lajoie, M., Bertrand, D., El-Mabrouk, N., Hotzel, I., Andresdottir, V., Smith, T.P.L. and Harris, R.S. (2008) The artiodactyls APOBEC3 innate immune repertoire shows evidence for a multifunctional domain organization that existed in the ancestor of placental mammals. *BMC Molecular Biology*, 9: 104-124.

Lau, P.P., Xiong, W.J., Zhu, H.J., Chen, S.H. and Chan, L. (1991) Apolipoprotein B mRNA editing is an intranuclear event that occurs posttranscriptionally coincident with splicing and polyadenylation. *The Journal of Biological Chemistry*, 266: 20550-20554.

Lawton, M.P.C. (2005) 'Herp biochemistry results – what do they mean and why do I need them'. *Exotics – Reptiles*, The North American Veterinary Conference, 1285-1286.

Linton, M.F., Farese, R.V., Chiesa, G., Grass, D.S., Chin, P., Hammer, R.E., Hobbs, H.H. and Young, S.G. (1993) Transgenic mice expressing high plasma concentrations of human apolipoprotein B100 and lipoprotein[a]. *Journal of Clinical Investigation*, 92: 3029-3037.

Lo, C-M., Nordskog, B.K., Naulia, A.M., Zheng, S., von Lehmden, S.B., Yang, Q., Lee, D., Swift, L.L., Davidson, N.O. and Tso, P. (2008) Why does the gut choose apolipoprotein B48 but not B100 for chylomicron formation? *American Journal of Physiology – Gastrointestinal and Liver Physiology*, 294: G344-G352.

Louis, A., Nguyen, N.T.T., Muffato, M. and Crollius, H.R., (2015). Genomicus update 2015: KaryoView and MatrixView provide a genome-wide perspective to multispecies comparative genomics. *Nucleic acids research*, 43(D1), pp.D682-D689.

Mahley, R.W. (1988) Apolipoprotein E: cholesterol transport protein with expanding role in cell biology. *Science*, 240: 622-630.

Mahley, R.W., Hui, D.Y., Innerarity, T.L. and Weisgraber, K.H. (1981) Two independent lipoprotein receptors on hepatic membranes of the dog, swine, and man. Apo-B,E and apo-E receptors. *Journal of Clinical Investigation*, 68: 1197-1206

Marschang, P. and Herz, J. (2003) Mouse models as tools for dissecting disorders of lipoprotein metabolism. *Seminars in Cell and Developmental Biology*, 14: 25-35.

Morrison, J.R., Paszty, C., Stevens, M.E., Hughes, S.D., Forte, T., Scott, J. and Rubin, E.M. (1996) Apolipoprotein B RNA editing enzyme-deficient mice are viable despite alterations in lipoprotein metabolism. *Proceedings of the National Academy of Sciences*, 93: 7154-7159.

Nakamuta, M., Tsai, A., Chan, L., Davidson, N.O. and Teng, B. (1999) Sequence elements required for apolipoprotein B mRNA editing enhancement activity from chicken enterocytes. *Biochemical and Biophysical Research Communications*, 254: 744-750.

Navaratnam, N., Morrison, J.R., Bhattycharya, S., Patel, D., Funahashi, T., Giannoni, F., Teng, B.-B., Davidson, N.O., and Scott, J. (1993). The p27 catalytic subunit of the apolipoprotein B mRNA editing enzyme is a cytidine deaminase. *Journal of Biological Chemistry*, 268: 20709-20712.

Nei, M. and Kumar, S. (2000) *Molecular Evolution and Phylogenetics*. Oxford University Press, New York.

Pham, P., Bransteitter, R. and Goodman, M.F. (2005) Reward versus risk: DNA cytidine deaminases triggering immunity and disease. *Biochemistry*, 44: 2703-2715.

Powell, C., Elsaeidi, F. and Goldman, D. (2012) Injury-dependent muller glia and ganglion cell reprogramming during tissue regeneration requires apobec2a and apobec2b. *The Journal of Neuroscience*, 32: 1096-1109.

Richardson, N., Navaratnam, N. and Scott, J. (1998) Secondary structure for the apolipoprotein B mRNA editing site: AU-binding proteins interact with a stem loop. *Journal of Biological Chemistry*, 273: 31707-31717.

Rogozin, I.B., Basu, M.K., Jordan, I.K., Pavlov, Y.I. and Koonin, E.V. (2005) APOBEC4, a new member of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases predicted by computational analysis. *Cell Cycle*, 4: 1281-1285.

Rogozin, I.B., Iyer, L.M., Liang, L., Glazko, G.V., Liston, V.G., Pavlov, Y.I., Aravind, L. and Pancer, Z. (2007) Evolution and diversification of lamprey antigen receptors: evidence for involvement of an AID/APOBEC family cytosine deaminase. *Nature Immunology*, 8: 647-656.

Saitou, N. and Nei, M. (1987) The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 4: 406-425.

Segrest, J.P., Jones, M.K., De Loof, H. and Dashti, N. (2001) Structure of apolipoprotein B-100 in low density lipoproteins. *The Journal of Lipid Research*, 42: 1346-1367.

Severi, F., Chicca, A. and Conticello, S.G. (2011) Analysis of reptilian APOBEC1 suggests that RNA editing may not be its ancestral function. *Molecular Biology and Evolution*, 28: 1125-1129.

Shah, R.R., Knott, T.J., Legros, J.E., Navaratnam, N., Greeve, J.C., and Scott, J. (1991) Sequence requirements for the editing of apolipoprotein B mRNA. *Journal of Biological Chemistry*, 266: 16301-16304.

Smith, T., Toda, T. and Kummerow, F.A. (1985) Plasma lipid peroxidation in hyperlipidemic chickens. *Atherosclerosis*, 51: 119-122.

Srivastava, R.A., Tang, J., Baumann, D. and Schonfeld, G. (1992) Hormonal and nutritional stimuli modulate apolipoprotein B mRNA editing in mouse liver. *Biochemical and Biophysical Research Communications*, 188: 135-41.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, 28: 2731-2739.

Tarugi, P., Albertazzi, L., Nicolini, S. and Calandra, S. (1990) Absence of apolipoprotein B-48 in the chick, *Gallus domesticus*. *The Journal of Lipid Research*, 31: 417-427.

Taylor, J. S., Braasch, I. Frickey, T., Meyer, A. and Van de Peer, Y. (2003) Genome duplication, a trait shared by 22000 species of ray-finned fish. *Genome Research*, 13: 382-390.

Teng, B. and Davidson, N.O. (1992) Evolution of intestinal apolipoprotein B mRNA editing. Chicken apolipoprotein B mRNA is not edited, but chicken enterocytes contain *in vitro* editing enhancement factor(s). *Journal of Biological Chemistry*, 267: 21265-21272.

Teng, B., Burant, C.F. and Davidson, N.O. (1993) Molecular cloning of an apolipoprotein B messenger RNA editing protein. *Science*, 260: 1816-1819.

Veniant, M.M., Kim, E., McCormick, S., Boren, J., Nielsen, L.B., Raabe, M. and Young, S.G. (1999) Insights into apolipoprotein B biology from transgenic and gene-targeted mice. *The Journal of Nutrition*, 129: 451S-455S.

Wedekind, J.E., Dance, G.S., Sowden, M.P. and Smith, H.C. (2003) Messenger RNA editing in mammals: new members of the APOBEC family seeking roles in the family business. *Trends in Genetics*, 19: 207-216.

Young, S.G., Cham, C.M., Pitas, R.E., Burri, B.J., Connolly, A., Flynn, L., Pappu, A.S., Wong, J.S., Hamilton, R.L. and Farese, R.V. (1995) A genetic model for absent chylomicron formation: mice producing apolipoprotein B in the liver, but not in the intestine. *Journal of Clinical Investigation*, 96: 2932-2946.

Zhang, J. and Webb, D.M. (2004) Rapid evolution of primate antiviral enzyme APOBEC3G. *Human Molecular Genetics*, 13: 1785-1791.

FIGURES

Figure 1: A neighbor-joining phylogenetic tree constructed using the full-length deduced amino acid sequences from four members of the Apobec family. The number succeeding the animal name indicates which of the four members of the Apobec has been previously assigned to the sequence within the NCBI database. An * indicates that the sequence used is an Apobec-like gene. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed. Each branch forms a distinct cluster and has been circled accordingly. Scientific names and accession numbers can be found in Table 1 (Supplementary material).

Figure 2: A neighbor-joining phylogenetic tree constructed using the amino acid sequences (approximately 2500 amino acids) encoded by exon 26 of Apob from sixteen mammalian species and twenty one non-mammalian vertebrates. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (10000 replicates) is shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the p-distance method and are in the units of the number of amino acid differences per site. The dotted line separates mammals (above) from oviparous vertebrates. Scientific names and accession numbers can be found in Table 1 (Supplementary material).

Figure 3: The *apob* gene sequences of sixteen mammals aligned with that of twenty one non-mammalian vertebrates. The codon involved in Apobec1 mRNA editing is shown, as is the 5' efficiency sequence, A-rich region, efficiency sequence, spacer element and mooring region. Species names and accession numbers can be found in Table 1 (Supplementary material).

Figure 4: Genomic synteny map comparing orthologues of the APOBEC1 locus and the genes flanking it in humans (*Homo sapiens*), chimpanzee (*Pan troglodytes*), mouse (*Mus musculus*), sheep (*Ovis aires*), dog (*Canis familiaris*), elephant (*Loxodonta Africana*), opossum (*Monodelphis domestica*), Chinese turtle (*Pelodiscus sinensis*), zebra finch (*Taeniopygia guttata*), flycatcher (*Ficedula albicollis*), duck (*Anas platyrhynchos*), chicken (*Gallus gallus*), turkey (*Meleagris gallopavo*), lizard (*Anolis carolinensis*), frog (*Xenopus tropicalis*), spotted gar (*Lepisosteus oculatus*), stickleback (*Gasterosteus aculeatus*), tetradon (*Tetraodon nigroviridis*), fugu (*Takifugu rubripes*), tilapia (*Oreochromis niloticus*), medaka (*Oryzias latipes*), Amazon molly (*Poecilia Formosa*), platyfish (*Xiphophorus maculatus*), Atlantic cod (*Gadus morhua*) and zebrafish (*Danio rerio*). This map was obtained by using the genome browser Genomicus (v86.01) AlignView with the reference gene being Human APOBEC1 (ENSF00000111701), root species Euteleostomi (~420 million years). Note that the direction of arrows indicates the gene orientation only in the reference species. Orthologs of each gene in other species are shown in a single column in the same colour. Genes are named after their human orthologs according to HUGO Gene Nomenclature Committee (HGNC).The chromosome number or genome fragment contig ID are indicated.

Figure 5: Alignment of genomic regions from human (*Homo sapiens*) chromosome 12: 7,080,001 – 8,080,000, stickleback (*Gasterosteus aculeatus*) linkage group XX: 11,580,000 - 12,100,000 and zebrafish (*Danio rerio*) chromosome 16: 13,800,000 – 14,540,000. Shaded boxes indicate common orthologues between all three species. Orthologous gene locations between zebrafish and stickleback are indicated by connected lines. Genes are named after their human orthologs according to HUGO Gene Nomenclature Committee (HGNC).

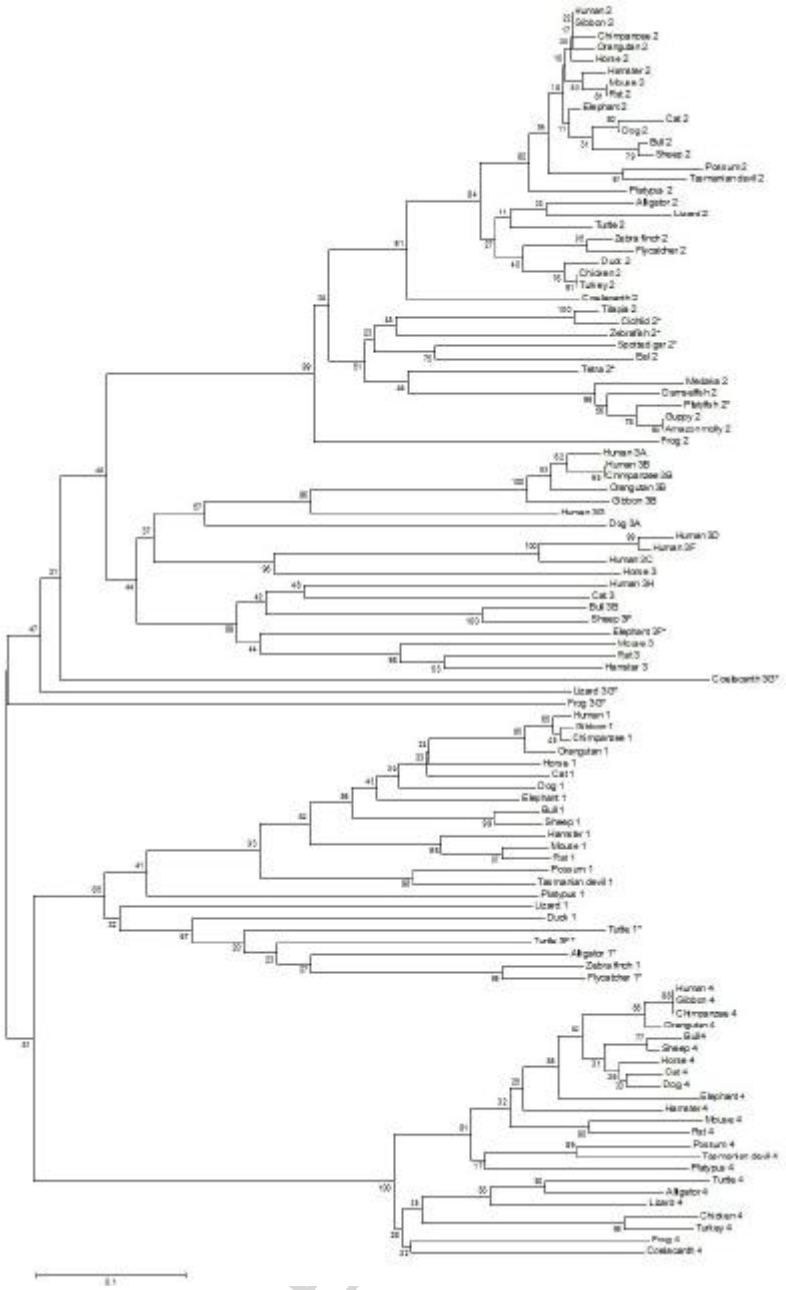


Figure 1

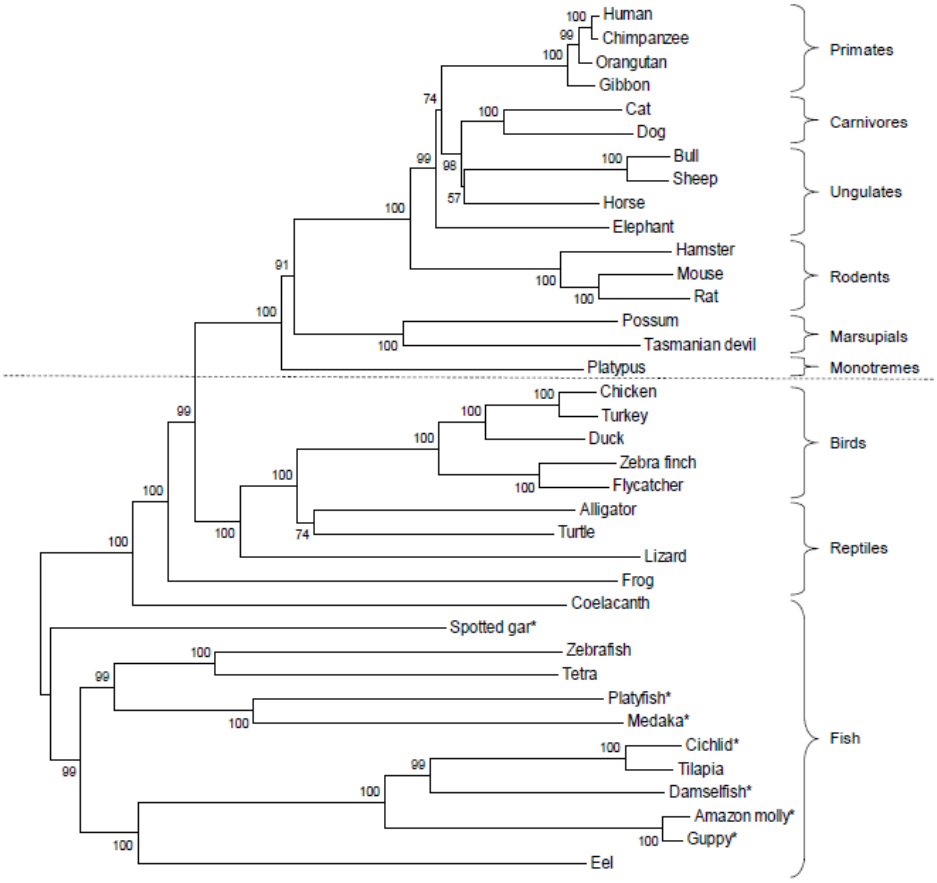


Figure 2

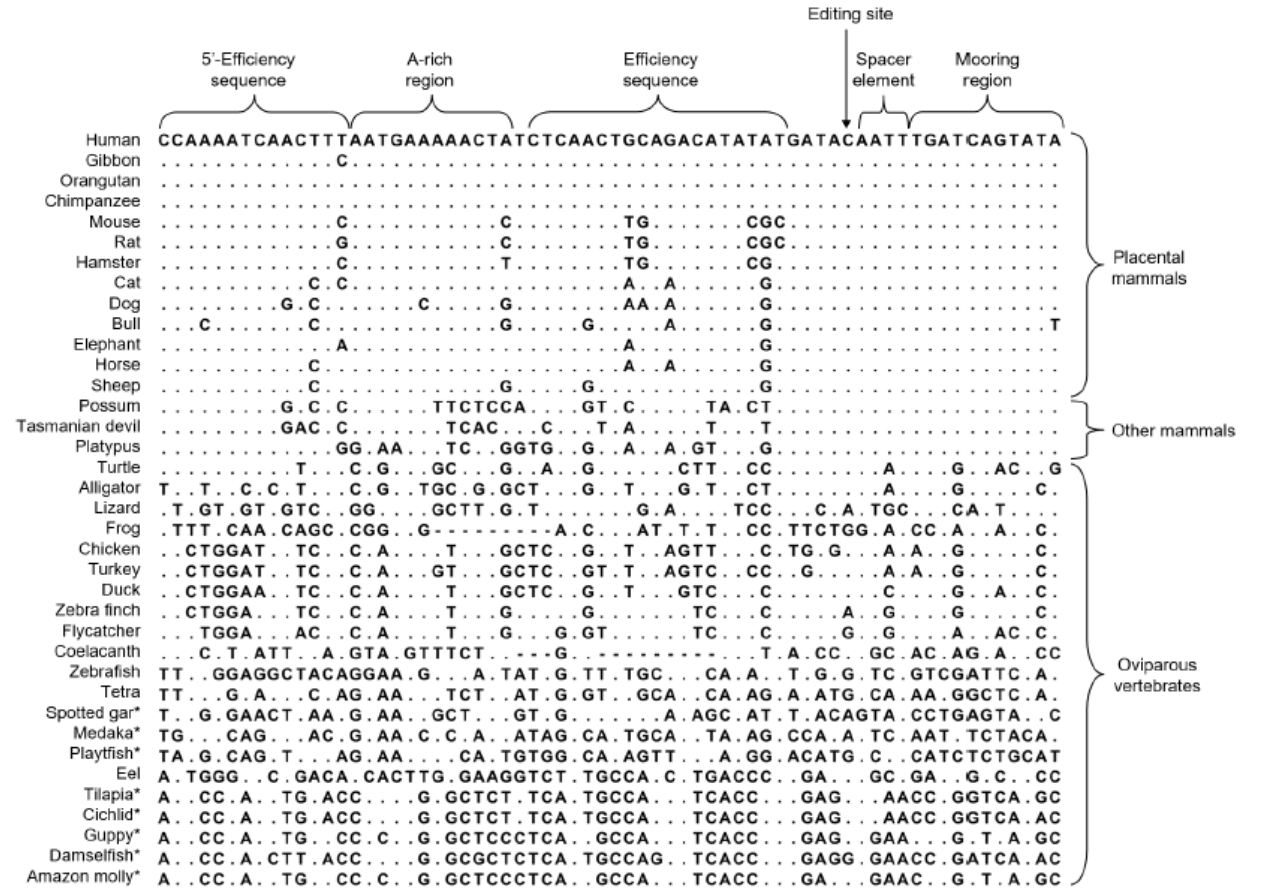


Figure 3

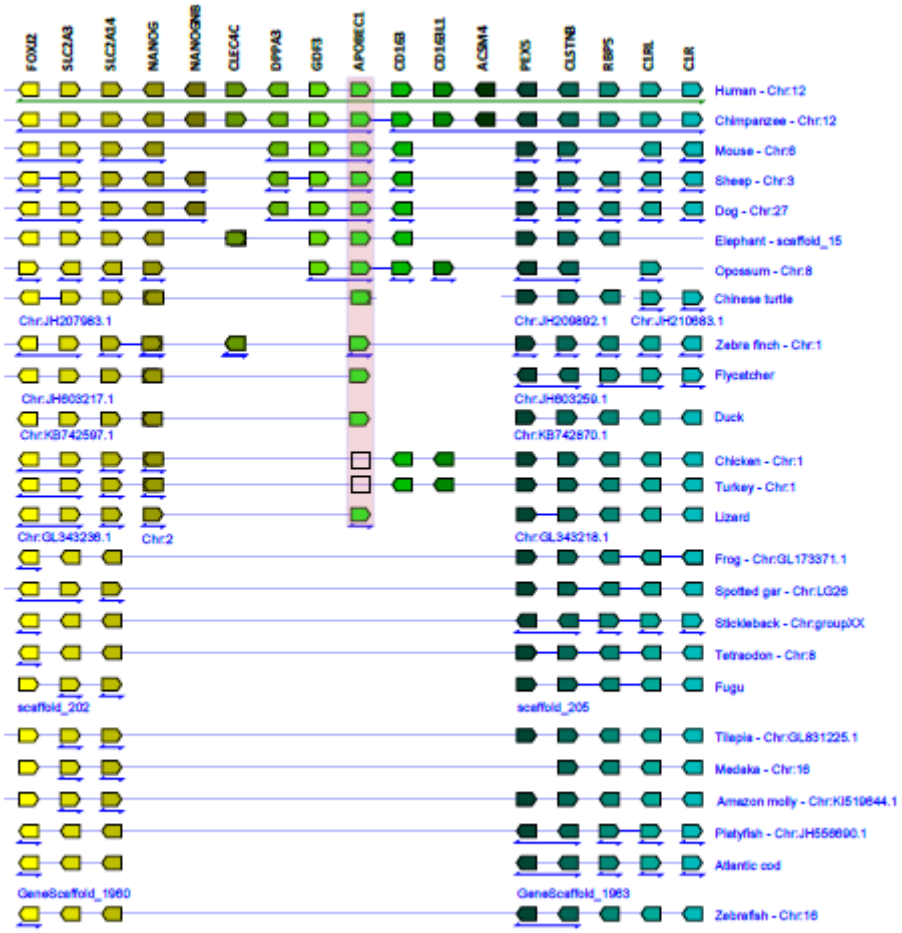


Figure 4

Figure 5

Highlights

- No evidence for Apob editing in anguillids, an evolutionarily important basal teleost
- Lack of editing may cause hypertriglyceridemia evident in non-mammalian vertebrates
- Link between the presence of only full-length Apob and changes in Apoe expression