

## Profiling G protein-coupled receptors of *Fasciola hepatica* identifies orphan rhodopsins unique to phylum Platyhelminthes



Paul McVeigh<sup>a,\*</sup>, Erin McCammick<sup>a</sup>, Paul McCusker<sup>a</sup>, Duncan Wells<sup>a</sup>, Jane Hodgkinson<sup>b</sup>, Steve Paterson<sup>c</sup>, Angela Mousley<sup>a</sup>, Nikki J. Marks<sup>a</sup>, Aaron G. Maule<sup>a</sup>

<sup>a</sup> Parasitology & Pathogen Biology, The Institute for Global Food Security, School of Biological Sciences, Queen's University Belfast, Medical Biology Centre, 97 Lisburn Road, Belfast, BT9 7BL, UK

<sup>b</sup> Institute of Infection and Global Health, University of Liverpool, Liverpool, UK

<sup>c</sup> Institute of Integrative Biology, University of Liverpool, Liverpool, UK

### ARTICLE INFO

#### Keywords:

Anthelmintic  
Deorphanization  
Flukicide  
Genome  
Invertebrate  
Nervous system  
Neuropeptide  
RNA-Seq

### ABSTRACT

G protein-coupled receptors (GPCRs) are established drug targets. Despite their considerable appeal as targets for next-generation anthelmintics, poor understanding of their diversity and function in parasitic helminths has thwarted progress towards GPCR-targeted anti-parasite drugs. This study facilitates GPCR research in the liver fluke, *Fasciola hepatica*, by generating the first profile of GPCRs from the *F. hepatica* genome. Our dataset describes 147 high confidence GPCRs, representing the largest cohort of GPCRs, and the largest set of *in silico* ligand-receptor predictions, yet reported in any parasitic helminth. All GPCRs fall within the established GRAFS nomenclature; comprising three glutamate, 135 rhodopsin, two adhesion, five frizzled, one smoothed, and one secretin GPCR. Stringent annotation pipelines identified 18 highly diverged rhodopsins in *F. hepatica* that maintained core rhodopsin signatures, but lacked significant similarity with non-flatworm sequences, providing a new sub-group of potential flukicide targets. These facilitated identification of a larger cohort of 76 related sequences from available flatworm genomes, representing new members of existing groups (PROF1/Srfb, Rho-L, Rho-R, SrfA, Srfc) of flatworm-specific rhodopsins. These receptors imply flatworm specific GPCR functions, and/or co-evolution with unique flatworm ligands, and could facilitate the development of exquisitely selective anthelmintics. Ligand binding domain sequence conservation relative to deorphanised rhodopsins enabled high confidence ligand-receptor matching of seventeen receptors activated by acetylcholine, neuropeptide F/Y, octopamine or serotonin. RNA-Seq analyses showed expression of 101 GPCRs across various developmental stages, with the majority expressed most highly in the pathogenic intra-mammalian juvenile parasites. These data identify a broad complement of GPCRs in *F. hepatica*, including rhodopsins likely to have key functions in neuromuscular control and sensory perception, as well as frizzled and adhesion/secretin families implicated, in other species, in growth, development and reproduction. This catalogue of liver fluke GPCRs provides a platform for new avenues into our understanding of flatworm biology and anthelmintic discovery.

### 1. Introduction

*Fasciola* spp. liver fluke are pathogens of veterinary ruminants that threaten the sustainability of global meat and dairy production. Infection with *Fasciola* (fasciolosis/fascioliasis) inhibits animal productivity through liver condemnation, reduced meat and milk yields, and reduced fertility (for recent impact surveys see [Abunna et al. \(2010\)](#), [Sariözkan and Yalçın \(2011\)](#), [Howell et al. \(2015\)](#) and [Habarugira et al. \(2016\)](#)). *Fasciola* spp. also infect humans, with fascioliasis considered a neglected tropical disease ([Hotez et al., 2008](#)). Anthelmintic chemotherapy currently carries the burden of fluke

control, since there are no liver fluke vaccines ([Toet et al., 2014](#)). Six flukicidal active compounds are available for general use, with on-farm resistance reported for all except oxclozanide ([Kelley et al., 2016](#)). Resistance to the frontline flukicide, triclabendazole, also exists in human *F. hepatica* infections ([Winkelhagen et al., 2012](#); [Cabada et al., 2016](#)). Given the absence of alternative control methods, new flukicides are essential for secure future treatment of veterinary and medical liver fluke infections.

The helminth neuromuscular system is a prime source of molecular targets for new anthelmintics ([Martin and Robertson, 2010](#); [McVeigh et al., 2012](#); [Ribeiro and Patocka, 2013](#)), not least because many

\* Corresponding author.

E-mail address: [paul.mcveigh@qub.ac.uk](mailto:paul.mcveigh@qub.ac.uk) (P. McVeigh).

existing anthelmintics (dichlorvos, levamisole, morantel, piperazine, pyrantel, macrocyclic lactones, paraherquamide, amino acetonitrile derivatives) act upon receptors or enzymes associated with classical neurotransmission in nematodes (Wolstenholme, 2011; McVeigh et al., 2012). G protein-coupled receptors (GPCRs) that transduce signals from both peptidergic and classical neurotransmitters are of broad importance to helminth neuromuscular function. Despite industry efforts to exploit helminth GPCRs in the context of anthelmintic discovery (Lowery et al., 2003), only a single current anthelmintic (emodepside) has been attributed GPCR-directed activity as part of its mode of action (Saeger et al., 2001; Harder et al., 2003; Buxton et al., 2011). GPCRs are druggable targets, since 33% of human prescription medicines have a GPCR-based mode of action (Santos et al., 2017).

Despite two *F. hepatica* genomes (Cwiklinski et al., 2015; McNulty et al., 2017), no GPCR sequences have been reported from *F. hepatica*. In contrast, GPCRs have been profiled in the genomes of trematodes (*Schistosoma mansoni* and *Schistosoma haematobium* (Zamanian et al., 2011; Campos et al., 2014)), cestodes (*Echinococcus multilocularis*, *E. granulosus*, *Taenia solium* and *Hymenolepis microstoma* (Tsai et al., 2013)), and planaria (*Schmidtea mediterranea*, *Girardia tigrina* (Omar et al., 2007; Zamanian et al., 2011; Saberi et al., 2016)). These datasets illustrated clear differences in the GPCR complements of individual flatworm classes and species, with reduced complements in parasitic flatworms compared to planarians.

This study profiles the GPCR complement of the temperate liver fluke *F. hepatica* for the first time, permitting comparisons with previously characterised species that inform evolutionarily and functionally conserved elements of flatworm GPCR signalling. We have identified and classified 147 GPCRs by GRAFS family (glutamate, rhodopsin, adhesion, frizzled, secretin) assignment (Fredriksson et al., 2003), the majority of which are expressed in *Fasciola* RNA-Seq datasets. These include clear orthologues of GPCRs activated by known neurotransmitters, within which we performed the deepest *in silico* ligand-receptor matching analyses to date for any parasitic helminth. The latter predicted ligands for 17 *F. hepatica* GPCRs, designating these as primary targets for deorphanisation. Intriguingly, the dataset included a set of flatworm-expanded GPCRs lacking orthologues outside of phylum Platyhelminthes. Evolution of such GPCRs across the parasitic flatworm classes may have been driven by flatworm-specific functional requirements or co-evolution with flatworm ligands, either of which could help support novel anthelmintic discovery. This dataset provides the first description of GPCRs in liver fluke, laying a foundation for future advances in GPCR-directed functional genomics and flucicide discovery.

## 2. Materials and methods

### 2.1. Liver fluke sequence databases

We exploited two *F. hepatica* genome assemblies available from WormBase ParaSite (Howe et al., 2017), generated by Liverpool University ([http://parasite.wormbase.org/Fasciola\\_hepatica\\_prjeb6687/Info/Index/](http://parasite.wormbase.org/Fasciola_hepatica_prjeb6687/Info/Index/)) (Cwiklinski et al., 2015), and Washington University, St Louis ([http://parasite.wormbase.org/Fasciola\\_hepatica\\_prjna179522/Info/Index/](http://parasite.wormbase.org/Fasciola_hepatica_prjna179522/Info/Index/)) (McNulty et al., 2017).

### 2.2. Identification of GPCR-like sequences from *F. hepatica*

Fig. 1 summarises our GPCR discovery methodology, which employed Hidden Markov Models (HMMs) constructed from protein multiple sequence alignments (MSAs) of previously described *S. mansoni* and *S. mediterranea* GPCR sequences (Zamanian et al., 2011). Individual HMMs were constructed for each GRAFS family (Fredriksson et al., 2003). Alignments were generated in Mega v7 ([www.megasoftware.net](http://www.megasoftware.net)) (Kumar et al., 2016) using the Muscle algorithm with default parameters. HMMER v3 (<http://hmmer.org>) was employed

to construct family-specific HMMs (*hmmbuild*) from alignments and these were searched (*hmmsearch*) against a predicted protein dataset from *F. hepatica* genome PRJEB6687 consisting of 33,454 sequences (Cwiklinski et al., 2015); default parameters were used for *hmmsearch* and *hmmbuild*. Returned sequences were filtered for duplicates and ordered relative to the *hmmsearch* scoring system, enabling the classification of hits according to the GRAFS family to which they showed most similarity (i.e. highest score, lowest E value). All remaining returns were then used as BLAST queries (BLASTp and tBLASTn with default parameters) to identify matching, or additional, sequences originating from the PRJEB6687 and PRJNA179522 genomes (Fig. 1). Where sequences appeared in both genomes, we kept the longest annotated sequence (S1 Table).

### 2.3. GPCR annotation

Sequences resulting from HMM searches were filtered by transmembrane (TM) domain composition, using hmmtop (<http://www.sacs.ucsf.edu/cgi-bin/hmmtop.py>) (Tusnday and Simon, 1998, 2001). Sequences containing  $\geq 4$  TMs were analysed as described below.

#### 2.3.1. Homology analyses

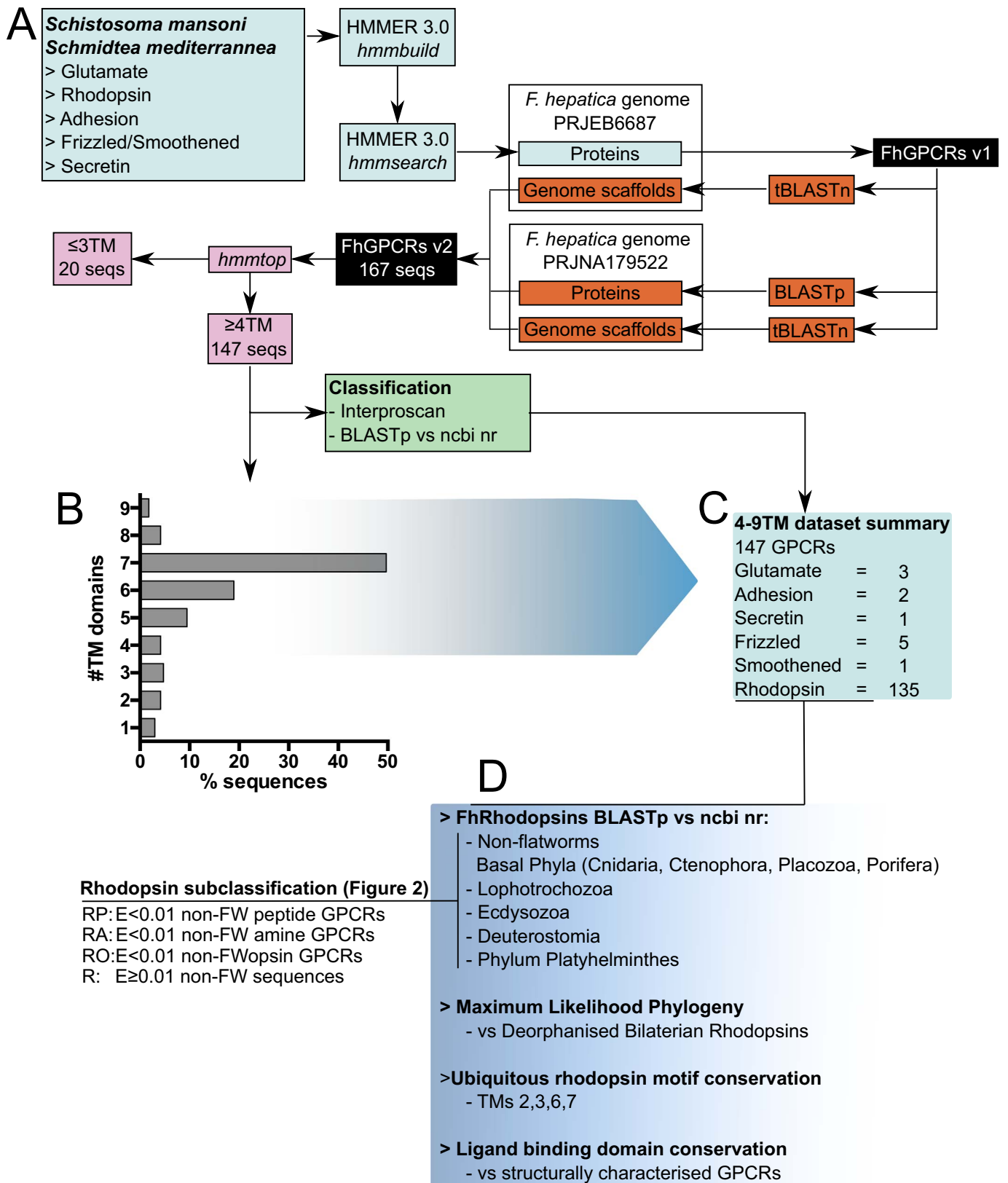
All GPCRs were used as BLASTp (Altschul et al., 1990) queries, to identify their closest (highest scoring) match in the ncbi non-redundant (nr) protein sequence dataset (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>), with default settings and the “Organism” field set to exclude Platyhelminthes (taxid: 6157). All GPCRs were additionally searched against more phylogenetically limited datasets, by using the “Organism” field to limit the BLASTp searches to: (i) Basal phyla, Ctenophora (taxid:10197), Porifera (taxid:6040), Placozoa (taxid:10226), Cnidaria (taxid:6073); (ii) Superphylum Lophotrochozoa (taxid:1206795), excluding phylum Platyhelminthes (taxid: 6157); (iii) Superphylum Ecdysozoa (taxid: 1206794); (iv) Superphylum Deuterostomia (taxid: 33511). For BLASTp searches against other flatworms, we performed local BLAST+ (Camacho et al., 2008) on the WBPS9 release of WormBase Parasite, which included predicted protein datasets from 30 flatworm species. In all cases, we recorded the single highest scoring hit, or recorded “no significant similarity found” in cases where no hits were returned (Table S1); sequences generating both GPCR hits and “no significant similarity” were retained. Where the top hit was not to a GPCR, that sequence was removed from the dataset.

#### 2.3.2. Domain composition

GPCR identities were confirmed using InterProScan Sequence Search ([www.ebi.ac.uk/interpro/search/sequence-search](http://www.ebi.ac.uk/interpro/search/sequence-search)) (Jones et al., 2014) and/or HMMER HMMScan ([www.ebi.ac.uk/Tools/hmmer/search/hmmScan](http://www.ebi.ac.uk/Tools/hmmer/search/hmmScan)) (Finn et al., 2015), with default parameters. Again, sequences returning non-GPCR domains were omitted from the dataset, with all others retained.

#### 2.3.3. Motif identification

As an additional measure of confidence in our identifications, we analysed the presence/absence of key motifs diagnostic of receptor families and subfamilies. These analyses were performed for rhodopsins generally, the ligand binding domains (LBDs) of rhodopsin receptors for acetylcholine (ACh), neuropeptide F/Y (NPF/Y), octopamine and serotonin (5-hydroxytryptamine, 5HT), and for the LBDs of glutamate and frizzled/smoothed families. Motifs were identified via protein multiple sequence alignment (MSA) of GPCRs, performed in MAFFT ([www.mafft.cbrc.jp/alignment/server/](http://www.mafft.cbrc.jp/alignment/server/)) (Katoh et al., 2017), using “E-INS-i” parameters, for sequences with multiple conserved domains. Only identical amino acids were accepted at each site, with conservation expressed as % identity across all sites. Motif illustrations were generated using WebLogo 3 (<http://weblogo.threeplusone.com>) (Crooks et al., 2004).



**Fig. 1. Methods for discovery and annotation of *Fasciola hepatica* G protein coupled receptors (FhGPCRs).** (A) Hidden Markov Models (HMMs) representing glutamate, rhodopsin, adhesion, frizzled/smoothened and secretin families, and two rhodopsin subfamilies, were built from protein multiple sequence alignments of *Schistosoma mansoni* and *Schmidtea mediterranea* GPCRs. HMMs were built and searched respectively using the *hmmbuild* and *hmmsearch* modules of HMMER v3.0. Searches were performed against two publicly available *F. hepatica* genomes using *hmmsearch* and basic local alignment search tool (BLAST) tools. Each putative FhGPCR sequence was assessed for transmembrane (TM) domain composition with *hmmtop* before classification using tools including BLASTp, Interproscan and CLANS. (B) The largest proportion (49%) of FhGPCRs carried the full complement of 7 TMs, with 88% of sequences bearing at least 4 TMs. (C) GRAFS composition of 147 FhGPCRs carrying ≥4 TMs. (D) Rhodopsins were subject to further classification, including BLASTp vs datasets representing major non-flatworm animal phyla and superphyla. These rhodopsin homology classifications fed back into phylogenetic analyses versus deorphanised bilaterian GPCRs to confirm their putative ligand selectivity, with a final analysis of ligand binding domain composition comparing conservation of ligand interacting residues for characterised GPCRs reported in the literature with our *F. hepatica* assignments.

**Table 1**

Comparison of the *Fasciola hepatica* G-protein coupled receptor (GPCR) complement with those reported from other flatworms. Species complements are shown in the context of GRAFS nomenclature (Fredriksson et al., 2003). <sup>a</sup> Zamanian et al. (2011); <sup>b</sup> Campos et al. (2014); <sup>c</sup> Tsai et al., 2013; <sup>d</sup> Saberi et al., 2016. Saberi et al. (2016) described 566 GPCRs in *Schmidtea mediterranea*, of which 516 fall within GRAFS nomenclature.

	<i>F. hepatica</i>	<i>S. mansoni</i> <sup>a</sup>	<i>S. mansoni</i> <sup>b</sup>	<i>S. haematobium</i> <sup>b</sup>	<i>E. multilocularis</i> <sup>c</sup>	<i>S. mediterranea</i> <sup>a</sup>	<i>S. mediterranea</i> <sup>d</sup>
Glutamate	3	2	2	2	5	9	11
Rhodopsin	135	105	59	53	48	418	461
Adhesion	2	3	–	–	4	9	14
Frizzled/Smoothened	6	5	4	4	5	11	10
Secretin	1	2	5	5	1	1	20
<b>Total</b>	<b>147</b>	<b>117</b>	<b>64</b>	<b>64</b>	<b>83</b>	<b>448</b>	<b>516*</b>

### 2.3.4. Phylogenetic reconstruction

Maximum likelihood (ML) phylogenetic trees were constructed using PhyML (<http://www.phylogeny.fr>) (Dereeper et al., 2008), from protein MSA generated in MAFFT ([www.mafft.cbrc.jp/alignment/server/](http://www.mafft.cbrc.jp/alignment/server/)). Alignments were manually edited (in Mega v7) to include only TM domains, by removing extra-membrane blocks aligned with human glutamate, rhodopsin, adhesion or frizzled proteins. Trees were constructed from these TM-focused alignments in PhyML using default parameters, with branch support assessment using the approximate likelihood ratio test (aLRT), under “SH-like” parameters. Trees, exported from PhyML in newick format were drawn and annotated in FigTree v1.4.2 (<http://tree.bio.ed.ac.uk/software/figtree/>).

### 2.4. RNA-seq analyses

Expression of *F. hepatica* GPCRs was investigated in publically available and in-house generated RNA-Seq datasets. These included developmentally staged Illumina transcriptome reads associated with the Cwiklinski et al. (2015) *F. hepatica* genome (reads accessed from the European Nucleotide Archive at <http://www.ebi.ac.uk/ena/data/search?query=PRJEB6904>). These samples originated from distinct developmental stages of US Pacific Northwest Wild Strain *F. hepatica* (Baldwin Aquatics), including egg ( $n = 2$ ), metacercariae (met;  $n = 4$ ), *in vitro* NEJs 1 h post-excystment (NEJ1h;  $n = 1$ ), *in vitro* NEJs 3 h post-excystment (NEJ3h;  $n = 2$ ), *in vitro* NEJs 24 h post-excystment (NEJ24h;  $n = 2$ ), *ex-vivo* liver-stage juveniles (juv1;  $n = 1$ ) and *ex-vivo* adult parasites (Ad;  $n = 3$ ). Our in-house datasets were generated from *ex vivo* liver stage *F. hepatica* juveniles (Italian strain, Ridgeway Research Ltd, UK), recovered from rat (Sprague Dawley) hosts at 21 days following oral administration of metacercariae (juv2;  $n = 3$ ). All animal use was approved by Queen's University Belfast's Animal Welfare and Ethical Review Body, and performed under Home Office project license PPL2764.

Total RNA, extracted with Trizol (ThermoFisher Scientific) from each of the 3 independent biological replicates, was quantified and quality checked on an Agilent Bioanalyzer, converted into paired-end sequencing libraries and sequenced on an Illumina HiSeq2000 by the Centre for Genomic Research at the University of Liverpool, UK. RNA samples were spiked prior to library construction with the ERCC RNA Spike-In Mix (ThermoFisher Scientific) (Jiang et al., 2011). All read samples were analysed using the TopHat, Cufflinks, Cuffmerge, Cuffdiff pipeline with default parameters, (Langmead et al., 2009; Trapnell et al., 2009, 2010, 2012a, 2012b; Roberts et al., 2011), with mapping against PRJEB6687 genome sequence and annotation files (accessed from WormBase Parasite; <http://parasite.wormbase.org/ftp.html>). Data were expressed as number of fragments mapped per million mapped reads per kilobase of exon model (FPKM). In juv2 datasets we discarded GPCRs represented by fewer than 0.5 FPKM (the minimum linear sensitivity that we detected with our ERCC spike in); for the staged datasets, we included only receptors represented by  $\geq 0.5$  FPKM in at least one life stage. Heatmaps were generated with heatmapper (<http://www.heatmapper.ca/>) (Babicki et al., 2016) set for Average Linkage, and Pearson Distance Measurement.

## 3. Results and discussion

### 3.1. A first look at GPCRs in the *F. hepatica* genome

This study represents the first description of the GPCR complement of the temperate liver fluke, *F. hepatica*. Using HMM-led methods to examine available *F. hepatica* genome datasets, we identified 166 GPCR-like sequences in *F. hepatica* (Figs. 1 and S1 Table). Fig. 1B shows that 49.7% contained 7 TM domains, with 88% of sequences containing at least four TMs. The remainder of this manuscript focuses on 147 sequences containing  $\geq 4$  TM domains (S1 Table; S2 Text). Twenty-two sequences containing  $\leq 3$  TMs were not analysed further (Fig. 1).

Our  $\geq 4$  TM dataset (147 sequences) was comprised of three glutamate, 135 rhodopsin, two adhesion, five frizzled, one smoothened, and one secretin GPCR. Sequence coverage was generally good in terms of TM and extracellular domain representation, so we did not attempt to extend truncated sequences into full-length receptors. The overall dataset contained excellent representation of seven TM domains, while N-terminal extracellular LBDs and cysteine-rich domains (CRD) were also detected (in glutamate, frizzled/smoothened, adhesion families). However, we could not identify N-terminal secretory signal peptides in any sequence, suggesting incomplete sequence coverage at extreme N-termini. Rhodopsins are designated by ubiquitously conserved motifs on TMs 2, 3, 6 and 7. All rhodopsin sequences contained at least one of these motifs (Fig. 2 and S3 Table), including in the highly diverged flatworm-specific rhodopsins described below.

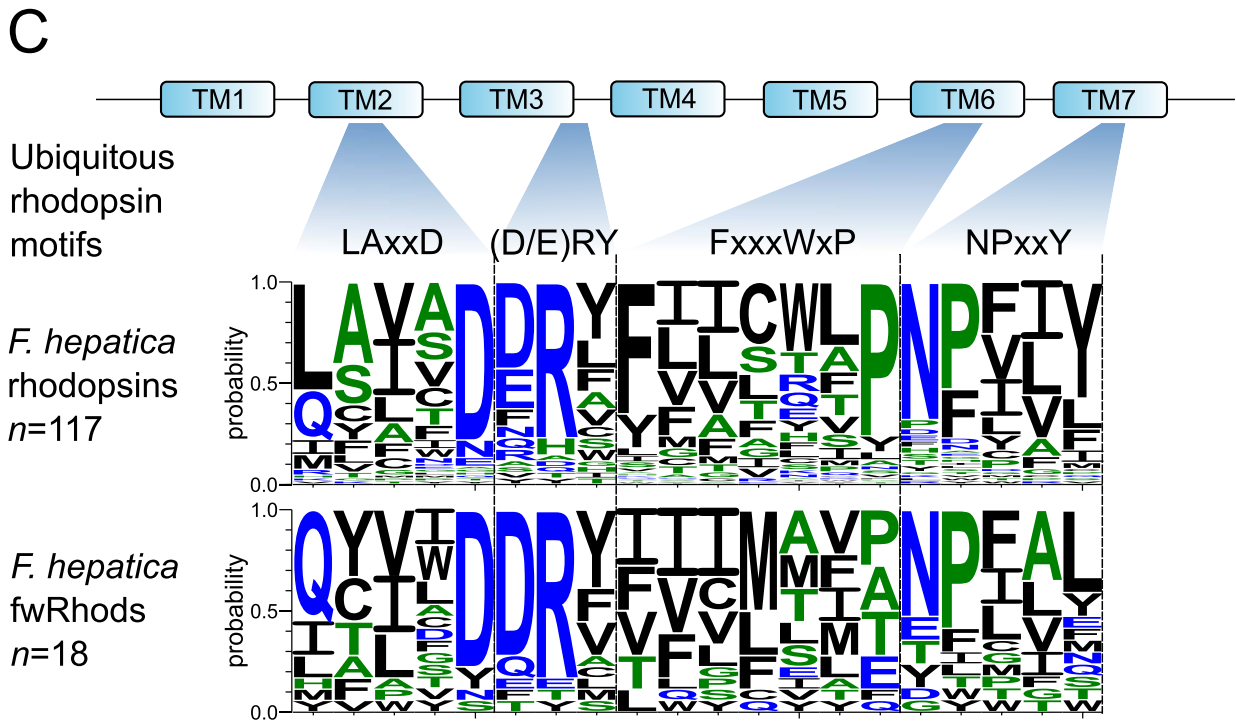
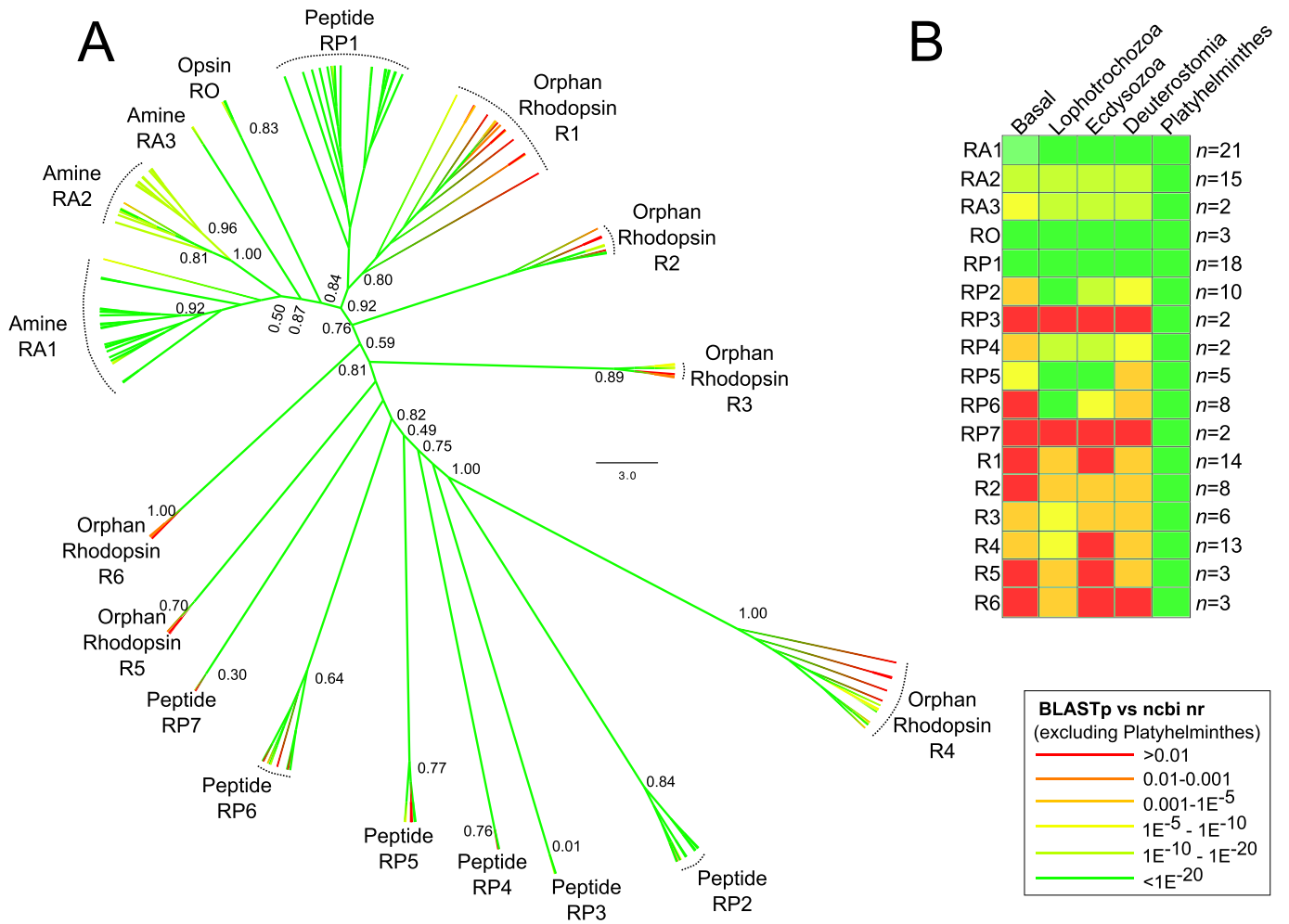
Table 1 compares the *F. hepatica* GPCR complement with other flatworms, illustrating that *F. hepatica* has the largest GPCR complement reported from any parasitic flatworm to date. The bulk of the expansion involves rhodopsins, while the other GRAFS families are comparable between *F. hepatica* and other flatworm parasites.

### 3.2. Stringent annotation of flatworm-specific orphan rhodopsin GPCRs in *F. hepatica*

Encompassing 135 sequences, the rhodopsin family is the largest of the GRAFS classifications in *F. hepatica*. Rhodopsins comprise four subfamilies ( $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\delta$ ) (Lagerström and Schiöth, 2008); we identified members of both  $\alpha$  and  $\beta$  groups, with nucleotide-activated (P2Y) receptors ( $\gamma$  group), and olfactory ( $\delta$  group) receptors absent from our dataset (Figs. 1 and 2; S1 Table). The *F. hepatica*  $\alpha$  subfamily contained 38 amine receptors and three opsins, with the  $\beta$  subfamily comprised of at least 47 peptide receptors. Homology-based annotations were supported by an ML phylogeny (Fig. 2A), which clearly delineated between amine and opsin  $\alpha$  clades, and the peptide-activated  $\beta$ -rhodopsin clades. Amine and peptide receptors were further delineated by additional phylogenetic and structural analyses, permitting high-confidence assignment of putative ligands to 16 GPCRs (see section 3.4).

Six clades contained an additional 44 rhodopsin sequences with low scoring (median  $E = 5.6e^{-5}$ ) similarity matches to a range of disparate  $\alpha$  and  $\beta$  rhodopsins. Due to the subsequent difficulty in designating these clades as amine, peptide or opsin, we labelled them orphan rhodopsins (“R” clades in Fig. 2A). Eighteen GPCRs within the orphan





(caption on next page)

**Fig. 2. Phylogenetic classification of *Fasciola hepatica* rhodopsin G protein-coupled receptors.** (A) Maximum-likelihood cladogram of *F. hepatica* rhodopsins. Phylogeny delineated clades containing rhodopsins with distinct homologies (RA, amine; RP, peptide; RO, opsin; R, orphan rhodopsin). The orphan clades contained sequences with generally low BLASTp similarity to their closest non-flatworm BLASTp hit, but concentrated within them were 18 sequences with exceptionally low ( $E > 0.01$ ) BLASTp similarity to non-flatworm sequences (fwRhods). The tree was midpoint rooted and was generated from a multiple protein sequence alignment trimmed to TM domains I–VII. Numbers at nodes indicate statistical support from approximate likelihood ratio test (aLRT). Tip colours are coded according to the E-value scale (as indicated) of that GPCR's closest BLASTp match in the ncbi nr database, excluding phylum Platyhelminthes. (B) Summary of sequence similarity comparisons between GPCRs within each rhodopsin clade, and their closest BLASTp hits in four major phylogenetic groups (1. Basal: Cnidaria, Ctenophora, Porifera, Placozoa; 2. Superphylum Lophotrochozoa, omitting Platyhelminthes; 3. Superphylum Ecdysozoa; 4. Superphylum Deuterostomia; 5. Phylum Platyhelminthes). BLASTp E-value (median) is summarised in each case, colour coded as a heat map on the same colour scale as (A). The number of GPCRs comprising each *F. hepatica* clade ( $n$ ) is also indicated. (C) Sequence diversity within ubiquitous rhodopsin motifs of the majority (117) of the *F. hepatica* rhodopsins (upper panel), compared to those motifs in 18 *F. hepatica* fwRhods (lower panel). The mammalian consensus motifs are illustrated above the top panel, along with an illustration of the location of each motif within the rhodopsin 7TM domain structure. Some variability is visible within the TM2 and TM6 motifs, but TM3 and TM7 motifs are well conserved. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

clades displayed exceptionally low similarity scores relative to non-flatworm sequences (Fig. 2A and B). Seven returned no-significant hits in BLASTp searches against non-flatworm members of the ncbi nr dataset (the most diverse sequence dataset available to the research community), and the remaining eleven scored  $E > 0.01$ . Domain analysis (InterPro) identified rhodopsin domains (IPR000276 or IPR019430) in thirteen of these (S1 Table, S3 Table), confirming their identity as rhodopsin-like GPCRs. More troublesome to classify were five that, in addition to lacking significant BLASTp identity to non-flatworm sequences, also lacked any identifiable protein domains/motifs (with the exception of TM domains). We annotated these as rhodopsins because: (i) They did not contain motifs/domains representative of any other protein family; (ii) They displayed topological similarity to GPCRs (ten had seven TM domains, seven had six TMs, one had five TM domains); (iii) They contained at least two of the conserved rhodopsin motifs in TM domains 2, 3, 6 and 7 similar to those seen in the rest of the *F. hepatica* rhodopsins (Fig. 2C; S4 Table). As highly diverged rhodopsins with little or no sequence similarity versus host species, these 18 *F. hepatica* receptors have obvious appeal as potential targets for flukicidal compounds with exquisite selectivity for parasite receptors over those of the host. This potential is contingent on future work demonstrating essential functionality for these receptors; showing their wider expression across flatworm parasites would enable consideration of anthelmintics with multi-species activity. To investigate the latter question, we used BLASTp to search the 18 *F. hepatica* rhodopsins against other available genomes representing phylum Platyhelminthes.

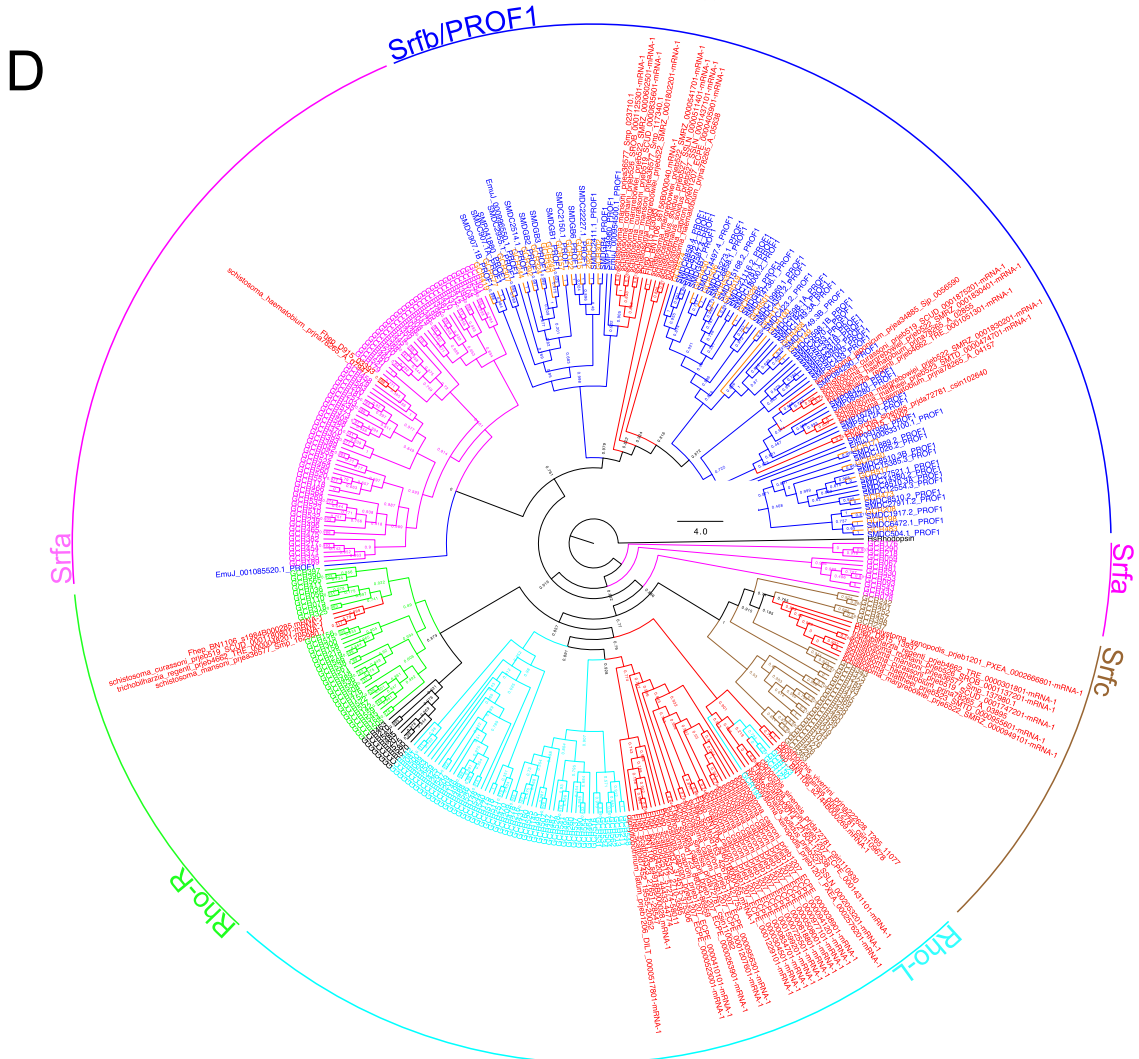
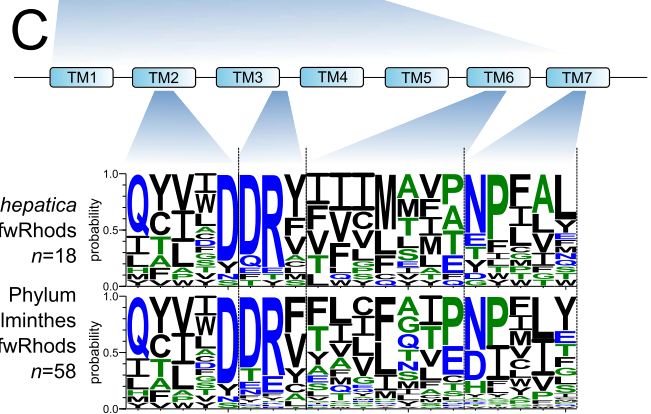
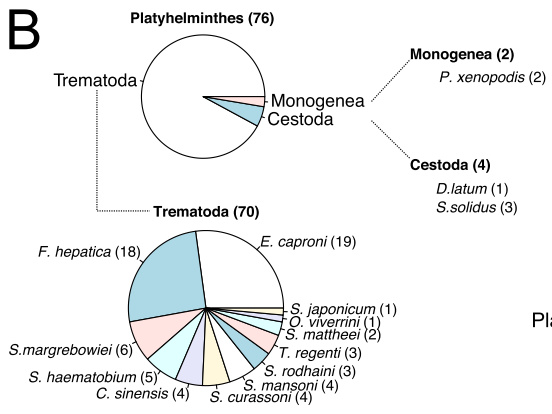
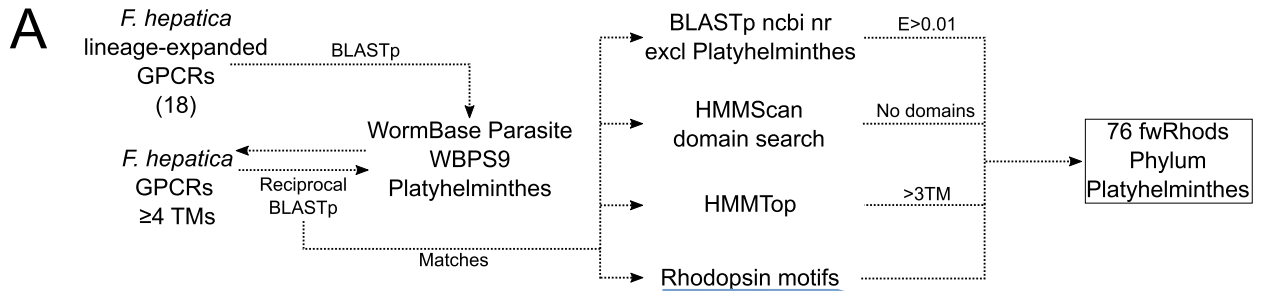
### 3.3. An orphan family of lineage-expanded rhodopsins in flatworm genomes

Although lacking similarity against non-flatworm datasets, each of the 18 lineage-expanded *F. hepatica* rhodopsins returned high-scoring hits in BLASTp searches against the genomes of other flatworms (WormBase Parasite release WBPS9). All returns were subsequently filtered through a stringent five-step pipeline (Fig. 3A) consisting of: (i) Removal of duplicate sequences; (ii) Exclusion of sequences containing fewer than four TM domains; (iii) A requirement for reciprocal BLASTp against the *F. hepatica* genome to return a top hit scoring  $E < 0.001$  to one of the original 18 *F. hepatica* queries; (iv) A requirement for BLASTp against ncbi nr non flatworm sequences to return a top hit scoring  $E > 0.01$ ; (v) Removal of sequences lacking conservation of the ubiquitous rhodopsin motifs seen in the divergent *F. hepatica* rhodopsins (Figs. 2C and 3C). The latter motifs were largely absent from cestode rhodopsins (with the exception of a single sequence from *Diphyllobothrium latum*, and three sequences from *Schistocephalus solidus*), and present in only two sequences from a single monogenean (*Protopolystoma xenopodis*). This left our final dataset consisting of 76 “flatworm-specific” rhodopsins (fwRhods; Fig. 3B, Table S4) in phylum Platyhelminthes, heavily biased towards trematodes (70 sequences). Nineteen sequences from nine species of cestode were omitted from the final dataset despite meeting the inclusion criteria in most respects, because they lacked conservation of ubiquitous rhodopsin motifs (filtering step (v)). Although their further characterisation was beyond the

scope of this study, they warrant more detailed examination in future studies as potential cestode-specific rhodopsins. Note that our filtering pipeline also excluded initial hits from *Gyrodactylus salaris* (Monogenea), and the turbellarians *Macrostomum lignano* and *S. mediterranea*. Individual species complements of fwRhods showed some consistency (Fig. 3B); the trematodes *F. hepatica* and *Echinostoma caproni* (both phylum Platyhelminthes, order Echinostomida) bore 18 and 19 sequences, respectively, most species of family Schistosomatidae contained 3–4 sequences each. The inclusion of two cestode species and a single monogenean may be an indication of the existence of distantly related rhodopsins in those lineages, rather than a true measure of the extent of cestode and monogenean fwRhods diversity. Proper classification of these groups will require further, Class-focused study.

Our method for identification of fwRhods is supported by a similar BLAST-driven approach used to identify highly diverged “hidden orthologues” in flatworms (Martin-Duran et al., 2017), as well as by similar, less stringent, methods used to identify PROF1 GPCRs (Zamanian et al., 2011). It should be noted that the existence of sequences lacking sequence similarity to genes of other species is not a new finding. “Taxonomically-restricted genes” comprise 10–20% of every sequenced eukaryote genome, and may be essential for phylum-specific morphological and molecular diversity (Khalturin et al., 2009). We also considered how our fwRhods compare to previously reported groups of flatworm restricted GPCRs in *S. mansoni*, *S. mediterranea* and *E. granulosus* (Zamanian et al., 2011; Tsai et al., 2013; Saberi et al., 2016). Phylogenetic comparisons (Fig. 3D) demonstrated that the previously described *Schmidtea* Srfb cluster (Saberi et al., 2016) and the PROF1 clade (*E. multilocularis*, *Schmidtea*, *S. mansoni* (Zamanian et al., 2011; Tsai et al., 2013) are equivalent, and likely represent a single group. Our phylogeny added 23 fwRhods to this clade, including three from *F. hepatica* (BN1106\_s6156B000040, D915\_03083, D915\_13002). Fig. 3D designated the remaining fwRhods within additional pre-existing groups (Saberi et al., 2016), placing 34 within Rho-L (including eight from *F. hepatica*), nine in Srfc (one from *F. hepatica*), four in Rho-R (one from *F. hepatica*) and two in Srfb (one from *F. hepatica*). Four fwRhods sequences were omitted from this tree due to poor alignment.

There is no set definition for lineage specificity in the flatworm GPCR literature, with the previous studies describing PROF1 (Zamanian et al., 2011), Srfb/b/c and RhoL/R (Saberi et al., 2016) receptors employing distinct methods and criteria (we have employed a similar, but more stringent, E value-driven approach to the former). An additional complicating factor is that many flatworm GPCRs described as taxonomically restricted still return high scoring matches from BLASTp searches of non-flatworm sequence datasets. For example, applying our BLASTp  $E \geq 0.01$  cutoff (modified from Pearson, 2013) to these published groups, would exclude 57 of the 62 PROF1s described from *S. mansoni* and *S. mediterranea* (most of the excluded sequences in this case can be explained by expansion in the ncbi nr dataset since their description in 2011), and 287 of the 318 RhoL/R and Srfb/b/c flatworm-specific clusters in *S. mediterranea*. This indicates the difficulty in interpreting existing definitions of “lineage specificity” or “taxonomic restriction” amongst flatworm GPCRs, and we therefore feel justified in applying our own simple, but more stringent definition for



(caption on next page)

**Fig. 3. Identification of flatworm-specific rhodopsins (fwRhods) in genomes from phylum Platyhelminthes.** (A) The 18 *Fasciola hepatica* GPCRs in our dataset that had poor BLASTp similarity ( $E > 0.01$ ) to non-flatworm sequences in the ncbi nr dataset (lsGPCRs), were used as queries in BLASTp searches of flatworm genomes in Parasite (release WBFS9). All hits scoring  $E < 0.01$  were back-searched by BLASTp against our *F. hepatica* GPCR dataset. Sequences scoring  $E < 0.01$  against one of the original *F. hepatica* GPCRs were retained as matches. These sequences were then filtered to identify those lacking matches in ncbi nr, lacking non-GPCR protein domains, possessing at least 4 transmembrane (TM) domains, and containing rhodopsin motifs consistent with those seen in the majority of *F. hepatica* rhodopsins (see C). (B) This process identified 76 fwRhods in phylum Platyhelminthes, the majority (70) of which were from class Trematoda. Small numbers were returned from classes Cestoda and Monogenea. Note that no fwRhods fitting these criteria were identified in class Turbellaria. (C) Sequence diversity within ubiquitous rhodopsin motifs of 18 *F. hepatica* fwRhods (upper panel), compared to those motifs in the 58 fwRhods identified in the wider phylum (lower panel); motifs are broadly similar between *F. hepatica* and the rest of the phylum. (D) Maximum likelihood phylogeny of 76 fwRhods, alongside flatworm-specific rhodopsins described previously (70 platyhelminth rhodopsin orphan family 1 (PROF1) (Zamanian et al., 2011; Tsai et al., 2013), and 245 *S. mediterranea* G protein coupled receptor [GPCRs, comprising RhoL, RhoR, SrfA, Srfb and Srfc families, reported as lacking non-flatworm homologues (Saberi et al., 2016)] with branches coloured to indicate Family (magenta, SrfA; dark blue, Srfb/PROF1; brown, Srfc; cyan, Rho-L; green, Rho-R; red, fwRhod). Tree was rooted to a human rhodopsin (P08100) and was generated from an alignment trimmed to transmembrane domains I-VII. Numbers at nodes indicate statistical support from approximate likelihood ratio test (aLRT). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

taxonomically restricted GPCRs from *Fasciola*, and their orthologues in other flatworms. Despite taking a slightly different approach to previous work, the taxonomically-restricted nature of our fwRhods was validated in every case by comparative analysis with other tools. The WormBase Parasite community resource provides comparative genomics analyses for every gene in available parasite genomes. These are driven by Ensembl Compara pipelines (Vilella et al., 2009; Howe et al., 2017) that identify orthologues and paralogues for each parasite gene represented by a gene model. These tools confirm that all of the sequences we have designated as fwRhod in *F. hepatica* and other flatworms, lack orthologues outside of phylum Platyhelminthes, and our phylogenetic analyses confirm that they represent new members of existing groups.

We have established the existence of a group of rhodopsin GPCRs that appear restricted to, and expanded in, phylum Platyhelminthes. By definition these receptors are orphan (i.e. their native ligands are unknown), so key experiments must focus on identifying their ligands and functions. Such experiments can exploit the expanding molecular toolbox for flatworm parasites, which in *F. hepatica* includes RNA interference (RNAi) (McGonigle et al., 2008; Rinaldi et al., 2008; Dell'Oca et al., 2014; McVeigh et al., 2014) interfaced with enhanced *in vitro* maintenance methods, and motility, growth/development and survival assays (McGonigle et al., 2008; McCammick et al., 2016; McCusker et al., 2016). Our phylogeny (Fig. 2A) suggests that fwRhods are more similar to peptide than amine receptors. If their heterologous expression can be achieved, one approach to characterisation would be to screen them with the growing canon of peptide ligands from flatworms (McVeigh et al., 2009; Collins et al., 2010; Koziol et al., 2016), as well as from other genera, in a receptor activation assay. Subsequent localisation of their spatial expression patterns would provide additional data that would inform function.

### 3.4. Predicting ligands for *F. hepatica* rhodopsin GPCRs

In addition to the flatworm-specific fwRhod sequences described above, for which the ligands and functions remain cryptic, we also identified many rhodopsins with clear similarity to previously annotated GPCRs. Fig. 2A shows the phylogenetic delineation of these sequences into amine-, opsin- and peptide-like receptors, distinctions that are supported by BLASTp comparisons with general (ncbi nr) and lineage-specific (superphylum level) datasets, as well as by gross domain structure (InterProScan) (S1 Table). These data provided a foundation for the deeper classification of putative ligand-receptor matches.

The structure and function of GPCR LBDs can be studied using molecular modelling to predict interactions with receptor-bound ligands. These predictions can then be validated by targeted mutagenesis of residues within the LBD, measuring impacts with downstream signalling assays. Such experiments have been performed in model vertebrates and invertebrates, enabling identification of evolutionarily conserved binding residues/motifs. These data inform the assignment of putative ligands to newly discovered receptors. Since mutagenesis experiments have not yet been performed in flatworm GPCRs, we employed a comparative approach to identify 17 *F. hepatica* rhodopsins

with LBD motifs diagnostic of receptors for NPF/Y, 5-HT, octopamine (Oct) or acetylcholine (ACh) (Fig. 4; S4 Table), thus enabling *in silico* ligand-receptor matching of these GPCRs.

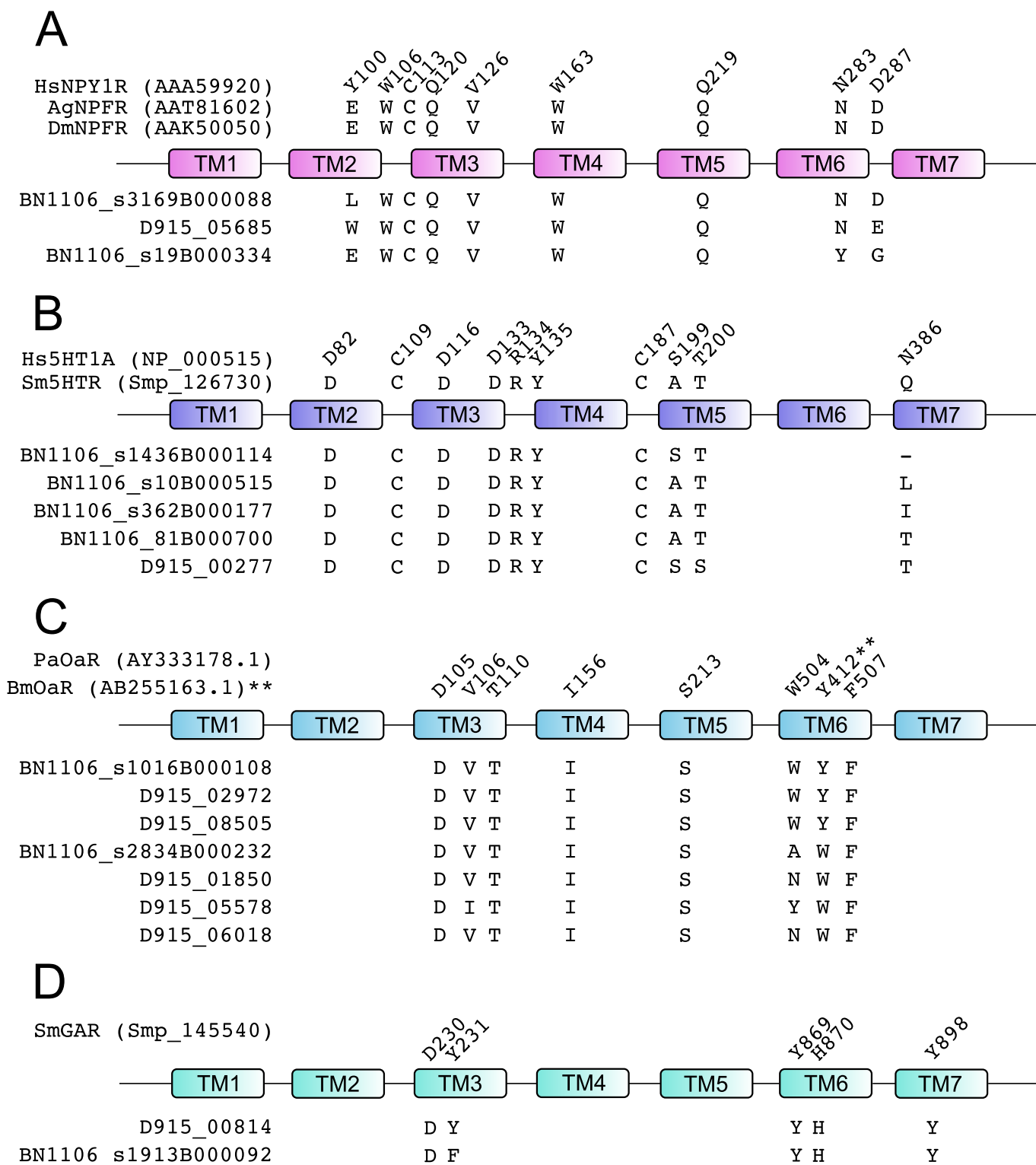
Comparison of *F. hepatica* rhodopsins by structural alignment with LBD residues conserved across vertebrate NPY and dipteran NPF receptors (Sautel et al., 1995, 1996; Berglund et al., 2002; Åkerberg et al., 2010; Fällmar et al., 2011; Vogel et al., 2013) identified three peptide receptors with more than 75% identity across 9 ligand-interacting positions (Fig. 4A). The two highest scoring GPCRs (BN1106\_s3169B000088 and D915\_05685) are also found, in our phylogenetic analysis (S5 Figure) in the same clade as the deorphanized NPF/Y receptors of human (HsNPYR2), *Glossina mortisans* (Glomo-NPFR) and *S. mediterranea* (SmedNPYR1). These data designate these three *F. hepatica* GPCRs as prime candidates for further work to deorphanize and confirm these receptors as NPF/Y-activated, and to probe the biology of NPF/Y receptors in parasitic flatworms. A single NPF/Y receptor has been functionally characterised in *S. mediterranea*, displaying a role in the maintenance of sexual maturity (Saberi et al., 2016). If related functions are conserved in liver fluke NPF/Y receptors they could have appeal as therapeutic targets in adult fluke that could interrupt parasite transmission, although their utility for the control of acute fasciolosis, caused by migrating juveniles, would be open to question.

Broad phylogenetic comparison of our peptide receptor set with a comprehensive collection of deorphanized bilaterian rhodopsin GPCRs (S5 Figure), identified *F. hepatica* receptors similar to those for myomodulin, FLP, luquin and Neuropeptide KY (NKY). These ligands have all been predicted or demonstrated in previous biochemical or *in silico* studies of flatworm neuropeptides (McVeigh et al., 2009; Collins et al., 2010; Koziol et al., 2016). We also uncovered *F. hepatica* GPCRs with phylogenetic similarity to allatotropin, allatostatin, thyrotropin-releasing hormone and sex peptide receptors. These ligands have not yet been reported in flatworms, although the existence of allatostatin-like receptors in flatworms is supported by the inter-phyla activity of arthropod allatostatins in helminth (including flatworm) neuromuscular assays (Mousley et al., 2005).

No *F. hepatica* neuropeptide sequences have been published yet, but our unpublished data suggest the presence of at least 36 neuropeptide genes in the *F. hepatica* genome (Duncan Wells, Queen's University Belfast, personal communication). These ligands would facilitate deorphanisation of heterologously-expressed peptide GPCRs (S1 Table). This is essential work, as although two planarian peptide receptors have been deorphanised (Omar et al., 2007; Saberi et al., 2016), no flatworm parasite peptide GPCRs have been ligand matched. Receptor deorphanisation provides a starting point for drug discovery, by enabling development of agonists or antagonists that modulate the interaction of a GPCR with its cognate ligand. Such compounds could form the basis of ligand series for screening pipelines to support the discovery of new potential flukicides (Yoshida et al., 2012; Stockert and Devi, 2015).

Serotonin (5-hydroxytryptamine, 5-HT) is abundant throughout flatworm nervous systems, and is considered the primary flatworm excitatory neurotransmitter (Ribeiro et al., 2005). Deorphanized GPCRs activated by 5-HT have been described in turbellarians and trematodes,

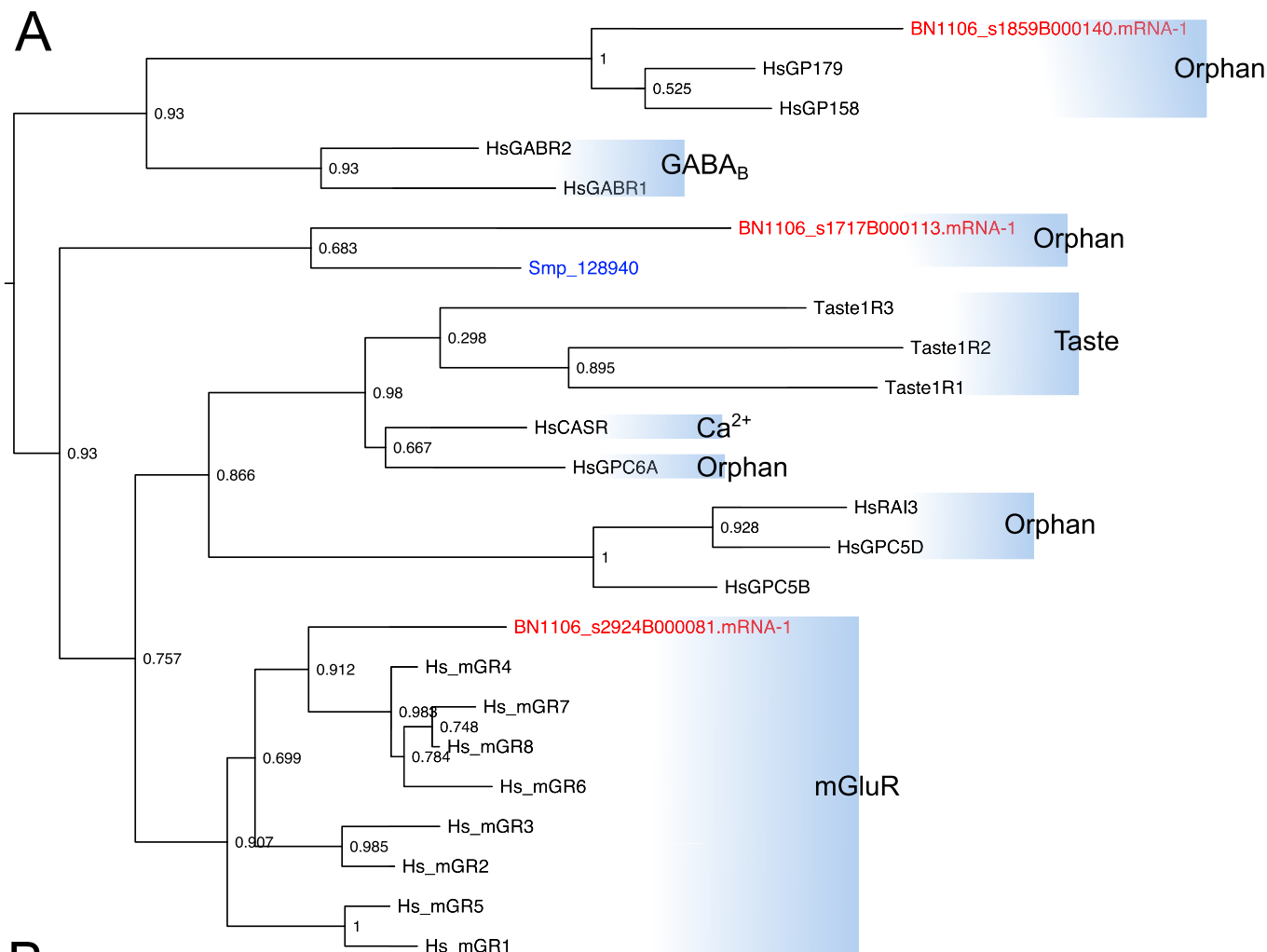




**Fig. 4. Conservation of ligand-interacting residues between 17 *Fasciola hepatica* G protein-coupled receptors (GPCRs) and structurally characterised homologues from other species.** (A) Neuropeptide F/Y receptor ligand binding residues as characterised by mutagenesis in human neuropeptide Y receptor NPY1R (Sautel et al., 1995, 1996; Berglund et al., 2002; Åkerberg et al., 2010; Fällmar et al., 2011), and conserved in *Anopheles gambiae* (Ag) and *Drosophila melanogaster* (Dm) neuropeptide F receptors (NPFR) (Vogel et al., 2013). Numbering relative to HsNPY1R. (B) Serotonin (5-hydroxytryptamine; 5HT) receptor ligand binding residues as characterised by mutagenesis in human 5HT receptor (Hs5HT1A) (Nakamura et al., 2015), and conserved in *Schistosoma mansoni* 5HTR (Patočka et al., 2014). Numbering relative to Hs5HT1A. (C) Octopamine receptor (OaR) ligand binding residues as characterised by homology modelling of the *Periplaneta americana* (Pa) (Hirashima and Huang, 2008), and mutational analysis of the *Bombyx mori* (Bm) (Huang et al., 2007) octopamine receptor ligand binding domain. Numbering relative to PaOAR, except for Y412 which is shown relative to BmOAR. (D) Acetylcholine receptor ligand binding residues as characterised by homology modelling of the *S. mansoni* G protein-coupled acetylcholine receptor (SmGAR) (MacDonald et al., 2016); numbering relative to SmGAR. In each case, only *F. hepatica* sequences displaying at least 75% identity across the stated ligand binding residues are shown. Relative positions of residues across seven transmembrane domains (TM1-7) are shown. TM diagrams are not to scale.

with an *S. mansoni* 5-HT receptor (Sm5HTR) involved in neuromuscular control (Nishimura et al., 2009; Zamanian et al., 2012; Patocka et al., 2014). Five *F. hepatica* rhodopsins (Fig. 4B) bore appreciable ( $\geq 80\%$ ) positional identity in amino acids shown to be key ligand-interacting residues in the human 5HT1A LBD (Wacker et al., 2013; Wang et al.,

2013a). Notably, these residues were also conserved in the deorphanized *S. mansoni* 5-HT receptor (Sm5HTR, Smp\_126730; Patocka et al., 2014). Three of the sequences (BN1106\_s362B000177, BN1106\_s81B000700 and BN1106\_s10B000515) also resembled Sm5HTR in our phylogenetic analysis, identifying them as likely 5-HT



**B**

MmGABA <sub>B1</sub>	S246	S269	S270	E465	D471
BN1106_s2924B000081	-	-	-	-	-
BN1106_s1717B000113	S	D	A	D	N
BN1106_s1859B000140	F	L	N	C	Y
Smp_128940	S	A	S	D	N
Smp_052660	S	T	T	E	D

Glu binding region:	COOH	αNH <sub>2</sub>	αNH <sub>2</sub>	αNH <sub>2</sub>	αNH <sub>2</sub>	αNH <sub>2</sub>	αNH <sub>2</sub>	αNH <sub>2</sub>	COOH			
mGluR1	Y74	R78	S	S165	T188	D208	Y236	E292	G293	D318	D	K409
mGluR2	R57	R	Y	S145	T168	D188	Y216	R	S	D295	R	K377
mGluR3	R64	R68	Y150	S151	T174	D194	Y222	R277	S	D301	Q306	K
mGluR4	K74	R78	G	S159	T182	D	Y	N	E	D	K	K
mGluR7	N	R	G	S	T	D	Y	N	D	D	K	K
mGluR8	K71	R75	A	S	T	D	Y	N	E	D309	K	K
BN1106_s2924B000081	-	-	-	-	-	-	-	-	-	-	-	-
BN1106_s1717B000113	I	F	I	S	D	A	Y	S	S	N	D	A
BN1106_s1859B000140	-	L	-	-	-	-	A	D	V	N	-	P
Smp_128940	N	L	Q	S	A	I	H	T	S	N	T	S
Smp_052660	R	R	Y	S	T	D	Y	G	M	D	E	K

(caption on next page)

**Fig. 5. *Fasciola hepatica* glutamate G-protein coupled receptors (GPCRs) display divergent phylogeny and ligand binding domain (LBD) composition.** (A) Maximum likelihood phylogeny containing three *F. hepatica* glutamate receptors, alongside representative receptors from the various recognised GPCR Class C subgroups (subclasses indicated by blue boxes: Ca<sup>2+</sup>, Ca<sup>2+</sup>-sensing receptor; GABA<sub>B</sub>,  $\gamma$ -aminobutyric acid type B receptors; mGluR, metabotropic glutamate receptors; Orphan, receptors with no known ligand; Taste, vertebrate taste receptors). Two previously reported *Schistosoma mansoni* glutamate receptors are also included; *F. hepatica* sequences are coloured red, *S. mansoni* are coloured blue, all others are black. Node numbers indicate statistical support as determined by approximate likelihood ratio test (aLRT). Tree was midpoint rooted. (B) Conservation of ligand-interacting residues between vertebrate GABA<sub>B</sub> and metabotropic glutamate receptors (mGluR), and *F. hepatica* class C GPCRs. Agonist-interacting residues were identified by multiple protein sequence alignment of *F. hepatica* glutamate receptors against mutationally-identified ligand interacting residues (those causing a significant reduction in receptor signalling activity), from mouse GABA<sub>B</sub> receptor (top panel), or selected human mGluR subtypes (lower panel). Identical amino acids in *F. hepatica*/*S. mansoni* GPCRs are represented by white text on black background, functionally conserved amino acids by black text on grey background. In lower panel, mutations causing a significant reduction in mGluR receptor activity are bold and numbered, with the region of the glutamate molecule bound by each residue indicated (COOH, C-terminus; NH<sub>2</sub>, N-terminus). For references see (Galvez et al., 2000; Wang et al., 2009; Wellendorph & Bräuner-Osborne, 2009; Geng et al., 2013). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

receptors. The remaining two (D915\_00277 and BN1106\_s1436B000114) appeared phylogenetically more similar to an *S. mansoni* dopamine receptor (Smp\_127310; Taman and Ribeiro, 2009). These annotations provide rational starting points for receptor deorphanization using functional genomic and/or heterologous expression tools. We found that *F. hepatica* dopamine-like receptors, identified by phylogeny (S5 Figure), displayed poor conservation (max 56% overall identity) to the human D2 LBD (Kalani et al., 2004). Due to this lack of selectivity, we did not annotate any *F. hepatica* GPCRs as dopamine receptors.

Although common in other invertebrates, octopamine has not yet been directly demonstrated as a neurotransmitter in flatworms. Evidence for its presence is indirect, based on tyramine  $\beta$ -hydroxylase (octopamine's biosynthetic enzyme) activity in cestodes and planaria (Ribeiro and Webb, 1983; Nishimura et al., 2008). Three rhodopsins (Fig. 4C) showed 100% conservation of the arthropod octopamine LBD, as determined from *Periplaneta americana* and *Bombyx mori* (Huang et al., 2007; Hirashima and Huang, 2008), with an additional four showing 88% conservation. Of these seven rhodopsins, four resolved in close phylogenetic proximity to *Drosophila* mushroom body octopamine receptors (D915\_02972), *Drosophila* octopamine beta-receptors (D915\_08505 and BN1106\_s1016B000108) (S5 Figure) or a *Drosophila* tyramine receptor (D915\_05578), denoting these as high-confidence octopamine receptors. These data provide further evidence in support of a functional role for this enigmatic classical neurotransmitter in flatworms.

Acetylcholine has species-specific impacts on flatworm neuromuscular preparations *in vitro*, with myoinhibitory effects in *Fasciola* (McVeigh and Maule, 2017). Two putative muscarinic acetylcholine receptors (mAChRs) shared highest LBD identity with a Rat M3 ACh receptor (Fig. 4D) (Kruse et al., 2012). Although these were only 67% identical to the rat sequence, the five ligand-interacting residues within their LBDs were 100% identical to those of a deorphanised *S. mansoni* mAChR, known to be involved in neuromuscular coordination (SmGAR) (MacDonald et al., 2016). These receptors (D915\_00814 and BN1106\_s1913B000092) were also the most similar to SmGAR in our phylogeny (S5 Figure) so we consider them amongst our high confidence candidates for deorphanization.

### 3.5. *F. hepatica* glutamate receptors bear divergent glutamate binding domains

At least three glutamate-like GPCRs exist in *F. hepatica* (Fig. 5A and S1 Table). All three are defined by significant BLASTp similarity (median E = 2.3e<sup>-34</sup>) to metabotropic glutamate receptors (mGluRs), and/or by the presence of InterPro GPCR family 3 (Class C) domains IPR017978, IPR000162 or IPR000337. Phylogenetic analysis of these GPCRs was performed alongside receptors representative of the various Class C subgroups (Fig. 5) (Wellendorph & Bräuner-Osborne, 2009), including Ca<sup>2+</sup>-sensing receptors,  $\gamma$ -aminobutyric acid type B (GABA<sub>B</sub>) receptors, metabotropic glutamate (mGluR) receptors, and vertebrate taste receptors; for reference we also included previously reported mGluRs from *S. mansoni* (Zamanian et al., 2011; Campos et al., 2014). One *F. hepatica* GPCR (BN1106\_s2924B000081) resolved alongside the

mGluR clade, supporting designation as an mGluR. A second *F. hepatica* glutamate receptor (BN1106\_s1717B000113) has a close *S. mansoni* orthologue (Smp\_128940), both of which reside in an orphan outgroup that is of uncertain provenance. The third *F. hepatica* glutamate receptor resides within another orphan group with human GPR158 and GPR179, two closely-related class C GPCRs expressed respectively in the human brain and retina (Orlandi et al., 2013). Although these receptors have been linked with specific disease states (Orhan et al., 2013; Patel et al., 2015), their ligands remain unknown.

Divergence within the LBD can inform the ligand selectivity of Class C receptors (Wang et al., 2009; Mitri et al., 2004; Zamanian et al., 2011). To further classify the two orphan glutamate GPCRs described above, we generated multiple sequence alignments to analyse the conservation of established agonist-interacting residues between mammalian mGluR and GABA<sub>B</sub> receptors and our *F. hepatica* GPCRs. These analyses identified no significant conservation of either mGluR or GABA<sub>B</sub> LBD residues (Fig. 6B). Fig. 6B also includes the previously reported *S. mansoni* glutamate receptors, where Smp\_052660 contained a relatively well-conserved LBD with Smp\_062660 appearing more atypical. Since all three *F. hepatica* glutamate GPCRs bear atypical LBDs with respect to both GABA<sub>B</sub> and mGluR, it remains difficult to define unequivocally their ligand selectivity on the basis of conserved motifs. Nevertheless, the lack of *in silico* evidence for *F. hepatica* GABA<sub>B</sub> GPCRs reflects the dominance of GABA<sub>A</sub>-like pharmacology, which suggests that flatworm GABA signal transduction is probably entirely mediated by ionotropic receptors (Mendonça-Silva et al., 2004; Ribeiro et al., 2005).

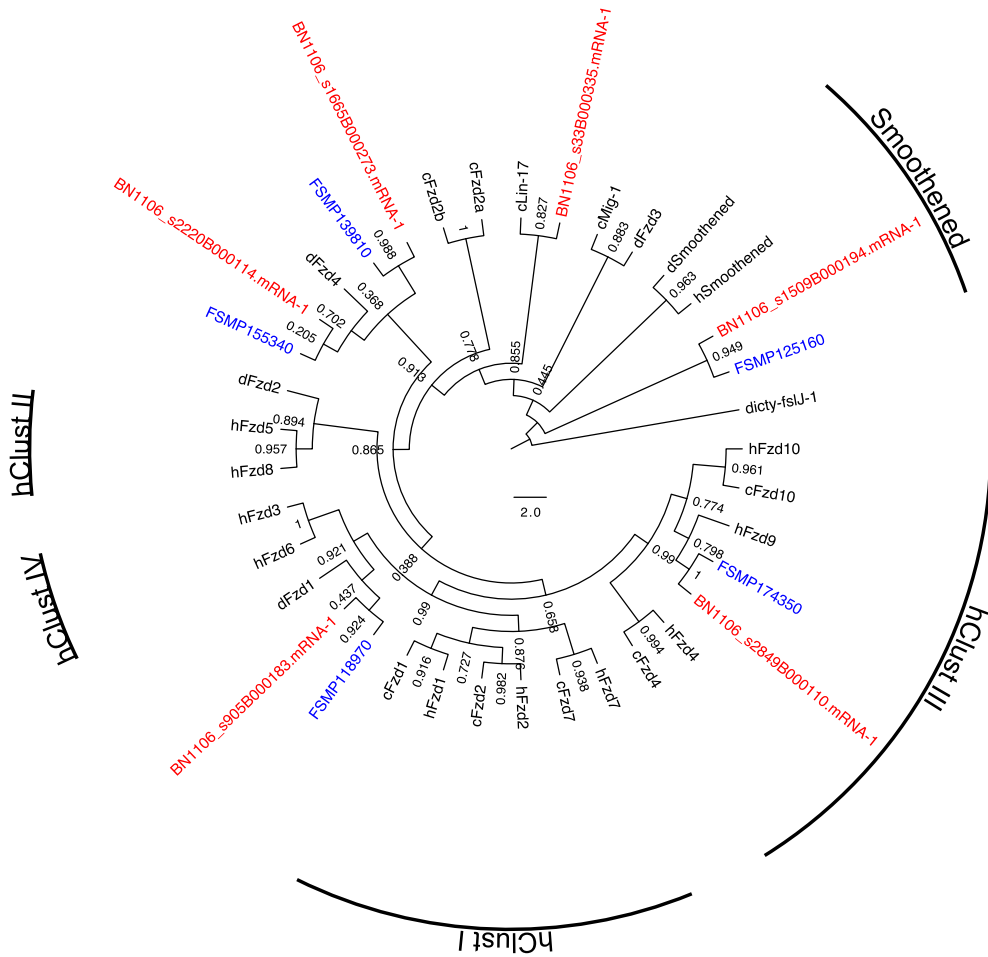
### 3.6. The Wnt binding domain is conserved in *F. hepatica* frizzled/smoothened receptors

Ten frizzled (fzd) GPCRs and a single smoothened (smo) GPCR are recognised in the human genome. In *F. hepatica* we identified five fzd-like sequences and one smo-like sequence (Fig. 6; Table S1; Table S6). All of these show high scoring similarity to annotated sequences in the ncbi nr dataset (median E = 3.8e<sup>-83</sup>), and all five fzd contain InterPro domain IPR000539, with the single smo containing domain IPR026544 (Table S1). Phylogenetic analysis of these alongside vertebrate and invertebrate receptors placed all in close proximity to existing fzd/smo groups (Fig. 6A). Four *F. hepatica* fzd had individual direct orthologues with the four known *S. mansoni* fzd GPCRs (Zamanian et al., 2011; Campos et al., 2014).

Frizzled receptors are activated by cysteine-rich glycoprotein ligands known as Wnts (Wingless and Int-1), and are involved in developmental signalling through at least three different signalling pathways. Crystallography of mouse fz8, docked with *Xenopus* wnt8, identified 14 amino acids within the fz8 CRD that make contact with the Wnt8 ligand (Janda et al., 2012). Positional conservation of these residues is apparent when fz8 is aligned with the five *F. hepatica* fzd sequences (Fig. 6B; S6 Table), suggesting conservation of the wnt-frizzled interaction between liver fluke and vertebrates.

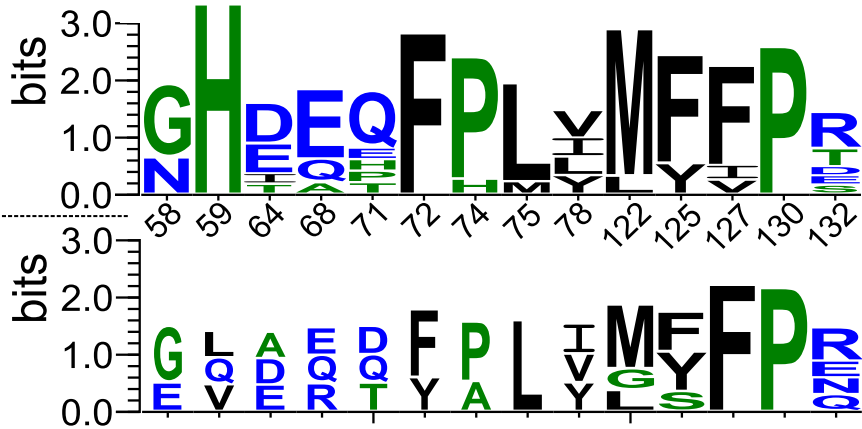
Two Wnt ligands have been described in *S. mansoni* (Li et al., 2010; Ta et al., 2015); our BLAST searches identified at least three Wnt-like sequences in the *F. hepatica* genome (BN1106\_s198B000330.mRNA-1,

A



B

Ligand-contacting amino acids  
*M. musculus* fz1-10,  
numbering relative to fz8



C

Amino acid position relative to Dro-wnt-1

	93	104	146	154	156	185	233	235	242	247	397	413	423	427	428	443	445	450	451	455	458	467
Dro-Wnt-1	ECQ	NCS	ACS	CTC	GCS	CKC	SCT	TCW	FCE	QCN	CGLMCC	CAC	CCEVKCKLC	TCL								
BN1106_s737B000430	---	---	ACS	CDC	GCS	CKC	ACS	TCW	YCR	LCN	CNRLCC	CDC	CCEVRCKVC	RCK								
BN1106_s198B000330	---	---	ACS	CGC	GCS	CKC	SCE	TCW	FCH	RCV	CEYLCC	CNC	CCKVICQTC	TCN								
BN1106_s1256B000163	ECQ	NCT	SCK	CGC	GCG	CRP	L--	---	YCQ	ECL	CVHLCC	CRC	CCRVECQTC	VCN								

(caption on next page)



**Fig. 6. Frizzled/smoothened seven transmembrane receptors and wnt ligands in *Fasciola hepatica*.** (A) Maximum likelihood phylogeny containing six *F. hepatica* frizzled/smoothened receptors, alongside those from *Schistosoma mansoni*, *Drosophila melanogaster*, *Caenorhabditis elegans* and *Homo sapiens* (identified by FSMP, d, c and h, respectively). *F. hepatica* sequences are coloured red, *S. mansoni* are coloured blue, all other species are coloured black. Radial labels indicate human frizzled clusters (hClust) I-IV, and the smoothened clade. Node numbers indicate statistical support as determined by approximate likelihood ratio test (aLRT). Tree was rooted against a *Dictyostelium* frizzled sequence (dicty-fslJ-1). Tree composition adapted from Zamanian et al. (2011). (B) WebLogo comparison of ligand interacting residues between mouse fz1-10 (top panel) and *F. hepatica* frizzled receptors. Numbering in top panel x-axis is relative to mouse fz8 (Janda et al., 2012). (C) Three wnt-like sequences exist in *F. hepatica*. Shading indicates positions of 22 characteristic Cys residues, positions numbered relative to *D. melanogaster* wnt-1 (Dro-wnt-1). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

BN1106\_s1256B000163.mRNA-1, BN1106\_s737B000430.mRNA-1; Fig. 6C). These showed conservation of the 23 cysteine residues that are diagnostic of Wnt glycoproteins (Willert and Nusse, 2012). Norrin, a non-Wnt protein ligand, can also activate Fz4, and the canonical  $\beta$ -catenin pathway. The amino acids involved in norrin binding to the fz4 CRD have also been determined (Smallwood et al., 2007), but we did not observe conservation of these in any of the *F. hepatica* fzd. Similarly, BLASTp searches of human norrin (Uniprot Q00604) against the *F. hepatica* genome did not return significant hits, suggesting that the norrin-fz signalling axis may not function in liver fluke. Smoothened receptors are structurally similar to frizzleds, but operate in a ligand-independent fashion within hedgehog signalling pathways that control several developmental processes (Ayers and Thérond, 2010). Model organism genomes typically contain only one smoothened gene (SMO); this was the case in *S. mansoni* and *S. mediterranea* (Zamanian et al., 2011), and here we have identified a single *F. hepatica* smoothened (BN1106\_s1509B000194; Fig. 6; Table S1).

Fzd/smo GPCRs are involved broadly in the control of cellular development. Our discovery of fzd/smo GPCRs, and their Wnt ligands, in *F. hepatica* opens avenues towards probing molecular aspects of development and differentiation in the putative stem cells/neoblasts of liver fluke (McCusker et al., 2016). Neoblasts are the cells that impart the regenerative capacity of free-living turbellarian flatworms (Gehrke and Srivastava, 2016), and neoblast-like cells also represent the only proliferating cells in several parasitic species (Collins et al., 2013; Wang et al., 2013b; Kozioł et al., 2014). Therefore, these cells are important in understanding fundamental fluke biology and represent potential repositories of unique anthelmintic targets, capable of inhibiting worm growth or development. The presence of both receptor and ligand sequences will permit functional genomic dissection of Wnt-Frizzled ligand-receptor signalling networks, aimed at elucidating their roles in the development and differentiation of liver fluke neoblast-like cells. These FhGPCRs will enable comparisons between the biology of parasitic and free-living flatworms, where Wnt signalling is known to be essential for anterior-posterior polarity in regenerating planaria (Gurley et al., 2008; Petersen and Reddien, 2008).

### 3.7. Class B (adhesion and secretin) receptors

Class B receptors incorporate both adhesions and secretins. Adhesions are characterised by a long N-terminal extracellular domain (ECD) that includes several functional motifs. These ECDs are autoproteolytically cleaved into two subunits that subsequently reassemble into a functional dimer (Lagerström and Schiöth, 2008). We identified two adhesion sequences in the *F. hepatica* genome (S7 Figure, S1 Table), both of which (scaffold181\_78723–79604, and BN1106\_s537B000355) contained GPCR class B InterPro domain IPR000832 and displayed closest BLASTp similarity ( $E = 5.6e^{-7}$ ) to latrophilin-like receptors. These data suggest that both are adhesions, rather than secretins. We also identified a single secretin-like sequence in our 4-9TM dataset (BN1106\_s1217B000278.mRNA-1), which also contained GPCR class B InterPro domain IPR000832, but showed closest BLASTp similarity to a pigment dispersal factor (PDF) receptor ( $E = 9e^{-46}$ ). Phylogenetic analysis of these receptors alongside human Class B receptors supports the definition of scaffold181\_78723–79604 and BN1106\_s537B000355 as adhesions, with BN1106\_s1217B000278 appearing within the clade of secretin receptors.

Deorphanization of a handful of adhesions matches them with a

complex assortment of ligands including collagen, transmembrane glycoproteins, complement proteins and FMRFamide-like neuropeptides (Langenhan et al., 2013). This assortment of potential ligands, and their expression in almost every organ system has led to the proposal of a diverse range of functions for vertebrate adhesions. The *F. hepatica* adhesion complement of two GPCRs is greatly reduced compared to the 33 receptors known in humans; in other flatworms 14, 4 and 1 adhesions have been described in *S. mediterranea*, *E. multilocularis* and *S. mansoni*, respectively (Zamanian et al., 2011; Tsai et al., 2013; Saberi et al., 2016). Functional characterisation will be a challenging task given the wide range of possible functions to be assayed; an appealing starting point would be to investigate roles in neoblast motility prior to differentiation, given that mammalian adhesion GPCRs are involved in the control of cellular migration (Langenhan et al., 2013).

### 3.8. Developmental expression

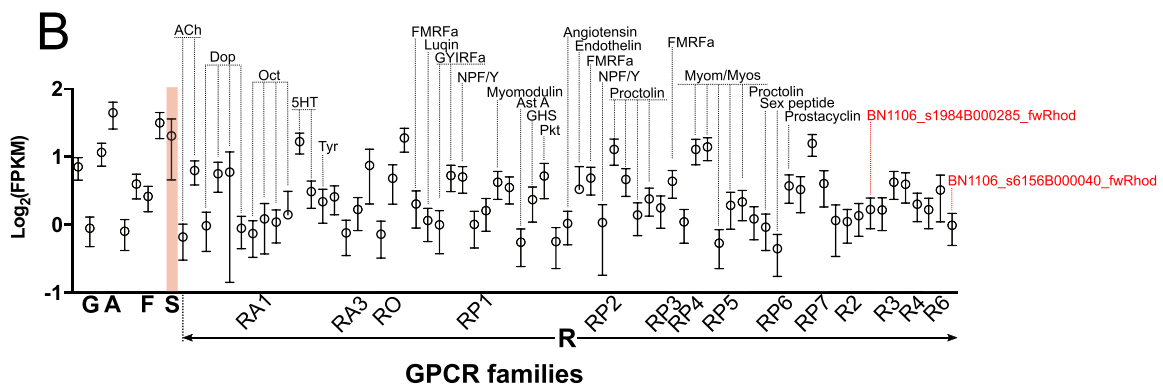
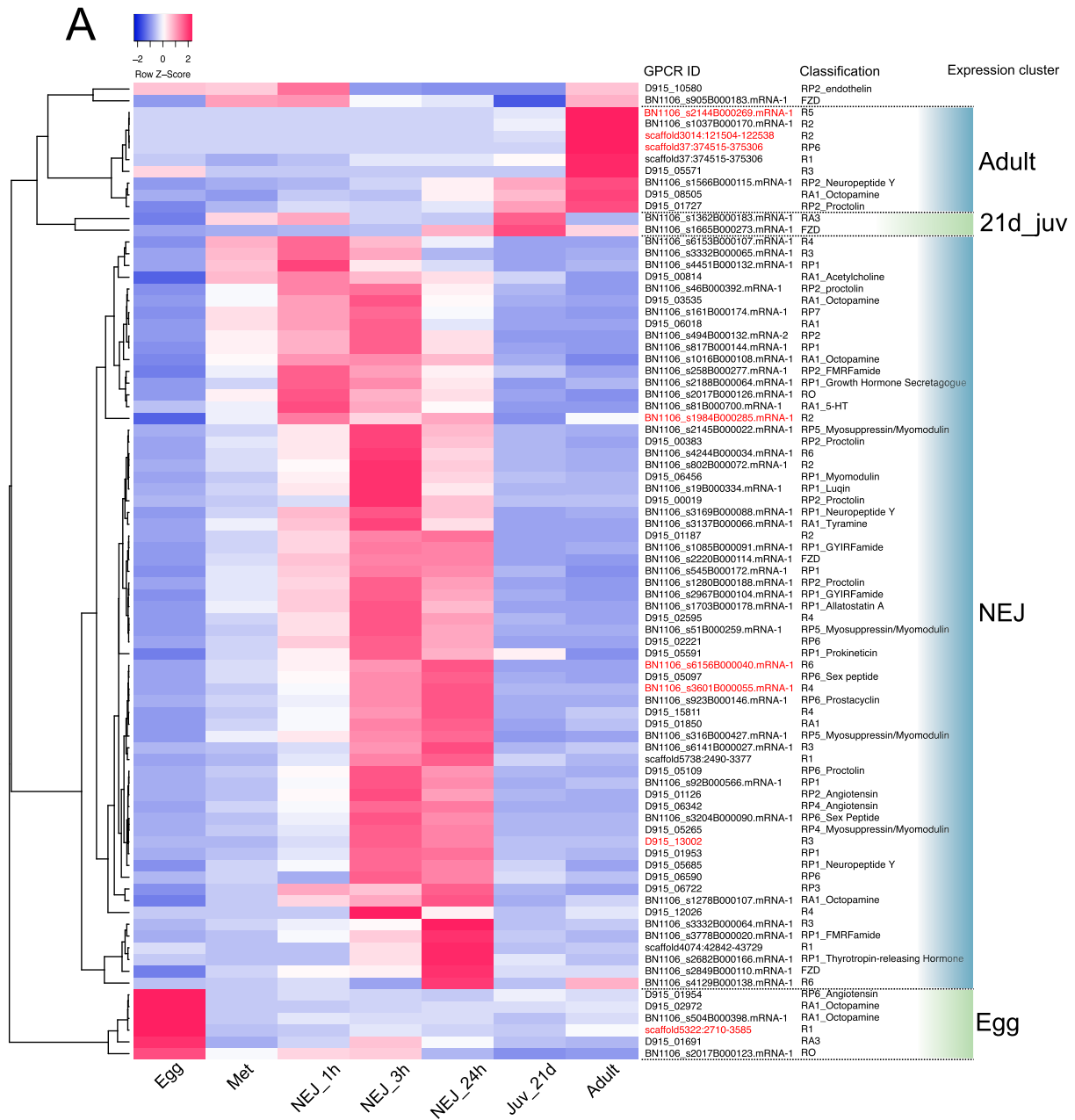
Using RNA-Seq methods, we were able to confirm the expression of 101 GPCRs across libraries representing several *F. hepatica* life-stages. These datasets included publicly available reads from individual developmental stages (Cwiklinski et al., 2015), and a transcriptome that we generated in-house for 21-day liver stage ex-vivo juveniles (juv2). Since these datasets were generated independently and clearly display distinct sequence diversities, we avoided any further direct comparisons between Cwiklinski juv1 and our juv2 datasets. Each dataset is analysed separately, below.

Fig. 7A illustrates detection of 83 GPCRs across Cwiklinski's developmentally staged RNA-Seq datasets. These comprised four FZD, thirteen aminergic rhodopsins, two opsins, 41 peptidergic rhodopsins, and 23 orphan rhodopsins. The latter included nine fwRhods. Clustering within Fig. 7A's expression heatmap shows clear developmental regulation of GPCR expression, outlining nine GPCRs with relatively higher expression in adults, two with higher expression in 21d juveniles, 64 GPCRs preferentially expressed in either 1 h, 3 h or 24 h NEJs, and six receptors expressed most highly in eggs. GPCR classes appear to be randomly distributed across these expression clusters, giving little opportunity to infer function from expression. Adult-expressed GPCRs include five orphan fwRhods, three peptide receptors including a putative NPF/Y receptor, and a predicted octopamine-gated aminergic rhodopsin. The majority of expressed GPCRs occurred in the NEJ-focused expression cluster. Given data implicating GPCRs in motility, growth/development and sensory perception (McVeigh et al., 2012), it is no surprise to find high levels of GPCR expression in the NEJs, which must navigate and burrow their way from the gut lumen into the liver parenchyma, while also sustaining rapid growth from the start of the infection process. The high expression in these stages, of receptors that we predict to be activated by myomodulators such as ACh, FMRFamide, GYIRFamide, myomodulin, myosuppressin and 5-HT, provide tentative support for these predictions. The focused expression of six GPCRs in eggs suggests potential roles in the control of cellular proliferation and fate determination processes that occur during embryonation. This complement did not include frizzled or adhesion GPCRs that are traditionally implicated in the control of development, instead consisting of rhodopsins (including an angiotensin-like peptide receptor, two octopamine-like amine receptors, one opsin receptor and one fwRhod receptor).

Focusing on the pathogenic 21-day juvenile stage, we detected 76 GPCRs in our juv2 datasets, and 29 in the corresponding juv1 samples

from Cwiklinski's dataset (Fig. 7B). Our juv2 dataset included three glutamate, one adhesion, four frizzled, one smoothed, and 67 rhodopsins. The identity of the receptors expressed here again attest to the

key role of neuromuscular co-ordination in this highly motile life stage, which must penetrate and migrate through the liver parenchyma *en route* to the bile ducts. Amongst the receptors expressed in this stage



(caption on next page)

**Fig. 7. Expression profiling of 101 G protein-coupled receptors (GPCRs) in *Fasciola hepatica* life stages.** (A) Expression heatmap generated from  $\log_2$  FPKM values of 83 GPCRs identified from developmentally staged RNA-seq libraries. Life stages are represented in columns (Egg; Met, metacercariae; NEJ\_1 h, newly-excysted juvenile collected 1 h post-excystment; NEJ\_3 h, NEJ collected 3 h post-excystment; NEJ\_24 h, NEJ collected 24 h post-excystment; Juv\_21d, liver stage juvenile parasites collected from murine livers 21 days following oral administration of metacercariae; Adult, adult parasites collected from the bile ducts of bovine livers). Rows indicate individual GPCRs, as denoted by the ID and phylogeny columns. The latter indicates receptor classification and predicted ligand where available (see S1 Table). Expression cluster column indicates clusters of GPCRs with highest expression focused in particular life stages. (B) Detection of 76 GPCRs in Illumina RNA-Seq libraries generated from *F. hepatica* 21-day liver-stage juveniles, recovered *ex vivo* from rat infections. Data show expression of three glutamate (G), one adhesion (A), four frizzled (F), one smoothed (S) and 67 rhodopsin (R) GPCRs. The rhodopsins include representatives of amine (RA1, RA3), opsin (RO), peptide (RP1-7), and orphan (R2,3,4,6). Data points (each at  $n = 3$ ) represent mean  $\log_2$  FPKM  $\pm$  95% confidence intervals, as calculated by cuffdiff. In both panels, flatworm rhodopsins (fwRhods) are marked in red text. ACh, acetylcholine; AstA, Allatostatin A; Dop, dopamine; FMRFa, FMRFamide; GHS, growth hormone secretagogue; GYIRFa, GYIRFamide; Myom, myomodulin; Myos, myosuppressin; NPF/Y, neuropeptide F/Y; Oct, octopamine; Pkt, prokineticin; Tyr, tyramine; 5HT, 5-hydroxytryptamine. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

and thought to have a role in neuromuscular function are several activated by classical neurotransmitters including ACh, dopamine and 5-HT. The peptide receptors include some with phylogenetic similarity to receptors for myoactive flatworm peptides (FMRFamide, GYIRFamide, NPF), as well as receptors from other invertebrates activated by peptide ligands known to have excitatory effects on flatworms (allatostatin A, myomodulin, proctolin) (Mousley et al., 2004). The presence of highly expressed GPCRs with probable neuromuscular functions in liver stage juveniles, points to the importance of studying these receptors with a view to flukicide discovery. The damage caused by migrating juvenile fluke requires that new flukicides are effective against this stage. The neuromuscular GPCRs expressed in migrating juveniles provide compelling targets for new drugs.

#### 4. Conclusions

GPCRs are targets for 33% of human pharmaceuticals (Santos et al., 2017), illustrating the appeal of GPCRs as putative anthelmintic targets. This study provides the first description of the *F. hepatica* GPCR complement permitting consideration of a GPCR target-based screening approach to flukicide discovery. To facilitate the deorphanization experiments that will precede compound screening efforts, we have described a set of high confidence rhodopsin ligand-receptor pairs. We identified these GPCRs, including receptors for ACh, octopamine, 5HT and NPF/Y, through phylogenetic comparison with existing deorphanised receptors and positional conservation of ligand-interacting residues within ligand binding domains. Our additional descriptions of new members of existing flatworm-specific rhodopsin groups in *Fasciola* and other species support the potential for synthetic ligands to be parasite-selective anthelmintics.

#### Appendix A. Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.jipddr.2018.01.001>.

#### References

Abunna, F., Asfaw, L., Megersa, B., Regassa, A., 2010. Bovine fasciolosis: coprological, abattoir survey and its economic impact due to liver condemnation at Soddo municipal abattoir, Southern Ethiopia. *Trop. Anim. Health Prod.* 42, 289–292.

Åkerberg, H., Fällmar, H., Sjödin, P., Boukharta, L., Gutiérrez-de-Terán, H., Lundell, I., Mohell, N., Larhammar, D., 2010. Mutagenesis of human neuropeptide Y/peptide YY receptor Y2 reveals additional differences to Y1 in interactions with highly conserved ligand positions. *Regul. Pept.* 163, 120–910.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410.

Ayers, K.L., Théron, P.P., 2010. Evaluating smoothed as a G-protein-coupled receptor for hedgehog signalling. *Trends Cell Biol.* 20, 287–298.

Babicki, S., Arndt, D., Marcu, A., Liang, Y., Grant, J.R., Maciejewski, A., Wishart, D.S., 2016. Heatmapper: web-enabled heat mapping for all. *Nucleic Acids Res.* 44 (W1), W147–W153.

Berglund, M.M., Fredriksson, R., Salaneck, E., Larhammar, D., 2002. Reciprocal mutations of neuropeptide Y receptor Y2 in human and chicken identify amino acids important for antagonist binding. *FEBS Lett.* 518, 5–9.

Buxton, S.K., Neveu, C., Robertson, A.P., Martin, R.J., 2011. On the mode of action of emodepsid: slow effects on membrane potential and voltage-activated currents in *Ascaris suum*. *Br. J. Pharmacol.* 164, 453–470.

Cabada, M.M., Lopez, M., Cruz, M., Delgado, J.R., Hill, V., White Jr., C., 2016. Treatment

failure after multiple courses of triclabendazole among patients with fasciolosis in Cusco, Peru: a case series. *PLoS Neglected Trop. Dis.* 10, e0004361.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L., 2008. BLAST+: architecture and applications. *BMC Bioinf.* 10, 421.

Campos, T.D., Young, N.D., Korhonen, P.K., Hall, R.S., Lonie, A., Gasser, R.B., 2014. Identification of G protein-coupled receptors in *Schistosoma haematobium* and *S. mansoni* by comparative genomics. *Parasites Vectors* 7, 242. <http://dx.doi.org/10.1186/1756-3305-7-242>.

Collins 3rd, J.J., Hou, X., Romanova, E.V., Lambrus, B.G., Miller, C.M., Saberi, A., Sweedler, J.V., Newmark, P.A., 2010. Genome-wide analyses reveal a role for peptide hormones in planarian germline development. *PLoS Biol.* 8 (10), e1000509.

Collins 3rd, J.J., Wang, B., Lambrus, B.G., Tharp, M.E., Iyer, H., Newmark, P.A., 2013. Adult somatic stem cells in the human parasite *Schistosoma mansoni*. *Nature* 494, 476–479.

Crooks, G.E., Hon, G., Chandonia, J.M., Brenner, S.E., 2004. WebLogo: a sequence logo generator. *Genome Res.* 14, 1188–1190.

Cwiklinski, K., Dalton, J.P., Dufresne, P.J., La Course, J., Williams, D.J.L., Hodgkinson, J., Paterson, S., 2015. The *Fasciola hepatica* genome: gene duplication and polymorphism reveals adaptation to the host environment and the capacity for rapid evolution. *Genome Biol.* 16, 71. <http://dx.doi.org/10.1186/s13059-015-0632-2>.

Dell'Oca, N., Basika, T., Corvo, I., Castillo, E., Brindley, P.J., Rinaldi, G., Tort, J.F., 2014. RNA interference in *Fasciola hepatica* newly excysted juveniles: long dsRNA induces more persistent silencing than siRNA. *Mol. Biochem. Parasitol.* 197, 28–35.

Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J.F., Guindon, S., Lefort, V., Lescot, M., Claverie, J.M., Gascuel, O., 2008. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.* 36, W465–W469.

Fällmar, H., Åkerberg, H., Gutiérrez-de-Terán, H., Lundell, I., Mohell, N., Larhammar, D., 2011. Identification of positions in the human neuropeptide Y/peptide YY receptor Y2 that contribute to pharmacological differences between receptor subtypes. *Neuropeptides* 45, 293–300.

Finn, R.D., Clements, J., Arndt, W., Miller, B.L., Wheeler, T.J., Schreiber, F., Bateman, A., Eddy, S.R., 2015. HMMER web server: 2015 update. *Nucleic Acids Res.* 43 (1), W30–W38.

Fredriksson, R., Lagerström, M.C., Lundin, L.G., Schiöth, H.B., 2003. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Mol. Pharmacol.* 63, 1256–1272.

Galvez, T., Prezeau, L., Milioti, G., Franek, M., Joly, C., Froestl, W., Bettler, B., Bertrand, H.G., Blahos, J., Pin, J.P., 2000. Mapping the agonist binding site of GABA<sub>B</sub> type 1 subunit sheds light on the activation process of GABA<sub>B</sub> receptors. *J. Biol. Chem.* 275, 41166–41174.

Gehrke, A.R., Srivastava, M., 2016. Neoblasts and the evolution of whole-body regeneration. *Curr. Opin. Genet. Dev.* 40, 131–137.

Geng, Y., Bush, M., Mosyak, L., Wang, F., Fan, Q.R., 2013. Structural mechanism of ligand activation in human GABA(B) receptor. *Nature* 504 (7479), 254–259.

Gurley, K.A., Rink, J.C., Sanchez-Alvarado, A., 2008. Beta-catenin defines head versus tail identity during planarian regeneration and homeostasis. *Science* 319, 323–327.

Habarugira, G., Mbasinga, G., Mushonga, B., Teedzai, C., Kandiwa, E., Ojok, L., 2016. Pathological findings of condemned bovine liver specimens and associated economic loss at Nyabugogo abattoir, Kigali, Rwanda. *Acta Trop.* 164, 27–32.

Harder, A., Schmitt-Wrede, H.P., Krücken, J., Marinovski, P., Wunderlich, F., Willson, J., Amliwala, K., Holden-Dye, L., Walker, R., 2003. Cyclooctadepsipeptides – an anthelmintically active class of compounds exhibiting a novel mode of action. *Int. J. Antimicrob. Agents* 22, 318–331.

Hirashima, A., Huang, H., 2008. Homology modeling, agonist binding site identification, and docking in octopamine receptor of *Periplaneta americana*. *Comput. Biol. Chem.* 32, 185–190.

Hotez, P.J., Brindley, P.J., Bethony, J.M., King, C.H., Pearce, E.J., Jacobson, J., 2008. Helminth infections: the great neglected tropical diseases. *J. Clin. Invest.* 118, 1311–1321.

Howe, K.L., Bolt, B.J., Shafie, M., Kersey, P., Berriman, M., 2017. WormBase ParaSite – a comprehensive resource for helminth genomics. *Mol. Biochem. Parasitol.* 215, 2–10.

Howell, A., Baylis, M., Smith, R., Pinchbeck, G., Williams, D., 2015. Epidemiology and impact of *Fasciola hepatica* exposure in high-yielding dairy herds. *Prev. Vet. Med.* 121, 41–48.

Huang, J., Hamasaki, T., Ozoe, F., Ohta, H., Enomoto, K., Kataoka, H., Sawa, Y., Hirota, A., Ozoe, Y., 2007. Identification of critical structural determinants responsible for octopamine binding to the alpha-adrenergic-like *Bombyx mori* octopamine receptor. *Biochemistry* 46, 5896–5903.

Janda, C.Y., Waghay, D., Levin, A.M., Thomas, C., Garcia, K.C., 2012. Structural basis of wnt recognition by frizzled. *Science* 337, 59–64.



- Jiang, L., Schlesinger, F., Davis, C.A., Zhang, Y., Li, R., Salit, M., Gingeras, T.R., Oliver, B., 2011. Synthetic spike-in standards for RNA-Seq experiments. *Genome Res.* 21, 1543–1551.
- Jones, P., Binns, D., Chang, H.Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A.F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S.Y., Lopez, R., Hunter, S., 2014. InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30 (9), 1236–1240.
- Kalani, M.Y., Vaidehi, N., Hall, S.E., Trabanino, R.J., Freddolino, P.L., Kalani, M.A., Floriano, W.B., Kam, V.W., Goddard 3rd, W.A., 2004. The predicted 3D structure of the human D2 dopamine receptor and the binding site and binding affinities for agonists and antagonists. *Proc. Natl. Acad. Sci. U.S.A.* 101, 3815–3820.
- Katoh, K., Rozewicki, J., Yamada, K.D., 2017. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Briefings Bioinf.* <http://dx.doi.org/10.1093/bib/bbx108>.
- Kelley, J.M., Elliott, T.P., Beddow, T., Anderson, G., Skuce, P., Spithill, T.W., 2016. Current threat of triclabendazole resistance in *Fasciola hepatica*. *Trends Parasitol.* 32, 458–469.
- Khalturin, K., Hemmrich, G., Fraune, S., Augustin, R., Bosch, T.C.G., 2009. More than just orphans: are taxonomically-restricted genes important in evolution? *Trends Genet.* 25, 404–413.
- Kozioł, U., Rauschendorfer, T., Zanon Rodriguez, L., Krohne, G., Brehm, K., 2014. The unique stem cell system of the immortal larvae of the human parasite *Echinococcus multilocularis*. *EvoDevo* 5 (1), 10. <http://dx.doi.org/10.1186/2041-9139-5-10>.
- Kozioł, U., Kozioł, M., Preza, M., Costáble, A., Brehm, K., Castillo, E., 2016. De novo discovery of neuropeptides in the genomes of parasitic flatworms using a novel comparative approach. *Int. J. Parasitol.* 46, 709–721.
- Kruse, A.C., Hu, J., Pan, A.C., Arlow, D.H., Rosenbaum, D.M., Rosemond, E., Green, H.F., Liu, T., Chae, P.S., Dror, R.O., Shaw, D.E., Weis, W.L., Wess, J., Kobilka, B.K., 2012. Structure and dynamics of the M3 muscarinic acetylcholine receptor. *Nature* 482, 552–556.
- Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874.
- Lagerström, M.C., Schiöth, H.B., 2008. Structural diversity of G protein-coupled receptors and significance for drug discovery. *Nat. Rev. Drug Discov.* 7, 339–357.
- Langenhan, T., Aust, G., Hamann, J., 2013. Sticky signaling – adhesion class G protein-coupled receptors take the stage. *Sci. Signal.* 6 (276). <http://dx.doi.org/10.1126/scisignal.2003825>. re3.
- Langmead, B., Trapnell, C., Pop, M., Salzberg, S.L., 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10 (3), R25.
- Li, H.F., Wang, X.B., Jin, Y.P., Xia, Y.X., Feng, X.G., Yang, J.M., Qi, X.Y., Yuan, C.X., Lin, J.J., 2010. Wnt4, the first member of the Wnt family identified in *Schistosoma japonicum*, regulates worm development by the canonical pathway. *Parasitol. Res.* 107, 795–805.
- Lowery, D.E., Geary, T.G., Kubiak, T.M., Larsen, M.J., 2003. G protein-coupled receptor-like receptors and modulators thereof. *Pharmacia and Upjohn Company, United States*. Patent US 6632621 B1.
- MacDonald, K., Kimber, M.J., Day, T.A., Ribeiro, P., 2016. A constitutively active G protein-coupled acetylcholine receptor regulates motility of larval *Schistosoma mansoni*. *Mol. Biochem. Parasitol.* 202, 29–37.
- Martin, R.J., Robertson, A.P., 2010. Control of nematode parasites with agents acting on neuro-musculature systems: lessons for neuropeptide ligand discovery. *Adv. Exp. Med. Biol.* 692, 138–154.
- Martin-Duran, J.M., Ryan, J.F., Vellutini, B.C., Pang, K., Hejnal, A., 2017. Increased taxon sampling reveals thousands of hidden orthologs in flatworms. *Genome Res.* 27, 1263–1272.
- McCammick, E.M., McVeigh, P., McCusker, P., Timson, D.J., Morphew, R.M., Brophy, P.M., Marks, N.J., Mousley, A., Maule, A.G., 2016. Calmodulin disruption impacts growth and motility in juvenile liver fluke. *Parasites Vectors* 9, 46.
- McCusker, P., McVeigh, P., Rathinasamy, V., Toet, H., McCammick, E., O'Connor, A., Marks, N.J., Mousley, A., Brennan, G.P., Halton, D.W., Spithill, T.W., 2016. Stimulating neoblast-like cell proliferation in juvenile *Fasciola hepatica* supports growth and progression towards the adult phenotype in vitro. *PLoS Neglected Trop. Dis.* 10 (9), e0004994.
- McGonigle, L., Mousley, A., Marks, N.J., Brennan, G.P., Dalton, J.P., Spithill, T.W., Day, T.A., Maule, A.G., 2008. The silencing of cysteine proteases in *Fasciola hepatica* newly excysted juveniles using RNA interference reduced gut penetration. *Int. J. Parasitol.* 38, 149–155.
- McNulty, S.N., Tort, J.F., Rinaldi, G., Fischer, K., Rosa, B.A., Smircich, P., Fontenla, S., Choi, Y.J., Tyagi, R., Hallsworth-Pepin, K., Mann, V.H., Kammili, L., Latham, P.S., Dell'Oca, N., Dominguez, F., Carmona, C., Fischer, P.U., Brindley, P.J., Mitreva, M., 2017. Genomes of *Fasciola hepatica* from the Americas reveal colonization with *Neorickettsia* endobacteria related to the agents of Potomac horse and human senetsu fevers. *PLoS Genet.* 13, e1006537.
- McVeigh, P., Maule, A.G., 2017. Flatworm Neurobiology in the Postgenomic Era. Oxford Handbook of Invertebrate Neurobiology. Oxford University Press.
- McVeigh, P., Mair, G.R., Atkinson, L., Ladurner, P., Zamanian, M., Novozhilova, E., Marks, N.J., Day, T.A., Maule, A.G., 2009. Discovery of multiple neuropeptide families in the phylum Platyhelminthes. *Int. J. Parasitol.* 39, 1243–1252.
- McVeigh, P., Atkinson, L., Marks, N.J., Mousley, A., Dalzell, J.J., Sluder, A., Hammerland, L., Maule, A.G., 2012. Parasite neuropeptide biology: seeding rational drug target selection? *Int. J. Parasitol.: Drugs Drug. Res.* 2, 76–91.
- McVeigh, P., McCammick, E.M., McCusker, P., Morphew, R.M., Mousley, A., Abidi, A., Saifullah, K.M., Muthusamy, R., Gopalakrishnan, R., Spithill, T.W., Dalton, J.P., Brophy, P.M., Marks, N.J., Maule, A.G., 2014. RNAi dynamics in juvenile *Fasciola* spp. liver flukes reveals the persistence of gene silencing in vitro. *PLoS Neglected Trop. Dis.* 8 (9), e3185.
- Mendonça-Silva, D.L., Gardino, P.F., Kubrusly, R.C., De Mello, F.G., Noël, F., 2004. Characterization of a GABAergic neurotransmission in adult *Schistosoma mansoni*. *Parasitology* 129, 137–146.
- Mitri, C., Parmentier, M.L., Pin, J.P., Bockaert, J., Grau, Y., 2004. Divergent evolution in metabotropic glutamate receptors. A new receptor activated by an endogenous ligand different from glutamate in insects. *J. Biol. Chem.* 279, 9313–9320.
- Mousley, A., Marks, N.J., Halton, D.W., Geary, T.G., Thompson, D.P., Maule, A.G., 2004. Arthropod FMRFamide-related peptides modulate muscle activity in helminths. *Int. J. Parasitol.* 34, 755–768.
- Mousley, A., Moffett, C.L., Duve, H., Thorpe, A., Halton, D.W., Geary, T.G., Thompson, D.P., Maule, A.G., Marks, N.J., 2005. Expression and bioactivity of allatostatin-like neuropeptides in helminths. *Int. J. Parasitol.* 35, 1557–1567.
- Nakamura, Y., Ishii, J., Kondo, A., 2015. Applications of yeast-based signaling sensor for characterization of antagonist and analysis of site-directed mutants of the human serotonin 1A receptor. *Biotechnol. Bioeng.* 112 (9), 1906–1915.
- Nishimura, K., Kitamura, Y., Inoue, T., Umeson, Y., Yoshimoto, K., Taniguchi, T., Agata, K., 2008. Characterization of tyramine beta-hydroxylase in planarian *Dugesia japonica*: cloning and expression. *Neurochem. Int.* 53, 184–192.
- Nishimura, K., Unemura, K., Tsushima, J., Yamauchi, Y., Taniguchi, T., Kaneko, S., Agata, K., Kitamura, Y., 2009. Identification of a novel planarian G-protein-coupled receptor that responds to serotonin in *Xenopus laevis* oocytes. *Biol. Pharm. Bull.* 32, 1672–1677.
- Omar, H.H., Humphries, J.E., Larsen, M.J., Kubiak, T.M., Geary, T.G., Maule, A.G., Kimber, M.J., Day, T.A., 2007. Identification of a platyhelminth neuropeptide receptor. *Int. J. Parasitol.* 37, 725–733.
- Orhan, E., Prezeau, L., El Shamieh, S., Bujakowska, K.M., Michiels, C., Zagar, Y., Vol, C., Bhattacharya, S.S., Sahel, J.A., Sennlaub, F., Audo, I., Zeitz, C., 2013. Further insights into GPR179: expression, localization and associated pathogenic mechanisms leading to complete congenital stationary night blindness. *Invest. Ophthalmol. Vis. Sci.* 54, 8041–8050.
- Orlandi, C., Masuho, I., Posokhova, E., Cao, Y., Ray, T., Hasan, N., 2013. GPR158 and GPR179: a subfamily of orphan GPCRs as a new class of signaling modulators. *Faseb. J.* 27 (1) Supplement 1095.2.
- Patel, N., Itakura, T., Jeong, S., Liao, C.P., Roy-Burman, P., Zandi, E., 2015. Expression and functional role of orphan receptor GPR158 in prostate cancer growth and progression. *PLoS One* 10 (2), e0117758.
- Patocka, N., Sharma, N., Rashid, M., Ribeiro, P., 2014. Serotonin signaling in *Schistosoma mansoni*: a serotonin-activated G protein-coupled receptor controls parasite movement. *PLoS Pathog.* 10 (1), e1003878.
- Pearson, W.R., 2013. An introduction to sequence similarity (“homology”) searching. *Curr. Protoc. Bioinformatics*. <http://dx.doi.org/10.1002/0471250953.bi0301s42>. (Chapter 3)Unit3.1.
- Petersen, C.P., Reddien, P.W., 2008. Smed-betacatenin-1 is required for anteroposterior blastema polarity in planarian regeneration. *Science* 319, 327–330.
- Ribeiro, P., Patocka, N., 2013. Neurotransmitter transporters in schistosomes: structure, function and prospects for drug discovery. *Parasitol. Int.* 62, 629–638.
- Ribeiro, P., Webb, R.A., 1983. The occurrence and synthesis of octopamine and catecholamines in the cestode *Hymenolepis diminuta*. *Mol. Biochem. Parasitol.* 7, 53–62.
- Ribeiro, P., El-Shehaby, F., Patocka, N., 2005. Classical transmitters and their receptors in flatworms. *Parasitology* 131, S19–S40.
- Rinaldi, G., Morales, M.E., Cancela, M., Castillo, E., Brindley, P.J., Tort, J.F., 2008. Development of functional genomic tools in trematodes: RNA interference and luciferase reporter gene activity in *Fasciola hepatica*. *PLoS Neglected Trop. Dis.* 2 (7), e260.
- Roberts, A., Trapnell, C., Donaghey, J., Rinn, J.L., Pachter, L., 2011. Improving RNA-Seq expression estimates by correcting for fragment bias. *Genome Biol.* 12 (3), R22. <http://dx.doi.org/10.1186/gb-2011-12-3-r22>.
- Saberi, A., Jamal, A., Beets, I., Schoofs, L., Newmark, P.A., 2016. GPCRs direct germline development and somatic gonad function in planarians. *PLoS Biol.* 14 (5), e1002457. Doi: 10.1371/journal.pbio.1002457.
- Saeger, B., Schmitt-Wrede, H.P., Dehnhardt, M., Bente, W.P.M., Krücken, J., Harder, A., Von Samson-Himmelstjerna, G., Wiegand, H., Wunderlich, F., 2001. Latrophilin-like receptor from the parasitic nematode *Haemonchus contortus* as a target for the anthelmintic decapeptide PF1022A. *Faseb. J.* 15, 1332–1334.
- Santos, R., Ursu, O., Gaulton, A., Bento, A.P., Donadi, R.S., Bologa, C.G., Karlsson, A., Al-Lazikani, B., Hersey, A., Oprea, T.I., Overington, J.P., 2017. A comprehensive map of molecular drug targets. *Nat. Rev. Drug Discov.* 16, 19–34.
- Sariozkan, S., Yalçın, C., 2011. Estimating the total cost of bovine fasciolosis in Turkey. *Ann. Trop. Med. Parasitol.* 105, 439–444.
- Sautel, M., Martinez, R., Munoz, M., Peitsch, M.C., Beck-Sickinger, A.G., Walker, P., 1995. Role of a hydrophobic pocket of the human Y1 neuropeptide Y receptor in ligand binding. *Mol. Cell. Endocrinol.* 112, 215–221.
- Sautel, M., Rudolf, K., Wittneben, H., Herzog, H., Martinez, R., Munoz, M., Eberlein, W., Engel, W., Walker, P., Beck-Sickinger, A.G., 1996. Neuropeptide Y and the non-peptide antagonist BIBP 3226 share an overlapping binding site at the human Y1 receptor. *Mol. Pharmacol.* 50, 285–292.
- Smallwood, P.M., Williams, J., Xu, Q., Leahy, D.J., Nathans, J., 2007. Mutational analysis of Nrr1-Frizzled4 recognition. *J. Biol. Chem.* 282, 4057–4068.
- Stockert, J.A., Devi, L.A., 2015. Advancements in therapeutically targeting orphan GPCRs. *Front. Pharmacol.* 6, 100. <http://dx.doi.org/10.3389/fphar.2015.00100>.
- Ta, N., Feng, X., Deng, L., Fu, Z., Hong, Y., Liu, J., Li, H., Lu, K., Lin, J., Yuan, C., 2015. Characterization and expression analysis of Wnt5 in *Schistosoma japonicum* at different developmental stages. *Parasitol. Res.* 114, 3261–3269.
- Taman, A., Ribeiro, P., 2009. Investigation of a dopamine receptor in *Schistosoma mansoni*: functional studies and immunolocalization. *Mol. Biochem. Parasitol.* 168, 24–33.



- Toet, H., Piedrafita, D.M., Spithill, T.W., 2014. Liver fluke vaccines in ruminants: strategies, progress and future opportunities. *Int. J. Parasitol.* 44, 915–927.
- Trapnell, C., Pachter, L., Salzberg, S.L., 2009. TopHat: discovering splice junctions with RNA-seq. *Bioinformatics* 25, 1105–1111.
- Trapnell, C., Williams, B., Pertea, G., Mortazavi, A., Kwan, G., van Baren, J., Salzberg, S.L., Wold, B.J., Pachter, L., 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28 (5), 511–515.
- Trapnell, C., Hendrickson, D., Sauvageau, M., Goff, L., Rinn, J.L., Pachter, L., 2012a. Differential analysis of gene regulation at transcript resolution with RNA-Seq. *Nat. Biotechnol.* 31 (1), 46–53.
- Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimental, H., Salzberg, S.L., Rinn, J.L., Pachter, L., 2012b. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* 7 (3), 562–578.
- Tsai, I.J., Zarowiecki, M., Holroyd, N., Garciarrubio, A., Sanchez-Flores, A., Brooks, K.L., Tracey, A., Bobes, R.J., Fragoso, G., Sciotto, E., Aslett, M., Beasley, H., Bennett, H.M., Cai, X., Camicia, F., Clark, R., Cucher, M., De Silva, N., Day, T.A., Deplazes, P., Estrada, K., Fernandez, C., Holland, P.W.H., Hou, J., Hu, S., Huckvale, T., Hung, S.S., Kamenetzky, L., Keane, J.A., Kiss, F., Koziol, U., Lambert, O., Liu, K., Luo, X., Luo, Y., Macchiaroli, N., Nichol, S., Paps, J., Parkinson, J., Pouchkina-Stantcheva, N., Riddiford, N., Rosezvit, M., Salinas, G., Wasmuth, J.D., Zamanian, M., Zheng, Y., *Taenia solium* Genome Consortium, Cai, J., Soberon, X., Olson, P.D., Lacleste, J.P., Brehm, K., Berriman, M., 2013. The genomes of four tapeworm species reveal adaptations to parasitism. *Nature* 496, 57–63.
- Tusnday, G.E., Simon, I., 1998. Principals governing amino acid composition of integral membrane proteins: application to topology prediction. *J. Mol. Biol.* 283, 489–506.
- Tusnday, G.E., Simon, I., 2001. The HMMTOP transmembrane topology prediction server. *Bioinformatics* 17, 849–850.
- Vilella, A.J., Severin, J., Ureta-Vidal, A., Heng, L., Durbin, R., Birney, E., 2009. EnsemblCompara GeneTrees: complete, duplication-aware phylogenetic trees in vertebrates. *Genome Res.* 19 (2), 327–335.
- Vogel, K.J., Brown, M.R., Strand, M.R., 2013. Phylogenetic investigation of peptide hormone and growth factor receptors in five dipteran genomes. *Front. Endocrinol.* 4, 193.
- Wacker, D., Wang, C., Katritch, V., Han, G.W., Huang, X.P., Vardy, E., McCorvy, J.D., Jiang, Y., Chu, M., Siu, F.Y., Liu, W., Xu, H.E., Cherezov, V., Roth, B.L., Stevens, R.C., 2013. Structural features for functional selectivity at serotonin receptors. *Science* 340, 615–619.
- Wang, C., Jiang, Y., Ma, J., Wu, H., Wacker, D., Katritch, V., Han, G.W., Liu, W., Huang, X.P., Vardy, E., McCorvy, J.D., Gao, X., Zhou, X.E., Melcher, K., Zhang, C., Bai, F., Yang, H., Yang, L., Wellendorph, P., Brauner-Osborne, H., 2009. Molecular basis for amino acid sensing by family C G-protein-coupled receptors. *Br. J. Pharmacol.* 156, 869–884.
- Wang, C., Jiang, Y., Ma, J., Wu, H., Wacker, D., Katritch, V., Han, G.W., Liu, W., Huang, X.P., Vardy, E., McCorvey, J.D., Gao, X., Zhou, X.E., Melcher, K., Zhang, C., Bai, F., Yang, H., Yang, L., Jiang, H., Roth, B.L., Cherezov, V., Stevens, R.C., Xu, H.E., 2013a. Structural basis for molecular recognition at serotonin receptors. *Science* 340, 610–614.
- Wang, B., Collins 3rd, J.J., Newmark, P.A., 2013b. Functional genomic characterization of neoblast-like stem cells in larval *Schistosoma mansoni*. *Elife* 2, e00768.
- Wellendorph, P., Bräuner-Osborne, H., 2009. Molecular basis for amino acid sensing by family C G-protein-coupled receptors. *Br. J. Pharmacol.* 156 (6), 869–884.
- Willert, K., Nusse, R., 2012. Wnt proteins. *Cold Spring Harb. Perspect. Biol.* 4 (9), a007864.
- Winkelhagen, A.J.S., Mank, T., de Vries, P.J., Soetekouw, R., 2012. Apparent triclabendazole-resistant human *Fasciola hepatica* infection, The Netherlands. *Emerg. Infect. Dis.* 18, 1028–1029.
- Wolstenholme, A.J., 2011. Ion channels and receptors as targets for the control of parasitic nematodes. *Int. J. Parasitol. Drugs Drug Resist* 1 (1), 2–13.
- Yoshida, M., Miyazato, M., Kangawa, K., 2012. Orphan GPCRs and methods for identifying their ligands. *Methods Enzymol.* 514, 33–44.
- Zamanian, M., Kimber, M.J., McVeigh, P., Carlson, S.A., Maule, A.G., Day, T.A., 2011. The repertoire of G protein-coupled receptors in the human parasite *Schistosoma mansoni* and the model organism *Schmidtea mediterranea*. *BMC Genom.* 12, 596.
- Zamanian, M., Agbedanu, P.N., Wheeler, N.J., McVeigh, P., Kimber, M.J., Day, T.A., 2012. Novel RNAi-mediated approach to G protein-coupled receptor deorphanization: proof of principle and characterization of a planarian 5-HT receptor. *PLoS One* 7 (7), e40787.