Available online at www.sciencedirect.com
ScienceDirect
Developmental Biology 309 (2007) 18–31
DEVELOPMENTAL BIOLOGY
www.elsevier.com/developmentalbiology

Three male germline-specific aldolase A isozymes are generated by alternative splicing and retrotransposition

Soumya A. Vemuganti [a,b], Timothy A. Bell [a,c], Cameron O. Scarlett [f,1], Carol E. Parker [f], Fernando Pardo-Manuel de Villena [a,c,d,e], Deborah A. O'Brien [a,b,d,*]

[a] *Laboratories for Reproductive Biology, University of North Carolina School of Medicine, Chapel Hill, NC 27599, USA*
[b] *Department of Cell and Developmental Biology, University of North Carolina School of Medicine, Chapel Hill, NC 27599, USA*
[c] *Department of Genetics, University of North Carolina School of Medicine, Chapel Hill, NC 27599, USA*
[d] *Lineberger Comprehensive Cancer Center, University of North Carolina School of Medicine, Chapel Hill, NC 27599, USA*
[e] *Carolina Center for Genome Sciences, University of North Carolina School of Medicine, Chapel Hill, NC 27599, USA*
[f] *UNC-Duke Michael Hooker Proteomics Center, University of North Carolina School of Medicine, Chapel Hill, NC 27599, USA*

Received for publication 27 April 2007; revised 9 June 2007; accepted 12 June 2007
Available online 18 June 2007

Abstract

Enzymes in the glycolytic pathway of mammalian sperm are modified extensively and are localized in the flagellum, where several are tightly bound to the fibrous sheath. This study provides the first evidence for three novel aldolase isozymes in mouse sperm, two encoded by *Aldoart1* and *Aldoart2* retrogenes on different chromosomes and another by *Aldoa_v2*, a splice variant of *Aldoa*. Phylogenetic analyses and comparative genomics indicate that the retrogenes and splice variant have remained functional and have been under positive selection for millions of years. Their expression is restricted to the male germline and is tightly regulated at both transcriptional and translational levels. All three isozymes are present only in spermatids and sperm and have distinctive features that may be important for localization in the flagellum and/or altered metabolic regulation. Both ALDOART1 and ALDOA_V2 have unusual N-terminal extensions not found in other aldolases. The N-terminal extension of ALDOA_V2 is highly conserved in mammals, suggesting a conserved function in sperm. We hypothesize that the N-terminal extensions are responsible for localizing components of the glycolytic pathway to the fibrous sheath and that this localization is required to provide sufficient ATP along the length of the flagellum to support sperm motility.

Keywords: Spermatogenesis; Sperm; Glycolysis; Aldolase isozymes; Retrogene; Retrotransposition; Alternative splicing; Fibrous sheath; LC-MALDI-MS

Introduction

Mammalian spermatogenesis includes three phases: a mitotic proliferation period that expands the number of spermatogonia, a prolonged meiotic prophase allowing spermatocytes to undergo recombination followed by two meiotic divisions, and a post-meiotic period where haploid spermatids differentiate into highly polarized sperm that are specialized to achieve fertilization. The program of gene expression that directs this developmental process has several distinct features and produces a large number of transcripts that are restricted to spermatogenic cells (DeJong, 2006; Eddy, 2002; Elliott and Grellscheid, 2006). Microarray analyses estimate that at least 4% of the mouse genome produces male germ cell-specific transcripts, predominantly during the meiotic and post-meiotic phases of spermatogenesis (Pang et al., 2006; Schultz et al., 2003; Shima et al., 2004). Essential processes that occur during these two phases include the generation and subsequent packaging of the haploid genome within an extremely condensed nucleus, formation of specialized sperm structures such as the acrosome and flagellum, organization of surface domains essential for fertilization, and the development of complex signaling and metabolic cascades

* Corresponding author. Department of Cell and Developmental Biology, University of North Carolina School of Medicine, Chapel Hill, NC 27599, USA. Fax: +1 919 966 1856.
E-mail address: dao@med.unc.edu (D.A. O'Brien).
[1] Present address: Analytical Instrumentation Center, School of Pharmacy, University of Wisconsin, 777 Highland Avenue, Madison, Wisconsin 53705, USA.

0012-1606/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.ydbio.2007.06.010

that regulate sperm function and gamete interaction. Multiple mechanisms contribute to this extensive diversification of gene function including gene duplication, retrotransposition and alternative splicing. Many conserved pathways, as well as germ cell-specific structures and processes, have unique variants with restricted expression during spermatogenesis.

Glycolysis is an important conserved pathway that has been modified substantially during mammalian spermatogenesis. This central metabolic pathway is composed of ten enzymes that convert glucose to pyruvate, with a net production of 2 ATPs per glucose molecule. Pyruvate can be further metabolized in the mitochondria via the Krebs cycle and oxidative phosphorylation, or alternatively converted to lactate by lactate dehydrogenase (LDH). Eight glycolytic enzymes and LDH have multiple isoforms with distinct patterns of expression in different tissues (Steinke et al., 2006). At least three of these gene families include members with restricted expression in germ cells. Two germ cell-specific isozymes with distinct enzymatic properties, glyceraldehyde 3-phosphate dehydrogenase-S (GAPDHS) and lactate dehydrogenase C (LDHC), are encoded by intron-containing genes that arose by gene duplication. In the mouse *Gapdhs* is expressed only in spermatids (Bunch et al., 1998; Welch et al., 1992), while *Ldhc* is expressed in spermatogenic cells and oocytes (Coonrod et al., 2006; Goldberg, 1977; Li et al., 1989). Phosphoglycerate kinase-2 (*Pgk2*), which encodes another germ cell-specific isozyme, is an intronless gene that evolved by retrotransposition (i.e., a retrogene) from the *Pgk1* gene (Boer et al., 1987; McCarrey and Thomas, 1987). Although *Pgk2* is transcribed in primary spermatocytes, the PGK2 protein is translated only in post-meiotic spermatids (Bluthmann et al., 1982; Vandeberg et al., 1981).

Other glycolytic enzymes, including hexokinase 1 (Mori et al., 1993, 1998), phosphoglucose isomerase (Buehr and McLaren, 1981), phosphofructokinase 1 (Yamada et al., 2004), aldolase (Gillis and Tamblyn, 1984) and enolase (Edwards and Grootegoed, 1983; Gitlits et al., 2000), have unique structural or functional properties in spermatogenic cells and sperm. The molecular basis for these distinctive properties has not been determined in most cases. One exception is HK1-S, the hexokinase 1 isozyme that is derived from alternative splicing and lacks the porin-binding domain responsible for binding to mitochondria (Mori et al., 1993, 1998).

Energy production in sperm is compartmentalized in distinct regions of the flagellum, with mitochondria and oxidative phosphorylation restricted to the middle piece and glycolysis localized in the longest segment known as the principal piece (Eddy et al., 2003). Although there are differences between species, mammalian sperm typically exhibit a high rate of glycolysis that is correlated with motility (Hoskins, 1973; Mann and Lutwak-Mann, 1981). Multiple studies indicate that glycolysis provides a significant proportion of ATP in both mouse and human sperm (Mukai and Okuno, 2004; Peterson and Freund, 1969; Williams and Ford, 2001). Gene targeting studies provide compelling evidence that glycolysis in spermatozoa (Miki et al., 2004) rather than mitochondrial ATP production (Narisawa et al., 2002) is essential for maintaining sperm motility and male fertility in the mouse. Distinctive features of sperm glycolytic enzymes may be important for localization in the principal piece and/or altered regulation of this key metabolic pathway.

Several sperm glycolytic enzymes are difficult to solubilize because they are tightly bound to the fibrous sheath, a cytoskeletal structure that defines the limits of the principal piece in the sperm flagellum. We found that GAPDHS, aldolase A (ALDOA), lactate dehydrogenase A (LDHA), and pyruvate kinase remain attached to the fibrous sheath throughout a rigorous isolation procedure (Bunch et al., 1998; Krisfalusi et al., 2006). GAPDHS is larger than other GAPDH family members and contains a novel proline-rich extension at the N-terminus (Welch et al., 1992, 2000) that may mediate binding to the fibrous sheath. Our proteomic and immunoblot analyses identified two ALDOA bands in mouse sperm, with the larger 50,000 molecular weight band always present in isolated fibrous sheaths (Krisfalusi et al., 2006). ALDOA, along with several other glycolytic enzymes, was also identified in fibrous sheaths isolated from human sperm (Kim et al., 2006). Consistent with this localization, an earlier study found that 90% of aldolase activity in bovine sperm could not be solubilized with detergents (Gillis and Tamblyn, 1984). That study also determined that sperm and muscle aldolases had distinct kinetic properties.

The fructose-1,6-bisphosphate aldolase gene family in vertebrates contains three well-characterized members: *Aldoa*, which is ubiquitously expressed with particularly high levels in the muscle, aldolase B (*Aldob*) which is expressed at high levels in the liver and kidney, and aldolase C (*Aldoc*) which is predominantly expressed in the nervous system (Penhoet et al., 1966, 1969; Rutter et al., 1968). All three of these isozymes have molecular weights of ~39,000, significantly smaller than the isoform detected in fibrous sheath. To determine the identity and origin of this larger ALDOA isoform, we used genomic, molecular and proteomic methods to examine all the aldolase variants expressed during mouse spermatogenesis and present in mature sperm. Our analyses identified three spermatogenic cell-specific aldolase isoforms in mouse, two encoded by retrogenes and a third resulting from alternative splicing of the *Aldoa* gene.

Materials and methods

Genomic analyses to identify Aldoa-related sequences

Ensembl Mouse BlastView (http://www.ensembl.org/Mus_musculus/blastview) was used to search the mouse genome (NCBI database, mouse build 34) for sequences highly related to the mRNA sequence of mouse *Aldoa* (accession number NM_007438). Chromosomal regions containing significant matches were aligned to the mRNA sequence of *Aldoa* using ClustalW (http://www.ebi.ac.uk/clustalw/). All chromosome numbers included refer to Assembly: NCBIm36, December 2005; Genebuild: Ensembl, April 2006; Database version: 42.36c. For each sequence, we identified all insertions and deletions resulting in a shift in open reading frame (ORF) and internal stop codons. We assume that sequences containing such insertions/deletions or stop codons are pseudogenes that do not encode functional aldolase enzymes. Intronless sequences that contain a full-length ORF with conserved start and stop codons were classified as putative retrogenes. These included sequences found on chromosome 4 and chromosome 12, now termed *Aldoart1* and *Aldoart2*. The Mouse Genome Informatics database (http://www.informatics.jax.org/) previously identified both of these regions as *Aldoa* pseudogenes. *Aldoart1* contains an extended ORF capable of encoding a protein with an additional 55 amino acids at the N-terminus. Further genomic comparisons identified a homologous

potential coding sequence within the first intron of the *Aldoa* gene on chromosome 7, indicating the existence of an additional *Aldoa* exon. Transcripts that include this newly identified exon 2 would result in a splice variant (*Aldoa_v2*) that also encodes a larger protein with a novel N-terminus.

Phylogenetic analyses

Phylogenetic analyses were performed with two sets of sequences. The first set comprises nucleotide sequences for the ORFs of mouse *Aldoa*, *Aldob*, *Aldoc*, *Aldoart1*, *Aldoart2* and their orthologs in rat, human and chimpanzee. All these sequences were obtained from Ensembl. To identify orthologous sequences of the predicted mouse retrogenes, we used the *Aldoart1* and *Aldoart2* sequence as queries in BLAST searches in each of these additional species. Putative matches were then tested for conservation of synteny with sequences flanking the mouse retrogenes and for presence of a conserved ORF. The second set of sequences includes the 165 nucleotides of mouse *Aldoart1* encoding 55 additional amino acids at the N-terminus, the 162 nucleotides of *Aldoa_v2* that encode a similar 54-amino-acid sequence at the N-terminus, and the homologous sequences in rat, human, chimpanzee and macaque identified through BLAST searches in the appropriate species. Each set of sequences was aligned separately using ClustalW. All phylogenetic analyses were performed using the PHYLIP phylogeny inference software package, version 3.6 (http://evolution.genetics.washington.edu/phylip.html). For each set of sequences, we generated 100 bootstrapped datasets using the Seqboot program. We then determined the phylogeny of the aldolase gene family with a distance matrix method (Neighbor) using distances obtained with the Dnadist program, a maximum likelihood method without a molecular clock (Dnaml), and a maximum parsimony method (Dnapars). The Consense program was used to construct a majority-rule consensus tree.

EST database searches

Sequences for the full-length *Aldoa* mRNA, individual exons of this gene including the proposed alternative exon 2, *Aldoart1* and *Aldoart2* were used as queries in BLAST searches against the mouse EST database, using the NCBI nucleotide–nucleotide BLAST (blastn) software. The EST database includes libraries from several cell types isolated from testis including type A spermatogonia, type B spermatogonia, pachytene spermatocytes, round spermatids and Sertoli cells (McCarrey et al., 1999). EST libraries are also available from cell lines produced from Leydig cells. To query specific EST libraries, searches were completed using either tissue type or testicular cell type as a limiting field (i.e., muscle, brain, testis or testicle, spermatogonia, spermatocyte and spermatid). Positive hits from libraries derived from pooled tissues were excluded. ESTs with very high levels of identity for substantial contiguous sequences (>98% identity for >100 bp) were scored as positive. Positive ESTs were then blasted against the mouse genomic database using Ensembl Mouse BlastView, and the best genomic sequence match (with at least 98% identity) was used to identify the gene from which the EST was transcribed.

Tissue and cell isolations

Outbred CD-1 mice and Sprague–Dawley rats were obtained from Charles River (Raleigh, NC) and inbred strains (C57BL/6J and A/J) were obtained from the Jackson Laboratory (Bar Harbor, ME). All procedures involving animals were approved by the University of North Carolina at Chapel Hill Animal Care and Use Committee and conducted in accordance with the Guide for the Care and Use of Laboratory Animals (Institute for Laboratory Animal Research, National Academy of Sciences). Testes and other tissues were processed immediately or quick frozen in liquid nitrogen and stored at −70 °C. Mixed germ cells were isolated by sequential enzymatic dissociation of testes from 8-day-old, 17-day-old, or adult mice (O'Brien, 1993). Spermatogonia (92% purity, including type A through type B spermatogonia) were isolated from mixed germ cells from 8-day-old mice by differential plating (Tsuruta et al., 2000). Sertoli cells were isolated from 17-day-old mice as described previously (O'Brien et al., 1989). Pachytene spermatocytes, round spermatids, and condensing spermatids were purified from adult mixed germ cell suspensions by unit gravity sedimentation (O'Brien, 1993). Both pachytene spermatocytes and round spermatids (steps 1–8) had purities >90%. Condensing spermatids contained 30–40% nucleated cells (steps 9–16 spermatids) and cytoplasts derived primarily from these same cells.

Mouse and rat sperm were collected from the cauda epididymis following a 15-min incubation of clipped tissue at 37 °C in phosphate-buffered saline with protease inhibitors (PBS+PI) containing 140 mM NaCl, 10 mM phosphate buffer (pH 7.4) and Complete protease inhibitor cocktail (Roche Diagnostics, Mannheim, Germany). Cryopreserved human sperm samples from healthy donors were provided by the Andrology Laboratory, Department of Obstetrics and Gynecology, University of North Carolina School of Medicine. These samples were washed twice with PBS to remove seminal plasma.

RT-PCR to detect transcription of Aldoa-related genes

Total RNA was isolated using Trizol (Invitrogen, Carlsbad, CA) from tissues or cells pooled from at least three mice, including the brain, heart, ovary, kidney, spleen, liver, juvenile (8–34 days of age) and adult testis, and purified testicular cells. Genomic DNA contamination was removed from RNA preparations using the Qiagen RNeasy Midi Kit (Qiagen Incorporation, Valencia, CA). RNA was quantified using the NanoDrop spectrophotometer (NanoDrop Technologies, Wilmington, DE). Equal loading and quality were confirmed using the RNA 6000 Nano Assay (Agilent Technologies, Waldbronn, Germany).

Reverse transcription followed by gene-specific polymerase chain reaction (cMaster RT_{plus}PCR System, Eppendorf, Hamburg, Germany) was used to amplify transcripts from total RNA samples. Primers were designed to amplify and incorporate α-[^{32}P]-dCTP into the same size product (530 bp) from *Aldoa* (including both the ubiquitously expressed *Aldoa_v1* transcript and the predicted *Aldoa_v2* transcript), *Aldoart1* and *Aldoart2*. The forward primer sequence was 5′AGAAGGCAGATGTGGACGTCC3′ and the reverse primer sequence was 5′ GGCACTGTGCGACGAAGTGCTGTGAC3′. Both primers are 100% identical to matching sequences in *Aldoa*, *Aldoa_v2*, *Aldoart1*, and *Aldoart2*. Arrows in Fig. 5A denote the location of these primers in each gene. Following PCR amplification with both primer sets, the products were digested with the *Hae*III restriction enzyme (New England BioLabs, Ipswich, MA). Unique *Hae*III restriction sites in each amplified product produced fragments with distinct sizes that were resolved by electrophoresis on 5% acrylamide gels at 95 W for 2 h. Appropriate controls (a cDNA clone for *Aldoa*, and BAC clones for *Aldoart1* [RP24-191C1] (Invitrogen, Carlsbad, CA) and *Aldoart2* [RP23-63H11]) were used as template in parallel PCR reactions to confirm the expected digestion pattern. Gels were exposed to Super RX X-ray film (Fujifilm, Tokyo, Japan) using intensifying screens to detect incorporation of α-[^{32}P]-dCTP into amplified products.

For semi-quantitative analysis of RT-PCR products, we first determined the linear phase for PCR amplification by varying the amount of input RNA and the number of amplification cycles. Once conditions for the linear phase of amplification were established, RT-PCR was repeated as described above with RNAs (75 ng/reaction) from isolated testicular cells (pachytene spermatocytes, round spermatids, condensing spermatids and Sertoli cells) using triplicate reactions from two preparations of each cell type. For each reaction, incorporation of α-[^{32}P]-dCTP into *Hae*III-digested products was measured by densitometry using the Fluor-S MultiImager (Bio-Rad, Hercules, CA). Corrections were made for varying amounts of α-[^{32}P]-dCTP incorporated into each gene product due to nucleotide differences and for *Hae*III-digested PCR fragments that were outside the resolution range of our gels.

A second set of primers (arrows in Fig. 6A) was used to amplify and incorporate α-[^{32}P]-dCTP into products from the N-terminal regions of *Aldoa* and *Aldoart1*. The forward primer sequence for this reaction was 5′CCGCG-TTCGCTCCTTAGT3′ and the reverse primer sequence was 5′ATCTGCAGC-CAGGATGCCC3′. Both primers are 100% identical to the matching sequences in the *Aldoa_v1* and *Aldoa_v2* transcripts and will amplify different sized RT-PCR products from each transcript. However, the forward primer contained one mismatched residue when amplifying *Aldoart1* (5′CCG/ ACGTTCGCTCCTTAGT3′). Products from *Aldoa_v2* and *Aldoart1* can be distinguished following *Hae*III digestion, since only the *Aldoa_v2* product is cleaved by this enzyme. [^{32}P]-labeled RT-PCR products were resolved on 5% polyacrylamide gels and detected by exposure to film using intensifying screens.

Reverse transcription of RNA from CD-1 testis using an oligo d(T) primer generated a testis cDNA sample used for full-length subcloning of *Aldoa*,

Aldoa_v2, *Aldoart1*, and *Aldoart2*. All four *aldolase A* genes were PCR amplified by gene-specific primer sets. The following primer pairs were using for PCR amplification of each *Aldoa*-related transcript: *Aldoa* (forward primer: 5′ACGAGGTTCTGGTGACCCTA3′, reverse primer: 5′GTGATGGGAAAGAGCCTGAA3′), *Aldoa_v2* (forward primer: 5′GAATTCATGGCAACGCACAG3′, reverse primer: 5′CTCGAGTTCAATAGCAAGTGG3′), *Aldoart1* (forward primer: 5′CGGAATTCATGGCAACGCACAGGCA3′, reverse primer: 5′AGCGTCGACACATGAGGGCA3′), and *Aldoart2* (forward primer: 5′GAATTCATGTCTTACCCCTACC3′, reverse primer: 5′CTCGAGACCTCTGCTCAGTA3′). Amplified products were subcloned into the pCR®4-TOPO® vector using the TOPO TA Cloning® Kit for Sequencing (Invitrogen, Carlsbad, CA). Positive clones were confirmed by *Eco*RI restriction digest and were directly sequenced using M13 forward and reverse primers by the UNC-CH Genome Analysis Facility. All three newly identified mouse aldolase A genes were deposited into GenBank using the following accession numbers: *Aldoa_v2* (EF662059), *Aldoart1* (EF662061), and *Aldoart2* (EF662062).

Immunoblotting of ALDOA-related proteins

Proteins were extracted from tissues or isolated cells by homogenization in lysis buffer (2% SDS, 100 mM DTT, 125 mM Tris pH 6.8, 18% glycerol). The lysates were centrifuged at 16,000×*g* for 10 min at 4 °C. The protein concentration in the supernatants was determined using the micro-BCA assay according to the manufacturer's instructions (Pierce Biotechnology, Rockford, IL). Samples with equal protein amounts were separated by SDS polyacrylamide gel electrophoresis (SDS-PAGE) on 10% polyacrylamide gels and electrophoretically transferred to Immobilon-P PVDF (polyvinylidene fluoride) membranes (Millipore Corp, Bedford, MA). Protein transfer and equal loading were confirmed by staining the membranes with 0.1% Coomassie blue R250 in 45% methanol, 10% acetic acid. Membranes were destained, rinsed with TBS-T (140 mM NaCl, 3 mM KCl, 0.05% Tween-20, 25 mM Tris–HCl, pH 7.4) and incubated in blocking buffer (5% nonfat dry milk in TBS-T) overnight at 4 °C. Membranes were then incubated for 1 h at room temperature with a polyclonal antibody raised against purified rabbit skeletal muscle aldolase (Polysciences, Warrington, PA) diluted 1:2000 in blocking buffer. After three 15-min washes with TBS-T, the membranes were incubated for 20 min at room temperature with secondary antibody (affinity-purified horseradish peroxidase-conjugated rabbit anti-goat IgG, KPL, Gaithersburg, MD) diluted 1:10,000 in blocking buffer. TBS-T washes were repeated and immunoreactive proteins were detected by enhanced chemiluminescence using the SuperSignal West Pico substrate (Pierce Biotechnology, Rockford, IL) and HyBlot CL autoradiography film (Denville Scientific, Metuchen, NJ).

Proteomic analysis of mouse sperm proteins

Mouse sperm were sonicated to separate sperm heads and tails and were further fractionated to obtain purified fibrous sheath preparations as described previously (Krisfalusi et al., 2006). Briefly, fragmented sperm tails were pelleted at 32,500×*g* after removing the sperm heads by low-speed centrifugation. Soluble sperm proteins released by sonication were precipitated with 10% trichloroacetic acid. Fibrous sheaths were isolated from the tail pellets by sequential incubations in Triton X-100 extraction buffer (1% Triton X-100, 2 mM dithiothreitol [DTT], 1 mM EDTA, 1 mM EGTA, 2 mM PMSF, 50 mM Tris–HCl, pH 9.0 and PI), potassium thiocyanate extraction buffer (0.6 M potassium thiocyanate, 2 mM DTT, 50 mM Tris–HCl, pH 8.0), and urea extraction buffer (6 M urea, 20 mM DTT, 50 mM Tris–HCl, pH 8.0). The final urea-insoluble pellet contained purified fibrous sheaths.

Proteins from the soluble, sperm tail and fibrous sheath fractions were resuspended in 2× SDS sample buffer (4% SDS, 18% glycerol, 125 mM Tris–HCl, pH 6.8) containing 50 mM TCEP reducing agent (*tris*(2-carboxyethyl) phosphine, Pierce Biotechnology, Rockford, IL), heated at 95 °C for 10 min, and separated by SDS-PAGE using pre-cast 10% Tris–HCl Ready Gels (Bio-Rad, Hercules, CA). The Benchmark™ Protein ladder (Invitrogen, Carlsbad, CA) was also resolved on these gels to determine approximate molecular weights of unknown proteins. SDS gels were fixed for 1 h at room temperature with gentle agitation in 45% methanol, 10% acetic acid in deionized water. Fixed gels were stained with SYPRO Ruby (Molecular Probes, Eugene, OR) overnight at room temperature with gentle agitation, followed by two 30-min washes in 10% methanol, 7% acetic acid in deionized water. Gels were stored in SYPRO Ruby stain until processed.

Gel bands corresponding to aldolase proteins identified by immunoblotting were cut from gels and processed using an in-gel digestion protocol as described previously (Parker et al., 2005). Peptides from the digested proteins were resuspended in 5% acetonitrile/95% water/0.1% trifluoroacetic acid (Buffer A) to facilitate liquid chromatography. Following digestion, the samples were subjected to MALDI-MS (matrix-assisted laser desorption ionization-mass spectrometry) and MS/MS analysis using the Applied Biosystems 4700 Proteomics Analyzer (Applied Biosystems Incorporated, Foster City, CA). The spectra were then searched against the NCBI and MSDB databases using the MASCOT search engine. After protein identification, samples were analyzed by liquid chromatography (LC)-MALDI-MS. Briefly, peptides in the sample (5 μl) were separated on a reverse-phase C18 nano-HPLC column (Dionex Corp., Sunnyvale, CA) using an Agilent 1100 Capillary HPLC system (Agilent Technologies Inc., Santa Clara, CA) with a post-pump split to deliver 250 nl/min through a manual injection valve. The column was coupled to a Probot Microfraction Collector (Dionex Corp., Sunnyvale, CA). A gradient from 5–65% Buffer B (95% acetonitrile/0.09% TFA) was run and the column eluate was spotted along with an equivalent volume of 60% acetonitrile supplied by the sheath flow from the Probot syringe pump. After drying, the fractions were overlaid with a saturated solution of α-cyano-4-hydroxycinamic acid in 50% acetonitrile/20 mM ammonium citrate. Spectra were collected and examined for the presence of aldolase isoform-specific peptide masses. Peptides within 100 ppm of predicted masses were subjected to MS/MS to confirm their identity.

Results

Genomic analyses predict two Aldoa-related retrogenes and a novel Aldoa splice variant

Ensembl BLAST searches with the coding sequence of the *Aldoa* gene on chromosome 7 detected related sequences on ten additional mouse chromosomes (Supplemental Table 1). These sequences lack introns and are consistent with a previous report of several *Aldoa* pseudogenes in mouse derived from multiple retrotransposition events (Cortinas and Lessa, 2001). In contrast, we did not detect retroposed sequences related to *Aldob* or *Aldoc* in comparable BLAST searches. Two of the *Aldoa*-related sequences have conserved full-length ORFs (including start and stop codons), suggesting that they may be functional retrogenes. These predicted retrogenes are located on chromosome 4 (*Aldoart1*; chromosome 4, minus strand, 72,337,671–72,339,071, Ensembl gene ID: ENSMUSG00000059343) and chromosome 12 (*Aldoart2*; chromosome 12, plus strand, 56,483,303–56,484,873, Ensembl gene ID: ENSMUSG00000063129). The start and end positions of the retrogenes were determined based on sequence alignments to *Aldoa*, combined with the identification of ESTs spanning both the 5′-UTR and 3′-UTR of each gene (see below).

Aldoart2 encodes a protein of the same size as ALDOA (Fig. 1, white boxes corresponding to homologous coding exons of *Aldoa*). The 5′- and 3′-UTRs of *Aldoart2* contain regions with homology to the 5′- and 3′-UTRs of *Aldoa* (boxes with diagonal lines), along with unique sequences (boxes with horizontal lines) that are proximal and distal to the *Aldoa*-homologous regions. *Aldoart1* is predicted to encode a larger protein because it contains sequence (black box, chromosome 4, minus strand, 72,338,786–72,338,943) that extends the ORF by 165 nucleotides at the N-terminus. This additional sequence encodes a

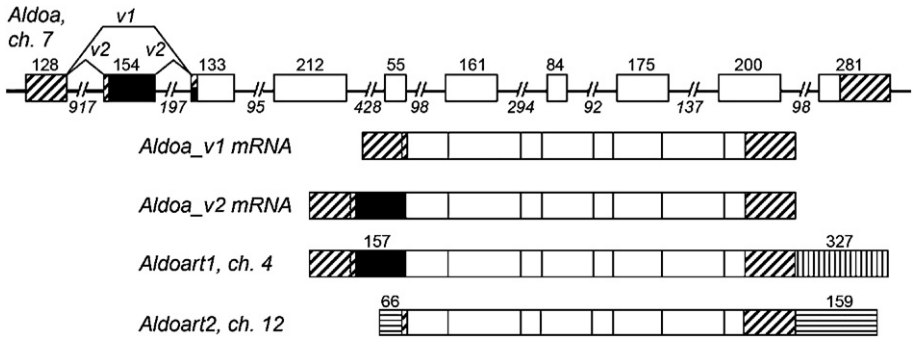

Fig. 1. Genomic organization and mRNA structure of the mouse *Aldoa* gene and related retrogenes. The *Aldoa* gene on chromosome 7 is shown at the top, with numbers indicating the lengths of exons (boxes) and introns (lines). *Aldoart1* and *Aldoart2* are intronless retrogenes on chromosomes 4 and 12, respectively. Coding sequences are indicated by white and black boxes, with black boxes representing novel N-terminal sequences. Boxes with diagonal lines denote homologous UTRs in *Aldoa*, *Aldoart1* and *Aldoart2*. Boxes with vertical or horizontal lines denote UTR sequences found only in *Aldoart1* or *Aldoart2*, respectively. The lengths of novel regions are noted.

55-amino-acid N-terminal extension that shows no homology with any of the previously described aldolase isoforms. The remaining *Aldoart1* coding sequence (white boxes), 5′-UTR and initial segment of the 3′-UTR (boxes with diagonal lines) are homologous to *Aldoa.* Only the terminal portion of the *Aldoart1* 3′-UTR has unique sequence that is not homologous to *Aldoa* or *Aldoart2*.

The *Aldoart1* sequence that encodes the novel N-terminal extension is 90% identical to a region within the first intron of *Aldoa*, providing evidence that *Aldoa* contains a previously unidentified exon. As shown at the top of Fig. 1, alternative splicing of *Aldoa* can generate both the ubiquitously transcribed mRNA with eight coding exons (*Aldoa_v1*) and a longer transcript (*Aldoa_v2*) from nine exons, including the newly identified exon 2 (154 nucleotides, chromosome 7, minus strand, 126,590,193–126,590,347). The first 21 nucleotides of exon 3 are part of the 5′-UTR in *Aldoa_v1*, but are part of the coding sequence in *Aldoa_v2*. The ORF of *Aldoa_v2* contains an additional 162 nucleotides encoding a 54-amino-acid N-terminal extension. There is an in-frame stop codon immediately upstream of the predicted start codon in both *Aldoa_v2* and *Aldoart1*.

Aldoart1 and Aldoart2 are derived from Aldoa and arose in the rodent lineage

Phylogenetic analyses of coding sequences conclusively demonstrate that *Aldoart1* and *Aldoart2* belong to the *Aldoa* subfamily (Fig. 2). Both *Aldoart1* and *Aldoart2* arose by retrotransposition in the rodent lineage after the primate–rodent split. All phylogenetic methods robustly predict that *Aldoart2* arose prior to the divergence of rat and mouse. This prediction was confirmed by the identification of the orthologous sequence for *Aldoart2* on rat chromosome 6 (plus strand,

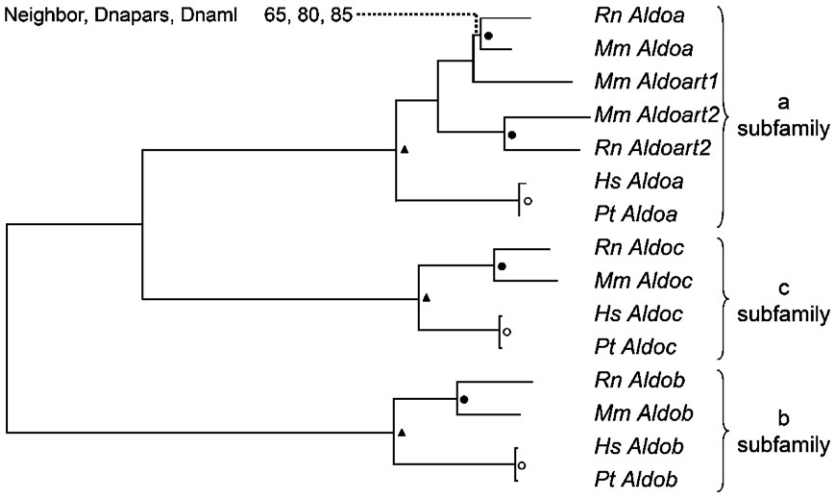

Fig. 2. Phylogenetic relationships between aldolase genes and retrogenes. The tree depicts the relationships between orthologs in mouse (*Mm*), rat (*Rn*), human (*Hs*) and chimpanzee (*Pt*), and is based on a consensus cladogram from three consensus trees obtained by three different phylogenetic methods (Neighbor, Dnapars and Dnam). The figure shows the number of times out of 100 that a branch is observed in each method for the less well-supported branch. All remaining branches are supported by more than 90% of trees in all three methods. Triangles denote the primate/rodent split. Open circles represent the split between the human and chimpanzee lineages. Filled circles represent the split between the mouse and rat lineages. The length of the branches is proportional to the ones obtained in a representative tree using the Neighbor method.

position 75,812,016–75,813,706). The rat *Aldoart2* has a conserved ORF with 94% identity to mouse *Aldoart2* (not shown), supporting the hypothesis that it is a functional retrogene in both species. It is uncertain whether *Aldoart1* arose before or after the rat–mouse split. Although phylogenetic analyses support a retrotransposition event just before the divergence of these two lineages, we were unable to find evidence for a homologous sequence anywhere in the rat genome, including the region of conserved synteny (90.5 Mb on rat chromosome 5). We conclude that the ORF and critical amino acids of both retrogenes have been conserved for a minimum of 15 million years in the mouse lineage for *Aldoart1* and for a much longer period in both the mouse and rat lineages in the case of *Aldoart2*. However, the retrogenes have evolved with higher substitution rates than other aldolase genes in rat and mouse, as indicated by the relative length of the retrogene branches in the phylogenetic tree (Fig. 2).

Aldoa_v2 is conserved in multiple mammalian species

We also performed phylogenetic analyses with the nucleotide sequences encoding the novel N-terminal extensions of *Aldoa_v2* from mouse (Fig. 3A, black segments of exons 2 and 3) and four other mammalian species, and mouse *Aldoart1*. The tree (Fig. 3B) replicates the known divergence order of these species and supports the sequence of events depicted in Fig. 2. The most remarkable feature of this tree is the striking difference in substitution rate in *Aldoart1* compared to *Aldoa_v2*. When amino acid sequences are compared (Fig. 3C), the N-terminal extensions of ALDOART1 and mouse ALDOA_V2 have 12-amino-acid differences (78% identity). The retention of an ORF with numerous substitutions confirms that the evolution of *Aldoart1* is not simply the result of purifying selection, but most likely involves positive selection for novel functions.

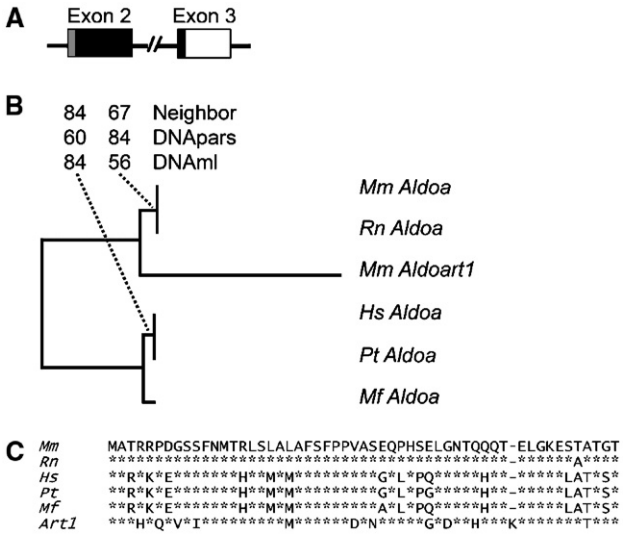

Fig. 3. Phylogenetic relationships between novel N-terminal sequences in the aldolase A family. (A) This diagram shows the two exons that generate the novel *Aldoa* splice variant (*Aldoa_v2*). Exon 2 includes a fragment of the 5′-UTR (shown as a grey box) and 144 nucleotides that encode the first 48 amino acids of the N-terminal extension of ALDO_V2 (black box). Exon 3 contains 21 nucleotides that encode the remaining 7 amino acids of the novel N-terminal sequence (black box). (B) The tree depicts the relationships between homologous sequences encoding the N-terminal extensions of *Aldoa_v2* and *Aldoart1*. The tree is based on a consensus cladogram from three consensus trees obtained by three different phylogenetic methods (Neighbor, Dnapars and Dnaml). The figure also shows the number of times out of 100 that a branch is observed in each method for the two less well-supported branches. All remaining branches are supported by more than 90% of trees in all three methods. (C) Conservation of the amino acid sequence of the N-terminal extensions across species. Amino acid alignment of the N-terminal extension of ALDOA_V2 in mouse (*Mm*), rat (*Rn*), macaque (*Mf*), chimpanzee (*Pt*), and human (*Hs*), and mouse ALDOART1 (*Art1*). An asterisk (*) denotes residues that are identical while a dash (–) denotes the absence of that residue.

The ALDOA-related isozymes have distinct amino acid sequences

Excluding the N-terminal extensions, the predicted amino acid sequences of ALDOART1 and ALDOART2 are 95.3% and 95.1% identical to ALDOA, respectively, while these isoforms are 92.3% identical to each other (Fig. 4). Active site residues (Arakaki et al., 2004) are conserved among all aldolase isoforms (Fig. 4, boxed residues). Both ALDOART1 and ALDOART2 contain unique amino acids that are not found in any other ALDO-related proteins (indicated by white letters on a black background). When the predicted amino acid sequences of mouse and rat *Aldoart2* are aligned, eight of the unique amino acids are conserved, including Y3, D130, G133, Q140, K156, S246, P266, and E348. ALDOART1 and ALDOART2 also contain residues that are identical to ALDOB and/or ALDOC (residues highlighted in grey). In ALDOART1 there are three adjacent amino acids (MGN, amino acids 95–97) between the first two active site residues that are identical to the homologous sequence in ALDOB (amino acids 39–41), but are different from the homologous IAK sequence in ALDOA. This region contains 2 amino acids (I39 in ALDOA and G40 in ALDOB, arrows in Fig. 4) that have been identified as isozyme-specific residues based on sequence conservation (Pezza et al., 2003). ALDOART1 also contains F112 in its sequence (arrow), which aligns with another ALDOB-specific residue (F58 in ALDOB).

ESTs support the restricted expression of Aldoa-related retrogenes

We searched mouse EST databases in our initial assessment of whether the predicted *Aldoa*-related retrogenes and splice variant are transcribed (Table 1). As expected, *in silico* analyses detected abundant transcripts for *Aldoa* in muscle, *Aldob* in liver, and *Aldoc* in brain. Transcripts for *Aldoa*, but not *Aldob* or *Aldoc*, were detected in testis. Multiple *Aldoart1* and *Aldoart2* ESTs were identified, mainly in testis libraries. Only 1 of 10 *Aldoart1* ESTs was detected in a library of unknown origin and 3 of 33 *Aldoart2* ESTs were detected in a spleen library. However, this spleen library also contains a similar number of ESTs from *Prm1*, a male germ cell-specific transcript (accession numbers AA290514, AI509226, and AI661727). ESTs specific for the alternatively spliced exon 2 of *Aldoa_v2* were found in libraries from testis, retina, a

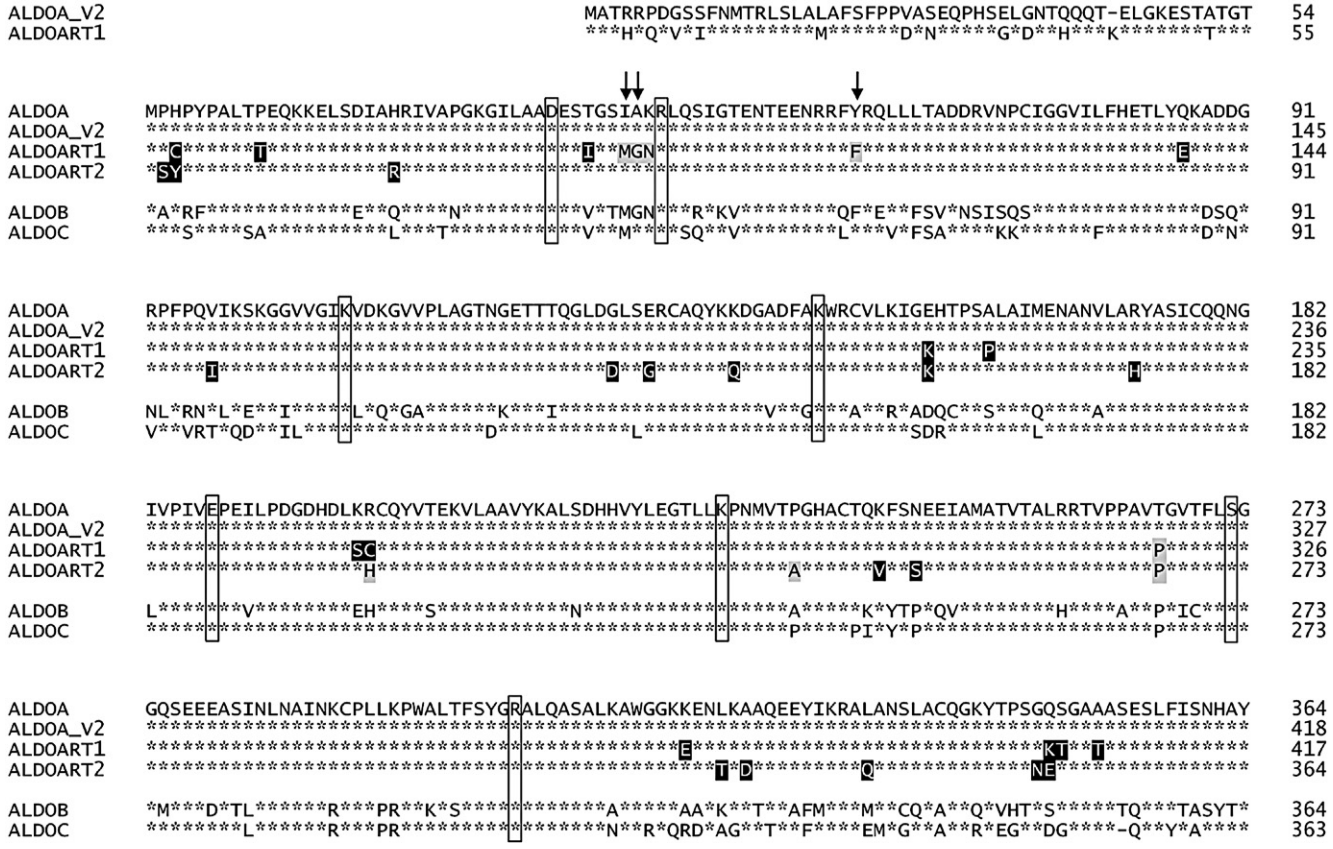

Fig. 4. Amino acid alignment for the mouse aldolase genes. Asterisks (∗) denotes identical amino acid residues. ALDOA_V2 and ALDOART1 have novel N-terminal extensions. Boxed amino acids denote active site residues, which are conserved in all members of the gene family. Amino acids that are unique to ALDOART1 or ALDOART2 are highlighted in black with white lettering. Additional residues in each retrogene that match ALDOB or ALDOC rather than ALDOA are highlighted in grey. Three isozyme-specific residues in ALDOART1 that match ALDOB are noted by black arrows.

mammary tumor cell line (RCB0527) and other carcinomas. Each EST was blasted against the mouse genome to confirm the gene from which it was transcribed (>98% identity for >100 bp). EST analyses indicate that the *Aldoart2* ortholog in rat is also expressed in the testis.

EST libraries are available for germ cells isolated at distinct stages of spermatogenesis and for testicular somatic cells, including Sertoli cells from the seminiferous epithelium and cell lines derived from the interstitial Leydig cells that are responsible for testosterone production. *Aldoa* transcripts were detected in Leydig and Sertoli cells and at earlier stages of spermatogenesis, including type B spermatogonia and pachytene spermatocytes (Table 1). *Aldoart1* and *Aldoart2* transcripts were detected only in round spermatids, which undergo haploid differentiation following meiosis. Frequency analysis supports the abundant expression of *Aldoart1* and *Aldoart2* in round spermatids (Table 2) compared to other testicular cell types (Table 2) and other tissues (Table 3). These EST analyses provide support for the expression of *Aldoa*-related retrogenes and the *Aldoa_v2* splice variant, with restricted expression of the retrogenes primarily during the later stages of spermatogenesis.

ESTs preferentially span the 3′ and 5′ ends of transcripts. The 3′ ends of the *Aldoart1* and *Aldoart2* transcripts and the 5′ end of *Aldoart2* were determined directly from the alignment of multiple ESTs to the corresponding genomic sequences. Since the 5′-UTRs of *Aldoart1* and *Aldoa* are homologous, we predict that *Aldoart1* has a transcription start site that is comparable to *Aldoa*, although available 5′-derived ESTs do not contain the first 88 bp of the predicted 5′-UTR sequence.

Table 1
Total number of aldolase ESTs in selected tissues and isolated testicular cells

	Selected tissues					Isolated testicular cells					
	All	Brain	Liver	Musc	Testis	L	S	A	B	P	R
Aldoc	>100	>100	1[a]	3[a]	0	0	0	0	0	0	0
Aldob	>100	0	>100	21	0	0	0	0	0	0	0
Aldoa	>100	>100	12	>100	8	1[a]	8[b]	0	3	1	0
Aldoa_v2[c]	17	0	0	0	1	0	0	0	0	0	0
Aldoart1	10	0	0	0	7	0	0	0	0	0	2
Aldoart2	33	0	0	0	27	0	0	0	0	0	3

ESTs were analyzed from selected mouse tissues and isolated testicular cells, including muscle (Musc), Leydig cells (L), Sertoli cells (S), type A spermatogonial (A), type B spermatogonial (B), pachytene spermatocytes (P) and round spermatids (R). ESTs from pooled libraries were excluded in this analysis.

[a] BLAST hits from cell lines.

[b] BLAST hits from primary Sertoli cell cultures.

[c] Refers to exon 2.

Table 2
Expression frequency of Aldoa-related genes

	Testis	Leydig	Sertoli	B Gonia	Pachytene	R Spermatid
Aldoa	0.8 (0.3)	11.4 (11.4)	7.5 (2.6)	3.7 (2.1)	1.1 (1.1)	0.0
Aldoart1	0.6 (0.2)	0.0	0.0	0.0	0.0	5.0 (3.5)
Aldoart2	2.3 (0.4)	0.0	0.0	0.0	0.0	7.5 (4.3)

Data are presented as the number of ESTs for each gene per 10,000 sequenced ESTs in libraries derived from testis and isolated cells including type B spermatogonia (B Gonia), pachytene spermatocytes (Pachytene) and round spermatids (R Spermatid). Standard errors are shown in parentheses.

RT-PCR confirms restricted transcription of Aldoa_v2, Aldoart1 and Aldoart2 in the testis

Although there is high sequence similarity between *Aldoa*, *Aldoart1* and *Aldoart2*, we developed an RT-PCR strategy using a single primer pair followed by *Hae*III digestion to distinguish among transcripts from these genes (Fig. 5A, also see Materials and methods). Following RT-PCR and *Hae*III digestion, we observed restriction fragments characteristic of *Aldoa* transcripts (*Aldoa_v1*or *Aldoa_v2*) in all tissues examined (Fig. 5B). Two distinct cleavage products each from *Aldoart1* and *Aldoart2* mRNAs were seen only in testis. Products were not amplified when the RT enzyme was omitted, confirming that the observed products were not derived from contaminating genomic DNA (data not shown). Analysis of RNAs isolated from the testes of juvenile mice (middle panel, Fig. 5B) indicated that *Aldoart1* and *Aldoart2* transcripts are both expressed beginning at 20 days of age, coincident with the appearance of round spermatids in the developing testis. PCR products specific for each of these retrogenes were present in testis RNAs isolated throughout subsequent pubertal development between 22 and 34 days of age (not shown). The appearance of *Aldoart1* and *Aldoart2* transcripts at day 20 was confirmed by three independent PCR replicates.

Semi-quantitative RT-PCR using the same primer set (right panel, Fig. 5B) followed by densitometry of the resulting products (Fig. 5C) showed that *Aldoart2* mRNA levels were elevated compared to other *Aldoa*-related transcripts, both in testis and in isolated round spermatids (R) and condensing spermatids (C). *Aldoart1* mRNA was detected in both round and condensing spermatids, at levels comparable to the expression of *Aldoa* and/or *Aldoa_v2* transcripts (Figs. 5B, C). When analyzed in the same linear range of PCR amplification, the levels of all *Aldoa*-related transcripts were reduced or absent in pachytene spermatocytes (P) and Sertoli cells. Minor contamination of isolated pachytene spermatocytes with round spermatids may account for the *Aldoart2* signal detected in this cell population.

Expression of the novel N-terminal coding sequences in *Aldoa_v2* and *Aldoart1* was confirmed using a similar RT-PCR strategy to amplify regions near the 5′ end of the *Aldoa* and *Aldoart1* transcripts (Fig. 6A). Primers used in these assays do not amplify a product from the *Aldoart2* transcript. *Hae*III restriction digestion cleaves only the *Aldoa_v2* RT-PCR product, generating distinct bands from *Aldoa_v1*, *Aldoa_v2* and *Aldoart1* transcripts that are resolved by gel electrophoresis. This assay detected expression of both *Aldoa_v2* and *Aldoart1* transcripts only in testis (Fig. 6B). *Aldoart1* transcripts first appeared at 20 days of age (Fig. 6C), confirming our previous analyses (Fig. 5B). In contrast, *Aldoa_v2* transcripts were detected throughout postnatal development of the testis between 8 and 34 days of age (Fig. 6C), indicating that this splice variant is expressed during earlier stages of spermatogenesis.

Proteins encoded by Aldoa_v2, Aldoart1 and Aldoart2 are present in mouse sperm

Our initial proteomic and Western analyses identified two ALDOA bands, with apparent molecular weights of ~39,000 and 50,000, in fibrous sheaths isolated from mouse sperm (Krisfalusi et al., 2006). Western blotting detected a ~39,000 molecular weight aldolase isoform in the soluble fraction following sonication of sperm (soluble, Fig. 7B) and in isolated sperm tails (tail P, Fig. 7B). Small amounts of this protein were sometimes detected in the fibrous sheath fraction (Krisfalusi et al., 2006). In contrast, the 50,000 molecular weight aldolase band was consistently enriched in isolated fibrous sheaths (FS, Figs. 7A and B). Detailed proteomic analyses of these bands identified unique peptide sequences that match the three ALDOA-related isoforms expressed during spermatogenesis (Fig. 7A, Supplemental Figs. S1–S3). Peptide sequences identified in the 39,000 molecular weight band from the soluble sperm fraction (Fig. 7A) matched the predicted amino acid sequence of ALDOART2 (Supplemental Fig. S1). We did not detect peptides unique to the somatic ALDOA isoform in this band, even though *in silico* tryptic digestion predicts that there are at least six peptides with sufficient difference to be assigned only to ALDOA.

Both *Aldoa_v2* and the *Aldoart1* retrogene encode larger proteins with 54 and 55 additional amino acids at the N-terminus, respectively. Our proteomic analyses identified peptide sequences matching each of these novel N-terminal sequences in the 50,000 molecular weight aldolase band from sperm or isolated fibrous sheaths (Fig. 7A, Supplemental Figs. S2B and S3B). Additional peptide sequences corresponding to other regions of ALDOA_V2 or ALDOART1 were also identified (Supplemental Figs. S2 and S3, respectively).

The larger ALDOA-related isoforms are synthesized late during spermatogenesis

Western analysis with a polyclonal antibody raised against rabbit muscle aldolase detected a 50,000 molecular weight band in mouse sperm (Fig. 8A) and testis (Fig. 8B), but not in several other tissues (Fig. 8A). We did not detect this band in mammary gland (virgin or lactating) or retina (data not shown), suggesting

Table 3
Aldoa-related gene expression in other tissues

	Tissue	Number of ESTs	Frequency[a]
Aldoart1	Unknown	1	0.5 (0.5)
Aldoart2	Spleen	3	0.4 (0.2)

[a] Number of ESTs for each gene per 10,000 sequenced ESTs in each library, with standard errors in parentheses.

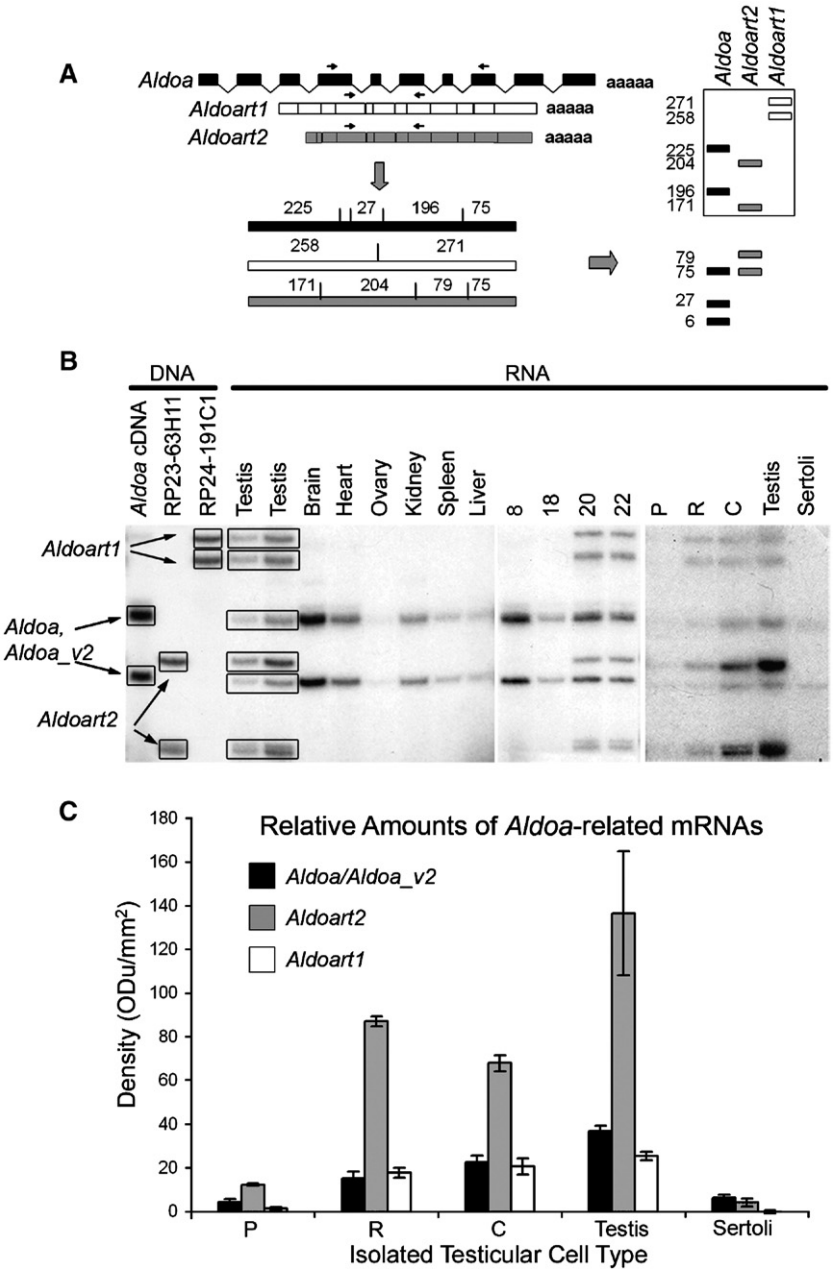

Fig. 5. Restricted expression of *Aldoart1* and *Aldoart2* during the late stages of spermatogenesis. (A) Diagram of the RT-PCR approach used to distinguish *Aldoa*-related transcripts. Arrows denote a single primer set which amplified a 530-bp product from *Aldoa/Aldoa_v2*, *Aldoart1* and *Aldoart2* cDNAs. *Hae*III digestion of these products produced distinct fragments that were resolved by electrophoresis on 5% acrylamide gels. (B) The expected fragments were amplified from DNA-positive controls, including an *Aldoa* cDNA subclone and BAC (bacterial artificial chromosome) clones for the intronless retrogenes (clone RP24-191C1 for *Aldoart1*, clone RP23-63H11 for *Aldoart2*). RT-PCR of total RNA samples detected *Aldoart1* and *Aldoart2* transcripts in testis, but not in other tissues. Transcripts from the *Aldoa* gene (*Aldoa_v1* or *Aldoa_v2*) were present in all tissues examined. RNAs for the second testis lane and remaining tissues were isolated from (C57BL/6J × A/J) F1 mice, while RNAs for the first testis lane, testes from juvenile animals between 8 and 22 days of age (8, 18, 20, 22) and isolated testicular cells were from CD-1 mice. During postnatal development, *Aldoart1* and *Aldoart2* are first expressed in the testis at 20 days of age. Semi-quantitative RT-PCR conditions were used to compare transcript levels in isolated pachytene spermatocytes (P), round spermatids (R), condensing spermatids (C) and Sertoli cells. (C) Analysis of the semi-quantitative RT-PCR products by densitometry. Bars represent mean densities ± SEM calculated from 3–6 replicated PCR reactions.

that *Aldoa_v2* transcripts detected by EST analysis in these tissues may represent artifacts or may not be translated. Despite the appearance of *Aldoart1* transcripts in round spermatids and *Aldoa_v2* transcripts during earlier stages of spermatogenesis (Figs. 5 and 6), the larger ALDOA-related protein band was first detected late in spermatogenesis in condensing spermatids (C in Fig. 8B). This protein was not detected in germ cells at earlier stages of spermatogenesis or in Sertoli cells. Therefore, expression of *Aldoa_v2* and *Aldoart1* during spermatogenesis appears to be regulated at both transcriptional and translational levels.

A ~50,000 molecular weight aldolase is also present in rat and human sperm (Fig. 8C). Alternatively spliced *Aldoa_v2* transcripts are likely to encode this isoform, since we did not

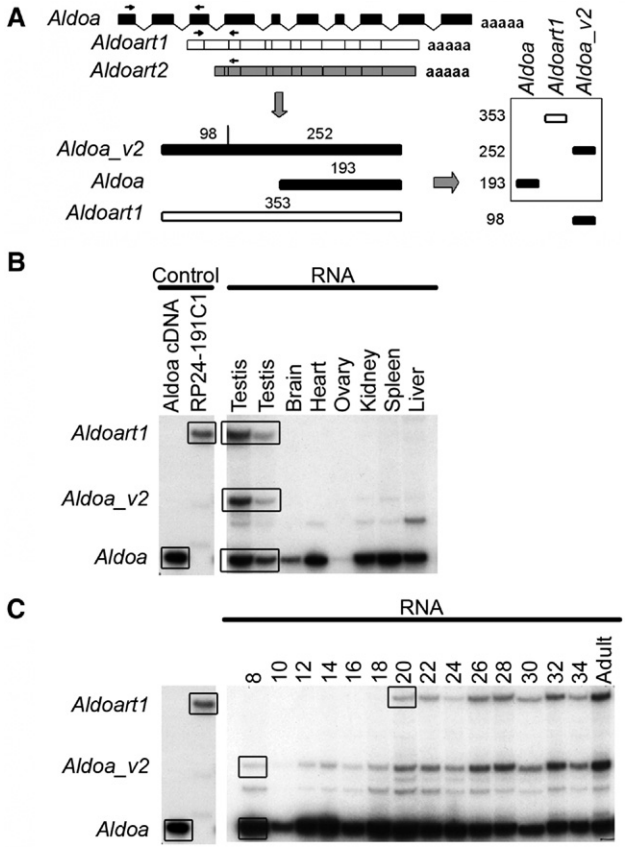

Fig. 6. Restricted expression of *Aldoa_v2* in testis. (A) Diagram of the RT-PCR strategy used to distinguish between *Aldoa*, *Aldoa_v2*, *Aldoart1*. Arrows denote a single primer set which amplified 350-, 193- and 353-bp products from *Aldoa_v2*, *Aldoa*, and *Aldoart1*, respectively. *Hae*III digestion cleaved only the *Aldoa_v2* product, resulting in distinct bands for each transcript on 5% acrylamide gels. (B) This strategy detected *Aldoa_v2* transcripts only in testis (CD-1 testis in the first RNA lane, other tissue RNAs from $(C57BL/6J \times A/J)F_1$ mice). (C) Testis RNAs from juvenile CD-1 mice (8–34 days of age) were also analyzed to compare the appearance of *Aldoa_v2* and *Aldoart1* transcripts during postnatal development. RNAs were isolated from pooled testes and RT-PCR was repeated three times to confirm reproducibility. Control DNAs for the assays in panels B and C included an *Aldoa* cDNA subclone and a BAC clone for *Aldoart1* (clone RP24-191C1).

identify *Aldoart1* retrogenes in either of these species. This result, along with the high level of sequence conservation across species (Fig. 3C), strongly supports a functional role for the extra amino acid residues located at the N-terminus of ALDOA_V2 and ALDOART1.

Discussion

This study identifies three new aldolase isozymes in mouse with restricted expression during male germ cell development. All three isozymes first appear late during the haploid phase of spermatogenesis and are retained in mature sperm. ALDOART1 and ALDOART2 are encoded by *Aldoa*-related retrogenes which arose in the rodent lineage. ALDOA_V2 is a splice variant of *Aldoa* that is conserved in several mammalian species, including human. Both ALDOART1 and ALDOA_V2 have novel N-terminal extensions that may be responsible for tight

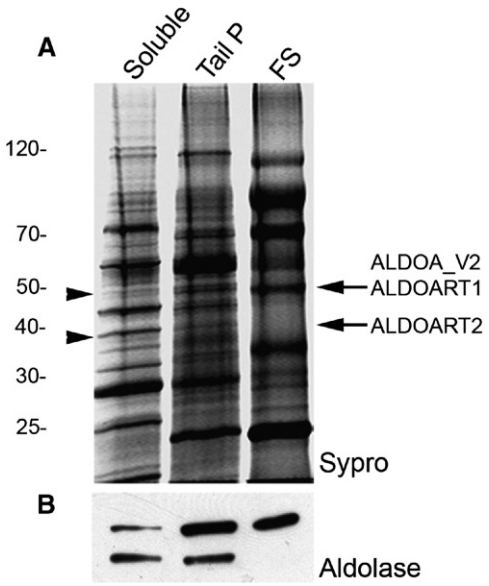

Fig. 7. Proteomic and Western analyses identify novel ALDOA-related proteins in mouse sperm. Proteins in the soluble fraction following sonication (soluble), the tail pellet (tail P), and isolated fibrous sheaths (FS) were separated by SDS-PAGE. Protein from 10^6 sperm was loaded in the soluble and tail P lanes, while the FS lane contains protein from 10^7 sperm. (A) Bands indicated by arrows on this representative Sypro Ruby-stained gel were isolated and analyzed by mass spectrometry. MALDI identified the bands as aldolase-A related proteins. LC-MALDI identified peptides specific for ALDOART2 in the soluble ~39,000 molecular weight protein and peptides specific for ALDOA_V2 and ALDOART1 in the ~50,000 molecular weight band enriched in the tail P and FS fractions (see Supplemental Figs. S1–S3). (B) Western analysis of the three fractions using a polyclonal antibody raised against rabbit skeletal muscle aldolase.

binding to the fibrous sheath. We hypothesize that these N-terminal extensions are important for the proper localization of glycolytic ATP production, required for sperm motility. Our data exemplify the diverse mechanisms that generate genes with restricted expression patterns and novel functions in the germ-

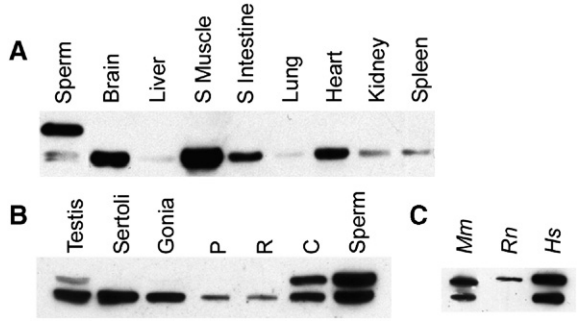

Fig. 8. Larger ALDOA-related isoforms are first translated in condensing spermatids and are conserved in mouse, rat and human sperm. Following SDS-PAGE, Western blots were probed with a polyclonal antibody raised against rabbit skeletal muscle aldolase. Equal protein loads were confirmed by Coomassie blue staining of the blots (not shown). (A) A ~50,000 molecular weight ALDOA-related band was detected in mouse sperm, but not in eight somatic tissues including skeletal muscle (S Muscle) and small intestine (S Intestine). (B) The larger ALDOA-related band was also detected in testis and condensing spermatids (C), but not in Sertoli cells, spermatogonia (Gonia), pachytene spermatocytes (P) or round spermatids (R). (C) A larger ALDOA-related band is present in mouse (*Mm*), rat (*Rn*) and human (*Hs*) sperm.

line, and illustrate the role of positive selection in maintaining expression of newly evolved genes with unique attributes.

Retrotransposition is a common feature of mammalian evolution, producing >4800 retroposed sequences in the mouse genome (Sakai et al., 2007; Zhang et al., 2004) and ~8000 in the human genome (Nishimune and Tanaka, 2006; Zhang and Gerstein, 2004; Zhang et al., 2004). LINE-1 transposable elements facilitate this process, initiating reverse transcription of an mRNA and insertion of the resulting intronless sequence into the genome at a new location (Kazazian, 2004). Only those retrotransposition events that occur in the germline are heritable and, therefore, can be exploited by natural selection to generate new gene family members. Although most retroposed sequences are pseudogenes incapable of coding functional proteins, EST analyses indicate that >1000 human retroposed sequences may be transcribed (Vinckenbosch et al., 2006). Further studies have confirmed that at least 100 of these sequences are functional retrogenes, with many exhibiting specific roles during spermatogenesis (Emerson et al., 2004; Marques et al., 2005; Nishimune and Tanaka, 2006; Vinckenbosch et al., 2006). We identified 17 retroposed sequences derived from the ancestral *Aldoa* gene located on mouse chromosome 7 (Supplemental Table 1), but none from *Aldob* or *Aldoc*, consistent with the observation that genes with multiple retroposed copies are frequently housekeeping genes that are highly expressed in the germline (Zhang and Gerstein, 2004; Zhang et al., 2004). Only *Aldoart1* and *Aldoart2*, which arose from independent retrotransposition events in the rodent lineage, have full-length conserved ORFs. These retrogenes, like *Pgk2* (Boer et al., 1987; McCarrey and Thomas, 1987), contribute to the diversity of the glycolytic pathway in the male germline.

It has been proposed that expression of retrogenes during spermatogenesis is facilitated by a transcriptionally permissive environment (Babushok et al., 2006; Kleene et al., 1998), particularly during the haploid stages when multiple components of the RNA polymerase II transcription machinery are substantially upregulated (Schmidt and Schibler, 1995; Schmidt et al., 1997). Both *Aldoart1* and *Aldoart2* mRNAs appear in the testis at 20 days of age and are selectively expressed in spermatids, suggesting that they may share common regulatory mechanisms. Although the promoters of these retrogenes have not been defined, comparison of upstream flanking sequences indicates that they are not derived from *Aldoa*, their ubiquitously expressed progenitor gene.

Unique constituents with restricted expression in spermatogenic cells have been identified in the multi-protein complexes required for transcriptional initiation (DeJong, 2006; Elliott and Grellscheid, 2006; Hochheimer and Tjian, 2003; Sergeant et al., 2007), alternative splicing (Elliott and Grellscheid, 2006; Liu et al., 2007; Sergeant et al., 2007), and 3′-processing of mRNAs (Liu et al., 2007). Alterations in these processes contribute to the complexity of gene expression during spermatogenesis. Analysis of human ESTs indicates that the testis, along with the brain and liver, have the highest frequencies of alternative splicing (Yeo et al., 2004). In this study, identification of the *Aldoart1* retrogene led to the discovery of an alternatively spliced exon that encodes an N-terminal extension of ALDOA. The *Aldoa_v2* splice variant, which served as the template for the *Aldoart1* retrogene, is expressed predominantly in testis. The restricted expression of this alternatively spliced exon and its conservation across distantly related species provide evidence that the encoded N-terminal sequence has an important functional role in male germ cells.

Western analysis demonstrated a delay in translation of the larger aldolases encoded by *Aldoart1* and the *Aldoa_v2* splice variant. Storage of mRNAs and delayed translation are common features of mammalian spermatogenesis, particularly during the haploid period of differentiation where extensive morphogenesis continues for a week or longer after transcription ceases (Eddy, 2002; Kleene, 1996). Although *Aldoa_v2* is transcribed throughout spermatogenesis and *Aldoart1* is initially transcribed in round spermatids, the 50,000 molecular weight aldolase band was first detected in condensing spermatids, which appear in the testis during postnatal development approximately 1 week later than the round spermatids. GAPDHS (Bunch et al., 1998; Mori et al., 1992) and PGK2 (McCarrey et al., 1992, 1996), two additional spermatogenesis-specific isoforms in the glycolytic pathway, are also synthesized in condensing spermatids following significant periods of translation repression. Both global and mRNA-specific mechanisms may regulate the translation of *Aldoart1* and *Aldoa_v2*, along with many other mRNAs transcribed during meiosis or the initial haploid phase of spermatogenesis (Kleene, 2003).

Novel features of the newly identified ALDOA-related isozymes, including their restricted expression and sequence diversity, suggest neofunctionalization and positive selection within the *Aldo* gene family. ALDOA, ALDOB and ALDOC, the three known aldolase isozymes in vertebrates, catalyze the reversible cleavage of fructose-1,6-bisphosphate to glyceraldehyde 3-phosphate and dihydroxyacetone phosphate (Penhoet and Rutter, 1975; Rutter et al., 1963). ALDOA and ALDOC preferentially catalyze the cleavage reaction essential for glycolysis, while the catalytic properties of ALDOB favor the reverse reaction required for gluconeogenesis (Penhoet, 1969; Penhoet and Rutter, 1971; Penhoet et al., 1969; Rutter et al., 1963). ALDOB also cleaves fructose-1-phosphate, a reaction that is essential for fructose metabolism (Cori et al., 1951; Penhoet et al., 1966). Amino acid differences in the ALDOA-related isozymes may contribute to distinctive enzymatic or structural properties, as reported for bovine sperm aldolases (Gillis and Tamblyn, 1984). Three of the amino acids in *Aldoart1* (M95, G96, F112) that match ALDOB rather than ALDOA have been identified as isozyme-specific residues (Pezza et al., 2003). The *Aldoart1* MGN sequence between the first two active site residues that is identical to ALDOB is particularly interesting since it is in a region that contributes to the distinct properties of ALDOA and ALDOB (Kitajima et al., 1990; Kusakabe et al., 1994; Motoki et al., 1993). Further studies will be necessary to assess the catalytic properties of the ALDOA-related isozymes. Although the potential for metabolizing fructose via an ALDOB-like isozyme would appear to be beneficial for sperm, fructose is apparently unable to support hyperactivated sperm motility and *in vitro* fertilization in the

mouse (Cooper, 1984; Fraser and Quinn, 1981; Hoppe, 1976). It is interesting to note that most of the isozyme-specific residues identified for the aldolases are not within hydrogen bonding distance of the active site and are localized on the surface where they may exert long-distance conformational effects or mediate tissue-specific protein interactions (Pezza et al., 2003). Several of these residues are localized in the flexible C-terminal region that has been shown to contribute to differences between the aldolase isozymes (Berthiaume et al., 1993; Motoki et al., 1993). Multiple unique residues are present within the C-terminal regions of both *Aldoart1* and *Aldoart2* where they could contribute to novel enzymatic properties.

Both ALDOART1 and ALDOA_V2, the larger aldolase isozymes, are present in purified fibrous sheath preparations. Like GAPDHS (Bunch et al., 1998), these isozymes have distinctive N-terminal extensions that could mediate targeting to the fibrous sheath and/or tight binding to this cytoskeletal structure. Aldolase functions as a tetramer, and heterotetramers of ALDOA with ALDOB or ALDOC have been identified in various tissues (Penhoet and Rutter, 1971; Penhoet et al., 1966). Our proteomic analyses did not detect ALDOA in mouse sperm. The restricted expression of the newly identified ALDOA-related isoforms in spermatids suggests that functional tetramers containing one or more of these isoforms are formed late during spermatogenesis and are tethered to the forming fibrous sheath by the N-terminal extensions of ALDOART1 and ALDOA_V2. Binding of aldolase tetramers to an insoluble structure such as the fibrous sheath could affect their conformation and modulate their kinetic properties.

Compartmentalization of the glycolytic enzymes in the principal piece may facilitate energy production along the entire length of the sperm flagellum, providing a localized supply of ATP to the dynein ATPases in the axoneme to maintain motility. Enzymatic complexes are common in nature and direct binding between glycolytic enzymes has been suggested (Ouporov et al., 2001; Vertessy et al., 1997; Westhoff and Kamp, 1997). Recent proteomic studies identified multiple glycolytic enzymes in the flagella of *Chlamydomonas reinhardtii*, including some that are associated with the axoneme (Mitchell et al., 2005; Pazour et al., 2005). These studies suggest that there may be a more general requirement for glycolytic enzyme localization and ATP production in flagella. Some glycolytic enzymes, including ALDOA and GAPDH, are co-localized at the Z-discs and M-lines in *Drosophila* flight muscles (Sullivan et al., 2003; Wojtas et al., 1997). Flies that lack glycerol-3-phosphate dehydrogenase lose the ability to fly and the localization of both ALDOA and GAPDH in the flight muscle, providing another link between the co-localization and function of glycolytic enzymes (Sullivan et al., 2003; Wojtas et al., 1997). Similarly, the co-localization of glycolytic enzymes in the principal piece of mammalian sperm may result in higher catalytic activity to support sperm motility.

There is ample evidence that the glycolytic pathway is extensively modified during mammalian spermatogenesis and is required for sperm function and male fertility. This study indicates that both retrotransposition and alternative splicing are responsible for the restricted expression of three novel ALDOA-related isoforms during the haploid phase of mouse spermatogenesis. All three are present in mature sperm and two isoforms with N-terminal extensions are tightly bound to the fibrous sheath. Further studies to assess the functional properties of the ALDOA-related isozymes and other sperm-specific glycolytic enzymes may provide new insights regarding potential genetic causes of infertility and the rational design of contraceptives. In addition, detailed genomic and molecular analyses of novel glycolytic enzymes in the male germline offer a promising approach to identify elements that regulate gene expression during spermatogenesis and to increase our understanding of the causes underlying the differential rates and non-random locations of retrotransposition. The fate of most retrotransposition events is to fade into evolutionary history as pseudogenes. It is, therefore, remarkable that *Aldoart1* and *Aldoart2* show a striking combination of negative selection at all active site residues (Fig. 4), convergent evolution at isozyme-specific residues (Fig. 4), and positive selection at many other amino acids (Figs. 2 and 3B). Since these two retrogenes are only expressed during the final stages of spermatogenesis and their products are present in sperm, we hypothesize that reproductive performance is the most likely force that has shaped their evolution.

Acknowledgments

This research was supported by NICHD/NIH through cooperative agreement U54 HD35041 as part of the Specialized Cooperative Centers Program in Reproduction and Infertility Research (DAO and UNC-Duke Michael Hooker Proteomics Center), NIGMS/NIH as part of the Center of Excellence in Systems Biology (FPMV, 1P50 GM076468) and the National Science Foundation (FPMV, MCB-0133526). We thank David Robinette and Nedyalka Dicheva for their technical assistance with proteomic methods. We also thank Michael O'Rand for the human sperm samples, Ellen Weiss for the mouse retina samples, and Subashini Chandrasekharan for the mouse mammary gland samples.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ydbio.2007.06.010.

References

Arakaki, T.L., Pezza, J.A., Cronin, M.A., Hopkins, C.E., Zimmer, D.B., Tolan, D.R., Allen, K.N., 2004. Structure of human brain fructose 1,6-(*bis*) phosphate aldolase: linking isozyme structure with function. Protein Sci. 13, 3077–3084.

Babushok, D.V., Ostertag, E.M., Courtney, C.E., Choi, J.M., Kazazian Jr., H.H., 2006. L1 integration in a transgenic mouse model. Genome Res. 16, 240–250.

Berthiaume, L., Tolan, D.R., Sygusch, J., 1993. Differential usage of the carboxyl-terminal region among aldolase isozymes. J. Biol. Chem. 268, 10826–10835.

Bluthmann, H., Cicurel, L., Kuntz, G.W., Haedenkamp, G., Illmensee, K., 1982. Immunohistochemical localization of mouse testis-specific phosphoglycerate kinase (PGK-2) by monoclonal antibodies. EMBO J. 1, 479–484.

Boer, P.H., Adra, C.N., Lau, Y.F., McBurney, M.W., 1987. The testis-specific

phosphoglycerate kinase gene pgk-2 is a recruited retroposon. Mol. Cell. Biol. 7, 3107–3112.

Buehr, M., McLaren, A., 1981. An electrophoretically detectable modification of glucosephosphate isomerase in mouse spermatozoa. J. Reprod. Fertil. 63, 169–173.

Bunch, D.O., Welch, J.E., Magyar, P.L., Eddy, E.M., O'Brien, D.A., 1998. Glyceraldehyde 3-phosphate dehydrogenase-S protein distribution during mouse spermatogenesis. Biol. Reprod. 58, 834–841.

Coonrod, S., Vitale, A., Duan, C., Bristol-Gould, S., Herr, J., Goldberg, E., 2006. Testis-specific lactate dehydrogenase (LDH-C4; Ldh3) in murine oocytes and preimplantation embryos. J. Androl. 27, 502–509.

Cooper, T.G., 1984. The onset and maintenance of hyperactivated motility of spermatozoa in the mouse. Gamete Res. 9, 55–74.

Cori, G.T., Ochoa, S., Slein, M.W., Cori, C.F., 1951. The metabolism of fructose in liver. Isolation of fructose-2-phosphate and inorganic pyrophosphate. Biochim. Biophys. Acta 7, 304–317.

Cortinas, M.N., Lessa, E.P., 2001. Molecular evolution of aldolase A pseudogenes in mice: multiple origins, subsequent duplications, and heterogeneity of evolutionary rates. Mol. Biol. Evol. 18, 1643–1653.

DeJong, J., 2006. Basic mechanisms for the control of germ cell gene expression. Gene 366, 39–50.

Eddy, E.M., 2002. Male germ cell gene expression. Recent Prog. Horm. Res. 57, 103–128.

Eddy, E.M., Toshimori, K., O'Brien, D.A., 2003. Fibrous sheath of mammalian spermatozoa. Microsc. Res. Tech. 61, 103–115.

Edwards, Y.H., Grootegoed, J.A., 1983. A sperm-specific enolase. J. Reprod. Fertil. 68, 305–310.

Elliott, D.J., Grellscheid, S.N., 2006. Alternative RNA splicing regulation in the testis. Reproduction 132, 811–819.

Emerson, J.J., Kaessmann, H., Betran, E., Long, M., 2004. Extensive gene traffic on the mammalian X chromosome. Science 303, 537–540.

Fraser, L.R., Quinn, P.J., 1981. A glycolytic product is obligatory for initiation of the sperm acrosome reaction and whiplash motility required for fertilization in the mouse. J. Reprod. Fertil. 61, 25–35.

Gillis, B.A., Tamblyn, T.M., 1984. Association of bovine sperm aldolase with sperm subcellular components. Biol. Reprod. 31, 25–35.

Gitlits, V.M., Toh, B.H., Loveland, K.L., Sentry, J.W., 2000. The glycolytic enzyme enolase is present in sperm tail and displays nucleotide-dependent association with microtubules. Eur. J. Cell Biol. 79, 104–111.

Goldberg, E., 1977. Isozymes in testes and spermatozoa. Isozymes. Curr. Topics Biol. Med. Res. 1, 79–124.

Hochheimer, A., Tjian, R., 2003. Diversified transcription initiation complexes expand promoter selectivity and tissue-specific gene expression. Genes Dev. 17, 1309–1320.

Hoppe, P.C., 1976. Glucose requirement for mouse sperm capacitation in vitro. Biol. Reprod. 15, 39–45.

Hoskins, D.D., 1973. Adenine nucleotide mediation of fructolysis and motility in bovine epididymal spermatozoa. J. Biol. Chem. 248, 1135–1140.

Kazazian Jr., H.H., 2004. Mobile elements: drivers of genome evolution. Science 303, 1626–1632.

Kim, Y.H., Haidl, G., Schaefer, M., Egner, U., Herr, J.C., 2006. Compartmentalization of a unique ADP/ATP carrier protein SFEC (Sperm Flagellar Energy Carrier, AAC4) with glycolytic enzymes in the fibrous sheath of the human sperm flagellar principal piece. Dev. Biol. 302, 463–476.

Kitajima, Y., Takasaki, Y., Takahashi, I., Hori, K., 1990. Construction and properties of active chimeric enzymes between human aldolases A and B. Analysis of molecular regions which determine isozyme-specific functions. J. Biol. Chem. 265, 17493–17498.

Kleene, K.C., 1996. Patterns of translational regulation in the mammalian testis. Mol. Reprod. Dev. 43, 268–281.

Kleene, K.C., 2003. Patterns, mechanisms, and functions of translation regulation in mammalian spermatogenic cells. Cytogenet. Genome Res. 103, 217–224.

Kleene, K.C., Mulligan, E., Steiger, D., Donohue, K., Mastrangelo, M.A., 1998. The mouse gene encoding the testis-specific isoform of Poly(A) binding protein (Pabp2) is an expressed retroposon: intimations that gene expression in spermatogenic cells facilitates the creation of new genes. J. Mol. Evol. 47, 275–281.

Krisfalusi, M., Miki, K., Magyar, P.L., O'Brien D, A., 2006. Multiple glycolytic enzymes are tightly bound to the fibrous sheath of mouse spermatozoa. Biol. Reprod. 75, 270–278.

Kusakabe, T., Motoki, K., Sugimoto, Y., Takasaki, Y., Hori, K., 1994. Human aldolase B: liver-specific properties of the isozyme depend on type B isozyme group-specific sequences. Protein Eng. 7, 1387–1393.

Li, S.S., O'Brien, D.A., Hou, 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.

Liu, D., Brockman, J.M., Dass, B., Hutchins, L.N., Singh, P., McCarrey, J.R., MacDonald, C.C., Graber, J.H., 2007. Systematic variation in mRNA 3′-processing signals during mouse spermatogenesis. Nucleic Acids Res. 35, 234–246.

Mann, T., Lutwak-Mann, C., 1981. Male Reproductive Function and Semen. Springer-Verlag, New York, NY.

Marques, A.C., Dupanloup, I., Vinckenbosch, N., Reymond, A., Kaessmann, H., 2005. Emergence of young human genes after a burst of retroposition in primates. PLoS Biol. 3, e357.

McCarrey, J.R., Thomas, K., 1987. Human testis-specific *PGK* gene lacks introns and possesses characteristics of a processed gene. Nature 326, 501–505.

McCarrey, J.R., Berg, W.M., Paragioudakis, S.J., Zhang, P.L., Dilworth, D.D., Arnold, B.L., Rossi, J.J., 1992. Differential transcription of *pgk* genes during spermatogenesis in the mouse. Dev. Biol. 154, 160–168.

McCarrey, J.R., Kumari, M., Aivaliotis, M.J., Wang, Z., Zhang, P., Marshall, F., Vandeberg, J.L., 1996. Analysis of the cDNA and encoded protein of the human testis-specific *pgk-2* gene. Dev. Genet. 19, 321–332.

McCarrey, J.R., O'Brien, D.A., Skinner, M.K., 1999. Construction and preliminary characterization of a series of mouse and rat testis cDNA libraries. J. Androl. 20, 635–639.

Miki, K., Qu, W., Goulding, E.H., Willis, W.D., Bunch, D.O., Strader, L.F., Perreault, S.D., Eddy, E.M., O'Brien, D.A., 2004. Glyceraldehyde 3-phosphate dehydrogenase-S, a sperm-specific glycolytic enzyme, is required for sperm motility and male fertility. Proc. Natl. Acad. Sci. U. S. A. 101, 16501–16506.

Mitchell, B.F., Pedersen, L.B., Feely, M., Rosenbaum, J.L., Mitchell, D.R., 2005. ATP production in *Chlamydomonas reinhardtii* flagella by glycolytic enzymes. Mol. Biol. Cell 16, 4509–4518.

Mori, C., Welch, J.E., Sakai, Y., Eddy, E.M., 1992. In situ localization of spermatogenic cell-specific glyceraldehyde 3-phosphate dehydrogenase (gapd-s) messenger ribonucleic acid in mice. Biol. Reprod. 46, 859–868.

Mori, C., Welch, J.E., Fulcher, K.D., O'Brien, D.A., Eddy, E.M., 1993. Unique hexokinase messenger ribonucleic acids lacking the porin-binding domain are developmentally expressed in mouse spermatogenic cells. Biol. Reprod. 49, 191–203.

Mori, C., Nakamura, N., Welch, J.E., Gotoh, H., Goulding, E.H., Fujioka, M., Eddy, E.M., 1998. Mouse spermatogenic cell-specific type 1 hexokinase (mHk1-s) transcripts are expressed by alternative splicing from the *mHk1* gene and the HK1-S protein is localized mainly in the sperm tail. Mol. Reprod. Dev. 49, 374–385.

Motoki, K., Kitajima, Y., Hori, K., 1993. Isozyme-specific modules on human aldolase A molecule. Isozyme group-specific sequences 1 and 4 are required for showing characteristics as aldolase A. J. Biol. Chem. 268, 1677–1683.

Mukai, C., Okuno, M., 2004. Glycolysis plays a major role for adenosine triphosphate supplementation in mouse sperm flagellar movement. Biol. Reprod. 71, 540–547.

Narisawa, S., Hecht, N.B., Goldberg, E., Boatright, K.M., Reed, J.C., Millan, J.L., 2002. Testis-specific cytochrome *c*-null mice produce functional sperm but undergo early testicular atrophy. Mol. Cell. Biol. 22, 5554–5562.

Nishimune, Y., Tanaka, H., 2006. Infertility caused by polymorphisms or mutations in spermatogenesis-specific genes. J. Androl. 27, 326–334.

O'Brien, D.A., 1993. Isolation, separation, and short-term culture of spermatogenic cells. In: Chapin, R.E., Heindel, J.J. (Eds.), Male reproductive toxicology. Methods in Toxicology, pp. 246–264.

O'Brien, D.A., Gabel, C.A., Rockett, D.L., Eddy, E.M., 1989. Receptor-mediated endocytosis and differential synthesis of mannose 6-phosphate receptors in isolated spermatogenic and Sertoli cells. Endocrinology 125, 2973–2984.

Ouporov, I.V., Knull, H.R., Huber, A., Thomasson, K.A., 2001. Brownian dynamics simulations of aldolase binding glyceraldehyde 3-phosphate dehydrogenase and the possibility of substrate channeling. Biophys. J. 80, 2527–2535.

Pang, A.L., Johnson, W., Ravindranath, N., Dym, M., Rennert, O.M., Chan, W.Y., 2006. Expression profiling of purified male germ cells: stage-specific expression patterns related to meiosis and postmeiotic development. Physiol. Genomics 24, 75–85.

Parker, C.E., Mocanu, V., Warren, M.R., Greer, S.F., Scarlett, C.O., Borchers, C.H., 2005. Mass spectrometric determination of protein ubiquitination. In: Patterson, C., Cyr, D.M. (Eds.), Ubiquitin-Proteasome Protocols, Methods in Molecular Biology Series. Humana Press, Patterson, NJ, pp. 117–152.

Pazour, G.J., Agrin, N., Leszyk, J., Witman, G.B., 2005. Proteomic analysis of a eukaryotic cilium. J. Cell Biol. 170, 103–113.

Penhoet, E.E., 1969. Isolation of fructose diphosphate aldolases A, B, and C. Biochemistry 8, 4391–4395.

Penhoet, E.E., Rutter, W.J., 1971. Catalytic and immunochemical properties of homomeric and heteromeric combinations of aldolase subunits. J. Biol. Chem. 246, 318–323.

Penhoet, E.E., Rutter, W.J., 1975. Detection and isolation of mammalian fructose-diphosphate aldolases. Methods Enzymol. 42, 240–249.

Penhoet, E., Rajkumar, T., Rutter, W.J., 1966. Multiple forms of fructose diphosphate aldolase in mammalian tissues. Proc. Natl. Acad. Sci. U. S. A. 56, 1275–1282.

Penhoet, E.E., Kochman, M., Rutter, W.J., 1969. Molecular and catalytic properties of aldolase C. Biochemistry 8, 4396–4402.

Peterson, R.N., Freund, M., 1969. Glycolysis by washed suspensions of human spermatozoa. Effect of substrate, substrate concentration, and changes in medium composition on the rate of glycolysis. Biol. Reprod. 1, 238–246.

Pezza, J.A., Choi, K.H., Berardini, T.Z., Beernink, P.T., Allen, K.N., Tolan, D.R., 2003. Spatial clustering of isozyme-specific residues reveals unlikely determinants of isozyme specificity in fructose-1,6-bisphosphate aldolase. J. Biol. Chem. 278, 17307–17313.

Rutter, W.J., Blostein, R.E., Woodfin, B.M., Weber, C.S., 1963. Enzyme variants and metabolic diversification. Adv. Enzyme Regul. 17, 39–56.

Rutter, W.J., Rajkumar, T., Penhoet, E., Kochman, M., Valentine, R., 1968. Aldolase variants: structure and physiological significance. Ann. N. Y. Acad. Sci. 151, 102–117.

Sakai, H., Koyanagi, K.O., Imanishi, T., Itoh, T., Gojobori, T., 2007. Frequent emergence and functional resurrection of processed pseudogenes in the human and mouse genomes. Gene 389, 196–203.

Schmidt, E.E., Schibler, U., 1995. High accumulation of components of the RNA polymerase II transcription machinery in rodent spermatids. Development 121, 2373–2383.

Schmidt, E.E., Ohbayashi, T., Makino, Y., Tamura, T., Schibler, U., 1997. Spermatid-specific overexpression of the TATA-binding protein gene involves recruitment of two potent testis-specific promoters. J. Biol. Chem. 272, 5326–5334.

Schultz, N., Hamra, F.K., Garbers, D.L., 2003. A multitude of genes expressed solely in meiotic or postmeiotic spermatogenic cells offers a myriad of contraceptive targets. Proc. Natl. Acad. Sci. U. S. A. 100, 12201–12206.

Sergeant, K.A., Bourgeois, C.F., Dalgliesh, C., Venables, J.P., Stevenin, J., Elliott, D.J., 2007. Alternative RNA splicing complexes containing the scaffold attachment factor SAFB2. J. Cell Sci. 120, 309–319.

Shima, J.E., McLean, D.J., McCarrey, J.R., Griswold, M.D., 2004. The murine testicular transcriptome: characterizing gene expression in the testis during the progression of spermatogenesis. Biol. Reprod. 71, 319–330.

Steinke, D., Hoegg, S., Brinkmann, H., Meyer, A., 2006. Three rounds (1R/2R/3R) of genome duplications and the evolution of the glycolytic pathway in vertebrates. BMC Biol. 4, 16.

Sullivan, D.T., MacIntyre, R., Fuda, N., Fiori, J., Barrilla, J., Ramizel, L., 2003. Analysis of glycolytic enzyme co-localization in *Drosophila* flight muscle. J. Exp. Biol. 206, 2031–2038.

Tsuruta, J.K., Eddy, E.M., O'Brien, D.A., 2000. Insulin-like growth factor-II/cation-independent mannose 6-phosphate receptor mediates paracrine interactions during spermatogonial development. Biol. Reprod. 63, 1006–1013.

Vandeberg, J.L., Lee, C.Y., Goldberg, E., 1981. Immunohistochemical localization of phosphoglycerate kinase isozymes in mouse testes. J. Exp. Zool. 217, 435–441.

Vertessy, B.G., Orosz, F., Kovacs, J., Ovadi, J., 1997. Alternative binding of two sequential glycolytic enzymes to microtubules. Molecular studies in the phosphofructokinase/aldolase/microtubule system. J. Biol. Chem. 272, 25542–25546.

Vinckenbosch, N., Dupanloup, I., Kaessmann, H., 2006. Evolutionary fate of retroposed gene copies in the human genome. Proc. Natl. Acad. Sci. U. S. A. 103, 3220–3225.

Welch, J.E., Schatte, E.C., O'Brien, D.A., Eddy, E.M., 1992. Expression of a glyceraldehyde 3-phosphate dehydrogenase gene specific to mouse spermatogenic cells. Biol. Reprod. 46, 869–878.

Welch, J.E., Brown, P.L., O'Brien, D.A., Magyar, P.L., Bunch, D.O., Mori, C., Eddy, E.M., 2000. Human glyceraldehyde 3-phosphate dehydrogenase-2 gene is expressed specifically in spermatogenic cells. J. Androl. 21, 328–338.

Westhoff, D., Kamp, G., 1997. Glyceraldehyde 3-phosphate dehydrogenase is bound to the fibrous sheath of mammalian spermatozoa. J. Cell Sci. 110, 1821–1829.

Williams, A.C., Ford, W.C., 2001. The role of glucose in supporting motility and capacitation in human spermatozoa. J. Androl. 22, 680–695.

Wojtas, K., Slepecky, N., von Kalm, L., Sullivan, D., 1997. Flight muscle function in *Drosophila* requires colocalization of glycolytic enzymes. Mol. Biol. Cell 8, 1665–1675.

Yamada, S., Nakajima, H., Kuehn, M.R., 2004. Novel testis- and embryo-specific isoforms of the phosphofructokinase-1 muscle type gene. Biochem. Biophys. Res. Commun. 316, 580–587.

Yeo, G., Holste, D., Kreiman, G., Burge, C.B., 2004. Variation in alternative splicing across human tissues. Genome Biol. 5, R74.

Zhang, Z., Gerstein, M., 2004. Large-scale analysis of pseudogenes in the human genome. Curr. Opin. Genet. Dev. 14, 328–335.

Zhang, Z., Carriero, N., Gerstein, M., 2004. Comparative analysis of processed pseudogenes in the mouse and human genomes. Trends Genet. 20, 62–67.