Evolution of Vertebrate Solute Carrier Family 9B Genes and Proteins (SLC9B): Evidence for a Marsupial Origin for Testis Specific SLC9B1 from an Ancestral Vertebrate SLC9B2 Gene

Roger S Holmes1*,2, Kimberly D Spradling-Reeves2 and Laura A Cox2

1Eskitis Institute for Drug Discovery and School of Natural Sciences, Griffith University, Nathan, QLD, Australia
2Department of Genetics and Southwest National Primate Research Center, Texas Biomedical Research Institute, San Antonio, Texas, USA

Abstract

SLC9B genes and proteins are members of the sodium/lithium hydrogen antiporter family which function as solute exchangers within cellular membranes of mammalian tissues. SLC9B2 and SLC9B1 amino acid sequences and structures and SLC9B-like gene locations were examined using bioinformatic data from several vertebrate genome projects. Vertebrate SLC9B2 sequences shared 50-98% identity as compared with ~50% identities with mammalian SLC9B1 sequences. Sequence alignments, key amino acid residues and conserved predicted transmembrane structures were also studied. Mammalian SLC9B2 and SLC9B1 genes usually contained 11 or 12 coding exons with differential tissue expression patterns: SLC9B2, broad tissue distribution; and SLC9B1, being testis specific. Transcription factor binding sites and CpG islands within the human SLC9B2 and SLC9B1 gene promoters were identified. Phylogenetic analyses suggested that SLC9B1 originated in an ancestral marsupial genome from a SLC9B2 gene duplication event.

Keywords: Solute carrier; SLC9B2, SLC9B1; NHA2; NHA1; Sodium/hydrogen exchanger; Testis-specific; Transmembrane; Marsupial origin; Gene evolution

Abbreviations: SLC: Sodium Lithium Countertransport Activity; NHE: Sodium Hydrogen Exchanger; BLAST: Basic Local Alignment Search Tool; BLAT: Blast-Like Alignment Tool; NCBI: National Center for Biotechnology Information; KO: Knock Out; GI: Gastro-Intestinal; AceView: NCBI Based Representation of Public mRNAs; SWISS-MODEL: Automated Protein Structure Homology-modeling Server

Introduction

SLC9/NHE gene family members (sodium lithium carrier 9/ sodium hydrogen exchanger) contribute to the maintenance of cellular and intracellular pH homeostasis and are predominantly responsible for the absorption of Na+ ions in the kidney and the GI tract [1-3]. SLC9 comprises three subfamilies: SLC9A, with 9 members (SLC9A1-A9), of which SLC9A1-9A5 are localized in plasma membranes and SLC9A6-SLC9A9 in subcellular organelles; SLC9B, with two members, SLC9B1 and SLC9B2 [2], which are the subject of this study; and SLC9C, with two members located in sperm, SLC9C1 and SLC9C2 [2,4,5].

A mitochondrial inner membrane Na+/H+ exchanger (NHA2 or NHEDC2, designated as SLC9B2), was originally identified in yeast (Saccharomyces cerevisiae) and bacteria (Escherichia coli), with the human homologue showing ubiquitous expression, particularly in tissues with high mitochondrial content, including the kidney distal convoluted tubule [6,7]. SLC9B2 has been identified as a contributor to sodium-lithium counter-transport activity (SLC), and may serve as a candidate gene for essential hypertension in human populations, particularly in urban black populations [7,8]. Moreover, genetic linkage studies of SLC activity in baboons have reported association with chromosome 5, the homologue to human chromosome 4 [9]. Human SLC9B2 has also been shown to reverse the Na+/H+ exchanger-null phenotype when expressed in Na+/H+ exchanger deficient yeast [10], which supports a major role for SLC9B2 as a sodium/hydrogen exchanger. A second gene (NHA1 or NHEDC1; designated as SLC9B1) has been located in tandem with SLC9B2 on human chromosome 4 which is specifically expressed in testis in mammals [7,11]. This is in contrast to Drosophila, for which two SLC9B-like genes (designated as NHA1 and NHA2) are widely expressed in epithelial tissues and play crucial roles in organismal ion homeostasis and as Na+/H+ exchangers [12].

This paper reports the predicted gene structures and amino acid sequences for several vertebrate SLCB2 and mammalian SLCB1 genes and proteins, the predicted structures for mammalian SLCB1 and SLCB2 proteins and the structural, phylogenetic and evolutionary relationships for these genes and enzymes. The results suggest that the mammalian SLCB1 gene arose from the duplication event of an ancestral mammalian SLCB2 gene with the appearance corresponding to the emergence of marsupial mammals during vertebrate evolution.

Methods

SLC9B gene and protein identification (SLC9B1 and SLC9B2)

BLAST studies were undertaken using web tools from the National Center for Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov) [13]. Protein BLAST analyses used vertebrate SLC9B2 and mammalian SLC9B1 amino acid sequences previously described [2,7,10,11] (Table 1; Tables S1 and S2). Predicted SLC9B-like protein sequences were obtained in each case and subjected to protein and gene structure analyses.

BLAT analyses were undertaken for each of the predicted SLCB1 and SLCB2 amino acid sequences using the UC Santa Cruz Genome Browser (http://genome.ucsc.edu) with the default settings to obtain the predicted locations for each of the vertebrate SLCB-like genes, including exon boundary locations and gene sizes [14]. The structures for the major human SLCB2 and SLCB1 transcripts were obtained

*Corresponding author: Roger S Holmes, Eskitis Institute for Drug Discovery and School of Natural Sciences, Griffith University, Nathan QLD 4111, Australia, Tel: +61 7 37356008; E-mail: r.holmes@griffith.edu.au

Received May 18, 2016; Accepted June 03, 2016; Published June 10, 2016

Citation: Holmes RS, Spradling-Reeves KD, Cox LA (2016) Evolution of Vertebrate Solute Carrier Family 9B Genes and Proteins (SLC9B): Evidence for a Marsupial Origin for Testis Specific SLC9B1 from an Ancestral Vertebrate SLC9B2 Gene. J Phylogen Evolution Biol 4: 167. doi:10.4172/2329-9002.1000167

Copyright: © 2016 Holmes RS, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
using the AceView website (http://www.ncbi.nlm.nih.gov/ieb/research/acembly/) [15].

Alignments, structures and predicted properties of mammalian SLC9B2 and SLC9B1 proteins

Amino acid sequence alignments of vertebrate SLC9B2 and SLC9B1 proteins were undertaken using Clustal Omega [16]. Predicted secondary and tertiary structures for human and other mammalian SLC9B2 and SLC9B1 proteins were obtained using the SWISS-MODEL web-server [17] and the reported tertiary structures for bacterial (Thermus thermophilus) Na/H antiporter [18] (PDB:4bwza) with modeling residue ranges of 114-512 for human

| Gene     | Species         | RefSeq ID | GenBank ID | UniPROT Amino | Chromosome | Coding Exons | Gene Size | Subunit | p1 | % Identity | % Identity |
|----------|-----------------|-----------|------------|---------------|------------|--------------|-----------|---------|----|-------------|-------------|
| SLC9B2   | -               | NCBI ID   | -          | ID Acids      | Location   | (Strand) bps | MW        | Human   | SLC9B2 | Human SLC9B1 |
| Human    | Homo sapiens    | NM_178833.5 | BC009732 | Q8UD5       | 537 | 4:103,947,530-103,988,707 | 11 (-ye) | 41,178   | 57,564 | 6.3 | 100 | 53         |
| Rhesus   | Macaca mulatta  | 'XP_001110438.1 | CH471198 | F6QDG5       | 537 | 597,274,361-97,312,198 | 11 (-ye) | 37,838   | 57,468 | 6   | 98 | 52         |
| Mouse    | Mus musculus    | IXP_519445 | BC009977 | Q5BK2R2     | 537 | 3:135,317,104-135,336,575 | 12 (+ve) | 19,472   | 58,934 | 6.8 | 81 | 55         |
| Rat      | Rattus norvegicus | NM_001109385.2 | BC167815 | B2RYK8      | 537 | 2:259,014,927-259,035,569 | 12 (+ve) | 20,643   | 58,738 | 6.5 | 82 | 53         |
| Pig      | Sus scrofa      | NP_003129331.2 | GAC001000293 | F1S111     | 537 | 8:126,983,541-127,018,193 | 11 (+ve) | 34,653   | 57,298 | 5.7 | 84 | 54         |
| Opossum  | Monodelphis domestica | NP_007495925.1 | na | F7ESXS      | 656 | 5:46,800,244-46,865,285 | 12 (+ve) | 65,042   | 71,611 | 5   | 52 | 42         |
| Chicken  | Gallus gallus   | 'XP_004936029.1 | na | na          | 568 | 4:60,564,448-60,583,177 | 12 (-ve) | 18,730   | 60,576 | 5.1 | 66 | 53         |
| Zebra fish | Dania rerio    | 'XP_699710.2 | na | 564 | 1:44,078,058-44,094,584 | 12 (-ve) | 16,528   | 6M888 | 8.9 | 56 | 47         |
| Coelacanth | Latimeria chalumnae | NP_00009735.1 | na | na          | 550 | 1H128666:67,204-118,949 | 11 (+ve) | 51,746   | 58,525 | 8   | 65 | 52         |
| Sea squirt | Ciona intestinalis | 'XP_002128513.1 | na | na          | 535 | 14p:1,088,299-1,094,369 | 13 (-ye) | 6,071    | 56,556 | 5.3 | 43 | 41         |
| SLC9B2   | -               | NCBI ID   | -          | ID Acids      | Location   | (Strand) bps | MW        | Human   | SLC9B2 | Human SLC9B1 |
| Human    | Homo sapiens    | NM_139173 | 6CO22079 | 441114     | 515 | 4:103,822,277-103,912,868 | 11 (+ve) | 80,592   | 56,054 | 8.3 | 53 | 100        |
| Rhesus   | Macaca mulatta  | NP_00110263.1 | na | F6URU9      | 511 | 5:97,140,216-97,227,958 | 11 (+ve) | 87,743   | 55,504 | 9   | 51 | 92         |
| Mouse    | Mus musculus    | NM_028946.3 | AK029525 | G5E812      | 565 | 3:135,355,009-135,397,805 | 12 (+ve) | 42,797   | 61,986 | 6.8 | 50 | 64         |
| Rat      | Rattus norvegicus | NM_001173773.1 | na | F1LPN2      | 519 | 2:259,055,735-259,096,954 | 12 (+ve) | 41,220   | 56,581 | 6.4 | 51 | 64         |
| Pig      | Sus scrofa      | na         | CU862009 | F1S110      | 514 | 8:127,051,477-127,103,437 | 11 (+ve) | 51,961   | 55,503 | 8.5 | 53 | 67         |
| Opossum  | Monodelphis domestica | IXP_007496035.1 | na | F6V2D6      | 495 | 2:259,014,927-259,035,569 | 12 (+ve) | 20,643   | 58,738 | 6.5 | 82 | 53         |

RefSeq: the reference amino acid sequence; 'predicted Ensembl amino acid sequence; na-not available; GenBank IDs are derived NCBI http://www.ncbi.nlm.nih.gov/genbank/; Ensembl ID was derived from Ensembl genome database http://www.ensembl.org; UNIPROT refers to UniprotKB/Swiss-Prot IDs for individual proteins (see http://kr.expasy.org); bps refers to base pairs of nucleotide sequences; pl refers to theoretical isoelectric points; the number of coding exons are listed

**Table 1:** Vertebrate SLC9B2 and mammalian SLC9B1 genes and proteins.
Citation: Holmes RS, Reeves KDS, Cox LA (2016) Evolution of Vertebrate Solute Carrier Family 9B Genes and Proteins (SLC9B): Evidence for a Marsupial Origin for Testis Specific SLC9B1 from an Ancestral Vertebrate SLC9B2 Gene. J Phylogen Evolution Biol 4: 167. doi:10.4172/2329-9002.1000167

Page 3 of 8

Results and Discussion

Alignments and possible roles of SLC9B2 amino acid sequences

The deduced amino acid sequences for baboon (Papio anubis), chicken (Gallus gallus), zebra fish (Danio rerio) and coelacanth (Latimeria chalumnae) SLC9B2 proteins are shown in Figure 1 together with the previously reported sequence for human and mouse SLC9B2 proteins [2,7] (Table 1). Alignments of human with other vertebrate sequences examined were between 52-98% identical, suggesting that these are likely to be members of the same family of genes, whereas comparisons of sequence identities with vertebrate SLC9B1 proteins exhibited lower levels of sequence identities, suggesting that these are members of a distinct but related gene family (Table 1; Tables S1 and Table S2).

The amino acid sequences for vertebrate SLC9B2 proteins contained between 537 (for human SLC9B2 isoform 1) and 564 (for zebra fish SLC9B2) residues (Figure 1). Previous studies have reported several key regions and residues for structurally related SLC9 Na+/H+ exchange proteins. These included an N-terminus region (residues 1-86), predicted to be external to the inner-mitochondrial membrane; hydrophobic transmembrane segments which may anchor the enzyme to the inner-mitochondrial membrane; and a C-terminal region (residues 513-537), predicted to be localized within the mitochondrial matrix. The N-terminal and C-terminal regions showed lower levels of amino acid sequence conservation which are in contrast to 13 identified hydrophobic regions (TM1-TM13), which are more highly conserved (Figure 1; Table 2). These regions were identified as transmembrane segments, also designated as 9B2:1/TM1-9B2:13/TM13 hydrophobic zones (Figure 2; Table 2). Several intra-hydrophobic sequences were also highly conserved, including the N-terminus region prior to 9B2:1/TM1 (residues 77-84 designated pre9B2:1); ITM1 (residues 107-112); ITM2 (residues 136-138); ITM4, containing a basic amino acid cluster (residues 193-208); ITM5 (residues 235-240); and ITM6 containing the active site (residues 255-311), particularly the double aspartate sequence (278-279) [23]. In addition, the C-terminus region, immediately following the 9B2:13/TM13 transmembrane sequence (residues 514-519), was highly conserved as well (Table 2). Mutation analyses of a plasmid containing human SLC9B2 expressed in a transfected salt sensitive strain of yeast cells demonstrated reduced ability to remove intracellular sodium when mutants in Val161 and Phe357 were used [23]. These residues were conserved among the mammalian SLC9B2 sequences examined although conservative substitutions were observed in some lower vertebrate SLC9B2 sequences (Figure 1).

The roles for these conserved amino acid sequences located outside of the hydrophobic and transmembrane regions for SLC9B2 remain to be determined; however, it is proposed that the presence of multiple proline and glycine residues in these sequences may contribute to the sharp turns required for maintaining these transmembrane structures, particularly for the ITM1 and ITM6 sequences, which contained double or triple glycine residues (112-113 and 258-260, respectively).
Citation: Holmes RS, Reeves KDS, Cox LA (2016) Evolution of Vertebrate Solute Carrier Family 9B Genes and Proteins (SLC9B): Evidence for a Marsupial Origin for Testis Specific SLC9B1 from an Ancestral Vertebrate SLC9B2 Gene. J Phylogen Evolution Biol 4: 167. doi:10.4172/2329-9002.1000167

| Function      | Name          | Location          | Residues | Sequence              |
|---------------|---------------|-------------------|----------|-----------------------|
| N-terminus    | N-terminus    | Outside membrane  | 1...86   | See Figure 1          |
| Book end to TM1| PreTM1        | Outside membrane  | 77...84  | CPPHRGLLDR           |
| TM helix 1    | TM1/9B2:1     | Transmembrane     | 87...107 | VITNVTIVLLWAVWSITGS  |
| Interhelix spacer | ITM1    | Inside membrane   | 108...114| ECLPGGN              |
| TM helix 2    | TM2/9B2:2     | Transmembrane     | 115...135| LEGIIILFYCAUGKLLGLI  |
| Interhelix spacer | ITM2    | Outside membrane  | 136...138| KLP                  |
| TM helix 3    | TM3/9B2:3     | Transmembrane     | 139...156| TLPLLPSUGMLLAGFLI    |
| Interhelix spacer | ITM3    | Inside membrane   | 157...171| RNIPVINDVVQHKHKR     |
| TM helix 4    | TM4/9B2:4     | Transmembrane     | 172...191| WSSUSIALVILVRAGL     |
| Interhelix spacer | ITM4    | Outside membrane  | 192...206| LSDKALKLKKGVCVR      |
| TM helix 5    | TM5/9B2:5     | Transmembrane     | 207...227| LSMGPDVEACTSALLAHYLLG|
| Interhelix spacer | ITM5    | Inside membrane   | 228...230| GLP                  |
| TM helix 6    | TM6/9B2:6     | Transmembrane     | 231...252| WQWGFILGFVLGAVSPAVVPS|
| Active site zone | ITM6    | Outside membrane  | 253...279| PALLIQGGYGVEGKVPRIMAAGSFDD|
| TM helix 7    | TM7/9B2:7     | Transmembrane     | 280...294| ILAITGFTNCGLIAF      |
| Interhelix spacer | ITM7    | Inside membrane   | 295...305| STGSTVSTGTVFVAR      |
| TM helix 8    | TM8/9B2:8     | Transmembrane     | 306...327| GVELEVIGVATGVSLGFLIFICILF|
| Interhelix spacer | ITM8    | Outside membrane  | 328...340| PSQRQDKLVCKRT        |
| TM helix 9    | TM9/9B2:9     | Transmembrane     | 328...340| FLVLGLSVLAVFSVHGFPGSGGLCTLMAFL|
| Interhelix spacer | ITM9    | Inside membrane   | 341...373| AGMGWTSEKAEVEK       |
| TM helix 10   | TM10/9B2:10   | Transmembrane     | 374...408| HAVAWIDQPLFLGLGAEV   |
| Interhelix spacer | ITM10   | Outside membrane  | 409...419| SIASLRPETYG          |
| TM helix 11   | TM11/9B2:11   | Transmembrane     | 420...440| LCATVGIALVLRTTLMV    |
| Interhelix spacer | ITM11   | Inside membrane   | 441...458| CFAFGLNKKEKISFASIW   |
| TM helix 12   | TM12/9B2:12   | Transmembrane     | 459...476| PKATVQAIGSVALDTAR    |
| Interhelix spacer | ITM12   | Outside membrane  | 477...492| SGHEKQLDVEGMDLTS     |
| TM helix 13   | TM13/9B2:13   | Transmembrane     | 493...513| VAFLLSITAPISSLGLGLG  |
| C-terminus    | postTM13      | Inside membrane   | 514...537| PRLIQKVEHQNKDKEEVEQGETSVGV |

Predicted key residues are identified based upon their positioning within transmembrane and/or inter-transmembrane structures; conserved residues are underlined; predicted Glycine Zipper sequences were identified designated as GZ1 (GFILGFVLG), GZ2 (GGGYGVEKG) and GZ3 (GVATGSVLG); a highly basic amino acid residue region was observed within ITM4, shown in blue; single and double Gly (G) and Pro (P) sequences were identified which may support the formation of transmembrane helices; the Asp-Asp (DD) active site sequence is shown in red.

Table 2: Conserved key residues for human SLC9B2.

---

The transmembrane structures for the human SLC9B2 and SLC9B1 subunits are based on results [31]. Amino acid residues and predicted transmembrane structures are numbered in order from N-terminus as 9B2:1, 9B2:2 etc. for SLC9B2 and 9B1:1, 9B1:2 etc. for human SLC9B1. The arrow indicates position of the active site. Predicted transmembrane regions are shown as red bars. Pink and blue colors indicate external and internal inter-membrane regions.

Figure 2: Comparisons of the predicted transmembrane structures for human SLC9B2 and SLC9B1.
Gene Locations, exonic structures and regulatory sequences for vertebrate SLC9B2 and mammalian SLC9B1 genes

Table 1 and Supplementary Tables S1 and S2 summarize the predicted locations for vertebrate SLC9B2 and mammalian SLC9B1 genes based upon BLAT interrogations of several vertebrate genomes using the reported sequences for human SLC9B2 [7,28,29] and SLC9B1 [7,11] and the predicted sequences for other vertebrate SLC9B2 and mammalian SLC9B1 proteins [11,14]. The vertebrate SLC9B2 and mammalian SLC9B1 genes examined were transcribed on either strand, depending on the species. Figure 1 summarizes the predicted exonic start sites for human, baboon, mouse, chicken, zebra fish and coelacanth SLC9B2 genes, based on the ‘a’ isoform, with each having 11 or 12 coding exons, in identical or similar positions to those predicted for the human SLC9B2 gene.

Figure 4 shows the predicted structures for the major human SLC9B2 and SLC9B1 transcripts. In each case, two major transcripts were observed, including the reference SLC9B2 sequences (AK172823 and BC047447) which were 3601 and 2945 bps in length; and the reference SLC9B1 sequences (BC136966 and AY461581) which were 1879 and 1755 bps in length. The two major human SLC9B2 transcript isoforms, designated as ‘a’ and ‘b’, encode identical proteins with 537 amino acids with 11 coding exons and 1 or 2 non-coding exons, respectively (Table 1; Figure 4). The two major human SLC9B1 transcripts, also designated as ‘a’ and ‘b’, encoded proteins with 515 and 475 amino acids, with 11 and 10 coding exons respectively and 2 non-coding exons (Table 1; Figure 4). The first human SLC9B1 intron is much larger in comparison with intron 1 for human SLC9B2, which explains the observation that primate SLC9B1 genes were double the size for the primate SLC9B2 genes (Table S1; Figure 4).

The human SLC9B2 promoter region contained several potential transcription factor binding sites, including FOXD3, which promotes neural crest development; FOXJ2, which serves as a transcriptional activator; and HNF3B, which is involved in the development of liver, kidney and notochord (Table S3). It would appear that the SLC9B2 gene promoter is well endowed with gene regulatory sequences which may contribute to the high levels of SLC9B2 expression in mammalian neural, liver and kidney cells and to the maintenance of this expression during development. In addition, the human SLC9B2 promoter region contained a large CpG island (CpG92), which may play an important role in developmental regulation.

and the pre9B2:1 region, which contained a double proline sequence (79-80) (Figure 1; Table 2). A distinctive intra-hydrophobic region was observed for ITM4, which exhibited high basic amino acid content, namely K195, K198, K199 and K201. This positive ion cluster may assist in the transfer of salt ions across the inner-mitochondrial membrane or form a salt-ionic with the double aspartate sequence (278-279DD), previously reported as the active site for SLC9B2 [3,7].

Predicted structures for vertebrate SLC9B2 and human SLC9B1

Predicted secondary structures for vertebrate SLC9B2 and human SLC9B1 sequences were examined, which were consistent with the presence of 13 hydrophobic regions, including at least 10 transmembrane structures, symmetrically placed on either side of the double aspartate (278-279DD) active site sequence (Figures 1-3). Predicted tertiary structures for human SLC9B2 and SLC9B1 were also examined which were based on the reported 3D structure for the sodium proton antiporter protein from Thermus thermophilus (also called NapA) (Figure S2) [18]. These transmembrane structures may be compared to pfam02080 domains [24], for which five transmembrane helices are found on either side of a 10 helical structure, predicted to form a pore involved in sodium proton exchange. The aspartate-alanine antiporter (asPT) from Tetratococcus halophilus has also been shown to function with similar domain architecture [25].

A similar secondary structure was observed for human SLC9B1, although with a shorter N-terminus sequence (Figures 2 and 3) and a notable absence of 3 glycine zipper sequences (GlyxxxGlyxxxGly), observed for human SLC9B2 (GZ1: 231-239; GZ2: 257-266; GZ3: 313-321) (Table 2; Figure 3). Glycine zipper sequences have been previously observed for human SLC9B2 and SLC9B1 proteins [11,14]. The vertebrate SLC9B2 and mammalian SLC9B1 genes examined were transcribed on either strand, depending on the species. Figure 1 summarizes the predicted exonic start sites for human, baboon, mouse, chicken, zebra fish and coelacanth SLC9B2 genes, based on the ‘a’ isoform, with each having 11 or 12 coding exons, in identical or similar positions to those predicted for the human SLC9B2 gene.

Figure 4 shows the predicted structures for the major human SLC9B2 and SLC9B1 transcripts. In each case, two major transcripts were observed, including the reference SLC9B2 sequences (AK172823 and BC047447) which were 3601 and 2945 bps in length; and the reference SLC9B1 sequences (BC136966 and AY461581) which were 1879 and 1755 bps in length. The two major human SLC9B2 transcript isoforms, designated as ‘a’ and ‘b’, encode identical proteins with 537 amino acids with 11 coding exons and 1 or 2 non-coding exons, respectively (Table 1; Figure 4). The two major human SLC9B1 transcripts, also designated as ‘a’ and ‘b’, encoded proteins with 515 and 475 amino acids, with 11 and 10 coding exons respectively and 2 non-coding exons (Table 1; Figure 4). The first human SLC9B1 intron is much larger in comparison with intron 1 for human SLC9B2, which explains the observation that primate SLC9B1 genes were double the size for the primate SLC9B2 genes (Table S1; Figure 4).

The human SLC9B2 promoter region contained several potential transcription factor binding sites, including FOXD3, which promotes neural crest development; FOXJ2, which serves as a transcriptional activator; and HNF3B, which is involved in the development of liver, kidney and notochord (Table S3). It would appear that the SLC9B2 gene promoter is well endowed with gene regulatory sequences which may contribute to the high levels of SLC9B2 expression in mammalian neural, liver and kidney cells and to the maintenance of this expression during development. In addition, the human SLC9B2 promoter region contained a large CpG island (CpG92), which may play an important role in developmental regulation.

See Table 1 for sources of human SLC9B2 and SLC9B1 sequences; *shows identical residues for SLC9B2 subunits; similar alternate residues; dissimilar alternate residues; predicted transmembrane residues are shown in blue and numbered in sequence TM1, TM2 etc.; predicted interhelical segments are numbered in sequence (ITM1, ITM2 etc.); active site residues are in green; predicted glycine zipper sequences are shown as GZ1, GZ2 and GZ3 for human SLC9B2; bold underlined font shows residues corresponding to known or predicted exon start sites; exon numbers refer to human SLC9B2 gene exons; proline residues are shaded in pink; glycine residues in turquoise

Figure 3: Amino acid sequence alignments for human SLC9B2 and SLC9B1 sequences.
role in SLC9B2 gene silencing in association with methylation of the gene promoter as reported for other genes during carcinogenesis [30].

SLC9B2 and SLC9B1 tissue expression

Supplementary Figure S2 compares the tissue expression patterns for human and primate SLC9B2 and SLC9B1 genes respectively. Of particular interest are the distinct tissue expression profiles observed for these genes, with SLC9B2 having a wider tissue expression pattern and showing much higher levels of expression in kidney, brain and liver, but with SLC9B1 showing a testis-specific pattern in both human and rhesus macaque tissues. This is consistent with previous studies [2,3,11,14]. The higher levels and wider distribution patterns of expression for SLC9B2 have been associated with the Na(+)/H(+) exchanger, sodium-lithium counter transport, cation/proton antiporter and salt homeostasis roles across tissue organelle lipid bilayers, including the inner mitochondrial membranes of kidney, which has led to this being designated as a candidate gene for essential hypertension [7]. No specific role for the unique testis-specific SLC9B1 expression has been identified [11,31].

Phylogeny and divergence of vertebrate SLC9B2 and mammalian SLC9B1 genes and proteins

A phylogenetic tree (Figure 5) was calculated by the progressive alignment of 26 vertebrate and invertebrate SLC9B2 amino acid sequences with 17 mammalian SLC9B1 sequences which was ‘rooted’ with the coelacanth (Latimeria chalumnae) SLC9B2 sequence (Table 1 and Table S1, S2). The phylogram showing clustering of the SLC9B-like sequences into groups consistent with their evolutionary relatedness as well as groups for the vertebrate SLC9B2 and mammalian SLC9B1 sequences. These groups were significantly different from each other (with bootstrap values >93). It is apparent from this study of vertebrate SLC9B-like genes and proteins that SLC9B2 represents the vertebrate ancestral form while SLC9B1 appeared early in marsupial and eutherian mammalian evolution for which a proposed common ancestor for these genes may have predated or coincided with the appearance of marsupial mammals during vertebrate evolution. In addition, the close localization of the SLC9B2 and SLC9B1 genes on human chromosome 4 as well as their similarities in sequence and exonic structure suggested that the mammalian SLC9B1 gene evolved following a gene duplication event of the SLC9B2 gene, similar to that reported for many gene families [32-34].

Conclusions

The results of the present study indicated that vertebrate SLC9B2 and mammalian SLC9B1 genes and encoded proteins represent a distinct gene and protein family of SLC9B-like proteins which share conserved sequences and active site residues with those reported for SLC9/NHE gene family members (sodium lithium carrier 9/sodium hydrogen exchanger), which contribute to the maintenance of cellular and intracellular pH homeostasis and are predominantly responsible for the absorption of Na+ ions in the kidney and the GI tract [1-3].

SLC9B1 is encoded by a single gene (SLC9B1) among the marsupial and eutherian mammalian genomes studied and are highly expressed in human testis and usually contained 11 or 12 coding exons on the negative or positive strand, depending on the species. SLC9B2 is also encoded by a single gene (SLC9B2) which is more widely expressed in human tissue, with highest levels being observed in kidney, liver and brain, being coexpressed with a proximal gene encoding 3-hydroxybutyrate dehydrogenase (BDH2). Several transcription factor binding sites were localized within the human SLC9B2 gene promoter region, including FOXD3, FOXJ2 and HNF3B, which regulate gene expression, and may contribute significantly to the high level of gene expression in kidney, liver and neural cells. In addition, this region contains a large CpG island which may also contribute to SLC9B2 gene regulation during development.

Predicted secondary and tertiary structures for vertebrate SLC9B2 and mammalian SLC9B1 proteins showed a strong similarity with similar Na/H antisotper originally identified in yeast (Saccharomyces cerevisiae) and bacterial (Thermus thermophilus) Na/H antiporter proteins [17] (PDB:4bwzA). Several major structural domains were apparent for vertebrate SLC9B2 and mammalian SLC9B1 proteins, including a N-terminal tail of varying lengths predicted to be external to the inner-mitochondrial membrane; hydrophobic transmembrane segments which may anchor the enzyme to the inner-mitochondrial membrane; and a C-terminal region (residues 513-537), predicted to be localized within the mitochondrial matrix. The N-terminal and C-terminal regions showed lower levels of amino acid sequence conservation which are in contrast to 13 identified hydrophobic regions (9B2:1-9B2:13), which are more highly conserved.

A phylogenetic study used 17 mammalian SLC9B1 and 26 SLC9B2 protein sequences which indicated that the SLC9B1 gene appeared early in marsupial mammalian evolution, prior to or coincident with the appearance of marsupial mammals, and derived from a gene
The tree is labeled with the SLC9B-like name and the name of the animal and is ‘rooted’ with the coelacanth (Latimeria chalumnae) SLC9B2 sequence, which was used to ‘root’ the tree. Note the 2 major groups corresponding to the SLC9B2 and SLC9B1 gene families. A genetic distance scale is shown. The number of times a clade (sequences common to a node or branch) occurred in the bootstrap replicates are shown. Only replicate values of 0.9 or more, which are highly significant, are shown with 100 bootstrap replicates performed in each case. A proposed duplication event is shown arising from an ancestral invertebrate SLC9B2-like gene, generating the mammalian SLC9B1 gene.

Figure 5: Phylogenetic tree of vertebrate SLC9B2 and mammalian SLC9B1 amino acid sequences.

duplication event of the ancestral vertebrate (and invertebrate)

SLC9B2 gene, which originated much earlier in evolution, since it has been reported in bird, reptile, amphibian, and fish genomes.

Acknowledgement

Grant support acknowledgements (to LAC): C06 RR013556, R01 HL118556 and P51 OD011133.

Disclosure

No conflicts of interest are reported.

References

1. Orlowski J, Grinstein S (2004) Diversity of the mammalian sodium/proton exchanger SLC9 gene family. Pflugers Arch 447: 549-565.
2. Brett CL, Donowitz M, Rao R (2005) Evolutionary origins of eukaryotic sodium/proton exchangers. Am J Physiol Cell Physiol 288: C223-239.
3. Donowitz M, Ming Tse C, Fuster D (2013) SLC9/NHE gene family, a plasma membrane and organelar family of Na+/H+ exchangers. Mol Aspects Med 34: 236-251.
4. Wang D, King SM, Quill TA, Doolittle LK, Garbers DL (2003) A new sperm-specific Na+/H+ exchanger required for sperm motility and fertility. Nat Cell Biol 5: 1117-1122.
5. Liu T, Huang JC, Zuo WL, Lu CL, Chen M, et al. (2010) A novel testis-specific Na+/H+ exchanger is involved in sperm motility and fertility. Front Biosci 2: 566-581.
6. Numata M, Pettecca K, Lake N, Orlowski J (1998) The identification of a mitochondrial Na+/H+ exchanger. J Biol Chem 272: 6951-6959.
7. Xiang M, Feng M, Muend S, Rao R (2007) A human Na+/H+ antiporter sharing evolutionary origins with bacterial NhaA may be a candidate gene for essential hypertension. Proc Natl Acad Sci U S A 104: 18677-18681.
8. Lindhorst J, Alexander N, Bilgaut J, Rayner B (2007) Differences in hypertension between blacks and whites: an overview. Cardiovasc J Afr 18: 241-247.
9. Kammerer CM, Cox LA, Mahaney MC, Rogers J, Shade RE (2001) Sodium-lithium countertransport activity is linked to chromosome 5 in baboons. Hypertension 37: 398-402.
10. Fuster DG, Alexander RT (2014) Traditional and emerging roles for the SLC9 Na+/H+ exchangers. Pflugers Arch 466: 61-76.
Citation: Holmes RS, Reeves KDS, Cox LA (2016) Evolution of Vertebrate Solute Carrier Family 9B Genes and Proteins (SLC9B): Evidence for a Marsupial Origin for Testis Specific SLC9B1 from an Ancestral Vertebrate SLC9B2 Gene. J Phylogen Evolution Biol 4: 167. doi:10.4172/2329-9002.1000167

11. Ye G, Chen C, Han D, Xiong X, Kong Y, et al. (2006) Cloning of a novel human NHEDC1 (Na+/H+ exchanger like domain containing gene expressed specifically in testis. Mol Biol Rep 33: 175-180.

12. Chintapalli VR, Kato A, Henderson L, Hirata T, Woods DJ, et al. (2015) Transport proteins NHA1 and NHA2 are essential for survival, but have distinct transport modalities. Proc Natl Acad Sci USA 112: 11720-11725.

13. Altschul F, Vyas V, Connfield A, Goodin S, Ravikumar TS, et al. (1990) Basic local alignment search tool. J Mol Biol 215: 403-410.

14. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, et al. (2002) The human genome browser at UCSC. Genome Res 12: 994-1006.

15. Thiery-Mieg D, Thierry-Mieg J (2006) AceView: A comprehensive cDNA-supported gene and transcripts annotation. Genome Biol 7: 1-14.

16. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, et al. (2011) Fast, scalable generation of high quality protein multiple sequence alignments using Clustal Omega. Mol Systems Biol 7: 539.

17. Schwede T, Kopp J, Guex N, Pletsch MC (2003) SWISS-MODEL: An automated protein homology-modelling server. Nucleic Acids Res 31: 3381-3385.

18. Lee C, Kang HJ, von Ballmoos C, Newstead S, Uzdavinys P, et al. (2013) A two-domain elevator mechanism for sodium/proton antiport. Nature 501: 573-577.

19. Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, et al. (2011) CDD: a conserved domain database for the functional annotation of proteins. Nucleic Acids Res 39: D338-D345.

20. Dereeher A, Guignon V, Blanc G, Audic S, Buffet S, et al. (2008) Phylogeny, fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res 36: W465-W469.

21. Edgar RC (2004) MUSCLE: a multiple sequence alignments method with a reduced time and space complexity. BMC Bioinformatics 5: 113.

22. Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52: 696-704.

23. Huang X, Morse LR, Xu Y, Zahradka J, Schrora H, et al. (2010) Mutational analysis of NHAcc/NHA2 in Saccharomyces cerevisiae. Biochim Biophys Acta 1800: 1241-1247.

24. Finn RD, Miller BL, Clements J, Bateman A (2014) iPfam: a database of protein family and domain interactions found in the Protein Data Bank. Nucleic Acids Res 42: D364-D373.

25. Nanatani K, Ohonishi F, Yoneyama H, Nakajima T, Abe K (2005) Membrane topology of the electrogenic aspartate-alanine antiporter AspT of Tetragnococcus halophilus. Biochem Biophys Res Commun 328: 20-26.

26. Kim S, Jeon TJ, Oberai A, Yanf D, Schmidt JJ, et al. (2005) Transmembrane glycine zippers: physiological and pathological roles in membrane proteins. Proc Natl Acad Sci USA 102: 14278-14283.

27. Dong H, Sharma M, Zhou HX, Cross TA (2012) Glycines: roles in -helical membrane protein structures and a potential indicator of native conformation. Biochemistry 51: 4779-4789.

28. Yun CH, Tse CM, Nath SK, Levine SA, Brant SR (1995) Mammalian Na+/H+ exchanger gene family: structure and function studies. Am J Physiol 269: G1-G11.

29. Malakooti J, Dahdal RY, Schmidt L, Loyden T, Dudeja PK, et al. (1999) Molecular cloning, tissue distribution, and functional expression of the human Na+(+)/H+(+) exchanger NHE2. Am J Physiol 277: G383-G390.

30. Kang XC, Chen ML, Yang F, Gao BQ, Yang QH, et al. (2016) Promoter methylation and expression of SOCS1 affect clinical outcome and epithelial-mesenchymal transition in colorectal cancer. Biomed Pharmacother 80: 23-29.

31. Tusnady GE, Simon I (2001) The HHMTPM transmembrane topology prediction server. Bioinformatics 17: 849-850.

32. Holmes RS (2013) Evolution of mammalian KELL blood group glycoproteins and genes (KEL): evidence for a marsupial origin from an ancestral M13 Type II endopeptidase gene. Phyl Ecol Biol 1: 112.

33. Storz JF (2016) Gene duplication and evolutionary innovations in hemoglobin-oxygen transport. Physiology 31: 223-232.

34. Vervoort M, Muelemesser D, Behague J, Kemler P (2016) Evolution of Prdm genes in animals: insights from comparative genomics. Mol Biol Evol 33: 679-698.