

In Silico Analysis of the Fucosylation-Associated Genome of the Human Blood Fluke *Schistosoma mansoni*: Cloning and Characterization of the Fucosyltransferase Multigene Family

Nathan A. Peterson², Tavis K. Anderson³, Timothy P. Yoshino^{1*}

1 Department of Pathobiological Sciences, University of Wisconsin, Madison, Wisconsin, United States of America, **2** Department of Entomology, University of Wisconsin, Madison, Wisconsin, United States of America, **3** Virus and Prion Research Unit, National Animal Disease Center, United States Department of Agriculture, Agricultural Research Service, Ames, Iowa, United States of America

Abstract

Fucosylated glycans of the parasitic flatworm *Schistosoma mansoni* play key roles in its development and immunobiology. In the present study we used a genome-wide homology-based bioinformatics approach to search for genes that contribute to fucosylated glycan expression in *S. mansoni*, specifically the α 2-, α 3-, and α 6-fucosyltransferases (FucTs), which transfer L-fucose from a GDP-L-fucose donor to an oligosaccharide acceptor. We identified and *in silico* characterized several novel schistosome FucT homologs, including six α 3-FucTs and six α 6-FucTs, as well as two protein O-FucTs that catalyze the unrelated transfer of L-fucose to serine and threonine residues of epidermal growth factor- and thrombospondin-type repeats. No α 2-FucTs were observed. Primary sequence analyses identified key conserved FucT motifs as well as characteristic transmembrane domains, consistent with their putative roles as fucosyltransferases. Most genes exhibit alternative splicing, with multiple transcript variants generated. A phylogenetic analysis demonstrated that schistosome α 3- and α 6-FucTs form monophyletic clades within their respective gene families, suggesting multiple gene duplications following the separation of the schistosome lineage from the main evolutionary tree. Quantitative decreases in steady-state transcript levels of some FucTs during early larval development suggest a possible mechanism for differential expression of fucosylated glycans in schistosomes. This study systematically identifies the complete repertoire of FucT homologs in *S. mansoni* and provides fundamental information regarding their genomic organization, genetic variation, developmental expression, and evolutionary history.

Citation: Peterson NA, Anderson TK, Yoshino TP (2013) *In Silico* Analysis of the Fucosylation-Associated Genome of the Human Blood Fluke *Schistosoma mansoni*: Cloning and Characterization of the Fucosyltransferase Multigene Family. PLoS ONE 8(5): e63299. doi:10.1371/journal.pone.0063299

Editor: Geoffrey N. Gobert, Queensland Institute of Medical Research, Australia

Received: December 9, 2012; **Accepted:** March 30, 2013; **Published:** May 16, 2013

Copyright: © 2013 Peterson 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.

Funding: This work was supported by the National Institutes of Health grants AI015503 and AI061436 to TPY and NIH-NIAID schistosome supply contract no. HHSN272201000005I. NAP was supported by a NRSA predoctoral fellowship through the Cellular and Molecular Parasitology Training Grant NIH T32 AI007414. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: yoshinot@vetmed.wisc.edu

Introduction

Chronic schistosomiasis in mammalian hosts, including humans, results from granulomatous inflammation in response to parasite eggs that accumulate in host tissues [1]. Previous studies in *Schistosoma mansoni* suggest that surface-expressed and secreted/excreted carbohydrates are key elements that drive this pathogenesis, with oligosaccharides playing roles in egg sequestration, Th2 immune biasing, granuloma formation and modulation, and strong antibody responses in human hosts [2,3]. Likewise, in snail intermediate hosts, carbohydrates serve as ligands to plasma-associated lectins and mediate hemocyte encapsulation and cytotoxic reactive oxygen species responses [4,5]. A common determinant in many of these host-parasite interactions is the deoxyhexose sugar L-fucose, which comprises as much as 40% of the total structural carbohydrates in larvae of *S. mansoni* [6].

Alpha2- and α 3-linked fucoses are major constituents of a diverse group of immunologically important LacdiNAc (LDN; GalNAc β 1-4GlcNAc)-derived glycotopes, including F-LDN,

LDN-F, F-LDN-F, LDN-DF and DF-LDN-DF, as well as the Lewis X glycotope, which constitute the non-reducing ends of protein- and lipid-conjugated oligosaccharides in both mammalian and snail host-associated developmental stages [7–10]. Many aspects of schistosome pathogenesis mentioned above are attributable to these glycotopes. Additionally, α 3- and α 6-linked fucoses occur alone or in combination as substituents of the chitobiose core (-GlcNAc β 1-4GlcNAc β 1-) in paucimannosidic and complex type N-glycans [11]. Such core modifications, especially α 3-fucosylation, induce strong allergic responses in mammalian hosts and account for the interspecific immunological cross-reactivity observed among plant, insect, and helminth glycoproteins [12,13]. Importantly, studies indicate that the expression of fucosylated glycotopes in *S. mansoni* is stage- and gender-specifically regulated [7,9,10,14], yet the mechanisms of this regulation, including the underlying enzymatic machinery, remains largely unknown. To better understand the developmental expression of immunologically important fucosylated glycans in *S. mansoni*, basic information is needed regarding the repertoire of Golgi-localized fucosyltrans-

ferases (FucTs), specifically the α 2-, α 3- and α 6-FucTs, which transfer L-fucose from a guanosine diphosphate (GDP)-L-fucose donor to an oligosaccharide acceptor, creating α 2, α 3, or α 6 linkages.

While α 2-, α 3-, and α 6-linked fucoses are prevalent in schistosome fucoconjugates and all three fucosylation activities have been observed in extracts of various developmental stages [15–18], only α 3-FucTs have been described in *S. mansoni*. These include homologs “*SmFucT*” [19] and “*SmFucTA*” [20], which are herein referred to as *FucT-VII* and *FucTA*, respectively. More recently, Fitzpatrick et al. [21] identified eight distinct Pfam-annotated putative α 3-FucT genes, including one corresponding to *FucTA*, in the *S. mansoni* GeneDB database. However, to date, the full extent of the α 3-FucT multigene family in *S. mansoni* is still unknown, and most of these predicted genes have yet to be substantively characterized. Furthermore, nothing is known about the α 2- and α 6-FucT genes in *S. mansoni* despite the wide distribution and abundance of glycans displaying α 2- and α 6-linked fucose. In the present study, we used a homology-based genome-wide bioinformatics approach to identify and *in silico* characterize the complete repertoire of schistosome FucT homologs. In addition to the α 2-, α 3- and α 6- FucTs, our investigation included the protein O-FucTs, which are not associated with glycoptope expression but instead transfer L-fucose to serine and threonine residues of epidermal growth factor- and thrombospondin-type repeats. To our knowledge, this is the most comprehensive study to date regarding the genomic organization, alternative splicing, and molecular phylogenetics of the FucTs in *S. mansoni*. Additionally, given the prominence of fucosylated glycans expressed at the host-parasite interface and our interest in their presumed role in innate immune responses in the snail hosts, we also performed an analysis of α 3-FucT gene expression during *in vitro* miracidium-to-primary sporocyst development.

Materials and Methods

Isolation and Cultivation of *S. mansoni* Larva

Ethics statement. Research protocols involving mice, including routine maintenance and care, were reviewed and approved by the Institutional Animal Care and Use Committee (IACUC) at the University of Wisconsin-Madison under assurance no. A3368-01.

Adult and larval *S. mansoni* (NMRI strain) were obtained and axenically cultivated as previously described by Yoshino and Laursen [22]. Briefly, adults and eggs were harvested from the hepatic portal veins and livers, respectively, of infected mice 7–8 weeks post-exposure to infective cercariae. Livers were homogenized to release the trapped eggs, and miracidial hatching was triggered by incubation in artificial pond water [23]. Miracidia were either used immediately or induced to transform by cultivation at 26°C in Chermis’s Balanced Salt Solution (CBSS; 47.9 mM NaCl/2.0 mM KCl/0.5 mM Na₂HPO₄/1.8 mM MgSO₄·7 H₂O/3.6 mM CaCl₂·2 H₂O/0.6 mM NaHCO₃; [24]) containing glucose and trehalose (1 g/L each) as well as penicillin and streptomycin (CBSS⁺). Within 24 h of cultivation, most miracidia had fully transformed to primary sporocysts. In this study, sporocyst cultures were maintained for 2 and 10 days before material extraction. For 10-day cultivations, the CBSS⁺ culture medium was changed on days 2 and 7.

FucT Gene Identification

The amino acid sequences of previously characterized α 2-, α 3-, α 6-, and protein O-FucTs of *S. mansoni*, *Homo sapiens*, *Drosophila melanogaster* and *Caenorhabditis elegans*, as well as the unique dual-

function β 3-galactosyltransferase/ α 2-FucT PgtA (also called FucB) of *Dictyostelium discoideum*, were downloaded from Reference Sequence (RefSeq) and GenBank online databases at the National Center for Biotechnology Information (NCBI) (accession numbers in Table S5). These sequences were used as queries in a genome-wide basic local alignment search tool [25] screen of genomic scaffolds and predicted genes in the *Schistosoma mansoni* database (SchistoDB).

Primer Design

FucT oligonucleotide primers used in reverse transcriptase (RT)-PCR, rapid amplification of cDNA ends (RACE), and real-time quantitative (q)PCR reactions were designed using Vector NTI Advance 11.0 software (Invitrogen, Eugene, OR, USA) and the Integrated DNA Technologies (IDT) SciTools suite (www.idtdna.com/scitools/scitools.aspx) based on available SchistoDB-derived genomic sequence information and original data obtained by this study. Custom DNA oligonucleotides were purchased from IDT (Coralville, IA, USA). Primer sequences used in this study are provided in Tables S1, S2, S3, S4.

Reverse Transcriptase-PCR and Rapid Amplification of cDNA Ends for FucT Transcript Sequencing

Unless otherwise stated, all kits and reagents were used according to the manufacturers’ recommendations. All protocols involving thermal cycling were executed on a Mastercycler[®] ep Thermal Cycler (Eppendorf North America, Hauppauge, NY, USA). Miracidia, 2-day *in vitro*-cultivated primary sporocysts, and mixed-sex adult worms were washed five times with artificial pond water, CBSS, and mammalian phosphate-buffered saline (pH 7.4), respectively, and total RNA was extracted using TRIzol[®] Reagent (Invitrogen). Genomic DNA contamination was removed from raw RNA extracts by TURBO[™] DNase treatment (Applied Biosystems, Foster City, CA, USA), and the DNA-free RNA was converted to RT-PCR-ready cDNA using the SuperScript[®] III First-Strand Synthesis System (Invitrogen). Reverse transcriptase-PCR reactions (25 μ L/rxn) contained GoTaq[®] amplification reagents (Promega, Madison, WI, USA), with reaction mixtures generally comprising 2.5 U GoTaq[®] Flexi DNA Polymerase, 1X Green GoTaq[®] Flexi Reaction Buffer, 400 nM each forward and reverse gene-specific primers (Table S1), 1.6 mM dNTP mix (400 μ M each), 1.5 mM MgCl₂, and 75–350 ng RNA input-equivalents RT-PCR-ready cDNA. The thermal profile was as follows: initial denaturation at 94°C/3 min; 40 cycles of 94°C/30 sec, 56–60°C/30 sec and 72°C/3 min; and final extension at 72°C/10 min. Some reactions required further optimization of cDNA input, annealing temperature, and cycle duration. Amplification products were fractionated by 1% agarose gel electrophoresis, and ethidium bromide-stained bands were excised and purified using a QIAquick Gel Extraction Kit (Qiagen, Germantown, MD, USA). Amplicons were inserted into pCR[®]4-TOPO[®] vector (TOPO[®] TA Cloning[®] Kit for Sequencing, Invitrogen) and transformed into One Shot[®] TOP10 Chemically Competent *E. coli*, which were then incubated overnight at 37°C on LB (1.0% tryptone/0.5% yeast extract/1.0% NaCl) agar (1.5%) containing 50 μ g/mL kanamycin. Positive transformants were picked and grown overnight at 37°C in LB broth containing 100 μ g/mL ampicillin, and plasmids were isolated using a QIAprep Spin Miniprep Kit (Qiagen). To verify the presence of an appropriate insert, plasmids were restriction-digested with EcoRI endonuclease (Promega, Madison, WI), and restriction fragments were analyzed by electrophoretic fractionation and ultraviolet transillumination. Insert-bearing plasmids were used as templates for dideoxy sequencing (BigDye Terminator v3.1; Applied Biosystems), and

reaction products were purified using Agencourt® CleanSEQ® magnetic beads (Beckman Coulter, Brea, CA, USA). Following cleanup, insert sequences were read by the DNA Sequence Laboratory at the University of Wisconsin Biotechnology Center (Madison, WI, USA) using a 3730×1 Automated DNA Sequencer (Applied Biosystems).

Following RT-PCR confirmation of gene transcription, TRIzol®-derived DNA-free total parasite RNA was converted to 5' and 3' RACE-ready cDNAs using a SMART™/SMARTer™ RACE cDNA Amplification Kit (Clontech, Mountain View, CA, USA), and cDNA ends were PCR-amplified using the Advantage® 2 PCR Kit (Clontech) with 200 nM gene-specific primers (Table S2), 240 nM universal primer mix (RACE kit component), and 20–100 ng RNA input-equivalents of RACE-ready cDNA (50 µL/rxn total volume). The thermal profile for RACE PCR reactions included initial denaturation at 94°C/3 min, 25–30 cycles of 94°C/30 sec, 58–62°C/30 sec and 72°C/3 min, and final extension at 72°C/10 min. Further optimization was required in some cases. Often, nested PCR was performed using the above thermal profile and recipe but with 200 nM nested gene-specific and universal primers and 2.5 µL diluted “outer” PCR reaction (1/50 dilution with Tricine-EDTA Buffer; Clontech). Amplification products were isolated, cloned, and sequenced as described above.

Reverse transcriptase-PCR and RACE sequence data were assembled and edited using Vector NTI Advance 11.0 software. Complete coding sequences (CDSs) were verified by RT-PCR amplification and subsequent sequencing (as above) using primers designed to encompass the full open reading frames (ORFs) (Table S3).

Phylogenetic Analysis of *S. mansoni* FucT Genes

Amino acid sequences representing the known diversity of $\alpha 2$ -, $\alpha 3$ -, $\alpha 6$ -, and protein O-FucTs from *Dictyostelium*, *Caenorhabditis*, *Drosophila*, *Danio*, *Mus*, and humans were compiled with our data from *S. mansoni* (Table S5). Alignments were generated using default settings in MUSCLE v3.6 [26], with manual correction in Mesquite [27]. An initial neighbor-joining tree was constructed using FastTree v2.0.1 [28] with a Jukes-Cantor+CAT model to serve as a guide tree for Bayesian phylogenetic inference. Analyses were then performed using mixed amino-acid models within MrBayes v3.1.2 [29] with two parallel runs of four Markov Chain Monte Carlo (MCMC) chains, each for five million generations, with subsampling every 100th generation. Two independent replicates were conducted to determine that analyses were not trapped at local optima. Convergence of the MCMC chains was explored graphically using the online program AWTY [30], in addition to assessing stationarity of molecular evolutionary parameters by effective sample sizes >400 in Tracer v1.5 [31]. Trees prior to stationarity were burned-in, and the remaining were used to assess posterior probabilities for nodal support. The bifunctional $\beta 3$ -galactosyltransferase/ $\alpha 2$ -FucT PgtA of *Dictyostelium discoideum* was used as an outgroup to facilitate inferences regarding FucT evolution.

Three predicted FucTs from *Schistosoma japonicum* (GenBank accession/SchistoDB annotation numbers CAX72936.1, CAX73054.1, and Sjp_0036210), in addition to the above data set, were used to construct a second phylogeny. FucT amino acid sequences were aligned, and a maximum-likelihood (ML) tree was inferred using the RAxML v7.3.4 [32] program, employing a general time-reversible (GTR) model of nucleotide substitution with Γ -distributed rate variation among sites. Statistical support for individual nodes within the best-scoring tree was estimated

using the rapid bootstrap algorithm (1,000 replications) in RAxML.

Real-time Quantitative PCR Analysis of $\alpha 3$ -FucT mRNA Expression in Miracidia and Primary Sporocysts

Real-time qPCR was performed according to recommendations by Applied Biosystems (<http://www.appliedbiosystems.com/absite/us/en/home/applications-technologies/real-time-pcr/>), including strict criteria for qPCR primer design, validation, and optimization. Relative transcript abundance in miracidia and primary sporocysts was assessed using the comparative C_T ($\Delta\Delta C_T$) method, in which an endogenous calibrator gene is used to normalize the C_T values for a gene of interest. To identify appropriate calibrators for this study, the *Schistosoma mansoni* Serial Analysis of Gene Expression (SAGE) Database [33] was screened for genes whose transcript abundances are stable between miracidia and primary sporocysts (R-value <4; [34]). Based on the available SAGE data, ATP synthase f chain (herein termed “*ATP_{sf}*”; SAGE tag 195 corresponding to Smp_140480 at SchistoDB) and the *GroES* chaperonin (SAGE tag 132 corresponding to Smp_097380) were selected. The compatibility of calibrator and $\alpha 3$ -FucT primers under normal reaction conditions was assessed by plotting ΔC_T at various dilutions of cDNA input and determining the slope of the resultant line; primer efficiencies were deemed compatible if the absolute value of the slope was less than 0.1. Primers used in this study for qPCR are listed in Table S4.

Miracidia and *in vitro*-cultivated primary sporocysts were washed five times with artificial pond water and CBSS, respectively, and total RNA was extracted using TRIzol® Reagent. The raw RNA was decontaminated by TURBO™ DNase treatment and quantitatively converted to first-strand cDNA using the Superscript™III-First-Strand Synthesis System. Real-time qPCR reactions (50 µL/rxn) were prepared in triplicate, comprising 1X SYBR Green PCR Master Mix (Applied Biosystems), 20 ng RNA input-equivalents parasite cDNA, and 200 nM each forward and reverse gene-specific primers. Reactions were run on an ABI 7300 Real-Time PCR System (Applied Biosystems) with the following cycle profile: initial denaturation at 95°C/10 min followed by 40 cycles of 95°C/15 sec and 60°C/1 min. PCR product accumulation was monitored in real time, and amplification fidelity was assessed by post-cycling thermal dissociation and electrophoretic fractionation.

To best assess $\alpha 3$ -FucT expression by the $\Delta\Delta C_T$ method, the geometric mean of *ATP_{sf}* and *GroES* C_T values was used to normalize $\alpha 3$ -FucT C_T such that $\Delta C_T = C_{T-FucT} - C_{T-GeoMean(ATP_{sf}, GroES)}$. Heteroscedastic Student's T-tests were employed to compare $\alpha 3$ -FucT expression in miracidia versus primary sporocysts across three independent biological replicates, with significance set at $p \leq 0.05$.

Results and Discussion

Composition, Genomic Organization, and Splicing of Schistosome FucT Genes

An exhaustive homology search of the *Schistosoma mansoni* database (SchistoDB; [35]) in conjunction with comprehensive sequence analyses generated a non-redundant list of 15 genes with predicted roles in $\alpha 3$ -, $\alpha 6$ - and protein O-fucosylation (see Table 1 for genes and corresponding SchistoDB annotations). Seven genes were classified as putatively involved in $\alpha 3$ -fucosylation (herein termed *FucTs A,B,C,D,E,F,G*), six in $\alpha 6$ -fucosylation (*FucTs H,I,J,K,L,M*), and two in protein O-fucosylation (*POFucTs A,B*). No genes with a predicted role in $\alpha 2$ -fucosylation were identified.

Of the 15 FucT homologs described here, only *FucTA* had been previously cloned and characterized [20]. Homology-based searches failed to detect any sequences in the SchistoDB corresponding to *FucT-VII*, the only other FucT homolog reported from *S. mansoni* [19]. Notably, subsequent attempts to clone the *FucT-VII* CDS from miracidia, primary sporocysts, and adult worms were unsuccessful despite the use of numerous primer sets and various amplification parameters (see discussion below).

Previously, Fitzpatrick et al. [21] used a similar bioinformatics approach (Pfam searches) to identify predicted α 3-FucTs for inclusion in a microarray analysis of gene expression in *S. mansoni*. Their approach generated eight unique contigs/sequences (seven corresponding to present *FucTs A-G*, plus one more), but no further sequence analyses were conducted to validate them as complete α 3-FucT-coding genes. The eighth putative FucT (SchistoDB annotation Smp_194990) was not incorporated in their microarray analysis and its transcription was not confirmed. In the present study, Smp_194990 was excluded from downstream analyses because it is ostensibly incomplete, comprising four exons that constitute just 804 nt of a potential FucT CDS. While it is possible the remaining coding segments reside within a genomic sequencing gap, the nearest gap is \sim 32 kb downstream and represents a distance much longer than the introns of other schistosome FucTs. Altogether, 13 such “gene fragments” (all α 3-FucT-like) were discarded as probable pseudogenes and not analyzed further.

To confirm FucT gene transcription in *S. mansoni* and obtain full-length CDSs, transcript sequences were RT-PCR- and RACE-amplified using cDNA derived from miracidia, primary sporocysts, and adults. Complete sequences were submitted to the NCBI GenBank (see Table S5 for accession numbers). In most cases, *in silico* translation of gene transcripts yielded a single prevailing ORF. However, translation of *FucTG* revealed two tandemly situated ORFs corresponding to different segments of a single α 3-FucT protein, indicating premature termination of translation. Indeed, sequence analyses determined that exon 8 of *FucTG* encodes a premature termination codon (PTC) while exon

9 forces a downstream frameshift (see Figure S2), both resulting in the omission of a major segment of the FucT catalytic domain. Moreover, *in silico* analyses of FucTG membrane topology suggested the absence of an N-terminal transmembrane domain (TMD), a structural element observed in every known eukaryotic α 3-FucT [36]. Altogether, our data suggest that *FucTG* is a pseudogene.

Nucleotide sequence data were mapped onto the corresponding SchistoDB-derived genomic scaffolds to examine FucT gene organization (Table 1; Figure S1). With the exceptions of *FucTE* and *FucTF*, which are tandemly situated on scaffold Smp_scaff000060, α 3-FucT genes occur on separate genomic scaffolds. In contrast, schistosome α 6-FucTs are distributed between just two scaffolds, with *FucTs I-M* arrayed on Smp_scaff000066 and *FucTH* on Smp_scaff000594. Scaffold Smp_scaff000066 features SchistoDB-annotated gene predictions corresponding to α 6-FucTs *J-L* (Table 1), but no obvious matches for *FucTI* and *FucTM*. The scaffold does include a fourth SchistoDB annotation for a putative FucT gene (Smp_138740), however all attempts by RT-PCR and RACE to confirm its expression failed. Notably, sequence comparisons show that *FucTI* and *FucTM* are almost identical to Smp_138740, differing in their CDS regions by 12 and 15 nt, respectively ($>$ 99% identity in both comparisons). Moreover, the CDSs of *FucTs I, J, and M* are very similar, differing from one another by just 11 (*I* vs. *J*), 23 (*I* vs. *M*), and 20 (*J* vs. *M*) nt. Given the high similarity among these sequences (both exonic and intronic) and the close proximity of Smp_138740 to *FucTJ* within the genome, it is conceivable that *FucTs I, J, and M* were incorrectly consolidated during contig assembly into two tandem genes. Indeed, the 5' half of the *FucTI* mRNA transcript (including the untranslated region) is identical to that of *FucTJ* while its 3' end is almost identical to the 3'-coding segments of upstream Smp_138740.

Also apparent by genomic overlay, schistosome FucTs are invariably multiexonic, with CDSs spanning 5–8 exons among the α 3-FucTs, 9 or 11 exons in the α 6-FucTs, and 9–10 exons in the protein O-FucTs (Figure S1). The observation that *FucTA* is

Table 1. Summary of FucT genomic organization in *Schistosoma mansoni*.

Gene	Gene ID ^a	Scaffold ID ^a	Approx. size (bp)	No. of exons	ORF length (nt) ^b	Prot. length (aa) ^b
<i>FucTA</i>	Smp_148850	Smp_scaff000127	$>$ 16,057	8	1,278	426
<i>FucTB</i>	Smp_099090	Smp_scaff000622	\geq 7,903	6	1,248	416
<i>FucTC</i>	Smp_154410	Smp_scaff000171	\geq 24,925	10	1,389	463
<i>FucTD</i>	Smp_054300	Smp_scaff000144	\geq 9,546	7	1,194	398
<i>FucTE</i>	Smp_137740	Smp_scaff000060	$>$ 21,917	8	1,284	428
<i>FucTF</i>	Smp_137730	Smp_scaff000060	$>$ 10,507	7	1,302	434
<i>FucTG</i>	Smp_129750	Smp_scaff000024	\geq 12,334	10	–	–
<i>FucTH</i>	Smp_175120	Smp_scaff000594	$>$ 38,746	11	1,797	599
<i>FucTI</i>	–	Smp_scaff000066	–	10	1,776	592
<i>FucTJ</i>	Smp_138730	Smp_scaff000066	$>$ 15,639	10	1,776	592
<i>FucTK</i>	Smp_138750	Smp_scaff000066	$>$ 37,696	10	1,737	579
<i>FucTL</i>	Smp_030650	Smp_scaff000066	$>$ 24,336	10	1,764	588
<i>FucTM</i>	–	Smp_scaff000066	–	10	1,776	592
<i>POFucTA</i>	Smp_065240	Smp_scaff000199	\geq 9,121	9	1,251	417
<i>POFucTB</i>	Smp_131810	Smp_scaff000033	$>$ 27,578	12	1,650	550

^aSmp gene and scaffold IDs refer to nomenclature in the SchistoDB [35].

^bORF and protein sizes are provided for the main/major transcripts. Alternative splicing may alter ORF length and protein coding.

doi:10.1371/journal.pone.0063299.t001

multiexonic contradicts a previous assertion by Trottein et al. [20] that its CDS is fully encoded by a single exon. Similar multiexonic organizations have been observed for FucTs of other invertebrates and plants [36–38], as well as a subset of vertebrate genes, including members of the *FUT10/11* superfamily, *FUTs* 7–8, and protein O-FucTs *POFUT1* and *POFUT2* [37,39–41]. In contrast, vertebrate *FUTs* 1–6 and *FUT9* are all monoexonic [36,42]. Notably, genomic overlay further revealed that the proportioning of schistosome α 3- and α 6-FucT CDSs among their multiple exons is roughly maintained (i.e., exon-exon junctions are well conserved within gene families). Similar conservation of ORF-exon architecture has been observed among vertebrate α 3-FucTs (e.g., human *FUTs* 3–6/9; [36]) and for the protein O-FucTs across a broad diversity of invertebrate and vertebrate taxa [37].

The significance of a multiexonic gene organization is perhaps most apparent in the context of alternative splicing. Variations in mRNA splicing were observed for all genes except α 3-FucTD and α 6-FucTM (Figure S2). It should be noted that due to the extreme similarity between the 5' regions of *FucTI* and *FucTJ*, it is unclear whether the observed splicing occurs in one or both genes. Also, many of these observations were derived from RT-PCR and RACE experiments that targeted specific sections of each transcript and not the entire ORF. For this reason, the relationships among alternative splice events (i.e., whether events are co-dependent in the formation of particular isoforms) are largely unknown. All modes of alternative splicing were observed: exon skipping (e.g., *FucTF*), intron retention (e.g., *FucTA*), mutual exclusion (e.g., exons 1 and 2 of *FucTH*), and use of alternate splice donor and acceptor sites (e.g., *FucTH*). An *in silico* analysis to define the consequences of alternative splicing determined that most variant splice events would alter protein coding by introducing a PTC, forcing a downstream frameshift, effecting an in-frame deletion or addition, or omitting the prototypical start or stop codon. Few alternative events are predicted to leave the prototypical ORF unchanged. Additional studies are necessary to determine the true biochemical effects of such variations.

In general, alternative splicing is the primary mechanism by which organisms generate greater mRNA structural complexity, thus expanding proteome diversity, facilitating post-transcriptional gene regulation (e.g., introduction of a PTC that results in nonsense-mediated decay), and enhancing untranslated region variability (affecting cis-regulatory elements that control translation efficiency, stability, and localization) (reviewed by [43]). In terms of fucosylation in *Schistosoma*, alternative splicing among the FucT genes might generate additional FucT diversity (possibly modifying acceptor specificity, localization, or catalytic efficiency) or cause transcripts to be targeted for nonsense-mediated decay, perhaps effecting a reduction in FucT protein expression. Alternative splicing also has complex roles at the cellular and organismal levels in the modulation of physiological activities during development and differentiation and in response to environmental stresses [43]. Indeed, the developmental regulation of alternative splicing in *S. mansoni* has been well documented. For example, Ram et al. [44] observed cercariae-specific intron retention in heat-shock transcription factor (*HSF*) transcripts that introduces a PTC and prohibits HSF translation, thus inhibiting downstream expression of the molecular chaperone heat-shock protein 70 (*HSP70*). In adults, the impeding intronic sequences are removed, allowing functional HSF protein to be synthesized and ultimately promoting *HSP70* expression. In another study, DeMarco et al. [45] demonstrated by semi-quantitative RT-PCR that several protein-coding variants of the schistosome transcriptional cofactor *CA150* are expressed in different ratios between male and female adults. In the present study, the results of

similar RT-PCR experiments suggest that at least a subset of the schistosome α 3-FucT genes may also be differentially spliced among miracidia, primary sporocysts, and adults (unpublished data). Ultimately, the regulated expression of particular isoforms may have a role in generating the observed stage- and gender-specific patterns of fucosylation in *S. mansoni*.

In silico Characterization of Schistosome FucT Proteins

To provide support for the putative functions of the 15 schistosome FucT homologs, their predicted amino acid sequences were compared within their respective gene families and against previously well characterized FucTs, and proteins were examined for the presence of key FucT-associated primary sequence elements. By definition, Golgi-resident α 2-, α 3-, and α 6-FucTs are type-II transmembrane proteins featuring a single TMD, which is flanked by a short cytoplasmic N-terminal tail and a luminal C-terminus that comprises a globular catalytic domain and a flexible hypervariable stem [46].

Alignment of the schistosome putative α 3-FucTs (including FucT-VII; primary sequences obtained by *in silico* translation) against homologs of *Caenorhabditis*, *Drosophila*, and humans revealed the presence of an N-terminal hypervariable region and a well-conserved C-terminal constant domain in each protein (Figure 1). Additionally, the alignment identified five α 3-FucT-specific motifs (I-V; [41]) that have roles in protein folding and catalytic activity (motif I; [47,48]), acceptor specificity (motif II; [49,50]), and GDP-L-fucose binding (motifs IV and V; [36,51]). Motifs III-V are well conserved for all sequences, but motifs I and II differ both between taxa and within the schistosome gene family. These variations likely reflect differences in acceptor utilization, evidenced by interspecific variations in their glycomes [52–54]. Analyses of transmembrane topology using TMHMM 2.0 [55] and the Tmpred server [56] suggest that all schistosome α 3-FucTs (excluding FucTG) are type-II transmembrane proteins featuring a single N-terminal TMD. In all cases, the TMD is offset from the C-terminal catalytic domain by a hypervariable stem region, which varies in length among family members (e.g., 13 aa in FucT-VII versus 118 in FucTC). Stem length in any Golgi-resident glycosyltransferase is thought to contribute to acceptor specificity by positioning the catalytic domain at a particular distance from the Golgi membrane and providing constraints on the sterics of the glycosylation reaction [57]. Overall, the schistosome α 3-FucT gene family shares ~7–10% identity with *Caenorhabditis*, *Drosophila* and human homologs, while alignment of the schistosome α 3-FucTs alone indicates ~10% identity (~17% if FucT-VII is omitted). Pairwise comparisons revealed 28.9–68.7% identity among schistosome genes.

Similar observations were made regarding the schistosome α 6-FucTs, which also comprise N-terminal variable and C-terminal constant regions (Figure 2). Alignment of these proteins against invertebrate and vertebrate homologs revealed the presence of three motifs that are well conserved across the α 2-, α 6- and protein O-FucT gene families [36,51,58]. Importantly, motif I residues Arg-420 and Arg-421, which are thought to be essential for GDP-L-fucose binding [59], are maintained in all schistosome α 6-FucTs. Analyses of transmembrane topology strongly indicate that schistosome α 6-FucTs also have a single N-terminal TMD with a type II orientation. Overall, schistosome α 6-FucTs are ~13–15% identical to *Caenorhabditis*, *Drosophila* and human homologs, while ~27% sequence identity exists within the schistosome gene family (73% if FucTH is excluded from the analysis). Intrafamilial pairwise alignments demonstrated as much as 99% identity (FucTI vs. FucTJ).

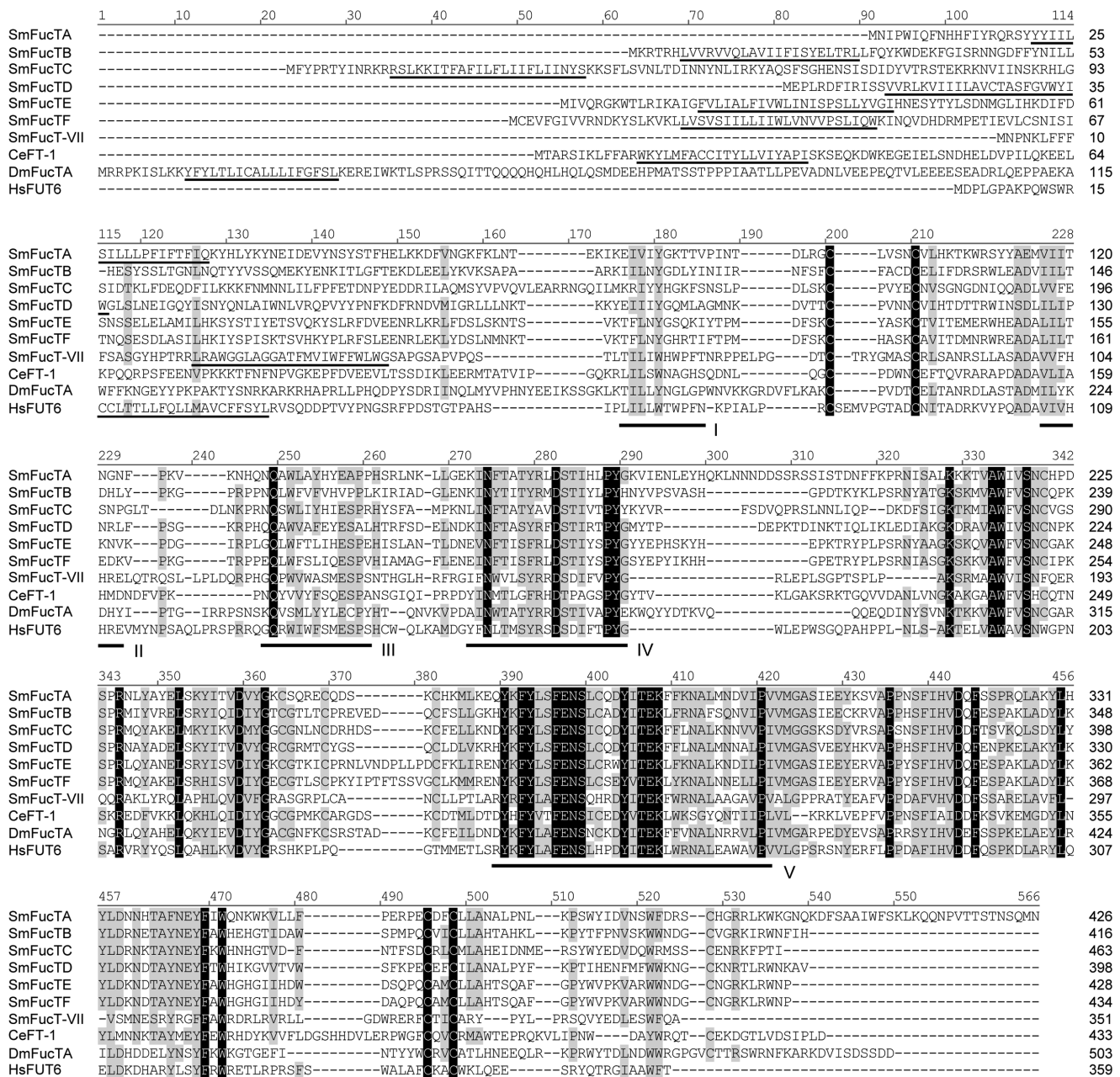


Figure 1. Amino acid alignment of α 3-fucosyltransferases (α 3-FucTs). Schistosome α 3-FucTs are compared to FT-1, FucTA, and FUT6 of *Caenorhabditis elegans*, *Drosophila melanogaster* and humans, respectively (NCBI accession numbers in Table S5). Alignment position is indicated above each block, and sequence length is reported to the right of each line. Positions exhibiting greater than 70% conservation are highlighted in gray, and identities are shown in black. An N-terminal TMD (underlined) was identified for each FucT using TMHMM 2.0 and/or the Tmpred online server (settings: min = 14/max = 23). The positions of five α 3-FucT-specific motifs (motifs I-V; [41]) are indicated. Motifs IV and V were previously described in human α 3-FucTs as motifs I and II [36,51], and present motifs I and II were formerly termed motif III and “acceptor-binding motif”, respectively [49,50]. Vector NTI Advance 11.0 software alignment settings: BLOSUM45 matrix with Cys and Trp weights adjusted to 99, gap opening penalty = 12, gap extension penalty = 0.1, gap separation penalty range = 0, no residue-specific or hydrophobic residue gaps; manual editing was necessary to fully align conserved motifs.
doi:10.1371/journal.pone.0063299.g001

Unlike α 2-, α 3- and α 6-FucTs, protein O-FucTs (comprising distinct O-FucT1 and O-FucT2 gene families) are predominantly ER-localized soluble proteins that transfer L-fucose to serine and threonine residues of epidermal growth factor- and thrombospondin-type repeats [60–62]. Amino-terminal signal peptides, which initially target proteins to the ER, have been described for both O-FucT1 and O-FucT2 proteins [37,61,62], and C-terminal

ER-retention/retrieval signals have been observed in most O-FucT1s studied to date [37,63,64]. Alignments of schistosome PO-FucTA and PO-FucTB against O-FucT1 and O-FucT2 orthologs, respectively, show that both proteins are well conserved in *S. mansoni* (~25–33% identity in pairwise alignments; Figures 3 and 4) and that both proteins contain three key motifs that are putatively involved in GDP-L-fucose binding [36,51,58]. A search

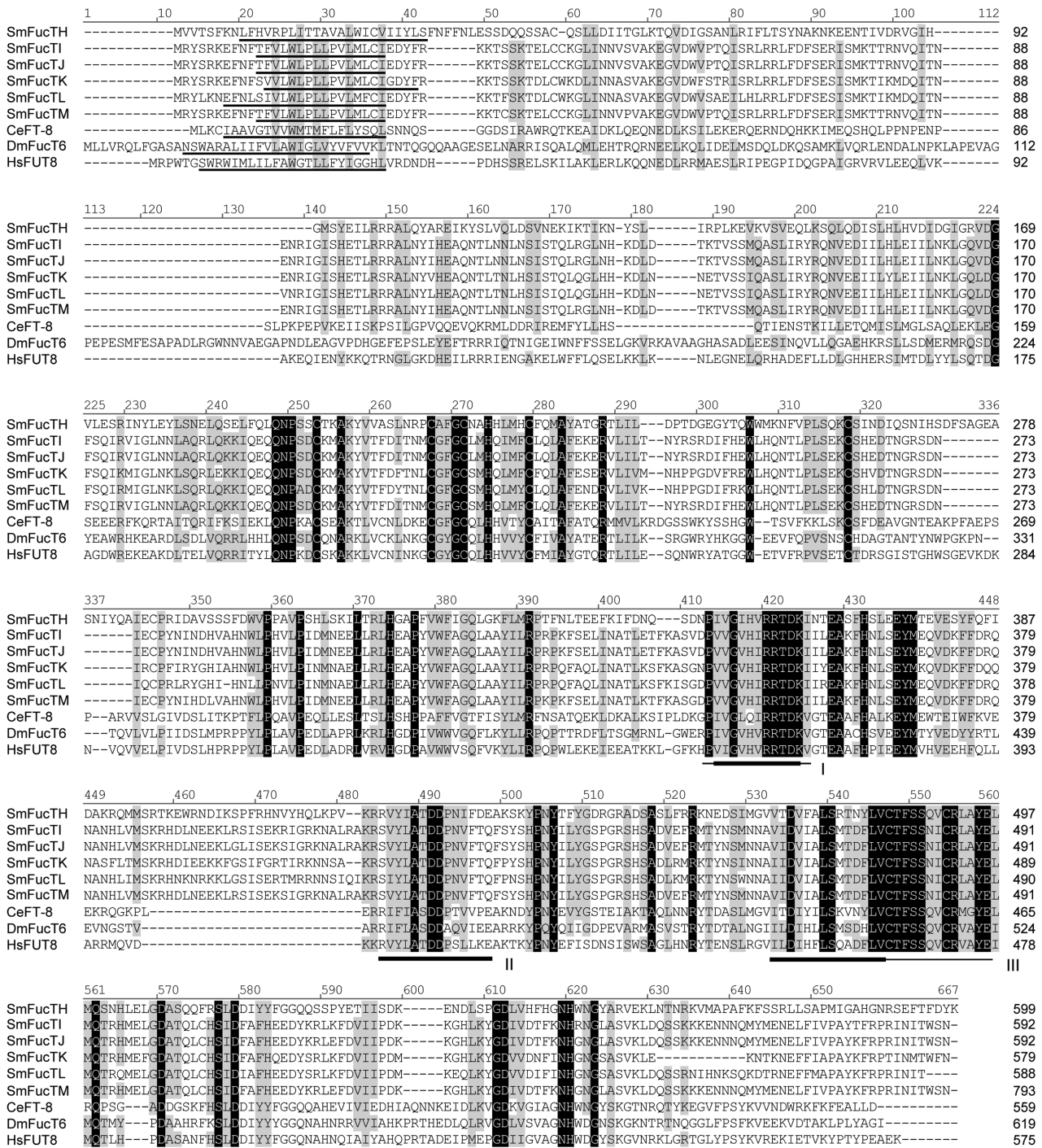


Figure 2. Amino acid alignment of $\alpha 6$ -fucosyltransferases ($\alpha 6$ -FucTs). Schistosome $\alpha 6$ -FucTs are compared to FT-8, FucT6, and FUT8 of *Caenorhabditis elegans*, *Drosophila melanogaster* and humans, respectively (NCBI accession numbers in Table S5). Alignment position is indicated above each block, and sequence length is reported at the end of each line. Positions of identity and positions exhibiting at least 70% conservation are highlighted in black and gray, respectively. An N-terminal TMD (underlined) was identified for each FucT using TMHMM 2.0 and/or the TMPred online server (settings: min = 14/max = 23). The positions of three hydrophobic cluster analysis-derived motifs (I-III), which are well conserved among $\alpha 2$ -, $\alpha 6$ - and protein O-FucTs [36,51,58], are indicated below the alignment blocks. The reported lengths of motifs I and III differ between previous studies, which is reflected in the thickness of the indicator line (thick line, [36]; thin line, [58]). Vector NTI Advance 11.0 software alignment settings: BLOSUM45 matrix with Cys and Trp weights adjusted to 99, gap opening penalty = 12, gap extension penalty = 0.1, gap separation penalty range = 0, no residue-specific or hydrophobic residue gaps.
doi:10.1371/journal.pone.0063299.g002

for N-terminal signal peptides was conducted using the Simple Modular Architecture Research Tool (SMART; [65]) and the

Phobius transmembrane topology and signal peptide prediction server [66], and signal peptides were successfully identified in all

sequences except schistosome POFucTA. Consistent with the expectation that protein O-FucTs are soluble, the Phobius output indicated the absence of a TMD in all cases. Protein sequences were also examined for a C-terminal ER-retention/retrieval signal such as the soluble protein motif KDEL or similar tetrapeptides, e.g., HEEL and RDEF of *Drosophila*, *Mus*, and human orthologs, or the membrane protein dibasic motifs KKxx and RRx [61,67,68]). Unlike most of its O-FucT1 orthologs, schistosome POFucTA lacks a KDEL-like tetrapeptide. Strikingly, the C-terminus of POFucTA terminates with a KK tandem repeat that is reminiscent of a membrane protein-associated dibasic ER sorting signal. However, lysine residues are inaptly spaced from the terminus and are therefore unlikely to participate in retrograde transport. It is unclear given the absence of both an N-terminal signal sequence and a retention/retrieval signal if or how POFucTA is initially targeted to and later retained in the ER, where it is predicted to function in protein O-fucosylation. In contrast to POFucTA, schistosome POFucTB features an N-terminal signal sequence and a classical C-terminal KDEL ER-retention/retrieval tetrapeptide. Importantly, this is the first report of an ER-retention/retrieval signal in a protein O-FucT2 to date.

Phylogenetic Analysis of Schistosome FucTs

A phylogenetic analysis was conducted to examine the relationship between the schistosome FucT homologs and 62 previously characterized $\alpha 2$ -, $\alpha 3$ -, $\alpha 6$ -, and protein O-FucTs of *Dictyostelium*, *Caenorhabditis*, *Drosophila*, *Danio*, *Mus* and humans (Figure 5; see Table S5 for NCBI accession numbers). The resultant tree, which is rooted on a bifunctional $\beta 3$ -galactosyltransferase/ $\alpha 2$ -FucT from *Dictyostelium* (PgtA; [69,70]), resolves the data into five major clades that correspond to the $\alpha 2$ -, $\alpha 3$ -, $\alpha 6$ -, O-FucT1 and O-FucT2 gene families, and it supports previous work demonstrating that gene duplication and functional divergence among FucTs is a relatively ancient event [58], with members of each gene family being represented across invertebrate and vertebrate taxa.

The formation of a superclade comprising the $\alpha 2$ -, $\alpha 6$ -, and protein O-FucT lineages is consistent with observations by Martínez-Duncker et al. [58] that these distinct functional groups constitute a single superfamily of FucTs. Indeed, three well-conserved motifs, which are thought to have a role in GDP-L-fucose binding, are shared across the $\alpha 2$ -, $\alpha 6$ - and protein O-FucT families [36,51,58]. Analogous motifs are also conserved among the $\alpha 3$ -FucTs (motifs IV and V; [36]), however the relationship of these motifs to motifs I-III of the $\alpha 2$ -/ $\alpha 6$ -/O-FucT superfamily is unclear. The topology of the $\alpha 2$ -/ $\alpha 6$ -/O-FucT superfamily clade suggests two main lineages, the $\alpha 2$ -/ $\alpha 6$ - and protein O-FucT lines, that were derived by duplication of an ancestral gene followed by functional divergence. For the O-FucT lineage this meant the loss of an N-terminal TMD and alteration of acceptor specificity (oligosaccharides to proteins) and subcellular localization (Golgi to ER). More recently, the ancestral $\alpha 2$ -/ $\alpha 6$ - and protein O-FucT genes underwent a second round of duplication and functional divergence to form the modern $\alpha 2$ -, $\alpha 6$ -, O-FucT1, and O-FucT2 gene families. The schistosome FucTs clustered predictably within the $\alpha 2$ -/ $\alpha 6$ -/O-FucT superfamily, with *Schistosoma* being represented in all but the $\alpha 2$ -FucT clade. POFucTA and POFucTB are clear orthologs of the O-FucT1 and O-FucT2 gene families, respectively, while the schistosome $\alpha 6$ -FucT homologs constitute their own monophyletic group within the $\alpha 6$ - lineage.

Consistent with observations by Roos et al. [71] and Mollicone et al. [41], the present phylogeny predicts two distinct lineages within the $\alpha 3$ -FucT clade, one comprising the FUT10/11 gene

superfamily and the other encompassing human FUTs 3–7/FUT9 and their orthologs, as well as distinct groups from *Schistosoma* and *Caenorhabditis*. With the exception of FucT-VII, schistosome $\alpha 3$ -FucTs comprise a monophyletic group. The close proximity of schistosome FucT-VII to murine Fut7 and its relative distance from the main schistosome lineage supports previous claims that *FucT-VII* is an artifact of *in vitro* contamination or a product of a recent horizontal gene transfer from mouse to schistosome [36]. Indeed, type-VII FucTs, including schistosome FucT-VII, are associated with $\alpha 3$ -fucosylation of sialylated Lewis-type oligosaccharide acceptors [19,39], which have never been observed in *S. mansoni*. Oriol et al. [36] favored the notion of lateral gene transfer because Marques et al. [19] reportedly detected *FucT-VII* mRNA expression in snail-derived larvae and hamster-reared adults. However, given failed attempts in the present study to PCR-amplify *FucT-VII* and locate the relevant gene sequence in the SchistoDB database, it is perhaps more likely that *FucT-VII* was derived in the previous work by *in vitro* contamination.

In contrast to the schistosome $\alpha 3$ -FucT gene family, *Drosophila* $\alpha 3$ -FucTs exhibit a polyphyletic distribution, with two of these genes appearing more closely related to the FUT10/11 superfamily and two others divided between the schistosome and *Caenorhabditis* lineages. Mollicone et al. [41] concluded that *Drosophila* FUT10/11-like $\alpha 3$ -FucTs share a common ancestor with the present FUT10/11 superfamily. Because schistosomes apparently lack a FUT10/11-like gene, the origin of the FUT10/11 superfamily is likely more recent than the separation of schistosomes from the insect and vertebrate lineages.

Interestingly, the layering of gene organization data onto the present tree is highly informative regarding $\alpha 3$ -FucT gene evolution. Invertebrate $\alpha 3$ -FucTs and FUT10/11 superfamily genes are all multiexonic, while vertebrate *FUTs* 3–7 and *FUT9* feature a much-simplified genomic organization (encoded by just one exon; two in *FUT7*). All of the bi- and monoexonic vertebrate genes form a monophyletic clade, suggesting that *FUTs* 3–7/9 were derived by successive duplication after the retrotransposition of a single ancestral gene. The introduction of a single intron into *FUT7* is perhaps a more recent event in the evolution of the vertebrate $\alpha 3$ -FucTs. Furthermore, because this simplified gene organization is conserved across vertebrate taxa (*Danio*, *Mus*, and humans) and not among invertebrates, it can be concluded that retrotransposition must have occurred after the separation of vertebrate and invertebrate lineages but early in vertebrate evolution. Notably, Marques et al. [19] observed that schistosome *FucT-VII* is monoexonic, which is consistent with its monophyletic relationship with *Mus* and human *FUT7* genes in the present phylogenetic tree and further supports hypotheses that *FucT-VII* is a product of *in vitro* contamination or horizontal gene transfer from mice.

Monophyly within the schistosome $\alpha 3$ - and $\alpha 6$ -FucT gene families, in conjunction with the tandem organization of some genes in the genome and their conserved ORF-exon architecture, suggests that multiplicity among the $\alpha 3$ - and $\alpha 6$ -FucTs likely derived by successive segment duplications (rather than retrotranspositions) and that such duplications, especially among the $\alpha 6$ -FucTs, are relatively recent events. Indeed, Silva et al. [72], using a phylogenomic approach to identify lineage-specific gene duplications, concluded that expansion of the FucT gene family was among the most significant to have occurred in *S. mansoni*. In general, the downstream consequences for duplicated genes include nonfunctionalization (collection of degenerative mutations), neofunctionalization (attainment of a new function), and subfunctionalization (partitioning of the original function between duplicate genes) (reviewed by Hurles [73]). All three outcomes are

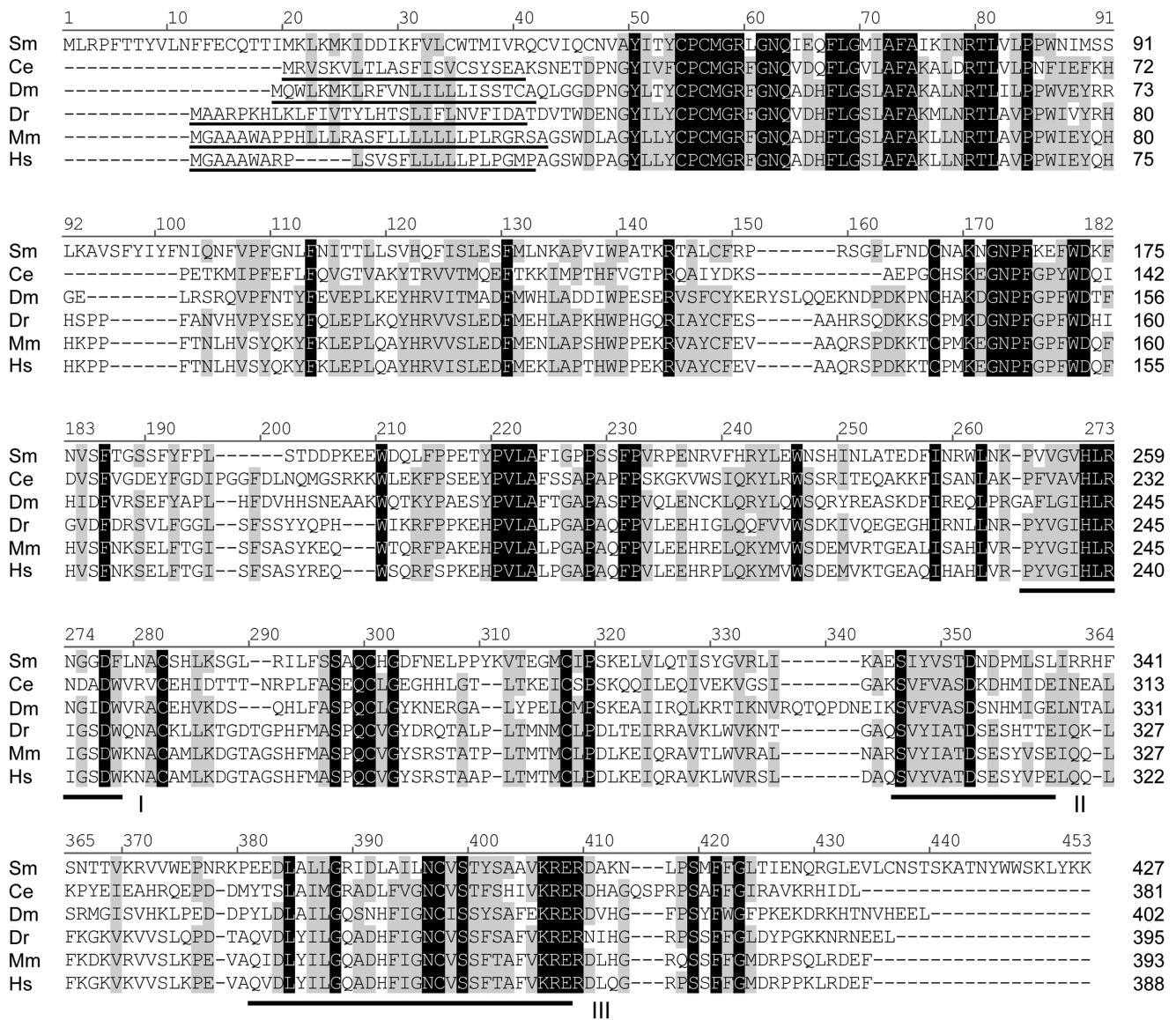


Figure 3. Amino acid alignment of protein O-fucosyltransferase 1 (O-FucT1) sequences. Schistosome POFucTA is compared to O-FucT1 sequences of *Caenorhabditis elegans* (Ce), *Drosophila melanogaster* (Dm), *Danio rerio* (Dr), *Mus musculus* (Mm), and humans (Hs) (NCBI accession numbers in Table S5). Alignment position is indicated above each block, and sequence length is reported to the right of each line. Positions of identity and positions exhibiting at least 70% conservation are highlighted in black and gray, respectively. The positions of three hydrophobic cluster analysis-derived motifs (I-III), which are shared features among $\alpha 2$ -, $\alpha 6$ - and protein O-FucTs [36,51,58], are indicated below the alignment blocks. Amino-terminal signal peptides (underlined) were identified using SMART and/or Phobius online tools. Vector NTI Advance 11.0 software alignment settings: BLOSUM45 matrix, gap opening penalty=12, gap extension penalty=0.1, gap separation penalty range=0, no residue-specific or hydrophobic residue gap.
doi:10.1371/journal.pone.0063299.g003

evident within the $\alpha 3$ -FucT gene family. The *FucTG* pseudogene is a likely example of nonfunctionalization, while neofunctionalization and subfunctionalization could be evidenced by gene-specific variations in acceptor specificity (as observed in *Caenorhabditis*; [38]) and developmentally regulated gene expression (described below; also see [21]), respectively. Furthermore, given the observed lack of $\alpha 2$ -FucT homologs in the schistosome genome and the expression of a unique Fuc $\alpha 1$ -2Fuc linkage [74], it is possible that one (or more) of the $\alpha 3$ -/ $\alpha 6$ - paralogs has neofunctionalized to add fucose in an $\alpha 2$ linkage instead of (or in addition to) the predicted $\alpha 3$ / $\alpha 6$ linkage. Indeed, previous studies have demonstrated that some forms of human FUT3 have the ability to

generate $\alpha 2$ linkages in addition to the usual $\alpha 3$ and $\alpha 4$ linkages [75,76].

Finally, a similar phylogenetic analysis incorporating three unverified homologs from a second human-infective schistosome, *S. japonicum* (predicted genes herein referenced by GenBank accession/SchistoDB annotation numbers CAX72936.1, CAX73054.1, and Sjp_0036210), generated a topologically concordant tree in which the putative *S. japonicum* FucTs form monophyletic groups with those of *S. mansoni* (Figure S3). Moreover, the phylogeny identifies *CAX72936.1* and *Sjp_0036210* as orthologs of *FucTH* and *FucTB*, respectively, and indicates that *CAX73054.1* shares a single ancestral node with

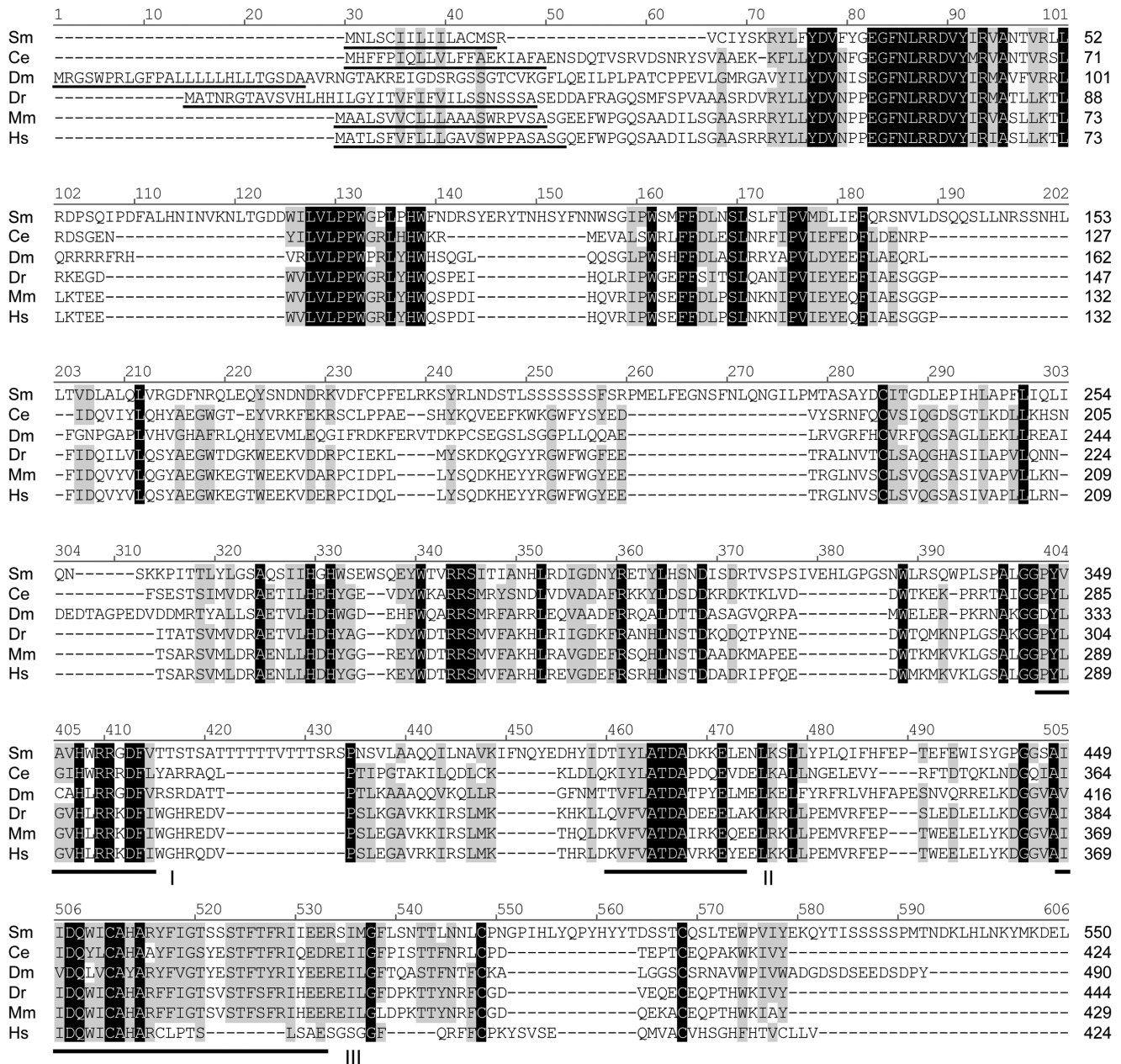


Figure 4. Amino acid alignment of protein O-fucosyltransferase 2 (O-FucT2) sequences. Schistosome PO-FucTB is compared to O-FucT2 sequences of *Caenorhabditis elegans* (Ce), *Drosophila melanogaster* (Dm), *Danio rerio* (Dr), *Mus musculus* (Mm), and humans (Hs) (NCBI accession numbers in Table S5). Alignment position is indicated above each block, and sequence length is reported to the right of each line. Identical and conserved (>70%) positions are highlighted in black and gray, respectively. The positions of three hydrophobic cluster analysis-derived motifs (I-III), which are shared features among members of the $\alpha 2$ -/ $\alpha 6$ -O-FucT superfamily [36,51,58], are indicated below the alignment blocks. An N-terminal signal peptide (underlined) was identified in all sequences using the SMART and/or Phobius servers. Note, the human O-FucT2 RefSeq protein used in the present analysis lacks motif III, however another version of human PO-FUT2 that includes this motif is available at NCBI (GenBank accession number AAH64623.1). Vector NTI Advance 11.0 software alignment settings: BLOSUM45 matrix, gap opening penalty = 12, gap extension penalty = 0.1, gap separation penalty range = 0, no residue-specific or hydrophobic residue gaps. doi:10.1371/journal.pone.0063299.g004

FucTs I-M. Interestingly, the topology within the schistosome $\alpha 6$ -clade implies a significant expansion of the $\alpha 6$ -FucT gene family in *S. mansoni* following the evolutionary separation of *S. mansoni* and *S. japonicum*. However, the complete repertoire of *S. japonicum* FucT genes has yet to be resolved and future investigations may identify additional $\alpha 6$ - (and $\alpha 3$ -) orthologs.

Real-time Quantitative PCR Analysis of $\alpha 3$ -FucT mRNA Expression in Miracidia and Primary Sporocysts

Given recent data demonstrating the abundant expression of fucosylated glycotopes in snail-associated schistosome larvae [10] and their predicted immunomodulatory roles in snail hosts, $\alpha 3$ -FucT transcript expression was assayed in miracidia and 2- and 10-day *in vitro*-cultivated primary sporocysts (Figure 6). Real-time

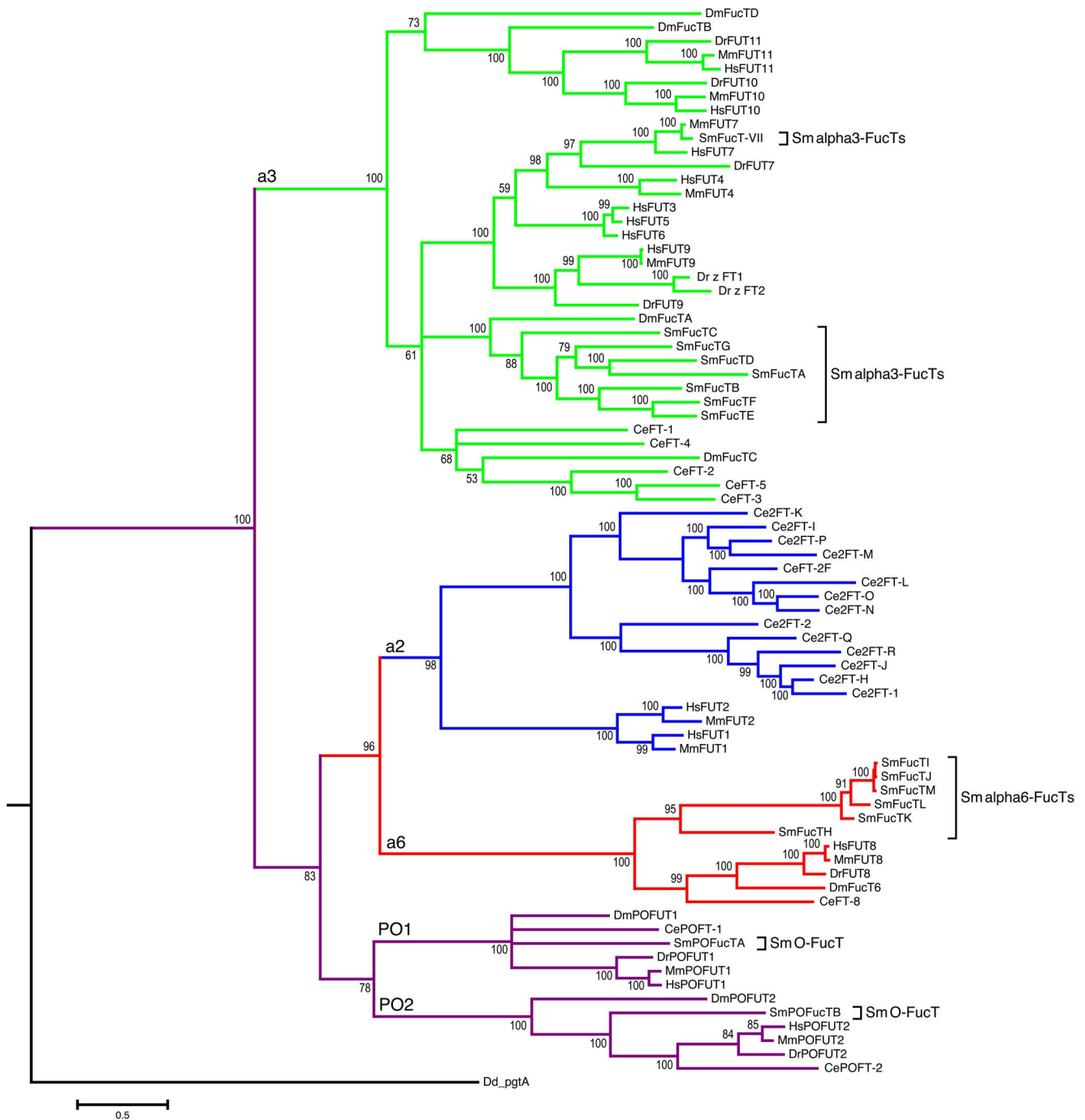


Figure 5. Phylogenetic tree of fucosyltransferases (FucTs). In addition to the FucTs of *Schistosoma mansoni* (Sm), the $\alpha 2$ -, $\alpha 3$ -, $\alpha 6$ - and protein O-FucTs from *Caenorhabditis elegans* (Ce), *Drosophila melanogaster* (Dm), *Danio rerio* (Dr), *Mus musculus* (Mm), and humans (Hs) were selected to represent the diversity of known, well-characterized FucTs (NCBI accession numbers and references in Table S5), and used to construct a molecular phylogeny rooted on the bifunctional $\beta 3$ -galactosyltransferase/ $\alpha 2$ -FucT PgtA of *Dictyostelium discoideum* (Dd). Posterior probabilities $\geq 50\%$ are indicated at each node, and genetic divergence (substitutions per site) is represented by the scale bar. doi:10.1371/journal.pone.0066329.g005

qPCR data indicate that *FucTA* and *FucTE* transcript abundance in larvae decreases as much as 74% during the miracidium-to-primary sporocyst transformation and remains low during *in vitro* cultivation. In contrast, expression of *FucTB* appears to stay high throughout larval transformation and only declines with extended cultivation (~40% reduction), while *FucTC* expression exhibits the opposite trend, initially dropping ~50% and then returning to

near miracidial levels by day 10 in culture. No significant variations in transcript abundance were observed for *FucTD* and *FucTF* in these experiments. During the miracidium-to-primary sporocyst transformation, miracidia shed their ciliated epidermal plates and the associated glycocalyx and form a syncytial tegument [77]. Peterson et al. [10] showed that miracidial epidermal plates are dominated by fucosylated glycotopes, many of which are lost

with the epidermal plates. Perhaps *FucTA* and *FucTE* are highly expressed in the epidermal plates and significant portions of their transcript populations are released with the plates. Similarly, the *FucTC* transcripts might also be lost during epidermal plate shedding, but transcription may increase in primary sporocysts during extended larval cultivation.

Previously, Fitzpatrick et al. [21] conducted a microarray analysis of FucT gene expression in all stages of the schistosome lifecycle, including miracidia and 2-day *in vitro*-cultivated primary sporocysts. Their results demonstrated that transcript levels for *FucTs D-F* are relatively high in the intramolluscan larval stages, with *FucTD* equally expressed between larvae, *FucTE* declining ~50% during transformation, and *FucTF* peaking in primary sporocysts. Transcript levels for *FucTA* and *FucTB* were relatively low in both miracidia and primary sporocysts. *FucTC* expression was not assessed. In all cases, FucT transcript abundance is elevated in adults, and for some genes (e.g., *FucTB*) array data indicate significant differences in expression between males and females. Unfortunately, the above observations in intramolluscan larvae are incongruous with the results of the present study. Disparities may be due to methodological differences between studies, however Fitzpatrick et al. [24] appropriately validated the array data for *FucTA* and *FucTD* using methods similar to those employed here (real-time qPCR but with normalization against alpha tubulin). Regardless of which dataset more accurately describes α 3-FucT gene expression in *S. mansoni*, both suggest that expression is developmentally regulated, which possibly contributes to the observed stage- and gender-specific expression of fucosylated glycotopes.

Conclusions

The present study used a genome-wide homology-based bioinformatics approach to identify and *in silico* characterize the complete repertoire of FucT homologs that presumably contribute to fucosylation, especially for the synthesis of terminal glycans, in *S. mansoni*. Our search yielded 15 complete genes, including seven α 3-FucTs, six α 6-FucTs and two protein O-FucTs. Why schistosomes encode such a large number of FucT homologs remains unclear, however it is thought that such duplicative expansions are an adaptive response to the parasitic lifestyle and imply important roles for these genes in schistosome development and immunobiology [72]. Notably, this level of redundancy also

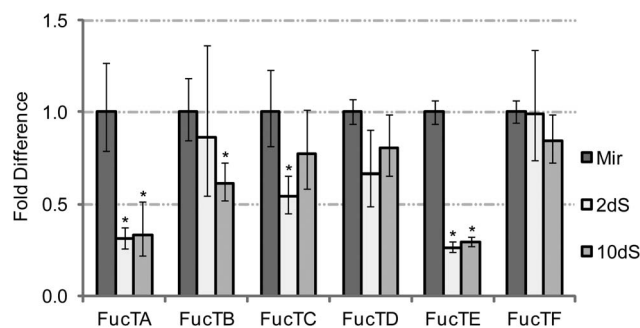


Figure 6. Alpha3-fucosyltransferase (α 3-FucT) gene transcription in larvae of *Schistosoma mansoni*. Real-time qPCR was used to examine steady-state levels of α 3-FucT transcription in miracidia (Mir) and 2- and 10-day *in vitro*-cultivated primary sporocysts (2 dS and 10 dS, respectively). Transcript abundance in primary sporocysts was assessed relative to miracidia, which was arbitrarily set at 1. Data represent the average of three independent biological replicates, and asterisks (*) indicate significantly altered gene transcription ($p < 0.05$). doi:10.1371/journal.pone.0063299.g006

exists in the non-parasitic nematode *Caenorhabditis* (see Figure 5; Table S5), which Oriol et al. [36] attributes to the evolutionary selection of fucosylation over sialylation as a means of terminating glycosylation. Indeed, sialic acid is absent from the *Caenorhabditis* glycans [78]. Likewise, sialic acid does not occur among the glycans of *Schistosoma* [79–81], indicating that fucosylation is the only means of terminal modification in schistosome glycosylation.

The observed redundancy in the α 6-FucT gene family alone is quite interesting given that most invertebrate and vertebrate species examined to date feature just one such gene. The singular known function of these genes is to add fucose in an α 6 linkage to the proximal GlcNAc of the N-glycan chitobiose core [18]. It is possible that some of the schistosome α 6- paralogs have neofunctionalized or are functionally compartmentalized (as with the α 3-FucTs of humans [82]), with FucTs featuring distinct expression patterns (i.e., tissue and stage specificity) and subtle variations in substrate utilization. Future studies should assess the tissue localization and stage-specificity of α 6-FucT expression across the schistosome lifecycle.

Strikingly, despite the prominence of a unique Fuc α 1-2Fuc linkage in schistosome glycoconjugates, no α 2-FucT homologs were identified in the present study. Because genomic sequence assembly for *S. mansoni* is not yet complete, with gaps still scattered throughout the genome [83], it is possible that α 2-FucT homologs are encoded but were not detected due to insufficient sequence information. However, given the uniqueness of the Fuc α 1-2Fuc linkage and the apparent lack of Fuc α 1-2Gal linkages in *S. mansoni*, the absence of a conventional α 2-FucT is not confounding. Alternatively, as described above, one of the predicted α 3- or α 6-FucTs may have neofunctionalized to create α 2 linkages or a novel enzyme may exist, unrelated to currently recognized α 2-FucTs, that serves this function. Future investigations should re-examine the schistosome genome for α 2-FucT sequences as new data are generated, as well as functionally test the present FucTs for α 2-fucosylation activity.

While protein O-FucTs are not directly involved in fucosylated glycotope expression, O-FucT1 and O-FucT2 play essential roles in diverse developmental and physiological processes. It is likely that these enzymes play many of the same roles in *S. mansoni*. A cursory search of the NCBI RefSeq database yielded schistosome homologs of O-FucT substrate-coding genes, including notch (accession number XM_002574857.1; [84]), ADAMTS5 peptidase (XM_002571852.1, [85]), properdin (XM_002580039.1; [86]), and f-spondin (XM_002581166.1; [86]). Future research should examine the role of schistosome O-FucTs in modifying these substrates and determine their significance in the context of schistosome development and immunobiology.

Unfortunately, attempts in this study to biochemically define the schistosome FucT homologs were unsuccessful, and assertions regarding their putative functions are based solely on our *in silico* analyses and are thus inherently speculative. Additional biochemical studies clearly are required to demonstrate FucT activities and link them to glycotope expression. However, the present work has provided an essential framework that will serve to inform and motivate future investigations exploring the role of fucosylation in schistosome development and immunobiology. Furthermore, this study highlighted several possible gene targets for the development of novel anti-schistosomal vaccines and chemotherapeutics.

Supporting Information

Figure S1 Diagrammatic depiction of fucosyltransferase (FucT) genomic organization in *Schistosoma mansoni*. The mRNA transcript sequences of schistosome FucTs were

mapped onto SchistoDB-derived genomic scaffolds. Exons (boxes, numbered below) and introns (connecting lines) are drawn to scale (bar = 1000 nt) with FucT-coding elements (segments of the prototypical ORF), including exons and a subset of retained introns, depicted as black boxes and non-coding exons depicted as gray boxes. Caret marks indicate gaps in the genomic sequence, and dotted lines represent introns of unknown length (spacing instead based on closely related FucTs). (TIF)

Figure S2 Diagrammatic depiction of fucosyltransferase (FucT) gene alternative splicing in *Schistosoma mansoni*. Alternative splicing, including exon skipping, intron retention, mutual exclusivity among exons and use of alternate splice donor and acceptor sites, was observed during transcript sequencing. Bent connectors indicate splicing between exons (boxes, numbered above), with solid lines representing splicing in the main/major full-length FucT-coding transcripts and dotted lines representing alternative splice events. Exons are drawn to scale (bar = 500 nt) and spacing of exons is arbitrary. Interexonic boxes represent retained introns (estimated lengths in parentheses), with solid outlines signifying retention in the main/major transcript and dotted lines indicating retention in other isoforms. Positions of the prototypical start and stop codons (AUG and UAA/UGA/UAG, respectively) are shown. Colors convey the *in silico* consequences of splicing: black, conservation of the prototypical ORF; red, introduction of a premature termination codon; orange, induction of a downstream frameshift; green, in-frame deletion/addition; blue, omission of the prototypical start/stop codon. Exon 4 of *FucTE* is tandemly duplicated (exon 5, white box); it is unknown if splice isoforms show preference for one copy versus the other. (TIF)

Figure S3 Phylogeny of fucosyltransferases (FucTs), including FucT homologs of *Schistosoma japonicum*. A phylogenetic tree was constructed using the maximum likelihood method and a GTR+ Γ substitution model implemented in RAxML v.7.3.4. The FucTs of *Schistosoma mansoni* (Sm; marked

by bars on right) as well as the $\alpha 2$ -, $\alpha 3$ -, and $\alpha 6$ - and protein O-FucTs from *Caenorhabditis elegans* (Ce), *Drosophila melanogaster* (Dm), *Danio rerio* (Dr), *Mus musculus* (Mm), and humans (Hs) were selected to represent the known FucT diversity (see for accession numbers). Additionally, three predicted FucTs of *Schistosoma japonicum* (labeled with GenBank accession/SchistoDB annotation numbers CAX72936.1, CAX73054.1, and Sjp_0036210) were included. The tree was rooted on the bifunctional $\beta 3$ -galactosyltransferase/ $\alpha 2$ -FucT PgtA of *Dictyostelium discoideum* (Dd). Numbers above or below branches indicate bootstrap support (%) estimated from 1,000 resamplings of the sequence data; bootstrap values $\leq 50\%$ are not shown. Genetic divergence (substitutions per site) is represented by the scale bar. (TIFF)

Table S1 Primers used for reverse transcriptase-PCR confirmation of fucosyltransferase gene transcription. (DOCX)

Table S2 Primers used for 5' and 3' rapid amplification of cDNA ends (RACE) of fucosyltransferase gene transcripts. (DOCX)

Table S3 Primers used for reverse transcriptase-PCR amplification of fucosyltransferase complete coding sequences. (DOCX)

Table S4 Primers used for quantitative PCR analyses of $\alpha 3$ -fucosyltransferase gene transcript expression. (DOCX)

Table S5 NCBI RefSeq/GenBank accession numbers (number.version) of referenced fucosyltransferase genes. (DOCX)

Author Contributions

Conceived and designed the experiments: NAP TPY. Performed the experiments: NAP TKA. Analyzed the data: NAP TKA. Contributed reagents/materials/analysis tools: TPY TKA. Wrote the paper: NAP. Reviewed and revised ms: TPY TKA NAP.

References

- Wilson MS, Mentink-Kane MM, Pesce JT, Ramalingam TR, Thompson R, et al. (2007) Immunopathology of schistosomiasis. *Immunol Cell Biol* 85: 148–154.
- Thomas PG, Harn DA (2004) Immune biasing by helminth glycans. *Cell Microbiol* 6: 13–22.
- Hokke CH, Yazdanbakhsh M (2005) Schistosome glycans and innate immunity. *Parasit Immunol* 27: 257–264.
- Hahn UK, Bender RC, Bayne C (2000) Production of reactive oxygen species by haemocytes of *Biomphalaria glabrata*: carbohydrate-specific stimulation. *Dev Comp Immunol* 24: 531–541.
- Castillo MG, Wu XJ, Dinguirard N, Nyame AK, Cummings RD, et al. (2007) Surface membrane proteins of *Biomphalaria glabrata* embryonic cells bind fucosyl determinants on the tegumental surface of *Schistosoma mansoni* primary sporocysts. *J Parasitol* 93: 832–840.
- Xu X, Stack RJ, Rao N, Caulfield J (1994) *Schistosoma mansoni*: fractionation and characterization of the glycocalyx and glycogen-like material from cercariae. *Exp Parasitol* 49: 399–409.
- van Remoortere A, Hokke CH, van Dam GJ, van Die I, Deelder AM, et al. (2000) Various stages of *Schistosoma* express Lewis^x, LacdiNAc, GalNAc β 1–4[Fuc α 1–3]GlcNAc, and GalNAc β 1–4[Fuc α 1–2Fuc α 1–3]GlcNAc carbohydrate epitopes: detection with monoclonal antibodies that are characterized by enzymatically synthesized neoglycoproteins. *Glycobiology* 10: 601–609.
- Wuhrer M, Kantelhardt SR, Dennis RD, Doenhoff MJ, Lochnit G, et al. (2002) Characterization of glycosphingolipids from *Schistosoma mansoni* eggs carrying Fuc(α 1–3)GalNAc-, GalNAc(β 1–4)[Fuc(α 1–3)]GalNAc- and Gal(β 1–4)[Fuc(α 1–3)]GalNAc- (Lewis X) terminal structures. *Eur J Biochem* 269: 481–493.
- Robijn ML, Wuhrer M, Kornelis D, Deelder AM, Geyer R, et al. (2005) Mapping fucosylated epitopes on glycoproteins and glycolipids of *Schistosoma mansoni* cercariae, adult worms, and eggs. *Parasitology* 130: 67–77.
- Peterson NA, Hokke CH, Deelder AM, Yoshino TP (2009) Glycotope analysis in miracidia and primary sporocysts of *Schistosoma mansoni*: differential expression during the miracidium-to-sporocyst transformation. *Int J Parasitol* 39: 1331–1344.
- Khoo KH, Chatterjee D, Caulfield JP, Morris HR, Dell A (1997) Structural mapping of the glycans from the egg glycoproteins of *Schistosoma mansoni* and *Schistosoma japonicum*: identification of novel core structures and terminal sequences. *Glycobiology* 7: 663–677.
- van Die I, Gomord V, Kooyman FN, van den Berg TK, Cummings RD, et al. (1999) Core alpha13fucose is a common modification of N-glycans in parasitic helminths and constitutes an important epitope for IgE from *Haemonchus contortus* infected sheep. *FEBS Letters* 463: 189–193.
- Paschinger K, Rendic D, Lochnit G, Jantsch V, Wilson IB (2004) Molecular basis of anti-horseradish peroxidase staining in *Caenorhabditis elegans*. *J Biol Chem* 279: 49588–49598.
- Wuhrer M, Koeleman CA, Fitzpatrick JM, Hoffmann KF, Deelder AM, et al. (2006) Gender-specific expression of complex-type N-glycans in schistosomes. *Glycobiology* 16: 991–1006.
- DeBose-Boyd R, Nyame AK, Cummings RD (1996) *Schistosoma mansoni*: characterization of an alpha 1–3 fucosyltransferase in adult parasites. *Exp Parasitol* 82: 1–10.
- Hokke CH, Neeleman AP, Koeleman CA, van den Eijnden DH (1998) Identification of an $\alpha 3$ -fucosyltransferase and a novel $\alpha 2$ -fucosyltransferase activity in cercariae of the schistosome *Trichobilharzia ocellata*: biosynthesis of the Fuc α 12Fuc α 13[Gal(NAc) β 14]GlcNAc sequence. *Glycobiology* 8: 393–406.
- Marques ET, Ichikawa Y, Strand M, August JT, Hart GW, et al. (2001) Fucosyltransferases in *Schistosoma mansoni* development. *Glycobiology* 11: 249–259.
- Paschinger K, Staudacher E, Stemmer U, Fabini G, Wilson IB (2005) Fucosyltransferase substrate specificity and the order of fucosylation in invertebrates. *Glycobiology* 15: 463–474.

19. Marques ET, Weiss JB, Strand M (1998) Molecular characterization of a fucosyltransferase encoded by *Schistosoma mansoni*. *Mol Biochem Parasitol* 93: 237–250.
20. Trottein F, Mollicone R, Fontaine J, de Mendonca R, Pillar F, et al. (2000) Molecular cloning of a putative alpha3-fucosyltransferase from *Schistosoma mansoni*. *Mol Biochem Parasitol* 107: 279–287.
21. Fitzpatrick JM, Peak E, Peraly S, Chalmers JW, Barrett J, et al. (2009) Anti-schistosomal intervention targets identified by lifecycle transcriptomic analyses. *PLoS Negl Trop Dis* 3: e543.
22. Yoshino TP, Laursen JR (1995) Production of *Schistosoma mansoni* daughter sporocysts from mother sporocysts maintained in syntenic culture with *Biomphalaria glabrata* embryonic (Bge) cells. *J Parasitol* 81: 714–722.
23. Nolan LE, Carriker JP (1946) Observations on the biology of the snail *Lymnaea stagnalis appressa* during twenty years of laboratory culture. *Am Midl Nat* 36: 467–493.
24. Chernin E (1963) Observations on hearts explanted *in vitro* from the snail *Austroloporbia glabrata*. *J Parasitol* 49: 353–364.
25. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25: 3389–3402.
26. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32: 1792–1797.
27. Maddison WP, Maddison DR (2011) Mesquite: a modular system for evolutionary analysis. Version 2.75 [http://mesquiteproject.org].
28. Price MN, Dehal PS, Arkin AP (2010) FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS ONE* 5: e9490.
29. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572–1574.
30. Nylander JAA, Wilgenbusch JC, Warren DL, Swofford DL (2008) AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics* 24: 581–583.
31. Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 7: 214. doi:10.1186/1471-2148-7-214.
32. Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22: 2688–2690.
33. Williams DL, Sayed AA, Bernier J, Birkeland SR, Cipriano MJ, et al. (2007) Profiling *Schistosoma mansoni* development using serial analysis of gene expression (SAGE). *Exp Parasitol* 117: 246–258.
34. Stekel DJ, Git Y, Falciani F (2000) The comparison of gene expression from multiple cDNA libraries. *Genome Res* 10: 2055–2061.
35. Zerlotini A, Heiges M, Wang H, Moraes RL, Domitini AJ, et al. (2009) SchistoDB: a *Schistosoma mansoni* genome resource. *Nucleic Acids Res* 37: D579–582.
36. Oriol R, Mollicone R, Cailleau A, Balanzino L, Breton C (1999) Divergent evolution of fucosyltransferase genes from vertebrates, invertebrates, and bacteria. *Glycobiology* 9: 323–334.
37. Loric C, Dupuy F, Rampal R, Dlugosz MA, Haltiwanger RS, et al. (2006) Molecular evolution of protein O-fucosyltransferase genes and splice variants. *Glycobiology* 16: 736–747.
38. Nguyen K, van Die I, Grundahl KM, Kawar ZS, Cummings RD (2007) Molecular cloning and characterization of the *Caenorhabditis elegans* α 1,3-fucosyltransferase family. *Glycobiology* 17: 586–599.
39. Natsuka S, Gersten KM, Zenita K, Kannagi R, Lowe JB (1994) Molecular cloning of a cDNA encoding a novel human leukocyte α -1,3-fucosyltransferase capable of synthesizing the sialyl Lewis x determinant. *J Biol Chem* 269: 16789–16794.
40. Javaud C, Dupuy F, Maftah A, Michalski JC, Oriol R, et al. (2000) Ancestral exonic organization of FUT8, the gene encoding the alpha6-fucosyltransferase, reveals successive peptide domains which suggest a particular three-dimensional core structure for the alpha6-fucosyltransferase family. *Mol Biol Evol* 17: 1661–1672.
41. Mollicone R, Moore SE, Bovin N, Garcia-Rosasco M, Candelier JJ, et al. (2009) Activity, splice variants, conserved peptide motifs, and phylogeny of two new α 1,3-fucosyltransferase families (FUT10 and FUT11). *J Biol Chem* 284: 4723–4738.
42. Abrantes J, Posada D, Guillon P, Esteves PJ, Le Pendu J (2009) Widespread gene conversion of alpha-2-fucosyltransferase genes in mammals. *J Mol Evol* 69: 22–31.
43. Kalsotra A, Cooper TA (2011) Functional consequences of developmentally regulated alternative splicing. *Nat Rev Genet* 12: 715–729.
44. Ram D, Ziv E, Lanter F, Lardans V, Schechter I (2004) Stage-specific alternative splicing of the heat-shock transcription factor during the life-cycle of *Schistosoma mansoni*. *Parasitology* 129: 587–596.
45. DeMarco R, Oliveira KC, Venancio TM, Verjovski-Almeida S (2007) Gender biased differential alternative splicing patterns of the transcriptional cofactor CA150 gene in *Schistosoma mansoni*. *Mol Biochem Parasitol* 150: 123–31.
46. Ma B, Simala-Grant JL, Taylor DE (2006) Fucosylation in prokaryotes and eukaryotes. *Glycobiology* 16: 158R–184R.
47. Xu Z, Vo L, Macher BA (1996) Structure-function analysis of human alpha1,3-fucosyltransferase. Amino acids involved in acceptor substrate specificity. *J Biol Chem* 271: 8818–8823.
48. Dupuy F, Germot A, Marendra M, Oriol R, Blancher A, et al. (2002) Alpha1,4-fucosyltransferase activity: a significant function in the primate lineage has appeared twice independently. *Mol Biol Evol* 19: 815–824.
49. Dupuy F, Petit JM, Mollicone R, Oriol R, Julien R, et al. (1999) A single amino acid in the hypervariable stem domain of vertebrate α 1,3/1,4-fucosyltransferases determines the type 1/type 2 transfer. Characterization of acceptor substrate specificity of the lewis enzyme by site-directed mutagenesis. *J Biol Chem* 274: 12257–12262.
50. Dupuy F, Germot A, Julien R, Maftah A (2004) Structure/function study of Lewis α 3- and α 3/4-fucosyltransferases: the α 1,4 fucosylation requires an aromatic residue in the acceptor-binding domain. *Glycobiology* 14: 347–356.
51. Breton C, Oriol R, Imbert A (1998) Conserved structural features in eukaryotic and prokaryotic fucosyltransferases. *Glycobiology* 8: 87–94.
52. Hokke CH, Deelder AM, Hoffmann KF, Wuhrer M (2007) Glycomics-driven discoveries in schistosome research. *Exp Parasitol* 117: 275–283.
53. Paschinger K, Guttermigg M, Rendic D, Wilson IB (2008) The N-glycosylation pattern of *Caenorhabditis elegans*. *Carbohydr Res* 343: 2041–2049.
54. Cummings RD (2009) The repertoire of glycan determinants in the human glycome. *Mol Biosyst* 5: 1087–1104.
55. Sonnhammer ELL, von Heijne G, Krogh A (1998) A hidden Markov model for predicting transmembrane helices in protein sequences. In: Glasgow J, Littlejohn T, Major F, Lathrop R, Sankoff D, Sensen C, editors. Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology. Menlo Park, CA: Association for the Advancement of Artificial Intelligence Press. 175–182.
56. Hofmann K, Stoffel W (1993) TMbase - A database of membrane spanning proteins segments. *Biol Chem Hoppe-Seyler* 347: 166.
57. Breton C, Mucha J, Jeanneau C (2001) Structural and functional features of glycosyltransferases. *Biochimie* 83: 713–718.
58. Martinez-Duncker I, Mollicone R, Candelier JJ, Breton C, Oriol R (2003) A new superfamily of protein-O-fucosyltransferases, α 2-fucosyltransferases, and α 6-fucosyltransferases: phylogeny and identification of conserved peptide motifs. *Glycobiology* 13: 1C-5C.
59. Takahashi T, Ikeda Y, Tateishi A, Yamaguchi Y, Ishikawa M, et al. (2000) A sequence motif involved in the donor substrate binding by α 1,6-fucosyltransferase: the role of the conserved arginine residues. *Glycobiology* 10: 503–510.
60. Wang Y, Shao L, Shi S, Harris RJ, Spellman MW, et al. (2001) Modification of epidermal growth factor-like repeats with O-fucose. Molecular cloning and expression of a novel GDP-fucose protein O-fucosyltransferase. *J Biol Chem* 276: 40338–40345.
61. Luo Y, Haltiwanger RS (2005) O-fucosylation of notch occurs in the endoplasmic reticulum. *J Biol Chem* 280: 11289–11294.
62. Luo Y, Koles K, Vorndam W, Haltiwanger RS, Panin VM (2006) Protein O-fucosyltransferase 2 adds O-fucose to thrombospondin type 1 repeats. *J Biol Chem* 281: 9393–9399.
63. Okajima T, Xu A, Irvine KD (2003) Modulation of notch-ligand binding by protein O-fucosyltransferase 1 and fringe. *J Biol Chem* 278: 42340–42345.
64. Sasamura T, Ishikawa HO, Sasaki N, Higashi S, Kanai M, et al. (2007) The O-fucosyltransferase O-fut1 is an extracellular component that is essential for the constitutive endocytic trafficking of Notch in *Drosophila*. *Development* 134: 1347–1356.
65. Schultz J, Milpetz F, Bork P, Ponting CP (1998) SMART, a simple modular architecture research tool: identification of signaling domains. *Proc Natl Acad Sci USA* 95: 5857–5864.
66. Käll L, Krogh A, Sonnhammer EL (2007) Advantages of combined transmembrane topology and signal peptide prediction—the Phobius web server. *Nucleic Acids Res* 35: W429–W432.
67. Teasdale RD, Jackson MR (1996) Signal-mediated sorting of membrane proteins between the endoplasmic reticulum and the golgi apparatus. *Annu Rev Cell Dev Biol* 12: 27–54.
68. Okajima T, Xu A, Lei L, Irvine KD (2005) Chaperone activity of protein O-fucosyltransferase 1 promotes notch receptor folding. *Science* 307: 1599–1603.
69. van der Wel H, Morris HR, Panico M, Paxton T, North SJ, et al. (2001) A non-Golgi α 1,2-fucosyltransferase that modifies Skp1 in the cytoplasm of *Dictyostelium*. *J Biol Chem* 276: 33952–33963.
70. van der Wel H, Fisher SZ, West CM (2002) A bifunctional diglycosyltransferase forms the Fuc α 1,2Gal β 1,3-disaccharide on Skp1 in the cytoplasm of *Dictyostelium*. *J Biol Chem* 277: 46527–46534.
71. Roos C, Kolmer M, Mattila P, Renkonen R (2002) Composition of *Drosophila melanogaster* proteome involved in fucosylated glycan metabolism. *J Biol Chem* 277: 3168–3175.
72. Silva LL, Marcet-Houben M, Nahum LA, Zerlotini A, Gabaldón T, et al. (2012) The *Schistosoma mansoni* phylome: using evolutionary genomics to gain insight into a parasite's biology. *BMC Genomics* 13: 617.
73. Hurlburt M (2004) Gene duplication: the genomic trade in spare parts. *PLoS Biol* 2: E206.
74. Khoo KH, Sarda S, Xu X, Caulfield JP, McNeil MR, et al. (1995) A unique multifucosylated -3GalNAc β 14GlcNAc β 13Galz1- motif constitutes the repeating unit of the complex O-glycans derived from the cercarial glycocalyx of *Schistosoma mansoni*. *J Biol Chem* 270: 17114–17123.
75. Gallet PF, Vaujour H, Petit JM, Maftah A, Oulmouden A, et al. (1998) Heterologous expression of an engineered truncated form of human Lewis fucosyltransferase (Fuc-TIII) by the methylotrophic yeast *Pichia pastoris*. *Glycobiology* 8: 919–925.

76. Chandrasekaran EV, Chawda R, Rhodes JM, Xia J, Piskorz C, et al. (2001) Human lung adenocarcinoma alpha1,3/4-L-fucosyltransferase displays two molecular forms, high substrate affinity for clustered sialyl LacNAc type 1 units as well as mucin core 2 sialyl LacNAc type 2 unit and novel alpha1,2-L-fucosylating activity. *Glycobiology* 11: 353–363.
77. Chiang CP, Caulfield JP (1988) *Schistosoma mansoni*: ultrastructural demonstration of a glycocalyx that cross-reacts with antibodies raised against the cercarial glycocalyx. *Exp Parasitol* 67: 63–72.
78. Bacic A, Kahane I, Zuckerman BM (1990) *Panagrellus redivivus* and *Caenorhabditis elegans*: evidence for the absence of sialic acids. *Exp Parasitol* 71: 483–488.
79. Nyame K, Smith DF, Damian RT, Cummings RD (1989) Complex-type asparagine-linked oligosaccharides in glycoproteins synthesized by *Schistosoma mansoni* adult males contain terminal beta-linked N-acetylgalactosamine. *J Biol Chem* 264: 3235–3243.
80. Nanduri J, Dennis JE, Rosenberry TL, Mahmoud AA, Tartakoff AM (1991) Glycocalyx of bodies versus tails of *Schistosoma mansoni* cercariae. Lectin-binding, size, charge, and electron microscopic characterization. *J Biol Chem* 266: 1341–1347.
81. Makaaru CK, Damian RT, Smith DF, Cummings RD (1992) The human blood fluke *Schistosoma mansoni* synthesizes a novel type of glycosphingolipid. *J Biol Chem* 267: 2251–2257.
82. de Vries T, Knegtel RM, Holmes EH, Macher BA (2001) Fucosyltransferases: structure/function studies. *Glycobiology* 11: 119R-128R.
83. Mourão MM, Grunau C, LoVerde PT, Jones MK, Oliveira G (2012) Recent advances in *Schistosoma* genetics. *Parasite Immunol* 34: 151–162.
84. Moloney DJ, Panin VM, Johnston SH, Chen J, Shao L, et al. (2000) Fringe is a glycosyltransferase that modifies Notch. *Nature* 406: 369–375.
85. Wang LW, Leonhard-Melief C, Haltiwanger RS, Apte SS (2009) Post-translational modification of thrombospondin type-1 repeats in ADAMTS-like 1/punctin-1 by C-mannosylation of tryptophan. *J Biol Chem* 284: 30004–30015.
86. Gonzalez de Peredo A, Klein D, Macek B, Hess D, Peter-Katalinic J, et al. (2002) C-mannosylation and o-fucosylation of thrombospondin type 1 repeats. *Mol Cell Proteomics* 1: 11–8.