**SUPPLEMENTARY INFORMATION**

# Evolution of substrate specificity for the bile salt transporter ABST (SLC10A2)

Daniël A. Lionarons[*], James L. Boyer[*†] and Shi-Ying Cai[*]

[*]Department of Internal Medicine and Liver Center, Yale University School of Medicine, New Haven, CT 06520; and [†]Mount Desert Island Biological Laboratory, Salisbury Cove, ME 04672

## Western blotting and immunofluorescent labeling methods

COS-7 cells and HEK293T cells were maintained at low passage number in growth medium. On day 1, cells were seeded on a 12-well plate. On day 2, cells approached confluence and were transiently transfected using Lipofectamine 2000 (Invitrogen). Specifically, for each well in a 12-well plate, 1 μg of pcDNA3-skAsbt-FLAG construct or 1 μg of pcDNA3 empty vector or 25 ng pcDNA3-lpAsbt-FLAG + 975 ng pcDNA3 or 150 ng pcDNA3-lpAsbt-FLAG + 850 ng pcDNA3 were used for transfection. On day 3, cells were harvested. For Western blotting, protein was extracted using M-PER Mammalian Protein Extraction Reagent (Pierce). Fifty μg of protein per lane was resolved in a 10% SDS-PAGE and transferred onto PVDF membrane. An anti-FLAG M2 antibody was used as primary antibody (1:500 dilution), followed by HRP conjugated anti-mouse IgG secondary antibody. The immunocomplex was visualized using ECL reagents and recorded on X-ray film. For immunofluorescent labeling, cells were plated on a poly-Lysine coated coverslip. After transfection, cells were fixed using 4% paraformaldehyde. An anti-FLAG M2 antibody (1:500 dilution) was used as primary antibody, followed by fluorescent-dye (Alex-488) conjugated anti-mouse IgG secondary antibody. Nuclei were stained with Topro-3. Fluorecent images were taken with a Zeiss LSM 510 confocal microscope.

**Table S1.** Oligonucleotide sequences used for molecular cloning and realtime RT-PCR detection of lpAsbt, skAsbt and hASBT.

| Name | Sequence (5´ -> 3´) | Gene |
|---|---|---|
| Slc10a2d-F1 | NTGYCARTTYGGNATHATGCC | Slc10a2 |
| Slc10a2d-R2 | NSWRTADATNARNGGRAANGYRAA | Slc10a2 |
| skAsbt-F9 | TGGAGACATGGATTTAAGCATAAGC | skate Asbt |
| skAsbt-R10 | GAGGCAAAGAGGCATCATTCC | skate Asbt |
| skAsbt-P11 | Fam-TGACCACATGCTCCACAGTGCTGGG | skate Asbt |
| sAsbt-R12 | GCTGAATTCACTTGTCATCGTCATCCTTGTAATCCAGGTTGGTACCTGCGTCC | skate Asbt |
| sAsbt-F13 | CGCATAAGCTTGCTGTGAAACTGCGAAGACCGG | skate Asbt |
| sAsbt-R14 | GCAGAATTCGCATTATTATCTGTAATGACCATGGC | skate Asbt |
| lSlc10a2-F1 | GCCAGTTCGGCTTCATGCCC | lamprey Asbt |
| lSlc10a2-R2 | GAGGCAGAGCGGCATCATCC | lamprey Asbt |
| lSlc10a2-R3 | AAGGCGCTGTAGATGAGCGG | lamprey Asbt |
| lpAsbt-F4 | CAGAACAGCCAGCTGTGCACCACC | lamprey Asbt |
| lpAsbt-F5 | CTTCACTCCCGAGCAGCTCGTGC | lamprey Asbt |
| lpAsbt-F6 | GCTCAAGCTTCCAAGATGCCGTACGAAGACTACGAGG | lamprey Asbt |
| lpAsbt-R7 | GCTGAATTCCCGCAGCGTCTACAGAGAAGTG | lamprey Asbt |
| lpAsbt-R10 | GCTGAATTCTACTTGTCATCGTCATCCTTGTAATCCAGAGAAGTGCCGGCGATCTCCC | lamprey Asbt |
| lampAsbt-F5 | GGACCTCAGCATCAGCATGAC | lamprey Asbt |
| lampAsbt-P6 | Fam-TTCTACGCTGCTGGGCATGG | lamprey Asbt |
| lampAsbt-R7 | AGGCAGAGCGGCATCATC | lamprey Asbt |
| hASBT-F1 | GCAAGGATCCAGCAGCAGACCCAGCAATG | human ASBT |
| hASBT-R2 | GCTGAATTCGTCTGTTTTGTCCACTTGATGTC | human ASBT |

**Table S2.** Amino acid sequence identity among ASBT/Asbt's and other members in SLC10A family. Sequences were aligned using the Clustal Omega algorithm and sequence identity was calculated using the Jalview tool. Accession numbers of sequences used are listed in Table S3. (mAsbt, mouse Asbt; skAsbt, skate Asbt; lpAsbt, lamprey Asbt; hNTCP, human $Na^+$-taurocholate cotransporting polypeptide (SLC10A1); hSOAT, human sodium-dependent organic anion transporter (SLC10A6); nmSlc10a, *N. meningitidis* Slc10a.)

| **% identity** | hASBT | mAsbt | skAsbt | lpAsbt | hSOAT | hNTCP | SLC10A4 | SLC10A5 | SLC10A3 | nmSlc10a |
|---|---|---|---|---|---|---|---|---|---|---|
| mAsbt | 81 | | | | | | | | | |
| skAsbt | 64 | 66 | | | | | | | | |
| lpAsbt | 58 | 57 | 60 | | | | | | | |
| hSOAT | 45 | 48 | 47 | 49 | | | | | | |
| hNTCP | 34 | 32 | 35 | 34 | 32 | | | | | |
| SLC10A4 | 34 | 35 | 35 | 35 | 31 | 35 | | | | |
| SLC10A5 | 28 | 29 | 35 | 27 | 24 | 25 | 23 | | | |
| SLC10A3 | 26 | 26 | 29 | 25 | 28 | 29 | 28 | 37 | | |
| nmSlc10a | 25 | 26 | 26 | 24 | 23 | 24 | 24 | 25 | 26 | |
| SLC10A7 | 21 | 19 | 18 | 18 | 20 | 20 | - | - | - | 20 |

**Table S3.** GenBank, Ensembl genome database (Ensembl release 66) and Protein Data Bank accession numbers of confirmed and predicted protein sequences used for amino acid alignment and phylogenetic analysis.

| Reference name | Gene symbol | Species | Accession number |
|---|---|---|---|
| human_NTCP/SLC10A1 | SLC10A1 | Homo sapiens | NP_003040.1 |
| mouse Ntcp/Slc10a1 | Slc10a1 | Mus musculus | NP_001171032.1 |
| human_ASBT | SLC10A2 | Homo sapiens | NP_000443.1 |
| chimpanzee_Asbt | Slc10a2 | Pan troglodytes | XP_522716.2 |
| dog_Asbt | Slc10a2 | Canis lupus familiaris | NP_001002968.1 |
| rabbit_Asbt | Slc10a2 | Oryctolagus cuniculus | NP_001076233.1 |
| rat_Asbt | Slc10a2 | Rattus norvegicus | NP_058918.1 |
| mouse_Asbt | Slc10a2 | Mus musculus | NP_035518.1 |
| chicken_Asbt | Slc10a2 | Gallus gallus | XP_425589.2 |
| zebrafish_Asbt | Slc10a2 | Danio rerio | NP_956652.1 |
| skate_Asbt | Slc10a2 | Leucoraja erinacea | JX014267 |
| lamprey_Asbt | Slc10a2 | Petromyzon marinus | JX014266 |
| human_SLC10A3 | SLC10A3 | Homo sapiens | NP_062822.1 |
| mouse_Slc10a3 | Slc10a3 | Mus musculus | NP_663381 |
| human_SLC10A4 | SLC10A4 | Homo sapiens | NP_689892.1 |
| mouse_Slc10a4 | Slc10a4 | Mus musculus | NP_775579.2 |
| human_SLC10A5 | SLC10A5 | Homo sapiens | NP_001010893.1 |
| mouse_Slc10a5 | Slc10a5 | Mus musculus | NP_001010834.1 |
| human_SOAT/SLC10A6 | SLC10A6 | Homo sapiens | NP_932069.1 |
| chimpanzee_Soat/Slc10a6 | Slc10a6 | Pan troglodytes | XP_526626.2 |
| dog_Soat/Slc10a6 | Slc10a6 | Canis lupus familiaris | XP_851303.1 |
| mouse_Soat/Slc10a6 | Slc10a6 | Mus musculus | NP_083691.1 |
| cod_Soat/Slc10a6 | Slc10a6 | Gadus morhua | ENSGMOP00000002764 |
| human_SLC10A7 | SLC10A7 | Homo sapiens | NP_001025169.1 |
| mouse_Slc10a7 | Slc10a7 | Mus musculus | NP_084012.1 |
| n.meningitidis_Slc10a | Slc10a | Neisseria meningitidis | 3ZUY |
| ciona_Slc10a-like | Slc10a-like | Ciona intestinalis | ENSCINP00000003667 |
| ciona_Slc10a1-like | Slc10a1-like | Ciona intestinalis | XP_002124965.1 |
| ciona_Slc10a4-like | Slc10a4-like | Ciona intestinalis | ENSCINP00000008688 |
| ciona_Slc10a5-like (#2127302.1) | Slc10a5-like | Ciona intestinalis | XP_002127302.1 |
| ciona_Slc10a4-like (#2127506.1) | Slc10a4-like | Ciona intestinalis | XP_002127506.1 |
| ciona_Slc10a-like (#13851) | Slc10a-like | Ciona intestinalis | ENSCINP00000013851 |
| ciona_Slc10a2-like (#2120408.1) | Slc10a2-like | Ciona intestinalis | XP_002120408.1 |
| ciona_Slc10a2-like (#2120470.1) | Slc10a2-like | Ciona intestinalis | XP_002120470.1 |
| ciona_Slc10a-like (#28590) | Slc10a-like | Ciona intestinalis | ENSCINP00000028590 |
| ciona_Slc10a2-like (#2123186.1) | Slc10a2-like | Ciona intestinalis | XP_002123186.1 |
| ciona_Slc10a-like (#12686) | Slc10a-like | Ciona intestinalis | ENSCINP00000012686 |
| ciona-Slc10a-like (#15302) | Slc10a-like | Ciona intestinalis | ENSCINP00000015302 |
| ciona-Slc10a-like (#22065) | Slc10a-like | Ciona intestinalis | ENSCINP00000022065 |
| ciona_Slc10a7-like (#2131347.1) | Slc10a7-like | Ciona intestinalis | XP_002131347.1 |
| ciona_Slc10a-like (#9747) | Slc10a-like | Ciona intestinalis | ENSCINP00000009747 |
| ciona-Slc10a-like (#22024) | Slc10a-like | Ciona intestinalis | ENSCINP00000022024 |
| human_SLC17A5 | SLC17A5 | Homo sapiens | NP_036566 |

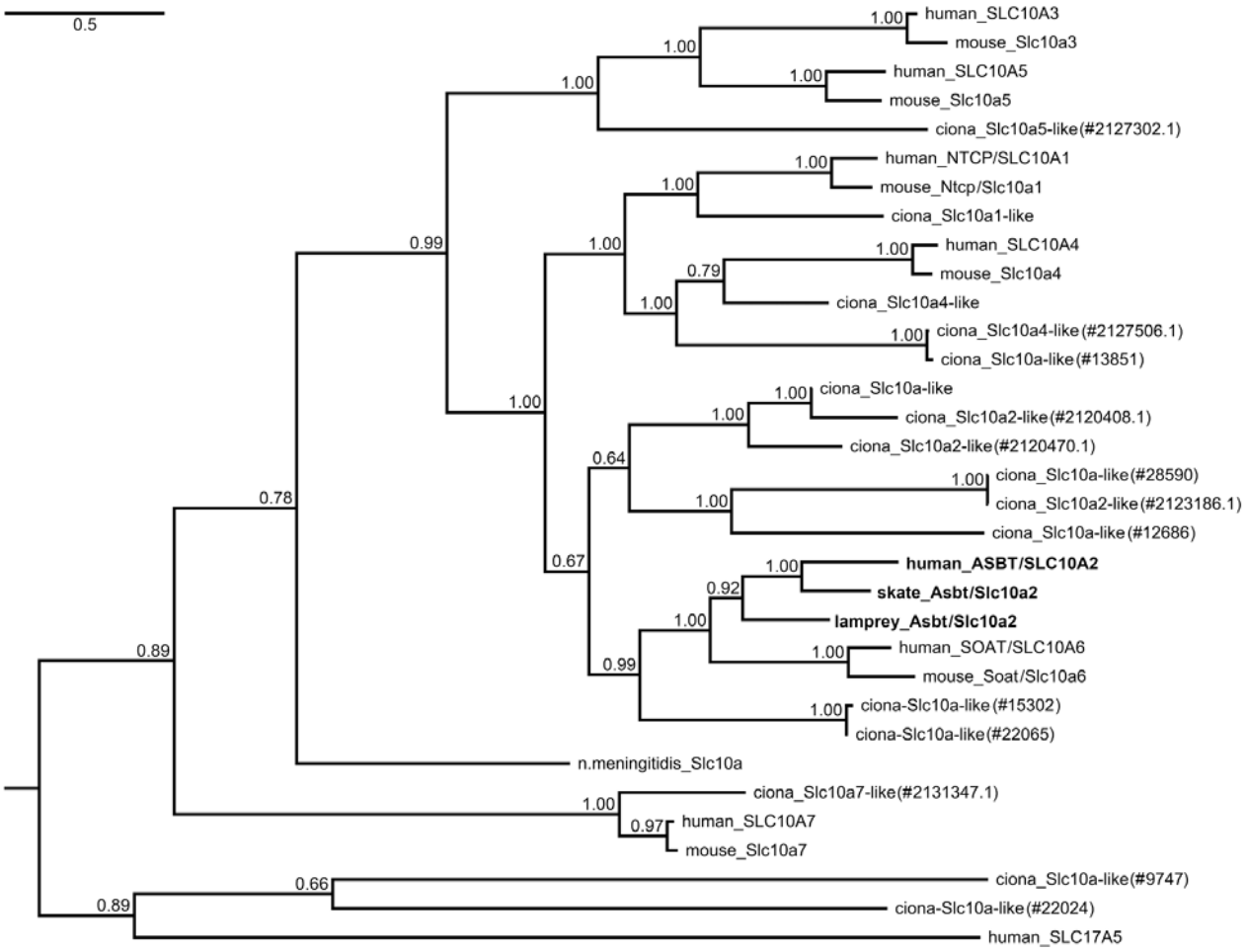


**Figure S1.** Phylogenetic tree of the Slc10a protein family, including all potential ciona Slc10a members retrieved using BLAST searches of the ciona genome. Phylogeny was inferred with Bayesian MCMC analysis. No ciona orthologs of Slc10a2 or Slc10a6 could be identified. However, several sequences were placed at the common branch of Slc10a2 and Slc10a6 subfamilies, suggesting these sequences originate from a common ancestral protein of Slc10a2 and Slc10a6. This leads us to propose that distinct Slc10a2 and Slc10a6 genes emerged from a gene duplication of an ancient Slc10a2/a6-like protein between ciona and lamprey evolution, around the beginning of vertebrate evolution. Posterior probabilities are indicated at nodes and branch length is expressed as number of expected substitutions per site. Accession numbers of analyzed sequences are listed in table S3.

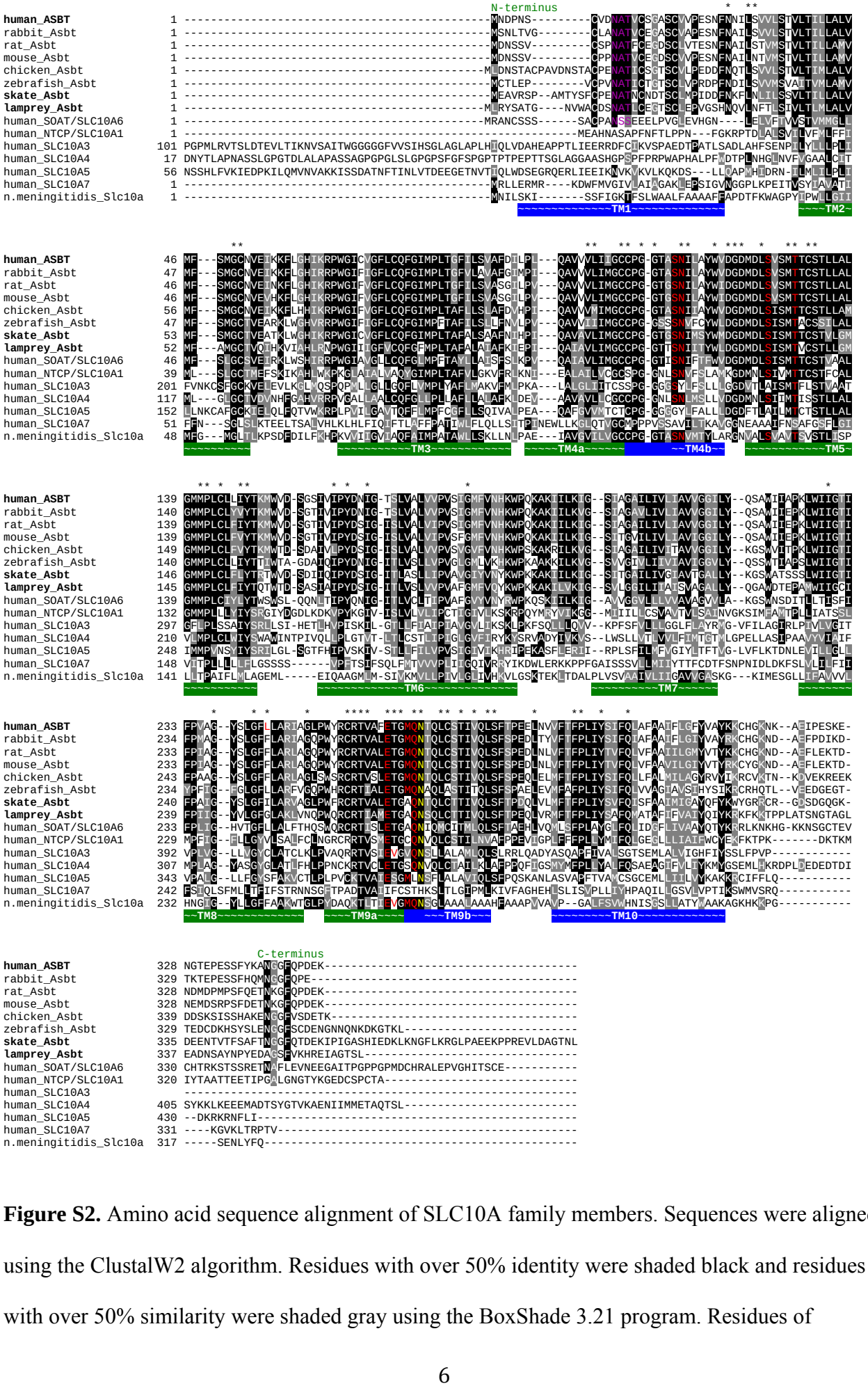


**Figure S2.** Amino acid sequence alignment of SLC10A family members. Sequences were aligned using the ClustalW2 algorithm. Residues with over 50% identity were shaded black and residues with over 50% similarity were shaded gray using the BoxShade 3.21 program. Residues of

hASBT that when mutated abolished $^{3}$H-TCA uptake (<25% of control) while membrane localization was sustained are marked with an asterisk (*) above (supplemental references 1-11). The secondary structure of a bacterial homologue from *Neisseria meningitis* (n.meningitidis Slc10a) is indicated below. Transmembrane helices (TM) colored blue are part of the substrate binding pocket, while TM helices colored green are not implicated in substrate binding. Residues that bind or interact with $Na^{+}$ in Slc10a_n.meningitidis are colored red and Asn-265, shown to interact with the 7-hydroxyl group of TCA in *N. meningitidis* Slc10a, is colored yellow. The N-linked glycosylation site in hASBT is colored purple. The N-termini of SLC10A3, A4 and A5 did not show similarity and are not completely shown. GenBank accession numbers are listed in table S3.

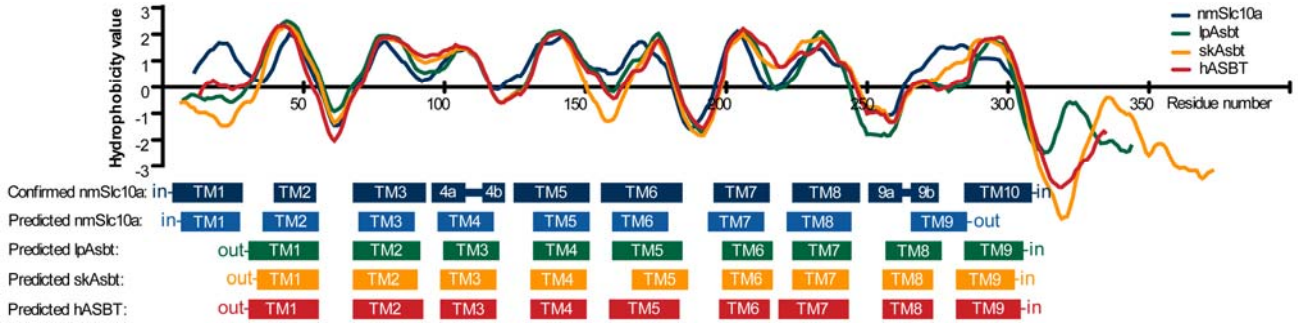


**Figure S3.** Hydrophobicity plot and predicted transmembrane topology of lpAsbt, skAsbt, hASBT and *N. meningitis* Slc10a (nmSlc10a). Protein sequences were aligned with the ClustalW2 algorithm and hydrophobicity values were calculated with the Toppred 0.01 program using the GES-scale (supplemental reference 12). Transmembrane (TM) helices of the crystal structure of nmSlc10a and putative TM helices predicted with the HMMTOP 2.0 program (supplemental reference 13) are indicated below, accompanied by cytosolic (in) or extracellular (out) localization of the N- and C-termini.

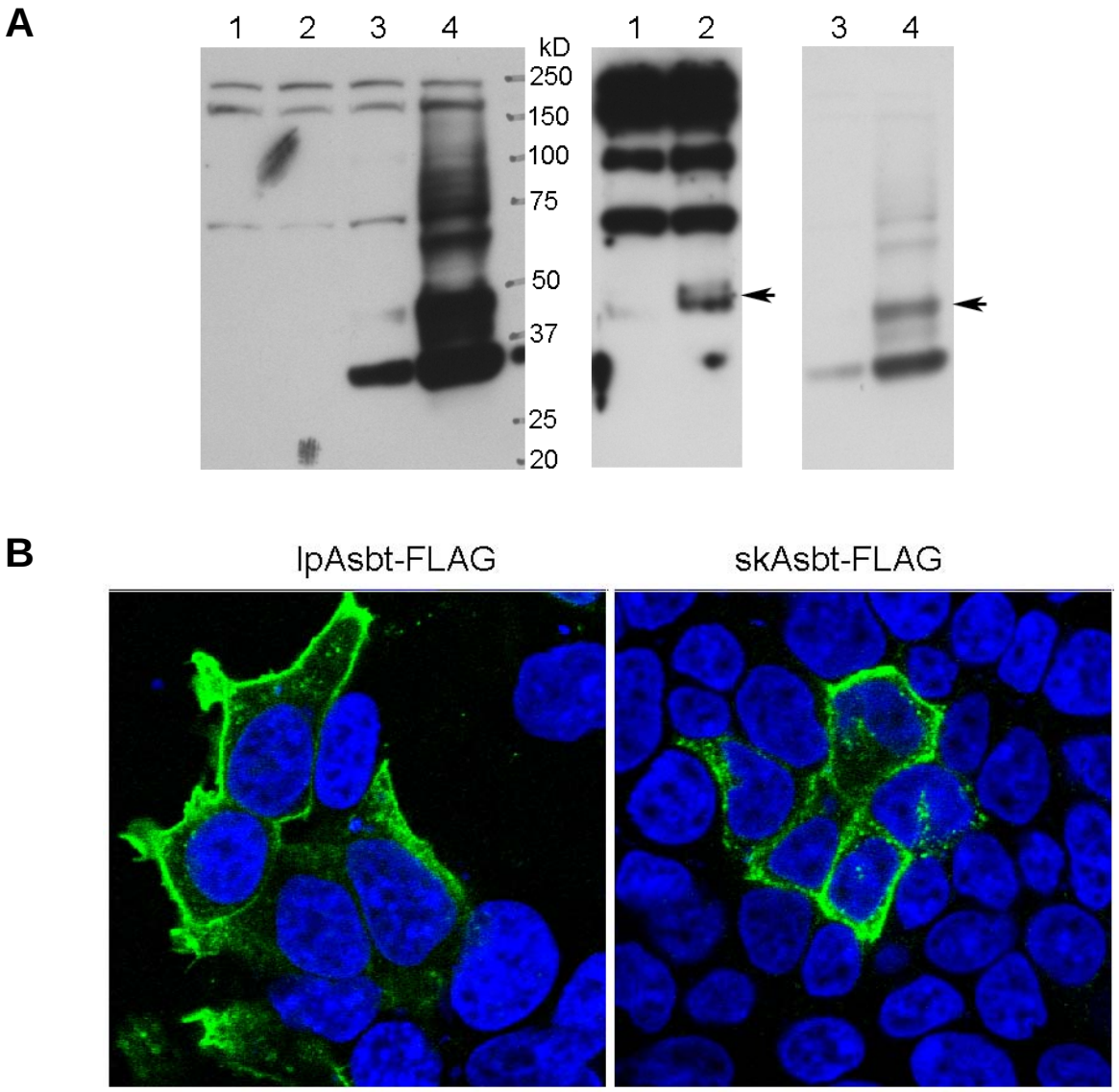


**Figure S4.** Expression of lpAsbt-FLAG and skAsbt-FLAG fusion proteins in transfected cells. (*A*) Western blot of lpAsbt-FLAG and skAsbt-FLAG expressed in transfected HEK293T cells. Anticipated sizes of lpAsbt-FLAG and skAsbt-FLAG were detected (pointed by arrows) in longer (Center) and shorter (Right) exposures. Lane 1, 1 μg pcDNA3 empty vector transfected cells; lane 2, 1 μg pcDNA3-skAsbt-FLAG construct transfected cells; lane 3, 25 ng pcDNA3-lpAsbt-FLAG + 975 ng pcDNA3 transfected cells; lane 4, 150 ng pcDNA3-lpAsbt-FLAG + 850 ng pcDNA3 transfected cells. (*B*) Immunofluorescence confocal microscopy demonstrates that both lpAsbt-FLAG and skAsbt-FLAG (green) can be expressed at least partially on plasma membrane in transfected HEK293T cells. Similar results were also obtained in transfected COS7 cells. Cell nuclei were stained using Topro-3 (blue)

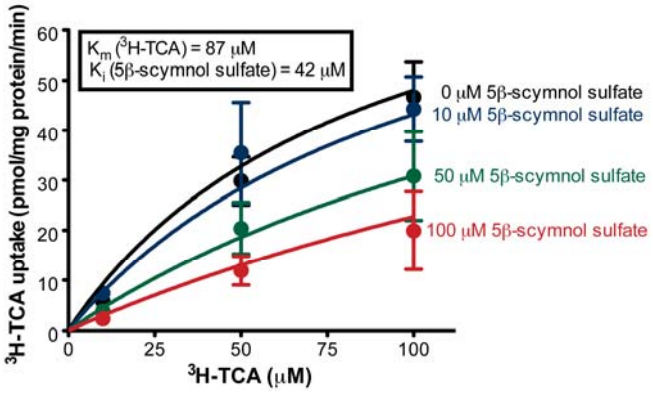


**Figure S5.** Competition of skAsbt transport function by 5β-scymnol sulfate. COS-7 cells were transfected with skAsbt and incubated for 10 minutes in uptake buffer supplemented with varying concentrations $^3$H-TCA and 5β-scymnol sulfate. Uptake of cells transfected with vector without insert was considered background and subtracted from measurements. Kinetic constants were calculated by curve fitting of the Michalis-Menten equation to the entire data set in Graphpad Prism 5 (Graphpad software). Data was normalized to total cell protein. Uptake of $^3$H-TCA had a $V_{max}$ of 90 ± 21 pmol/mg protein/min and a $K_m$ of 87 ± 38 μM. Competitive inhibition by 5β-scymnol sulfate had a $K_i$ of 42 ± 12 μM. These results indicate skAsbt has higher affinity for its endogenous bile salt 5β-scymnol sulfate than for the modern bile salt TCA. Values represent 3 independent experiments and are expressed as means ± SD.

**Supplemental References**

1. Wong MH, Oelkers P, Dawson PA. Identification of a mutation in the ileal sodium-dependent bile acid transporter gene that abolishes transport activity. J Biol Chem. 270(45):27228-34 (1995).

2. Oelkers P, Kirby LC, Heubi JE, Dawson PA. Primary bile acid malabsorption caused by mutations in the ileal sodium-dependent bile acid transporter gene (SLC10A2). J Clin Invest. 99(8):1880-7 (1997).

3. Zhang EY, Phelps MA, Banerjee A, Khantwal CM, Chang C, Helsper F, Swaan PW. Topology scanning and putative three-dimensional structure of the extracellular binding domains of the apical sodium-dependent bile acid transporter (SLC10A2). Biochemistry. 43(36):11380-92 (2004).

4. Banerjee A, Ray A, Chang C, Swaan PW. Site-directed mutagenesis and use of bile acid-MTS conjugates to probe the role of cysteines in the human apical sodium-dependent bile acid transporter (SLC10A2). Biochemistry. 44(24):8908-17 (2005).

5. Hussainzada, N., Banerjee, A. & Swaan, P. W. Transmembrane domain VII of the human apical sodium-dependent bile acid transporter ASBT (SLC10A2) lines the substrate translocation pathway. Mol. Pharmacol. 70, 1565–1574 (2006).

6. Hussainzada N, Khandewal A, Swaan PW. Conformational flexibility of helix VI is essential for substrate permeation of the human apical sodium-dependent bile acid transporter. Mol Pharmacol. 73(2):305-13 (2008).

7. Banerjee A, Hussainzada N, Khandelwal A, Swaan PW. Electrostatic and potential cation-pi forces may guide the interaction of extracellular loop III with Na+ and bile acids for human apical Na+-dependent bile acid transporter. Biochem J. 410(2):391-400 (2008).

8. Khantwal CM, Swaan PW. Cytosolic half of transmembrane domain IV of the human bile acid transporter hASBT (SLC10A2) forms part of the substrate translocation pathway. Biochemistry. 47(12):3606-14 (2008).

9. Hussainzada N, Da Silva TC, Zhang EY, Swaan PW. Conserved aspartic acid residues lining the extracellular loop 1 of sodium-coupled bile acid transporter ASBT Interact with Na+ and

7alpha-OH moieties on the ligand cholestane skeleton. J Biol Chem. 283(30):20653-63 (2008).

10. Hussainzada N, Claro Da Silva T, Swaan PW. The cytosolic half of helix III forms the substrate exit route during permeation events of the sodium/bile acid cotransporter ASBT. Biochemistry. 48(36):8528-39 (2009).

11. da Silva TC, Hussainzada N, Khantwal CM, Polli JE, Swaan PW. Transmembrane helix 1 contributes to substrate translocation and protein stability of bile acid transporter SLC10A2. J Biol Chem. 286(31):27322-32 (2011).

12. von Heijne, G. (1992) Membrane Protein Structure Prediction: Hydrophobicity Analysis and the 'Positive Inside' Rule. J.Mol.Biol. 225, 487-494.

13. G.E Tusnády and I. Simon (2001) The HMMTOP transmembrane topology prediction server" *Bioinformatics* **17**, 849-850