Schistosoma mansoni α-N-acetylgalactosaminidase (SmNAGAL) regulates coordinated parasite movement and egg production

Benjamin J. Hulme[1], Kathrin K. Geyer[1], Josephine E. Forde-Thomas[1], Gilda Padalino[1], Dylan W. Phillips[1], Wannaporn Ittiprasert[2], Shannon E. Karinshak[2], Victoria H. Mann[2], Iain W. Chalmers[1], Paul J. Brindley[2], Cornelis H. Hokke[3] and Karl F. Hoffmann[1*]

[1]Institute of Biological, Environmental and Rural Sciences (IBERS), Edward Llwyd Building, Aberystwyth University, Aberystwyth, United Kingdom

[2]Department of Microbiology, Immunology and Tropical Medicine, Research Center for Neglected Diseases of Poverty, School of Medicine and Health Sciences, George Washington University, Washington, DC, United States of America

[3]Department of Parasitology, Leiden University Center for Infectious Diseases (LU-CID), Leiden University Medical Center (LUMC), Leiden, The Netherlands

* Corresponding author

E-mail: krh@aber.ac.uk

Abstract

α-galactosidase (α-GAL) and α-N-acetylgalactosaminidase (α-NAGAL) are two glycosyl hydrolases responsible for maintaining cellular homeostasis by regulating glycan substrates on proteins and lipids. Mutations in the human genes encoding either enzyme lead to neurological and neuromuscular impairments seen in both Fabry- and Schindler/Kanzaki-diseases. Here, we investigate whether the parasitic blood fluke *Schistosoma mansoni,* responsible for the neglected tropical disease schistosomiasis, also contains functionally important α-GAL and α-NAGAL proteins. As infection, parasite maturation and host interactions are all governed by carefully-regulated glycosylation processes, inhibiting *S. mansoni*'s α-GAL and α-NAGAL activities could lead to the development of novel chemotherapeutics. Sequence and phylogenetic analyses of putative α-GAL/α-NAGAL protein types showed Smp_089290 to be the only *S. mansoni* protein to contain the functional amino acid residues necessary for α-GAL/α-NAGAL substrate cleavage. Both α-GAL and α-NAGAL enzymatic activities were higher in females compared to males ($p<0.05$; α-NAGAL > α-GAL), which was consistent with *smp_089290*'s female biased expression. Spatial localisation of *smp_089290* revealed accumulation in parenchymal cells, neuronal cells, and the vitellaria and mature vitellocytes of the adult schistosome. siRNA-mediated knockdown (>90%) of *smp_089290* in adult worms significantly inhibited α-NAGAL activity when compared to control worms (*siLuc* treated males, $p<0.01$; *siLuc* treated females, $p<0.05$). No significant reductions in α-GAL activities were observed in the same extracts. Despite this, decreases in α-NAGAL activities correlated with a significant inhibition in adult worm motility as well as in egg production. Programmed CRISPR/Cas9 editing of *smp_089290* in adult worms confirmed the egg reduction phenotype. Based on these results, Smp_089290 was determined to act

predominantly as an α-NAGAL (hereafter termed SmNAGAL) in schistosome parasites where it participates in coordinating movement and oviposition processes. Pharmacological inhibition of SmNAGAL may lead to the development of a novel anthelmintic class of compounds.

Author summary

Schistosomiasis is a parasitic disease caused by infection with blood flukes, which leads to acute and chronic pathology in millions of infected individuals located in deprived tropical and subtropical regions. Elucidating the function of schistosome genes has provided a clearer view on their roles in various molecular pathways, which are critical to successful parasitism. This information is invaluable when progressing novel drug and vaccine candidates. Here, we add to the existing knowledge of the *Schistosoma mansoni* parasitic glycan processing and modification machinery by functionally characterising a glycosyl hydrolase (*S. mansoni* α-N-acetylgalactosaminidase, SmNAGAL). We demonstrate that this protein is enzymatically active and important in coordinating parasite movement in adult male and female schistosomes. Additionally, we provide evidence that this protein regulates pathways associated with egg production in female schistosomes, which is responsible for inducing pathological reactions. Developing drugs that inhibit SmNAGAL enzymatic activity could provide a novel approach for controlling schistosomiasis.

Introduction

Schistosomiasis is a parasitic disease caused by infection with blood flukes of the genus *Schistosoma* and is estimated to affect more than 250 million people worldwide [1]. With up to 200,000 deaths per annum, schistosomiasis is considered one of the most devastating neglected tropical diseases (NTDs) of resource poor communities within developing countries [2]. Currently, the primary method used to control schistosomiasis is mediated by praziquantel chemotherapy [3] with recent data suggesting that this drug targets transient receptor potential (TRP) channels expressed in adult schistosomes [4, 5]. The continued use of praziquantel in mass drug administration (MDA) programmes has given rise to the fear of possible drug resistant blood flukes developing. Indeed, multiple studies have demonstrated that praziquantel insensitive schistosomes can be generated in the laboratory [6-8]. These findings, thus, highlight the need for the development of novel anti-schistosomal drugs, which can act as praziquantel replacements if praziquantel-resistant schistosomes eventually arise. Identifying and inhibiting processes critical to schistosome developmental biology and/or host interactions presents a strategy that may contribute to this objective.

It is well established that schistosomes carefully regulate their production of specific glycans and glycan elements as they transition between diverse environmental niches including water, mammals and snails [9-12]. Some parasite glycans, as moieties of glycoproteins or glycolipids, are immunomodulatory in nature and contribute to host (snail/mammal) as well as dioecious (male/female schistosome) interactions [12-15]. Regardless of their significance in orchestrating these heterospecific, conspecific or developmental activities, the gene products responsible for schistosome glycoprotein and glycolipid anabolism and degradation are only slowly being characterised [16, 17]. One such

glycogene family, critical to glycosylation homeostasis in other eukaryotes, is the glycosyl hydrolases (or glycoside hydrolases, GH). These enzymes hydrolyse glycosidic bonds between monosaccharide residues in a carbohydrate chain or bonds between the carbohydrate and a non-carbohydrate moiety (such as a lipid or protein). According to the latest survey of glycogene diversity (Carbohydrate-Active enZYmes, CAZY, database [18]), the GHs are comprised of approximately 171 families (further segregated into clans [19]), each differing in substrate specificity and molecular mechanism.

Within GH family 27 are the related lysosomal enzymes α-galactosidases (α-GAL) and α-N-acetylgalactosaminidases (α-NAGAL), which are clan D members responsible for the cleavage of terminal α-D-galactose or α-N-acetylgalactosamine residues from glycosylated substrates [20]. Molecular defects in *Homo sapiens α-gal* lead to a lysosomal storage condition called Fabry disease, which is characterised by a range of multi-organ pathologies due to an accumulation of galactose-containing substrates (mostly aberrant glycolipids) in tissues [21]. Mutations in *H. sapiens α-nagal* lead to a clinically heterogeneous lysosomal disorder called Schindler/Kanzaki disease, which can contribute to neuromuscular defects and skin disorders due to accumulation of both glycolipids and glycoproteins [22, 23]. While these data strongly support the indispensable nature of both α-GAL and α-NAGAL activities in *H. sapiens*, RNA interference (RNAi) – mediated knockdown studies of *gana-1* (an ortholog of both *H. sapiens α-gal* and *α-nagal*) in *Caenorhabditis elegans* failed to demonstrate essentiality [24]. As the significance in loss of α-GAL and α-NAGAL activities may phenotypically vary among species, we initiated an investigation exploring the identification and characterisation of *S. mansoni* genes with similarity to these two members of GH family 27.

Of the five putative *S. mansoni* α-GAL/α-NAGAL proteins identified, we present sequence-, bioinformatic, functional genomic- and enzymatic- based evidence suggesting that Smp_089290 is the only α-GAL/α-NAGAL ortholog present in the genome (v 7.0) of this blood fluke. In the adult schistosome, *smp_089290* transcript abundance and enzymatic activity (α-NAGAL > α-GAL) were both enriched in the female compared to the male. Spatial localisation of *smp_089290* by whole-mount *in situ* hybridisation demonstrated accumulation to the vitellaria and vitellocytes in adult female schistosomes as well as diffuse distribution throughout parenchymal cells in both adult male and female parasites. Furthermore, RNAi-mediated knockdown of *smp_089290* led to a significant reduction in adult worm motility and egg production (confirmed by CRISPR/Cas9 genome editing). Collectively, our results predict that Smp_089290/SmNAGAL is a predominant α-NAGAL necessary for coordinated schistosome movement and oviposition processes.

Materials and methods

Ethics statement

All mouse procedures performed at Aberystwyth University (AU) adhered to the United Kingdom Home Office Animals (Scientific Procedures) Act of 1986 (project licenses PPL 40/3700 and P3B8C46FD) as well as the European Union Animals Directive 2010/63/EU, and were approved by AU's Animal Welfare and Ethical Review Body (AWERB). The use of mice at George Washington University Medical School was approved by The Institutional Animal Care and Use Committee (IACUC).

Parasite material

A Puerto Rican strain of *S. mansoni* (NMRI, Naval Medical Research Institute) was used in this study. Mixed-sex adult worms were perfused from percutaneously infected TO (HsdOla:TO, Envigo, UK), MF-1 (HsdOla:MF1, Envigo, UK) or Swiss Webster (CFW, Charles River Laboratories, USA) mice infected 7 wk earlier with 180 cercariae [25]. These parasites were cultured (10 sex-separated adults or five adult pairs per well in a 48 well tissue culture plate) at 37°C in DMEM (Sigma-Aldrich) supplemented with 10% foetal calf serum, 2 mM L-glutamine and 100 µg/ml penicillin/streptomycin in an atmosphere of 5% CO_2 with a 70% media exchange performed after 24 hr and 48 hr for adult male/adult pairs and female worms respectively. Cultured worms were subsequently used for RNA interference, CRISPR/Cas9 genome editing, total RNA isolation/complementary DNA (cDNA) generation, dual RNA and genomic DNA (gDNA) extraction and enzyme assays.

Multiple sequence alignment and phylogenetic trees

The Pfam identifier PF16499 (melibiase-2) was used to identify *S. mansoni*, *Schistosoma haematobium* and *Schistosoma japonicum* GH family 27 orthologs of *H. sapiens* α-GAL and α-NAGAL enzymes using WormBase ParaSite [26]. Only those *Schistosoma* entries that contained >70% of the residues within the melibiase-2 domain identified by Pfam, UniProt and CAZypedia were retained [18, 27, 28]. These included *S. mansoni* Smp_170840 (G4VLE1), *S. haematobium* MS3_10002 (A0A095A3Q8), *S. haematobium* MS3_10345 (A0A095B3P1), *S. mansoni* Smp_179250 (G4VLD7), *S. mansoni* Smp_247760 (A0A5K4F2K9), *S. mansoni* Smp_247750 (A0A5K4F165), *S. haematobium* MS3_10001 (A0A095CF43), *S. japonicum* EWB00_005283 (A0A4Z2D326), *S. japonicum* EWB00_005285 (A0A4Z2D213), *S. japonicum* EWB00_005284 (A0A4Z2D211), *S. haematobium* MS3_11280 (A0A095BSM8) and *S. mansoni* Smp_089290 (G4VLE3). A multiple sequence alignment of these *Schistosoma* melibiase-2 domain containing proteins with human α-GAL (*H. sapiens*, Hs_GAL, P06280), human α-NAGAL (*H. sapiens*, Hs_NAGAL, P17050), chicken α-NAGAL (*Gallus gallus*, Gg_NAGAL, Q90744) and *C. elegans* α-NAGAL (Ce_GANA-1, Q21801) was generated using MUSCLE v3.8 [29]. Amino acids included in the alignment were based on the positions of ligand binding and catalytic aspartic acid (D) residues of the melibiase-2 domains of human α-GAL and α-NAGAL [30-32].

Phylogenetic analysis of melibiase-2 domain containing proteins from a broad range of species across phyla was conducted using Bayesian inference and Maximum Likelihood approaches. MUSCLE v3.8 software was used to align proteins from *S. mansoni*, *S. haematobium*, *S. japonicum* (all noted previously) and additional sequences including *H. sapiens* Hs_GAL (P06280), *H. sapiens* Hs_NAGAL (P17050), *Drosophila melanogaster*

206 Dm_CG7997 (Q7K127), *D. melanogaster* Dm_CG5731 (Q8MYY3), *C. elegans* Ce_GANA-1
207 (Q21801), *Arabidopsis thaliana* At_AGAL1 (Q9FT97), *A. thaliana* At_AGAL2 (Q8RX86), *A.*
208 *thaliana* At_AGAL3 (Q8VXZ7), *Schizosaccharomyces pombe* Sp_MEL1 (Q9URZ0), *Coffea*
209 *arabica* Ca_GAL (Q42656), *G. gallus* Gg_NAGAL (Q90744), *Rattus norvegicus* Rn_NAGAL
210 (Q66H12), *Mus musculus* Mm_GAL (P51569) and *M. musculus* Mm_NAGAL (Q9QWR8). The
211 alignment was submitted to Gblocks Server 0.91b which removed non-conserved regions to
212 allow phylogenetic comparisons [33]. In total, 208 amino acid residues were used for this
213 phylogenetic analysis. Bayesian inference analyses were performed using MrBayes v3.2.6
214 [34], the WAG substitution model [35], 500,000 generations and a sample frequency of 100.
215 As the analysis proceeded, log likelihood values were generated and plotted against the
216 number of generations. The analysis was stopped upon examination of the average standard
217 deviation of split frequencies indicating when the log likelihood values reached a stationary
218 distribution. Maximum Likelihood analyses were performed using MEGA X [36] with the JTT
219 model [37] and 500 bootstrap replicates. The final Bayesian consensus phylogram was
220 generated using FigTree v1.4.3 [38]. The Bayesian inference tree was edited using Adobe
221 Illustrator CS4 software where Bayesian posterior probability support values (outside
222 parentheses) and Maximum Likelihood percentage bootstrap support values (inside
223 parentheses) were superimposed on corresponding nodes.
224

225 ## Homology modelling of Smp_089290

226 The three-dimensional structure of Smp_089290 was derived by homology modelling
227 using MODELLER [39]. The α-GAL/α-NAGAL template selected for Smp_089290 modelling was
228 the three-dimensional structure of *H. sapiens* α-NAGAL (Protein Databank identification code

4DO4) [31]. The sequence identity between Smp_089290 and the human α-NAGAL was 40% with a sequence coverage of 97%, hence well within the acceptable range for comparative modelling [40]. The stereochemical quality of the Smp_089290 model was assessed by RAMPAGE (Ramachandran Plot Analysis) [41], ProSA-web (Protein Structure Analysis) [42] and Verify 3D [43].

Total RNA isolation and cDNA generation

S. mansoni total RNA was isolated using the Direct-zol RNA MiniPrep Kit (Zymo Research) with slight modifications. Briefly, worms were removed from culture, transferred to 2 ml Eppendorf tubes, and homogenised in QIAZOL reagent (Qiagen) with a 5 mm diameter stainless steel bead (Qiagen) for a total of 6 min (2 x 3 min with 1 min on ice in between homogenisations) at 50 Hz using a TissueLyser LT (Qiagen). Thereafter, the RNA was treated with DNase I to remove contaminating gDNA. RNA was eluted into collection tubes by adding 30 µl of DNase/RNase-free H_2O directly to the Direct-zol column matrix and subsequently centrifuged at 17,000 x *g* for 60 sec. Yields of total RNA samples were assessed using a NanoDrop ND-1000 UV-Vis spectrophotometer; RNAs were concentrated when necessary in a Concentrator plus (Eppendorf) and re-quantified. Schistosome cDNAs were reverse transcribed from the total RNAs using the SensiFAST cDNA Synthesis Kit (Bioline) according to the manufacturer's instructions.

cDNA cloning and sequencing of *smp_089290*

PCR primers used to amplify the full coding sequence of Smp_089290 included forward (5´ – ATG GCT ACC GTA CCA CCG – 3´) and reverse (5´ – CTA TAA CAA TGT CTG AAA

CAG TCC ATC – 3′) pairs. PCR amplification utilised cDNA prepared from total RNA obtained from mixed-sex adult worms [44]. The amplification was performed in 25 µl containing Biomix reaction mix (Bioline) and ultra-stable Taq DNA polymerase (Bioline). Amplicons were subjected to electrophoresis through 1% w/v agarose, stained and visualised with SYBR Safe dye (Thermo Fisher Scientific) under UV light. Products of the predicted size of 1463 bp were ligated into pGEM-T Easy (Promega) after which α-Select Bronze Efficiency Competent Cells (Bioline) were transformed with the ligation products. Blue/white screening of *E. coli* colonies was accomplished on LB agar plates supplemented with 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) and Isopropyl β-D-1-thiogalactopyranoside (IPTG). Recombinant (white) clones were isolated from a 5 ml High Salt (HS) Luria Bertani (LB) culture containing ampicillin (50 µg/ml) using the QIAprep Spin Miniprep Kit (Qiagen). Inserts in pGEM-T Easy were sequenced at Aberystwyth University's Translational Genomics Facility and the sequence trace was analysed using FinchTV 1.4.0 software [45]. The full coding sequence of Smp_089290 is deposited under the GenBank accession number MZ508282.

DNA microarray analysis

The *S. mansoni* long oligonucleotide DNA microarray was designed and constructed by Fitzpatrick and colleagues [44]. The DNA microarray consists of 35,437 *S. mansoni* oligonucleotide 50-mers as well as 2,195 controls and is deposited in the ArrayExpress database under the accession number A-MEXP-830. *S. mansoni* lifecycle stages profiled by the DNA microarray included egg, miracidia, mother (2 day) sporocysts, daughter sporocysts, cercariae, 3 hr schistosomula, 24 hr schistosomula, 3 day schistosomula, 6 day schistosomula, 3 wk worms, 5 wk worms, 7 wk worms, male 7 wk worms and female 7 wk worms. Normalised

fluorescence intensity values (available via ArrayExpress under the experimental accession number E-MEXP-2094), corresponding to 50-mers representing exons 2 and 3 of *smp_089290*, enabled the quantification of gene expression profiles across these developmental stages of the schistosome.

RNA-Seq meta data analysis

The RNA-Seq meta database was created by Lu and colleagues [46] by normalising gene expression values derived from RNA-Seq data produced by various publications. Gene expression values for each lifecycle stage were obtained from the following reports: Anderson *et al*. [47] for egg, Wang *et al*. [48] for miracidia and sporocysts, Protasio *et al*. [49] for cercariae, 3 hr and 24 hr schistosomula, Protasio *et al*. [50] for 21 day juvenile male, 21 day juvenile female, 28 day juvenile male, 28 day juvenile female, 35 day adult male, 35 day adult female, 38 day adult male and 38 day adult female and Lu *et al*. [51] for 42 day adult male and 42 day adult female. The normalised gene expression values for a gene of interest were obtained by entering the gene ID into the 'schisto_xyz' search engine [52] and subsequently plotted as a gene expression profile.

RNA interference (RNAi)

Small interfering RNA (siRNA) duplexes were designed (si*Smp_089290*: sense strand 5'-CUA AUG AAA UCG UUG CAG A-3'; anti-sense strand 5'-UCU GCA ACG AUU UCA UUA G-3') based on the cDNA sequence verified *smp_089290* amplicon. An siRNA duplex designed for firefly luciferase (si*Luc*) without significant homology to gene products in the *S. mansoni* genome assembly (v7.0) served as a negative control [53]. Briefly, sets of 10 worms (sex-

298 separated) or five adult pairs were transferred to 0.4 cm pathway electroporation cuvettes
299 (Invitrogen) containing DMEM supplemented with 2 mM L-glutamine and 100 µg/ml
300 penicillin/streptomycin. siRNA duplexes (5 µg) were subsequently added and worms were
301 electroporated with a single pulse at 125 V (LV mode) for 20 ms using an ECM 830 Square
302 Wave Electroporation System (BTX, Harvard Apparatus, Holliston, MA). Worms were
303 subsequently transferred to a 48 well tissue culture plate and cultured for up to seven days.
304

305 ## Preparation of CRISPR/Cas9 plasmid constructs

306 Two CRISPR target sites (single guide RNA; sgRNA sequence) for Cas9-catalysed gene
307 editing for *smp_089290* were designed by Breaking-Cas [54] with default parameters
308 compatible for the protospacer adjacent motif (PAM) of Cas9 from *Streptococcus pyogenes*
309 (NGG) [55, 56]. Two sgRNAs targeting the coding regions exons 1 (SmNAGALX1: 5´ – CUA CCG
310 UAC CAC CGA UGG GU – 3´) and 2 (SmNAGALX2: 5´ – UUG UAA UCU AUG GCG UAU GC – 3´)
311 were used in this study. The sgRNAs contained >40% GC-content, no self-complementarity
312 and no off-target sites against the *S. mansoni* genome assembly (v7.0) as predicted by
313 Breaking-Cas software. A 20 nucleotide (nt) 'Scramble' sgRNA designed with low homology to
314 the *S. mansoni* genome assembly (v7.0) and lack of an adjacent PAM site (necessary for Cas9
315 function) served as a non-targeting control (5´ – GCA CUA CCA GAG CUA ACU CA – 3´).
316 CRISPR/Cas9 plasmid constructs were assembled using the pLenti-Cas-Guide construction Kit
317 (GE100010, OriGene, Maryland, USA) and each sgRNA was ligated into the pLenti-Cas-sgRNA
318 backbone as per the manufacturer's instructions. Expression of the sgRNA (SmNAGALX1,
319 SmNAGALX2 or Scramble containing plasmids) was driven by the mammalian U6 promoter
320 and expression of Cas9 from *S. pyogenes* with nuclear localisation signals was driven by the

human cytomegalovirus (CMV) promoter. Each CRISPR/Cas9 plasmid construct was independently transformed into DH5-α Chemically Competent Cells (GoldBio) with chloramphenicol (25 µg/ml) as drug selection. The chloramphenicol resistant transformants were confirmed for corrected orientation of sgRNA in the vector backbone by Sanger direct sequencing. To amplify the plasmid DNA, the transformants were first cultured in a 10 ml HSLB culture containing chloramphenicol (25 µg/ml) in a shaking incubator at 37°C, 225 rpm for approximately 7 hr. Thereafter, 10 ml of the culture was added to 240 ml of fresh HSLB culture containing chloramphenicol (25 µg/ml) and cultured overnight, as above. Plasmid DNA was recovered from 150 ml of this overnight culture using the GenElute™ HP Plasmid Midiprep Kit (Sigma-Aldrich) as per the manufacturer's instructions except for a modified elution step; here, accomplished with 800 µl of nuclease-free H_2O. Plasmid DNAs were quantified using the NanoDrop ND-1000 UV-Vis spectrophotometer (Thermo-Fisher Scientific), after which the DNA was precipitated and subsequently dissolved in Opti-MEM™ Reduced Serum Medium (Thermo-Fisher Scientific) and stored at -20°C.

CRISPR/Cas9 mediated genome editing

A mixture of 24 µg of CRISPR/Cas9 plasmid DNA (either SmNAGALX1, SmNAGALX2 or Scramble) reconstituted in a total volume of 400 µl Opti-MEM™ Reduced Serum Medium was dispensed into a 0.4 cm pathway electroporation cuvette (Invitrogen). Five pairs of worms were transferred to the cuvettes, which had been chilled on wet ice, and subjected to square wave electroporation with a single pulse of 125 V (LV mode) for 20 ms (BTX, Harvard Apparatus). Worms were subsequently transferred to a 48 well tissue culture plate and cultured for up to seven days at 37°C in an atmosphere of 5% CO_2. Upon completion of the

seven day experiment, RNA and gDNA was extracted from adult male and female worms using both RNAzol reagent (Sigma-Aldrich) and DNAzol reagent (Invitrogen) based on the manufacturer's instructions. One ml of RNAzol reagent was added to the parasite material, then subsequently homogenised four times with a 5 mm diameter stainless steel bead (Qiagen) for 3 min at 50 Hz using a TissueLyser LT (Qiagen) with 1 min incubation on ice in between homogenisations. Once the tissues of the schistosomes were completely disrupted, the protocol steps were followed until completion. RNA samples were immediately used for cDNA synthesis, as above. gDNA samples were amplified by PCR using primers encompassing the programmed double-strand break (DSB) site at exon 1 (Forward: 5´ – CTT ATA GGT GTG CCA TAT TAA CGA T – 3´, Reverse: 5´ – ATG CAC TAC ATT CGA AAG ACA – 3´) or exon 2 (Forward: 5´ – AGT GTT CTC ATG CAG TTA TCC T – 3´, Reverse: 5´ – TCC ATG TCA GCT GAG ATC A – 3´) of *smnagal*. Correct amplification was verified by 1% w/v agarose gel electrophoresis. Thereafter, amplicons were subjected to the QIAseq 1-Step Amplicon Library Kit (Qiagen) for Illumina compatible next generation sequencing (NGS) library construction with GeneRead DNAseq Targeted Panels V2 (Qiagen) as per the manufacturer's instructions. Amplicon size of each NGS library was verified using a 2100 Bioanalyzer (Agilent, Santa Clara, CA). The NGS libraries were quantified using the GeneRead Library Quant Kit (Qiagen). NGS libraries were pooled at the Genewiz NGS facility (Genewiz, NJ) and processed with a MiSEQ configuration of 2x250 bp. The demultiplexing data generated by the Genewiz NGS facility was exported as Fastq files (*.qz* format). The mutations around the cut sites were analysed by CRISPResso2 software using default parameters [57, 58]. Background mutations (i.e. mutations not attributable to genome editing) were inferred by identifying mutations present in the control samples and the treatment samples. The unique mutations around the programmed DSB sites only reported in the high quality sequence reads of the target sample were designated as

368 CRISPR/Cas9-induced mutations. The software analysed and reported the percentage of
369 indels (insertions and deletions) and substitutions in the target sample (inferred by the
370 number of reads that are modified in comparison to the total reads), the percentages of each
371 mutation type and a list of each specific mutation showing which nts have been altered.
372 Sequence reads from the NGS libraries are available at the Sequence Read Archive (SRA)
373 under BioProject ID PRJNA743897.
374

375 Quantitative reverse transcription (RT) – PCR (qRT-PCR) analysis

376 cDNAs synthesised from freshly perfused, siRNA treated or CRISPR/Cas9 plasmid
377 treated schistosomes were used as templates for qRT-PCR to analyse transcript abundance.
378 *smp_089290* transcript levels were quantified relative to α-tubulin (*smat1*) using a
379 StepOnePlus Real-Time PCR System (Applied Biosystems) and SensiFAST SYBR Hi-ROX mix
380 (Bioline). Total reaction volume was 10 µl with 150 nM of each primer, 5 µl of SensiFAST SYBR
381 Hi-ROX mix, 2 µl of cDNA template and 2.7 µl of nuclease-free H_2O. qRT-PCR primers for
382 *smp_089290* were designed based on the sequence verified *smp_089290* amplicon and
383 included forward (5'-CAC GAC TGA TGG TGG TGG-3') and reverse (5'-CTC GAT ACA TCA TTA
384 TCC CGC T-3') pairs. *smat1* primers included forward (5'-GGC GGT GGT ACT GGT TCT GGG-3')
385 and reverse (5'-CAT TTA GCG CAC CAT CGA AGC-3') pairs. To determine *smp_089290*
386 knockdown (KD) in adult worms following RNAi and CRISPR/Cas9 programmed knockout,
387 *smp_089290* transcript levels in siRNA treated samples (at 48 hr post electroporation) and
388 CRISPR/Cas9 plasmid treated samples (at seven days post electroporation) were quantified
389 relative to *smat1* as described previously [59]. The same calculations were used to quantify
390 *smp_089290* transcript levels in freshly perfused male and female schistosomes. Melting

curves (dissociation curves) were generated for each qRT-PCR analysis to verify the amplification of one product only. A two-tailed student's *t*-test was used to test for significance between siRNA treatments. A Kruskal-Wallis ANOVA with Dunn post-hoc comparisons was used to test for significance between CRISPR/Cas9 plasmid treatments. A two-tailed student's *t*-test was used to test for significance between freshly perfused adult male and female worms.

Whole mount *in situ* hybridisation (WISH)

The full coding sequence of Smp_089290 was amplified using forward (5´ – ATG GCT ACC GTA CCA CCG – 3´) and reverse (5´ – CTA TAA CAA TGT CTG AAA CAG TCC ATC – 3´) primer pairs. The PCR products were subsequently cloned into the pJC53.2 vector [60] using standard cloning methods as mentioned previously. These recombinant plasmids were subsequently used to generate digoxigenin-labelled riboprobes using the Riboprobe System (Promega) with SP6 or T3 RNA polymerases and digoxigenin-labelled Uracil triphosphate (Roche) [61]. Antisense treatment riboprobes (generated by SP6 polymerase) and sense control riboprobes (generated by T3 polymerase) were processed as described [61] and stored at -20°C until needed for the WISH staining protocol. Riboprobes were used within 2 wk of their initial storage at -20°C.

Adult male and female worms were relaxed and separated by incubation (15 min) in a 0.25% solution of the anaesthetic ethyl 3-aminobenzoate methanesulphonate (Sigma-Aldrich) dissolved in DMEM. Relaxed parasites were subsequently killed in a 0.6 M solution of $MgCl_2$ (1 min) and fixed for 4 hr in 4% formaldehyde in PBSTx (PBS + 0.3% Triton X-100). Fixed parasites were dehydrated in MeOH and stored at -20°C until needed. Parasite samples were

rehydrated in 1:1 MeOH:PBSTx (5-10 min) followed by incubation in PBSTx (5-10 min). Rehydrated parasite samples were bleached in formamide bleaching solution (0.5% v/v de-ionised formamide, 0.5% v/v SSC and 1.2% w/w H_2O_2, brought to a final volume of 10 ml with diethyl pyrocarbonate H_2O) for 90 min, rinsed twice in PBSTx, treated with proteinase K (10 µg/ml, Invitrogen) for 45 min at room temperature and post-fixed for 10 min in 4% formaldehyde in PBSTx. Parasite samples were processed as previously described [60, 61] with several modifications. Antisense treatment riboprobes and sense control riboprobes were mixed with hybridisation solution (50% v/v de-ionised formamide, 10% w/v dextran sulphate, 1.25% v/v SSC, 1 mg/ml yeast RNA, 1% v/v Tween-20, brought to a final volume of 40 ml with diethyl pyrocarbonate H_2O) and hybridised overnight at 52°C. Parasite samples were transferred to fresh colorimetric developing solution consisting of alkaline phosphatase buffer (100 mM Tris pH 9.5, 100 mM NaCl, 50 mM $MgCl_2$, 0.1% v/v Tween-20, brought to a final volume of 10 ml with 10% polyvinylalcohol solution) supplemented with 4.5 µl/ml NBT (Roche) and 3.5 µl/ml BCIP (Roche). All parasite samples treated with the antisense treatment riboprobe were developed at room temperature in the dark until the desired level of purple signal was reached (male samples = 2 hr, female samples = 45 min). In parallel, all parasite samples treated with the sense control riboprobe were developed for the same length of time. Once the desired level of signal was reached, the colorimetric developing solution was removed and worms were rinsed twice in PBSTx to stop any further development from occurring. Worms were dehydrated in 100% ethanol for 5 min and subsequently submerged in 80% glycerol in 1x PBS and incubated overnight at 4°C. Thereafter, stained worms were mounted onto microscope slides and examined with a light microscope (Leica LMD6000 Laser Microdissection Microscope).

Single-cell RNA-Seq (scRNA-Seq) analysis

Localisation of *smp_089290* (*smnagal*) found within the 68 adult worm clusters generated from available scRNA-Seq data [62] was depicted as uniform manifold approximation and projection (UMAP) plots using Seurat V3 [63].

α-GAL/α-NAGAL enzymatic activity measurements

Soluble worm antigen preparation (SWAP) was derived from worms (freshly perfused or electroporated with siRNAs) removed from culture after 1 wk. Worms were homogenised with a 5 mm diameter stainless steel bead (Qiagen) for 4 min at 50 Hz in 100 µl of 0.15 M McIlvaine buffer pH 4.6 [64] with EDTA-free protease inhibitors (Sigma-Aldrich) using a TissueLyser LT (Qiagen). After homogenisation, tubes were centrifuged at 21,100 x *g* for 30 min at 4°C and SWAP was collected and quantified by the Bradford method (Sigma-Aldrich). The enzymatic activity of SWAP derived from freshly perfused adult male and female worms, si*Luc* treated and si*Smp_089290* treated adult male and female worms was measured using 4-Nitrophenyl α-D-galactopyranoside (α-GAL colorimetric substrate) and 4-Nitrophenyl N-acetyl-α-D-galactosaminide (α-NAGAL colorimetric substrate) in separate reactions. Differing concentrations of human α-GAL (Fabrazyme, kindly provided by the Leiden University Medical Centre, LUMC) and commercially sourced α-NAGAL cloned from *Chryseobacterium meningosepticum* and expressed in *E. coli* (New England Biolabs, Ipswich, MA) were measured with corresponding substrates.

Enzyme assays were performed in standard flat-bottomed 96 well plates (STARLAB). α-GAL/α-NAGAL assays consisted of 100 µl of α-GAL or α-NAGAL substrate dissolved in 0.15 M McIlvaine buffer (pH 4.6) with the addition of differing concentrations of α-GAL or α-NAGAL

and 0.15 M McIlvaine buffer pH 4.6 to give a final volume of 125 µl per well. Concentrations of α-GAL used to generate a standard curve of activity were as follows: 0.95, 0.475, 0.19, 0.0475, 0.0095, 0.00475 and 0.0019 µg/ml. Concentrations of α-NAGAL used to generate a standard curve of activity were as follows: 20, 15, 10, 5, 2 and 1 µg/ml. SWAP reactions were set-up in the same way but α-GAL/α-NAGAL enzymes were replaced with different quantities of sample specific SWAP and incubated with both α-GAL and α-NAGAL substrates. For untreated adult male and female worms, 5 µg of SWAP was used per well to enable comparison between gender. For si*Luc* treated and si*Smp_089290* treated adult male and female worms, 6.45 µg and 2.44 µg of SWAP was used per well to enable comparison between siRNA treatments, respectively. Reaction proceeded for 60 min at 37°C and terminated by the addition of 70 µl of 0.4 M glycine (pH 10.4) [64]. Final absorbances were quantified at 410 nm using a POLARstar Omega microplate reader (BMG Labtech). Absorbance values produced from different SWAP treatments were compared to α-GAL/α-NAGAL standard curves to calculate α-GAL and α-NAGAL activities (µg/ml). A two-tailed student's *t*-test was used to test for significance between siRNA treated samples or between genders.

Motility analysis quantified by WormAssayGP2 and adult worm scoring matrix

The digital image processing-based system known as WormAssayGP2 was derived from Marcellino *et al*. [65] and implemented by us as previously described [66]. Individual wells containing up to 10 adult worms were recorded for 60 sec each day and analysed by the Lucas-Kanade algorithm. Once recording was completed, the data were quantified and stored as a *.csv* file which was further processed to calculate the mean motility for the control and

484 treatment groups. In parallel, adult worm motility was also scored using the WHO-TDR scoring
485 matrix [67]. All worms were scored daily from two days after electroporation until seven days
486 after electroporation. Worms were ranked between 4 – 0 based on their motility; 4 = normal
487 active/paired up, 3 = full body movement but slowed activity, 2 = minimal activity, occasional
488 movement of head and tail only, 1 = movement in the suckers only or slight contraction of
489 the body and 0 = total absence of motility. For both motility analyses, a General Linear Mixed-
490 Effects Model was fitted to each dataset ('NLME' package) and statistical differences were
491 determined by performing pairwise comparisons of the estimated marginal means of each
492 group per time point ('EMMEANS' package) in R. Video footage of adult worms was captured
493 using a NexiusZoom stereo microscope (Euromex) and edited with ImageFocus 4 software
494 (Euromex).
495

496 Enumeration of vitellocytes and egg volume measurements

497 The eggs released by adult female worms from RNAi and CRISPR/Cas9 genome editing
498 experiments were collected and fixed in 1 ml 10% neutral buffered formalin solution (Sigma-
499 Aldrich) for 24 hr at 4°C. Thereafter, eggs were enumerated using a Sedgewick Rafter
500 Counting Chamber [68]. Prior to visualisation by laser scanning confocal microscopy (LSCM),
501 stored eggs from RNAi experiments only were immersed in PBS supplemented with DAPI (4',6-
502 diamidino-2-phenylindole, 2 µg/ml). Fluorescence images (10 eggs per treatment) were
503 captured on a Leica TCS SP8 super resolution laser confocal microscope fitted with a 63 X
504 (water immersion) objective using the Leica Application Suite X (LAS X). Green (egg
505 autofluorescence) fluorescence was visualised with an argon or diode-pumped, solid state
506 (DPSS) laser at 488 nm. DAPI was visualised using a 405 nm blue diode laser. The number of

vitellocytes ($DAPI^+$ cells) and overall volume (mapped by the green autofluorescence) for individual eggs were calculated using IMARIS 7.3 software (Bitplane). A two-tailed student's *t*-test was used to test for significance between siRNA treated samples with regard to the total number of eggs produced and individual egg volume. A Kruskal-Wallis ANOVA with Dunn post-hoc comparisons was used to test for significance between CRISPR/Cas9 plasmid treated samples with regard to the total number of eggs produced. IMARIS 7.3 software (Bitplane) was also used to create a video showing the 360° horizontal rotation of a representative egg from each siRNA treatment.

Results

Schistosoma mansoni contains a single α-N-acetylgalactosaminidase (SmNAGAL)

α-N-acetylgalactosaminidase (α-NAGAL) is a member of the glycosyl hydrolase (GH) 27 family and contains both a melibiase-2 (PF16499) and a melibiase-2 C-terminal (PF17450) domain. As all enzymatically important (ligand binding and catalytic) amino acids are conserved within the melibiase-2 domain [32], we focused our interrogation of the *S. mansoni* genome (v7.0) for the presence of putative homologs that contained >70% of the residues within this domain. Our analysis revealed the presence of five schistosome members that met this criterion: Smp_170840, Smp_179250, Smp_247760, Smp_247750 and Smp_089290 (**Fig 1**).

Fig 1. *Schistosoma mansoni* contains five GH27 family members, but only Smp_089290 contains all residues necessary for α-NAGAL substrate binding and cleavage. A concatenated multiple sequence alignment of α-GAL and α-NAGAL proteins from *H. sapiens* (Hs), *G. gallus* (Gg), *C. elegans* (Ce), *S. japonicum* (EWB), *S. mansoni* (Smp), and *S. haematobium* (MS3) throughout the melibiase-2 domain (PF16499). Numbers located at the beginning of each sequence represent the amino acid position in the protein sequence. Ligand binding residues are highlighted green whereas non-conserved amino acids in the same position in other sequences are white. Catalytic aspartic acid (D) residues are highlighted light blue whilst non-conserved amino acids in the same position in other sequences are highlighted red. Amino acid residues which are missing from 'Ce_GANA-1', 'Smp_170840' and 'MS3_10001' are indicated with - whereas N/A is used for the amino acid position.

Among these five *S. mansoni* gene products, only one (Smp_089290, highlighted yellow, **Fig 1**) contained all 13 ligand binding residues (green shaded amino acids, **Fig 1**) as well as both catalytic aspartic acid residues (light blue Ds, **Fig 1**) critical for α-NAGAL activity [31]. MS3_11280 (*S. haematobium* homolog) and EWB00_005284 (*S. japonicum* homolog) also shared these diagnostic characteristics. Furthermore, while two additional *S. japonicum* homologs (EWB00_005283 and EWB00_005285) contained both catalytic aspartic acid residues, they did not possess all 13 ligand binding residues. These five schistosome proteins, containing the essential catalytic aspartic acid residues, clustered into a separate clade discrete from other putative schistosome GH27 family members (**Fig 2**).

Fig 2. *Schistosoma* melibiase-2 domain containing proteins with conserved catalytic aspartic acid residues cluster in a distinct clade of GH27 family members. Phylogenetic analyses were conducted using a concatenated multiple sequence alignment of *Schistosoma* proteins that contained >70% of the residues within the melibiase-2 domain and α-GAL/α-NAGAL protein types across phyla. Proteins were analysed using both Maximum Likelihood and Bayesian inference approaches. Branch lengths (indicated by scale bar) represent distance among different taxa as predicted by the Bayesian inference approach. Node labels outside parentheses represent Bayesian posterior probability support values whilst those within parentheses represent percentage bootstrap support values from Maximum Likelihood analysis. *Schistosoma* proteins conserving catalytic aspartic acid residues are highlighted in the blue dashed box. The phylogram includes protein sequences from *S. mansoni* (Smp), *S. haematobium* (MS3), *S. japonicum* (EWB), *A. thaliana* (At), *C. arabica* (Ca), *C. elegans* (Ce), *D.*

574 *melanogaster* (Dm), *G*. *gallus* (Gg), *H*. *sapiens* (Hs), *R*. *norvegicus* (Rn), *M*. *musculus* (Mm) and
575 *S*. *pombe* (Sp).
576
577 The predicted three-dimensional structure of Smp_089290 was derived by homology
578 modelling and passed all stereochemical quality assessments (**S1 Fig**). Homology modelling of
579 Smp_089290 as a monomer (HsNAGAL functions as a homodimer [31]) revealed the
580 positioning of these 13 ligand binding residues and two catalytic residues around the putative
581 active site (purple, **Fig 3**). Despite the melibiase-2 C-terminal domain (red, **Fig 3**) not
582 possessing any residues involved in ligand binding and substrate cleavage mechanisms, the
583 predicted Smp_089290 model and *H. sapiens* α-NAGAL structure are consistent across this
584 region. Both the predicted Smp_089290 model and *H. sapiens* α-NAGAL structure possess
585 eight anti-parallel β-strands (**Fig 3**) [32]. Overall, the homology modelling analysis predicted
586 that Smp_089290 likely possesses folding topology and spatial arrangements that closely
587 resemble typical α-GAL and α-NAGAL proteins.
588
589 **Fig 3. Comparisons of the catalytic active site pockets found in the Smp_089290 homology**
590 **model and the crystal structure of human α-NAGAL.** The grey dashed boxes depict a close-
591 up view of the catalytic active site pocket within the Smp_089290 homology model's and *H*.
592 *sapiens* α-NAGAL crystal structure's melibiase-2 domain (purple). The close-up views label all
593 13 ligand binding residues (green) and two catalytic Asp residues (light blue, underlined with
594 double line) with their corresponding amino acid positions in each of the protein sequences
595 shown. The red arrows show the location of the eight anti-parallel β-strands found near the
596 C-terminus of the Smp_089290 homology model's and the *H*. *sapiens* α-NAGAL crystal

structure's melibiase-2 C-terminal domain (red). Black arrows show the locations of the N and C-termini of the model/structure.

Together, these data suggest that the *S. mansoni* genome encodes a single gene product (Smp_089290) containing all essential catalytic amino acid residues for hydrolysis and release of α-N-acetylgalactosamine from glycosylated substrates. Therefore, this putative schistosome α-N-acetylgalactosaminidase (SmNAGAL, Smp_089290) was taken forward for further transcriptional, enzymatic, and functional genomics studies.

Smnagal is developmentally regulated, female-enriched and localised to vitellaria, mature vitellocytes and parenchymal cells

To begin deciphering SmNAGAL function, both transcriptomic- and enzymatic-based approaches were initiated (**Fig 4**). Meta-analysis of historical DNA microarray data across the *S. mansoni* lifecycle [44] was facilitated by two 50-mer oligonucleotide probes that retained 100% base-pair complementarity to exon 2 (CONTIG7235) and exon 3 (CONTIG6265), respectively, of *smnagal* (**Fig 4A**). For each of these two oligonucleotides, similar patterns of *smnagal* abundance were deduced. While *smnagal* expression was low in eggs, it increased in miracidia only to wane as schistosome development (sporocysts - cercariae) continued in the molluscan host (**Fig 4B**). Upon early intra-mammalian schistosome maturation (3 hr - 3 day schistosomula), *smnagal* transcription remained invariably low, until day six post schistosomula transformation. At this point and extending into more developmentally mature lifecycle forms (3 wk – 7 wk schistosomes), *smnagal* expression increased, reaching peak abundance in 7 wk old schistosomes. Here, female-enriched *smnagal* expression was clearly

observed in adult schistosomes, which was independently confirmed by qRT-PCR (**Fig 4C**). Additionally, female-biased expression of *smnagal* was observed when interrogating *S. mansoni* RNA-Seq meta data (**S2 Fig**) [46]. Further support for female biased expression of *smnagal* was obtained from the analysis of α-N-acetylgalactosaminidase activity in adult worm protein extracts (**Fig 4D**). In a direct comparison, female protein extracts contained significantly higher levels of α-NAGAL activity when compared to males. This female biased trend was also observed with α-GAL activity measured in the same extracts (**Fig 4D**).

Fig 4. *smnagal* (*smp_089290*) expression and α-NAGAL/α-GAL activities are female-enriched. (A) Diagrammatic representation of *smnagal*/*smp_089290* gene structure with 50-mer oligonucleotides mapped. Exons are depicted as red boxes, which are linked by lines representing introns. Numbers written inside each exon represent their position in the gene sequence. Numbers written below exons and above introns represent their length in base pairs. The 5' and 3' ends are shown above exon 1 and 3, respectively. The positions of oligonucleotide 50-mers corresponding to CONTIG7235 and CONTIG6265 are shown above exon 2 and 3 respectively. (B) DNA microarray analysis of *smp_089290* expression across 15 lifecycle stages. DNA microarray gene expression profile consisted of normalised mean fluorescence intensities of *smnagal*/*smp_089290* transcript abundance derived from oligonucleotides CONTIG7235 and CONTIG6265 as described previously [44]. (C) *smnagal*/*smp_089290* transcript levels in untreated adult male and female were quantified relative to *smat1* to validate normalised mean fluorescence intensities produced in 7 wk male and 7 wk female schistosomes. Statistical significance is indicated (Student's *t*-test, two tailed, unequal variance, * = $p<0.05$). (D) Equal quantities of SWAP (5 μg per well) were used for both sexes and measured for α-NAGAL and α-GAL activity on diagnostic α-NAGAL and α-GAL

644 substrates. Statistical significance is indicated (Student's *t*-test, two tailed, unequal variance,
645 * = $p<0.05$ and ** = $p<0.01$).
646
647 Expanding these temporal investigations of *smnagal* abundance and enzymatic
648 activities to spatial localisation, by WISH and scRNA-Seq, in adult worms revealed additional
649 gender-specific traits. An antisense RNA probe spanning the full length coding sequence of
650 *smp_089290* was used for each WISH localisation experiment (**Fig 5**) [61]. A negative control
651 was prepared using a sense *smp_089290* probe (**S3 Fig**); no specific staining was observed.
652 While the ovary (containing mature and immature oocytes) was not a rich source of *smnagal*
653 expression, the vitellarium (and mature vitellocytes passing through the vitello-oviduct) was
654 highly enriched for this putative α-NAGAL gene product (**Fig 5A**). These observations were
655 supported by scRNA-Seq approaches, which showed prominent *smnagal* abundance within
656 the vitellaria and mature vitellocytes and no expression within female gametes (**S4 Fig**). In
657 addition, low levels of *smnagal* expression were found within the parenchyma, which was
658 more clearly observed in the female scRNA-Seq plots (**S4 Fig**). Moderate *smnagal* expression
659 was also observed in some clusters of neuronal cells, tegument lineage cells and muscle cells
660 when inspecting the female scRNA-Seq plots. No appreciable expression was found in any
661 other female tissue examined. The WISH analysis of males revealed that *smnagal* expression
662 was predominantly localised to parenchymal cells widely distributed throughout the body (**Fig**
663 **5B**), which was confirmed by the male scRNA-Seq expression profile plots (**S5 Fig**). A lack of
664 staining was additionally observed throughout the primary reproductive organs (testes),
665 which was supported by scRNA-Seq profiles. WISH analyses also showed that cells lining the
666 tegument and intestine lacked intense *smnagal* expression in both genders, although this was
667 easier to deduce in males. Similar to females, small yet noticeable levels of *smnagal*

expression were evident in some neuronal cell clusters when inspecting male scRNA-Seq plots (**S5 Fig**). Appreciable expression was not observed elsewhere in any other male tissue examined.

Fig 5. *smnagal* expression is concentrated in the vitellarium, mature vitellocytes and parenchymal cells. Micrographs of the anterior, mid-section and posterior (10x magnification) of (A) female and (B) male schistosomes as well as anterior images with a higher magnification (40x magnification, area depicted by black dashed box). Structures labelled include egg (E), ovary (O), vitellarium (V), vitello-oviduct (VOD), intestine (I), oral sucker (OS), oesophagus (OES), ventral sucker (VS) and testes (TES). Black scale bars = 200 µm and red scale bars = 50 µm.

α-N-acetylgalactosaminidase activity is required for worm motility, egg production and development

The localisation of *smnagal* to adult parenchymal cells, neuronal clusters and mature vitellocytes as well as the well-documented neuromuscular defects characteristic of Schindler/Kanzaki disease (due to *α-nagal* deficiencies, [22, 23]) implicated key roles for this gene product in schistosome motility, oviposition and development. Therefore, to assess whether these processes were dependent upon α-NAGAL activity, functional genomics investigations of *smnagal*/*smp_089290* were conducted in adult schistosomes (**Fig 6**). RNA interference (RNAi) of *smnagal*/*smp_089290*, using small interfering RNAs (siRNAs), led to a highly significant knockdown (92%) of *smnagal* in adult male worms when compared to controls (**Fig 6A**). In parallel, genome editing approaches were implemented with

CRISPR/Cas9 plasmid constructs and used to complement RNAi experiments. The CRISPResso2 pipeline was utilised to quantify NHEJ-associated mutations within aligned MiSEQ library sequencing reads from genome-edited worms (**S1 Table**). Oligonucleotide primers designed for MiSEQ library generation were designed to encompass the programmed DSB sites for each sgRNA used (**S6 Fig**). All genome edited samples showed detectable levels of genome editing (0.25 – 0.31%) with substitutions being the most frequently observed mutation (**S2 Table**). Further analyses of modified sequence reads revealed all CRISPR/Cas9 plasmid treatment samples displayed NHEJ-associated mutations predicted to introduce frameshifts at the *smnagal* locus (likely resulting in the translation of substantially truncated proteins) or to ablate *smnagal* transcription (**S7 Fig**). Similar to RNAi, CRISPR/Cas9 genome editing led to significant knockdowns (46 – 60%) of *smnagal* in mixed-sex adult worms when compared to controls (**Fig 6B**). Reassuringly, *smnagal* depletion in siRNA-treated worms significantly reduced α-NAGAL activity in SWAP derived from both males and females when compared to si*Luc* controls (**Fig 6C**). However, RNAi-mediated *smnagal* knockdown did not result in a significant reduction in SWAP-derived α-GAL activity when compared to si*Luc* treated schistosome samples (**S8 Fig**).

Fig 6. RNAi and CRISPR/Cas9 approaches lead to *smnagal* knockdown and siRNA-treated adult worms contain reduced α-NAGAL activity. (A) 7 wk old adult male schistosomes were electroporated with 5 μg siRNA duplexes targeting luciferase (si*Luc*) and *smp_089290* (si*Smp_089290*). After 48 hr, total RNA was isolated and used to generate cDNA, which was subjected to qRT-PCR. Percent knockdown (KD) is indicated. Statistical significance is indicated (Student's *t*-test, two tailed, unequal variance, ** = $p<0.01$). (B) 7 wk old adult male and female schistosomes were electroporated with either lentiviral CRISPR/Cas9 plasmid

715 constructs targeting exon 1 (SmNAGALX1) or exon 2 (SmNAGALX2) of *smnagal*. Additionally,
716 both exons were targeted by electroporating a mixture of both plasmid constructs
717 (SmNAGALX1/X2). Electroporations with CRISPR/Cas9 plasmid DNA containing a Scramble
718 sgRNA were used as a control. After seven days, total RNA was isolated and used to synthesise
719 cDNA for qRT-PCR analysis. Percent knockdown (KD) is indicated. Statistical significance is
720 indicated (Kruskal-Wallis ANOVA with Dunn post-hoc comparisons, * = p<0.05). (C) α-NAGAL
721 activity was measured in SWAP derived from si*Luc* treated and si*Smp_089290* treated adult
722 male and female worms (6.45 µg for males and 2.44 µg for females) as described above.
723 Statistical significance is indicated (Student's *t*-test, two tailed, unequal variance, * = p<0.05
724 and ** = p<0.01).
725
726 Having established that *smnagal* encodes a functional α-NAGAL and given that RNAi
727 and CRISPR/Cas9 depleted this transcript from intracellular RNA pools, motility and egg-laying
728 phenotypes of si*Smnagal* treated schistosome pairs were subsequently examined and
729 quantified. Regardless of the quantification metric used (WormAssayGP2 [65, 66] or WHO-
730 TDR standards [69]), a clear motility defect was observed in both male and female
731 schistosomes when *smnagal* was depleted by RNAi (**Fig 7A**, **S9** and **S10 Figs**). This motility
732 defect was apparent by day two post RNAi and maintained until day 7, when the assay was
733 terminated. Upon completion of the RNAi assays (day seven), eggs were collected from *in*
734 *vitro* cultures and quantified for number, volume and retention of mature vitellocytes. Here,
735 a significant reduction in the quantity of deposited eggs was associated with *smnagal*
736 deficiency (**Fig 7B**). The CRISPR/Cas9 genome editing approach further supported these
737 observations with all *smnagal*-edited worms displaying a significantly reduced number of
738 deposited eggs when compared to controls (**Fig 7C**). Notably, eggs derived from siRNA-treated

739 female worms were also significantly smaller than those collected from wells of si*Luc* treated
740 worms (**Fig 7D**) and contained less-compact, abnormal (less mature) vitellocytes (**Fig 7E**, **S11**
741 and **S12 Figs**).
742
743 **Fig 7. Reductions in *smnagal/smp_089290* affect adult worm motility and egg production**
744 **processes.** (A) Motility of si*Luc* treated and si*Smnagal* treated adult male and female worms
745 was analysed daily for up to seven days after electroporation using WormAssayGP2 as
746 described in the Materials and methods. Statistical significance is indicated (General Linear
747 Mixed-Effects Model, NLME and EMMEANS R packages, * = $p<0.05$ and ** = $p<0.01$). (B) The
748 total number of eggs produced by si*Luc* treated and si*Smnagal* treated adult female worms
749 were collected seven days after electroporation and enumerated. Statistical significance is
750 indicated (Student's *t*-test, two tailed, unequal variance, * = $p<0.05$). (C) The total number of
751 eggs produced by adult worm pairs electroporated with SmNAGALX1 CRISPR/Cas9 plasmid,
752 SmNAGALX2 CRISPR/Cas9 plasmid and SmNAGALX1/X2 CRISPR/Cas9 plasmid were counted
753 seven days after electroporation. Electroporations with lentiviral CRISPR/Cas9 plasmid
754 constructs containing a Scramble sgRNA served as a control. Statistical significance is
755 indicated (Kruskal-Wallis ANOVA with Dunn post-hoc comparisons, * = $p<0.05$). (D) Volumes
756 of eggs produced by si*Luc* treated and si*Smnagal* treated adult female worms were calculated
757 as described in the Materials and methods. Statistical significance is indicated (Student's *t*-
758 test, two tailed, unequal variance, ** = $p<0.01$). (E) Representative images of fluorescence in
759 eggs collected from wells of si*Luc* and si*Smp_089290* treated worm pairs. Blue = $DAPI^+$ cells
760 (405 nm blue diode laser) and white scale bars = 20 μm. Average number of $DAPI^+$ cells per
761 egg calculated by IMARIS 7.3 software for each siRNA treatment is shown below each image.

762 DAPI^{+} cells in eggs derived from wells of si*Smp_089290* treated worm pairs could not be
763 determined and are labelled as 'nd'.
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785

Discussion

The identification and inhibition of gene products responsible for essential developmental or gender-associated processes provide a pathway by which schistosome drug discovery can progress rationally. To fast track such investigations, an extensive collection of putative *S. mansoni* drug candidates is currently available within the TDR Targets database [70]. Alongside this resource, multiple reports have recently described how genome sequencing outputs can be effectively leveraged by both cheminformatics and functional genomics for characterising next-generation schistosome drug targets and chemotherapeutics [66, 71-76]. Complementary evidence is provided here to support an essential role for SmNAGAL in the regulation of worm movement and reproductive processes.

SmNAGAL (Smp_089290) encoded by *S. mansoni*, MS3_11280 encoded by *S. haematobium* and EWB00_005284 encoded by *S. japonicum* were the only melibiase-2 domain-containing *Schistosoma* proteins conserving all functionally important amino acid residues necessary for the hydrolysis of α-galactose and α-N-acetylgalactosamine residues from glycolipid and glycoproteins (**Fig 1**). Closer inspection of these residues showed these three proteins possessed the exact same ligand binding residues to *H. sapiens* α-NAGAL, *G. gallus* α-NAGAL and *C. elegans* gana-1, which are all α-NAGAL enzymes with hydrolytic activity against terminal α-N-acetylgalactosamine and α-galactose moieties [24, 31, 32]. Accordingly, it is likely that Smp_089290, MS3_11280 and EWB00_005284 would exhibit enzymatic activity towards both terminal α-N-acetylgalactosamine and α-galactose residues unlike α-GAL enzymes, which can only cleave terminal α-galactose residues. Studies characterising human α-GAL/α-NAGAL activity provide evidence that *H. sapiens* α-GAL cannot use α-N-acetylgalactosamine as a substrate due to steric hindrance mediated by Glu^{203} and Leu^{206} [30,

31]. In contrast, *H. sapiens* α-NAGAL, Smp_089290, MS3_11280 and EWB00_005284 all possess Ser and Ala at homologous positions ($Ser^{188, 168, 168, 197}$ and $Ala^{191, 171, 171, 200}$), which are required/essential for using α-N-acetylgalactosamine as a substrate [31]. Additionally, EWB00_005283 and EWB00_005285 encoded by *S. japonicum* contain both catalytic aspartic acid residues (**Fig 1**), which suggests that *S. japonicum* possesses three melibiase-2 domain-containing proteins capable of enzymatic activity. However, neither of these *S. japonicum* homologs possess all 13 ligand binding residues, which may influence affinity to target substrates. Although mutagenesis studies performed on *Pichia pastoris* α-GAL and α-NAGAL showed Trp^{16} to be essential for enzymatic activity [77], little is known on the essentiality of the other ligand binding residues. Nevertheless, phylogenetic analyses reinforced that Smp_089290, MS3_11280, EWB00_005284, EWB00_005283 and EWB00_005285 are schistosome α-NAGALs as all demonstrated stronger relations to representative α-NAGAL proteins compared to representative α-GAL proteins (**Fig 2**). Whether the other GH27 clan D family members identified here contribute to the α-GAL activities measured within adult worm extracts (**Fig 4D**) has yet to be determined.

Interrogating DNA microarray (**Fig 4B**) and RNA-Seq meta-analysis (**S2 Fig**) databases provided the first insight to *smnagal*'s temporal expression profile across the developmental stages of the schistosome. In both cases, expression of *smnagal* increased throughout intra-mammalian schistosome development until full adult worm maturation, suggesting a potential role for SmNAGAL in adult schistosome development within the definitive human host. Confirmation of *smnagal* expression in adult male and female worm stages by qRT-PCR analyses (**Fig 4C**) and dominant α-NAGAL activity as shown by enzymatic assays (**Fig 4D**) suggests SmNAGAL preferentially cleaves off α-N-acetylgalactosamine (and not α-galactose) residues from glycan substrates (functionally confirmed by RNAi, **Fig 6** and **S8 Fig**). Homology

833 modelling suggests SmNAGAL possesses a pattern of α-helices and β-strands throughout its
834 N-terminal domain (**Fig 3**) comparable to the $(\beta/\alpha)_8$ barrel structure commonly observed in
835 GH27 family members [30-32, 78-80]. Furthermore, all enzymatically important residues were
836 shown to be arranged in an exposed catalytic active pocket, which suggests SmNAGAL utilises
837 a double displacement mechanism for substrate binding and cleavage (**Fig 3**). This type of
838 cleavage mechanism is commonly utilised by retaining GH enzymes (yielding a product that
839 possesses the same anomeric configuration as the cleaved substrate) and involves two
840 nucleophilic attacks on the 1-carbon of the substrate [32, 81-84].

841 Amongst a variety of glycans and glycoconjugates in adult schistosomes is the O-
842 glycopeptide Galβ1-3GalNAcα1-Ser/Thr (also known as the oncofetal Thomsen-Friedenreich
843 antigen or TF antigen), which is an abundant α-N-acetylgalactosamine-containing structure
844 [85]. The TF antigen is present on the surface syncytium and may be involved in protecting
845 tegumental structures that are essential for schistosome survival within the vasculature of
846 the human host [85-87]. Furthermore, the TF antigen has been suggested to interfere with
847 the functions of host Kupffer cells and hepatocytes [85, 88]. Therefore, SmNAGAL activity
848 required for hydrolysis of the GalNAcα1-Ser/Thr linkage during O-glycopeptide degradation
849 may have an impact on adult worm tegument metabolism and host interactions. Expression
850 of *smnagal* in adult female and male tegument lineage cells identified by scRNA-Seq analyses
851 (**S4** and **S5 Figs**) further supports a potential role for SmNAGAL in tegument metabolism and
852 host interactions. Parenchymal expression of *smnagal* (**Fig 5**) and noticeable abundance in
853 other cell types might be explained by the reported localisation of α-NAGAL enzymatic activity
854 in lysosomes [89]. Lysosomes are found throughout many different *S. mansoni* tissues [90],
855 which suggests other functional roles for SmNAGAL in addition to those identified in this
856 study. Clearly, these observations require further exploration.

Knockdown of *smnagal* correlated with striking motility defects (**Fig 7A** and **S9 Fig**) as early as day two post siRNA treatment. Smaller reductions in motility observed in si*SmNAGAL* treated female worms may be due to the higher expression of *smnagal* in female worms compared to male worms. Higher mRNA turnover rates and transcript abundances have been shown to be a limiting factor of siRNA efficiency when targeting DGKE and ARHGAP27 kinases in HeLa and HepG2 cells [91]. Regardless of this potential discrepancy of RNAi efficiency between the sexes, the motility defect became more severe at day three (**Fig 7A**, **S9** and **S10 Figs**) and was maintained for the seven day experiment. The observed abnormal motility defects were consistent with the neurological and neuromuscular impairments associated with the human lysosomal storage disorder known as Schindler/Kanzaki disease (human α-NAGAL deficiency) [22, 23, 31]. Symptoms of Schindler/Kanzaki disease include a wide range of clinical neurological/neuromuscular deficits due to the accumulation of substrates possessing α-N-acetylgalactosamine residues, which are grouped into three distinct types [31]. The most severe form, type I, is characterised by stiff movements (spasticity) caused from involuntary muscle spasms, developmental retrogression, decorticate posturing, profound psychomotor retardation and muscular hypotonia, which begins in infancy [23, 92-94]. Worms substantially depleted of *smnagal*/SmNAGAL activity display motility defects consistent with the spasticity associated with type I Schindler/Kanzaki disease. Due to the neurological nature of Schindler/Kanzaki disease, the onset of impaired motility phenotypes may be due to *smnagal* depletion within neuronal cell clusters in both adult female and male worms (scRNA-Seq expression profiles; **S4** and **S5 Figs**). Similarly, female muscle cells also possess moderate *smnagal* levels, which suggests the onset of RNAi-mediated motility phenotypes could be driven by depletion in muscle cell activity and, therefore, collectively characterised as a neuromuscular impairment. However, minimal *smnagal* expression

observed in male muscle cells suggests the abnormal motility phenotype is exclusively associated with neurological impairments in this sex. Regardless of the differential molecular mechanisms involved, it is clear that SmNAGAL also contributes to coordinated movement in adult schistosomes.

A fundamental key difference between the male and female schistosome centres on production of eggs, which leads to the pathology and transmission of schistosomiasis [95, 96]. An important organ necessary for egg production is the vitellarium, which extends throughout the majority of the female worm and is involved in the production of mature vitellocytes [97, 98]. Vitellocyte maturation within the vitellarium progresses through four stages involving cell division and differentiation to ultimately produce mature vitellocytes (also referred to as stage 4 vitellocytes) [99]. Mature vitellocytes are transported through the vitello-oviduct and surround the ovum, which is initially produced in the ovary [100-102]. The rigid insoluble eggshell is subsequently synthesised by phenol oxidase-mediated protein cross-linking (quinone tanning) as a result of increased tyrosinase activity originating from late/mature vitellocytes [103-107]. Subsequently, the egg enters the uterus and is expelled through the gonopore of the female into the blood [107]. The elevated expression of *smnagal* within the vitellaria and mature vitellocytes traversing the vitello-oviduct (**Fig 5A** and **S4 Fig**) suggest SmNAGAL may be involved in aspects of egg production. The detection of *smnagal* expression within the egg (**Fig 4B** and **S2 Fig**) may also be explained by the presence of mature vitellocytes and strongly supports a role for SmNAGAL in vitellogenesis and oviposition, which was subsequently confirmed by both RNAi and programmed gene knockout (**Figs 7B** and **7C**). Here, diminished numbers of eggs produced by *smnagal* depleted adults exhibited two predominant abnormalities; lack of typical vitellocyte structuring and spacing (**Fig 7E**, **S11** and **S12 Figs**) and reductions in egg volume (**Fig 7D**). *smnagal* deficiency, however, does not

influence vitellocyte incorporation as DAPI$^+$ cells are observed in all eggs obtained from si*SmNAGAL* treatment groups. Instead, *smnagal* depletion seems to affect vitellocyte development. Immature (stage 1) vitellocytes of *S. mansoni* are undifferentiated cells with large, irregularly roundish nuclei with diffuse chromatin scattered through the nucleoplasm [108]. In contrast, mature (stage 4) vitellocytes appear more condensed and uniform in shape and size. This observation is comparable to the diffuse DAPI$^+$ material observed in abnormal *S. mansoni* eggs derived from *smnagal* depleted schistosomes. Therefore, it is likely that vitellocytes produced by adult female worms depleted of *smnagal* do not progress beyond stage 1 and the diffuse DAPI$^+$ material observed in eggs represents an accumulation of immature vitellocytes. Furthermore, the atypical spacing between neighbouring immature vitellocytes likely contributes to the reductions in egg volume observed in eggs from si*SmNAGAL* treatment groups (approximately 15,000 µm^3 per egg) when compared to eggs from si*Luc* treatment groups (approximately 35,000 µm^3 per egg) (**Fig 7D**). Uneven and patchy auto-fluorescence was also consistently observed in eggs from si*SmNAGAL* treatment groups (**S12 Fig**), which may be explained by fewer mature vitellocytes containing functionally active tyrosinase being packaged in the *in vitro* laid eggs (IVLEs) [103-105]. In some cases, *smnagal* deficiency also contributed to additional observed phenotypes such as abnormal shaped eggs and incomplete development of the lateral spine (**S12 Fig**). Similar egg phenotypes have also been noted following inhibition of tyrosinase in schistosomes [109]. Therefore, in addition to adult worm motility, SmNAGAL clearly participates in oviposition. In light of this evidence, *S. japonicum* may need three enzymatically active α-NAGAL proteins (EWB00_005284, EWB00_005283 and EWB00_005285, **Fig 1**) due to markedly higher rate of oviposition (>2000 eggs per day per worm pair) compared to *S. mansoni* and *S. haematobium* (>300 and >200 eggs per day per worm pair, respectively) [110, 111]. Further characterisation of these other

929 schistosome homologs could provide further insight into the differential rate and absolute
930 numbers of eggs produced by the three main, human-infecting schistosome species.

931 The use of RNAi to characterise *S. mansoni* genes of interest [66], such as *smnagal*,
932 has been established for several years and is currently considered the functional genomics
933 gold standard for this parasite. However, CRISPR/Cas9 genome editing approaches to
934 characterise schistosome gene function are increasingly being explored as recently
935 exemplified by studies of *S. mansoni omega-1* [112], *ache* [113], and *sult-or* [114]. Our use of
936 this technology to edit *smnagal* contains some broadly-overlapping similarities to these
937 previous studies. For example, the overall percentages of modified sequence reads observed
938 in *smnagal*-edited worms (0.25 – 0.31%) were comparable to those observed in *omega-1*-
939 edited eggs (approximately 4.5%) [112], *ache*-edited eggs (0.0295 – 0.12%) [113], *sult-or*-
940 edited worms (0.3 – 2.0%) and *sult-or*-edited sporocysts (0.1 – 0.2%) [114]. Furthermore, the
941 predominant types of mutation in *smnagal*-edited worms were substitutions (**S2 Table**),
942 which is consistent with *omega-1*-edited eggs [112] and *ache*-edited eggs [113]. The
943 CRISPR/Cas9 investigation targeting the liver fluke granulin (*ov-grn-1*) locus in *Opisthorchis*
944 *viverrini* also showed substitutions (98.7%) to be introduced at a higher rate than insertions
945 (0.6%) or deletions (0.7%) [115]. However, this observation was not consistent for the *sult-or*
946 investigation, which only showed deletions attributable to genome editing [114]. Additional
947 investigations are necessary in *S. mansoni* and other platyhelminths to confirm if
948 predominant substitution rates resulting from programmed genome editing are a conserved
949 feature. It is notable that worms treated with CRISPR/Cas9 plasmids targeting *smnagal* exon
950 1 (SmNAGALX1 and dual SmNAGALX1/X2) exhibited modified sequence reads with complex
951 rearrangements consisting of insertions with deletions (**S2 Table**), which was not previously
952 reported in the other *S. mansoni* genome editing studies to date. This may be due to the

infrequency of this type of indel being introduced when DSBs are repaired after CRISPR/Cas9 genome editing [116]. Nonetheless, low percentages of *S*. *mansoni* genome editing quantified within *omega-1*-edited eggs and *ache*-edited eggs led to pronounced phenotypes [112, 113]. This observation was also consistent for *smnagal*-edited worms, which showed significant reductions in IVLE production (**Fig 7C**) similar to si*SmNAGAL* treated worms (**Fig 7B**). Further comparisons can be made between *smnagal*-edited worms and *omega-1*-edited eggs in which both manipulated groups showed significant reductions in target transcript abundance. However, this is in contrast to the CRISPR/Cas9 study targeting *sult-or*, which showed no mRNA knockdown or expected phenotypes despite NHEJ-associated deletions predicted to cause frameshifts that ablate *sult-or* transcription [114]. Collectively, these results suggest that phenotypic effects/reductions in mRNA abundance associated with RNAi-treated *S*. *mansoni* parasites targeting a particular transcript may not always be equivalent to those found in CRISPR/Cas9-edited *S*. *mansoni* parasites that target the associated gene locus. Furthermore, the overall percentages of genome editing may be underestimated in *smnagal*-edited worms due to the presence of large deletions completely removing the primer regions over exon 1 and 2 and, thus, some mutations remain undetected by amplicon sequencing and bioinformatics analysis of the alleles. This situation has been reported with *S*. *stercoralis* [117] and *C*. *elegans* [118] and speculated to occur in the reports on *omega-1* [112], *sult-or* [114] and *ache* [113].

In addition to the presence of large deletions, the *omega-1* investigation suggested that several non-synonymous substitutions may have disrupted the ribonuclease catalytic site and contributed to the mutant phenotypes. Similarly, the *smnagal* genome edited-associated phenotypes observed (*smnagal* knockdown and reduced egg production) may also be due to NHEJ-associated indels/substitutions producing frameshifts (leading to the translation of

977 substantially truncated proteins), ablating *smnagal* transcription or leading to translated
978 proteins that lack all the functional amino acid residues (**S7 Fig**) [32, 119]. The continued
979 refinement of CRISPR/Cas9 technology in *S. mansoni* will help resolve some of these
980 outstanding queries.

981 While the specific SmNAGAL targets within adult schistosomes have yet to be
982 identified, our functional characterisation of this glycogene product suggests that
983 glycoproteins/glycolipids containing α-N-acetylgalactosamine residues are critical for
984 coordinated worm movement and egg production. These traits suggest that SmNAGAL is an
985 essential schistosome gene product, representing a novel parasite vulnerability for exploiting
986 further as a next-generation anthelminthic target for controlling schistosomiasis. As an
987 important component of these investigations, we additionally confirm that *smnagal* is
988 susceptible to somatic genome editing and contribute to the growing literature on utilising
989 the CRISPR/Cas9 system in *S. mansoni* as a tool for functional genomics in parasitic
990 platyhelminths [120]. An enhanced understanding of SmNAGAL or other *S. mansoni* glycan
991 machinery components in lifecycle functions or host interactions will aid the search for
992 urgently-needed, next-generation interventions.

993

994

995

996

997

998

999

1000

1001

Acknowledgments

We acknowledge all members of the Hoffmann laboratory and Ms Julie Hirst for assisting in the maintenance of the schistosome lifecycle at Aberystwyth University. We thank Dr Toby Wilkinson (The University of Edinburgh, UK) for assisting in the statistical analysis of the WormAssayGP2 and adult worm scoring matrix data. At George Washington University Medical School, schistosome-infected mice and snails were provided by the NIAID Schistosomiasis Resource Center of the Biomedical Research Institute, Rockville, Maryland through NIH-NIAID Contract HHSN272201000005I for distribution through BEI Resources. Benjamin J. Hulme was supported from a kind donation provided by David & Eleanor James.

References

1. McManus DP, Dunne DW, Sacko M, Utzinger J, Vennervald BJ, Zhou XN. Schistosomiasis. Nat Rev Dis Primers. 2018;4(1):13. doi: 10.1038/s41572-018-0013-8. PubMed PMID: 30093684.
2. Thetiot-Laurent SA, Boissier J, Robert A, Meunier B. Schistosomiasis chemotherapy. Angew Chem Int Ed Engl. 2013;52(31):7936-56. doi: 10.1002/anie.201208390. PubMed PMID: 23813602.
3. Karanja DM, Boyer AE, Strand M, Colley DG, Nahlen BL, Ouma JH, et al. Studies on schistosomiasis in western Kenya: II. Efficacy of praziquantel for treatment of schistosomiasis in persons coinfected with human immunodeficiency virus-1. Am J Trop Med Hyg. 1998;59(2):307-11. doi: 10.4269/ajtmh.1998.59.307. PubMed PMID: 9715952.
4. Park SK, Gunaratne GS, Chulkov EG, Moehring F, McCusker P, Dosa PI, et al. The anthelmintic drug praziquantel activates a schistosome transient receptor potential channel. J Biol Chem. 2019;294(49):18873-80. doi: 10.1074/jbc.AC119.011093. PubMed PMID: 31653697; PubMed Central PMCID: PMCPMC6901322.
5. Park SK, Friedrich L, Yahya NA, Rohr C, Chulkov EG, Maillard D, et al. Mechanism of praziquantel action at a parasitic flatworm ion channel. bioRxiv. 2021:2021.03.09.434291. doi: 10.1101/2021.03.09.434291.
6. Couto FF, Coelho PM, Araujo N, Kusel JR, Katz N, Jannotti-Passos LK, et al. Schistosoma mansoni: a method for inducing resistance to praziquantel using infected Biomphalaria glabrata snails. Mem Inst Oswaldo Cruz. 2011;106(2):153-7. doi: 10.1590/s0074-02762011000200006. PubMed PMID: 21537673.
7. Fallon PG, Doenhoff MJ. Drug-resistant schistosomiasis: resistance to praziquantel and oxamniquine induced in Schistosoma mansoni in mice is drug specific. Am J Trop Med Hyg. 1994;51(1):83-8. doi: 10.4269/ajtmh.1994.51.83. PubMed PMID: 8059919.
8. Ismail MM, Taha SA, Farghaly AM, el-Azony AS. Laboratory induced resistance to praziquantel in experimental schistosomiasis. J Egypt Soc Parasitol. 1994;24(3):685-95. PubMed PMID: 7844435.
9. Hokke CH, Deelder AM, Hoffmann KF, Wuhrer M. Glycomics-driven discoveries in schistosome research. Exp Parasitol. 2007;117(3):275-83. doi: 10.1016/j.exppara.2007.06.003. PubMed PMID: 17659278.
10. Hokke CH, Fitzpatrick JM, Hoffmann KF. Integrating transcriptome, proteome and glycome analyses of Schistosoma biology. Trends Parasitol. 2007;23(4):165-74. doi: 10.1016/j.pt.2007.02.007. PubMed PMID: 17336161.
11. Smit CH, van Diepen A, Nguyen DL, Wuhrer M, Hoffmann KF, Deelder AM, et al. Glycomic Analysis of Life Stages of the Human Parasite Schistosoma mansoni Reveals Developmental Expression Profiles of Functional and Antigenic Glycan Motifs. Mol Cell Proteomics. 2015;14(7):1750-69. doi: 10.1074/mcp.M115.048280. PubMed PMID: 25883177; PubMed Central PMCID: PMCPMC4587318.
12. Wuhrer M, Koeleman CA, Fitzpatrick JM, Hoffmann KF, Deelder AM, Hokke CH. Gender-specific expression of complex-type N-glycans in schistosomes. Glycobiology. 2006;16(10):991-1006. doi: 10.1093/glycob/cwl020. PubMed PMID: 16825488.
13. Everts B, Perona-Wright G, Smits HH, Hokke CH, van der Ham AJ, Fitzsimmons CM, et al. Omega-1, a glycoprotein secreted by Schistosoma mansoni eggs, drives Th2 responses. J Exp Med. 2009;206(8):1673-80. doi: 10.1084/jem.20082460. PubMed PMID: 19635864; PubMed Central PMCID: PMCPMC2722183.
14. Meevissen MH, Driessen NN, Smits HH, Versteegh R, van Vliet SJ, van Kooyk Y, et al. Specific glycan elements determine differential binding of individual egg glycoproteins of the human parasite Schistosoma mansoni by host C-type lectin receptors. Int J Parasitol. 2012;42(3):269-77. doi: 10.1016/j.ijpara.2012.01.004. PubMed PMID: 22673410.

15. Peterson NA, Hokke CH, Deelder AM, Yoshino TP. Glycotope analysis in miracidia and primary sporocysts of Schistosoma mansoni: differential expression during the miracidium-to-sporocyst transformation. Int J Parasitol. 2009;39(12):1331-44. doi: 10.1016/j.ijpara.2009.06.002. PubMed PMID: 19545571; PubMed Central PMCID: PMCPMC3740939.
16. Mickum ML, Prasanphanich NS, Heimburg-Molinaro J, Leon KE, Cummings RD. Deciphering the glycogenome of schistosomes. Front Genet. 2014;5:262. doi: 10.3389/fgene.2014.00262. PubMed PMID: 25147556; PubMed Central PMCID: PMCPMC4122909.
17. van Noort K, Nguyen DL, Kriechbaumer V, Hawes C, Hokke CH, Schots A, et al. Functional characterization of Schistosoma mansoni fucosyltransferases in Nicotiana benthamiana plants. Sci Rep. 2020;10(1):18528. doi: 10.1038/s41598-020-74485-z. PubMed PMID: 33116178; PubMed Central PMCID: PMCPMC7595089.
18. Lombard V, Ramulu HG, Drula E, Coutinho PM, Henrissat B. Carbohydrate-Active EnZymes database (CAZy), Glycoside Hydrolase family classification 2021 [May 6, 2021]. Available from: http://www.cazy.org/Glycoside-Hydrolases.html.
19. Henrissat B, Davies G. Structural and sequence-based classification of glycoside hydrolases. Curr Opin Struct Biol. 1997;7(5):637-44. doi: 10.1016/s0959-440x(97)80072-3. PubMed PMID: 9345621.
20. Garman SC. Structural studies on α-GAL and α-NAGAL: The atomic basis of Fabry and Schindler diseases. Biocatalysis and Biotransformation. 2006;24(1-2):129-36. doi: 10.1080/10242420600598194.
21. Chan B, Adam DN. A Review of Fabry Disease. Skin Therapy Lett. 2018;23(2):4-6. PubMed PMID: 29562089.
22. Schindler D, Bishop DF, Wolfe DE, Wang AM, Egge H, Lemieux RU, et al. Neuroaxonal dystrophy due to lysosomal alpha-N-acetylgalactosaminidase deficiency. N Engl J Med. 1989;320(26):1735-40. doi: 10.1056/NEJM198906293202606. PubMed PMID: 2733734.
23. van Diggelen OP, Schindler D, Kleijer WJ, Huijmans JM, Galjaard H, Linden HU, et al. Lysosomal alpha-N-acetylgalactosaminidase deficiency: a new inherited metabolic disease. Lancet. 1987;2(8562):804. doi: 10.1016/s0140-6736(87)92542-6. PubMed PMID: 2889023.
24. Hujova J, Sikora J, Dobrovolny R, Poupetova H, Ledvinova J, Kostrouchova M, et al. Characterization of gana-1, a Caenorhabditis elegans gene encoding a single ortholog of vertebrate alpha-galactosidase and alpha-N-acetylgalactosaminidase. BMC Cell Biol. 2005;6(1):5. doi: 10.1186/1471-2121-6-5. PubMed PMID: 15676072; PubMed Central PMCID: PMCPMC548690.
25. Smithers SR, Terry RJ. The infection of laboratory hosts with cercariae of Schistosoma mansoni and the recovery of the adult worms. Parasitology. 1965;55(4):695-700. doi: 10.1017/s0031182000086248. PubMed PMID: 4957633.
26. Howe KL, Bolt BJ, Shafie M, Kersey P, Berriman M. WormBase ParaSite - a comprehensive resource for helminth genomics. Mol Biochem Parasitol. 2017;215:2-10. doi: 10.1016/j.molbiopara.2016.11.005. PubMed PMID: 27899279; PubMed Central PMCID: PMCPMC5486357.
27. Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, et al. The Pfam protein families database. Nucleic Acids Res. 2012;40(Database issue):D290-301. doi: 10.1093/nar/gkr1065. PubMed PMID: 22127870; PubMed Central PMCID: PMCPMC3245129.
28. UniProt Consortium. The Universal Protein Resource (UniProt) in 2010. Nucleic Acids Res. 2010;38(Database issue):D142-8. doi: 10.1093/nar/gkp846. PubMed PMID: 19843607; PubMed Central PMCID: PMCPMC2808944.
29. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic acids research. 2004;32(5):1792-7. Epub 2004/03/23. doi: 10.1093/nar/gkh340. PubMed PMID: 15034147; PubMed Central PMCID: PMC390337.
30. Garman SC, Garboczi DN. The molecular defect leading to Fabry disease: structure of human alpha-galactosidase. J Mol Biol. 2004;337(2):319-35. doi: 10.1016/j.jmb.2004.01.035. PubMed PMID: 15003450.

31. Clark NE, Garman SC. The 1.9 a structure of human alpha-N-acetylgalactosaminidase: The molecular basis of Schindler and Kanzaki diseases. J Mol Biol. 2009;393(2):435-47. doi: 10.1016/j.jmb.2009.08.021. PubMed PMID: 19683538; PubMed Central PMCID: PMCPMC2771859.
32. Garman SC, Hannick L, Zhu A, Garboczi DN. The 1.9 A structure of alpha-N-acetylgalactosaminidase: molecular basis of glycosidase deficiency diseases. Structure. 2002;10(3):425-34. doi: 10.1016/s0969-2126(02)00726-8. PubMed PMID: 12005440.
33. Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol. 2000;17(4):540-52. doi: 10.1093/oxfordjournals.molbev.a026334. PubMed PMID: 10742046.
34. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Hohna S, et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol. 2012;61(3):539-42. doi: 10.1093/sysbio/sys029. PubMed PMID: 22357727; PubMed Central PMCID: PMCPMC3329765.
35. Whelan S, Goldman N. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol Biol Evol. 2001;18(5):691-9. doi: 10.1093/oxfordjournals.molbev.a003851. PubMed PMID: 11319253.
36. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Mol Biol Evol. 2018;35(6):1547-9. doi: 10.1093/molbev/msy096. PubMed PMID: 29722887; PubMed Central PMCID: PMCPMC5967553.
37. Jones DT, Taylor WR, Thornton JM. The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci. 1992;8(3):275-82. doi: 10.1093/bioinformatics/8.3.275. PubMed PMID: 1633570.
38. Rambaut A. FigTree v1.4.3. 2016. Available from: http://tree.bio.ed.ac.uk/software/figtree/.
39. Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen MY, et al. Comparative protein structure modeling using Modeller. Curr Protoc Bioinformatics. 2006;Chapter 5:Unit-5 6. doi: 10.1002/0471250953.bi0506s15. PubMed PMID: 18428767; PubMed Central PMCID: PMCPMC4186674.
40. Baker D, Sali A. Protein structure prediction and structural genomics. Science. 2001;294(5540):93-6. doi: 10.1126/science.1065659. PubMed PMID: 11588250.
41. Lovell SC, Davis IW, Arendall WB, 3rd, de Bakker PI, Word JM, Prisant MG, et al. Structure validation by Calpha geometry: phi,psi and Cbeta deviation. Proteins. 2003;50(3):437-50. doi: 10.1002/prot.10286. PubMed PMID: 12557186.
42. Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res. 2007;35(Web Server issue):W407-10. doi: 10.1093/nar/gkm290. PubMed PMID: 17517781; PubMed Central PMCID: PMCPMC1933241.
43. Bowie JU, Luthy R, Eisenberg D. A method to identify protein sequences that fold into a known three-dimensional structure. Science. 1991;253(5016):164-70. doi: 10.1126/science.1853201. PubMed PMID: 1853201.
44. Fitzpatrick JM, Peak E, Perally S, Chalmers IW, Barrett J, Yoshino TP, et al. Anti-schistosomal intervention targets identified by lifecycle transcriptomic analyses. PLoS Negl Trop Dis. 2009;3(11):e543. doi: 10.1371/journal.pntd.0000543. PubMed PMID: 19885392; PubMed Central PMCID: PMCPMC2764848.
45. Geospiza. FinchTV 1.4.0. 2018. Available from: https://digitalworldbiology.com/FinchTV.
46. Lu Z, Zhang Y, Berriman M. A web portal for gene expression across all life stages of Schistosoma mansoni. bioRxiv. 2018:308213. doi: 10.1101/308213.
47. Anderson L, Amaral MS, Beckedorff F, Silva LF, Dazzani B, Oliveira KC, et al. Schistosoma mansoni Egg, Adult Male and Female Comparative Gene Expression Analysis and Identification of Novel Genes by RNA-Seq. PLoS Negl Trop Dis. 2015;9(12):e0004334. doi: 10.1371/journal.pntd.0004334. PubMed PMID: 26719891; PubMed Central PMCID: PMCPMC4699917.

48. Wang B, Collins JJ, 3rd, Newmark PA. Functional genomic characterization of neoblast-like stem cells in larval Schistosoma mansoni. Elife. 2013;2:e00768. doi: 10.7554/eLife.00768. PubMed PMID: 23908765; PubMed Central PMCID: PMCPMC3728622.

49. Protasio AV, Tsai IJ, Babbage A, Nichol S, Hunt M, Aslett MA, et al. A systematically improved high quality genome and transcriptome of the human blood fluke Schistosoma mansoni. PLoS Negl Trop Dis. 2012;6(1):e1455. doi: 10.1371/journal.pntd.0001455. PubMed PMID: 22253936; PubMed Central PMCID: PMCPMC3254664.

50. Protasio AV, van Dongen S, Collins J, Quintais L, Ribeiro DM, Sessler F, et al. MiR-277/4989 regulate transcriptional landscape during juvenile to adult transition in the parasitic helminth Schistosoma mansoni. PLoS Negl Trop Dis. 2017;11(5):e0005559. doi: 10.1371/journal.pntd.0005559. PubMed PMID: 28542189; PubMed Central PMCID: PMCPMC5459504.

51. Lu Z, Sessler F, Holroyd N, Hahnel S, Quack T, Berriman M, et al. Schistosome sex matters: a deep view into gonad-specific and pairing-dependent transcriptomes reveals a complex gender interplay. Sci Rep. 2016;6:31150. doi: 10.1038/srep31150. PubMed PMID: 27499125; PubMed Central PMCID: PMCPMC4976352.

52. Lu Z, Zhang Y, Berriman M. Schisto_xyz search engine for RNA-Seq meta data analysis. 2018. Available from: http://schisto.xyz/.

53. Geyer KK, Rodriguez Lopez CM, Chalmers IW, Munshi SE, Truscott M, Heald J, et al. Cytosine methylation regulates oviposition in the pathogenic blood fluke Schistosoma mansoni. Nat Commun. 2011;2:424. doi: 10.1038/ncomms1433. PubMed PMID: 21829186; PubMed Central PMCID: PMCPMC3265374.

54. Oliveros JC, Franch M, Tabas-Madrid D, San-Leon D, Montoliu L, Cubas P, et al. Breaking-Cas-interactive design of guide RNAs for CRISPR-Cas experiments for ENSEMBL genomes. Nucleic Acids Res. 2016;44(W1):W267-71. doi: 10.1093/nar/gkw407. PubMed PMID: 27166368; PubMed Central PMCID: PMCPMC4987939.

55. Heler R, Samai P, Modell JW, Weiner C, Goldberg GW, Bikard D, et al. Cas9 specifies functional viral targets during CRISPR-Cas adaptation. Nature. 2015;519(7542):199-202. doi: 10.1038/nature14245. PubMed PMID: 25707807; PubMed Central PMCID: PMCPMC4385744.

56. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. Nat Protoc. 2013;8(11):2281-308. doi: 10.1038/nprot.2013.143. PubMed PMID: 24157548; PubMed Central PMCID: PMCPMC3969860.

57. Clement K, Rees H, Canver MC, Gehrke JM, Farouni R, Hsu JY, et al. CRISPResso2 provides accurate and rapid genome editing sequence analysis. Nat Biotechnol. 2019;37(3):224-6. doi: 10.1038/s41587-019-0032-3. PubMed PMID: 30809026; PubMed Central PMCID: PMCPMC6533916.

58. Pinello L, Canver MC, Hoban MD, Orkin SH, Kohn DB, Bauer DE, et al. Analyzing CRISPR genome-editing experiments with CRISPResso. Nat Biotechnol. 2016;34(7):695-7. doi: 10.1038/nbt.3583. PubMed PMID: 27404874; PubMed Central PMCID: PMCPMC5242601.

59. Chalmers IW, McArdle AJ, Coulson RM, Wagner MA, Schmid R, Hirai H, et al. Developmentally regulated expression, alternative splicing and distinct sub-groupings in members of the Schistosoma mansoni venom allergen-like (SmVAL) gene family. BMC Genomics. 2008;9:89. doi: 10.1186/1471-2164-9-89. PubMed PMID: 18294395; PubMed Central PMCID: PMCPMC2270263.

60. Collins JJ, 3rd, Hou X, Romanova EV, Lambrus BG, Miller CM, Saberi A, et al. Genome-wide analyses reveal a role for peptide hormones in planarian germline development. PLoS Biol. 2010;8(10):e1000509. doi: 10.1371/journal.pbio.1000509. PubMed PMID: 20967238; PubMed Central PMCID: PMCPMC2953531.

61. Cogswell AA, Collins JJ, 3rd, Newmark PA, Williams DL. Whole mount in situ hybridization methodology for Schistosoma mansoni. Mol Biochem Parasitol. 2011;178(1-2):46-50. doi: 10.1016/j.molbiopara.2011.03.001. PubMed PMID: 21397637; PubMed Central PMCID: PMCPMC3102561.

62. Wendt G, Zhao L, Chen R, Liu C, O'Donoghue AJ, Caffrey CR, et al. A single-cell RNA-seq atlas of Schistosoma mansoni identifies a key regulator of blood feeding. Science. 2020;369(6511):1644-9. doi: 10.1126/science.abb7709. PubMed PMID: 32973030; PubMed Central PMCID: PMCPMC7875187.
63. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, 3rd, et al. Comprehensive Integration of Single-Cell Data. Cell. 2019;177(7):1888-902 e21. doi: 10.1016/j.cell.2019.05.031. PubMed PMID: 31178118; PubMed Central PMCID: PMCPMC6687398.
64. Blom D, Speijer D, Linthorst GE, Donker-Koopman WG, Strijland A, Aerts JMFG. Recombinant enzyme therapy for Fabry disease: Absence of editing of human alpha-galactosidase A mRNA. American Journal of Human Genetics. 2003;72(1):23-31. doi: Doi 10.1086/345309. PubMed PMID: WOS:000180186800003.
65. Marcellino C, Gut J, Lim KC, Singh R, McKerrow J, Sakanari J. WormAssay: a novel computer application for whole-plate motion-based screening of macroscopic parasites. PLoS Negl Trop Dis. 2012;6(1):e1494. doi: 10.1371/journal.pntd.0001494. PubMed PMID: 22303493; PubMed Central PMCID: PMCPMC3269415.
66. Wang J, Paz C, Padalino G, Coghlan A, Lu Z, Gradinaru I, et al. Large-scale RNAi screening uncovers therapeutic targets in the parasite Schistosoma mansoni. Science. 2020;369(6511):1649-53. doi: 10.1126/science.abb7699. PubMed PMID: 32973031; PubMed Central PMCID: PMCPMC7877197.
67. Ramirez B, Bickle Q, Yousif F, Fakorede F, Mouries MA, Nwaka S. Schistosomes: challenges in compound screening. Expert Opin Drug Discov. 2007;2(s1):S53-61. doi: 10.1517/17460441.2.S1.S53. PubMed PMID: 23489033.
68. Jaromin-Glen K, Klapec T, Lagod G, Karamon J, Malicki J, Skowronska A, et al. Division of methods for counting helminths' eggs and the problem of efficiency of these methods. Ann Agric Environ Med. 2017;24(1):1-7. doi: 10.5604/12321966.1233891. PubMed PMID: 28378963.
69. Ramirez B, Bickle Q, Yousif F, Fakorede F, Mouries MA, Nwaka S. Schistosomes: challenges in compound screening. Expert Opin Drug Discov. 2007;2(s1):S53-61. Epub 2007/10/01. doi: 10.1517/17460441.2.S1.S53. PubMed PMID: 23489033.
70. Aguero F, Al-Lazikani B, Aslett M, Berriman M, Buckner FS, Campbell RK, et al. Genomic-scale prioritization of drug targets: the TDR Targets database. Nat Rev Drug Discov. 2008;7(11):900-7. doi: 10.1038/nrd2684. PubMed PMID: 18927591; PubMed Central PMCID: PMCPMC3184002.
71. Calixto NM, Dos Santos DB, Bezerra JCB, Silva LA. In silico repositioning of approved drugs against Schistosoma mansoni energy metabolism targets. PLoS One. 2018;13(12):e0203340. doi: 10.1371/journal.pone.0203340. PubMed PMID: 30596650; PubMed Central PMCID: PMCPMC6312253.
72. Giuliani S, Silva AC, Borba J, Ramos PIP, Paveley RA, Muratov EN, et al. Computationally-guided drug repurposing enables the discovery of kinase targets and inhibitors as new schistosomicidal agents. PLoS Comput Biol. 2018;14(10):e1006515. doi: 10.1371/journal.pcbi.1006515. PubMed PMID: 30346968; PubMed Central PMCID: PMCPMC6211772.
73. Gouveia MJ, Brindley PJ, Gartner F, Costa J, Vale N. Drug Repurposing for Schistosomiasis: Combinations of Drugs or Biomolecules. Pharmaceuticals (Basel). 2018;11(1). doi: 10.3390/ph11010015. PubMed PMID: 29401734; PubMed Central PMCID: PMCPMC5874711.
74. Padalino G, Ferla S, Brancale A, Chalmers IW, Hoffmann KF. Combining bioinformatics, cheminformatics, functional genomics and whole organism approaches for identifying epigenetic drug targets in Schistosoma mansoni. Int J Parasitol Drugs Drug Resist. 2018;8(3):559-70. doi: 10.1016/j.ijpddr.2018.10.005. PubMed PMID: 30455056; PubMed Central PMCID: PMCPMC6288008.
75. Pasche V, Laleu B, Keiser J. Screening a repurposing library, the Medicines for Malaria Venture Stasis Box, against Schistosoma mansoni. Parasit Vectors. 2018;11(1):298. doi: 10.1186/s13071-018-2855-z. PubMed PMID: 29764454; PubMed Central PMCID: PMCPMC5952519.

1274 76. Pasche V, Laleu B, Keiser J. Early Antischistosomal Leads Identified from in Vitro and in Vivo
1275 Screening of the Medicines for Malaria Venture Pathogen Box. ACS Infect Dis. 2019;5(1):102-10. doi:
1276 10.1021/acsinfecdis.8b00220. PubMed PMID: 30398059.
1277 77. Zhu A, Monahan C, Wang ZK. Trp-16 is essential for the activity of alpha-galactosidase and
1278 alpha-N-acetylgalactosaminidase. Biochim Biophys Acta. 1996;1297(1):99-104. doi: 10.1016/0167-
1279 4838(96)00108-2. PubMed PMID: 8841386.
1280 78. Fujimoto Z, Kaneko S, Momma M, Kobayashi H, Mizuno H. Crystal structure of rice alpha-
1281 galactosidase complexed with D-galactose. J Biol Chem. 2003;278(22):20313-8. doi:
1282 10.1074/jbc.M302292200. PubMed PMID: 12657636.
1283 79. Golubev AM, Nagem RA, Brandao Neto JR, Neustroev KN, Eneyskaya EV, Kulminskaya AA, et
1284 al. Crystal structure of alpha-galactosidase from Trichoderma reesei and its complex with galactose:
1285 implications for catalytic mechanism. J Mol Biol. 2004;339(2):413-22. doi:
1286 10.1016/j.jmb.2004.03.062. PubMed PMID: 15136043.
1287 80. Kytidou K, Beekwilder J, Artola M, van Meel E, Wilbers RHP, Moolenaar GF, et al. Nicotiana
1288 benthamiana alpha-galactosidase A1.1 can functionally complement human alpha-galactosidase A
1289 deficiency associated with Fabry disease. J Biol Chem. 2018;293(26):10042-58. doi:
1290 10.1074/jbc.RA118.001774. PubMed PMID: 29674318; PubMed Central PMCID: PMCPMC6028973.
1291 81. Davies G, Henrissat B. Structures and mechanisms of glycosyl hydrolases. Structure.
1292 1995;3(9):853-9. doi: 10.1016/S0969-2126(01)00220-9. PubMed PMID: 8535779.
1293 82. Kim WD, Kobayashi O, Kaneko S, Sakakibara Y, Park GG, Kusakabe I, et al. alpha-
1294 Galactosidase from cultured rice (Oryza sativa L. var. Nipponbare) cells. Phytochemistry.
1295 2002;61(6):621-30. doi: 10.1016/s0031-9422(02)00368-0. PubMed PMID: 12423882.
1296 83. McCarter JD, Withers SG. Mechanisms of enzymatic glycoside hydrolysis. Curr Opin Struct
1297 Biol. 1994;4(6):885-92. doi: 10.1016/0959-440x(94)90271-2. PubMed PMID: 7712292.
1298 84. Olsson ML, Hill CA, de la Vega H, Liu QP, Stroud MR, Valdinocci J, et al. Universal red blood
1299 cells--enzymatic conversion of blood group A and B antigens. Transfus Clin Biol. 2004;11(1):33-9. doi:
1300 10.1016/j.tracli.2003.12.002. PubMed PMID: 14980547.
1301 85. Thors C, Jansson B, Helin H, Linder E. Thomsen-Friedenreich oncofetal antigen in
1302 Schistosoma mansoni : localization and immunogenicity in experimental mouse infection.
1303 Parasitology. 2006;132(Pt 1):73-81. doi: 10.1017/S003118200500867X. PubMed PMID: 16393356.
1304 86. Hockley DJ, McLaren DJ, Ward BJ, Nermut MV. A freeze-fracture study of the tegumental
1305 membrane of Schistosoma mansoni (Platyhelminthes:Trematoda). Tissue Cell. 1975;7(3):485-96. doi:
1306 10.1016/0040-8166(75)90020-8. PubMed PMID: 1179410.
1307 87. Wendt GR, Collins JN, Pei J, Pearson MS, Bennett HM, Loukas A, et al. Flatworm-specific
1308 transcriptional regulators promote the specification of tegumental progenitors in Schistosoma
1309 mansoni. Elife. 2018;7. doi: 10.7554/eLife.33221. PubMed PMID: 29557781; PubMed Central PMCID:
1310 PMCPMC5927768.
1311 88. Fadden AJ, Holt OJ, Drickamer K. Molecular characterization of the rat Kupffer cell
1312 glycoprotein receptor. Glycobiology. 2003;13(7):529-37. doi: 10.1093/glycob/cwg068. PubMed
1313 PMID: 12672702.
1314 89. Andreotti G, Guarracino MR, Cammisa M, Correra A, Cubellis MV. Prediction of the
1315 responsiveness to pharmacological chaperones: lysosomal human alpha-galactosidase, a case of
1316 study. Orphanet J Rare Dis. 2010;5:36. doi: 10.1186/1750-1172-5-36. PubMed PMID: 21138548;
1317 PubMed Central PMCID: PMCPMC3016270.
1318 90. Nawaratna SS, McManus DP, Moertel L, Gobert GN, Jones MK. Gene Atlasing of digestive
1319 and reproductive tissues in Schistosoma mansoni. PLoS Negl Trop Dis. 2011;5(4):e1043. doi:
1320 10.1371/journal.pntd.0001043. PubMed PMID: 21541360; PubMed Central PMCID:
1321 PMCPMC3082511.
1322 91. Larsson E, Sander C, Marks D. mRNA turnover rate limits siRNA and microRNA efficacy. Mol
1323 Syst Biol. 2010;6:433. doi: 10.1038/msb.2010.89. PubMed PMID: 21081925; PubMed Central PMCID:
1324 PMCPMC3010119.

92. Bakker HD, de Sonnaville ML, Vreken P, Abeling NG, Groener JE, Keulemans JL, et al. Human alpha-N-acetylgalactosaminidase (alpha-NAGA) deficiency: no association with neuroaxonal dystrophy? Eur J Hum Genet. 2001;9(2):91-6. doi: 10.1038/sj.ejhg.5200598. PubMed PMID: 11313741.
93. Sarbu M, Robu A, Peter-Katalinic J, Zamfir AD. Automated chip-nanoelectrospray mass spectrometry for glycourinomics in Schindler disease type I. Carbohydr Res. 2014;398:90-100. doi: 10.1016/j.carres.2014.08.014. PubMed PMID: 25243357.
94. van Diggelen OP, Schindler D, Willemsen R, Boer M, Kleijer WJ, Huijmans JG, et al. alpha-N-acetylgalactosaminidase deficiency, a new lysosomal storage disorder. J Inherit Metab Dis. 1988;11(4):349-57. doi: 10.1007/BF01800424. PubMed PMID: 3149698.
95. Atkinson KH, Atkinson BG. Biochemical basis for the continuous copulation of female Schistosoma mansoni. Nature. 1980;283(5746):478-9. doi: 10.1038/283478a0. PubMed PMID: 7352026.
96. Gryseels B, Polman K, Clerinx J, Kestens L. Human schistosomiasis. Lancet. 2006;368(9541):1106-18. doi: 10.1016/S0140-6736(06)69440-3. PubMed PMID: 16997665.
97. Popiel I, Basch PF. Schistosoma mansoni: cholesterol uptake by paired and unpaired worms. Exp Parasitol. 1986;61(3):343-7. doi: 10.1016/0014-4894(86)90189-x. PubMed PMID: 3709750.
98. Wang J, Collins JJ, 3rd. Identification of new markers for the Schistosoma mansoni vitelline lineage. Int J Parasitol. 2016;46(7):405-10. doi: 10.1016/j.ijpara.2016.03.004. PubMed PMID: 27056273; PubMed Central PMCID: PMCPMC4917872.
99. Kunz W. Schistosome male-female interaction: induction of germ-cell differentiation. Trends Parasitol. 2001;17(5):227-31. doi: 10.1016/s1471-4922(01)01893-1. PubMed PMID: 11323306.
100. Erasmus DA. Schistosoma mansoni: development of the vitelline cell, its role in drug sequestration, and changes induced by Astiban. Exp Parasitol. 1975;38(2):240-56. doi: 10.1016/0014-4894(75)90027-2. PubMed PMID: 1175727.
101. Shaw MK. Schistosoma mansoni: vitelline gland development in females from single sex infections. J Helminthol. 1987;61(3):253-9. doi: 10.1017/s0022149x00010117. PubMed PMID: 3117875.
102. Smyth JD, Clegg JA. Egg-shell formation in trematodes and cestodes. Exp Parasitol. 1959;8(3):286-323. doi: 10.1016/0014-4894(59)90027-x. PubMed PMID: 13663910.
103. Ashton PD, Harrop R, Shah B, Wilson RA. The schistosome egg: development and secretions. Parasitology. 2001;122(Pt 3):329-38. doi: 10.1017/s0031182001007351. PubMed PMID: 11289069.
104. Dewalick S, Bexkens ML, van Balkom BW, Wu YP, Smit CH, Hokke CH, et al. The proteome of the insoluble Schistosoma mansoni eggshell skeleton. Int J Parasitol. 2011;41(5):523-32. doi: 10.1016/j.ijpara.2010.12.005. PubMed PMID: 21236260.
105. Fitzpatrick JM, Johansen MV, Johnston DA, Dunne DW, Hoffmann KF. Gender-associated gene expression in two related strains of Schistosoma japonicum. Mol Biochem Parasitol. 2004;136(2):191-209. doi: 10.1016/j.molbiopara.2004.03.014. PubMed PMID: 15478798.
106. Wang J, Chen R, Collins JJ, 3rd. Systematically improved in vitro culture conditions reveal new insights into the reproductive biology of the human parasite Schistosoma mansoni. PLoS Biol. 2019;17(5):e3000254. doi: 10.1371/journal.pbio.3000254. PubMed PMID: 31067225; PubMed Central PMCID: PMCPMC6505934.
107. Waite JH. Precursors of quinone tanning: dopa-containing proteins. Methods Enzymol. 1995;258:1-20. doi: 10.1016/0076-6879(95)58033-6. PubMed PMID: 8524142.
108. Erasmus DA, Popiel I, Shaw JR. A comparative study of the vitelline cell in Schistosoma mansoni, S. haematobium, S. japonicum and S. mattheei. Parasitology. 1982;84(Pt 2):283-7. doi: 10.1017/s0031182000044838. PubMed PMID: 7200223.
109. Fitzpatrick JM, Hirai Y, Hirai H, Hoffmann KF. Schistosome egg production is dependent upon the activities of two developmentally regulated tyrosinases. FASEB J. 2007;21(3):823-35. doi: 10.1096/fj.06-7314com. PubMed PMID: 17167065.

110. Cheever AW, Macedonia JG, Mosimann JE, Cheever EA. Kinetics of egg production and egg excretion by Schistosoma mansoni and S. japonicum in mice infected with a single pair of worms. Am J Trop Med Hyg. 1994;50(3):281-95. doi: 10.4269/ajtmh.1994.50.281. PubMed PMID: 8147487.
111. Cheever AW, Torky AH, Shirbiney M. The relation of worm burden to passage of Schistosoma haematobium eggs in the urine of infected patients. Am J Trop Med Hyg. 1975;24(2):284-8. doi: 10.4269/ajtmh.1975.24.284. PubMed PMID: 1119670.
112. Ittiprasert W, Mann VH, Karinshak SE, Coghlan A, Rinaldi G, Sankaranarayanan G, et al. Programmed genome editing of the omega-1 ribonuclease of the blood fluke, Schistosoma mansoni. Elife. 2019;8. doi: 10.7554/eLife.41337. PubMed PMID: 30644357; PubMed Central PMCID: PMCPMC6355194.
113. You H, Mayer JU, Johnston RL, Sivakumaran H, Ranasinghe S, Rivera V, et al. CRISPR/Cas9-mediated genome editing of Schistosoma mansoni acetylcholinesterase. FASEB J. 2021;35(1):e21205. doi: 10.1096/fj.202001745RR. PubMed PMID: 33337558.
114. Sankaranarayanan G, Coghlan A, Driguez P, Lotkowska ME, Sanders M, Holroyd N, et al. Large CRISPR-Cas-induced deletions in the oxamniquine resistance locus of the human parasite Schistosoma mansoni. Wellcome Open Res. 2020;5:178. doi: 10.12688/wellcomeopenres.16031.2. PubMed PMID: 32789192; PubMed Central PMCID: PMCPMC7405262.2.
115. Arunsan P, Ittiprasert W, Smout MJ, Cochran CJ, Mann VH, Chaiyadet S, et al. Programmed knockout mutation of liver fluke granulin attenuates virulence of infection-induced hepatobiliary morbidity. Elife. 2019;8. doi: 10.7554/eLife.41463. PubMed PMID: 30644359; PubMed Central PMCID: PMCPMC6355195.
116. Kosicki M, Tomberg K, Bradley A. Repair of double-strand breaks induced by CRISPR-Cas9 leads to large deletions and complex rearrangements. Nat Biotechnol. 2018;36(8):765-71. doi: 10.1038/nbt.4192. PubMed PMID: 30010673; PubMed Central PMCID: PMCPMC6390938.
117. Gang SS, Castelletto ML, Bryant AS, Yang E, Mancuso N, Lopez JB, et al. Targeted mutagenesis in a human-parasitic nematode. PLoS Pathog. 2017;13(10):e1006675. doi: 10.1371/journal.ppat.1006675. PubMed PMID: 29016680; PubMed Central PMCID: PMCPMC5650185.
118. Chiu H, Schwartz HT, Antoshechkin I, Sternberg PW. Transgene-free genome editing in Caenorhabditis elegans using CRISPR-Cas. Genetics. 2013;195(3):1167-71. doi: 10.1534/genetics.113.155879. PubMed PMID: 23979577; PubMed Central PMCID: PMCPMC3813845.
119. Zhu A, Wang ZK, Goldstein J. Identification of tyrosine 108 in coffee bean alpha-galactosidase as an essential residue for the enzyme activity. Biochim Biophys Acta. 1995;1247(2):260-4. doi: 10.1016/0167-4838(94)00228-9. PubMed PMID: 7696317.
120. Hoffmann KF, Brindley PJ, Berriman M. Medicine. Halting harmful helminths. Science. 2014;346(6206):168-9. doi: 10.1126/science.1261139. PubMed PMID: 25301604.

Supporting information

S1 Fig. Assessing the stereochemical quality of the Smp_089290 homology model using RAMPAGE Ramachandran plot analysis, ProSA-web and Verify 3D software. (A) The graphical representation of the RAMPAGE Ramachandran plot analysis for the Smp_089290 homology model. The plot depicts the torsional angles (phi, ϕ, x-axis and psi, Ψ, y-axis) of the amino acid residues in the homology model; this illustrates which combinations of angles for each atom is possible by considering their dimensions and van der Waals radii. Stable and unstable conformations of the model can, therefore, be plotted on the graph. (B) The graphical representation of the ProSA-web analysis for the Smp_089290 homology model. The z-score (-7.44, black dot) indicates the overall model quality, which is displayed in a plot that contains the z-scores of all experimentally determined protein chains in the current PDB database. The input structure is verified when it is found to be within the range of z-scores typically found for deposited proteins of similar size. (C) The graphical representation of the Verify 3D analysis for the Smp_089290 homology model. The analysis determines the compatibility of the atomic model (3D) of the input structure with its own amino acid sequence (1D) by assigning a structural class for each amino acid residue based on its location and environment (as part of an α-helix, β-strand or an interconnecting loop) and polarity. (D) A summary table of the analysis tool used (colour coded: RAMPAGE Ramachandran plot analysis = red, ProSA-web = orange and Verify 3D = yellow), the results obtained from the assessment of the Smp_089290 homology model and the expected values for verified structures.

S2 Fig. Quantification of *smp_089290* abundance across the *S. mansoni* lifecycle by RNA-Seq analysis reinforces female-biased expression. RNA-Seq meta data analysis of

smp_089290 abundance for 11 lifecycle stages during mixed-sex infections. Individual expression values for male and female are only plotted for lifecycle stages where expression was assessed in sex-separated samples by Protasio *et al*. [50] and Lu *et al*. [51] (i.e. from '21 day juveniles' to '>42 day adults'). A single expression value is plotted for lifecycle stages where expression was assessed in mixed-sex samples by Anderson *et al*. [47], Wang *et al*. [48] and Protasio *et al*. [49] (i.e. from 'egg' to '24 hr somules').

S3 Fig. Adult female and male schistosomes showed little to no specific staining when hybridised with the sense *smp_089290* probe. Images of the anterior, mid-section and posterior (10x magnification) of (A) female and (B) male schistosomes as well as anterior images with a higher magnification (40x magnification, area depicted by black dashed box). Structures labelled include egg (E), ovary (O), vitellarium (V), vitello-oviduct (VOD), intestine (I), oral sucker (OS), oesophagus (OES), ventral sucker (VS) and testes (TES). Black scale bars = 200 µm and red scale bars = 50 µm.

S4 Fig. scRNA-Seq expression profile of *smp_089290* in adult female *S. mansoni* shows localisation to vitellocytes, parenchymal cells and neuronal cells. Labelled UMAP projection plot of various cell clusters highlighted by black dashed regions. *smp_089290* scRNA-Seq expression values shown in this plot were generated from sexually mature adult female samples only. Expression values are normalised to a scale of 0 – 100 and colour coded (blue = low, red = high). Mature vitellocytes and late vitellocytes are labelled by red arrows.

S5 Fig. scRNA-Seq expression profile of *smp_089290* in adult male *S. mansoni* shows enrichment in parenchymal and neuronal cells. Labelled UMAP projection plot of various cell

clusters highlighted by black dashed regions. *smp_089290* scRNA-Seq transcript expression values shown in this plot were generated from adult male samples only. Expression values are normalised to a scale of 0 – 100 and colour coded (blue = low, red = high).

S6 Fig. Flanking oligonucleotide primers designed for MiSEQ library generation encompass DSB site of SmNAGALX1 and SmNAGALX2 sgRNAs. Diagrammatic representation of the genomic sequence of *smnagal* from nucleotide (nt) positions 1500 – 4000 (represented by double lines) based on WormBase ParaSite entry [26]. Exons are depicted as red boxes with nt positions indicated above 5´ and 3´ ends. Numbers written inside each exon represent their position in the gene sequence. Positions and nt sequences of SmNAGALX1_sgRNA and SmNAGALX2_sgRNA are shown below exon 1 and 2, respectively (represented by red arrows). DSB sites for SmNAGALX1_sgRNA and SmNAGALX2_sgRNA are indicated by yellow arrows, which are located three nts upstream of the PAM sequence highlighted in light blue (note: the PAM sequence is not part of the sgRNA sequence). Positions and nt sequences of SmNAGALX1_MISEQ (represented by orange lines) and SmNAGALX2_MISEQ (represented by green lines) primers are shown along the genomic sequence. Amplicon lengths of SmNAGALX1_MISEQ and SmNAGALX2_MISEQ products are underlined in orange and green, respectively.

S7 Fig. NHEJ-associated indels and representative substitutions observed in lentiviral CRISPR/Cas9 plasmid treated worms by CRISPResso2 analysis. Multiple sequence alignment (MSA) showing all insertions, deletions, insertions with deletions and representative substitutions (i.e. supported by more than one sequence read) identified in modified sequence reads by CRISPResso2 analysis. Modified sequence reads are aligned with the

original unmodified *smnagal* nucleotide (nt) sequence. Alignments are grouped as follows: (A) Experimental treatment groups targeting exon 1 and (B) Experimental treatment groups targeting exon 2. A colour code is used to show the number of sequence reads each indel/substitution appears in for SmNAGALX1 (red), SmNAGALX2 (dark blue) and dual SmNAGALX1/X2 plasmid treated worms (purple). Nts that are modified by each indel/substitution within the alignment are coloured green. Insertions appear as additional nts not found in the original sequence, deletions appear as lines (-) and substitutions appear as replaced nts located in the same positions as the original sequence. The nt sequences of SmNAGALX1_sgRNA and SmNAGALX2_sgRNA are highlighted in yellow within the original sequence. DSB sites for SmNAGALX1_sgRNA and SmNAGALX2_sgRNA are indicated by yellow arrows, which are located three nts upstream of the PAM sequence highlighted in light blue (note: the PAM sequence is outside and downstream of the sgRNA sequence). An additional colour code is used to distinguish between start (orange nts) and termination (red nts) codons found in the sequences, which have been highlighted in black.

S8 Fig. SWAP samples derived from *smp_089290* depleted adult male and female schistosomes show no significant reductions in α-GAL activity. (A) 6.45 µg of siRNA treated adult male-derived SWAP and (B) 2.44 µg of siRNA treated adult female-derived SWAP were measured for α-GAL using α-GAL colorimetric substrates. Final absorbances were quantified at 410 nm. Using linear trendline equations generated from α-GAL standard curves, α-GAL activity (µg/ml) were calculated for each sample. No statistical significance in α-GAL activity between samples was observed (Student's *t*-test, two tailed, unequal variance).

S9 Fig. *smnagal* deficiency in adult male and female worms leads to motility defects as assessed by WHO-TDR scoring matrix. The motility of individual worms (five adult pairs per well) were scored between 4 – 0 based on the WHO-TDR scoring matrix guidelines; 4 = normal active/paired up, 3 = slowed activity, 2 = minimal activity and occasional movement of head and tail, 1 = absence of motility apart from gut movements and 0 = total absence of motility. The total occurrences of each score was plotted per day for si*Luc* treated and si*SmNAGAL* treated adult female (A) and male (B) worms starting from day two after electroporation up until day seven (day of electroporation considered as day zero). Statistical significance is indicated (General Linear Mixed-Effects Model, NLME and EMMEANS R packages, ** = $p<0.01$). Day one was not included in the statistical analysis due to worms appearing stunned and immobile, likely due to electroporation manipulation. Six wells/biological replicates per siRNA treatment were used for this analysis.

S10 Fig. Movement of si*SmNAGAL* treated worms is substantially impaired when compared to si*Luc* treated worms on day three. Video footage of (A) si*Luc* treated and (B) si*SmNAGAL* treated adult worms was captured using a NexiusZoom stereo microscope (Euromex) and edited with ImageFocus 4 software (Euromex).

S11 Fig. Representative morphologies of eggs derived from siRNA treated female worms. IMARIS 7.3 software (Bitplane) was used to create a video showing the 360° horizontal rotation of a representative egg derived from (A) si*Luc* treated and (B) si*SmNAGAL* treated adult female worms. Blue = $DAPI^+$ cells, green = egg auto-fluorescence and white scale bars = 20 µm.

S12 Fig. Representative fluorescence micrographs of eggs collected from wells of si*SmNAGAL* treated adult female worms reveal a broad spectrum of abnormal morphologies. Images of eggs from si*Luc* treated worm pairs depicting (A) vitellocytes under the blue channel (i.e. excitation wavelength = 405 nm and emission wavelength = 461 nm) and (B) auto-fluorescence under the green channel (i.e. excitation wavelength = 488 nm and emission wavelength = 520 nm). Images of eggs from si*SmNAGAL* treated worm pairs depicting (C) vitellocytes under the blue channel and (D) auto-fluorescence under the green channel. Blue = $DAPI^+$ cells, green = egg auto-fluorescence and white scale bars = 20 µm.

S1 Table. Initial processing of MiSEQ deep-coverage sequence reads for CRISPResso2 analysis.

Primer pair set	Sample	Reads in inputs	Reads after pre-processing	Reads aligned
SmNAGALX1_MiSEQ primers (targeting exon 1)	Scramble	588918	588918	139154
	SmNAGALX1	1469667	1048675	429175
	Dual SmNAGALX1/X2	1081902	535084	258576
SmNAGALX2_MiSEQ primers (targeting exon 2)	Scramble	722655	683478	176041
	SmNAGALX2	1916911	1916911	103559
	Dual SmNAGALX1/X2	1838265	1683634	136129

The number of sequence reads at each of the three initial processing stages before indel characterisations can be made by further CRISPResso2 analysis is presented. These three initial processing stages are “Reads in inputs” (highlighted in red, first stage), “Reads after pre-processing” (highlighted in blue, second stage) and “Reads aligned” (highlighted in yellow, third stage). “Reads in inputs” refers to the total number of sequence reads from raw MiSEQ sequencing data. “Reads after pre-processing” refers to the number of sequence reads after PCR amplification or trimming artefacts are removed. “Reads aligned” refers to the number

of sequence reads that are of high quality (>60% homology to reference amplicon sequence), which are used for indel characterisations. The table also lists the primer pair set and sample (samples amplified by SmNAGALX1_MiSEQ and SmNAGALX2_MiSEQ primers are highlighted in orange and green, respectively) used for each barcoded MiSEQ amplicon library constructed.

S2 Table. Detectable frequencies of insertions, deletions, insertions with deletions and substitutions in *smnagal*-edited worms as quantified by CRISPResso2 analysis.

		Mutation frequencies (%)				
Primer Pair set	**Sample**	**Unmodified**	**Insertions**	**Deletions**	**Insertions with deletions**	**Substitutions**
SmNAGALX1_MiSEQ primers (targeting exon 1)	SmNAGALX1	99.7306	0.0032	0.0075	0.0014	0.2570
	Dual SmNAGALX1/X2	99.7400	0.0034	0.0073	0.0007	0.2506
SmNAGALX2_MiSEQ primers (targeting exon 2)	SmNAGALX2	99.6900	0.0057	0.0164	0	0.2870
	Dual SmNAGALX1/X2	99.7517	0.0081	0	0	0.2402

The mutation frequencies attributable to genome editing (i.e. insertions, deletions, insertions with deletions, and substitutions) in *smnagal*-edited worms as quantified by CRISPResso2 analysis is indicated. The percentage of unmodified sequence reads is included. The primer pair set and sample (samples amplified by SmNAGALX1_MiSEQ and SmNAGALX2_MiSEQ primers are highlighted in orange and green, respectively) used for each barcoded MiSEQ amplicon library constructed are indicated.

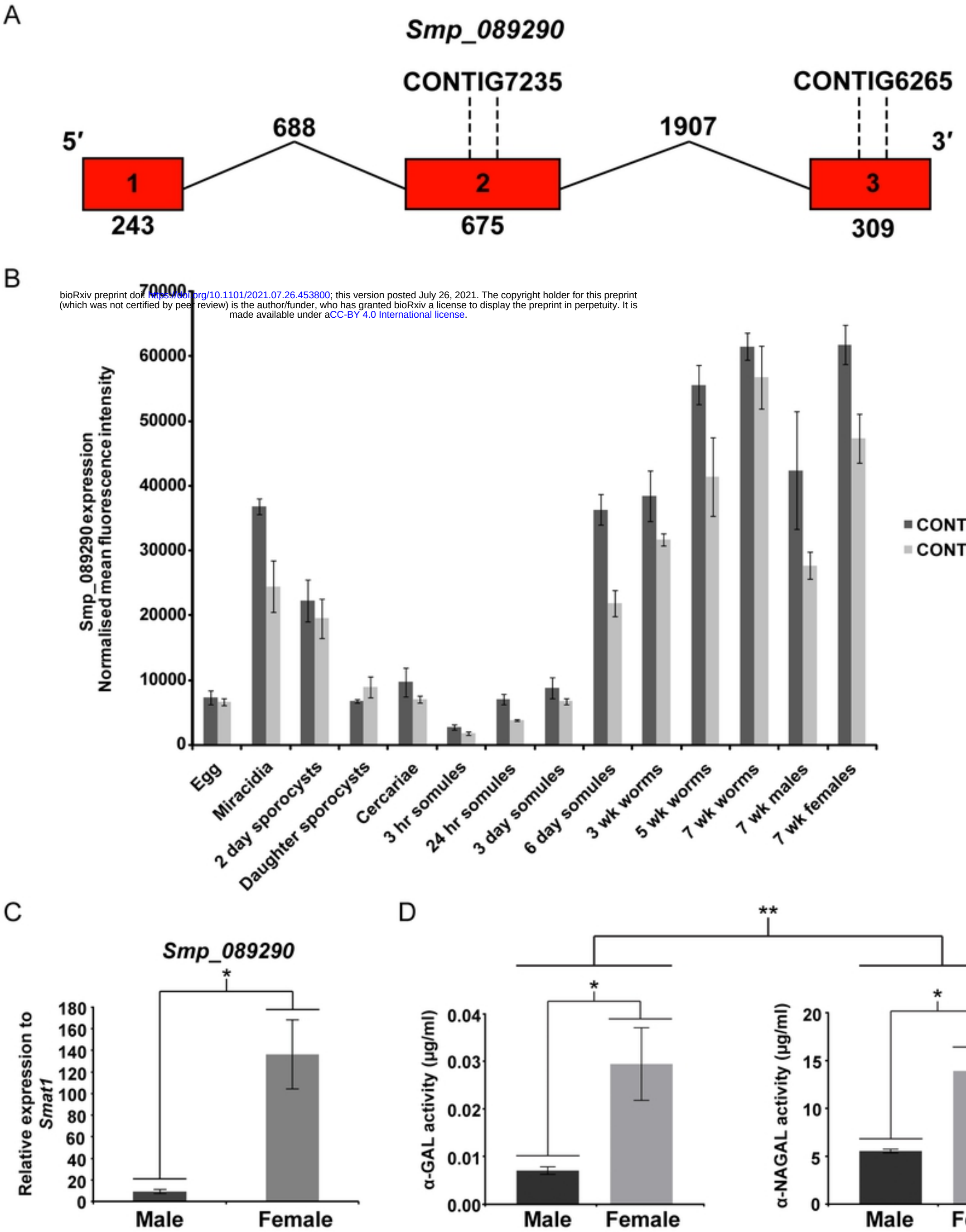


Figure 4

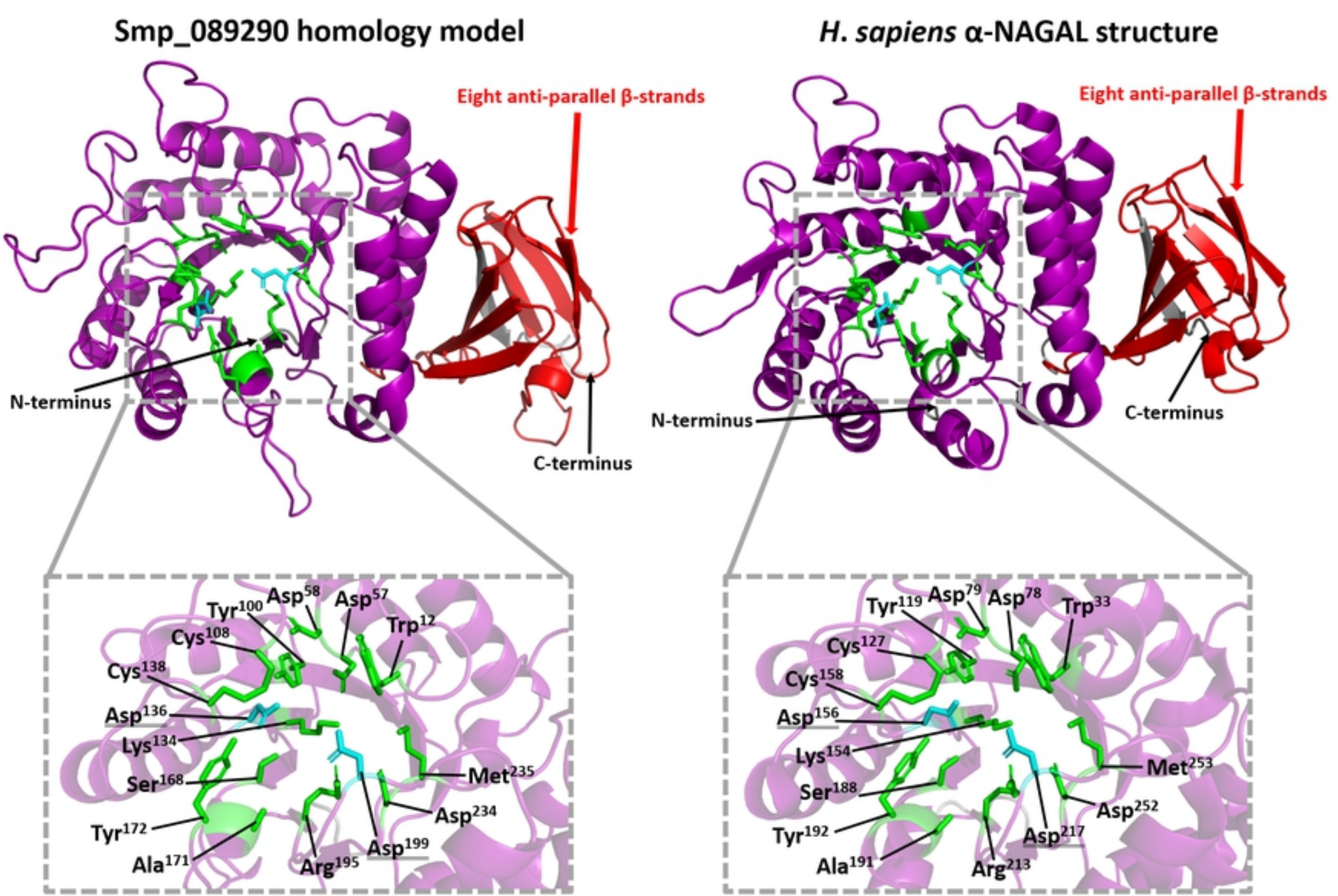

Smp_089290 homology model
Eight anti-parallel β-strands
N-terminus
C-terminus
Tyr100
Asp58
Asp57
Trp12
Cys108
Cys138
Asp136
Lys134
Ser168
Tyr172
Ala171
Arg195
Asp199
Asp234
Met235
H. sapiens α-NAGAL structure
Eight anti-parallel β-strands
N-terminus
C-terminus
Tyr119
Asp79
Asp78
Trp33
Cys127
Cys158
Asp156
Lys154
Ser188
Tyr192
Ala191
Arg213
Asp217
Asp252
Met253

Figure 3

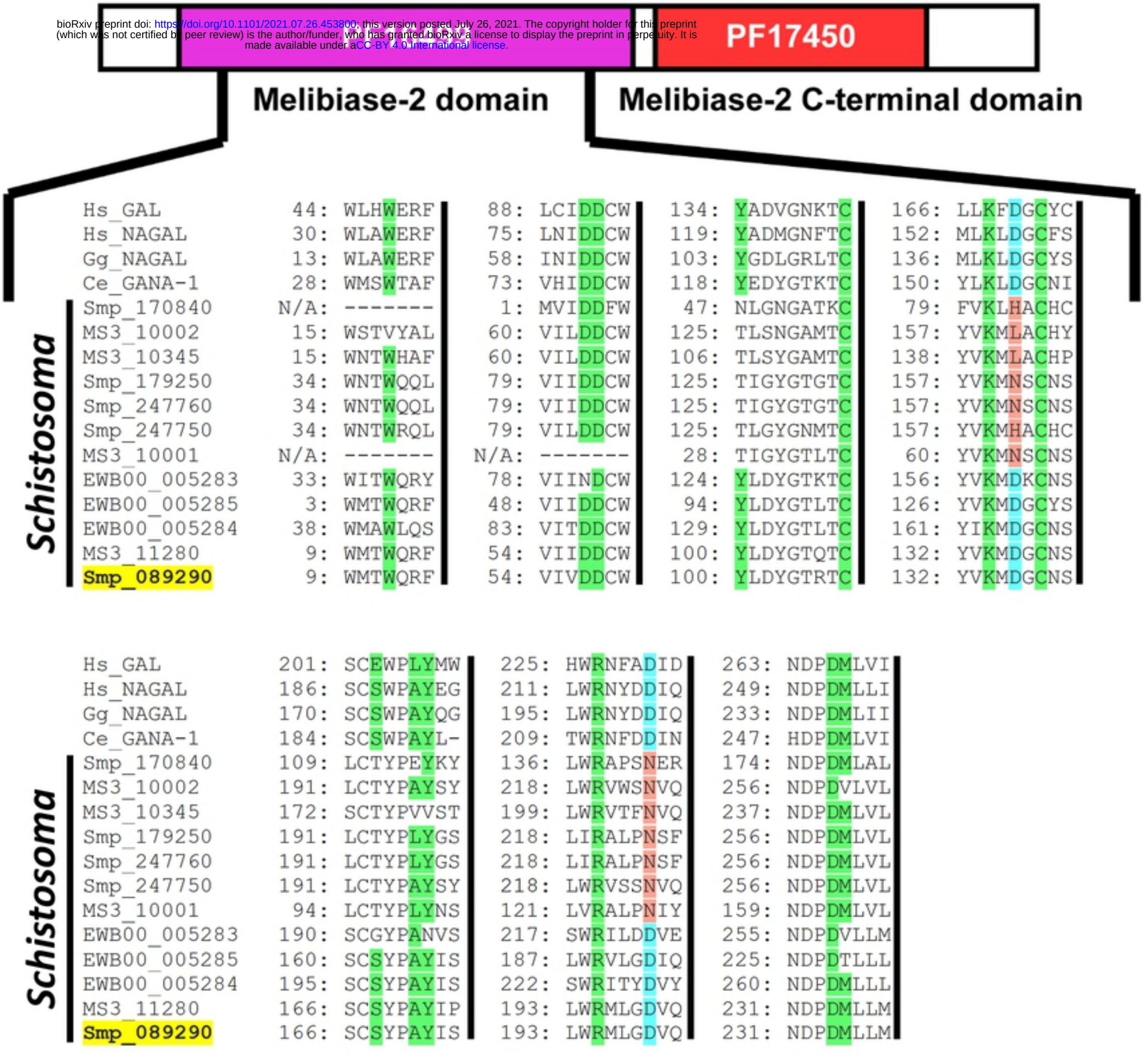
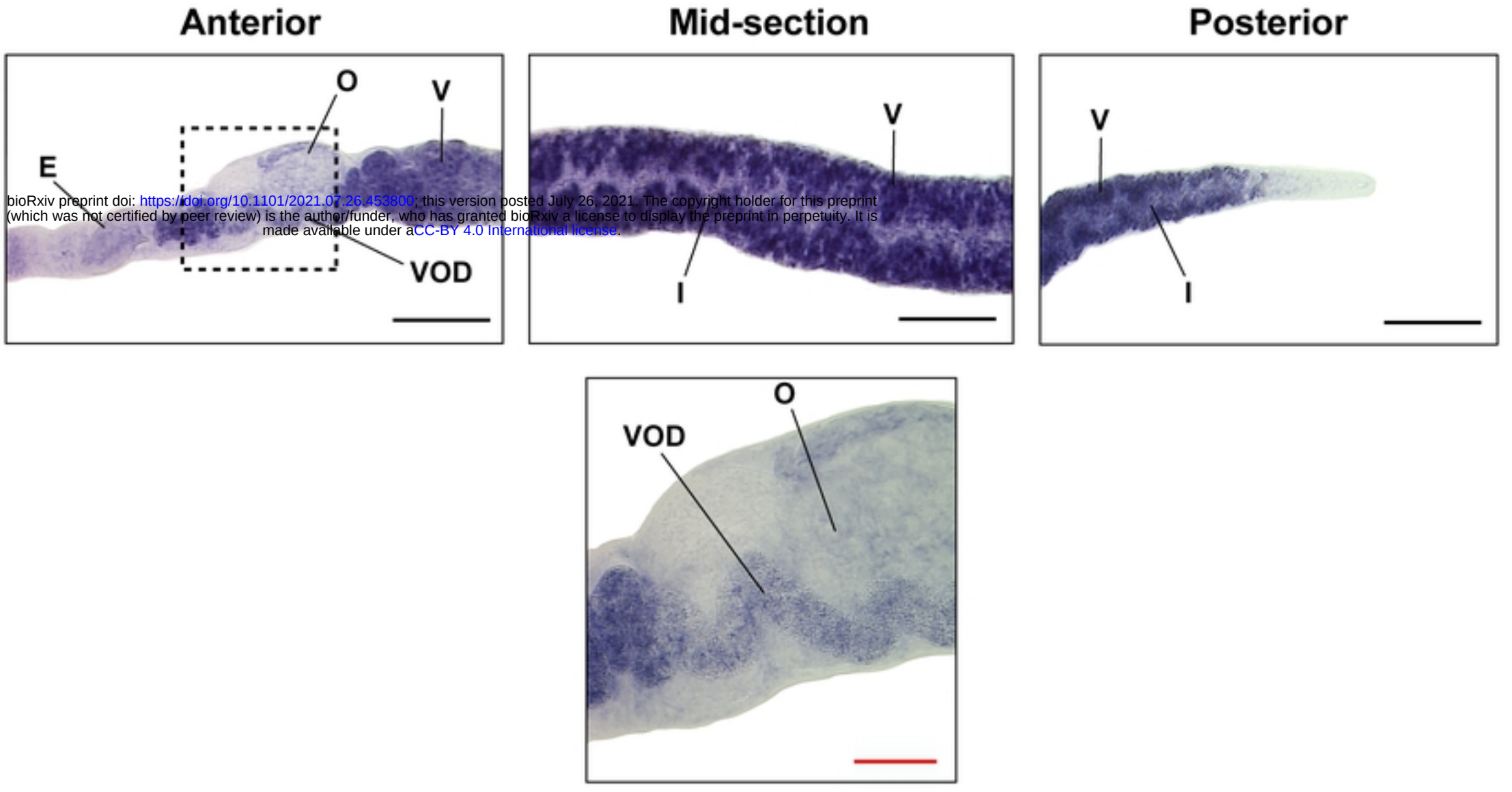


Figure 1

A



B

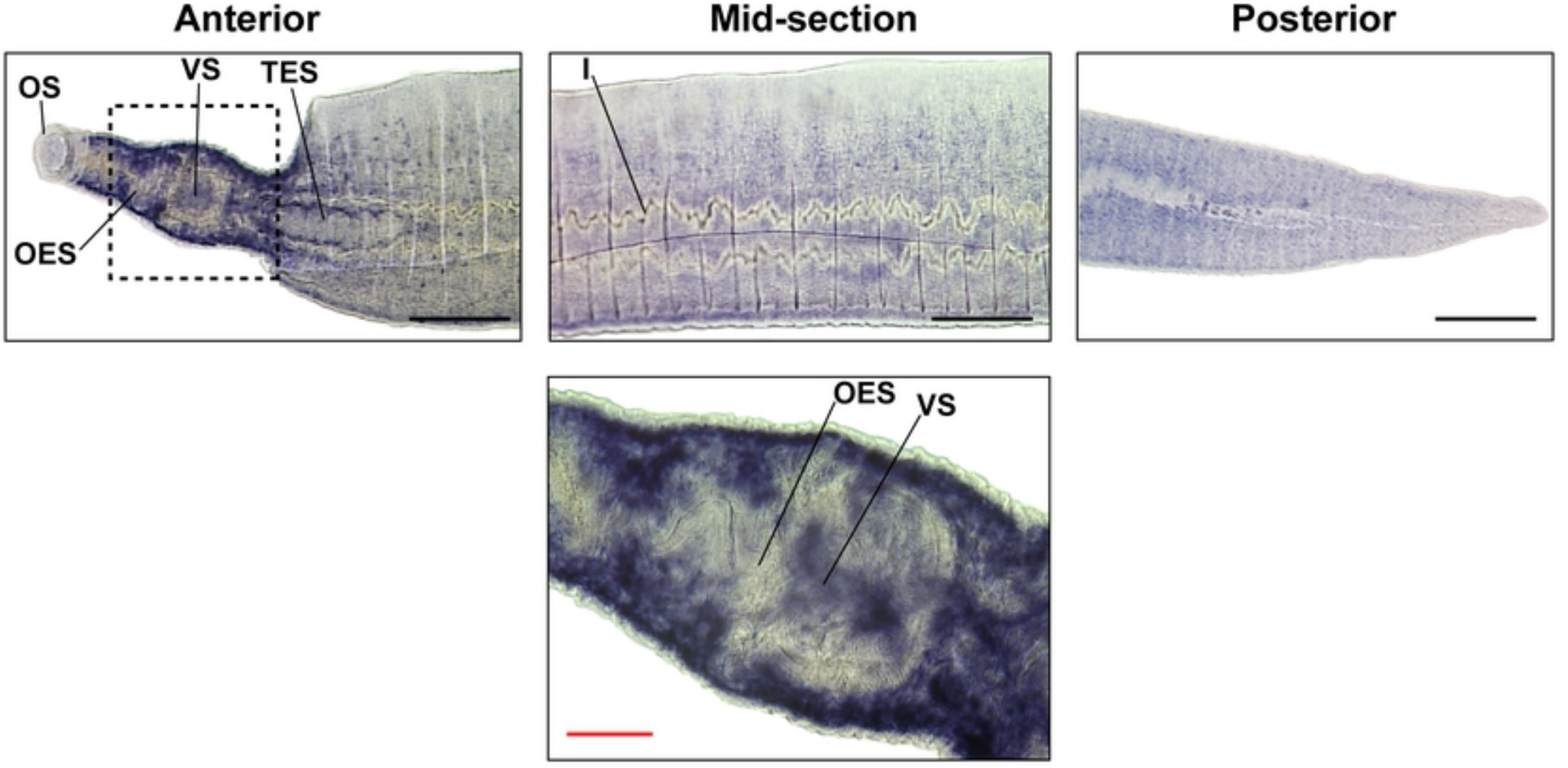


Figure 5

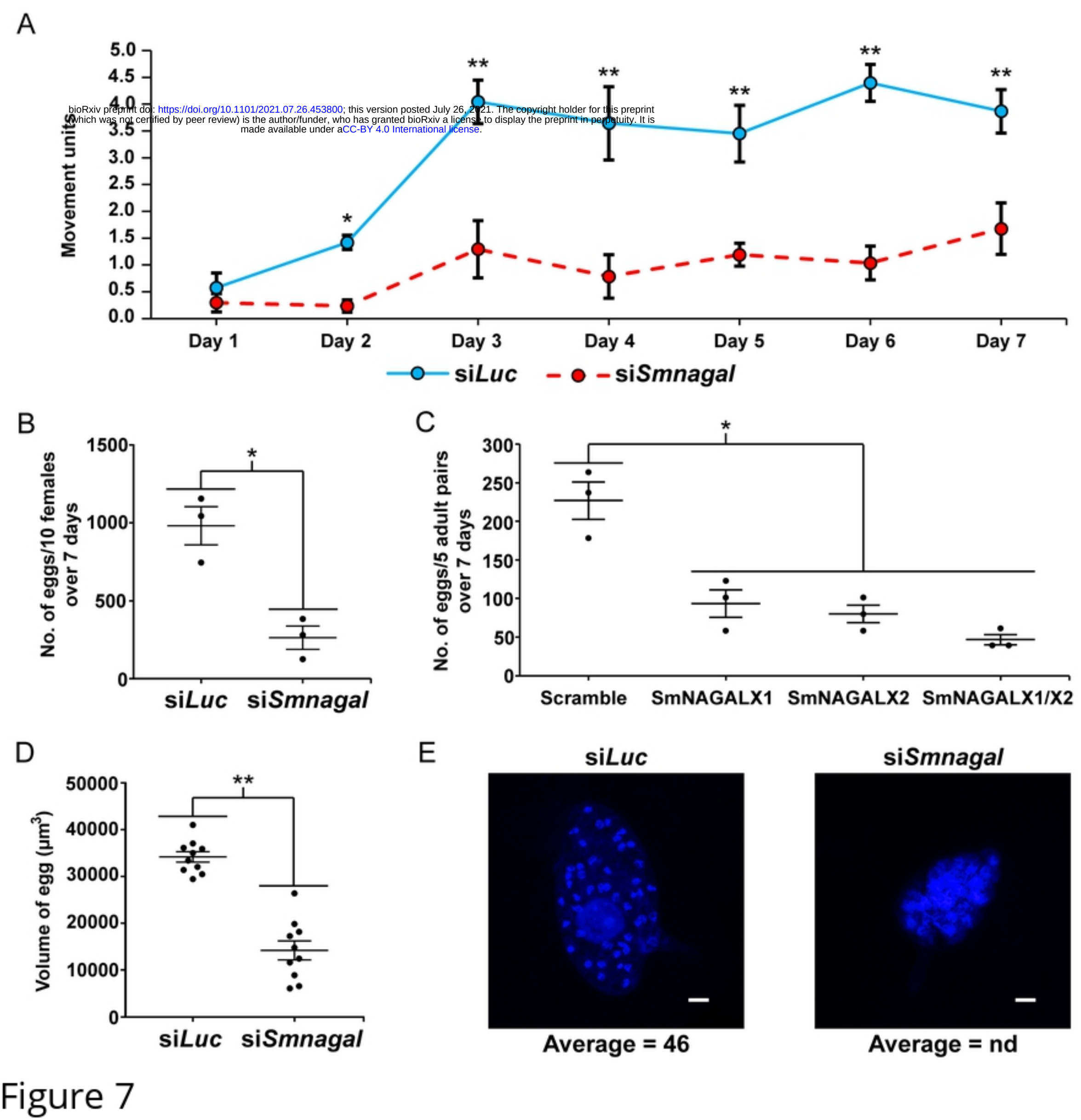


Figure 7

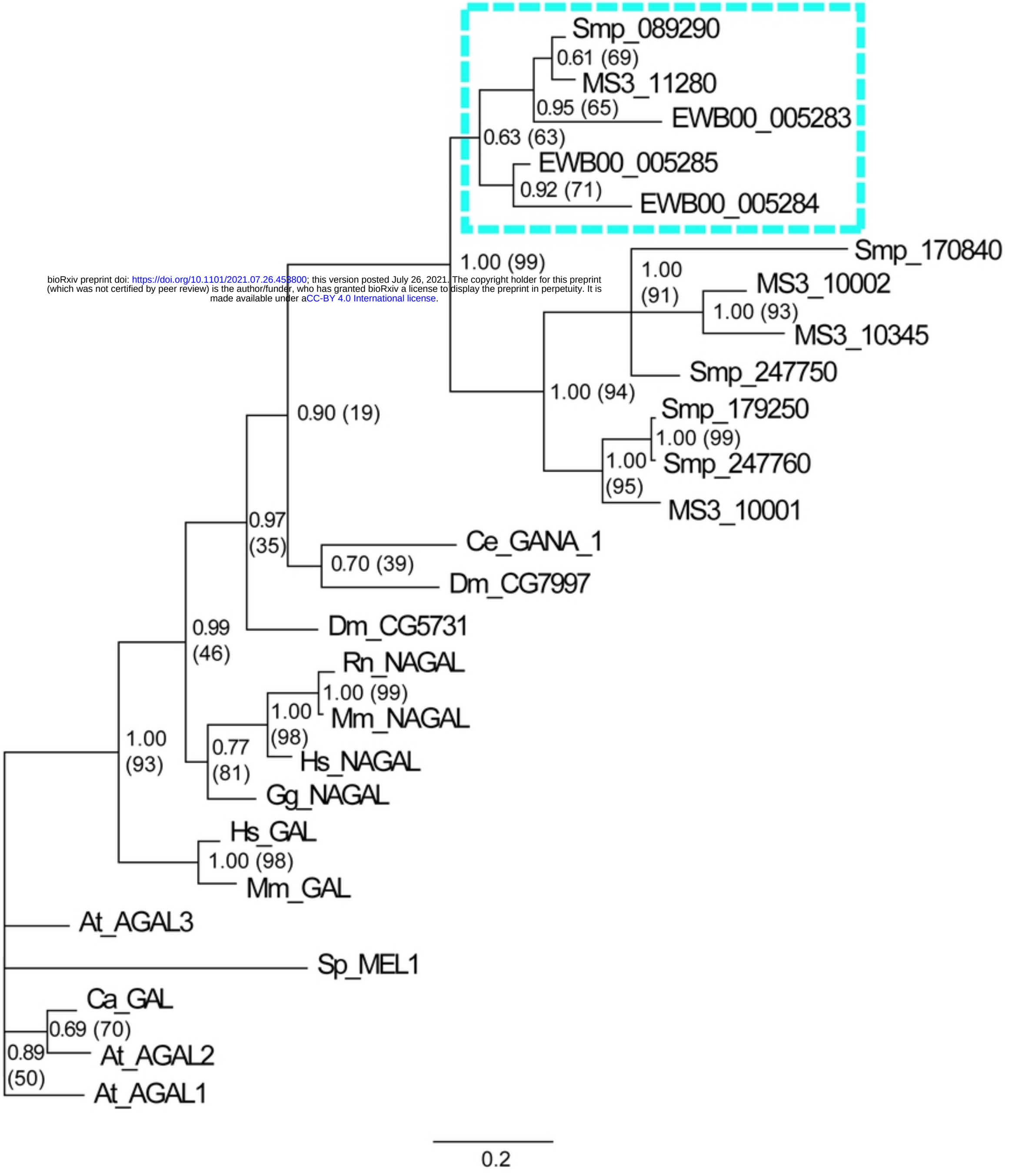


Figure 2

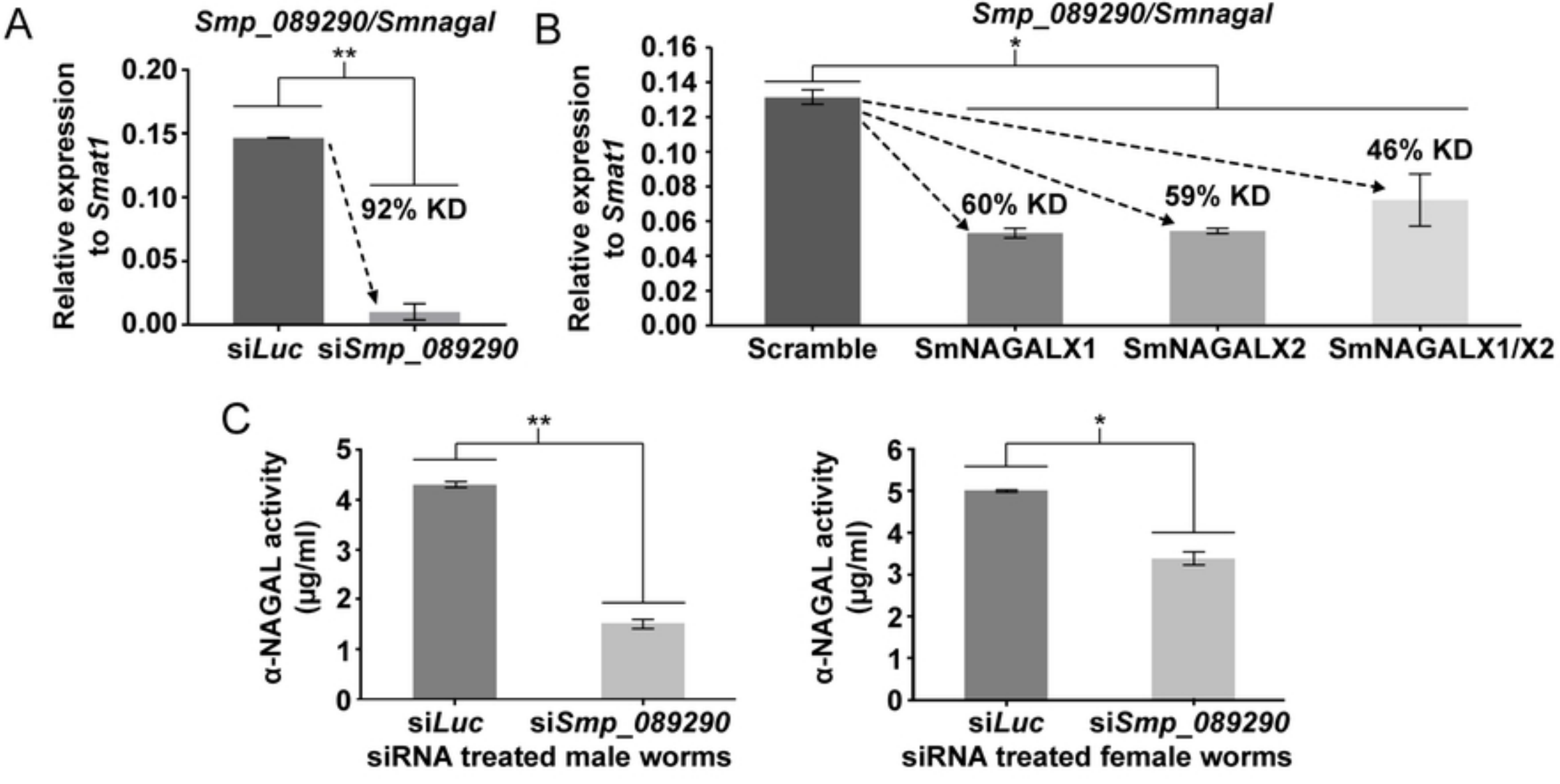

A
Smp_089290/Smnagal
**
Relative expression to Smat1
0.20
0.15
0.10
0.05
0.00
92% KD
siLuc
siSmp_089290
B
Smp_089290/Smnagal
*
Relative expression to Smat1
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
60% KD
59% KD
46% KD
Scramble
SmNAGALX1
SmNAGALX2
SmNAGALX1/X2
C
**
α-NAGAL activity (µg/ml)
5
4
3
2
1
0
siLuc
siSmp_089290
siRNA treated male worms
*
α-NAGAL activity (µg/ml)
6
5
4
3
2
1
0
siLuc
siSmp_089290
siRNA treated female worms

Figure 6