**Supplementary Table**

**Table S1: Primate *SLC9B2* and *SLC9B1* Genes and Proteins**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Gene** | **Species** | **RefSeq ID** | **UNIPROT ID** | **Amino acids** | **Chromosome Location** | **Coding Exons** | **Coding Exons** | **Subunit** | **pl** | **% Identity** | **% Identity** |
|  |  | **1NCBI ID** |  |  |  | **(Strans)** | **Gene size bps** |  |  | **Human SLC9B2** | **Human SLC9B1** |
| **SLC9B2** |  |  |  |  |  |  |  |  |  |  |  |
| Human | Homo sepians | NM\_178833.5 | Q86UD5 | 537 | 4:103,947,530-103,988,707 | 11 (-Ve) | 41,178 | 57,654 | 6.3 | 100 | 53 |
| Chimp | Pan troglodytes | NM\_001246497.1 | G2HJC8 | 537 | 4:105,574,811-105,616,209 | 11 (-Ve) | 41,339 | 57,666 | 61 | 99 | 53 |
| Gorilla | Gorilla gorilla | 1XP\_0040402661 | G3RED2 | 537 | 4:122,287,801-112,330,017 | 11 (-Ve) | 42,217 | 57,556 | 6.4 | 99 | 53 |
| Orangutan | Pango abelii | NM\_001133107.1 | Q5R6B8 | 537 | 4:107,488,639-107,532,144 | 11 (-Ve) | 43,506 | 57,541 | 6.1 | 99 | 53 |
| Gibbon | Nomascus leucogenys | 1XP\_003257515.1 | M3ZA39 | 537 | 9:87,878,572-87,918,230 | 11 (-Ve) | 39,659 | 57,570 | 6.3 | 99 | 53 |
| Rhesus | Macaca mulatta | 1XP\_001110438.1 | F6QDG5 | 537 | 5:97,274,361-97,312,198 | 11 (-Ve) | 37,838 | 57,468 | 6 | 99 | 55 |
| Baboon | Papio anubis | 1XP\_003899085.1 | A0A096MYF0 | 537 | 5:94,380,118-94,418,540 | 11 (-Ve) | 38,423 | 57,510 | 6.2 | 88 | 55 |
| Sqirrel Monkey | Saimiri boliviensis | 1XP\_010338407.1 | NA | 537 | 1jh378138:12,104,483-12,146,376 | 11 (+Ve) | 41,894 | 57,633 | 5.5 | 96 | 55 |
| Marmoset | Callithrix jacchus | 1XP\_008991153.1 | F6UF28 | 537 | 3:90,593,938-90,644,360 | 11 (+Ve) | 50,423 | 57,497 | 5.6 | 96 | 55 |
| **SLC9B1** |  |  |  |  |  |  |  |  |  |  |  |
| Human | Homo sepians | NM\_139173 | Q4ZJI4 | 515 | 4:103,822,277-103,912,868 | 11 (-Ve) | 90,592 | 56,054 | 8.3 | 53 | 100 |
| Chimp | Pan troglodytes | 1XP\_0094419191.1 | H2QPZ1 | 515 | 4:105,449,789-105,540,181 | 11 (-Ve) | 90,393 | 56,034 | 8.3 | 54 | 99 |
| Orangutan | Pango abelii | 1XP\_002815058.1 | H2PE09 | 515 | 4:107,357,812-107,543,080 | 11 (-Ve) | 95,629 | 56,231 | 8.5 | 55 | 97 |
| Gibbon | Nomascus leucogenys | 1XP\_003257516.1 | NA | 515 | 9:81,751,151-87,843,907 | 11 (-Ve) | 92,757 | 56,190 | 8.6 | 54 | 95 |
| Rhesus | Macaca mulatta | 1XP\_001110263.1 | F6URU9 | 511 | 5:97,140,216-97,227,958 | 11 (-Ve) | 87,743 | 55,504 | 9 | 51 | 92 |
| Baboon | Papio anubis | 1XP\_003899074.1 | A0A096MS92 | 514 | 5:94,250,012-94,339,835 | 11 (-Ve) | 89,824 | 55,775 | 9 | 54 | 93 |
| Sqirrel Monkey | Saimiri boliviensis | 1XP\_010338410.1 | NA | 511 | 1jh378138:12,180,688-12,238,879 | 11 (+Ve) | 58,192 | 55,195 | 8.6 | 56 | 87 |
| Marmoset | Callithrix jacchus | 1XP\_002745582.1 | NA | 516 | 3:90.677.898-90.745.779 | 11 (+Ve) | 67,882 | 56,112 | 8.2 | 55 | 87 |

RefSeq: the reference amino acid sequence; ¹predicted 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); the number of coding exons are listed; bps refers to base pairs of nucleotide sequences; pI refers to theoretical isoelectric points.

**Table S2: Vertebrate and Invertebrate *SLC9B2*, Mammalian *SLC9B1* and Bacterial *SLC9B*-Like Genes and Proteins**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Gene** | **Spades** | **RefSeqID** | **UNIPROT** | **Amino acids** | **Chromosome** | **Coding Dons** | **Alternate** |
| ***SLC982*** |  | **'NCBI ID** | **ID** |  | **location** | **(strand)** | **Name** |
| Rat | *Rotten norveglcus* | NP 001102855.1 | B2RYK8 | 547 | 2:259,014,927-259,035,569 | 12 *(\*ye)* | *NHA2* |
| Guinea pig | *Covio porcellus* | 'XP 003468081.2 | HOVGR2 | 536 | * 43:3,728,811.3,753,215 | 11 *(-ye)* | *NHA2* |
| Cow | *80$ toms* | NM\_001193108.1 | E18FC4 | 535 | 6:23,339,053.23,382,660 | 11 (eire) | *NHA2* |
| Sheen | *Ovis arks* | 1XP 004009709.1 | WSPQB3 | 535 | 6:21,912,060-21,955,129 | *11 (we)* | *NHA2* |
| Rabbit | *Oryctologus cunkulus* | 'XP\_002717252.1 | 615N82 | 533 | 15:46549,60446,589,566 | 11 *(\*ye)* | *NHA2* |
| Dog | *Canis tom/Ports* | 'XD 003640158.1 | na | 537 | 32:24,360,178-24,405,483 | 11 *(-ye)* | *NHA2* |
| Cat | *Felix cot us* | 'XP\_003985172 | M3X5X0 | 534 | 81:119,467,346-119,503,460 | 11 *(\*ve)* | *NHA2* |
| Bat | *Myotis lucifugus* | 'XD 006094161.1 | G1P9W5 | 537 | '61429896:2,424,536-2,452,761 | 11 (-ve) | *NHA2* |
| Turtle | *Ouysmys pieta beat* | 'XP\_005287919.1 | na | 544 | * 111584529:187,678.206,874 | 11 *(are)* | *NHA2* |
| Frog | *Xenopus tropkolis* | NM 001100259.1 | A4IHB9 | 539 | * GL172747:1,218,051-1,240,083 | *11(•ve)* | *NHA2* |
| -Mania | *Oreochromls nilotkus* | 'X0005458788 | na | 576 | * 1.66:15,097,428-15,108,293 | 12 *(we)* | *NHA2* |
| Fruit *fly* | *Drosophila melonogoster* | NM\_135236.3 | 095U27 | 550 | 21:6,861,037.6,862,823 | 3 (At) | *NHAI* |
|  |  | NM 170043 | 49VC57 | 584 | 3R:22,981,459.22,993,561 | 9 *(-ye)* | *NHA2* |
| *SLC981* |  |  |  |  |  |  |  |
| Rat | *Rattus norvegkus* | NM 001173773.1 | F1LPN2 | 519 | 2:259,055,735-259,096,954 | 12 *(•ve)* | *NHA I* |
| Guinea pig | *Covia pa/calks* | 'XP 003468080.2 | HOVTV1 | 516 | '43:3,665,355.3,703,908 | 11 *(.ve)* | *NHAI* |
| Sheep | *Ovis cries* | 'XD 004023736.1 | WSPQW1 | 513 | 6:21,993,146-22,040,539 | *11(we)* | *NHA* |
| Rabbit | *Otyctologus cunkulus* | 'XP 008265775.1 | G1SPF2 | 513 | 15:46,632,43146,709,867 | *11 (we)* | *NHA I* |
| Dog | *Canis familiar's* | na | F1P170 | 512 | 32:24,292,982-24,442,179 | 11 (-ve) | *NHAI* |
| Cat | *Felix cotus* | 'XP 003985215.1 | M3X698 | 525 | 81:119,536,829-119,593,061 | 11 (eve) | *NHAl* |
| Horse | *Equus cobollus* | 'AP\_005608644.1 | F6SVC3 | 513 | 3:37,008,963-37,059,059 | 11 (we) | *NHAI* |
| ***NAPA*** |  |  |  |  |  |  |  |
| Bacterial | *Thymus thermophllus* | WP\_011173522.1 | Q721M4 | 386 | na | na | NAPA |

RefSeq: the reference amino acid sequence; ¹predicted 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 endopeptidase-like proteins (see http://kr.expasy.org); the number of coding exons are listed. Note the presence of 2 *SLC9B*-like genes and proteins from *Drosophila melanogaster*, previously investigated [12].

**Supplementary Table S3:** Identification of Transcription Factor Binding Sites (TFBS) and CpG Islands within the Human *SLC9B2* and *SLC9B1* Gene Promoters

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Human** | **TF8S** | **Name** | **Strand** | **Ow 4 Position** | **Function/ Rol<** | **UNIPROT ID** |
| **Gene** |  |  |  |  |  |  |
| ***SCC982*** | FOXO3 | ferttead box protein 01 | Dve) | 103.979,493496 | transciptIonai reortssolacthator. PIVII/Offt *newel* melt tall peveiopment | CatDus |
|  | FOXI2 | ronA\*44 box paten 1l | Ow) | 103.929444300 | tranvarodonai activator | Q91701011 |
|  | NNF39 | HepatOCieht node./ lector Platte | on) | 103474143497 | Devoopentnt of Men. Neaten, luny, notedoxe | Mead |
|  | MSX1 | Horneobco protein PASX4 | I•val | 103.979404492 | Transcriptional repent\*\*. Acts In crancriecial development | P78340 |
| ***Lass*** | CUM | Honteolfmt protein wail | Om) | 103.944544473 | TraniCripOoma repressoe | Q7Smt4 |
|  | WW2 | Feekteted box protein 22 | Ow) | 10494 1.162 .575 | Trenscriptionelactivata | Q9PDXI |
|  | FOX03 | fortineed box protein 03 | ("wl | 104941,145-368 | TI/aStriptiOnal «Ovate'. | 04132a |
|  | GATA1 | tryowoid transcription factor | 19e) | 103.944194394 | tranuriptkad ottemtor or repromee supponIng MANN'S development | 913976 |
|  | CPG | Gemini< Hie e1C44 Island |  |  |  |  |
| ***SLC982*** | 92 | 799 |  | 103.997.2094.007 |  |  |
| ***SLC981*** | 34 | 213 |  | 103444669451 |  |  |

The identification of TFBS and CpG islands within the *SLC9B2* and *SLC9B1* gene promoter regions was undertaken using the human genome browser (http://genome.ucsc.edu); UNIPROT refers to UniprotKB/Swiss-Prot IDs for individual TFBS sequences (see http://kr.expasy.org).