ARTICLE IN PRESS

International Journal for Parasitology xxx (xxxx) xxx

Contents lists available at ScienceDirect

International Journal for Parasitology

journal homepage: www.elsevier.com/locate/ijpara

Schistosoma mansoni venom allergen-like proteins: phylogenetic relationships, stage-specific transcription and tissue localization as predictors of immunological cross-reactivity

Leonardo P. Farias [a,b,1], Iain W. Chalmers [c,1], Samirah Perally [c], Henrique K. Rofatto [a], Colin J. Jackson [c], Martha Brown [c], Mariana I. Khouri [b], Mayra M.F. Barbosa [a,d], Paul J. Hensbergen [e], Cornelis H. Hokke [f], Luciana C.C. Leite [a], Karl F. Hoffmann [c,*]

[a] Centro de Biotecnologia, Instituto Butantan, Av. Vital Brasil, 1500, 05503-900 São Paulo, SP, Brazil
[b] Instituto Gonçalo Moniz, Fundação Oswaldo Cruz (FIOCRUZ), Rua Waldemar Falcão, Salvador, Bahia, Brazil
[c] Institute of Biological, Environmental and Rural Sciences, Aberystwyth University, SY23 3FG Aberystwyth, UK
[d] Programa de Pós-Graduação Interunidades em Biotecnologia, Universidade de São Paulo, São Paulo, SP, Brazil
[e] Center for Proteomics and Metabolomics, Leiden University Medical Centre, P.O. Box 9600, 2300 RC Leiden, The Netherlands
[f] Department of Parasitology, Leiden University Medical Centre, P.O. Box 9600, 2300 RC Leiden, The Netherlands

ARTICLE INFO

Article history:
Received 21 December 2018
Received in revised form 25 March 2019
Accepted 28 March 2019
Available online xxxx

Keywords:
Venom allergen-like
Immunological cross reactivity
Tissue localization
Schistosoma mansoni

ABSTRACT

Schistosoma mansoni venom allergen-like proteins (SmVALs) are part of a diverse protein superfamily partitioned into two groups (group 1 and group 2). Phylogenetic analyses of group 1 SmVALs revealed that members could be segregated into subclades (A–D); these subclades share similar gene expression patterns across the parasite lifecycle and immunological cross-reactivity. Furthermore, whole-mount in situ hybridization demonstrated that the phylogenetically, transcriptionally and immunologically-related SmVAL4, 10, 18 and 19 (subclade C) were all localized to the pre-acetabular glands of immature cercariae. Our results suggest that SmVAL group 1 phylogenetic relationships, stage-specific transcriptional profiles and tissue localization are predictive of immunological cross-reactivity.
© 2019 The Authors. Published by Elsevier Ltd on behalf of Australian Society for Parasitology. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Schistosomiasis, a disease caused by parasitic flatworms within the genus *Schistosoma*, affects 200 million people in sub-Saharan Africa, South America and the Far East, with a further 600 million at risk of developing the disease (McManus et al., 2018). Currently, treatment against schistosomiasis relies on a single drug, Praziquantel (PZQ), which may lead to drug resistance and, over the long term, reduced efficacy (Doenhoff and Pica-Mattoccia, 2006). Therefore, the search for new sustainable control strategies and intervention targets remains an important priority.

In light of this objective, discoveries obtained through post-genomics investigation (Hokke et al., 2007) have led to studies focused on the identification of novel schistosome antigens putatively involved in the parasite/host interface. As a result, remarkable progress in the understanding of schistosome secretions and tegumental composition has been achieved, revealing the presence of multigenic protein families including the tegument allergen-like proteins (TALs), Tetraspanins, micro-exon gene (MEGs), Ly6 proteins and the venom-allergen like proteins (VALs) (reviewed in Wilson (2012)). While several members of these families are being examined for vaccine potential (e.g. *Schistosoma mansoni* tetraspanin 2 (SmTSP2); (Tran et al., 2006) and Sm29; (Cardoso et al., 2008)), an important aspect that has been overlooked is the extent to which immunological cross-reactivity occurs with respect to members of a given protein family.

Cross-reactive immune responses related to a group of proteins are crucially important in many areas of biology. However, the majority of research in this area focuses on inter-species/strain cross-reactivity, as for example the overlap in allergens between different foods (Kazatsky and Wood, 2016) and overlapping responses between species/strains of pathogens (Lee and Wilson, 2015; Cao et al., 2016). Nevertheless, there is an increasing understanding that immunological cross-reactivity occurring in response to similar protein family members in the same species can be critical to provide immunity against parasites. In malaria, a study of the vaccine candidate merozoite surface protein 3 (MSP3) showed that six related MSP3 genes presented conserved transcript expression profiles, similar localization and serological cross-reactivity (Singh et al., 2009). Limiting/avoiding cross-reactivity is also a

* Corresponding author.
E-mail address: krh@aber.ac.uk (K.F. Hoffmann).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.ijpara.2019.03.003
0020-7519/© 2019 The Authors. Published by Elsevier Ltd on behalf of Australian Society for Parasitology.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Please cite this article as: L. P. Farias, I. W. Chalmers, S. Perally et al., *Schistosoma mansoni* venom allergen-like proteins: phylogenetic relationships, stage-specific transcription and tissue localization as predictors of immunological cross-reactivity, International Journal for Parasitology, https://doi.org/10.1016/j.ijpara.2019.03.003

key aspect in pathogen immune responses, in which multigenic families such as *P. falciparum* erythrocyte membrane protein 1 (PfEMP1) (Klein et al., 2014) and *Trypanosoma brucei* variant surface glycoproteins (VSG), are critical to pathogen survival (Cross, 1977). Importantly, in human schistosomiasis natural immunity has also been associated with cross-reactive antibody responses to related SmTAL proteins (Fitzsimmons et al., 2012). Given these findings, this study proposed to examine antigenic cross-reactivity within the *Schistosoma mansoni* VAL protein family, since a number of these members are present in parasite secretions and are also being investigated as anti-schistosomal vaccine candidates (Farias et al., 2011).

All SmVALs contain the SCP/TAPS (Sperm-coating protein/ Tpx-1/Ag5/PR-1/Sc7; Pfam PF00188) domain, which is a conserved α-β-α sandwich structure, ubiquitously found in family members across phyla. To date there have been 29 SCP/TAPS domain-containing proteins identified in *S. mansoni* (SmVAL1-28 from Chalmers et al. (2008) and SmVAL29 from Farias et al. (2012)). The SmVAL family is divided into two distinct sub-groups. All SmVALs from group 1 contain a predicted signal sequence and six highly conserved cysteine residues capable of forming disulphide bridges, indicating that these are secreted by the parasite. Group 2 SmVALs, which are fewer in number, consisting of only five members, are unlikely to be secreted due to their lack of a signal sequence as well as the six highly conserved cysteine residues. This is supported by the detection of multiple group 1 SmVALs in proteomic studies of parasite excretory/secretory (E/S) products including cercarial secretions (SmVAL4, 10 and 18; (Curwen et al., 2006; Hansell et al., 2008)), egg secretions (SmVAL2, 3, 5 and 9; (Cass et al., 2007)), miracidial/sporocyst secretions (SmVAL2, 3/23, 5/15, 9, 26/28, 27 and 29; (Wu et al., 2009; Farias et al., 2012)) and egg hatching fluid/secretions (SmVAL26/28; (Mathieson and Wilson, 2010; Farias et al., 2012)). Despite the abundance of group 1 SmVALs in these important life stages, the understanding of SmVAL function remains somewhat limited. A lipid-binding function was demonstrated for SmVAL4 (Kelleher et al., 2014), a plasminogen-binding function was revealed for SmVAL18 (Fernandes et al., 2017) and a host matrix metalloprotease modulatory function detected for SmVAL9 (Yoshino et al., 2014). The group 2 protein SmVAL6 has recently been shown to be a target for IgE, IgG4 and IgG1 responses in infected individuals, indicating that the SmVALs may influence the immune response (Farnell et al., 2015). However, these studies have limited their focus to individual SmVALs. Here, we demonstrate that phylogenetic, gene expression and spatial localization information can be used in combination to predict group 1 SmVAL immunological cross-reactivity. Further extending this analysis to the entire SmVAL family (and other multigenic families) may help assist in the selection of next generation of vaccine candidates.

The *S. mansoni* life cycle was maintained in hamsters or Tuck Ordinary (T.O.) mice (Harlan, USA) and experimental protocols were approved by both the Aberystwyth University (UK) animal welfare and ethical review body (AWERB, project license PPL 40/3700) and the Institutional Review Board on Animal Experimentation of the Butantan Institute, Brazil (CEUAIB; license no. 604-2009). *Schistosoma mansoni* cercariae, schistosomula, adult worms, eggs, miracidia and mother sporocysts (24 h and 96 h sporocysts) were obtained as previously described (Chalmers et al., 2008). Germ balls were obtained as previously described (Fernandes et al., 2017).

To examine whether antigenic cross-reactivity could occur between the SmVAL protein family members, we first compared the expression profiles of SmVALs at different stages of the parasite life cycle with respect to phylogenetic relatedness. Previous quantitative reverse transcription PCR (qRT-PCR) analysis provided gene expression quantification for 16 of the 29 *Smval* transcripts (1–13, 15, 21 and 23) (Chalmers et al., 2008). Herein we attempted to extend this analysis across the entire SmVAL family (data for SmVAL24 and 25 are absent due to amplification failures despite repeated attempts). Parasite materials, RNA extraction and cDNA synthesis as previously described (Chalmers et al., 2008) were used in qRT-PCR analyses of *Smvals 14, 16–20, 22* and *26–29* (qRT-PCR primers used can be found in Supplementary Table S1). In the cases of *Smval26, 27* and *28*, due to extremely high levels of sequence similarity (90–99% identity at the nucleotide level), a single combined expression profile representing these three transcripts was produced. This was also the case in a previous study for *Smval1/21, Smval3/23* and *Smval5/15* (Chalmers et al., 2008). Of the nine new expression profiles obtained, several displayed peak expressions in stages associated with parasite invasion or establishment in the mammalian host (e.g. *Smval14, 16, 18* and *19*), whilst others appear to be highly expressed in stages related to parasite invasion of the invertebrate host (e.g. *Smvals 20, 22, 26/27/28*, and *29*). The only exception was *Smval17*, which presented elevated expression in miracidia and cercariae stages (Supplementary Fig. S1).

The SmVAL expression profiles were then compared with SmVAL phylogenetic relationships to investigate whether closely related SmVALs (based on primary amino acid sequences) also shared common transcriptional patterns across the life cycle. To directly compare SmVAL expression profiles, data were converted to reflect relative abundance, in which the lifecycle stage with the highest expression value was designated as score 1.0, while expression in the other lifecycle stages was normalized accordingly. Expression data for *Smval1-13, 15, 21* and *23* was derived from data originally published by Chalmers et al. (2008). For phylogram construction, Bayesian phylogenetic analysis was performed as outlined in Chalmers et al. (2008) using an amino acid sequence alignment of the 30 SCP/TAPS domains with non-conserved regions removed. The phylogram showed four well-supported subclades (A–D) within the group 1 SmVALs (Fig. 1). An examination of qRT-PCR expression within these subclades revealed clusters of genes with very similar expression profiles, which is reflective of high sequence relatedness. For example, in subclade A, SmVAL5, 9, 15, 26, 27, 28 and 29 group together (posterior probability support value 1.00), all exhibiting high levels of expression within the miracidia life stage, yet minimal expression levels in most other life stages (Fig. 1). In subclade C, SmVAL4, SmVAL10, SmVAL18, SmVAL19 and SmVAL20 all group together (posterior probability support value 0.72) with all but SmVAL20 sharing a similar expression profile that peaks in the cercaria stage (Fig. 1). In contrast, subclade B presents a mixed profile, with one member (SmVAL2) showing peak expression in the miracidia stage, versus another (SmVAL22) in sporocysts, while others (SmVAL1 and 21) peak in cercariae and one (SmVAL14) shows elevated expression in the adult stage (Fig. 1). It is interesting to note that most group 1 proteins exhibit peak expression in the penetrating larval stages (miracidia and cercariae). The exceptions to this are SmVAL7, 8 and 12, which did not segregate into a distinct subclade, as well as the subclade B member SmVAL14.

Based on SmVAL sequence relatedness, phylogenetic relationships and expression profiles, we hypothesized that cross-reactivity would occur within the subclades, but not across different SmVAL family subclades. To investigate this, recombinant (r) SmVAL4 and 9 were produced in *Escherichia coli* as described in a previous study (Yoshino et al., 2014) and rSmVAL5 protein was generated in *Pichia pastoris* (Supplementary Fig. S2); antisera against each recombinant protein were raised in mice (Supplementary Data S1). These anti-sera were then used in western blotting experiments against recombinant proteins and native extracts to assess cross-reactivity as previously described (Rofatto et al., 2012). First, experimental confirmation that rSmVAL4 and rSmVAL5 were recognized by their respective antisera was

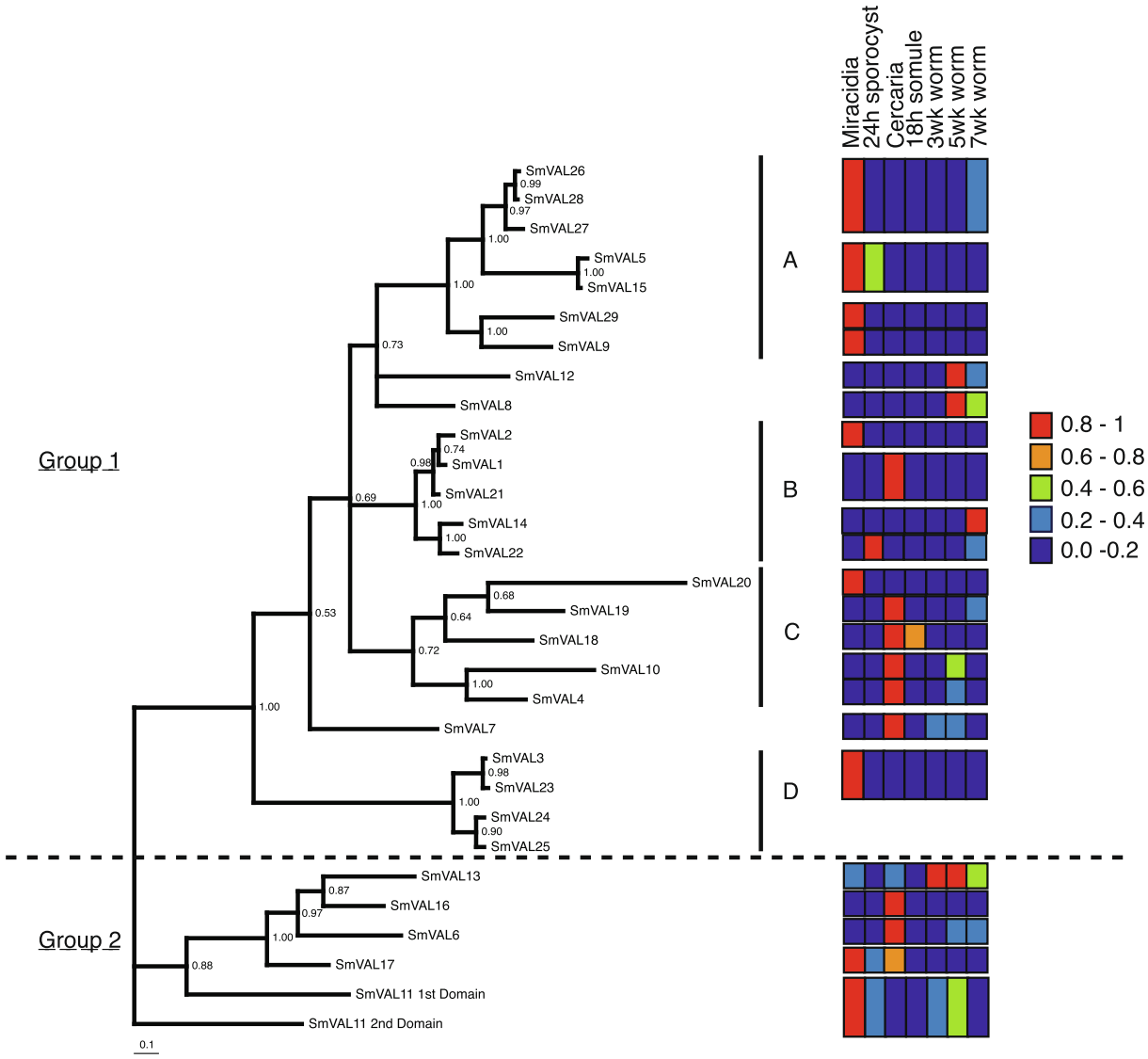


Fig. 1. *Schistosoma mansoni* Venom allergen-like proteins (SmVAL) subclade divisions and expression profile across the parasite life cycle. Rooted phylogram illustration detailing the 30 SmVAL SCP/TAPS (Sperm-coating protein/ Tpx-1/Ag5/PR-1/Sc7) domain-containing proteins and distinctive subclades (A–D), accompanied by a heat map showing transcription levels (subject to data availability) at different stages determined by quantitative reverse transcription (qRT)-PCR analysis. Where highly similar transcripts could not be analysed separately by qRT-PCR, the data is represented jointly. qRT-PCR results for each *Smval* were transformed into proportional data, where the value of the life cycle with the highest expression for that *Smval* was set at 1 and all other life cycle transcript levels compared with that figure. Red represents any lifecycle stage between 0.8 and 1 compared with the highest measured transcript abundance, while orange (0.6–0.8), green (0.4–0.6), light blue (0.2–0.4) and dark blue (0–0.2) represent lower comparative expression levels. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

obtained by one-dimensional SDS-PAGE of the recombinant proteins and Western blotting using anti-rSmVAL4 and anti-rSmVAL5 antisera (Fig. 2). While no cross-reactivity across subclades was detected, as anti-rSmVAL4 only recognized itself and anti-rSmVAL5 did not recognize rSmVAL4, there was evidence of cross-reactivity within the same subclade, with anti-rSmVAL5 antisera recognizing the SmVAL9 protein.

We then carried out two-dimensional gel electrophoresis (2-DE) Western blotting and mass spectrometric analysis to test for cross-reactivity within the A and C subclades by examining native SmVALs. Parasite protein extract preparations, electrophoresis, electroblotting, spot isolation, digestion and LC-MS/MS analysis are described in detail in Supplementary Data S1. For the subclade C analysis, proteins from cercariae secretions (0–3 h released products (RP)) were examined via 2-DE and blotted using anti-rSmVAL4 antisera. As predicted, due to the high sequence identity in this subclade (45–46%, complete protein sequence), anti-rSmVAL4 antisera recognized SmVAL4, SmVAL18 and SmVAL19 in the 0–3 h RP fraction (Fig. 3A, B, E and Supplementary Table S2). The results also showed two regions of cross-reactive signals (spots 2 and 5) that, upon LC-MS/MS analysis, yielded no significant hits to schistosome proteins (Fig. 3A and B). The predicted and observed molecular weight for some SmVALs differ significantly, likely due to protein glycosylation, as already demonstrated for SmVAL4, 10 and 18 (Jang-Lee et al., 2007; Farias et al., 2012). It is noteworthy that this is the first proteomics analysis reporting evidence of SmVAL19 protein detection in cercaria to schistosomula (0–3 h RP) secretions. Additionally, we observed a strong signal on the western blot (Fig. 3B, labeled as spot 5) in a pH region that resembles what was characterized as SmVAL10 protein by Curwen et al. (2006). However, we were unable to identify the SmVAL10 protein in our mass spectrometry analysis, likely due to low isoelectric focusing resolution.

To further evaluate cross-reactivity within subclade A, anti-rSmVAL5 antisera was applied to 2-DE blots of *S. mansoni* soluble egg antigen (SEA) preparations. SEA preparation, electrophoresis,

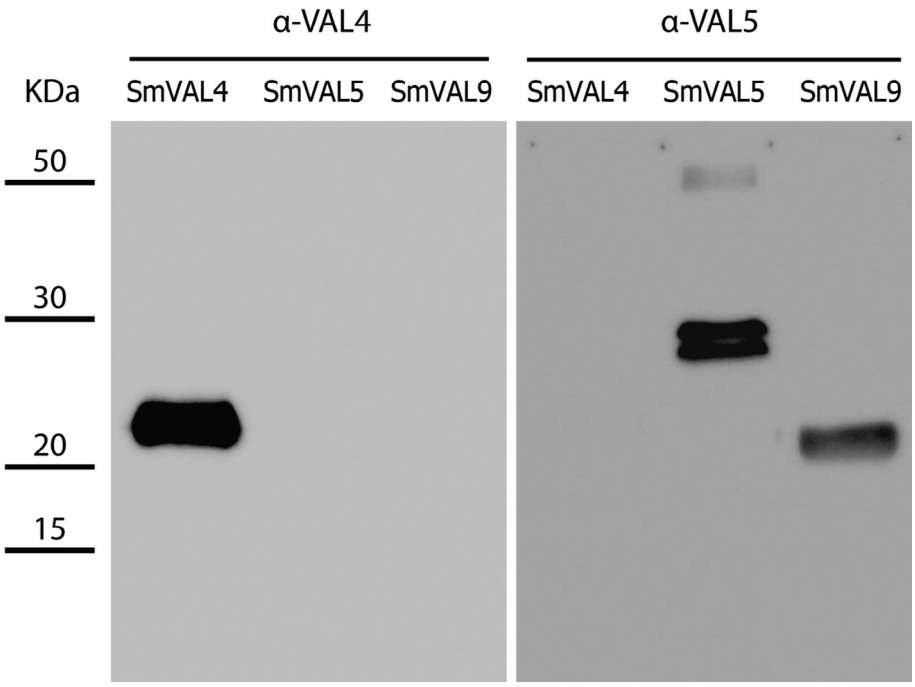


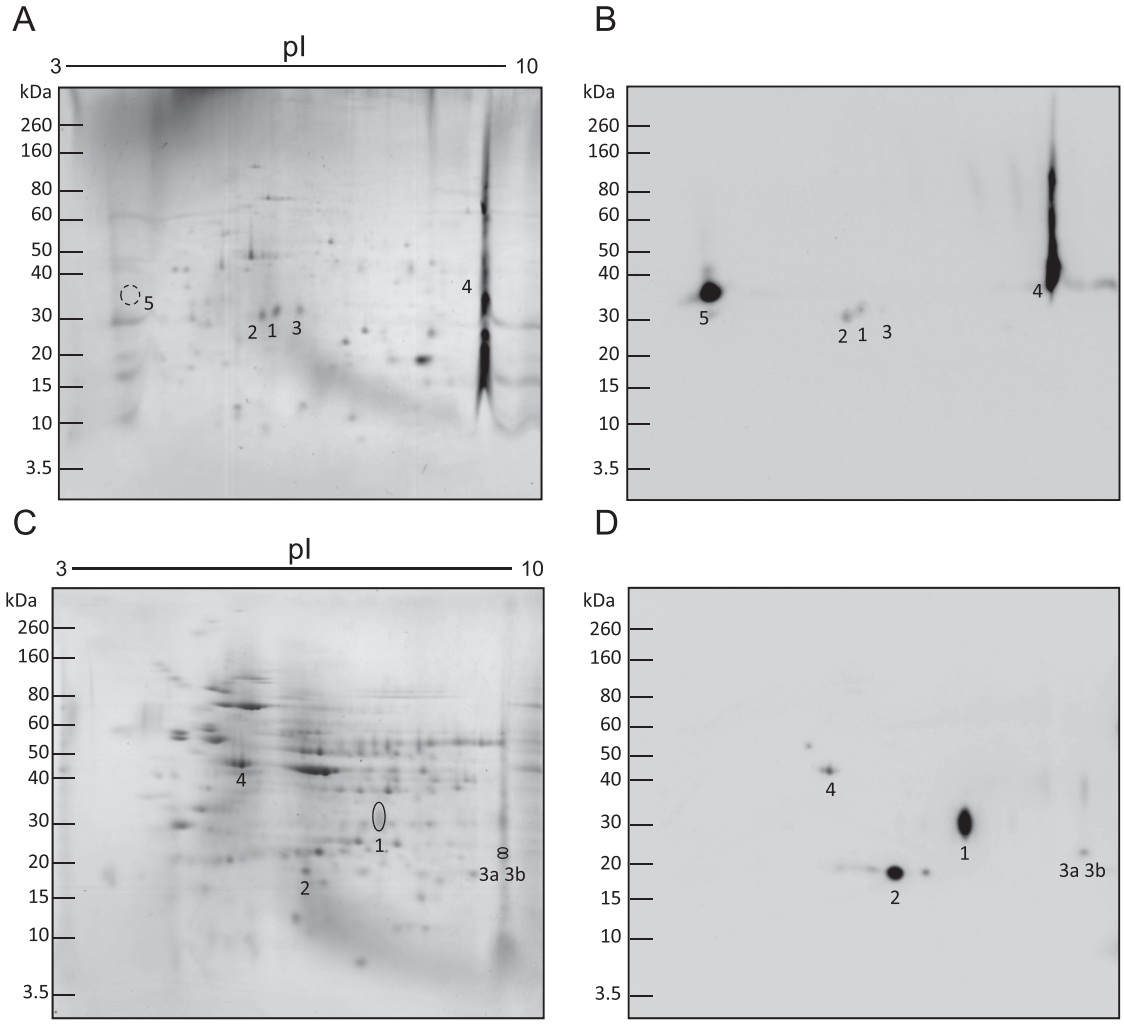
Fig. 2. Anti-SmVAL (*Schistosoma mansoni* Venom allergen-like proteins) antisera cross-react only with intra-subclade members. Evaluation of cross-reactivity using recombinant SmVALs 4, 5 and 9 and anti-recombinant (r)SmVAL4 and anti-rSmVAL5 antisera. Recombinant proteins were submitted to western blotting (1.5 μg/lane). Blots were probed with murine antisera (1:15,000) from animals immunized with rSmVAL4 or rSmVAL5 expressed in *Pichia pastoris*. Signal was developed using a horseradish peroxidase (HRP)-conjugated anti-mouse IgG secondary antibody (Sigma no. A4416, 1:10,000) and enhanced chemiluminescence (ECL) prime reagent (GE healthcare no. RPN2232). Blots were exposed for 1 min and images were captured using a CCD camera (BioSpectrum Multi Imaging Unit, UVP, USA).

electroblotting, spot isolation, digestion and LC-MS/MS analysis are described in detail in Supplementary Data S1. The anti-rSmVAL5 antisera recognized proteins with different levels of sequence identity, namely SmVAL9 (49%), SmVAL26/28 (62/63%), SmVAL27 (60%) and SmVAL29 (50%) (Fig. 3C, D, E and Supplementary Table S3). It should be noted that as SmVAL26 and SmVAL28 share almost identical primary sequences, neither of the peptides identified were discriminatory between these two SmVALs. Surprisingly, none of the proteins detected were identified as SmVAL5, again likely due to isoelectric focusing resolution or MS detection limitations.

To examine whether the subclade C members identified as cross-reactive shared the same localization within the parasite, we employed Whole-Mount In Situ Hybridization (WISH), which we believe is a stringent technique capable of accurately mapping spatial gene expression in parasite tissue. Prior to conducting WISH assays, we first evaluated the specificity of the SmVAL4, 10 and 18 RNA probes against their corresponding cDNAs by modified reverse northern blotting. Briefly, template cDNAs were generated via the amplification of previously cloned genes using PCR (Chalmers et al., 2008). Approximately 50 ng of the PCR products were run on a 1% agarose gel; after documentation, the gel was submitted to capillary blotting on Hybond-N+ charged nylon membranes (GE Healthcare, USA), followed by depurination, denaturation and neutralization, all carried out according to the manufacturer's recommendations. Blocking, probe hybridization and washing steps, as well as detection techniques, were performed under the same conditions described below for the WISH assays. Images were captured using an Image Quant LAS 4000 photodocumentation system (GE Healthcare, USA). Our analysis revealed that, contrary to our results regarding the anti-rSmVAL4 antibody, the SmVAL RNA probes specifically detected their related cDNAs, with no evidence of cross-reactivity observed (Supplementary Fig. S3).

Next, WISH was used to determine the localization of *Smval4, 10 18* and *19* transcripts in immature cercariae. Additionally, *Smval1* localization (a subclade B protein expressed in the cercaria) was examined to ascertain whether different subclades possessed unique localizations within the parasite. The protocols used for fixation, permeabilization, in situ hybridization and staining of germ balls and immature cercariae were based on protocols previously described (Fernandes et al., 2017), and recently used to localize *Smval4* and *24* in the pre-acetabular glands of immature cercariae. Specific antisense RNA probes were synthesized with digoxigenin (DIG) in vitro using T7 or Sp6 RNA polymerase (Promega, Madison, USA) from cDNA sequences previously cloned in a pGEM-T easy vector (Supplementary Table S1). A *Smval4* DIG-labeled sense probe was used as a negative control. All five probes presented clear evidence of staining in the cercaria stage, represented by a band across the middle of the body along the anterior and lateral edges of the pre-acetabular glands (Fig. 4A–E). Confocal microscopy employing Alexa Fluor 647-conjugated lectin (PNA) and FITC-phalloidin was used to provide additional interpretation of the positioning of pre- and post-acetabular glands with musculature counterstaining (Fig. 4G and H), which revealed that two pairs of pre-acetabular glands lie anterior to the acetabulum, while three pairs of post-acetabular glands were located posterior to it. In addition, detectable levels of these transcripts were observed in other development stages of germ balls (Supplementary Fig. S4); in some of these, it was possible to view the two pairs of pre-acetabular glands as indicated by four blue spots (Supplementary Fig. S4C, G, H and L).

Recently, transcriptomic, microarray and proteomic studies have demonstrated the potential of SmVALs as prospective targets for immune intervention (reviewed in Chalmers and Hoffmann (2012)). The present study attempted to extend previous analyses of the transcriptional expression profiles of the SmVAL family (*Smval14-29*, with the exception of *Smval15, 21, 23 and 24/25*).



E

Gel	Spots	Protein Identity	Expected MW / pI	Unique peptides
Fig. 3A	1, 3	SmVAL18	19.6 kDa / 5.5	4/4
Fig. 3A	4	SmVAL4 SmVAL19	18.5 kDa / 8.6 19.1 kDa / 8.3	1 1
Fig. 3C	1	SmVAL9 SmVAL29	18.3 kDa / 6.9 17.6 kDa / 7.6	3 3
Fig. 3C	2	SmVAL26/28	20.5 kDa / 6.0	2
Fig. 3C	3a,3b	SmVAL27	18.1 kDa / 9.0	4
Fig. 3C	4	Actin	41.6 kDa / 5.3	11

Fig. 3. Intra-clade cross-reactivity of anti-rSmVAL4 (recombinant *Schistosoma mansoni* Venom allergen-like 4 protein) and anti-rSmVAL5 antisera against native SmVALs. (A) two-dimensional gel electrophoresis (2-DE) of 0–3 h cercarial/schistosomula released products (50 μg) stained with colloidal Coomassie. (B) Western blot of a replica gel presented in (A) revealing protein spots that cross-react with anti-rSmVAL4 antibodies. (C) 2-DE of soluble egg antigens (SEA) (56 μg) stained with colloidal Coomassie. (D) Western blot from a replica gel presented in (C) showing spots that cross-react with anti-rSmVAL5 antisera. Positions of molecular mass standards (kDa) are indicated on the sides of the gels. (E) Spots submitted to LC-MS/MS analysis and SmVAL protein identification. The predicted MW and average pI for each SmVAL identified as defined by Expasy (https://web.expasy.org/compute_pi/) and the number of unique peptides identified during mass spectrometry analysis are displayed. The full mass spectrometric analysis for all protein identities is detailed in Supplementary Tables S2 and S3.

Again, we broadly distinguished gene products that were associated with invasion or establishment in the mammalian host from others that appeared to be highly expressed in parasite stages related to the invasion of the invertebrate host.

In order to select an appropriate vaccine candidate in the context of a multi-gene family, it is crucial to consider cross-reactivity among members. Ultimately, cross-reactivity will determine the extent by which different antigens are recognized by the

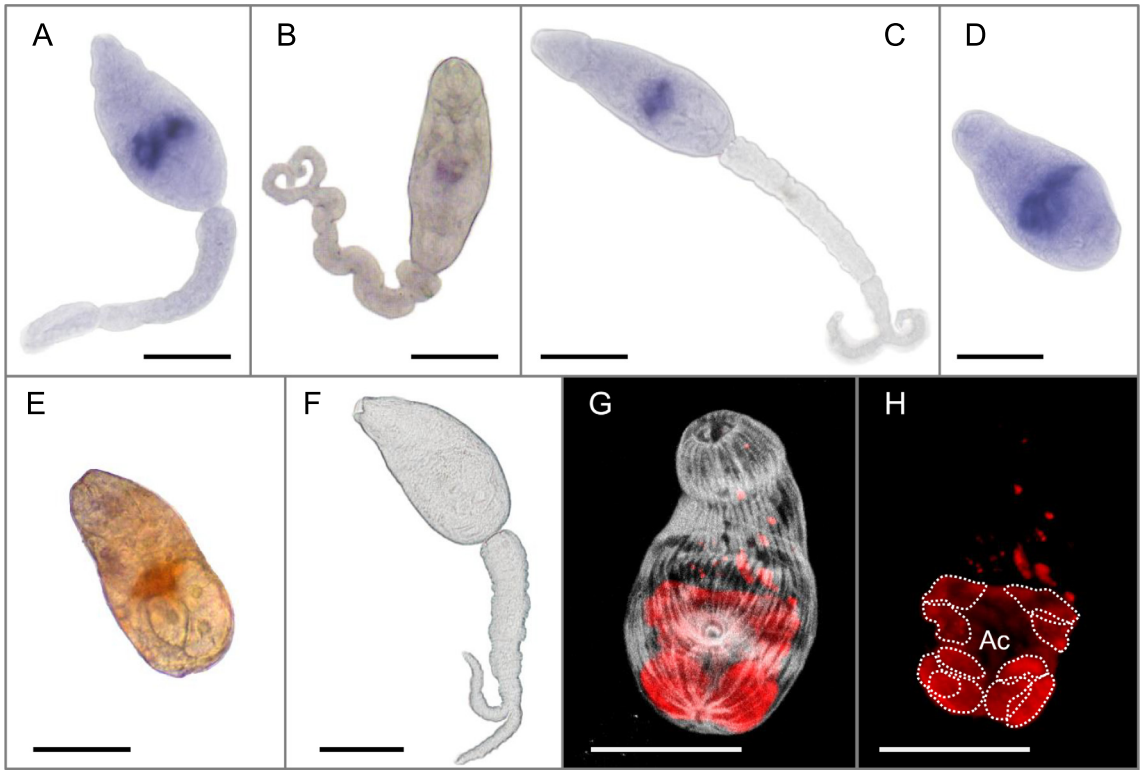


Fig. 4. Localization of *Schistosoma mansoni* Venom allergen-like protein (SmVAL) transcripts to the pre-acetabular glands in *S. mansoni* cercariae. Whole mount in situ hybridization for *Smval1, 4, 10 18* and *19* transcripts in immature cercariae. (A–E) focal expression associated with the pre-acetabular glands of immature cercariae hybridized with *Smval1, 4, 10, 18* and *19* probes, respectively. (F) Negative control hybridized with a sense probe of SmVAL4. (G and H) A phalloidin-FITC stained parasite incubated with PNA-647 in the same orientation to aid interpretation. Note the acetabulum (Ac) and the two and three pairs of pre- and post-acetabular glands, respectively. Scale bars = 50 μm.

host immune system, which has obvious implications in vaccine design. In the case of the schistosome, immunization with antigens sharing epitopes (i.e. cross-reacting) with egg antigens presents a potential risk, since this may sensitize to egg proteins, thereby inducing an immune response against disintegrating eggs that could consequently exacerbate the pathological process. We have previously demonstrated the absence of antibody cross-reactivity between anti-rSmVAL6 (a group 2 protein) and recombinants SmVAL4, 7 and 26 (all group 1 proteins), which crucially have low sequence similarity to SmVAL6 (Rofatto et al., 2012). Data presented here extends these findings, indicating the absence of cross-reactivity among different subclades of the group 1 SmVALs, while demonstrating cross-reactivity within these subclades, such as in subclade C (between SmVAL4, 18 and 19) and in subclade A (between SmVAL5, 9, 26/28, 27 and 29). Moreover, the detection of SmVAL19 in cercaria secretions further expands the repertoire of SmVALs that are likely involved in host invasion.

Within clade cross-reactivity may not occur at low sequence identity levels though, as previous experiments using anti-rSmVAL4 failed to detect a cross-reactive protein (such as the clade C protein SmVAL20) in egg or miracidia extracts using 1D Western Blot experiments (Farias et al., 2012). In addition, due to the use of recombinant proteins to raise the anti-sera for these experiments, we cannot exclude the possibility that additional/different cross-reactivity could be detected using proteins with native folding and post-translational modifications. Therefore, a complete understanding of schistosome protein cross-reactivity may only be achieved with the use of anti-sera raised against native proteins due to schistosome-specific glycosylation such as was detected for SmVAL9 (Yoshino et al., 2014).

Our results demonstrate that anti-SmVAL antibodies cross-react with members from the same phylogenetic subclade (i.e. those sharing at least 30% sequence identity), suggesting that cellular localization studies using these anti-SmVALs antibodies could lead to misinterpretations. To circumvent this limitation, we employed the WISH technique, which is more stringent than immunohistochemistry due to the high probe hybridization temperatures used (60–65 °C). The WISH assays determined the tissue localization of *Smval1, 4, 10, 18 and 19* transcripts in the pre-acetabular glands of stubby-tailed, young elongating-tail and immature cercariae. These data corroborate and extend previous DNA microarray and proteomics data (Curwen et al., 2006; Chalmers et al., 2008; Hansell et al., 2008; Parker-Manuel et al., 2011), which imply that these SmVAL proteins may play a role in the early stages of infection. In light of the identification of SjVAL1 in the head gland of cercariae (Chen et al., 2010), it is tempting to speculate that some *S. mansoni* orthologs could also be localized in this structure. Our data, however, exclude this possibility, at least for SmVAL1, 4, 10, 18 and 19. In addition, SjVAL1 localization data was obtained using polyclonal antibodies, and as demonstrated here, multi-gene family immunolocalization data should be interpreted carefully due to cross-reactivity issues.

The detection of SmVAL1, 4, 10, 18 and 19 provided here, SmVAL4 and 24/25 described by Fernandes et al. (2017) and, potentially, SmVAL21, 22 (Parker-Manuel et al., 2011), in the acetabular glands indicates that 37.5% of all group 1 SmVAL proteins are expressed in this tissue. Although the amount of SmVAL4, 10 and 18 based on spot volume (~3%) does not rival the five cercarial elastases (~34.4% of the total spot volume) (Wilson, 2012), the large number of group 1 SmVALs (nine) potentially expressed

Please cite this article as: L. P. Farias, I. W. Chalmers, S. Perally et al., *Schistosoma mansoni* venom allergen-like proteins: phylogenetic relationships, stage-specific transcription and tissue localization as predictors of immunological cross-reactivity, International Journal for Parasitology, https://doi.org/10.1016/j.ijpara.2019.03.003

in the acetabular glands calls attention to this location as a hotspot for this gene family. In conclusion, our results highlight the importance of establishing localization status, as well as determining the extent of cross-reactivity across the SmVAL family members in order to better define their vaccine potential.

Acknowledgements

This work was supported by grants from the Welcome Trust (UK) (WT084273/Z/07/Z) to KFH, Fundação Butantan, Fundação de Amparo à Pesquisa do Estado de São Paulo (Brazil) to LPF and LLC (2012/23124-4), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) to LCCL and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001, and by fellowships from Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, Brazil) to LPF (2008/57946-5) and HKF (2007/07685-8) and from CNPq to MIK (160861/2017-9). We thank Dra. Eliana Nakano and Ms. Patricia A. Miyasato for supplying the parasite stages and to Alexsander Seixas de Souza for confocal microscopy (FAPESP 00/11624-5) imaging support, all from Instituto Butantan, Brazil.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijpara.2019.03.003.

References

Cao, Y., Bansal, G.P., Merino, K., Kumar, N., 2016. Immunological cross-reactivity between malaria vaccine target antigen P48/45 in *Plasmodium vivax* and *P. falciparum* and cross-boosting of immune responses. PloS One 11, e0158212.

Cardoso, F.C., Macedo, G.C., Gava, E., Kitten, G.T., Mati, V.L., de Melo, A.L., Caliari, M. V., Almeida, G.T., Venancio, T.M., Verjovski-Almeida, S., Oliveira, S.C., 2008. *Schistosoma mansoni* tegument protein Sm29 is able to induce a Th1-type of immune response and protection against parasite infection. PLoS Negl. Trop. Dis. 2, e308.

Cass, C.L., Johnson, J.R., Califf, L.L., Xu, T., Hernandez, H.J., Stadecker, M.J., Yates 3rd, J. R., Williams, D.L., 2007. Proteomic analysis of *Schistosoma mansoni* egg secretions. Mol. Biochem. Parasitol. 155, 84–93.

Chalmers, I.W., Hoffmann, K.F., 2012. Platyhelminth Venom Allergen-Like (VAL) proteins: revealing structural diversity, class-specific features and biological associations across the phylum. Parasitology 139, 1231–1245.

Chalmers, I.W., McArdle, A.J., Coulson, R.M., Wagner, M.A., Schmid, R., Hirai, H., Hoffmann, K.F., 2008. Developmentally regulated expression, alternative splicing and distinct sub-groupings in members of the *Schistosoma mansoni* venom allergen-like (SmVAL) gene family. BMC Genomics 9, 89.

Chen, J., Hu, X., He, S., Wang, L., Hu, D., Wang, X., Zheng, M., Yang, Y., Liang, C., Xu, J., Yu, X., 2010. Expression and immune response analysis of *Schistosoma japonicum* VAL-1, a homologue of vespid venom allergens. Parasitol. Res. 106, 1413–1418.

Cross, G.A., 1977. Antigenic variation in trypanosomes. Am. J. Trop. Med. Hyg. 26, 240–244.

Curwen, R.S., Ashton, P.D., Sundaralingam, S., Wilson, R.A., 2006. Identification of novel proteases and immunomodulators in the secretions of schistosome cercariae that facilitate host entry. Mol. Cell Proteomics: MCP 5, 835–844.

Doenhoff, M.J., Pica-Mattoccia, L., 2006. Praziquantel for the treatment of schistosomiasis: its use for control in areas with endemic disease and prospects for drug resistance. Expert Rev. Anti Infect. Ther. 4, 199–210.

Farias, L.P., Rodrigues, D., Cunna, V., Rofatto, H.K., Faquim-Mauro, E.L., Leite, L.C., 2012. *Schistosoma mansoni* venom allergen like proteins present differential allergic responses in a murine model of airway inflammation. PLoS Negl. Trop. Dis. 6, e1510.

Farias, L.P., Tararam, C.A., Miyasato, P.A., Nishiyama Jr., M.Y., Oliveira, K.C., Kawano, T., Verjovski-Almeida, S., Leite, L.C., 2011. Screening the *Schistosoma mansoni* transcriptome for genes differentially expressed in the schistosomulum stage in search for vaccine candidates. Parasitol. Res. 108, 123–135.

Farnell, E.J., Tyagi, N., Ryan, S., Chalmers, I.W., Pinot de Moira, A., Jones, F.M., Wawrzyniak, J., Fitzsimmons, C.M., Tukahebwa, E.M., Furnham, N., Maizels, R. M., Dunne, D.W., 2015. Known allergen structures predict *Schistosoma mansoni* IgE-binding antigens in human infection. Front. Immunol. 6, 26.

Fernandes, R.S., Barbosa, T.C., Barbosa, M.M.F., Miyasato, P.A., Nakano, E., Leite, L.C. C., Farias, L.P., 2017. Stage and tissue expression patterns of *Schistosoma mansoni* venom allergen-like proteins SmVAL 4, 13, 16 and 24. Parasit. Vectors 10, 223.

Fitzsimmons, C.M., Jones, F.M., Pinot de Moira, A., Protasio, A.V., Khalife, J., Dickinson, H.A., Tukahebwa, E.M., Dunne, D.W., 2012. Progressive cross-reactivity in IgE responses: an explanation for the slow development of human immunity to schistosomiasis? Infect. Immun. 80, 4264–4270.

Hansell, E., Braschi, S., Medzihradszky, K.F., Sajid, M., Debnath, M., Ingram, J., Lim, K. C., McKerrow, J.H., 2008. Proteomic analysis of skin invasion by blood fluke larvae. PLoS Negl. Trop. Dis. 2, e262.

Hokke, C.H., Fitzpatrick, J.M., Hoffmann, K.F., 2007. Integrating transcriptome, proteome and glycome analyses of *Schistosoma* biology. Trends Parasitol. 23, 165–174.

Jang-Lee, J., Curwen, R.S., Ashton, P.D., Tissot, B., Mathieson, W., Panico, M., Dell, A., Wilson, R.A., Haslam, S.M., 2007. Glycomics analysis of *Schistosoma mansoni* egg and cercarial secretions. Mol. Cell. Proteomics 6, 1485–1499.

Kazatsky, A.M., Wood, R.A., 2016. Classification of food allergens and cross-reactivity. Curr. Allergy Asthma Rep. 16, 22.

Kelleher, A., Darwiche, R., Rezende, W.C., Farias, L.P., Leite, L.C., Schneiter, R., Asojo, O.A., 2014. *Schistosoma mansoni* venom allergen-like protein 4 (SmVAL4) is a novel lipid-binding SCP/TAPS protein that lacks the prototypical CAP motifs. Acta Crystallogr. D Biol. Crystallogr. 70, 2186–2196.

Klein, E.Y., Graham, A.L., Llinas, M., Levin, S., 2014. Cross-reactive immune responses as primary drivers of malaria chronicity. Infect. Immun. 82, 140–151.

Lee, P.S., Wilson, I.A., 2015. Structural characterization of viral epitopes recognized by broadly cross-reactive antibodies. Curr. Top. Microbiol. Immunol. 386, 323–341.

Mathieson, W., Wilson, R.A., 2010. A comparative proteomic study of the undeveloped and developed *Schistosoma mansoni* egg and its contents: the miracidium, hatch fluid and secretions. Int. J. Parasitol. 40, 617–628.

McManus, D.P., Dunne, D.W., Sacko, M., Utzinger, J., Vennervald, B.J., Zhou, X.N., 2018. Schistosomiasis. Nat. Rev. Dis. Primers 4, 13.

Parker-Manuel, S.J., Ivens, A.C., Dillon, G.P., Wilson, R.A., 2011. Gene expression patterns in larval *Schistosoma mansoni* associated with infection of the mammalian host. PLoS Negl. Trop. Dis. 5, e1274.

Rofatto, H.K., Parker-Manuel, S.J., Barbosa, T.C., Tararam, C.A., Alan Wilson, R., Leite, L.C., Farias, L.P., 2012. Tissue expression patterns of *Schistosoma mansoni* venom allergen-like proteins 6 and 7. Int. J. Parasitol. 42, 613–620.

Singh, S., Soe, S., Weisman, S., Barnwell, J.W., Perignon, J.L., Druilhe, P., 2009. A conserved multi-gene family induces cross-reactive antibodies effective in defense against *Plasmodium falciparum*. PloS One 4, e5410.

Tran, M.H., Pearson, M.S., Bethony, J.M., Smyth, D.J., Jones, M.K., Duke, M., Don, T.A., McManus, D.P., Correa-Oliveira, R., Loukas, A., 2006. Tetraspanins on the surface of *Schistosoma mansoni* are protective antigens against schistosomiasis. Nat. Med. 12, 835–840.

Wilson, R.A., 2012. Proteomics at the schistosome-mammalian host interface: any prospects for diagnostics or vaccines? Parasitology 139, 1178–1194.

Wu, X.J., Sabat, G., Brown, J.F., Zhang, M., Taft, A., Peterson, N., Harms, A., Yoshino, T. P., 2009. Proteomic analysis of *Schistosoma mansoni* proteins released during in vitro miracidium-to-sporocyst transformation. Mol. Biochem. Parasitol. 164, 32–44.

Yoshino, T.P., Brown, M., Wu, X.J., Jackson, C.J., Ocadiz-Ruiz, R., Chalmers, I.W., Kolb, M., Hokke, C.H., Hoffmann, K.F., 2014. Excreted/secreted *Schistosoma mansoni* venom allergen-like 9 (SmVAL9) modulates host extracellular matrix remodelling gene expression. Int. J. Parasitol. 44, 551–563.

1. Supplementary materials and methods

1.1. Cloning, expression and purification of recombinant Schistosoma mansoni venom allergen-like (SmVAL) proteins

The full-length open reading frames for SmVAL4, 5 and 9 was previously validated (Chalmers et al., 2008). These genes were amplified by PCR (oligonucleotides are listed on Supplementary Table S1) from parasite-derived cDNA (Chalmers et al., 2008) with the predicted native signal peptide removed, and then cloned into the pGEM-T Easy vector system (Promega, USA]). Sub-cloning, expression and purification conditions of SmVAL9 have been previously described (Yoshino et al., 2014) with SmVAL4 expression following the same procedures. The cDNA fragment corresponding to SmVAL5 (from M49 to R243) was cloned into the pPICZα-A vector (Invitrogen, USA), followed by the transformation in *Pichia pastoris*; the screening of clones and protein expression were as previously described (Farias et al., 2012). As this protein was expressed intracellularly, cells were harvested and resuspended in 200 ml of lysis buffer (100 mM Tris pH 6.4, 0.3 M NaCl), then disrupted with a bead beater (Biospec model-11079000) for 10 cycles of 1 min each (with a 1 min interval between each cycle). The crude homogenate was centrifuged at 20,000 *g* for 40 min. For protein purification, the inclusion bodies of recombinant SmVAL5 (rSmVAL5) were rinsed twice with washing buffer (lysis buffer, 2% Triton X-100, 2 M urea), and then resuspended in solubilization buffer (lysis buffer plus 5 mM beta-mercaptoethanol, 20 mM imidazole, 6 M guanidine hydrochloride). The recombinant protein was then purified by metal affinity chromatography using the Akta Prime system (GE Healthcare, USA) under denaturing conditions. Briefly, the sample was loaded onto a Ni2+-NTA column (5 ml bed volume) pre-equilibrated with the same buffer. The column was then washed with 10 bed volumes of equilibration buffer and eluted with 20-500 mM imidazole linear gradient. Fractions encompassing the main peak and the purity of the preparation were assessed by SDS-PAGE. The protein was then refolded by slow dialysis in PBS pH 7.4 (PBS) with 5% glycerol containing decreasing concentrations of guanidine (4, 3, 2, 1 M and no guanidine) (Supplementary Fig. S2).

1.2. Antisera preparation

Polyclonal mouse serum was produced against preparations of rSmVAL4, as previously described (Farias et al., 2012), and rSmVAL5, as described herein. Briefly, BALB/c mice were immunized four times at 7 day intervals with 10 µg of protein. For the initial dose, proteins were

formulated with TiterMax Gold adjuvant (Sigma, USA) as per the manufacturer's recommendations, while proteins were inoculated in PBS 1X for subsequent doses. Fifteen days after the final inoculation, the rodents were exsanguinated.

1.3. Two-dimensional (2-D) PAGE of native schistosome proteins

To obtain the proteins released by cercariae, mechanical transformation was performed by vortexing to stimulate the release of gland cell contents of ≈ 115,000 cercariae, which were then cultured for 3 h in 9 ml of RPMI 1640 medium (Invitrogen) at 37°C under 5% CO_2. Parasites were collected by centrifugation at 200 *g* at 4° C for 5 min and the supernatant (medium containing released proteins), termed 0-3 h RP, was stored at -20 °C with the addition of a protease inhibitor (Roche Diagnostics, Mannheim, Germany). Soluble protein extracts from eggs (SEA) were prepared in 40 mM Tris, pH 7.4, plus protease inhibitor (Roche Diagnostics, Germany) by sonication on ice (four cycles of 50 s, 40% amplitude, with 2 min intervals between each cycle). Samples were centrifuged at 17,000 g for 30 min at 4°C and the supernatant was recovered and stored at -20 °C. Protein concentrations were determined with a Bradford Assay Kit (Sigma-Aldrich, USA).

For 2-D polyacrylamide gel electrophoresis (2-DE), 0-3 h RP preparations were precipitated by adding 0.25 volume of 100% trichloroacetic acid to samples followed by freezing at -20°C for 1 h. Proteins were pelleted after centrifugation at 17,000 g for 20 min at 4°C. The protein pellet was washed twice in acetone in an ice bath. Once the protein pellet was air-dried, it was resuspended in isoelectric focusing (IEF) buffer (C1) and quantified by a BCA protein assay.

A total of 50 µg or 250 µg of protein sample was used to passively rehydrate and focus 7 cm pH 3-10 non-linear IPG strips (Biorad, UK) at 20°C for protein separation in the first dimension. IPG strips were focused between 10000-15000 Vh using a Biorad IEF focusing machine. Each IPG strip was equilibrated for 15 min in 3 ml of equilibration buffer (50 mM Tris-HCl, pH 8.8, 6 M urea, 30% v/v glycerol and 2% w/v SDS) with the addition of 20 mg/ml of dithiothreitol (DTT). The equilibration buffer containing DTT was removed, and IPG strips were then incubated for another 15 min in the same buffer with iodoacetimide (IAA) at 25 mg/ml. The focused proteins were separated in the second dimension on NuPage Novex pre-cast 4-12% Bis-Tris gradient gels (Invitrogen, USA) at 200 V for 35 min using the XCell SureLock® Mini-Cell electrophoresis system (Invitrogen, USA). Gels were run in duplicate to provide a reference gel for protein spot identification by western blotting. Reference gels

were stained with colloidal Coomassie and background staining was removed using 1% acetic acid. All Coomassie-stained gels were imaged with a GS-800 calibrated densitometer (Biorad, UK) set for Coomassie staining at 700 dpi. Key protein spots of interest were excised for in-gel tryptic digestion.

1.4. Western blotting

Resolved proteins on 1-D or 2-D PAGE gels were blotted onto polyvinyl difluoride (PVDF) membranes in NuPAGE transfer buffer (Invitrogen). Western blots were blocked in blocking buffer (50 mM Tris, pH 7.5, 0.3% Tween containing 5% skimmed milk powder) prior to incubation for 3 h with the primary antibody (1:3500 dilutions of anti-rSmVAL4, anti-rSmVAL5, or normal mouse serum) in antibody diluent (50 mM Tris, pH 7.5, 0.05% Tween, 5% skimmed milk powder and 150 mM NaCl). Blots were washed three times in 50 mM Tris-HCl, pH 7.5, before incubation for 1 h with secondary horseradish peroxidase-conjugated anti-mouse IgG antibody (Sigma) diluted 1:5000 in antibody diluent. After washing again in 50 mM Tris-HCl, pH 7.5, blots were developed using ECL-Plus reagent (GE Healthcare) and images were captured using a CCD camera (BioSpectrum Multi Imaging Unit, UVP, USA]).

1.5. In-gel tryptic digestion

Selected spots were cut from colloidal Coomassie-stained gels and stored in 10% ethanol/1% acetic acid at 4 °C until further use. For trypsinization, the gel plugs were cut into small pieces, and after two rounds of washing with 25 mM NH_4HCO_3 and dehydration with 100% acetonitrile, the gel particles were completely dried in a centrifugal vacuum concentrator (Eppendorf, Hamburg, Germany). Dried gel particles were re-swollen for 15 min on ice by addition of 15 µl of a trypsin solution (12.5 ng/µl in 25 mM NH_4HCO_3 sequencing grade modified trypsin, Promega, Madison, WI, USA). Subsequently, 20 µl of 25 mM NH_4HCO_3 were added and samples were kept on ice for an additional 30 min, followed by tryptic digestion overnight at 37 °C. The overlaying digestion solution was collected. Two additional rounds of extraction with 20 µl of 0.1% TFA were used to extract peptides from the gel plugs and the extracts were pooled.

1.6. Nano LC ESI MS/MS

Nanoflow LC was performed on an Ultimate LC system (Dionex, Sunnyvale, CA, USA). A volume of 10 µL of sample was injected onto a C18 PepMapTM 0.3 mm×5 mm trapping column (Dionex, USA) and washed with 100% A (2% acetonitrile in 0.1% formic acid in MQ water, v/v) at 20 µL/min for 15 min. Following valve switching, peptides were separated on a C18 PepMap 75 µm×150 mm column (Dionex) at a constant flow of 200 nL/min. The peptide elution gradient was from 10 to 60% B (95% acetonitrile in 0.1% formic acid in ultrapure water v/v) over 50 min. The nanoflow LC system was coupled to an HCTultra IonTrap (Bruker Daltonics, Bremen, Germany) using a nano-electrospray ionisation source. The spray voltage was set at 1.2 kV and the temperature of the heated capillary was set to 165 °C. Eluting peptides were analyzed using the data-dependent MS/MS mode over a 300–1500 m/z range. The five most abundant ions in an MS spectrum were selected for MS/MS analysis by collision-induced dissociation using helium as the collision gas.

1.7. Mass spectrometry data analysis

Peak lists were generated using DataAnalysis 4.0 software (Bruker Daltonics, USA) and exported as Mascot Generic (MGF) files. These files were searched against the *S. mansoni* predicted protein sequences database (v4.0h, www.genedb.org) using the Mascot (version 2.2.1) search algorithm (Matrix Science, London, UK) and data from one plug were merged using Mascot Deamon. A peptide mass tolerance of 0.6 Da (with # 13C = 1) and a MS/MS fragment tolerance of 0.5 Da was used. Trypsin was designated as the enzyme and up to one missed cleavage site was allowed. Carbamidomethylcysteine was selected as a fixed modification and oxidation of methionine as a variable modification. Only significant protein hits with at least one unique peptide with a score above 30 were selected.

References

Chalmers, I.W., McArdle, A.J., Coulson, R.M., Wagner, M.A., Schmid, R., Hirai, H., Hoffmann, K.F., 2008. Developmentally regulated expression, alternative splicing and distinct sub-groupings in members of the *Schistosoma mansoni* venom allergen-like (SmVAL) gene family. BMC Genomics 9, 89.

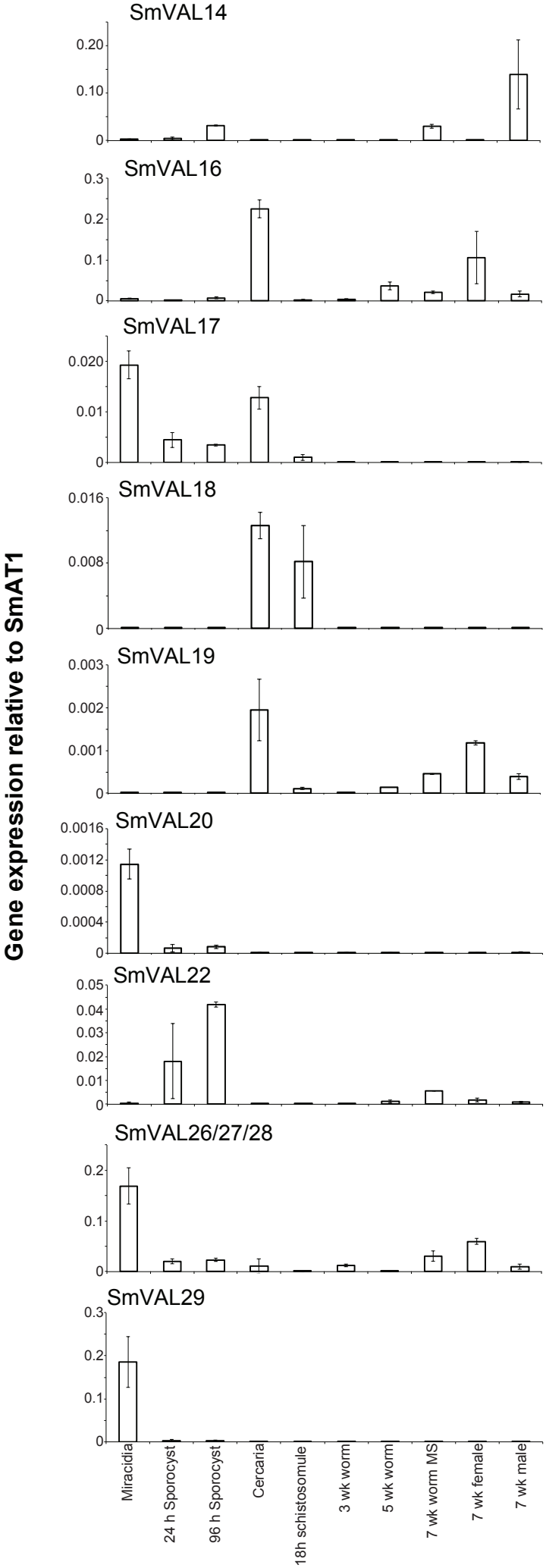
Farias, L.P., Rodrigues, D., Cunna, V., Rofatto, H.K., Faquim-Mauro, E.L., Leite, L.C., 2012. *Schistosoma mansoni* venom allergen like proteins present differential allergic responses in a murine model of airway inflammation. PLoS Negl Trop Dis 6, e1510.

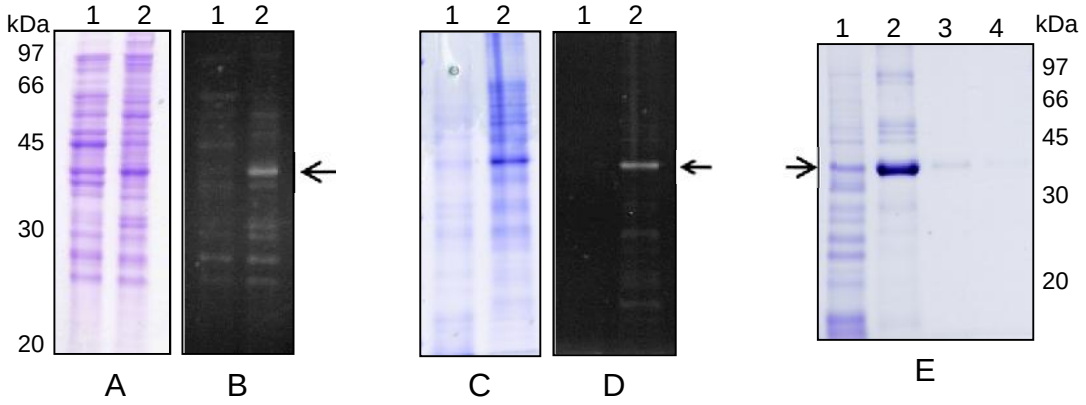
Yoshino, T.P., Brown, M., Wu, X.J., Jackson, C.J., Ocadiz-Ruiz, R., Chalmers, I.W., Kolb, M., Hokke, C.H., Hoffmann, K.F., 2014. Excreted/secreted *Schistosoma mansoni* venom allergen-like 9 (SmVAL9) modulates host extracellular matrix remodelling gene expression. Int J Parasitol 44, 551-563.

Supplementary Table S1. DNA sequences used to clone *Schistosoma mansoni* venom allergen-like (SmVAL) cDNAs in pET30a (+) and pPICZα-A vectors, to quantify transcript abundance of SmVAL14-29 by quantitative reverse transcription PCR (qRT-PCR) and to detect gene SmVAL spatial expression by Whole in-situ Hybridisation (WISH).

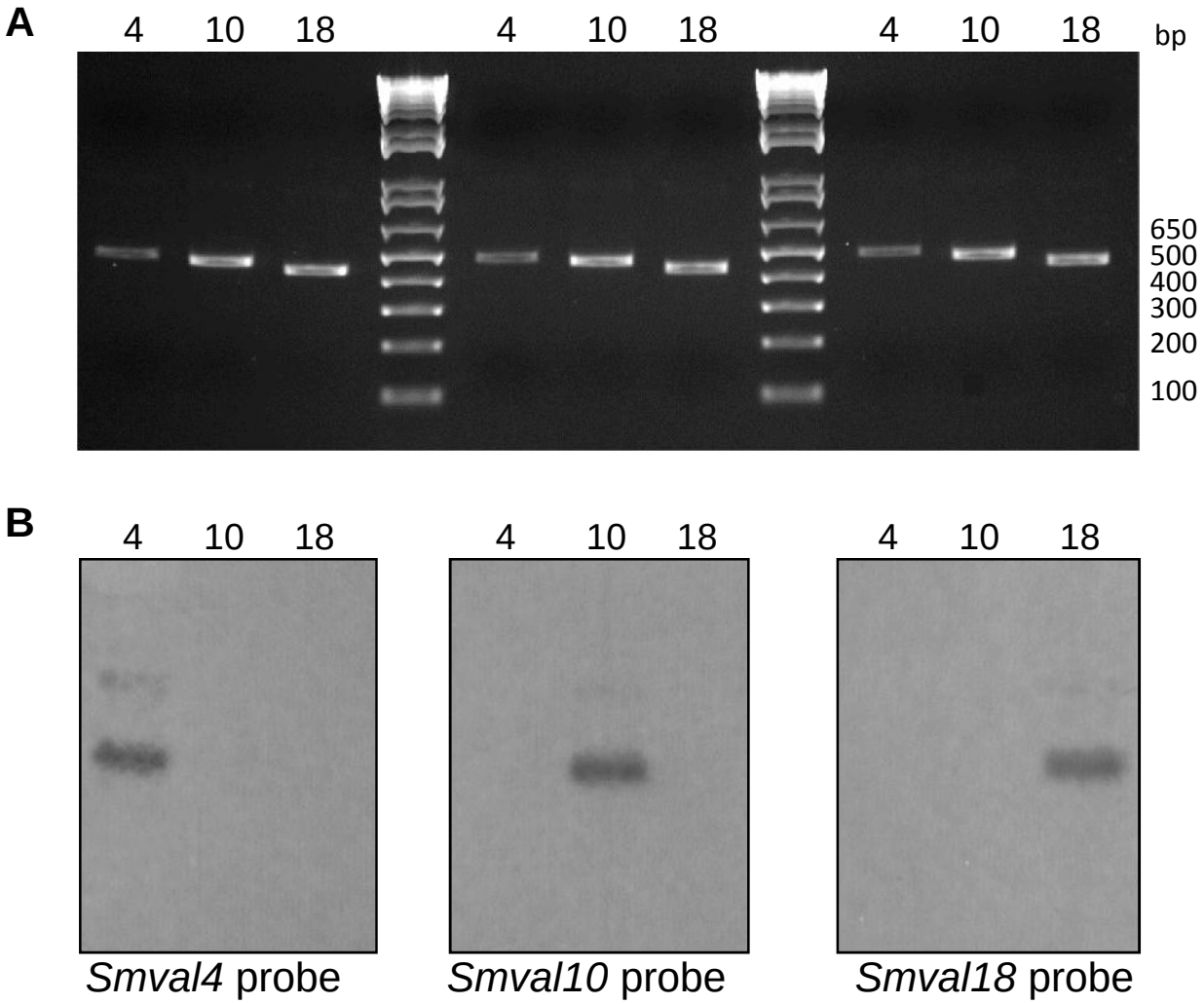
pET30a (+)			
SmVAL	Gene DB	Forward primer (5'-3')	Reverse Primer (5'-3')
4	Smp_002070	TCTAGAATGAAGTTATCGGAAGGACAACGAGC	CTCGAGTTCTGATTTATTACATTTACTTTC
9	Smp_176180	TCTAGAATGAAAATGAATGACACGATTCGTG	CTCGAGTGCAGTCCTATACGGTCTTTGTTC
pPICZ~-A			
5	Smp_179480	ACGAATTCATGGACAATGCCACCAGAGAGAAG	CTTCTAGACTTTCGTTTGTATGATTACACCT
qRT-PCR			
SmVAL	Gene DB	Forward primer (5'-3')	Reverse Primer (5'-3')
14	Smp_078490	CAACCATTAATTCAGCAGTGGA	CCTGGACCATAGTAACACACGA
16	Smp_124070	GTCGTGAAGCAACTCGAAAT	CCTGCTTTAATAGTTGACTTCC
17	Smp_163400	TGCACAAAAATATGCAGAACA	CATCTGCGCCTGTAAAATCT
18	Smp_001890	CCAAAACAACCTCCAGCAAA	CCACCGCTTTTTCAATTGTT
19	Smp_123090	TCATTGTTCGTTTGCACTTCA	TCATTCCAACGCAATACTTCC
20	Smp_127130	CAGCGGATAATTCGGATGTAA	ACAATGTCCAACTGAGCAACC
22	Smp_139450	CAAAATATGGCTTTTGCTTCAA	ACTGTCTTTAGAAATATTTTGCAC
26 27 28	Smp_154260 Smp_154290 Smp_176160	TTCTGTCTGATCAATGTCGT	CGGTTTTGATGTCTGTCGC
29	Smp_120670	ACTGGGCTCTTACACCTCGAT	CTCCAGGACCGTAATTGCATA
Templates for RNA probe synthesis			
1	aacggtggattagatgcgaaaagtgaagagttgctcaatctgcatagaaaatacagacaagatttagttgattgtaaagttgatg gacaacctccagctaaatatatgtcaccattgaagtggaatcacgatttagctcgacaagcacaatcattggcaaacaaatgcat cttacgtcacgacaaacgacattcgaatcaattcagttgggttggacaaaatatagctctccatccaaccattaagtcaggagtg gatgcttggttcaatgaacacaaattatacaattacaatacgaacaactgtcctcaatgtttgcattacacacaaatggcttggg ccaagaccacagacattggatgtggagttgcgaattgtccacggtacggtctatcaatcgtgtgtaattacggtccagggggtaa ttggaataatgagaaaccatatgaaatgaaaccacgtaatttgtgcccaaaagtgcaaaatattcctaaaaacagtttacaaact aacagtgctcatacacaacatggaccgaagccattaacgcaaagcgaagaaagggtgtcaacgaacgttcaaaatggacgtagag aatgcaatcaacgcgaacagttgagaagtagatattaa		
4	aagttatcggaaggacaacgagcaatctataatttccataaaaaagttcgtaaagatgtaaaaaattgcaggatacctggtcaac ctccagctaaaaatctaacaaagttgaaatggaataaactattagctaataaagctaaacaacaagctaaacgatgtaaatatga ttcaaatgatccaaatgattttattattggggattttgaatcaattggacaaaatttagccgattatccaacaattgaaggtgca atgaaagattggttagaagagtataaaaattataattttgaaaagaatcaatgtaatggtgattgtaaaaattataaacagatgg tttggaataccactgaagaaataggctgcggttatgaaaaatgtggaaagaactatttgattgtttgcaattatgcaccagggga ctctgaagatagaccatatgaagctaaaccagaaagtaaatgtaataaatcagaataa		
10	gataagtcatctacgaaagagttgattttcaattttcataataaaattcgtgaagatgtatttaaaggtgttttatctggtcaac ctaaagcgaaaaagatgtccaaactgaaatggaataaacttttagctaaattagctaaaggtcacgttcagaaatgtattcttga tagtggagatcttggaaaattatacgttggtaaatttgattctgtcggacaaactgtagcagaacatacttcaatacaaaacata ttagatacatggttagaagaaaaaaatgattacgatttagacaagaatacatgtgaaaatgaatgtggaaactataaacaattag tgtgggctaacacaacggatattggatgtgcttcaaataaatgtggtaacaggtatatggtcgtatgtaattacgctccaggggc tgatgatgaaaggccatatgaaaaagactcataa		
18	ccaaaacaacctccagcaaaatatatgtcaaaattacaatggaataaacatttagcagaaaaagctcaactcacagctagccgat gtgattactcttatgatagtccaagtgatatgcgttttgaagagttttcttcagtggcacaaaatattgctgatagtccaacaat tgaaaaagcggtggccagttggtttattgagtacaagaactattcttttgatgataacacatgcaaagatacatgtatgcagtat aaacagatggtaaaaggtgaagaaactgaaattggttgtggtgtacaaaaatgttcaaacagatttctagtagtatgtaattatt caccagcagctgaagaagataaacaaccgtatgaaaaaggtactcaagaaaattgcgatgatgttgatgatgcagaatattaa		
19	ctcgagatgaataagaatcatcccttgcattttgttatcttttcattgttcgtttgcacttcaaatggtgctaagaagaatctaa cacaaaaattagaaactatacatcgtttacatacatattatagaaattctattctactttgtaaagttccaacacaacctccagc tgaagatatggaagtattgcgttggaatgatattttagccaataatgctcaacaagtagcaaataaatgtgatttaaattttgat ttagtcaatgataaactactggaacattttgaatcgattggacaaaatgtagctgaatcggatacaattaaaaatgcaatggaaa attggttccgagaatatcataattat		

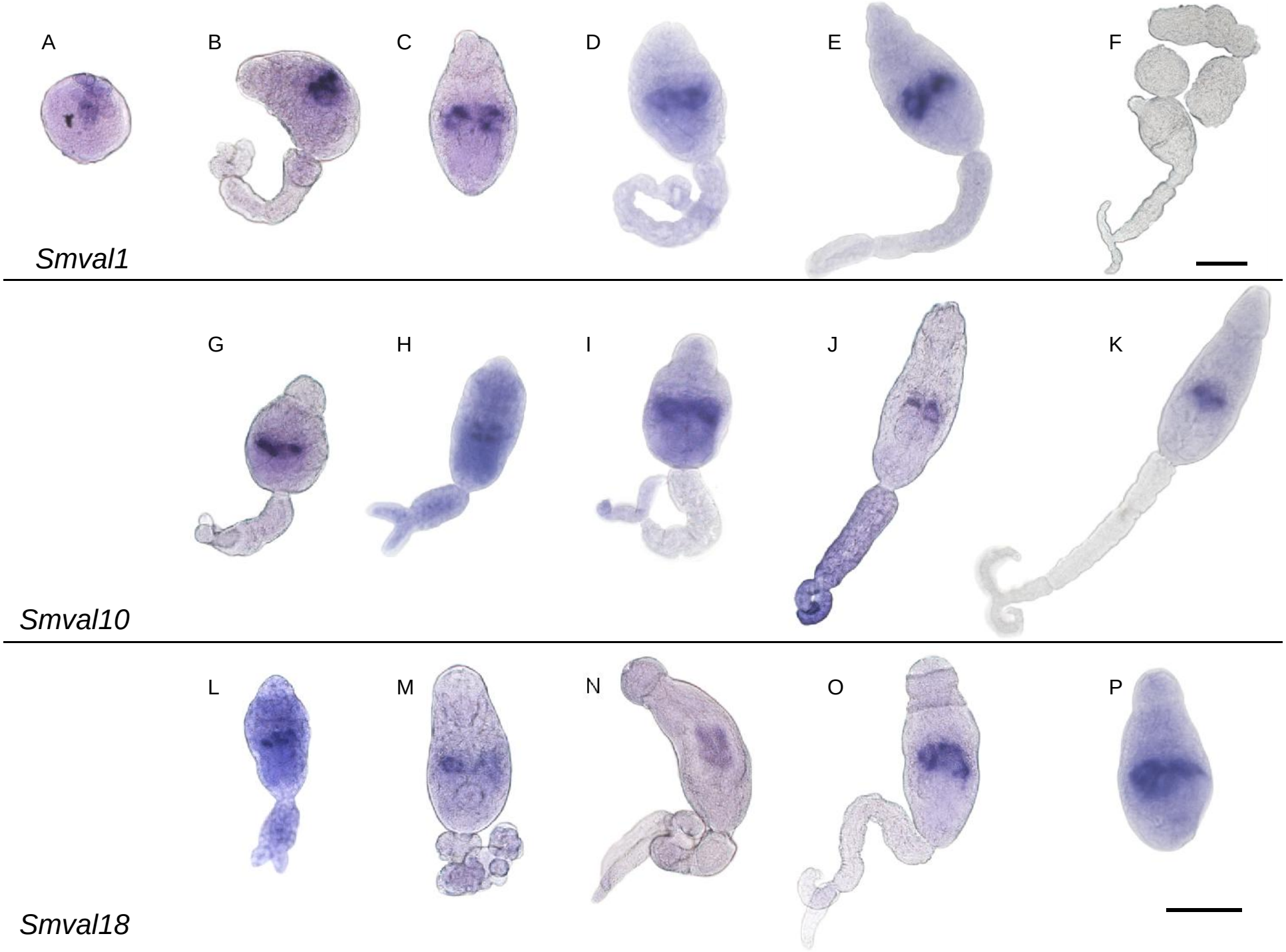
Restriction sites are underlined: *Xba*I (forward) and *Xho*I (reverse) for SmVAL4 and SmVAL9, and *EcoR*I (forward) and *Xba*I (reverse) for SmVAL5.


SmVAL14
0.20
0.10
0
SmVAL16
0.3
0.2
0.1
0
SmVAL17
0.020
0.01
0
SmVAL18
0.016
0.008
0
SmVAL19
0.003
0.002
0.001
0
SmVAL20
0.0016
0.0012
0.0008
0.0004
0
SmVAL22
0.05
0.04
0.03
0.02
0.01
0
SmVAL26/27/28
0.2
0.1
0
SmVAL29
0.3
0.2
0.1
0
Gene expression relative to SmAT1
Miracidia
24 h Sporocyst
96 h Sporocyst
Cercaria
18h schistosomule
3 wk worm
5 wk worm
7 wk worm MS
7 wk female
7 wk male



Supplementary Figure S2


A
4
10
18
4
10
18
4
10
18
bp
650
500
400
300
200
100
B
4
10
18
4
10
18
4
10
18
Smval4 probe
Smval10 probe
Smval18 probe


A
B
C
D
E
F
Smval1
G
H
I
J
K
Smval10
L
M
N
O
P
Smval18