Contents lists available at ScienceDirect

Chemico-Biological Interactions

journal homepage: www.elsevier.com/locate/chembioint

Opossum alcohol dehydrogenases: Sequences, structures, phylogeny and evolution Evidence for the tandem location of *ADH* genes on opossum chromosome 5

Roger S. Holmes*

School of Biomolecular and Physical Sciences, Griffith University, Kessels Road, Nathan, 4111 Brisbane, Qld, Australia

ARTICLE INFO

Article history: Available online 19 September 2008

Keywords: Alcohol dehydrogenase ADH Opossum Genes Gene evolution

ABSTRACT

BLAT (BLAST-Like Alignment Tool) analyses and interrogations of the recently published opossum genome were undertaken using previously reported rat ADH amino acid sequences. Evidence is presented for six opossum ADH genes localized on chromosome 5 and organized in a comparable ADH gene cluster to that reported for human and rat ADH genes. The predicted amino acid sequences and secondary structures for the opossum ADH subunits and the intron-exon boundaries for opossum ADH genes showed a high degree of similarity with other mammalian ADHs, and four opossum ADH classes were identified, namely ADH1, ADH3, ADH6 and ADH4 (for which three genes were observed: ADH4A, ADH4B and ADH4C). Previous biochemical analyses of opossum ADHs have reported the tissue distribution and properties for these enzymes: ADH1, the major liver enzyme; ADH3, widely distributed in opossum tissues with similar kinetic properties to mammalian class 3 ADHs; and ADH4, for which several forms were localized in extrahepatic tissues, especially in the digestive system and in the eye. These ADHs are likely to perform similar functions to those reported for other mammalian ADHs in the metabolism of ingested and endogenous alcohols and aldehydes. Phylogenetic analyses examined opossum, human, rat, chicken and cod ADHs, and supported the proposed designation of opossum ADHs as class I (ADH1), class III (ADH3), class IV (ADH4A, ADH4B and ADH4C) and class VI (ADH6). Percentage substitution rates were examined for ADHs during vertebrate evolution which indicated that ADH3 is evolving at a much slower rate to that of the other ADH classes.

© 2008 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Mammalian alcohol dehydrogenases (ADH; EC 1.1.1.) exist as a family of enzymes which metabolize a broad range of alcohol and carbonyl compounds and are encoded by at least six classes of *ADH* genes [1–3]. Further mammalian *ADH* gene multiplicity may occur as a result of recent gene duplication events during mammalian evolution generating multiple non-allelic *ADH* genes [4,5] or following *ADH* pseudogene formation [6]. ADH subunit multiplicity may also be generated by allelic variation of *ADH* genes [7–10] or by differential splicing events during *ADH* mRNA transcription see [11]. Although complex in nature, the nomenclature for mammalian ADH genes and encoded proteins has taken account of the extensive gene and protein multiplicity reported for this enzyme [1].

E-mail address: r.holmes@griffith.edu.au.

Mammalian liver class I ADHs in particular have been intensively investigated because of their roles in ethanol metabolism and in neurotransmitter, retinoid and bile acid metabolism see [12]. Human liver class I ADHs comprise 3 genetically distinct subunits $(\alpha, \beta \text{ and } \gamma)$ which form six homodimeric and heterodimeric isozymes [7]. Other ADH classes, with the exception of ADH4, are also localized in mammalian liver and include the class II enzyme ADH2, which is divided into two groups, human and rabbit forms and the rodent forms [13-15]. Class III ADH, ADH3, is widely distributed in mammalian tissues and functions in formaldehyde metabolism, involving S-hydroxymethyl glutathione, a conjugation product formed from glutathione and formaldehyde [16,17]; while ADH5 and ADH6 have been investigated only at the DNA and RNA level and analyzed for tissue distribution [12]. Mammalian class IV ADH, or ADH4, occurs in extra-hepatic tissues, particularly in stomach, intestine and eye [18-20], and may play a role in the first pass clearance of ingested alcohols and carbonyl compounds [10,21] and in retinoid metabolism [17,22]. Mammalian ADHs are zinc dependent dimeric enzymes which require NAD/NADH as coenzyme for





^{*} Tel.: +61733482834.

^{0009-2797/\$ –} see front matter @ 2008 Elsevier Ireland Ltd. All rights reserved. doi:10.1016/j.cbi.2008.09.009

catalytic activity, with each subunit consisting of about 375 amino acid residues [23,24].

Evolutionary studies on vertebrate ADHs have shown that class III ADH3 represents the primordial form and that both class I and class III ADHs are present in vertebrate fish and have co-existed for more than 500 million years [3,25]. These *ADH* genes have continued to diverge throughout vertebrate evolution and have undergone further gene duplication, generating additional *ADH* gene classes, as well as multiple *ADH* genes within an *ADH* class following more recent gene duplication.

This report outlines the predicted sequences, structures, phylogeny and evolution of *ADH* genes and enzymes in a South American marsupial, the gray short-tailed opossum (*Monodelphis domestica*), an established laboratory animal used to study the genetics of eye cancer and heart disease [26–28]. In silico methods were used to predict the amino acid sequences and secondary structures for opossum ADHs and gene locations for opossum *ADH* genes, using data from the recently released opossum genome sequence [29]. This paper extends previous biochemical genetic studies on opossum ADHs which reported the tissue distribution and biochemical properties of class I, class III and multiple class IV ADHs [30]. Phylogenetic analyses also describe the relationships and potential evolutionary origins of the opossum *ADH*1, *ADH*3, multiple *ADH4* and *ADH6* predicted gene and enzyme sequences with other previously reported mammalian and vertebrate ADHs.

2. Materials and methods

2.1. In silico opossum ADH gene and protein identification

BLAT (BLAST-Like Alignment Tool) *in silico* studies were undertaken using the UC Santa Cruz web browser (http://genome.ucsc.edu/cgi-bin/hgBlat) [31,32] with the default settings. UniProtKB/Swiss-Prot Database (http://au.expasy.org) and GenBank (http://www.ncbi.nlm.nih.gov/Genbank/) sequences for rat ADH1, ADH2, ADH3, ADH4, ADH5 and ADH6A (Table 1) were used to interrogate the opossum genome sequence. Gene locations, predicted gene structures and ADH protein subunit sequences were observed for each ADH examined for those regions showing identity with the respective opossum *ADH* gene products.

2.2. Predicted secondary structures for opossum ADH gene products

Predicted secondary structures for rat ADH1, ADH2, ADH3, ADH4, ADH5 and ADH6A and for opossum ADH1, ADH3,

Table 1

Alcohol dehydrogenase (ADH) genes and enzymes examined.

| Animal | ADH gene | ADH gene lineage | GenBank mRNA (or *N-scan ID) | UNIPROT ID | No. of amino acids | Chromosome location | Strand | Exons | Subunit | Alternate Gene Name |
|---------|-------------|---------------------|---------------------------------|---------------|-----------------------|----------------------------|----------|-------|---------|------------------------|
| Human | ADH1A | ADH1 | BX647987 | P07327 | 375 | 4: 100,419,608-100,427,845 | Negative | 9 | α | ADH1 |
| | ADH1B | ADH1 | BC033009 | P00325 | 375 | 4: 100,450,947-100,459,066 | Negative | 9 | β | ADH2 |
| | ADH1C | ADH1 | BC066227 | P00326 | 375 | 4: 100,479,759-100,488,026 | Negative | 9 | γ | ADH3 |
| | ADH2 | ADH2 | BC002319 | P08319 | 380 | 4: 100,266,770-100,282,954 | Negative | 9 | П | ADH4 |
| | ADH3 | ADH3 | AK226177 | P11766 | 374 | 4: 100,212,749-100,225,390 | Negative | 9 | Х | ADH5 |
| | ADH4 | ADH4 | X76342 | P40394 | 374 | 4: 100,555,660-100,569,813 | Negative | 9 | σ/μ | ADH7 |
| | ADH5 | ADH5 | BX647987 | P28332 | 368 | 4: 100,345,104-100,356,445 | Negative | 9 | | ADH6 |
| Rat | ADH1 | ADH1 | BC062403 | P06757 | 376 | 2: 235,801,479-235,810,238 | Positive | 9 | А | |
| | ADH2 | ADH2 | BC127504 | Q64563 | 377 | 2: 235,953,574-235,969,179 | Positive | 9 | | ADH4 |
| | ADH3 | ADH3 | AY310136 | | 374 | 2:235,981,154-235,990,634 | Positive | 9 | Х | |
| | ADH4 | ADH4 | *chr2.1432 | P41682 | 374 | 2:235,753,870-235,762,009 | Positive | 9 | | ADH7 |
| | ADH5 | ADH5 | BC083782 | Q5XI95 | 376 | 2: 235,916,404-235,937,398 | Positive | 9 | | ADH6 |
| | ADH6A | ADH6 | XM215715 | | 375 | 2: 235,830,413-235,848,639 | Positive | 9 | | |
| | ADH6B | ADH6 | XM227745 | | 349 | 2: 235,868,290-235,885,810 | Positive | 9 | | |
| Mouse | ADH1 | ADH1 | BC13477 | P00329 | 375 | 3: 137,942,695–137,952,902 | Positive | 9 | А | |
| | ADH2 | ADH2 | BC100729 | Q9QYY9 | 369 | 3: 138,081,089-138,092,137 | Positive | 9 | | ADH4 |
| | ADH3 | ADH3 | AK146949 | P28474 | 374 | 3: 138,108,251-138,117,768 | Positive | 9 | В | |
| | ADH4 | ADH4 | AK9588 | Q64437 | 374 | 3: 137,884,677–137,891,879 | Positive | 9 | С | ADH7 |
| Horse | ADH1S | ADH1 | NM1081945 | P00328 | 374 | 3: 59,314,912-59,324,682 | Negative | 9 | S | |
| | ADH1E | ADH1 | NM1082528 | P00327 | 375 | 3: 59,311,838–59,327,348 | Negative | 9 | E | |
| Opossum | ADH1 | ADH1 | *chr5.6.012 | | 375 | 5: 51,904,684–51,919,994 | Positive | 9 | | |
| | ADH3 | ADH3 | *chr5.6.013 | | 374 | 5: 52,074,904-52,084,744 | Positive | 9 | | |
| | ADH4A | ADH4 | XP1369839 | | 374 | 5: 51,768,316-51,783,274 | Positive | 9 | | |
| | ADH4B | ADH4 | ŶР1369773 | | 374 | 5: 51,644,038-51,659,611 | Positive | 9 | | |
| | ADH4C | ADH4 | ŶP1369808 | | 376 | 5: 51,702,617-51,735,349 | Positive | 9 | | |
| | ADH6 | ADH6 | ŴW1581967 | | 375 | 5: 51,971,321-51,977,316 | Positive | 8 | | |
| | $ADH\Psi 1$ | ADH4 | | | 61 | 5: 51,990,956-51,991,138 | Positive | 1 | | |
| | $ADH\Psi 2$ | ADH3 | | | 147 | 5: 52,003,384-52,005,501 | Positive | 6 | | |
| Chicken | ADH1 | ADH1 | U73654 | | 375 | 4: 61,555,669-61,565,489 | Negative | 9 | | |
| | ADH3 | ADH3 | AJ720203 | | 374 | 4: 61,561,878-61,565,489 | Negative | 9 | | |
| Cod | ADH1 | ADH1 | | P26325 | 375 | | | | | |
| | ADH3L | ADH3 | | P81601 | 376 | | | | | |
| | ADH3H | ADH3 | | P81600 | 376 | | | | | |

GenBank mRNA (or cDNA) IDs identify previously reported sequences (see http://www.ncbi.nlm.nih.gov/Genbank/); *N-scan IDs identify gene predictions using the N-SCAN gene structure prediction software provided by the Computational Genomics Lab at Washington University in St. Louis, MO, USA (see http://genome.ucs.edu); UNIPROT refer to UniprotKB/Swiss-Prot IDs for individual ADHs (see http://kr.expasy.org). Sources for ADH sequences are provided by the above. Opossum ADH4A, and ADH4C protein sequences (XP136839; XP1369773; and XP1369808) were obtained from a blast using predicted opossum ADH4 sequences (derived from a blat of the opossum genome using the rat ADH4 sequence http://genome.ucsc.edu) and web tools of the National Center for Biotechnology Information (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

ADH4A, ADH4B, ADH4C and ADH6 were obtained using the PSIPRED v2.5 web site tools provided by Brunel University (http://bioinf.cs.ucl.ac.uk/psipred/psiform.html) [33].

2.3. Alignment of mammalian ADH active site residues

Alignments of a key ADH active site binding region (residues 112–134 for rat ADH1) were undertaken using a ClustalW-technique [34] (http://www.ebi.ac.uk/clustalw/) and previously reported sequences for human ADH1A, ADH1B, ADH1C, ADH2, ADH3, ADH4 and ADH5; horse ADH1E and ADH1S; mouse ADH1, ADH2, ADH3 and ADH4; rat ADH1, ADH2, ADH3, ADH4, ADH5

and ADH6A; and predicted opossum ADH1, ADH3, ADH4A, ADH4B, ADH4C and ADH6 sequences (Table 1).

2.4. Phylogenetic studies and sequence divergence

Phylogenetic trees were constructed using a ClustalW-derived amino acid alignment of ADH protein sequences, obtained with default settings and corrected for multiple substitutions [34] (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. Alignment ambiguous regions, including the amino and carboxyl termini, were excluded prior

| | ANP N O A I | м |
|-----|---|--------------------|
| 01 | STAGKVIKCKAAVLWELNKPFSIEEVEVAPPKANEVRIKIIATGIC SDD VVAGLLAY-PVPIILGHEAAGIVESVGEGVTSVKPGDKVIPLF | Г 94 |
| R1 | STAGKVIKCKAAVLWEPHKPFTIEDIEVAPPKAHEVRIKWVATGVC SDD AVSGSLFT-PLPAVLGHEGAGIVESIGEGVTCVKPGDKVIPLF | S 94 |
| 04A | STAGKVIKCKAAVLWGPKOPFSIEEVEVAPPKAYEVRIKIIATGICRTDO AIKGILLA-NFPVILGHEGAGIVESIGEGVTTVKPGDKVIPLC | L 94 |
| 04B | SSAGKVIKCKAAVLWGLKOPFSIEEVEVAPPKACEVRIKILATGICRTDE AISGAMTT-KFPVIVGHEATGVVESIGEGVTSVKPGDKVIPLF | L 94 |
| R4 | DTAGKVIKCKAAVLWGTNOPFSIEDIEVAPPKAKEVRVKILATGICGTDD VIKGTMVS-KFPVIVGHEAVGIVESVGEEVTTVRPGDKVIPLF | L 94 |
| 04C | NTSGKVIKCKAAVLWGLEOPFSIEETEVAPPKAHEVRIKTLATGICETDN ATTGVMPA-KEPVIVGHEAVGIVESIGEGVTSLKPGDKVIPLC | V 94 |
| 03 | - A GOVIKCKA AVAWEAGKPLSTEETEVA PPKAHEVRIK VIATAVCITDA TASGADDI GMNCLLISHDGAGIVESVGEGVTKIK AGDTVIDIY | т 93 |
| R3 | ANOVIE CKAAVAWEAGKPLSTEELEVAPPOAHEVELKTIATAVCUTDA / TLASGADPEGCEPVILGHEGAGIVESVGEGVTKLKAGDTVI PLY | т 93 |
| R2 | GTOGKVITCKAA JAWKTDSPLCIEEIEVSPEKAHEVEIKVIATCVCPTDI ATN-PKKKALEPVVIGHECAGIVESVGPGVTNEKPGDKVIPEE | A 94 |
| RS | GTOGKUTECKATVLWKPGAPLATEETEVAPPKAKEVETKWVATGVC_TDT.HLDTOELSKECPMIMGHEGVGIVESVGEGVSSVFTGDKVILLC | T 95 |
| R6A | DTLGKTTTCRAATAWARNSPLSTEEVOVEPPKSGEVETKMTSSGTCGSDD MLKGELLA-NEPLTPGHEGAGTVESVGDGVCSVKPGDKVLTLT | T 94 |
| 06 | ETTGOVITCKAAVTWTIDAPMSIEDVEVDPPKAGEVEIKTISSGICGSDN. VLEGDEKV-PLPVIIGHEGAGIVESIGEGVSSVKPGDKVLTVF | 6 94 |
| | .* *** * * *** **. ***** ** ** *****.* .* | |
| | | |
| 01 | POCROCTVCKHPVGNLCKG-NALNHRDVTLKEGTTRFTCRGKPTNHFLSTSTFSEYTVVDEISVVKLDSSAPLEKVCLVGCGFSTGYGSAV | K 185 |
| R1 | POCGKCRICKHPESNLCCOTKNLTOPKGALLDGTSRFSCRGKPTHHFISTSTFSOYTVUDLAVAKIDAAAPLDKVCLIGCGFSTGYGSAV | 186 |
| 04A | POCGKCSSC INPNGNECYKADITGRGVLSDGTSRFTCKGKPVYHFSSTSTFTEYTVVELAVTKIDASAPLEKVCLIG.GFSTGYGAAM | K 184 |
| 04B | POCGRC8SCINPNGNLCVKADVTGKGVLSDGTTRFTCKGKPVVHFMNTSTFSEVTVVDESSVTKIDANAPPEKVCLIGCGFSTGYGAAM | K 184 |
| R4 | POCRECNPC RNPEGNICIRSDITGRGVLADGTTRFTCKDKPV0HEMNTSTFTEYTVLDESSVAKIDAEAPPEKACLIGCGFSTGYGAAV | K 184 |
| 04C | POCGICSNCI.KPDSNYCDMI.DIVGKGVI.SDGTSRFTCKGKPV/HYMNTSTFTEVTVVRDVAVAKIDAAAPPEKACI.FGCAFTTGYGAAT | 184 |
| 03 | POCGECKECRNPKTNLCOKIETOGKGLMPDNTSRFTCKGKOIFHEMGTSTFSEVTVVADISVAKIDPLAPLDKVCLLGCGISTGYGAAT | 185 |
| R3 | POCGECKECLNPKTNLCOKTEVTOGKGLMPDGTSRFTCKGKPTLHEMGTSTFSEYTVVADISVAKIDPSAPLDKVCLLGCGISTGYGAAV | 185 |
| R2 | POCKKCKLCLSPLTNLCGKLENEKYPTIDOELMEDRISETCKGRSIVHEMGVSSFSOYTVVSEANLARVDDEANLERVCLIGCGFTSGYGAAT | 189 |
| R5 | POCGECKTCLNSKNNTCTEIRLSKTHLASEGTSRITCKGKLMHOYIALGSFSEVTULKEISVANTDEGAPLEKUCTIGCGFATGYGAAT | 185 |
| R6A | POCRECINGULILIKGNECEKODVI.PCSGUMI.DGTSRESCEGEKIVHSERTSSETEVITVUPETAVUKIDDAAPMIDKUCI.ISCGEPTGYGAAY | 185 |
| 06 | POCEKCOSCI, HAKGNCCI, KEDVEH PUGI, MI, DGTSEFTCEGERKI, HNA FGTSTFTEYTYMHET SVIKKI, DEA A PI, EKVCI, LA CGETTGYGSA T | 185 |
| 00 | | . 105 |
| | N N N N N | |
| 01 | VAKVTPGSTCVVFGLGGVGLSVVIGCKAAGASRIIGVDINKD, FAKAKEVGATECVNPLDYKKPIODVLIEMTDNTIDFSFEVIGRLDTVTAAL | 280 |
| R1 | VAK VTPGSTCAVFGLG VGLSVVIGCKTAGAAKIIAVDINKD FAKAKELGATDCINPODYTKPIOEVLOEMTDGGVDFSFEVIGRLDTMTSAL | 281 |
| 04A | TAKYTPGSTCVVFGLGGVGLSVIIGCKVAGATRIIGVDLNKD FEKAKAVGATECISPKDVTKPISEVLKEMTGDSVGYTFEAVGRLETMTDAL | A 279 |
| 04B | TAKYTPGSTCAVFGLGGVGLSVIMGCKSAGASRIIGIDLNKS FEKAKAVGATECISPKDYTKPISEVLSEMTDNSVGYTFEVVGRLETMIDAL | A 279 |
| R4 | TAK VSPGSTCAVFGLG VGLSVVMGCKA AGASRIIGIDINKD FOKALD VGATECINPRDFTKPISEVLSDMTGNTVOYTFEVIGRLETM VDAL | 5 279 |
| 04C | TAK VTPGSTCAVFGLG GVGLSVI IGCKI AGASRI IGVDIN PR FEKAKAVGATECVNPKDHTKPISEVLKEMTGDSVRYTFEVTGNLDTMIDAL | A 279 |
| 03 | TAKVEPGSTCAIFGLGUGLAVIMGCKVAGASRIIGVDINKD FAKAKEFGATECINPODFKKSIOEVLVEMTDGGVDFSFECIGNVGVMRAAL | E 280 |
| R3 | TAKVEPGSTCAVFGLGVGLAVIMGCKVAGASRIIGIDINKD FAKAKEFGATECINPODFSKSIOEVLIEMTDGGVDFSFECIGNVKVMRSAL | E 280 |
| R2 | TAKYTPGSACAVFGLGCVGLSAVIGCKIAGASRIIAIDIN <mark>SE FPKAKAL</mark> GATDCLNPRDLDKPVODVITELTGGGVDFSLDCAGTAOTLKAAV | 284 |
| R5 | SAKVTPGSTCAVFGLGGVGLSVIIGCKAAGAARIIAVDINKD FAKAKTVGATDCVDPRDFEKPIEEVLSDMIDGGVDFCFEVTGNTEAVGAAL | 3 280 |
| R6A | SAKUTPGSTCVVFGLG <mark>GVGSAIVMGCKA</mark> SGASRIIGVDINEO <mark>SFPRARA</mark> LGVTDCLNPKKLEKPVQEVVKEMTGVGVDFAFEAIGOVDTMAAAW | <mark>1</mark> 280 |
| 06 | KARWTPGSTCVVFGLG <mark>GVGSSVVLGCKA</mark> AGAARIIGIDINEE LARAKALGVTDCLNPRNFK <mark>KPIQQVVVEMT</mark> GFGADFSFEAIGTIDTMWAAL | <mark>E</mark> 280 |
| | .*:*.**:*::**::**::::***::::***:**:*::*: | |
| | Ex7 M O M MM N | |
| 01 | SCNDAYGVCVIVGVPPGSQ-TISIDPLLLLTGRTWKGAVFGGFKSKDDVPKIVSDVLSKKFNLDPLITHVYNFDKINEGFDLLRSGKSI TVLT | F 374 |
| R1 | SCHSACGVSVIVGVPPSAQ-SLSVN <mark>PMSLLL</mark> GRTWKGAIFGGFKSKDAVPKLVADFMAKKFPLEPLITHVLPFEKINEAFDLLRAGKSI TVLT | F 375 |
| 04A | SCHMSYGTSVIVGLPPSAT-MCTYDPMLLFTGRTWKGSTFGGWKSKDDLPKIVTDFLAKKFDFDELITHVLPFNEIEEGFNLLYKGESI aVLV | M 373 |
| 04B | SCHLSYGTSVVVGAPPSSK-MLTYDPMLLFTGRTWKGCVFGGWKSKDDVPKLVSDFLAKKFDLDQLITHVLHFKDINEGFELLKKGESI SVIL | M 373 |
| R4 | SCHMNYGTSVVVGAPPSAK-MLSYDPMLLFTGRTWKGCVFGGWKSRDDVPKLVTEFLEKKFDLGQLITHTLPFHNISEGFELLSQGKSI | F 373 |
| 04C | SCHKNCGMSVVVGDPPASS-VLTFDPMLVFDGRTWKGC1FGGWKPVNDLPKLVSDFMAKKFNLDELVTHILPFDKIEEGFNLLKKGESI TVLT | FQH |
| 03 | <mark>A</mark> CHKGWGVSVVVGVAASGQ-EISTR <mark>PFQLV</mark> TGRTWKGTAFG <u>G</u> WKS <mark>VESVPKLVSEYM</mark> SKKIKV <mark>DEF</mark> VTHNMP <mark>FDQINEAFELMH</mark> TGKS <u>I</u> SVLK | L 373 |
| R3 | <mark>A</mark> AHKGWGVSVVVGVAASGE-EISTR <mark>PFQLV</mark> TGRTWKGTAFG <u>G</u> WKS <mark>VESVPKLVSEYM</mark> SKKIKV <mark>DEF</mark> VTGNLS <mark>FDQINKAFDLMH</mark> SGNS <u>I</u> TVLK | 1 373 |
| R2 | CTVVGWGSCTVVGAKVDEM-NIST <mark>VDMIL</mark> GRSVKGTFFGGWKS <mark>VDSVPNLVTDYK</mark> NKKFDLDLLVTHALP <mark>FDKINDAIDLMN</mark> OGKSITILT | F 376 |
| R5 | <mark>S</mark> CHKDHGVCVTVGALASFTSTLSIR <mark>SHLFF</mark> SGRILKGSILG <u>G</u> WKT <mark>KEEIPKLVSDYM</mark> AKKFNI <mark>DPL</mark> ITHTLT <mark>LSEANEAVQLMK</mark> SGQC <u>I</u> CVLL | ն 375 |
| R6A | SCNHSYGVCLIVGLAPSDT-HLSLE <mark>ASKIL</mark> SGKTLKGVCLG <u>D</u> YK <mark>TRDCIPQIVTDYL</mark> QNKINI <mark>DPL</mark> VTHQLP <mark>FSQLHKALELYH</mark> SGKT <u>I</u> CVLL | F 374 |
| 06 | SCNSSYGVCVIIGVAPEKS-QLAFNPMQLLSGRTLKGCFLGDFKTRDHVPLLVDDYMKNKINLDPLITHRLPFLKVNEGFDLLRSGKSVCVIS | F 374 |
| | . * ::* *: ** :*.:* : *: ::*: : *: *: *: *: *: | |

Fig. 1. Amino acid sequence alignments for rat and opossum alcohol dehydrogenases (ADHs)

See Table 1 for sources of ADH sequences; O1-opossum ADH1; R1-rat ADH1; O4A-opossum ADH4A; O4B-opossum ADH4B; O4C-opossum ADH4C; R4-rat ADH4; O3-opossum ADH3; R3-rat ADH3; R2-rat ADH2; R5-rat ADH5; R6A-rat ADH6A; O6-opossum ADH6; * shows identical residues; 2 alternate residues; 3 alternate residues; bold font shows known or predicted exon junctions; predicted exon boundaries and exon numbers are shown as **Ex1**| **Ex2**| etc; predicted β -sheet (gray shade) and α -helix (yellow shade) secondary structures are shown. Key residue identification is based on previous 3D studies of human and horse homologues and likely predicted roles for amino acid residues; [24,37–44]: A-residues binding active site Zinc (blue); C-Cysteine residues binding structural Zinc; I-inner active site substrate binding residues; M-mid region active site substrate binding residues; P-Ser/Thr residue involved in reaction mechanism (brown); R-Arg binding of S-hydroxymethyl glutathione by ADH3 (red); D-Asp/Glu charge clamp residue for dimer formation; N-coenzyme binding (pink). Initiation methionine is not shown. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

 Table 2

 Percentage identities for rat, opossum, chicken and cod ADH amino acid sequences.

| ADH | rat1 | opo1 | chi1 | cod1 | rat2 | rat3 | opo3 | chi3 | cod3L | cod3H | rat4 | opo4A | opo4B | opo4C | rat5 | rat6A | opo6 |
|-------|------|------|------|------|------|------|------|------|-------|-------|------|-------|-------|-------|------|-------|------|
| rat1 | 100 | 74 | 71 | 56 | 55 | 63 | 63 | 66 | 59 | 60 | 64 | 68 | 66 | 64 | 57 | 54 | 59 |
| opo1 | 74 | 100 | 72 | 54 | 52 | 61 | 60 | 62 | 56 | 57 | 68 | 69 | 69 | 63 | 57 | 55 | 60 |
| chi1 | 71 | 72 | 100 | 56 | 54 | 63 | 63 | 67 | 62 | 63 | 69 | 68 | 68 | 65 | 58 | 55 | 60 |
| cod1 | 56 | 54 | 56 | 100 | 52 | 62 | 61 | 61 | 61 | 61 | 51 | 54 | 53 | 52 | 51 | 46 | 47 |
| rat2 | 55 | 52 | 54 | 52 | 100 | 56 | 55 | 56 | 53 | 56 | 51 | 51 | 52 | 53 | 50 | 49 | 48 |
| rat3 | 63 | 61 | 63 | 62 | 56 | 100 | 90 | 87 | 77 | 81 | 60 | 60 | 61 | 60 | 57 | 51 | 52 |
| opo3 | 63 | 60 | 63 | 61 | 55 | 90 | 100 | 85 | 75 | 81 | 58 | 59 | 60 | 59 | 57 | 51 | 50 |
| chi3 | 66 | 62 | 67 | 61 | 56 | 87 | 85 | 100 | 76 | 81 | 61 | 61 | 63 | 62 | 56 | 52 | 55 |
| cod3L | 59 | 56 | 62 | 61 | 53 | 77 | 75 | 76 | 100 | 81 | 54 | 55 | 56 | 56 | 52 | 48 | 50 |
| cod3H | 60 | 57 | 63 | 61 | 56 | 81 | 81 | 81 | 81 | 100 | 56 | 59 | 58 | 58 | 53 | 49 | 51 |
| rat4 | 64 | 68 | 69 | 51 | 51 | 60 | 58 | 61 | 54 | 56 | 100 | 74 | 78 | 70 | 57 | 52 | 53 |
| opo4A | 68 | 69 | 68 | 54 | 51 | 60 | 59 | 61 | 55 | 59 | 74 | 100 | 82 | 74 | 56 | 56 | 54 |
| opo4B | 66 | 69 | 68 | 53 | 52 | 61 | 60 | 63 | 56 | 58 | 78 | 82 | 100 | 74 | 56 | 51 | 53 |
| opo4C | 64 | 63 | 65 | 52 | 53 | 60 | 59 | 62 | 56 | 58 | 70 | 74 | 74 | 100 | 57 | 51 | 53 |
| rat5 | 57 | 57 | 58 | 51 | 50 | 57 | 57 | 56 | 52 | 53 | 57 | 56 | 56 | 57 | 100 | 52 | 51 |
| rat6A | 54 | 55 | 55 | 46 | 49 | 51 | 51 | 52 | 48 | 49 | 52 | 56 | 51 | 51 | 52 | 100 | 64 |
| opo6 | 59 | 60 | 60 | 47 | 48 | 52 | 50 | 55 | 50 | 51 | 53 | 54 | 53 | 53 | 51 | 64 | 100 |

Numbers show the percentage of amino acid sequence identities: rat1-rat ADH1; opo1-opossum ADH1; chi1-chicken ADH1; cod1-cod ADH1; rat2-rat ADH2; rat3-rat ADH3; opo3-opossum ADH3; chi3-chicken ADH3; cod3L-cod ADH3L; cod3H-cod ADH3L; rat4-rat ADH4; opo4A-opossum ADH4A; opo4B-opossum ADH4B; opo4C-opossum ADH4C; rat5-rat ADH5; rat6A-rat ADH6A; opo6-opossum ADH6. Numbers in bold show higher sequence identities for ADHs from the same class.

to phylogenetic analysis yielding alignments of 298 residues of human, rat, opossum, chicken and cod ADH sequences (Table 1). Pairwise scores were calculated using the number of identities in the best alignment divided by the number of residues compared. Scores were initially calculated as percent identity scores and were converted to distances by dividing by 100 and subtracting from 1.0 to give the number of differences per site. The extent of divergence for the rat ADH1, ADH2, ADH3, ADH4, ADH5 and ADH6A subunits with the opossum ADH1, ADH3, ADH4 and ADH6 subunits were determined using the SIM-Alignment tool for Protein Sequences (http://au.expasy.org/tools/sim-prot.html) [35,36].

3. Results

3.1. Alignments of predicted opossum ADH amino acid sequences with rat ADH sequences

The deduced amino acid sequences for six predicted opossum ADH subunits (designated as ADH1, ADH3, ADH4A, ADH4B, ADH4C and ADH6) are shown in Fig. 1 together with previously reported sequences for rat ADH1, ADH2, ADH3, ADH4, ADH5 and ADH6A (see Table 1). The ADH alignments showed high levels of sequence identities for the corresponding ADH family sequences for rat and opossum ADH1-, ADH3-, ADH4- and ADH6-like sequences, respectively (Table 2). This supports a proposal that they are products of the same ADH gene family in each case. In contrast, the sequences for rat and opossum ADHs from different gene families showed lower levels of identity (51-69%) (Table 2). Comparisons of predicted opossum ADH sequences with rat ADH sequences also enabled identification of key residues which may contribute to catalysis and function (Figs. 1 and 2). These included active site residues (rat ADH1 numbers used) which bind to the catalytic zinc and/or substrate (Cys46; Ser48; His67; Cys175); structural zinc binding cysteines (residues 97, 100, 103 and 111); and coenzyme binding residues, Arg47, His51; Asp224; Lys228; Arg272 and Arg370 (Table 3) (see [37–43]). ADH residues aligning with active site substrate binding residues were also identified: inner pocket Ser48 was retained for opossum ADH1 and ADH6 but substituted for opossum ADH3 and ADH4 sequences with Thr48; and inner pocket Phe93 was also retained for the opossum ADH1, ADH4B and ADH6 sequences but substituted for the opossum ADH3 (Tyr92), ADH4A and ADH4C (Cys93) sequences. Other previously identified ADH1 substrate binding and coenzyme binding residues underwent several substitutions, although variation within families was lower than sequence comparisons of different ADH families (Table 3). Opossum ADH3 retained key ADH class III residues, Asp57 and Arg115, which support S-hydroxymethyl glutathione binding, whereas opossum ADH1 and ADH4 sequences lacked these residues (Fig. 1) [41]. Rat ADH2 retained Arg115 but substituted Asp57 with Lys56; rat ADH5 retained an acidic residue (Glu57) for Asp57 of rat ADH3 as well as Arg115; whereas rat ADH6A and opossum ADH6 have substituted both of these residues. Consecutive glycine residues (Gly262 and Gly263) were observed for rat ADH1, ADH2, ADH3 and ADH5 and opossum ADH3 sequences, but were absent from rat ADH4 and ADH6A and from opossum ADH1, ADH4 and ADH6 sequences.

Fig. 3 illustrates the alignment of human, horse, mouse rat and opossum ADH sequences for a region previously shown to bind substrates at the active site [37,42]. With the exception of rat ADH1, which has an extra amino acid in this region (119Gln) [45], the mammalian ADH1, ADH3 and ADH6 sequences examined showed a consistent gap of 4 residues, in comparison with the human, rat and mouse ADH2 sequences. In contrast, human and rat ADH5 and mammalian ADH4 sequences exhibited a gap of five amino acid residues in this region.

3.2. Predicted secondary structures for opossum ADHs

Predicted secondary structures for opossum ADH1, ADH3, ADH4 and ADH6 subunits and for rat ADH1, ADH2, ADH3, ADH4, ADH5 and ADH6A subunits, were compared in Fig. 1. Similar α -helix β -



Fig. 2. Schematic representation of rat and opossum alcohol dehydrogenase (ADH) genes on chromosomes 2 and 5, respectively. Kb refers to kilobases of DNA. Size of ADH genes is presented in larger font as compared with the distance between ADH genes in smaller font.

Table 3

Predicted key alcohol dehydrogenase (ADH) amino acid residues for rat (r) ADH1, ADH3, ADH4, ADH5 and ADH6A and for opossum ADH1, ADH3, ADH4A, ADH4B, ADH4C and ADH6.

| Predicted function | rADH1 | oADH1 | rADH3 | oADH3 | rADH4 | oADH4A | oADH4B | oADH4C | rADH5 | rADH2 | rADH6A | oADH6 |
|-----------------------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| S binding inner region | Ser48 | Ser48 | Thr46 | Thr46 | Thr48 | Thr48 | Thr48 | Thr48 | Thr48 | Thr48 | Ser48 | Ser14 |
| | Phe93 | Phe93 | Tyr92 | Tyr92 | Phe93 | Cys93 | Phe93 | Cys93 | Cys93 | Cys93 | Ile93 | Phe59 |
| S binding middle region | Phe141 | Phe140 | Phe139 | Phe139 | Phe139 | Phe139 | Phe139 | Tyr139 | Tyr140 | Phe144 | Ser150 | Ala106 |
| | Ile142 | Leu141 | Met140 | Met140 | Met140 | Ser140 | Met140 | Met140 | Ile141 | Met145 | Phe151 | Phe107 |
| | Val295 | Val294 | Val293 | Val293 | Ala293 | Leu293 | Ala293 | Asp293 | Ala294 | Ala298 | Leu294 | Val260 |
| | Leu310 | Leu309 | Val308 | Val308 | Phe308 | Phe308 | Phe308 | Phe308 | Phe310 | Ile311 | Leu309 | Leu275 |
| | Ile319 | Val318 | Ala317 | Ala317 | Val317 | Thr317 | Val317 | Ile317 | Ile309 | Phe320 | Cys318 | Phe284 |
| | Phe320 | Phe319 | Phe318 | Phe318 | Phe318 | Phe318 | Phe318 | Phe318 | Leu310 | Phe321 | Leu319 | Leu285 |
| S binding outer region | Leu57 | Leu57 | Pro56 | Pro56 | Met57 | Leu57 | Met57 | Met57 | Leu58 | Lys56 | Leu57 | Phe23 |
| | Phe58 | Ala58 | Glu57 | Leu57 | Val58 | Leu58 | Thr58 | Pro58 | Ser59 | Lys57 | Leu58 | Lys24 |
| | Leu110 | Leu110 | Leu109 | Leu109 | Leu110 | Phe110 | Leu110 | Tyr110 | Ile111 | Leu110 | Phe110 | Cys76 |
| | Asn116 | Leu116 | Arg115 | Arg115 | Leu116 | Ile116 | Val116 | Ile116 | Leu117 | Asn116 | Val116 | Val82 |
| | Leu117 | Asn117 | Val116 | Ile116 | Thr117 | Thr117 | Thr117 | Val117 | Ser118 | Phe117 | Leu117 | Arg83 |
| | Ser298 | Gly297 | Ser296 | Ser296 | Ser296 | Ser296 | Ser296 | Ala296 | Ser297 | Asp301 | Ser297 | Glu263 |
| | Met307 | Leu306 | Phe305 | Phe305 | Phe305 | Phe305 | Phe305 | Phe305 | His307 | Val308 | Ser306 | Met273 |
| Coenzyme binding | Arg47 | Arg47 | His45 | His45 | Gly47 | Arg47 | Arg47 | Arg47 | Gly47 | Pro47 | Gly47 | Gly13 |
| | Ala231 | Ala230 | Ala229 | Ala229 | Gln229 | Glu229 | Glu229 | Glu229 | Ala230 | Pro234 | Pro230 | Ala196 |
| | Arg271 | Arg270 | Asn269 | Asn269 | Arg269 | Arg269 | Arg269 | Asn269 | Asn271 | Thr275 | Gln271 | Thr237 |
| | Arg364 | Arg363 | His362 | His362 | Tyr362 | Tyr362 | Ser362 | Lys362 | Lys364 | Asn365 | His363 | Arg329 |
| Charge clamp +ve* | Lys105 | Lys103 | Lys100 | Lys100 | Arg104 | Lys99 | Lys99 | Lys105 | Lys102 | Lys101 | Arg98 | Lys65 |
| Charge clamp –ve | Glu257 | Glu256 | Glu255 | Glu255 | Asp255 | Glu255 | Glu255 | Glu255 | Asp256 | Glu260 | Glu256 | Glu222 |
| Active site Zn binding | Cys46 | Cys46 | Cys44 | Cys44 | Cys46 | Cys12 |
| | His67 | His67 | His66 | His66 | His67 | His67 | His67 | His67 | His68 | His67 | His67 | His33 |
| | Cys175 | Cys174 | Cys174 | Cys174 | Cys173 | Cys173 | Cys173 | Cys173 | Cys174 | Cys173 | Cys174 | Cys140 |
| Structural Zn binding | Cys97 | Cys97 | Cys96 | Cys96 | Cys97 | Cys97 | Cys97 | Cys97 | Cys98 | Cys97 | Cys97 | Cys63 |
| | Cys100 | Cys100 | Cys99 | Cys99 | Cys100 | Cys100 | Cys100 | Cys100 | Cys101 | Cys100 | Cys100 | Cys66 |
| | Cys103 | Cys103 | Cys102 | Cys102 | Cys103 | Cys103 | Cys103 | Cys103 | Cys104 | Cys103 | Cys103 | Cys69 |
| | Cys111 | Cys111 | Cys110 | Cys110 | Cys111 | Cys111 | Cys111 | Cys111 | Cys112 | Cys111 | Cys111 | Cys77 |
| S-hydroxymethyl glutathione | Leu57 | Leu57 | Asp55 | Asp55 | Met57 | Leu57 | Met57 | Met57 | Glu57 | Lys56 | Leu57 | Phe23 |
| binding | Lys115 | Asn114 | Arg115 | Arg115 | Asp115 | Asp115 | Asp115 | Asp115 | Arg116 | Arg115 | Asp115 | Asp81 |

Identification of predicted key catalytic and structural amino acid residues is based on 3D structural studies from several sources [24,37–44]; S refers to substrate; inner, middle and outer region refers to substrate binding regions of the ADH active site. See Fig. 1 for the complete amino acid sequences for rat and opossum ADHs.

sheet structures were observed for all of the rat and opossum ADH subunits examined and comparable structures were predicted in each case, which resemble the secondary structures previously reported [43]. These included key regions for the enzyme such as the α -helix located near the four cysteine residues binding the structural zinc atom and the subunit–subunit binding site at His105 (rat ADH1); the α -helix and β -sheet structures on either side of Cys46 (binding the catalytic zinc atom) and Arg47 (charge relay transfer role during catalysis); and the two α -helices on either side of active site Cys175 (also binding the catalytic zinc atom). Opossum and rat ADH3 however lacked a predicted β -sheet structure prior to the Zinc binding His67 in comparison with the other ADHs examined.

3.3. Predicted gene locations and exonic structures for opossum ADH1, ADH3 and ADH4 genes

Table 1 and Fig. 3 summarize the predicted locations for opossum ADH1, ADH3, ADH6 and three ADH4-like genes based upon BLAT interrogation of the opossum genome [29], using the reported sequences for rat ADH1 [47], ADH2 [14], ADH3 [48], ADH4 [20], ADH5 and ADH6 [12] and the UC Santa Cruz Web Browser [32]. All of the predicted opossum *ADH* genes were located on chromosome 5 in a large (441 kbs) cluster of six genes. In addition, two ADH-like pseudogenes, designated as ADH Ψ 1 and ADH Ψ 2, were also observed in this region with the predicted gene order of *ADH4B-ADH4C-ADH4A-ADH1-ADH6-ADH\Psi1-ADH\Psi2-ADH3. This was compared with the predicted gene locations for rat ADH*

genes, which were also located on one chromosome (chromosome 2) and located within a 237 kb cluster, with the following gene order: ADH4-ADH1-ADH6A-ADH5-ADH2-ADH3. The opossum ADH genes were apparently transcribed on the positive DNA strand, as for the rat ADH genes. This is in contrast with human, horse and chicken ADH genes which are transcribed on the negative strand. Predicted exonic start sites for the opossum and rat ADH genes were examined with each gene having nine exons in each case, whereas the opossum $ADH\Psi1$ and $ADH\Psi2$ pseudogenes contained fewer exons (one and six exons, respectively). BLAT analyses using the rat ADH2 and ADH5 sequences for interrogation of the opossum genome did not reveal any unique opossum ADH2-and ADH5-like sequences.

3.4. Phylogeny and divergence of vertebrate ADH sequences

A phylogenetic tree (Fig. 4) was calculated by the progressive alignment of human, rat, opossum, chicken and cod ADH amino acid sequences which showed clustering into six main groups (ADH1–ADH6). The three human ADH1 sequences were grouped together on the ADH1 branch of the phylogram, as were the three opossum ADH4 sequences on the mammalian ADH4 branch, indicating that these *ADH* genes are products of recent gene duplication events of the respective ancestral *ADH1* and *ADH4* genes, respectively. Table 2 summarizes the percentages of sequence identities for the ADHs examined. Opossum ADH4 subunits shared a higher level of identity with each other (74% and 81%) and with rat ADH4 (70–78%) than with those from other ADH classes. In addi-

| ADH1A | HUMAN | CLKNDVSNPQGTLQDGTSRFTC | 133 |
|-------|---------|----------------------------|-----|
| ADH1B | HUMAN | CLKNDLGN PRGTLQDGTRRFTC | 133 |
| ADH1C | HUMAN | CLKNDLGN PRGTLQDGTRRFTC | 133 |
| ADH1E | HORSE | CLKNDLSMPRGTMQDGTSRFTC | 133 |
| ADH1S | HORSE | CLKN-LSMPRGTMQDGTSRFTC | 132 |
| ADH1 | MOUSE | CSRSDLLMPRGTLREGTSRFSC | 133 |
| ADH1 | RAT | CCQTKNLTQ PKGALLDGTSRFSC | 134 |
| ADH1 | OPOSSUM | CKGNALNHRDVTLKEGTTRFTC | 133 |
| ADH4 | HUMAN | CIRSDITGRGVLADGTTRFTC | 132 |
| ADH4 | MOUSE | CIRSDLTGCGVLADGTTRFTC | 132 |
| ADH4 | RAT | CIRSDLTGRGVLADGTTRFTC | 132 |
| ADH4A | OPOSSUM | CVKADITGRGVLSDGTSRFTC | 132 |
| ADH4B | OPOSSUM | CVKADVTGKGVLSDGTTRFTC | 132 |
| ADH4C | OPOSSUM | CDMLDIVGKGVLSDGTSRFTC | 132 |
| ADH3 | HUMAN | COKIRVTOGKGLMPDGTSRFTC | 132 |
| ADH3 | MOUSE | COKIRVTOGKGLMPDGTSRFTC | 132 |
| ADH3 | RAT | CQKIRVTQGKGLMPDGTSRFTC | 132 |
| ADH3 | OPOSSUM | CQKIRITQGKGLMPDNTSRFTC | 132 |
| ADH2 | HUMAN | CGKISNLKSPASDQQLMEDKTSRFTC | 138 |
| ADH2 | MOUSE | CGKLRNFKYPTIDQELMEDRTSRFTC | 127 |
| ADH2 | RAT | CGKLRNFKYPTIDQELMEDRTSRFTS | 137 |
| ADH5 | HUMAN | CIQFKQSKTQLMSDGTSRFTC | 133 |
| ADH5 | RAT | CTEIRLSKTHLASEGTSRITC | 133 |
| ADH6A | RAT | CEKQDVLPCSGVMLDGTSRFSC | 132 |
| ADH6 | OPOSSUM | CLKEDVRHPVGLMLDGTSRFTC | 98 |
| | | * :* *::* | |

Fig. 3. Amino acid sequence alignments for a substrate binding region of human, horse, mouse, rat and opossum ADH1; human, mouse and rat ADH2; human, mouse, rat and opossum ADH3; human, mouse, rat and opossum ADH4; human and rat ADH5; and rat ADH6A and opossum ADH6.

See Table 1 for sources of ADH sequences; * shows identical residues for rat ADH1, ADH2, ADH3, ADH4, ADH5 and ADH6A, and for opossum ADH1, ADH3, ADH4A, ADH4B, ADH4C and ADH6; 2 alternate residues observed.

tion, opossum ADH1 shared a higher level of sequence identity with rat (74%) and chicken ADH1 (72%); opossum ADH3 shared a very high sequence identity with rat (90%) and chicken ADH3 (85%); and opossum ADH6 shared more sequence identity with rat

ADH6A (64%) than with other ADH classes. The average amino acid sequence divergence rates for the mammalian ADH classes were also calculated using the average genetic distances observed for these ADHs and the dates for the common ancestors of eutherian and metatherian mammals and birds (Table 4). The results were indicative of a lower amino acid substitution rate for ADH3 (\sim 0.03% per million years of evolution) as compared with ADH4 (\sim 0.06%), ADH1 (0.06–0.11%), ADH6 (0.08–0.09%), ADH2 (0.14–0.16%) and ADH5, which showed the fastest amino acid substitution rate for mammalian ADHs of \sim 0.18% per million years of mammalian evolution.

4. Discussion

Several ADHs have been previously reported in the gray shorttailed opossum, *Monodelphis domestica*, with similar properties with other mammalian ADHs [30]. Three classes of opossum ADHs were described: class I or opossum liver ADH1, with preferential activity towards ethanol as substrate and showing sensitivity to pyrazole inhibition; class III ADH, being inactive towards ethanol as substrate and insensitive to pyrazole inhibition; and class IV ADH, requiring high concentrations of ethanol for activity and with a preference towards medium chain alcohol substrates [30]. The PAGE-IEF results also showed that opossum ADH4 exhibited several forms of activity which are located in various opossum tissues, including the cornea, stomach and esophagus.

The release of the opossum genome sequence [29] enabled this current study of opossum ADHs using BLAT techniques to interrogate the genome and to predict sequences for *ADH* genes and encoded ADH subunits. The results provide evidence for six *ADH* genes and encoded ADH proteins in the opossum, for which the amino acid sequences are consistent with classification into four ADH classes: class I (*ADH1*); class III (*ADH3*); class IV (3 *ADH4* genes: *ADH4A*, *ADH4B* and *ADH4C*); and class VI (*ADH6*) (Fig. 1). The respective *ADH* genes are predicted to be localized within a gene cluster



Fig. 4. Phylogenetic tree of selected vertebrate alcohol dehydrogenase (ADH) sequences. Each branch of the tree is labeled with the gene name followed by the species name. Predicted common ancestral genes are identified: 1A (human and rat ADH1); 1B (opossum, rat and human ADH1) and 1C (mammalian ADH1and chicken ADH1); 2A (human and rat ADH2); 3A (human and rat ADH3); 3B (opossum, human and rat ADH3); 3C (chicken and mammalian ADH3); 3D (cod, chicken and mammalian ADH3); 4A (human and rat ADH4); 4B (human and rat ADH4 and opossum ADH4B); 5A (human and rat ADH5); and 6B (opossum and rat ADH6). Note the clustering into six ADH groups (ADH1–6) and the likely sequence of gene duplication events: $ADH3 \rightarrow ADH1 \rightarrow ADH2 \rightarrow ADH5/ADH6 \rightarrow ADH4$ during vertebrate ADH gene evolution.

Table 4

Genetic distance and amino acid substitution rate predictions for human, rat, opossum, chicken and cod alcohol dehydrogenases (ADHs).

| ADH gene common ancestor | Common ancestor MY ago* | Genetic distance | Percentage substitution rate/MY |
|--------------------------|-------------------------|------------------|---------------------------------|
| 1A | 84-99 | 0.09 ± 0.01 | 0.1 |
| 1B | 173–193 | 0.12 ± 0.01 | 0.06-0.07 |
| 1C | 300-320 | 0.18 ± 0.01 | 0.06 |
| 2A | 84–99 | 0.14 ± 0.04 | 0.14-0.16 |
| 3A | 84–99 | 0.03 ± 0.001 | 0.03 |
| 3B | 173–193 | 0.05 ± 0.005 | 0.03 |
| 3C | 300-320 | 0.07 ± 0.01 | 0.02 |
| 3D | 500 | 0.1 ± 0.02 | 0.02 |
| 4A | 84–99 | 0.06 ± 0.01 | 0.06 |
| 4B | 173–193 | 0.09 ± 0.01 | 0.05 |
| 5A | 84–99 | 0.17 ± 0.01 | 0.18 |
| 6B | 173–193 | 0.15 ± 0.01 | 0.08-0.09 |

Common ancestors are identified in Fig. 4 and include 1A (human and rat ADH1); 1B (opossum, rat and human ADH1) and 1C (mammalian ADH1and chicken ADH1); 2A (human and rat ADH2); 3A (human and rat ADH3); 3B (opossum, human and rat ADH3); 3C (chicken and mammalian ADH3); 3D (cod, chicken and mammalian ADH3); 4A (human and rat ADH4); 4B (human, rat and opossum ADH4); 5A (human and rat ADH5); 6B (opossum and rat ADH6). Substitution rate is presented as a percentage of amino acid substitutions per million years. MY – million years ago. Dates for common ancestors were obtained from [48–50].

on chromosome 5 (Fig. 3), with each gene containing nine exons, in identical or similar positions to those previously reported (or predicted) for other mammalian *ADH* genes (Fig. 1). Three opossum *ADH4* genes were also described and designated as *ADH4A*, *ADH4B* and *ADH4C*, which were located in tandem with the opossum *ADH1*, *ADH3* and *ADH6* genes (Table 1; Fig. 3). In addition, two ADH-like pseudogenes were observed (ADH Ψ 1 and ADH Ψ 2), within the ADH gene cluster on chromosome 5 (Table 1; Fig. 2). BLAT analyses of the opossum genome using rat ADH2 and ADH5 sequences were unsuccessful in locating opossum *ADH2*- and *ADH5*-like genes which may be explained by gaps in the published opossum genome sequence in the corresponding regions for these genes or an absence of these ADH genes on the opossum genome.

Predicted amino acid sequences for the opossum ADHs showed high levels of identity in each case with the corresponding class of rat ADHs (Fig. 1; Table 2): ADH1 (74%), ADH3 (90%), ADH4 (74% or 82%) and ADH6 (64%), which is consistent with these ADH genes being classified within the same mammalian ADH gene family. This is supported by the phylogenetic analyses of mammalian ADH genes (Fig. 4) which showed that the opossum ADH sequences clustered with the corresponding ADH class of genes for human and rat. The phylogenetic studies also examined human, rat, opossum and chicken ADH sequences with those for cod ADH3 (ADH3L and ADH3H) and ADH1 sequences. The results supported previous studies which concluded that ADH3 represents the primordial vertebrate ADH gene [25,3], from which subsequent gene duplication events have generated several ADH gene classes during vertebrate evolution. A likely order for ADH gene evolution is suggested by this study as follows: $ADH3 \rightarrow ADH1 \rightarrow ADH2 \rightarrow ADH5/ADH6 \rightarrow ADH4$, with further gene duplication events occurring during primate evolution for ADH1 (generating ADH1A, ADH1B and ADH1C genes) [4,5], and during marsupial evolution (generating the ADH4C, ADH4A and ADH4B genes on the opossum genome).

Given the sequence homologies with those for rat ADHs of the same ADH class, and the retention of family specific kinetic properties and the key amino acid residues described earlier, it is proposed that opossum ADHs contribute to metabolic functions previously reported for the eutherian mammalian ADH classes. For opossum ADH3, this may include metabolic roles in formaldehyde and long chain fatty alcohol metabolism [16,41]; for opossum ADH1, a major role in the clearance of aliphatic and aromatic alcohols from the body [10]; and for the opossum ADH4 enzymes, major roles in the first pass clearance of ingested aliphatic and aromatic alcohols within the digestive system [10], in the metabolism of retinoid compounds in extrahepatic tissues [17,22], and in the metabolism of lipid peroxidation products in the eye and other tissues of the body [18,46,47].

In summary, BLAT analyses of the recently published opossum genome [29] have been undertaken using the amino acid sequences reported for rat ADHs for interrogation of the genome. Evidence is reported for at least six opossum ADH genes which were localized on chromosome 5 in a comparable ADH gene cluster to that observed for human and rat ADH genes. In addition, the predicted amino acid sequences and secondary structures for the opossum ADH subunits showed a high degree of similarity with the corresponding classes of mammalian ADHs, and four opossum ADH classes were identified, namely ADH1, ADH3, ADH6 and three forms of ADH4 (designated as ADH4A, ADH4B and ADH4C). This is supported by a previous biochemical analysis of opossum ADHs which examined the tissue distribution and properties for these enzymes, showing that ADH1 is the major liver enzyme; ADH3 is widely distributed in opossum tissues and has similar kinetic properties to mammalian class 3 ADHs; and with several forms of ADH4 localized in extrahepatic tissues, especially in the digestive system and in the eye. Phylogenetic analyses undertaken with opossum, human, rat, chicken and cod ADHs, supported the proposed designation of opossum ADHs as class I (ADH1), class III (ADH3), class IV (ADH4A, ADH4B and ADH4C) and class VI (ADH6) and their differential functions in the metabolism of ingested and endogenous alcohols and aldehydes in tissues of the opossum. In addition, percentage substitution rates were examined for ADHs during vertebrate evolution which indicated that ADH3 is evolving at a much slower rate to that of the other ADH classes.

Conflict of interest statement

None.

Acknowledgements

I am grateful to Dr Laura Cox and Dr John VandeBerg of the Southwest Foundation for Biomedical Research in San Antonio Texas for helpful discussions and advice; and to Dr Ricard Albalat of the University of Barcelona Spain and Dr Héctor Riveros-Rosas of the National Autonomous University of Mexico for sharing their results on opossum ADHs following the 14th International Meeting on Enzymology and Molecular Biology of Carbonyl Metabolism.

References

[1] G. Deuster, J. Farrés, M.R. Felder, R.S. Holmes, J.-O. Höög, J. Parés, B.V. Plapp, S.-J. Yin, H. Jörnvall, Recommended nomenclature for the vertebrate alcohol dehydrogenase gene family, Biochem. Pharmacol. 58 (1999) 389–395.

- [2] H. Jörnvall, J.-O. Höög, B. Persson, X. Parés, Pharmacogenetics of the alcohol dehydrogenase system, Pharmacology 61 (2000) 184–191.
- [3] R. Gonzălez-Duarte, R. Albalat, Merging protein, gene and genomic data: the evolution of the MDR-ADH family, Heredity 95 (2005) 184–197.
- [4] T. Ikuta, S. Szeto, A. Yoshida, Three human alcohol dehydrogenase subunits: cDNA structure and molecular and evolutionary divergence, Proc. Natl. Acad. Sci. U.S.A. 83 (1986) 634–638.
- [5] A.E.O. Trezise, E.A. Godfrey, R.S. Holmes, I.R. Beacham, Cloning and sequencing of cDNA encoding baboon liver alcohol dehydrogenase: evidence for a common ancestral lineage with the human alcohol dehydrogenase β subunit and for class I gene duplications predating primate radiation, Proc. Natl. Acad. Sci. U.S.A. 86 (1989) 5454–5458.
- [6] M.H. Foglio, G. Deuster, Characterization of the functional gene encoding mouse class III alcohol dehydrogenase (glutathione-dependent formaldehyde dehydrogenase) and an unexpressed processed pseudogene with an intact open reading frame, Eur. J. Biochem. 237 (1996) 496–504.
- [7] M. Smith, D.A. Hopkinson, H. Harris, Developmental changes and polymorphism in human alcohol dehydrogenase, Ann. Human Genet. 34 (1971) 257–271.
- [8] R.S. Holmes, Genetics and ontogeny of alcohol dehydrogenase isozymes in the mouse: evidence for a cis-acting regulator gene (Adt-1) controlling C2 isozyme expression in reproductive tissues and close linkage of Adh-3 and Adt-1 on chromosome 3, Biochem. Genet. 17 (1979) 461–472.
- [9] J.C. Burnell, W.F. Bosron, Genetic polymorphism of human liver alcohol dehydrogenase and kinetic properties of the isoenzymes, in: K.E. Crow, R.D. Batt (Eds.), Human Metabolism of Alcohol, vol. 2, CRC Press, Boca Raton, pp. 65–75.
- [10] S.-J. Yin, S.-L. Lee, C.-T. Yao, C.-L. Lai, Functional roles of alcohol dehydrogenases in human alcohol metabolism, in: H. Weiner, E. Maser, R. Lindahl, B. Plapp (Eds.), Enzymology and Molecular Biology of Carbonyl Metabolism-13., Purdue University Press, 2007, pp. 134–143.
- [11] D. Thierry-Mieg, J. Thierry-Mieg, AceView: a comprehensive cDNAsupported gene and transcripts annotation, Genome Biol. 7 (2006) S12, http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly.
- [12] J.-O. Höög, M. Brandt, J.J. Hedberg, P. Stromberg, Mammalian alcohol dehydrogenase of higher classes: analyses of human ADH5 and rat ADH6, Chem. Biol. Interact. 130–132 (2001) 395–404.
- [13] C.C. Ditlow, B. Holmquist, M.M. Morelock, B.L. Vallee, Physical and enzymatic properties of a class II alcohol dehydrogenase isozyme of human liver: Π-ADH, Biochemistry 23 (1984) 6363–6368.
- [14] J.-O. Höög, Cloning and characterization of a novel rat alcohol dehydrogenase of class II type, FEBS Lett. 368 (1995) 445–448.
- [15] J.-O. Höög, S. Svensson, Mammalian class II alcohol dehydrogenase. A highly variable enzyme, Adv. Exp. Med. Biol. 414 (1997) 303-311.
- [16] M. Koivusalo, M. Baumann, L. Uotila, Evidence for the identity of glutathionedependent formaldehyde dehydrogenase and class III alcohol dehydrogenase, FEBS Lett. 257 (1989) 105–111.
- [17] Z.N. Yang, G.J. Davies, T.D. Hurley, C.L. Stone, T.-K. Li, W.F. Bosron, Catalytic efficiency of human alcohol dehydrogenases for retinol oxidation and retinal reduction, Alcohol. Clin. Exp. Res. 18 (1994) 587–591.
- [18] P. Julià, X. Parés, H. Jörnvall, Rat liver alcohol dehydrogenase of class III. Primary structure, functional consequences and relationships to other alcohol dehydrogenases, Eur. J. Biochem. 172 (1988) 73–83.
- [19] H. Von Bahr-Linsdstrom, J.-O. H. Jörnvall, J.-O. Höög, Cloning and characterization of the human ADH4 gene, Gene 103 (1991) 269–274.
- [20] X. Parés, E. Cederlund, A. Moreno, L. Hjelmqvist, J. Farrés, H. Jörnvall, Mammalian class IV alcohol dehydrogenase (stomach alcohol dehydrogenase): structure, origin and correlation with enzymology, Proc. Natl. Acad. Sci. U.S.A. 91 (1994) 1893–1897.
- [21] C.-L. Han, C.-S. Liao, C.-W. Wu, C.-L. Hwong, A.-R. Lee, S.-J. Yin, Contribution to first-pass metabolism of ethanol and inhibition by ethanol for retinol oxidation in human alcohol dehydrogenase family: implications for etiology of fetal alcohol syndrome and alcohol-related diseases, Eur. J. Biochem. 254 (1998) 25–31.
- [22] G. Deuster, L. Deltour, H.L. Ang, Evidence that class IV alcohol dehydrogenase may function in embryonic retinoic acid synthesis, Adv. Exp. Med. Biol. 414 (1997) 357–364.
- [23] B. Vallee, T.J. Bazzone, Isozymes of human liver alcohol dehydrogenase, in: M. Ratazzi, J.G. Scandalios, G.S. Whitt (Eds.), Isozymes: Current Topics in Biological and Medical Research, vol. 8, Alan R. Liss, New York, 1983, pp. 219–224.
- [24] H.-W. Sun, B.V. Plapp, Progressive sequence alignment and molecular evolution of the Zn-containing alcohol dehydrogenase family, J. Mol. Evol. 34 (1992) 522–535.
- [25] H. Jörnvall, E. Nordling, B. Persson, Multiplicity of eukaryote ADH and other MDR forms, Chem. -Biol. Interact. (2003) 143–144.
- [26] J.L. VandeBerg, The gray short-tailed opossum (Monodelphis domestica) as a model didephid species for marsupial genetic research, Aust. J. Zool. 37 (1990) 235–246.

- [27] J.L. VandeBerg, S. Williams-Blangero, G.B. Hubbard, E.S. Robinson, Susceptibility to ultra-violet indiced corneal sarcomas is highly heritable in a laboratory opossum model, Int. J. Cancer 56 (1994) 119–123.
- [28] D.L. Rainwater, C.M. Kammerer, A.T.K. Singh, P.H. Moore, M. Pousesh, W.R. Shelledy, J.L. Vandeberg, E.S. Robinson, Genetic control of lipoprotein phenotypes in the laboratory opossum, Genescreen 1 (2001) 117–124.
- [29] T.S. Mikkelsen, M.J. Wakefield, B. Aken, C.T. Amemiya, J.L. Chang, S. Duke, M. Garber, A.J. Gentles, L. Goodstadt, A. Heger, J. Jurka, M. Kamal, E. Mauceli, S.M.J. Searle, T. Sharpe, M.L. Baker, M.A. Batzer, P.V. Benos, K. Belov, M. Clamp, A. Cook, J. Cuff, R. Das, L. Davidow, J.E. Deakin, M.J. Fazzari, J.L. Glass, M. Grabherr, J.M. Greally, W. Gu, T.A. Hore, G.A. Huttley, M. Kleber, R.L. Jirtle, E. Koin, J.T. Lee, S. Mahony, M.A. Marra, R.D. Miller, R.D. Nicholls, M. Oda, A.T. Papenfuss, Z.E. Parra, D.D. Pollock, D.A. Ray, J.E. Schein, T.P. Speed, K. Thompson, J.L. VandeBerg, C.M. Wade, J.A. Walker, P.D. Waters, C. Webber, J.R. Weidman, X. Xie, M.C. Zody, J.A. Marshall Graves, C.P. Ponting, M. Breen, P.B. Samollow, E.S. Lander, K. Lindblad-Toh'K, Genome of the marsupial *Monodelphis domestica* reveals innovation in non-coding sequences, Nature 447 (2007) 167–177.
- [30] R.S. Holmes, R.A.H. van Oorschot, J.L. Vandeberg, Biochemical genetics of alcohol dehydrogenase isozymes in the gray short-tailed opossum, Biochem. Genet. 30 (1992) 215–231.
- [31] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search tool, J. Mol. Biol. 215 (1990) 403–410.
- [32] W.J. Kent, C.W. Sugnet, T.S. Furey, K.M. Roskin, T.H. Pringle, A.M. Zahler, D. Haussler, The human genome browser at UCSC, Genome Res. 12 (2002) 994–1006.
- [33] LJ. McGuffin, K. Bryson, D.T. Jones, The PSIPRED protein structure prediction server, Bioinformatics 16 (2000) 404–405.
- [34] R. Chenna, H. Sugawara, T. Koike, R. Lopez, T.J. Gibson, D.G. Higgins, J.D. Thompson, Multiple sequence alignment with the Clustal series of programs, Nucl. Acids Res. 31 (2003) 3497–3500.
- [35] M.C. Pietsch, Internet-based tools for automated comparative protein modeling, Biochem. Soc. Trans. 24 (1996) 274–279.
- [36] T. Schwede, J. Kopp, N. Guex, M.C. Peitsch, SWISS-MODEL: an automated protein homology-modelling server, Nucl. Acids Res. 31 (2003) 3381–3385.
- [37] H. Eklund, B.V. Plapp, J.-P. Samama, C.-I. Bränden, Binding of substrate in a ternary complex of horse liver alcohol dehydrogenase, J. Biol. Chem. 257 (1982) 14349–114358.
- [38] H. Eklund, P. Müller-Wille, E. Horjales, O. Futer, B. Holmquist, B.L. Vallee, Comparison of three classes of human liver alcohol dehydrogenase and related enzymes, Eur. J. Biochem. 193 (1990) 303–310.
- [39] T.D. Hurley, C.G. Steinmetz, P. Xie, Z.-N. Yang, Three-dimensional structures of human alcohol dehydrogenase isoenzymes reveal the molecular basis for their functional diversity, Adv. Exp. Med. Biol. 414 (1997) 291–302.
- [40] P.G. Xie, S.H. Parsons, D.C. Speckhard, W.F. Bosron, T.D. Hurley, X-ray structure of human class IV σσ alcohol dehydrogenase-structural basis of substrate specificity, J. Biol. Chem. 272 (1997) 18558–18563.
- [41] Z.N. Yang, W.F. Bosron, T.D. Hurley, Structure of human chi chi alcohol dehydrogenase: a glutathione-dependent formaldehyde dehydrogenase, J. Mol. Biol. 265 (1997) 330–343.
- [42] M.S. Niederhut, B.J. Gibbons, S. Merez-Miller, T.D. Hurley, Three-dimensional structures of the three human class I alcohol dehydrogenases, Prot. Sci. 10 (2001) 697–706.
- [43] S. Ramaswamy, H. Eklund, B.V. Plapp, Structures of horse liver alcohol dehydrogenase complexed with NAD+ and substituted benzyl alcohols, Biochemistry 33 (1994) 5230–5237.
- [44] H. Eklund, B. Nordström, E. Zeppezauer, G. Söderlund, I. Ohlsson, T. Boiwe, B.O. Söderberg, O. Tapia, C.I. Brändén, A. Akeson, Three-dimensional structure of horse liver alcohol dehydrogenase at 2.4 Å resolution, J. Mol. Biol. 102 (1976) 27–59.
- [45] D.W. Crabbe, H.J. Edenberg, Complete amino acid sequence of rat liver alcohol dehydrogenase deduced from the cDNA sequence, Gene 48 (1986) 287–291.
- [46] J.E. Downes, J.L. VandeBerg, R.S. Holmes, Regional distribution of mammalian aldehyde dehydrogenase and alcohol dehydrogenase, Cornea 11 (1992) 560–566.
- [47] P. Julià, X. Parés, X.H. Jörnvall, Rat liver alcohol dehydrogenase of class III. Primary structure, functional consequences and relationships to other alcohol dehydrogenases, Eur. J. Biochem. 172 (1988) 73–83.
- [48] S. Kumar, S.B. Hedges, A molecular timescale for vertebrate evolution, Nature 392 (1998) 917–920.
- [49] M.O. Woodburne, T.H. Rich, M.S. Springer, The evolution of tribospheny and the antiquity of mammalian clades, Mol. Phylogenet. Evol. 28 (2003) 360–385.
- [50] M.A. Nilsson, U. Arnason, P.B. Spencer, A. Janke, Marsupial relationships and a timeline for marsupial radiation in South Gondwana, Gene 340 (2004) 1189–1196.