In silico analyses of mammalian lactate dehydrogenases: human, mouse, opossum and platypus LDHs

Roger S Holmes[1, 3], and Erwin Goldberg[2]

[1]School of Biomolecular and Physical Sciences, Griffith University, Nathan, Qld, Australia; [2]Department of Biochemistry, Molecular Biology and Cell Biology, Northwestern University, Evanston. Illinois. 60201. USA

[3]Corresponding Author:

Roger S Holmes, D.Sc.
School of Biomolecular and Physical Sciences
Griffith University
Nathan. Qld. 4111.Australia.
Email: r.holmes@griffith.edu.au Phone: 61-7-3735-5077

Keywords: Mammals; data mining; sequence analyses; lactate dehydrogenase
Running Head: Mammalian Lactate Dehydrogenase Genes

Summary

Three major mammalian lactate dehydrogenase (LDH) genes and proteins have been extensively investigated, including *LDHA* (major muscle isozyme); *LDHB* (major heart isozyme); and *LDHC* (major sperm isozyme), and another (*LDH6B*) has been reported in humans. In this study, *in silico* methods were used to predict the amino acid sequences, structures and gene locations for *LDH* genes and proteins using genome sequence databanks for human, mouse, opossum and platypus mammalian species. Amino acid sequence alignments and predicted secondary and tertiary structures enabled observation (by similarity) of key residues previously reported for human and mouse LDHA, LDHB and LDHC subunits. The human genome contained at least 4 *LDH* genes encoding LDH A, B, C and 6B subunits, with the predicted *LDH6B* gene showing no evidence of introns. Two other human *LDH6*-like genes were also observed, including *LDH6A* (7 introns) and *LDH6C* (single exon). Human *LDHA*, *LDHC* and *LDH6A* genes were located in tandem on chromosome 11, while *LDH6B* and *LDH6C* genes were on chromosomes 15 and 12, respectively. *In silico* evidence was obtained for at least 13 human *LDH* pseudogenes located on 10 separate chromosomes of the human genome, of which seven were imbedded within introns of other genes involved in distinct but unrelated functions. Opossum *LDHC* and *LDH6B* genes were located in tandem with the opossum *LDHA* gene on chromosome 5 and contained 7 (*LDHA* and *LDHC*) or 8 (*LDH6B*) exons. An amino acid sequence prediction for the opossum LDH6B subunit gave an extended N-terminal sequence, similar to the human and mouse LDH6B sequences, which may support the export of this enzyme into mitochondria. The platypus genome contained at least 3 LDH genes encoding LDHA, LDHB and LDH6B subunits. Phylogenetic studies and sequence analyses indicated that *LDHA*, *LDHB* and *LDH6B* genes are present in all mammalian genomes examined, including a monotreme species (platypus), whereas the *LDHC* gene may have arisen more recently in marsupial mammals.

Keywords: Mammals; amino acid sequence; genomics; lactate dehydrogenase; opossum; platypus.

Running Head: Mammalian LDH Genes and Subunits

Published in:

Computational Biology and Chemistry (2009) 33:379-385

Introduction

Mammalian lactate dehydrogenase (LDH; E.C.1.1.1.27) comprises three major families of conserved enzymes that catalyse the reversible interconversion of pyruvate and lactate, a key metabolic step in glycolysis and other metabolic pathways (Everse & Kaplan, 1973) At least five LDH tetrameric isozymes are reported in somatic mammalian tissues, comprising LDHA and LDHB subunits, whereas the homotetrameric $LDHC_4$ isozyme is found only in mature testis and spermatozoa (Goldberg & Hawtrey, 1967; Goldberg, 1973; Li et al., 1989). The *LDHA*, *LDHB* and *LDHC* families of mammalian *LDH* genes and subunits have been extensively investigated, with human and mouse *LDHA* and *LDHC* genes located in tandem on chromosomes 11 and 7 respectively (Edwards et al., 1989), as compared with the *LDHB* gene, on chromosomes 12 (human) and 6 (mouse) (Takeno & Li, 1989). Phylogenetic studies have indicated that the *LDHC* gene has arisen from independent gene duplication events during vertebrate evolution, including separate *LDHB* gene duplications in fish and birds (pigeon) (Zinkham et al., 1969; Markert et al., 1975; Hiraoka et al., 1989; Quattro et al., 1993; Mannen et al., 1997), and an *LDHA* gene duplication during mammalian evolution (Millan et al., 1987).

Transcription studies have reported two other human *LDHA*-like genes, designated as *LDH6A* and *LDH6B*, which are expressed in brain and testis respectively, and located on chromosome 11 (*LDH6A* in tandem with human *LDHA* and *LDHC* genes) (Ota et al., 2004) and chromosome 15 (*LDH6B*, an intronless gene) (Wang et al., 2005). In this study, we have identified and characterized *in silico* new forms of mammalian LDHs and described predicted amino acid sequences, protein subunit structures, gene locations and exonic structures for human (*LDH6C*), mouse (*LDH6B*), opossum (*LDHA; LDHB; LDHC;* and *LDH6B*) and platypus (*LDHA*, *LDHB* and *LDH6B*) genes and proteins, as well as the phylogenetic relationships for mammalian *LDH* gene families. *In silico* evidence is also presented for N-terminal extensions of LDH6B subunit sequences which may support mitochondrial export and location of human, mouse and opossum LDH6B.

Materials and Methods

In silico mammalian LDH gene and protein identification.

BLAST (Basic Local Alignment Search Tool) studies were undertaken using web tools from the National Center for Biotechnology Information (NCBI) (http://blast.ncbi.nlm.nih.gov/Blast.cgi) (Altschul *et al*, 1997). Protein BLAST analyses used previously reported human LDHA (Tsujibo et al., 1985), LDHB (Takeno and Li, 1989), LDHC (Millan et al., 1987) and LDH6B (Ota et al., 2004) amino acid sequences . Non-redundant protein sequence databases for several mammalian genomes were examined using the blastp algorithm, including the human (International Human Genome Sequencing Consortium, 2001); mouse (*Mus musculus*) (Mouse Sequencing Consortium, 2002); opossum (Mikkelsen et al., 2007); and platypus (Platypus Genome Sequencing Consortium, 2008). This procedure produced multiple BLAST 'hits' for each of the protein databases which were individually examined and retained in FASTA format, and a record kept of the sequences for predicted mRNAs and encoded CES-like proteins . These records were derived from annotated genomic sequences using the gene prediction method: GNOMON and predicted sequences with high similarity scores for mammalian LDH. With some exceptions, predicted LDHA, LDHB, LDHC and LDH6B protein subunit sequences were obtained in each case and subjected to *in silico* analyses of predicted protein and gene structures. Other LDH sequences were obtained following BLAT (BLAST-Like Alignment Tool) *in silico* analysis using the human LDHA, LDHB, LDHC and LDH6B sequences to interrogate human, mouse, opossum and platypus genome sequences using the UC Santa Cruz gene browser [http://genome.ucsc.edu/cgi-bin/hgBlat] (Kent *et al.* 2003) with the default settings to obtain Ensembl generated protein sequences by applying the method of Hubbard et al (2002) (http://www.ensembl.org/index.html).

BLAT analyses were subsequently undertaken for each of the predicted LDH amino acid sequences using the UC Santa Cruz gene browser [http://genome.ucsc.edu/cgi-bin/hgBlat] (Kent *et al.* 2003) with the default settings to obtain the predicted locations for each of the mammalian *LDH* genes, including predicted exon boundary locations and gene sizes. For a study of predicted human LDH pseudogenes, BLAT analyses were undertaken of the human genome using human LDHA, LDHB, LDHC and LDH6B-like subunit sequences in each case (see Table 1; Figure 1). Predicted human LDH pseudogene structures were deduced following corrections for changes in sequence and size, and details recorded for each pseudogene , including the BLAT

score, percentage of identity with the LDH subunit sequence used and its location within the human genome. Structures for human LDHA, LDHB and LDHC isoforms (splicing variants) were obtained using the AceView website (http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/index.html?human) to examine predicted gene and protein structures using this database of human mRNA sequences (Thierry-Mieg and Thierry-Mieg, 2006).

Predicted Structures and Properties for Mammalian LDH Subunits.

Predicted secondary and tertiary structures for human and other mammalian LDH-like subunits were obtained using the PSIPRED v2.5 web site tools provided by Brunel University [http://bioinf.cs.ucl.ac.uk/psipred/psiform.html] (McGuffin *et al.* 2000) and the SWISS MODEL web tools [http://swissmodel.expasy.org/], respectively (Guex & Pietsch 1997; Kopp & Schwede 2004). Reported tertiary structures for human LDHA (PDB ID **1i10A**), human LDHB (PDB ID **1i0zA**) (Read et al., 2001) and mouse LDHC (PDB ID **9ldtA**)(Hogrefe et al., 1987) served as references for predicted opossum LDHA and LDH6B; LDHB; and LDHC tertiary structures, respectively. Modeling ranges for the opossum LDH residues were as follows: 2 to 332 (LDHA); 2 to 333 (LDHB); 2 to 331 (LDHC); and 51-381 (LDH6B).

Theoretical isoelectric points and molecular weights for mammalian LDH subunits were obtained using Expasy web tools (http://au.expasy.org/tools/pi_tool.html). *In silico* prediction of an LDH N-terminal protein region that may support a mitochondrial targeting sequence and the identification of a potential cleavage site was conducted using MITOPROT web based methods (Claros and Vincens, 1996) (ftp://ftp.biologie.ens.fr/pub/molbio) .

Phylogenetic Studies and Sequence Divergence

Phylogenetic trees were constructed using an amino acid alignment from a ClustalW-derived alignment of CES protein sequences, obtained with default settings and corrected for multiple substitutions (Chenna *et al* 2003; Larkin *et al*. 2007) [http://www.ebi.ac.uk/clustalw/]. An alignment score was calculated for each aligned sequence by first calculating a pairwise score for every pair of sequences aligned. The alignment ambiguous amino-terminus region was excluded prior to phylogenetic analysis yielding alignments of 332 residues for comparisons of mammalian LDHA, LDHB, LDHC and LDH6B sequences with chicken LDHA and LDHB sequences, which served as 'outgroup' sequences (see Table 1). Sequence identities for mammalian LDH subunits were determined using the SIM-Alignment tool for Protein Sequences [http://au.expasy.org/tools/sim-prot.html] (Pietsch 1995; Schwede *et al.* 2003).

Results and Discussion

Alignments of human LDHA, LDHB, LDHC, LDH6A, LDH6B and LDH6C amino acid sequences.

The amino acid sequences for human LDHA (Tsujibo et al., 1985), LDHB (Takeno and Li, 1989), LDHC (Millan et al., 1987) and LDH6B (Ota et al., 2004) and the *in silico* derived LDH6A and LDH6C human subunits are aligned in Figure 1 (see Table 1). Human LDH A, B, C and 6B subunits showed 71-75% sequence identities, indicating extensive conservation in amino acid sequences for these enzymes (Table 2). Major differences were observed however at the N-termini for the human LDH6B and LDH6C subunits, which showed an extension of 49 residues. MITOPROT computer based analyses of these sequences predicted a high probability for LDH6B and LDH6C subunit export into mitochondria (0.92 and 0.78, respectively), as well as a potential cleavage site at residue 31, in each case (Table 1; see Figure 1). The predicted mitochondrial N-terminal sequences were positively charged, with excess basic amino acid residues (3 and 2 respectively for LDH6B and LDH6C), contained no acidic residues and revealed a predicted amphiphilic α-helix, which are common features for mitochondrial leader sequences (Hanmen and Weiner, 1998). Key LDH catalytic residues were present in all six human LDH subunits, including the active site proton acceptor (His193), as well as coenzyme (Arg99 and Asn138) and substrate (Arg106; Arg169; Thr248) binding residues (Figure 1) (Read et al., 2001).

Alignments of mammalian LDHA, LDHB, LDHC and LDH6B amino acid sequences.

The amino acid sequences for predicted mouse LDH6B, opossum LDHA, LDHB, LDHC and LDH6B, and platypus LDHA, LDHB and LDH6B subunits are aligned with previously reported sequences for the corresponding human and mouse subunits (Tsujibo et al., 1985; Takeno and Li, 1989; Millan et al., 1987; Fukasawa and Li, 1987; Sakai et al., 1987; Hiraoka et al., 1990) (Figure 2; see Table 1). The predicted opossum and platypus LDH sequences

showed higher levels of identity with homologue sequences from human and mouse sources, particularly for the LDHA and LDHB sequences, which were 89-93% identical and 80-97% identical, respectively. Mammalian LDHC and LDH6B sequences, however, exhibited lower levels of identity, showing 65-74% identity for human, mouse and opossum LDHC sequences and 59-75% for human, mouse and platypus LDH6B sequences, respectively (Table 2). Mammalian LDH6B sequences showed evidence of N-terminus extensions for the predicted mouse, opossum and platypus subunits in comparison with LDHA, LDHB and LDHC sequences for all species examined (Figure 2). MITOPROT computer based analyses of these sequences predicted high probabilities for mouse and opossum LDH6B subunit export into mitochondria (0.98 and 0.79, respectively), as well as potential cleavage sites at residues 36 (mouse LDH6B) and 49 (opossum LDH6B) (Table 1; Figure 2). The platypus LDH6B sequence, however, differed significantly in this property, with the 53 residue N-terminus extension showing a lower probability as a mitochondrial signal peptide (0.27) (Table 1; Figure 2). Mitochondrial LDH (Brooks et al, 1999) has been previously proposed to play a role in the intracellular lactate shuttle and in lactate clearance by mitochondria, however the responsible LDH isozyme(s) have not been conclusively identified. The identification of a mitochondrial leader sequence for human, mouse and opossum LDH6B subunits may assist further investigations concerning a potential role for mammalian LDH in mitochondrial lactate clearance.

Each of the predicted mouse (LDH6B), opossum (LDHA; LDHB; LDHC; and LDH6B) and platypus (LDHA; LDHB; and LDH6B) sequences aligned closely with the corresponding human and mouse sequences, and all subunits (with one exception), showed sequence identity for the key active site residues previously described for human LDH subunits (see Read et al., 2001). The predicted platypus LDH6B sequence, however, contained an Arg residue in place of the key LDHA coenzyme binding residue (Asn138), which may significantly alter the kinetic properties for this enzyme.

Differences in the theoretical isoelectric points (pI) for opossum and platypus LDHA and LDHB subunits were observed, with LDHA showing higher pI values (7.1 and 8.2) than for the LDHB subunits (5.7 and 7.1), which is consistent with pI differences observed for other mammalian LDHs (Table 1). LDH6B subunits showed higher pI values than for the LDHA and LDHB subunits, which may be explained by the high basic amino acid content for the N-terminus peptide extensions, whereas theoretical pI values for mammalian LDHC subunits were intermediate between LDHB (lower pI) and LDHA/LDH6B (higher pI). Human, mouse and opossum LDHA, LDHB and LDHC subunits examined contained 331-334 amino acid sequence residues, whereas LDH6B subunits contained 381-385 amino acids due to the N-terminus extensions in each case.

Comparative Mammalian LDH Genomics

The NCBI AceView web browser currently defines the human *LDHA* gene by 7912 GenBank accessions isolated from a wide range of tissues (http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/index.html?human) (Thierry-Mieg and Thierry-Mieg, 2006). These human *LDHA* transcripts included 20 alternatively spliced variants (*LDHA* isoforms) which may result from differential 5' or 3' end truncations, exon shuffling, overlapping exons with different boundaries and alternative splicing (Figure 3). The human *LDHB* gene is also defined by a large number of GenBank accessions and the *LDHB* transcripts included 14 alternatively spliced variants also resulting from differential truncations of the 5' and 3' ends, exon shuffling and overlapping exons with different boundaries. In contrast, transcription of the human *LDHC* gene produced only 6 alternatively spliced mRNAs apparently resulting from exon shuffling (Figure 3). The differential roles for these splicing variants (LDH isoforms) for human LDHA, LDHB and LDHC isozymes have not been established.

Figures 1 and 2 show the locations of the intron-exon boundaries for the mammalian *LDH* gene products examined, and compares them with previously reported human and mouse *LDH* gene structures (Chung et al., 1985; Fusakawa and Li, 1987; Takeno and Li, 1989a,b) and their positioning within the aligned amino acid sequences. The mammalian *LDHA*, *LDHB* and *LDHC* genes examined, and the predicted human *LDH6A* gene, contained 7 exons in each case, with intron-exon boundaries in identical or comparable positions. In contrast, the human and mouse LDH6B genes were without intronic sequences, confirming a report for the human LDH6B gene (Wang et al., 2005), for which expression was observed in human testis. The predicted *LDH6B* genes in the opossum and platypus genomes, however, contained 8 exons, with the first exon encoding the predicted N-terminus extensions for these gene products, whereas the other 7 exons were localized in similar or identical positions to other mammalian *LDH* genes.

Table 1 describes the predicted locations for the mammalian *LDH* genes examined which showed that human, mouse and opossum *LDHA* and *LDHC* genes are located together within respective genomes on chromosomes 11, 7 and 5, respectively. The human and mouse LDHA and LDHC genes are very closely located together being separated by < 7 kilobases of DNA. The predicted human *LDH6A* gene is also part of this gene cluster on chromosome 11, as is the opossum *LDH6B* gene on chromosome 5 of the opossum genome. In addition, the platypus *LDHA* and *LDH6B* genes are apparently located on or near the same contiguous piece of DNA (Contig3116) suggesting that these genes are also closely located on the platypus genome. In contrast, the human, mouse and opossum *LDHB* genes are on a separate chromosome to that of the *LDHA*-like gene cluster (Table 1).

Table 3 compares the predicted locations, sizes, exon number and percentage identities for 13 proposed human *LDH* pseudogenes. Ten of these predicted pseudogenes showed a higher degree of identity with the human *LDHA* exonic sequences, and were designated as *LDHAps* genes; two showed higher sequence identity with human LDHB exonic sequences (designated as *LDHBps* genes); and one was more closely aligned with human LDHC gene exonic sequences (*LDHCps1*). These predicted genes are apparently located on 9 different chromosomes, and several of these were localized within intron sequences for other genes, which encode proteins responsible for distinct functions in the body, such as the *LYST* gene, encoding a lysosome trafficking regulatory protein (Barbosa et al., 1996); the *DYH6* gene, encoding dynein heavy chain 6 (Ota et al., 2004); the *MYO1E* gene, encoding myosin 1E (Bement et al., 1994); and the *S4A4* gene, encoding a solute carrier protein family 4 (Burnham et al., 1997). It is possible that *LDH* pseudogenes may perform as yet unknown regulatory functions for a range of genes (eg *LYST*; *DYH6; MYO1E;* and *S4A4*) or may serve as passive genetic elements within intronic sequences for these and other genes of the human genome.

Secondary and Tertiary Structures for Mammalian (and Chicken) LDH Sequences

Figures 1 and 2 show the secondary structures previously reported for human LDHA and LDHB (Read et al., 2001) and for mouse LDHC (Hogrefe et al., 1987) or predicted for mammalian LDHA, LDHB, LDHC and LDH6B subunit sequences, together with human LDH6A and LDH6C sequences. Predicted secondary structures for chicken LDHA and LDHB sequences were also examined as these were used as 'outgroup' LDH sequences for comparative analyses of mammalian *LDH* gene and protein structures. Similar α-helix β-sheet structures were observed for all mammalian and chicken LDH subunits examined, particularly near key residues or functional domains, including active site residues such as the active site proton acceptor (His193), as well as coenzyme (Arg99 and Asn138) and substrate (Arg106; Arg169; Thr248) binding residues (Read et al., 2001; Hogrefe et al., 1987). The obvious major difference in mammalian LDH secondary structure related to the N-terminus extensions for human LDH6B and LDH6C, and for mouse and opossum LDH6B, which contained an additional amphiphilic α-helix at the amino terminus, which may support being exported into mitochondria via these potential mitochondrial leader sequences (see Table 1). Although the platypus LDH6C N-terminal sequence contained a predicted α-helix, this did not extend into regions containing basic amino acid residues which may explain the lower probability for this sequence as a mitochondrial signal peptide (Table 1; Figure 2). Predictions of LDH secondary structures, however, may not fully reflect structures *in vivo* and serve only as a guide as to the comparative structures for mammalian LDH subunits.

Predicted tertiary structures for opossum LDHA, LDHB, LDHC and LDH6B subunits were examined and compared with previously reported tertiary structures for human LDHA and LDHB (Read et al., 2001), and for mouse LDHC (Hogrefe et al, 1987) (Figure 4). The predicted tertiary structures for opossum LDHA (residues 2 to 332) and LDH6B (residues 51 to 381) were sufficiently similar to the human LDHA structure to be based on the previously reported human LDHA-NADH-oxamate complex structure (Read et al., 2001) (Figure 4). In addition, the predicted structures for opossum LDHB and LDHC were sufficiently similar to the previously reported human LDHB-NADH-oxamate complex (residues 2-333) (Read et al., 2001) and mouse LDHC (residues 2 to 331) (Hogrefe et al., 1987), respectively. It is apparent from these predictions that LDHA, LDHB, LDHC and LDH6B subunits are highly conserved in mammals, and it is likely that LDH subunits in the opossum will resemble the corresponding LDHs in human.

Phylogeny of Mammalian LDH Subunits

A phylogenetic tree (Figure 5) was calculated by the progressive alignment of human LDHA, LDHB, LDHC and LDH6B amino acid sequences with the corresponding LDH sequences from mouse, opossum and the

platypus. Chicken LDHA and LDHB sequences were also included and served as an 'outgroup' for this analysis of mammalian LDHs. Four major clusters of mammalian and chicken LDHs were observed: the mammalian (and chicken) LDHA and LDHB gene clusters; the LDHC gene cluster of human, mouse and opossum; and the LDH6B cluster of human, mouse, opossum and platypus. This is consistent with the existence of four distinct mammalian LDH gene families: LDHA, encoding the major skeletal muscle isozyme; LDHB, encoding the major heart isozyme (Markert et al., 1975); LDHC, encoding the testis and sperm specific isozyme (Millan et al., 1987); and LDH6B, which awaits more detailed investigation. LDHA and LDHB have been described in all vertebrates examined and may be considered as the 'ancestral' genes for this enzyme (Holmes, 1972; Markert et al., 1975). In contrast, the LDHC gene has arisen independently from the LDHB gene in both teleost fish (Quattro et al., 1993) and in some birds (eg. pigeon) (Zinkham et al., 1969; Mannen et al., 1997), while in mammals, the LDHC gene has been apparently formed from an LDHA gene duplication event (Millan et al., 1987; Mannen et al., 1997). Biochemical studies have previously shown that LDHA, LDHB and LDHC isozymes are present in several Australian marsupials examined, including the pretty-faced wallaby (*Macropus parryi*), the koala (*Phascolarctos cincereus*) and the brush-tailed possum (*Trichosurus vulpecula*) (Holmes et al., 1973) whereas LDHC is apparently absent in monotreme mammals, the echidna (*Tachyglossus aculeatus*) and the platypus (Ornithorhynchus anatinus) (Baldwin and Temple-Smith, 1973). This study of LDH genes and proteins predicted from the South American gray short-tailed opossum (*Monodelphis domestica*) genome lends support to the distribution of LDHA, LDHB and LDHC genes and proteins among marsupials from both Australia and South America. The absence of an *LDHC*-like gene in the monotreme (platypus) genome, however, suggests that the proposed *LDH-A* gene duplication event leading to the appearance of the marsupial LDHC gene may have occurred following the separation of marsupial and monotreme common ancestors. In contrast, the mammalian *LDH6B* gene is apparently present throughout eutherian, marsupial and monotreme mammalian evolution but is apparently absent in the chicken genome (Table 1; Figure 5). A further LDHA gene duplication event is proposed forming the ancestral LDH6B gene at an earlier stage of mammalian evolution, prior to the separation of monotremes from the marsupial and eutherian mammalian common ancestors. This is supported by the higher levels of sequence identities observed for LDHA and LDH6B subunits (65-71%) as compared with LDHB and LDH6B subunits (57-62%), and the close locations observed for *LDHA* and *LDH6B* genes for the mammalian genomes examined.

Summary and Conclusions

Mammalian LDHs comprise at least four gene families encoding distinct subunits (A; B; C; 6B) which form tetrameric enzymes and catalyze a key step in carbohydrate metabolism in all tissues of the body. *LDH* genes are differentially expressed in mammalian tissues, with *LDHA* and *LDHB* genes exhibiting high expression levels in skeletal and heart muscle respectively, but with wide tissue expression patterns (Everse and Kaplan, 1973; Markert et al., 1975). In contrast, the *LDHC* gene is expressed predominantly in spermatocytes and the mature testis, and is required for male fertility (Odet et al., 2008). This isozyme plays an essential role in ATP production by glycolysis in spermatozoa. The human *LDH6B* gene has not been extensively studied, but has been shown to lack introns and to be expressed in testis (Wang et al., 2005).

In this study, we report *in silico* predictions for the amino acid sequences, structures and gene locations for *LDH* genes and proteins of four mammalian species, the human, mouse, opossum (a South American marsupial) and platypus (an Australian monotreme). The human genome contained at least 4 *LDH* genes encoding LDH A, B, C and 6B subunits, with the predicted *LDH6B* gene showing no evidence of introns. Two other human *LDH6*-like genes were observed, including an intronless *LDH6C* gene and a proposed *LDH6A* gene, which contained 7 introns. Human *LDHA*, *LDHC* and *LDH6A* genes were located in tandem on chromosome 11, while *LDH6B* and *LDH6C* genes were located on chromosomes 15 and 12, respectively. Several *LDH* pseudogenes were located elsewhere on the human genome, of which seven were apparently located within introns of other genes involved in distinct but unrelated functions. Opossum *LDHC* and *LDH6B* genes were located in tandem with the opossum *LDHA* gene on chromosome 5 and contained 7 (*LDHA* and *LDHC*) or 8 (*LDH6B*) exons. An amino acid sequence prediction for the opossum LDH6B subunit yielded an extended N-terminal sequence, similar to the human and mouse LDH6B sequences, which are proposed to support the export of these enzymes into mitochondria. The platypus genome contained at least 3 *LDH* genes encoding LDHA, LDHB and LDH6B subunits. Phylogenetic studies analyses indicated that *LDHA*, *LDHB* and *LDH6B* genes are present in all mammalian genomes examined, including a monotreme (platypus), whereas the

LDHC gene may have arisen more recently in marsupial mammals prior to the appearance of eutherian mammals.

Acknowledgements:

This project was supported in part by NIH HD05863 to E.G.

REFERENCES

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. 1997, Basic local alignment search tool. J. Mol. Biol. 215:403-410.
Baldwin, J., Temple-Smith, P., 1973. Distribution of LDHX in mammals: presence in marsupials and absence in the monotremes platypus and echidna. Comp. Biochem. Physiol. B 46, 805-811.
Barbosa, F.S., Nguyen, Q.A., Tschernev, V.T., Ashley, J.A., Detter, J.C., Blaydes, S.M., Brandt, S.J., Chotai, D., Hodgman, C., Solari, R.C., Lovett, M., Kingsmore, S.F., 1996. Identification of the homologous beige and Chediak-higashi syndrome genes. Nature 382, 262-265.
Bement, W.M., Wirth, J.A., Mooseeker, M.S., 1994. Cloning and mRNA expression of human unconventional myosin-IC. A homologue of amoeboid myosins-I with a single IQ motif and an SH3 domain. J. Mol. Biol. 243, 356-363.
Brooks, G.A., Dubouchaud, H., Brown, M., Sicurello, J.P., and Butz, C.E., 1999. Role of mitochondrial lactate dehydrogenase and lactate oxidation in the intracellular lactate shuttle. Proc. Natl. Acad. Sci. USA 96, 1129-1134.
Burnham, C.E., Amlal, H., Wang, Z., Shull, G.E., Soleimani, M., 1997. Cloning and functional expression of a human kidney Na+:HCO3-cotransporter. J. Biol. Chem. 272, 19111-19114.
Chenna R., Sugawara H., Koike T., Lopez, R., Gibson, T.J., Higgins, D.J., Thompson, J.D., 2003. Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Res. 31, 3497-3500.
Chung, F.Z., Tsujibo, H., Bhattacharyya, U., Sharief, F.S., Li, S.S.-L., 1985. Genomic organization of the human lactate dehydrogenase-A gene. Biochem. J. 231, 537-541.
Claros, M.G., Vincens, P., 1996. Computational method to predict mitochondrially imported proteins and their targeting sequences. Eur. J. Biochem. 241, 779-786.
Edwards, Y., West, L., Van Heyningen, V., Cowell, J., Goldberg, E., 1989. Regional localization of the sperm specific lactate dehydrogenase, LDHC, gene on human chromosomal 11. Ann. Human Genet. 53, 215-219.
Everse, J., Kaplan, N.O., 1973. Lactate dehydrogenases: structure and function. Adv. Enzym. Mol. Biol. 37, 61-133.
Fukasawa, K.M., Li, S.S.-L., 1987. Complete nucleotide sequence of the mouse lactate dehydrogenase-A functional gene: comparison of the exon-intron organization of dehydrogenase genes. Genetics 116, 99-105.
Goldberg, R., Hawtrey, C., 1967. The ontogeny of sperm specific lactate dehydrogenase in mice. J. Exp. Zool. 164, 309-316.
Goldberg, E., 1973. Molecular basis of multiple forms of LDH-X. J. Exp. Zool. 186, 273-278.
Gue,N., Peitsch, M.C., 1997. SWISS-MODEL and the Swiss-Pdb viewer. An environment for comparative protein modeling. Electrophoresis 18, 2714-2723.
Hanmen, P.K., Weiner, H., 1998. Mitochondrial leader sequence: structural similarities and sequence differences. J. Exp. Zool. 282, 280-283.
Hiraoka, K.A., Sharief, F.S., Yamg, Y.W., Li, W.H., Li, S.S.-L., 1990. The cDNA and protein sequences of mouse lactate dehydrogenase B. Molecular evolution of vertebrate lactate dehydrogenase genes A (muscle), B (heart) and C (testis). Eur. J. Biochem. 189, 215-220.
Hogrefe, H.H., Griffith, J.P., Rossmann, M.G., Goldberg, E., 1987. Characterization of the antigenic sites on the refined 3-A resolution structure of mouse testicular lactate dehydrogenase C4. J.Biol.Chem. 262, 13155-13162.
Holmes, R.S., 1972. Evolution of lactate dehydrogenase genes. FEBS Letters 28, 51-55.
Holmes, R.S., Cooper, D.W., VandeBerg, J.L., 1973. Marsupial and monotreme lactate dehydrogenase isozymes: phylogeny, ontogeny and homology with eutherian mammals. J. Exp Zool. 184, 127-148.
Hubbard, T.J.P., Aken, B.L., Beal1, K., Ballester, B., Caccamo, M., Chen, Y., Clarke, L., Coates, G., Cunningham, F., Cutts, T., Down, T., Dyer, S.C., Fitzgerald, S., Fernandez-Banet, J., Graf, S., Haider, S., Hammond, M., Herrero, J., Holland, R., Howe, K., Johnson, N., Kahari, A., Keefe, D., Kokocinski, F., Kulesha, E., Lawson, D., Longden, I., Melsopp, C., Megy, K., Meidl, P., Overduin, B., Parker, A., Prlic, A., Rice, S., Rios, D., Schuster, M., Sealy, I., Severin, J., Slater, G., Smedley, D., Spudich, G., Trevanion, S., Vilella, A., Vogel, J., White, S., Wood, M., Cox, T., Curwen, V., Durbin, F., Fernandez-Suarez, X.P., Flicek, P., Kasprzyk, A., Proctor, G., Searle, S., Smith, J., Ureta-Vidal, A., Birney, E.,2007. Ensembl. Nucleic Acids Res. 35, D610-D617.
International Human Genome Sequencing Consortium., 2001. Initial sequencing and analysis of the human genome. Nature 409, 860-921.

Kent W.J., Sugnet C.W., Furey T.S., Roskin, K.M., Pringle, T. H., Zahler, A. M., Haussler, D., 2002. The human genome browser at UCSC. Genome Res., 12, 994-1006.
Kopp, J., Schwede, T., 2004. The SWISS-MODEL repository of three dimensional protein structure homology models. Nucl. Acids res. 32, D230-D234.
Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins D.G., 2007. ClustalW2 and ClustalX version 2. Bioinformatics 23, 2947-2948.
Li, S.S., O'Brien, D.A., How, E.W., Versola, J., Rockett, D.L., Eddy, E.M., 1989. Differential activity and synthesis of lactate dehydrogenase isozymes A (muscle), B (heart) and C (testis) in mouse spermatogenic cells. Biol. Reprod. 40, 173-180.
McGuffin, L.J., Bryson, K., Jones, D.T., 2000. The PSIPRED protein structure prediction server. Bioinformatics 16, 404-5.
Mannen, H., Tsoi, S.C.-M., Krushkal, J.S., Li, W-H., Li, S.S-L., 1997. The cDNA cloning and molecular evolution of reptile and pigeon lactate dehydrogenase isozymes. Mol. Biol. Evol. 14, 1081-1087.
Markert, C.L., Shaklee, J.B., Whitt, G.S., 1975. Evolution of a gene. Multiple genes for LDH isozymes provide a model of the evolution of gene structure, function and regulation. Science 189, 102-114.
Mikkelsen, T.S., Wakefield, M.J., Aken, B., Amemiya, C.T., Chang, J.L., Duke, S., Garber, M., Gentles, A.J., Goodstadt, L., Heger, A., Jurka, J., Kamal, M., Mauceli, E., Searle, S.M.J., Sharpe, T., Baker, M.L., Batzer, M.A., Benos, P.V., Belov, K., Clamp, M., Cook, A., Cuff, J., Das, R., Davidow, L., Deakin, J.E., Fazzari, M.J., Glass, J.L., Grabherr, M., Greally, J.M., Gu, W., Hore, T.A., Huttley, G.A., Kleber, M., Jirtle, R.L., Koina, E., Lee, J.T., Mahony, S., Marra, M.A., Miller, R.D., Nicholls, R.D., Oda, M., Papenfuss, A.T., Parra, Z.E., Pollock, D.D., Ray, D.A., Schein, J.E., Speed, T.P., Thompson, K., VandeBerg, J.L., Wade, C.M., Walker, J.A., Waters, P.D., Webber, C., Weidman, J.R., Xie, X., Zody, M.C., Broad Institute Genome Sequencing Platform and Broad Institute Whole Genome Assembly Team, Marshall Graves, J.A., Ponting, C.P., Breen, M., Samollow, P.B., Lander, E.S., Lindblad-Toh, K., 2007. Genome of the marsupial Monodelphis domestica reveals innovation in non-coding sequences. Nature 447, 167-177.
Millan, J.L., Driscoll, C.E., Goldberg, E., 1987. Epitopes of human testis-specific lactate dehydrogenase deduced from a cDNA sequence. Proc. Natl. Acad. Sci. USA 84, 5311-5315.
Mouse Sequencing Consortium., 2002. Initial sequencing and comparative analysis of the mouse genome. Nature 420: 520-562.
Odet, F., Duan, C., Willis, W.D., Goulding, E.H., Kung, A., Eddy, E.M., Goldberg, E., 2008. Expression of the gene for lactate dehydrogenase C (*Ldhc*) is required for male fertility. Biol. Reprod. 79, 26-34.
Ota,T., Suzuki,Y., Nishikawa,T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,Sekine, M., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M., Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H., Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M., Hiraoka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S., Yosida, M., Hotuta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A., Hara, H., Tanase, T.O., Nomura, Y., Togiya, S., Komai, F., Hara, R.,Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A., Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T., Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S., Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y., Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T., Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T., Sugano, S. (2004) Complete sequencing and characterization of 21,243 full-length human cDNAs. Nature Genetics 36, 40-45.
Peitsch, M.C., 1995. Protein modeling by E-mail. Bio Tech. 13, 658-660.
Platypus Genome Sequencing Consortium., 2008. Genome analysis of the platypus reveals unique signatures of evolution. Nature 453, 175-183.
Quattro, J.M., Woods, H.A., and Powers, D.A., 1993. Sequence analysis of teleost retina-specific lactate dehydrogenase C: evolutionary implications for the vertebrate lactate dehydrogenase gene family. Proc. Natl Acad. Sci. USA 90, 242-246.
Read, J.A., Winter, V.J., Eszes, C.M., Sessions, R.B., Brady, R.L., 2001. Structural basis for altered activity of M- and H- isozyme forms of human lactate dehydrogenase. Protein Sci. 43, 175-185.

Sakai, I., Sharief, F.S., Li, S.S.-L., 1987. Molecular cloning and nucleotide sequence of the cDNA for sperm-specific lactate dehydrogenase-C from mouse. Biochem. J. 242, 619-622.
Schwede, T., Kopp, J., Guex, N., Peitsch, M.C., 2003. SWISS-MODEL: an automated protein homology-modeling server. Nucleic Acids Res. 31, 3381-3385.
Takeno, T., Li, S.S-L., 1989a. Structure of the human lactate dehydrogenase B gene. Biochem. J. 257, 921-924.
Takeno, T., Li, S.S.-L., 1989b. Human testicular lactate dehydrogenase gene is interrupted by six introns at positions homologous to those of LDH-A (muscle) and LDH-B (heart) genes. Biochem. Biophys. Res. Commun. 159, 579-583.
Thierry-Mieg, D., Thierry-Mieg, J., 2006, AceView: A comprehensive cDNA-supported gene and transcripts annotation. Genome Biol. 7, S12.
Tsujibo, H., Tiano, H.F., Li, S.S., 1985, Nucleotide sequences of the cDNA and an intronless pseudogene for human lactate dehydrogenase-A isozyme. Eur. J. Biochem. 147, 9-15.
Wang, H., Zhou, Z., Lu, L., Xu, Z., Sha, J., 2005. Cloning and characterization of a novel intronless lactate dehydrogenase gene from adult testis. Int. J. Mol. Med. 15, 949-953.
Zinkham, W.H., Isensee, H., Renwick, J.H., 1969. Linkage of lactate dehydrogenase B and C loci in pigeons. Science 164, 185-187.

Figure Legends:

Figure 1: Amino acid sequence alignments for human LDHA, LDHB, LDHC, LDH6A, LDH6B and LDH6C sequences
See Table 1 for sources of LDH sequences; * shows identical residues; Residues identified by MITOPROT as high probability mitochondrial leader sequences; conserved active site residues Arg99 and 106; Asn138; Arg169; His193; and Thr248 Helix (Human LDHA and LDHB or predicted helix); Sheet (Human LDHA and LDHB or predicted sheet). **Bold underlined font** shows known or predicted exon junctions (|)**.** A, B, C, 6A, 6B and 6C refer to the corresponding human LDH subunits.

Figure 2: Amino acid sequence alignments for human, mouse, opossum, platypus and chicken LDH sequences
See Table 1 for sources of LDH sequences; * shows identical residues; Residues identified by MITOPROT as high probability mitochondrial leader sequences; conserved active site residues Arg99 and 106; Asn138; Arg169; His193; and Thr248 Helix (Human LDHA and LDHB or predicted helix); Sheet (Human LDHA and LDHB or predicted sheet). **Bold underlined font** shows known or predicted exon junctions (|)**. LDHs examined included human (hu); mouse (mo); opossum (op); platypus (pl); and chicken (ch).**A, B, C and 6B refer to the corresponding LDH subunits.

Figure 3: Gene structures and splicing variants for human *LDHA*, *LDHB* and *LDHC* genes
Derived from the AceView website (Thierry-Mieg and Thierry-Mieg, 2006) http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/ Isoform variants (a, b, c etc) are shown with capped 5'- and validated 3'-ends for the predicted mRNA sequences. NM numbers refer to annotated RefSeq sequences for human *LDHA*, *LDHB* and *LDHC* genes. Scale refers to base pairs of nucleotide sequences.

Figure 4: Three dimensional structures for human LDHA and mouse LDHC subunits and predicted three dimensional structures for opossum LDHA, LDHB, LDHC and LDH6B subunits
Human LDHA and mouse LDHC 3-D structures and predicted opossum LDHA, LDHB, LDHC and LDH6B structures were obtained using the SWISS MODEL web site http://swissmodel.expasy.org/workspace/index.php? (see Table 1). The rainbow color code describes the 3-D structures from the N- (blue) to C-termini (red color). The structures are based on known 3-D structures for human LDHA and LDHB (Read et al., 2001) and mouse LDHC (Hogrefe et al., 1987) complexed with NADH and oxamate. Modeling ranges for the opossum LDH sequences were: LDHA: 2 to 332 based on PDB template 1i10A; LDHB: 2 to 333 based on PDB template 1i0zA ; LDHC: 2 to 331 based on PDB template 9ldtA ; and LDH6B: 51 to 381 based on template 1i10A .

Figure 5: Phylogenetic tree of mammalian CES6 and of human CES1, CES2, CES3 and CES5 sequences.
The tree is labeled with the LDH gene family number and the species name. Note the separation of the *LDH* genes into four *LDH* family clusters: LDHA; LDHB; LDHC; and LDH6B.

LEGENDS FOR TABLES

Table 1: Mammalian and chicken lactate dehydrogenase (*LDH*) genes and enzymes examined
GenBank mRNA (or cDNA) IDs identify previously reported sequences (see http://www.ncbi.nlm.nih.gov/Genbank/); [1]N-scan and [2]SGP IDs identify gene predictions using gene structure prediction software provided by the Computational Genomics Lab at Washington University in St. Louis, MO, USA (see http://genome.ucsc.edu); UNIPROT refers to UniprotKB/Swiss-Prot IDs for individual LDH subunits (see http://kr.expasy.org); [3]Mitochondrial export probabilities and predicted signal peptides were based on MITOPROT web based tools (see Methods); [4]Contig ID for platypus genome sequences; [5]Prediction software based ENSOANT IDs; Sources for LDH sequences were provided by the above sources.

Table 2: Percentage identities for mammalian and chicken LDH amino acid sequences

Numbers show the percentage of amino acid sequence identities. Numbers in **bold** show higher sequence identities for eutherian mammalian LDH sequences.

Table 3: Predicted human LDH pseudogenes

Predicted human LDH pseudogenes are named LDHAps, LDHBps or LDHCps in numerical order according to the subunit showing highest sequence identity and BLAT score using the UC Santa Cruz human genome web browser (http://genome.ucsc.edu). [1]BLAT score determined by using the relevant human LDH subunit sequence (A, B or C) to interrogate the human genome; [2]percentage identity of the derived human pseudogene sequence with the relevant LDH subunit sequence; [3]range of LDH subunit sequence corresponding to the derived LDH pseudogene sequence; [4]number of relevant LDH residues obtained for the derived LDH pseudogene sequence; [5]predicted pseudogene size (nucleotides); [6]predicted exon sequences observed; [7]GenBank or prediction software based ENSOANT IDs for the pseudogene sequence; [8]predicted colocation of the LDH pseudogene with another known human gene; [9]colocated gene function identified.

Figure 1

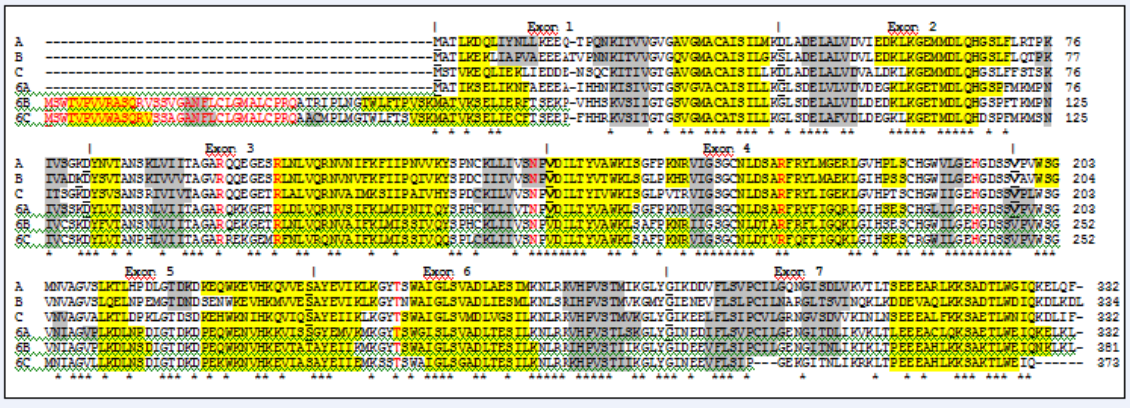

Figure 2

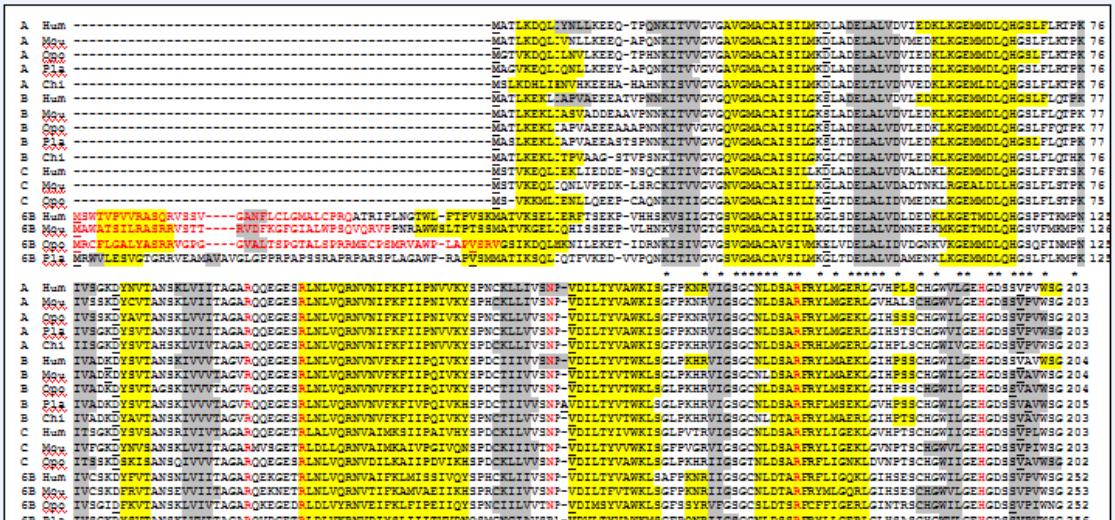

Figure 3

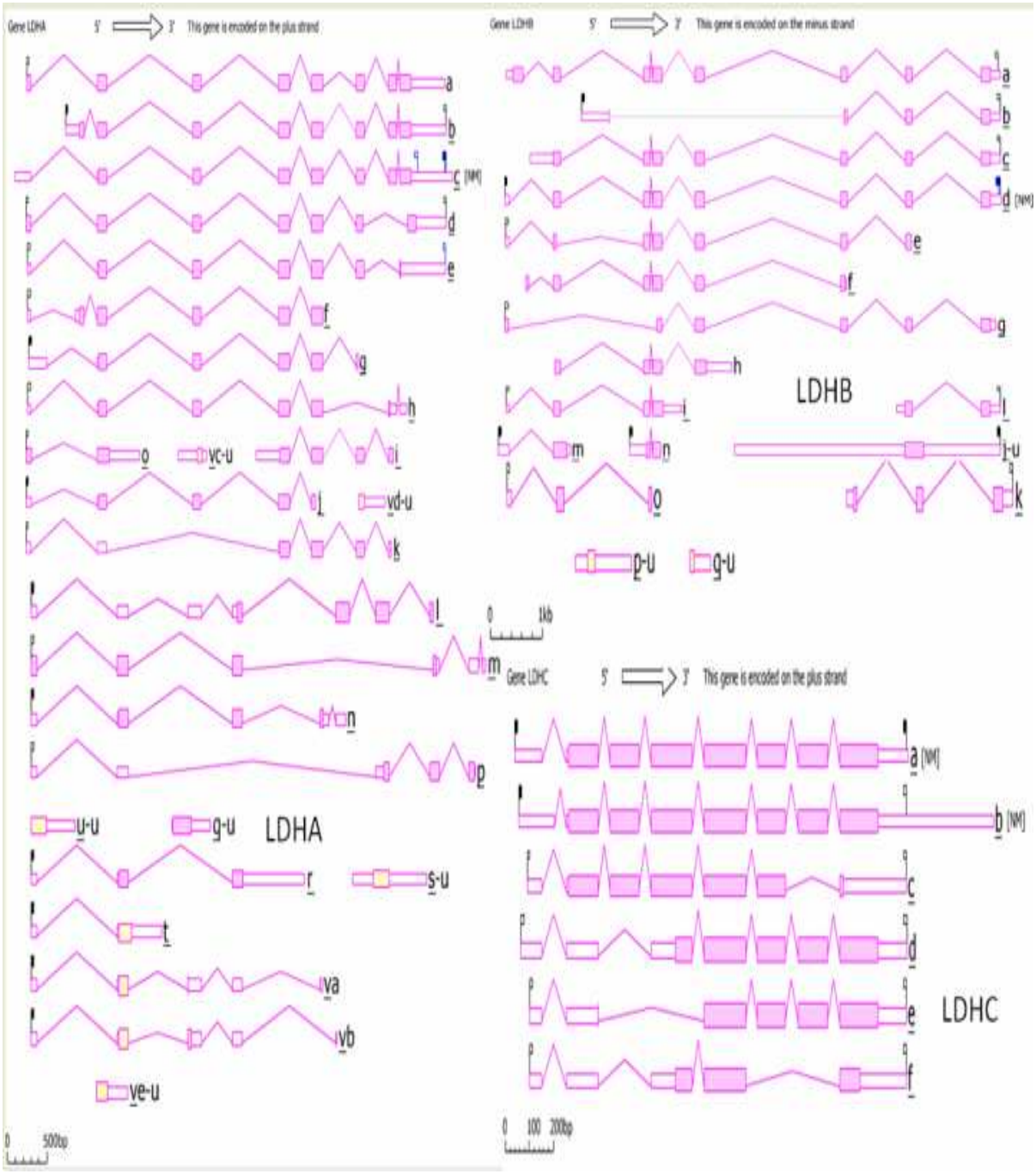
Gene LDHA
5'
3'
This gene is encoded on the plus strand
a
b
c [NM]
d
e
f
g
h
o
vc-u
i
vd-u
j
k
l
m
n
p
u-u
g-u
LDHA
r
s-u
t
va
vb
ve-u
0
500bp
Gene LDHB
This gene is encoded on the minus strand
a
b
c
d [NM]
e
f
g
h
i
LDHB
l
m
n
j-u
o
k
p-u
g-u
0
1kb
Gene LDHC
This gene is encoded on the plus strand
a [NM]
b [NM]
c
d
e
LDHC
f
0
100
200bp

Figure 4

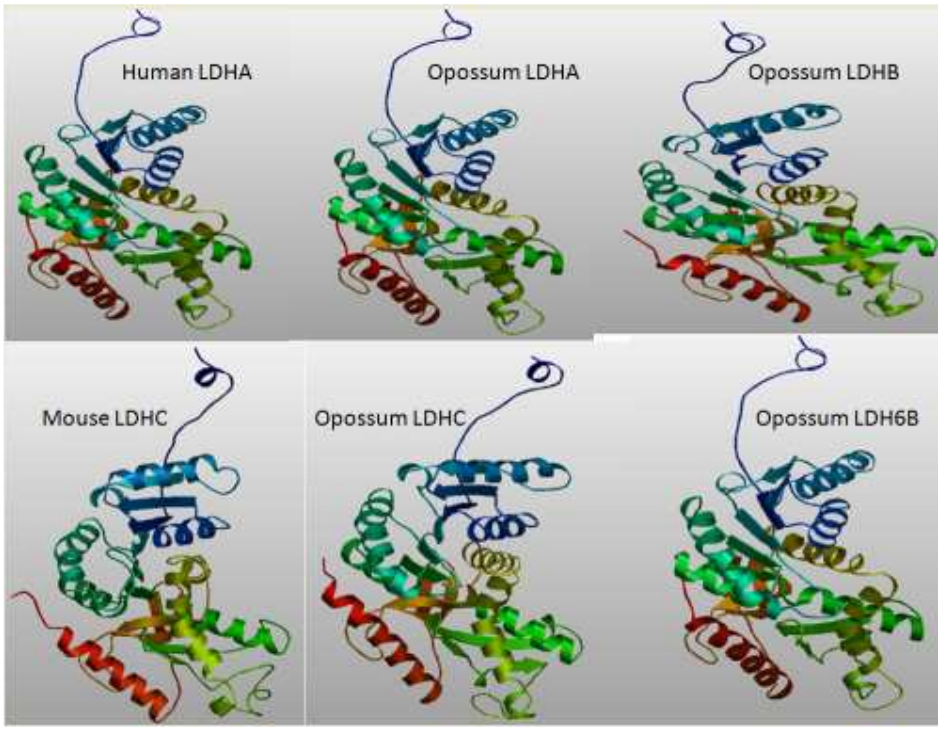

Figure 5

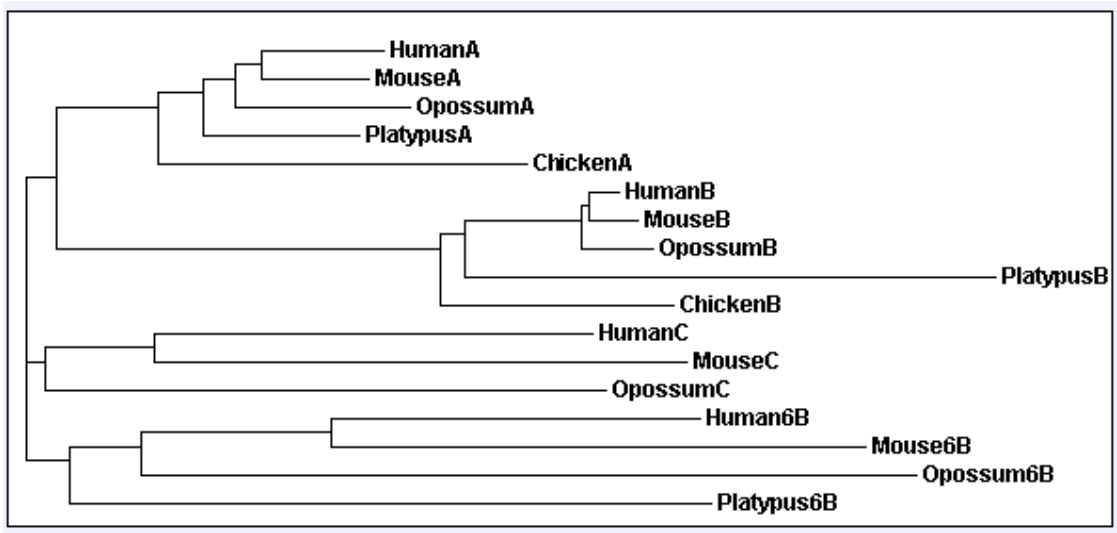

Table 1

Species	LDH Gene	GenBank ID	UNIPROT ID	NCBI RefSeq ID	Chromosome location	Strand	Amino Acids	Gene size kbs	Exons	pI	Subunit MW	[3]Mitochondrial Export Probability (Residues)
Human	LDHA	BC067223	P00338	NP005557	11: 18,374,966-18,385,401	positive	332	10,436	7	8.4	36,689	0.04 (Nil)
	LDHB	BC071860	P07195	NP002291	12: 21,679,746-21,698,872	negative	334	19,127	7	5.7	36,638	0.05 (Nil)
	LDHC	BC064388	P07864	NM002301	11: 18,390,841-18,429,247	positive	332	38,407	7	7.1	36,311	0.07 (Nil)
	LDH6A	BC014340		NP659409	11: 18,434,804-18,456,990	positive	332	22,187	7	6.5	36,507	0.14 (Nil)
	LDH6B	BC022034	Q9BYZ2	NP149972	15: 57,286,432-57,287,574	positive	381	1,143	1	8.9	41,943	**0.92 (1-33)**
	LDH6C			[1]SGP.12.853.1	12: 61,683,600-61,684,723	positive	373	1,124	1	8.6	41,157	**0.78 (1-33)**
Mouse	LDHA	BC004639	P06151	NP034829	7: 54,102,990-54,110,508	positive	332	7,519	7	7.6	36,499	0.02 (Nil)
	LDHB	BC046755	P16125	NP032518	6: 142,438,960-142,454,060	negative	334	15,101	7	5.7	36,572	0.09 (Nil)
	LDHC	BC049602	Q548Z6	NP038608	7: 54,117,140-54,133,244	positive	332	16,105	7	8.4	35,912	0.1 (Nil)
	LDH6B	BC019420		NP780558	17: 5,417,512-5,418,657	negative	382	1,146	1	9.3	42,049	**0.98 (1-37)**
Opossum	LDHA	AF070996	Q9XT87	NP1028147	5: 242,665,392-242,674,081	negative	332	8,690	7	7.1	36,358	0.02 (Nil)
	LDHB	AF070997	Q9XT86	[2]chr8.557.a	8: 93,264,241-93,287,176	positive	334	22,936	7	5.7	36,537	0.06 (Nil)
	LDHC	[2]chr5.25.018		XP1378365	5: 242,633,611-242,658,159	negative	331	24,549	7	6.8	36,303	0.1 (Nil)
	LDH6B	[2]chr5.25.016		XP1378357	5: 242,565,407-242,601,987	negative	381	36,581	8	8.7	41,871	**0.79 (1-50)**
Platypus	LDHA	AF545182			[4]11958: 1024-4697; [4]3118: 2244-4946	negative	332	6,377	7	8.2	36,451	0.04 (Nil)
	LDHB	[5]ENSOANT16632			[4]59108: 2671-2799; [4]8353: 5168-27343	positive	335	22,305	7	7.1	36,525	0.08 (Nil)
	LDH6B	[5]ENSOANT13298			[4]3118: 2264-26601	positive	385	17,338	8	8.8	41,920	0.27 (Nil)
Chicken	LDHA		P00340	NP990615	5: 13,645,367-13,649,740	positive	332	4,373	7	7.8	36,514	0.01 (Nil)
	LDHB		P00337	NP989508	1: 69,204,825-69,213,883	positive	333	9,059	7	7.1	36,318	0.08 (Nil)

Table 2

LDH Subunit	Hu A	Mo A	Op A	Pl A	Ch A	Hu B	MoB	Op B	Pl B	Ch B	Hu C	Mo C	Op C	Hu 6B	Mo 6B	Op 6B	Pl 6B
Human A	100	**93**	**90**	**92**	**84**	75	74	74	63	74	75	72	74	71	65	65	74
Mouse A	**93**	100	**92**	**91**	**84**	75	75	75	64	73	74	72	75	72	67	64	76
Opossum A	**90**	**92**	100	89	**82**	77	75	75	64	75	73	71	74	71	67	65	73
Platypus A	**92**	**91**	**89**	100	**85**	**75**	75	74	64	73	77	71	75	70	66	65	77
Chicken A	**84**	**84**	**82**	**85**	100	**74**	74	74	63	71	70	68	72	65	60	62	68
Human B	75	75	77	75	74	100	**97**	**97**	**81**	**90**	69	65	67	65	62	58	65
Mouse B	74	75	75	75	74	**97**	100	**96**	**81**	**89**	69	64	67	64	62	57	65
Opossum B	74	75	75	74	74	**97**	**96**	100	**80**	**89**	69	64	67	64	62	57	64
Platypus B	63	64	64	64	63	**81**	**81**	**80**	100	77	59	57	59	56	55	50	57
Chicken B	74	73	75	73	71	**90**	**89**	**89**	77	100	67	64	67	64	61	59	64
Human C	75	74	73	77	70	69	69	69	59	67	100	74	73	67	61	59	65
Mouse C	72	72	71	71	68	65	64	64	57	64	74	100	65	65	61	61	65
Opossum C	74	75	74	75	72	67	67	67	59	67	73	65	100	67	61	62	64
Human 6B	71	72	71	70	65	65	64	64	56	64	67	65	67	100	75	63	66
Mouse 6B	65	67	67	66	60	62	62	62	55	61	61	61	61	75	100	60	62
Opossum 6B	65	64	65	65	62	58	57	57	50	59	59	61	62	63	60	100	59
Platypus 6B	74	76	73	77	68	65	65	64	57	64	65	65	64	66	62	59	100

Table 3

Human LDH Pseudogene	BLAT [1]Score	% [2]Identity	Residue [3]Range	No. of [4]Residues	Gene Location	Strand	Gene [5]Size	[6]Exons	GenBank ID [7]Prediction	Gene [8]Colocation [9]Possible Function
LDHAps1	832	92	2-332	331	9: 14,911,338-14,912,329	negative	992	2	[7]ENST397561	
LDHAps2	739	88	2-332	316	1: 233,967,929-233,968,909	positive	981	4	[6]U67615	LYST: part of intron for lysosome trafficking regulator
LDHAps3	705	88	22-232	310	2: 41,900,453-41,901,398	positive	946	3	[7]ENST394996	
LDHAps4	702	86	1-332	317	4: 49,46,838-49,47,801	negative	964	3	[7]ENST400077	
LDHAps5	638	88	55-328	271	2: 84,857,909-84,858,721	negative	813	1	[7]ENST389394	DYH6: part of intron for dynein heavy chain 6
LDHAps6	498	90	128-332	203	10: 120,682,174-120,682,785	negative	612	3	[7]ENST402787	
LDHAps7	459	76	27-327	301	15: 57,286,657-57,287,559	positive	903	1	[7]ENST288235	MYO1E: part of intron for myosin 1E
LDHAps8	383	73	27-296	272	6: 1,520,274-1,521,085	negative	812	2	BC029130	BC029130: part of intron for unidentified gene
LDHAps9	383	73	27-296	270	6: 157,640,146-157,640,951	negative	806	2	[7]ENST339126	CF035: part of intron for membrane protein
LDHAps10	312	70	27-292	298	12: 61,683,825-61,684,723	positive	899	2	[7]ENST324626	AL833331: part of intron for testis mRNA encoding gene
LDHBps1	731	87	2-334	331	X: 75,471,647-75,472,645	negative	999	3	[7]ENST395646	
LDHBps2	199	77	186-306	119	13: 112,979,604-112,979,964	positive	361	2	[7]ENST404271	
LDHCps1	210	65	22-230	242	4: 72,518,995-72,522,295	negative	3301	6	BC030977	S4A4: part of intron for solute carrier protein family 4