

RESEARCH

Open Access



# Evolution of tetraspanin antigens in the zoonotic Asian blood fluke *Schistosoma japonicum*

Daniel A. J. Parsons<sup>1,2</sup>, Anthony J. Walker<sup>1</sup>, Aidan M. Emery<sup>2</sup>, Joanne P. Webster<sup>3</sup> and Scott P. Lawton<sup>4\*</sup>

## Abstract

**Background** Despite successful control efforts in China over the past 60 years, zoonotic schistosomiasis caused by *Schistosoma japonicum* remains a threat with transmission ongoing and the risk of localised resurgences prompting calls for a novel integrated control strategy, with an anti-schistosome vaccine as a core element. Anti-schistosome vaccine development and immunisation attempts in non-human mammalian host species, intended to interrupt transmission, and utilising various antigen targets, have yielded mixed success, with some studies highlighting variation in schistosome antigen coding genes (ACGs) as possible confounders of vaccine efficacy. Thus, robust selection of target ACGs, including assessment of their genetic diversity and antigenic variability, is paramount. Tetraspanins (TSPs), a family of tegument-surface antigens in schistosomes, interact directly with the host's immune system and are promising vaccine candidates. Here, for the first time to our knowledge, diversity in *S. japonicum* TSPs (SjTSPs) and the impact of diversifying selection and sequence variation on immunogenicity in these proteins were evaluated.

**Methods** SjTSP sequences, representing parasite populations from seven provinces across China, were gathered by baiting published short-read NGS data and were analysed using in silico methods to measure sequence variation and selection pressures and predict the impact of selection on variation in antigen protein structure, function and antigenic propensity.

**Results** Here, 27 SjTSPs were identified across three subfamilies, highlighting the diversity of TSPs in *S. japonicum*. Considerable variation was demonstrated for several SjTSPs between geographical regions/provinces, revealing that episodic, diversifying positive selection pressures promote amino acid variation/variability in the large extracellular loop (LEL) domain of certain SjTSPs. Accumulating polymorphisms in the LEL domain of SjTSP-2, -8 and -23 led to altered structural, functional and antibody binding characteristics, which are predicted to impact antibody recognition and possibly blunt the host's ability to respond to infection. Such changes, therefore, appear to represent a mechanism utilised by *S. japonicum* to evade the host's immune system.

**Conclusion** Whilst the genetic and antigenic geographic variability observed amongst certain SjTSPs could present challenges to vaccine development, here we demonstrate conservation amongst SjTSP-1, -13 and -14, revealing their likely improved utility as efficacious vaccine candidates. Importantly, our data highlight that robust evaluation of vaccine target variability in natural parasite populations should be a prerequisite for anti-schistosome vaccine development.

\*Correspondence:

Scott P. Lawton

Scott.Lawton@sruc.ac.uk

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

**Keywords** *Schistosoma japonicum*, China, Tetraspanin, Vaccine candidate, Selection pressures, Immune evasion, Antigenic variability

## Background

Across Southeast Asia, zoonotic schistosomiasis, caused by *Schistosoma japonicum*, is responsible for infection of humans, as well as ~40 species of wild and domesticated animals, where animal hosts play an important role in maintaining parasite transmission [1, 2]. The recent Healthy China 2030 strategic plan outlines a revised schistosomiasis elimination target of 2030, in line with the recent WHO NTD roadmap [3, 4]. However, despite well-concerted, multi-disciplinary control efforts over the last 60 years [5, 6], which have resulted in a reduction in human infections from > 11 million people to ~30,000, as well as a similar reduction in livestock infections [7, 8], the risk of resurgence remains ever present in China and other schistosome-endemic SE Asian countries and is expected to increase, in part due to persistent hotspots of infection, the perpetual presence of refugia, and anthropogenic changes to the climate and environment [9–12].

The development of an effective anti-schistosome vaccine, which has recently seen renewed interest globally, would provide an essential tool in moving towards end-game elimination in China. Concerns of reduced praziquantel (PZQ) susceptibility would be mitigated by implementing a vaccine in current control strategies, complementing and reducing reliance on frequent PZQ mass drug administration (MDA) and thereby simultaneously suppressing the risk of parasites evolving PZQ resistance [13–15], thus serving to reduce the possibility of future resurgences [10, 16–18].

To date, more than 100 vaccine candidate antigens have been identified and tested against various schistosome species within animal models [17]. A protein family of particular interest, the transmembrane, tegument- and oesophageal-associated tetraspanins (TSPs), have been shown to elicit variable levels of immune response in animal models, including murine and bovine definitive hosts [19–22].

Of the known *S. japonicum* TSPs (SjTSPs), several have been shown to interact directly with host immune components via their large extracellular loop (LEL) domains [23–26], with a number of TSPs shown to be crucial in tegument development, maturation and stability [27], and which are also involved in host immune evasion, transition to parasitism and survival in the definitive host [23, 24, 28–31]. Several TSPs have been suggested as anti-schistosome vaccine candidates, particularly SjTSP-2 and -23 [17, 23, 25, 29–32]. However, despite the immunological properties of SjTSPs supporting their use

as vaccine targets, only partial protection and variable efficacy have been demonstrated in animal-based SjTSP immunisation studies [19, 20, 30].

The TSP-LEL domain is highly polymorphic, particularly in the case of SjTSP-2, with such ‘hypervariability’ suggested to underly the variable vaccine efficacy [24, 26, 30, 33, 34]. Indeed, the polymorphic nature and genetic diversity of TSP-2 and TSP-23, demonstrated within and between schistosome species, have been suggested to limit their utility as vaccine targets [19, 24, 25, 30, 32, 35]. Schistosomes express multiple TSP variants through alternative splicing and allelic variation [24, 30, 34, 36], and hypervariability within the TSP-LEL domain is suggested to reflect the selective influence of the definitive host, resulting in antigenic variation [33].

Elucidating how antigen diversity and evolution influence the antigenic propensity of SjTSP proteins remains necessary to properly evaluate their utility as vaccine targets [32–34]; however, these aspects of SjTSP biology have not been explored. Here, by employing in silico approaches using TSPs retrieved from published *S. japonicum* genomic datasets, this study aimed to identify putative SjTSPs and determine their diversity and evolutionary history in *S. japonicum*. In particular, the geographic variability of SjTSP sequence from seven provinces across China was assessed, with the aim of identifying which members of the antigen family were most variable, localising particularly variable SjTSP regions and categorising selection pressures driving the measured SjTSP geographical sequence variation. The results obtained should have significant theoretical and applied implications, particularly regarding the identification of future anti-schistosome vaccine targets.

## Methods

### SjTSP sequence retrieval, characterisation and phylogenetic analysis

Owing to the paucity of available and accurately described tetraspanin sequence data from *S. japonicum*, the full SjTSP protein family remains poorly resolved. Thirty-three putative SjTSP mRNA sequences were retrieved from the *S. japonicum* genome assembly on WormBase ParaSite [37], and a further 29 were identified from literature sources (Table 1) [29, 32, 38]. Where coding sequences (cds) were absent from the literature and online databases, whole gene and protein sequences were queried in GeneWise [39] to retrieve the related cds.

**Table 1** Tetraspanin sequence sources and related domain information

Tetraspanin	Accession number	Amino acid residue positions			Source
		Total	SEL	LEL	
TSP-1	CAX70118.1	247	36–49	107–202	[85]
TSP-2	AEG74374.1	215	35–53	106–183	[30]
TSP-8	AAW24708.1	224	36–54	107–189	[86]
TSP-13	CAX69650.1	233	35–43	93–194	[85]
TSP-14	CAX69587.1	270	32–72	125–235	[85]
TSP-23	CAX72085.1	218	36–49	108–183	[85]
TSP-25	AY814252.1	287	101–115	166–252	[86]

SEL small extracellular loop, LEL large extracellular loop

Putative TSP protein sequences were validated using TMHMM 2.0 to confirm their membrane association, SMART to predict domain architecture (presence of the two extracellular loop domains [26] and PROTTTER to assess protein membrane topology [40–42]. From these structurally validated SjtTSPs, seven were selected for detailed analysis (Table 1) based upon an analysis of published information pertaining to each SjtTSPs structure, function, immunogenicity and vaccine candidacy [20, 25, 26, 32, 38].

Structurally validated SjtTSP protein sequences were aligned using the MEGA7-integrated MUSCLE tool [43, 44], and phylogenetic reconstruction was done using the distance-based neighbour joining (NJ) method [45] to assess the relationships between members of the *S. japonicum* tetraspanin family. Nodal support was determined using 1000 bootstrap replicates. Amino acid positions containing gaps and missing data were omitted from analysis.

#### Genetic variation of SjtTSPs and identification of selection pressures

The seven SjtTSP reference sequences (Table 1) were used to bait short-read NGS data, sequenced from ten pairs of adult worms from each of the seven sampled provinces across China (Yunnan, Sichuan, Hunan, Hubei, Jiangxi, Anhui and Zhejiang) (Additional File 1: Fig. S1) by Young et al. [34]. Short-read NGS data were reported by Young et al. [34] to have achieved up to 50-fold coverage of the *S. japonicum* reference genome for each of the seven sampled parasite populations, considered to be sufficient for draft genome assembly, and the identification of putative protein coding genes for population and phylogenetic analysis. Using a similar approach as seen in [46, 47], sequence read archive (SRA)-BLAST was employed to retrieve the complete cds for each provincial SjtTSP variant [48]. SjtTSP SRA-collected sequences were

aligned against SjtTSP reference sequence and assembled using the Contig Assembly Program 3 (CAP3) feature in Bioedit [49, 50]. Finally, using MUSCLE, multiple sequence alignments were constructed representing parasite isolates from each of the Chinese provinces for each SjtTSP.

Phylogeographic relationships, also analysed using MEGA7, were investigated using provincial SjtTSP cds alignments, inferred using the NJ method and 500 bootstrap replicates. *Schistosoma mansoni* and *Schistosoma haematobium* TSP orthologues were used to root each phylogeny. Orthologous Sm- and ShTSPs were gathered through a BLASTn similarity search [51] of each SjtTSP against the NCBI nucleotide collection, filtering for hits against each *Schistosoma* species.

Measures of genetic diversity were calculated using DNASP v6.12 [52] to gain insight into the geographic variation of SjtTSPs between distinct *S. japonicum* populations across China, such as the number of polymorphic sites, number of haplotypes (h), haplotype diversity (Hd) and nucleotide diversity ( $\pi$ ).

To detect the presence and extent of selection acting on the SjtTSPs, non-synonymous (dN) and synonymous (dS) substitutions, averaged across all codons in the alignment ( $\omega$ ) for the whole cds and LEL domain encoding regions, were determined using MEGA7. Conventional thresholds for inferring the occurrence and extent of selection were used (positive selection,  $\omega > 1$ ; purifying selection,  $\omega < 1$ ; neutrality,  $\omega = 1$ ). The effect of selection on protein-coding genes is often difficult to identify, as traditional models, such as  $\omega$ , predict selection across all sites in an alignment and thus lack statistical power, as selection may only occur at a small number of sites [53, 54]. To define individual nucleotide sites under selection, and thus increase the resolution of the  $\omega$  analysis, the Mixed Effects Model of Evolution (MEME) [55] and Fast, Unconstrained Bayesian AppRoximation (FUBAR) [54] analyses were performed via Datamonkey [56]. MEME applies a mixed-effects maximum-likelihood approach to test for episodic and diversifying selection; *P*-value thresholds were set to 0.5. FUBAR employs a Bayesian approach to infer positive or purifying selection at each site across the SjtTSP alignment. As default, posterior probability value  $> 0.9$  strongly indicates the presence of positive or purifying selection at a site.

#### Determining the effects of sequence variation on protein structure and function

Using BioEdit, SjtTSP sequences from each of the seven provinces were translated into proteins to determine the impact of accumulating polymorphisms on each SjtTSP amino acid sequence. Amino acid changes and B-cell antibody binding sites (epitopes), predicted using

BepiPred-2.0 [57], were mapped onto SjTSP protein secondary structures and visualised using PROTTTER.

To generate 3D protein models of each SjTSP, Protein Data Bank (PDB) models were generated from reference sequences using I-TASSER [58]. Confidence scores (c scores) estimated the quality of the models ( $> -5$  = high confidence). Subsequently, PyMOL [59] was used to render 3D structures of the I-TASSER-produced SjTSP reference PDB files, enabling comprehensive visualisation of key domains, antibody binding sites and sequence variation in relation to protein tertiary structure.

Structural and conformational variation between provincial SjTSPs was assessed by comparing protein structures within DALI [60]. DALI  $z$  scores that define structural homology between provincial SjTSP 3D protein models were then used to evaluate the evolutionary divergence and clustering between the proteins using principal component analyses (PCAs).

To determine whether amino acid substitutions impacted the function of each SjTSP protein, SjTSP reference sequences and site location information of amino acid substitutions were uploaded into PROVEAN [61], generating a PROVEAN score. A score  $< -2.5$  was used to predict whether an amino acid substitution alters SjTSP function at that individual site. TreeSAAP v3.2 was then used to further investigate selection and determine whether amino acid changes significantly altered the physicochemical properties of each SjTSP [62], producing  $z$  scores to establish the significance of amino acid replacement, type of deviation from neutrality and ranking the magnitude of changes in 31 physicochemical amino acid properties [63, 64]. Eight magnitude categories were analysed, with higher categories (5–8) denoting the most radical changes in protein biochemistry. Significantly positive or negative  $z$  scores ( $> \pm 1.645 = P < 0.05$ ) indicate the influence of positive or purifying selection, respectively [64]. Significant and radical changes were considered evidence of positive selection resulting in adaptation (magnitude categories 5–8,  $z$  scores  $> 3.09 = P < 0.001$ ).

#### Evaluating the impact of selection and provincial SjTSP variation on antigenicity and antibody binding affinity

SjTSP protein sequences were compared between provinces to determine the association between sequence diversity and antigenicity, using the Kolaskar and Tongaonkar (KT) method in the IEDB Antibody Epitope prediction tool [65]. The KT method is one of the most accurate, robust and widely used predictors of antigenicity of proteins. It applies an accurate semi-empirical approach by utilising the physicochemical properties of amino acid residues and relative frequency of experimentally determined epitopes to predict

antigenic determinants at each site directly from a protein sequence [66].

To elucidate effects of amino acid substitutions on antibody binding affinity in each SjTSP, mCSM-Ab was employed [67]. Utilising SjTSP 3D protein models and amino acid substitutions, mCSM-Ab employs a graph-based machine-learning approach to infer changes in structural signatures from each residue in the SjTSP model that resulted from amino acid variation. Predictions of the difference in Gibbs free energy ( $\Delta\Delta G$  in Kcal/mol) between reference SjTSP 3D protein models and provincial SjTSPs with amino acid substitutions were established to accurately assess the effects of mutations on the propensity for antibody binding.

## Results

### Phylogenetic relationship of TSP proteins in *Schistosoma japonicum*

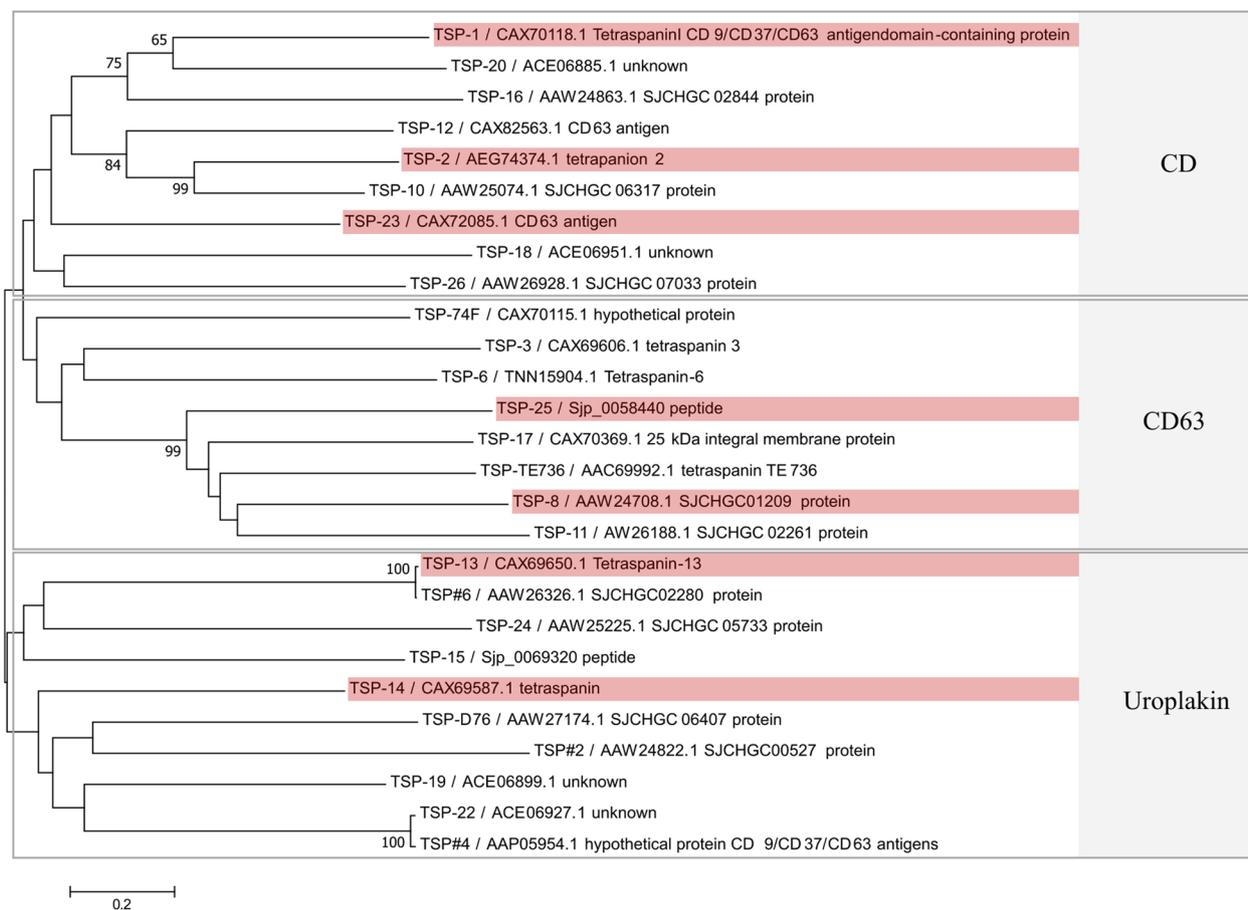
From 62 sequences annotated as SjTSPs, 27 were structurally validated as TSPs and were thus included for phylogenetic analysis. Phylogenetic reconstruction revealed that the TSP family in *S. japonicum* is highly divergent, with SjTSPs forming three monophyletic clades (Fig. 1). The clades resolved here correspond to the CD, CD63 and Uroplakin groupings, as described in other organisms, and in *S. japonicum* [29]. Deeper ancestral nodes, however, had low bootstrap support, providing uncertainty in the deeper tree topology. Only seven nodes demonstrated bootstrap support values  $> 50$  after 1000 iterations. Clustering of SjTSP proteins in the CD subfamily was best supported, with four of the seven well-supported nodes evident in this clade. The short branch lengths and strong bootstrap support demonstrated in Fig. 1 for SjTSP-13 and TSP-6, and SjTSP-22 and TSP-4, suggest these SjTSPs are the same proteins.

### Genetic differentiation and diversity of tetraspanin proteins between *Schistosoma japonicum* populations

The seven SjTSPs selected for phylogeographic analysis and assessment of genetic and antigen geographic variation were primarily chosen based on their recognition as potentially protective antigens and therefore as promising vaccine targets.

Within all phylogeographic analyses, deeper ancestral nodes were generally well supported by bootstrap values (Fig. 2), supporting the resolved branching. Reconstructed phylogeographic relationships mostly divide SjTSP sequences into two distinct clades, other than for SjTSP-13; however, there was no consistent pattern based upon geographical distance between provinces (Fig. 2).

Nucleotide diversity ( $\pi$ ) within *S. japonicum* populations was lowest for SjTSP-1, 13 and -14 (Table 2), and no allelic variants were found in the SRA data. In contrast,



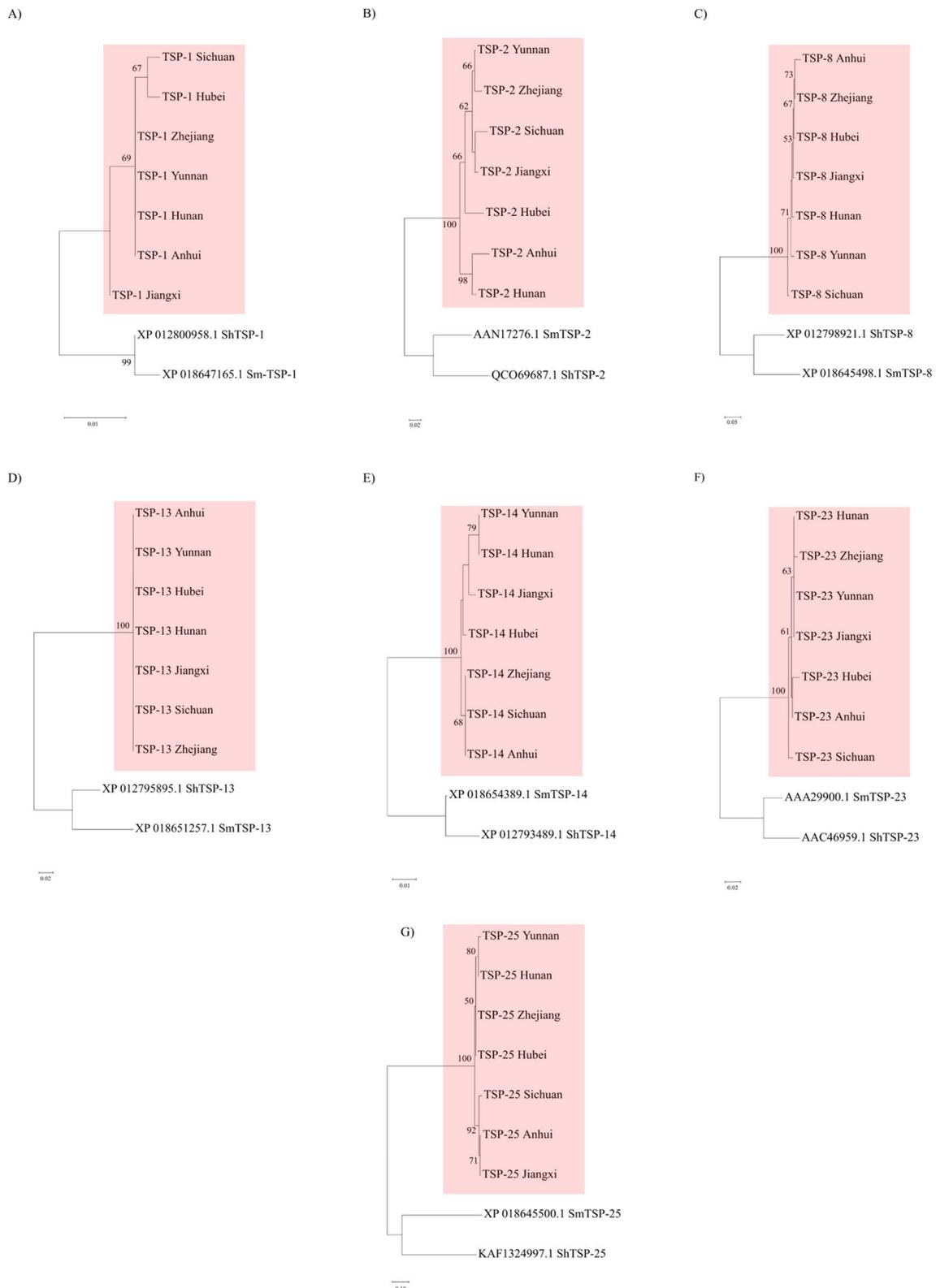
**Fig. 1** Phylogenetic analysis of 27 protein sequences from the tetraspanin (TSP) protein family in *Schistosoma japonicum*. Evolutionary history was inferred using the neighbour-joining method, with 1000 bootstrap replicates, using MEGA7. Bootstrap values with > 50% support are displayed at each node. Three major SjtSP subfamily clades are highlighted in grey; highlighted sequences (red) were utilised for further analyses

SjtTSP-2, -25 and -8 were the most genetically diverse across the seven provinces. The varied  $\pi$  is also reflected in the haplotype diversity (Hd) for the seven STSPs, which ranged between 0.29 and 1. SjtTSP-2 and -8 both displayed a Hd of 1, indicating each provincial sequence is a unique haplotype variant. SjtTSP-13 demonstrated the lowest Hd compared to the other SjtTSPs, with only two haplotypes shared across all seven analysed provinces (Table 2). Despite demonstrating one of the lowest  $\pi$  between provinces, SjtTSP-1 had a Hd of 0.95, with six unique haplotypes that differed in only a few amino acids (Table 2). Between four (SjtTSP-13) and 67 (SjtTSP-2) polymorphic nucleotide sites were discovered across each SjtTSP population. As such, SjtTSP-13, -1 and -14 were shown to be relatively invariable between the seven provinces compared to the other SjtTSPs (Table 2).

**Signatures of selection in SjtTSPs**

Substantial sequence variation was demonstrated within several SjtTSPs from different provinces (Table 2);

therefore, the impact of selection pressures was investigated for each SjtTSP using  $\omega$ . Mean dS and dN polymorphisms across the whole coding sequence (cds) were greatest for SjtTSP-2, at 0.228 and 0.102, respectively (Fig. 3A), and were consistent with the far greater number of polymorphic sites detected in SjtTSP-2 than the other SjtTSPs (Table 2). SjtTSP-25 has the second greatest dN and dS, both across the whole cds and LEL (Fig. 3), reflecting the number of detected polymorphic sites between provinces (Table 2). When assessing the whole cds, dS across SjtTSP-1 and 2 is substantially higher than dN (Fig. 3A). In contrast, all other SjtTSPs had higher dN than dS across each of their coding sequences. SjtTSP-8, in particular, had >twofold higher dN compared to dS across the whole cds. When considering only the region encoding the LEL domain, SjtTSP-2 also displayed the greatest dN and dS, where dN was more than twofold greater than dS (Fig. 3B), and relative to the remaining SjtTSPs, dN and dS values were comparatively low. Overall, SjtTSP-1, -2, -8 and -25 revealed an increase in



**Fig. 2** Phylogeographic analyses of provincial *Schistosoma japonicum* protein isolates (from Yunnan, Sichuan, Jiangxi, Zhejiang, Hunan, Hubei and Anhui) from seven TSP proteins: **A** SjtSP-2, **B** SjtSP-8, **C** SjtSP-23, **D** SjtSP-1, **E** SjtSP-13, **F** SjtSP-14, **G** SjtSP-25. Phylogenetic relationships were determined using the neighbour-joining method, with 500 bootstrap replicates, using MEGA7. Only bootstrap values with > 50% support are displayed at each node

**Table 2** Genetic diversity measurements for seven *Schistosoma japonicum* tetraspanin (TSP) metapopulations

Tetraspanin	Polymorphic sites	Nucleotide diversity ( $\pi$ )	$\pi$ S.D	Number of haplotypes	Haplotype diversity (Hd)
SjTSP-1	5	0.0027	0.00056	6	0.95
SjTSP-2	67	0.0425	0.00936	7	1.0
SjTSP-8	19	0.0099	0.00212	7	1.0
SjTSP-13	4	0.0016	0.00112	2	0.29
SjTSP-14	5	0.0034	0.00057	4	0.81
SjTSP-23	8	0.0044	0.00096	5	0.86
SjTSP-25	27	0.0137	0.00226	6	0.95

**Table 3** Signatures of selection, measured as mean dN/dS ( $\omega$ ) across all codons of the whole coding sequence and the region encoding the large extracellular loop (LEL) domain, within *Schistosoma japonicum* tetraspanins

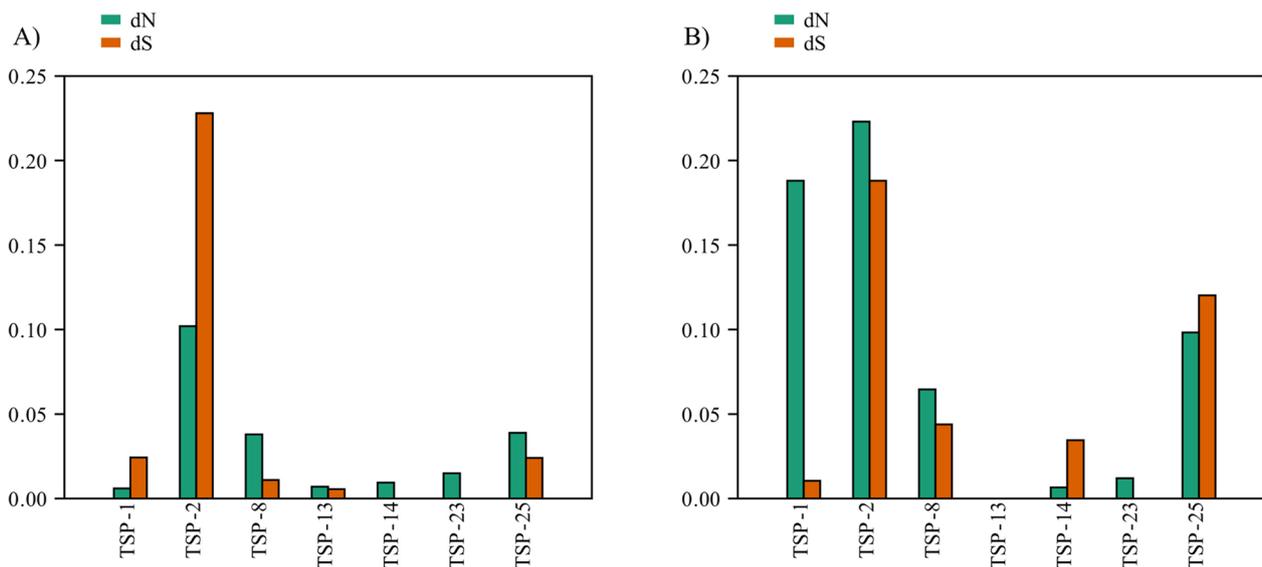
Tetraspanin	Whole coding sequence	Region encoding large extracellular loop (LEL) domain
SjTSP-1	0.247	0
SjTSP-2	0.447	1.186
SjTSP-8	3.455	1.472
SjTSP-13	1.25	0
SjTSP-14	0.009	0.192
SjTSP-23	0.015	0.012
SjTSP-25	1.621	0.818

dN, to varying degrees, when solely analysing the LEL cds (Fig. 3B), indicating the majority of nucleotide base changes that result in changes to the protein are accumulating in this important region of the gene.

Comparing the type and extent of selection for each SjTSP, using the  $\omega$  averaged across all codons, analysing the LEL only or across the whole cds, considerable variation in  $\omega$  was identified, both within and between SjTSPs (Table 3). The  $\omega$  analysis indicates positive selection is acting on the LEL of SjTSP-2 and -8. Conversely, average  $\omega$  values across the whole cds of SjTSP-2 revealed the presence of purifying selection, suggesting that positive selective pressures are disproportionately acting on the LEL domain. In contrast, SjTSP-25 showed signatures of positive selection across the whole cds, and despite

increased dN and dS across the LEL, purifying selection was detected. Interestingly, SjTSP-8 displayed positive selection when considering both the whole cds and that of the LEL, although  $\omega$  for the whole cds was ~2.3-fold greater than that of the LEL, suggesting that selection pressures have had a greater impact on regions outside the LEL domain in SjTSP-8.

MEME and FUBAR were utilised to determine specific sites where selection may be acting (Additional File 2: Table S1). For SjTSP-2, MEME detected 21 sites under positive selection, 17 of which were within the LEL. FUBAR detected six sites under positive selection in SjTSP-2 (four in the LEL) and 13 under purifying



**Fig. 3** Signatures of selection acting upon seven tetraspanin (TSP) genes in *Schistosoma japonicum*. **A** The mean distribution of non-synonymous (dN) and synonymous (dS) mutations per TSP, averaged among all codons across the whole gene, or **B** among all codons within the region encoding the large extracellular loop (LEL) domain only. Mutations were estimated across all codons using DNASP and were plotted using PAST [82]

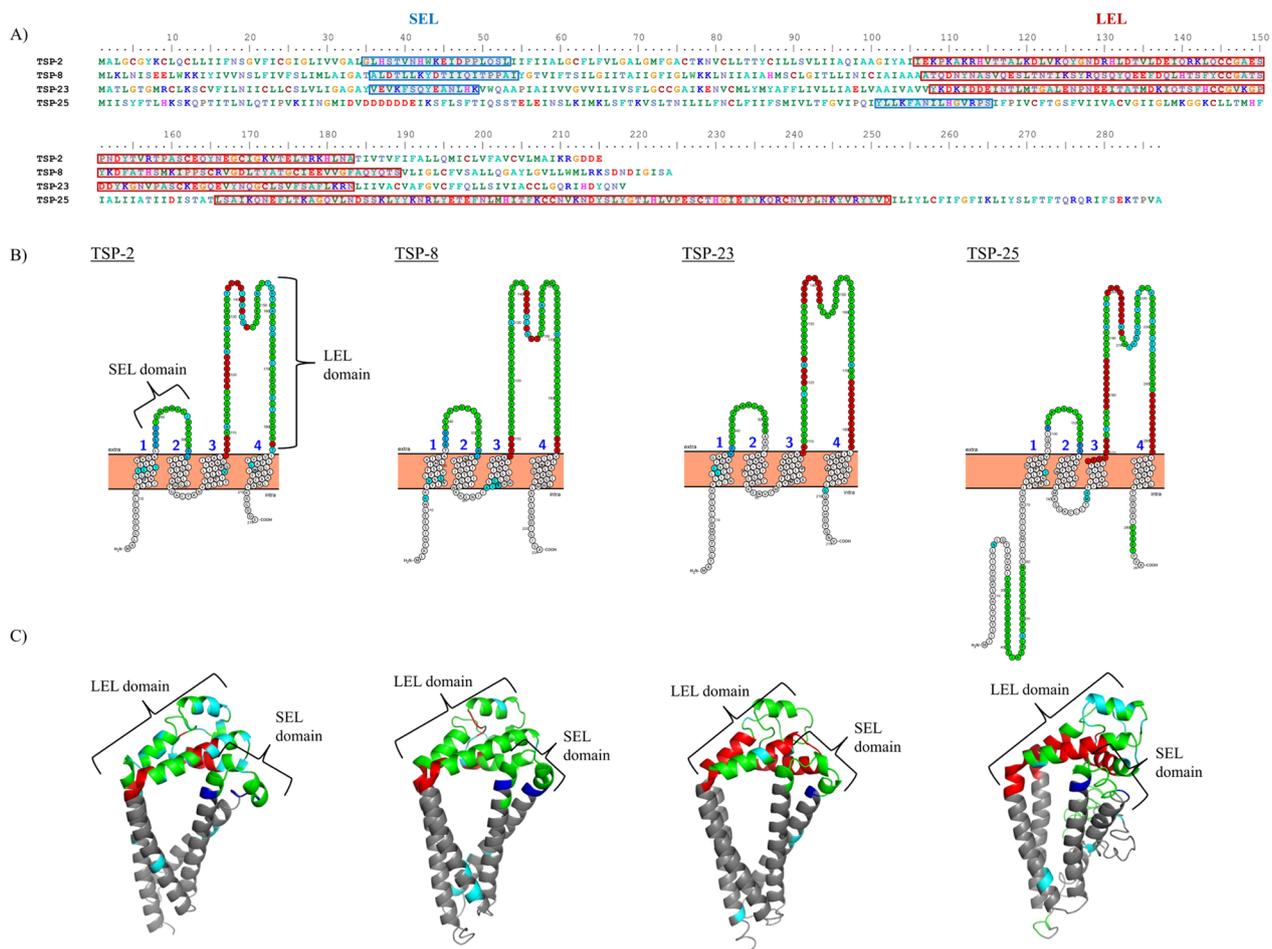
selection (five with in the LEL). MEME and FUBAR both predicted positive selection at four sites (124, 128, 143 and 158), all of which reside within the LEL of SjtTSP-2. Analysis of SjtTSP-8 revealed seven sites under positive selection, three of which were within the LEL, whereas FUBAR predicted three sites under positive and one under purifying selection. Agreement was seen for three sites (84, 130 and 155), two of which are in the LEL. For SjtTSP-23, six (MEME) and one site(s) (FUBAR) were considered under positive selection, with nucleotide site 169, contained within the LEL, called by both approaches. FUBAR failed to detect any purifying selection in SjtTSP-23.

**In silico analysis of SjtTSP protein structure**

All analysed SjtTSP proteins contained short cytoplasmic tails at the C- and N-termini, a small extracellular loop (SEL) domain, a LEL domain, four transmembrane

helices and (with the exception of SjtTSP-25) the conserved -CCG- (-Cys-Cys-Gly-) motif characteristic of the antigen family (Fig. 4). All seven 3D SjtTSP protein models had c scores  $\geq -2.13$  (SjtTSP-25), providing confidence in the I-TASSER-constructed SjtTSP protein models. SjtTSPs ranged from between 215 and 287 amino acids in length, with the SEL domains between 13 (SjtTSP-1 and SjtTSP-23) and 32 (SjtTSP-14) and LEL domains between 77 (SjtTSP-2) and 110 (SjtTSP-14) amino acids (Table 1). According to BepiPred, several predicted B-cell epitopes were <20 amino acids and thus of inadequate length to bind antibodies. All viable B-cell epitopes were predicted to exist in the extracellular loop domains (Fig. 4).

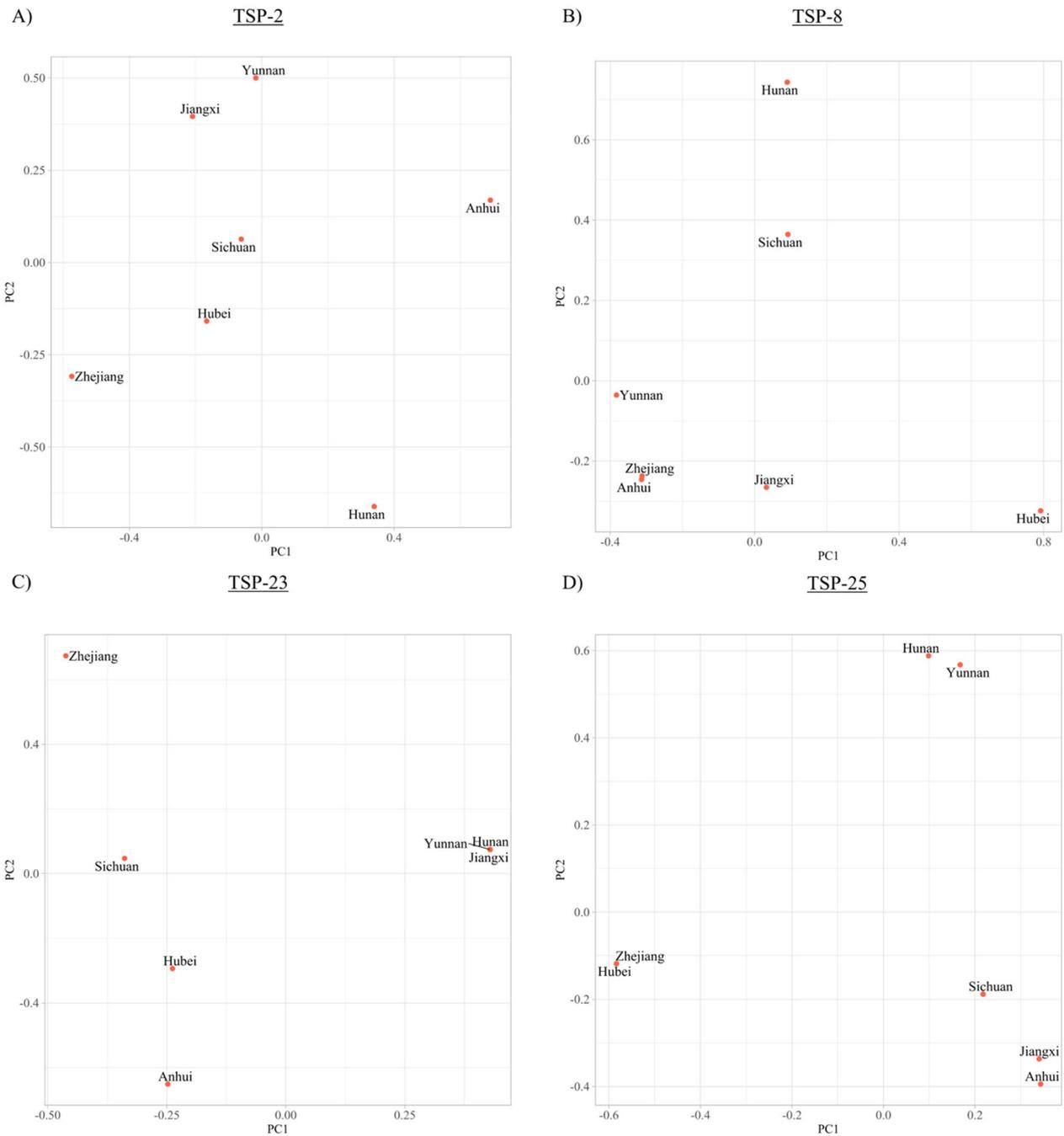
Polymorphisms in the provincial SjtTSPs translated to three amino acid substitutions in SjtTSP-1, 34 in SjtTSP-2, 17 in SjtTSP-8, two in SjtTSP-13, four in SjtTSP-14, six in SjtTSP-23 and 21 in SjtTSP-25. Due to limited nucleotide diversity (Table 2), and thus minimal



**Fig. 4** Primary (A), secondary (B) and tertiary structures (C) of *Schistosoma japonicum* tetraspanins (TSPs), SjtTSP-2, SjtTSP-8, SjtTSP-23 and SjtTSP-25. In A, B and C blue identifies the SEL domain and red the large extracellular loop (LEL) domain. In B and C green regions show the predicted epitopes and cyan the identified amino acid substitutions. Primary structure was visualised using BioEdit, secondary structure was predicted using PROTTTER, and PyMOL molecular visualisation software was used to render the SjtTSP tertiary structures

variation in their respective amino acid sequences, SjtTSP-1, -13 and -14 were omitted from further analysis. Considering the variability of SjtTSP 3D protein structures across the seven provinces, the structural similarity PCA revealed limited clustering, indicating

that across China, SjtTSP-2, -8 and -23 variants generally exhibit limited structural conservation. SjtTSP-23 sequences from Hunan, Jiangxi and Yunnan as well as SjtTSP-8 sequences from Anhui and Zhejiang did, however, demonstrate some structural similarities (Fig. 5). SjtTSP-2 provincial sequences revealed a distinct lack



**Fig. 5** Principal component analyses of structural similarity for *Schistosoma japonicum* tetraspanin (TSP) proteins between provinces for SjtTSP-2 (A), SjtTSP-8 (B), SjtTSP-23 (C) and SjtTSP-25 (D). Structural similarity was inferred from z scores produced by the DALI protein comparison server and then plotted using the Tidyverse package [83] in RStudio [84]

of clustering, outlining the extent of the structural variation across parasite populations from the seven provinces.

**Impact of selection on SjtSP structure and function**

Predicting the functional effects of amino acid substitutions on the protein-coding regions of the three-variable SjtSPs using PROVEAN revealed that SjtSP-2 and -8 had three function-altering mutations, whereas SjtSP-23 had one (Additional File 3: Fig. S2). Sites with altered functionality in SjtSP-2 were spread across the entire protein, at positions 24, 144 and 169, with the latter two sites residing in the LEL domain as well as the largest epitope in SjtSP-2 (Fig. 4). Similarly, all altered sites in SjtSP-8 (143, 145 & 147) resided within the LEL domain, although are outside predicted epitopes (Fig. 4). In contrast, a single function-altering mutation at site 209 within transmembrane domain 4 (TM4) was predicted for SjtSP-23. Further investigation of genetic divergence between SjtSPs, using TreeSAAP, enabled a comparison of amino acid substitutions between each provincial SjtSP relative to those evolving neutrally. Amino acid substitutions in SjtSP-2, -8 and -23 were predicted to

alter a total of 22 physicochemical properties, spanning conformational, chemical, energetic and physical characteristics of amino acids (Table 4). Z scores determined by TreeSAAP evaluated the significance of the amino acid replacements, establishing the type of selection pressure acting on each SjtSP. All 22 property changes were significant (Table 4). Five z scores reached the highest threshold (>3.09 = P>0.001); therefore, those property changes were considered useful for inferring those property changes will result from adaptive evolution.

Isoelectric point, root-mean-square (RMS) fluctuation displacement, chromatographic index and partial specific volume physicochemical properties were predicted to most affect the amino acid biochemistry of SjtSP-2 (Table 4). Based on the TreeSAAP z for SjtSP-2, three and four properties were determined to be influenced by positive and purifying selection, respectively. Seven of the nine altered physicochemical properties in SjtSP-8 were considered significant (P<0.05) and radically altered; of these, refractive index and composition were highly significant and under the influence of positive selection and were therefore considered evidence of adaptive molecular evolution [62]. For SjtSP-23, six of the seven altered

**Table 4** Significant TreeSAAP z scores denoting amino acid physicochemical property alterations across *Schistosoma japonicum* tetraspanin proteins; SjtSP-2, SjtSP-8 and SjtSP-23

Tetraspanin (TSP)	Amino acid physicochemical properties	Magnitude category (z scores)			
		5	6	7	8
SjtSP-2	Isoelectric point			2.219 *	1.92 *
	Mean RMS fluctuation displacement			-2.18 *	
	Partial specific volume		2.041 *		
	Chromatographic index		-1.911 *		
	Hydropathy	-1.679 *			
	Composition	-1.739 *			
SjtSP-8	Power to be at the N-terminal				2.385 **
	Refractive index			7.047 ***	
	Composition		7.677 ***		
	Equilibrium constant (ionisation of COOH)		2.349 **		
	Mean RMS fluctuation displacement		2.301 *		
	Average number of surrounding residues		2.228 *		
	Hydropathy		2.210 *		
	Compressibility	5.025 ***			
	Bulkiness	3.139 ***			
	Power to be at the middle of α-helix	2.036 *			
SjtSP-23	Polar requirement			3.238 ***	
	Refractive index			2.575 **	
	Coil tendencies			1.978 *	
	Composition		2.640 **		
	Compressibility	1.674 *	2.407 **		
	Total non-bonded energy		1.862 *		

Magnitude categories denoting 'radical' or 'substantial' property changes (5–8) are shown; \*P<0.05, \*\*P<0.01, \*\*\*P<0.001

amino acid properties were predicted to be significant and radically altered, although only polar requirement was deemed to be under positive adaptive selection.

These changes in SjtTSP physicochemical properties illustrate the phenotypic impact non-synonymous mutations induce in TSPs from *S. japonicum* populations across China. As established using PROVEAN, most function-altering amino acid changes accumulated within the LEL domain of each variable SjtTSP; therefore, the significantly altered physicochemical properties are assumed to also occur in the LEL domain.

#### Association of SjtTSP genetic variability, antigenicity and antibody binding affinity

Antigenicity profiles confirmed the substantial antigenic propensity of all seven analysed SjtTSPs (Fig. 6A/B/C and Additional File 4: Fig. S3). Interestingly, the exposed LEL domain of these SjtTSPs contained distinctly non-antigenic regions (antigenicity scores < 1), particularly in predicted B-cell epitopes. Considerable variation in the antigenicity profiles of provincial SjtTSP-2, -8 and -23 was demonstrated. The terminal epitope in SjtTSP-2 displayed the most variation in antigenicity between all provincial sequences, whereas the most profound change in antigenicity was demonstrated at residues 141–146 in the LEL domain of SjtTSP-8 from Yunnan. In contrast, SjtTSP-1, -13, -14 and -25 showed no variation in antigenicity throughout the entire protein (Additional File 4: Fig. S3).

When considering the three SjtTSPs displaying variable antigenicity between the provinces, the greatest antigenic fluctuations were predicted within the LEL domain, although certain transmembrane helical regions (e.g. 186–209aa in SjtTSP-2) also exhibited variation. Of particular interest is a peak in antigenicity within an area of the LEL domain flanked by predicted B-cell epitopes (140–150aa), which also appears to be a site of substantial antigenic variation between provinces, in SjtTSP-2 and -8.

The impact of selection, and resultant amino acid substitutions, is further exemplified through the variability in antibody binding affinity within SjtTSP-2, -8 and -23. The predicted antibody binding affinity for each SjtTSP is most variable at sites within the LEL domain and is relatively invariable outside the LEL domain, with  $\Delta\Delta G$  close to zero (Fig. 6D/E/F). The highest frequency of change occurred across the LEL domain of SjtTSP-2, with 10 decreases and 13 increases in affinity (Fig. 6D). A region of SjtTSP-2's LEL domain between residues 141–168 was especially variable, comprising 15 of the 23 predicted antibody binding affinity changes. The largest change in affinity within the LEL domain of SjtTSP-2 was a decrease at site 163 ( $\Delta\Delta G = -0.441$ ), although the greatest change in binding affinity across the three variable SjtTSPs was in the LEL domain of SjtTSP-8, which experienced a large

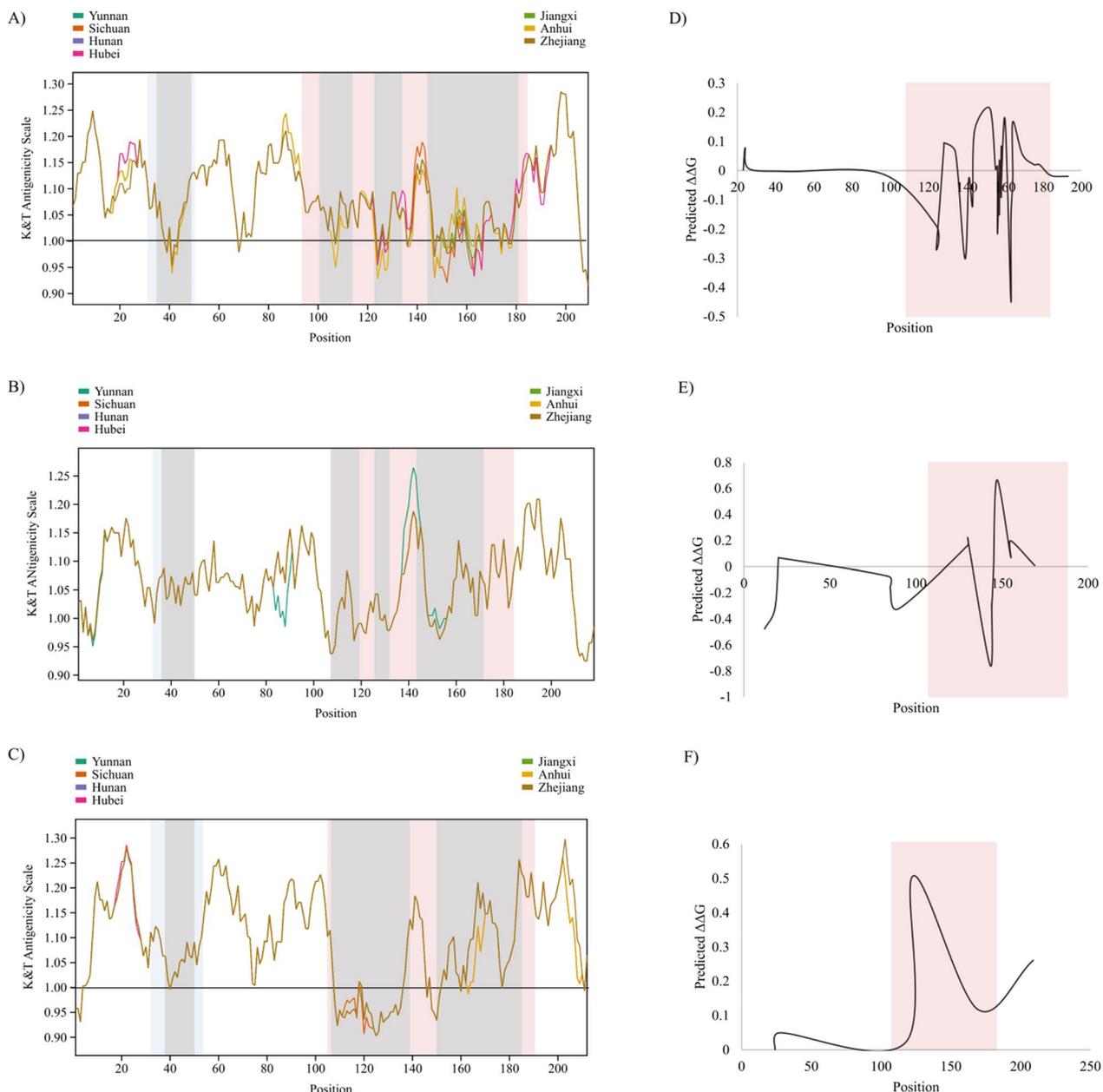
decrease in affinity at site 143 ( $\Delta\Delta G = -0.753$ ) and an increase in affinity at site 147 ( $\Delta\Delta G = 0.665$ ). The greatest alteration in SjtTSP-23 binding affinity also occurred in the LEL domain, with a sharp increase at site 123 ( $\Delta\Delta G = 0.507$ ). Also worth noting is the second largest increase in binding affinity, arising in TM4 at site 209 ( $\Delta\Delta G = 0.261$ ).

#### Discussion

Schistosomes have evolved multiple strategies to evade the host immune system, enabling them to survive in the hostile and complex immunological environment of the host's vasculature for many years [24]. Surface-exposed proteins, such as tegumental TSPs, are likely fundamental to this strategy [31]. Our findings suggest that schistosome TSPs present at the host-parasite interface are attractive targets for a host immune response and that selection pressures drive polymorphisms in the LEL domain of SjtTSPs that impose structural, functional and antibody binding consequences. As such, this diversification acts to reduce the impact of the host's immune response by limiting antibody recognition, which may reduce their suitability as vaccine targets.

Phylogenetic assessment revealed *S. japonicum* TSPs to be highly divergent, with 27 unique members identified across three major protein subfamilies (CD, CD63 and Uroplakin). The subfamilies resolved here broadly agree with those presented by Wu et al. [32] and Jiang et al. [25], with only minor discrepancies, such as TSP#6 falling into the Uroplakin subfamily rather than the CD63 family [29]. Wu et al. [32] reported that TSPs-1, -16 and -20 fell within the CD9-like group [68], as also found here.

The phylogeographic relationships of SjtTSPs-2, -8, -23, and -25 revealed considerable genetic diversity between provincial localities, indicating that between provinces these SjtTSPs are evolving independently. This is supported by the  $\omega$ , MEME and FUBAR analyses, which showed that gene regions encoding the extracellular loops of SjtTSP-2, -8 and -23 are particularly influenced by evolutionary selective pressures, likely resulting from exposure of the domains to the host's immune system. It appears that positive, diversifying selection, acting on specific sites within the LEL domain, promotes variation between TSP proteins from *S. japonicum* isolates from different Chinese provinces. Furthermore, the phylogeographic grouping of provincial SjtTSP sequences coupled with similarities in 3D protein structure suggests a link between genetic and structural variation, particularly in SjtTSP-8 from Anhui and Zhejiang. Proteins are often functionally constrained, limiting amino acid substitutions and structural variation to maintain function, although clustering of SjtTSP-23 from Hunan, Yunnan and Jiangxi, despite a lack of structural similarity from



**Fig. 6** Predicted antigenicity profiles determined for tetraspanins (TSPs) of *Schistosoma japonicum* from Chinese provinces; SjtSP-2 (A), SjtSP-8 (B) and SjtSP-23 (C) and predicted  $\Delta\Delta G$  (change in Gibbs free energy) across tetraspanin protein sequences for SjtSP-2 (D), SjtSP-8 (E) and SjtSP-23 (F). SjtSP proteins are considered antigenic where the antigenicity score is  $> 1$  (solid black line). The blue area marks the small extracellular loop domain (SEL) and the red area the large extracellular loop domains (LEL). The grey area denotes BepiPred predicted B-cell epitopes. A negative  $\Delta\Delta G$  value corresponded to an amino acid change predicted to reduce affinity, while a positive value was predicted to increase binding affinity. Antigenicity scores were plotted using the Tidyverse package [83] in RStudio [84]

Zhejiang and other provinces, shows that functional constraints and sequence similarity do not eliminate amino acid substitutions from driving structural changes in SjtSPs.

Prolonged selection on the LEL domain has led to formation of a hypervariable region [24, 33], whereby

selection appears to favour amino acid changes at certain antibody binding domain codons, as seen between residues 140–147 in the LEL domains of SjtSP-2 and -8. Host-derived selection pressure would promote such hypervariability, thus acting to limit recognition by host antibodies and providing *S. japonicum* with ‘escape’

mutations to facilitate immune evasion [24–26, 33]. This current study revealed distinct variation in SjtTSP-LEL antigenicity and antibody binding affinity in SjtTSPs-2, -8 and -23 between parasites from different provinces. Antibody binding affinity was predicted to be low in the hypervariable region (140–147) in SjtTSP-2-LEL and SjtTSP-8-LEL, but was greater thereafter, highlighting the somewhat indiscriminate effect of non-synonymous mutations and resulting amino acid changes on SjtTSP immunogenicity. Importantly, the variation identified here may indeed underestimate the total variation in TSP ACGs in *S. japonicum* populations across China, as studies have shown that, at least for TSP2, multiple allelic variants can occur at a single location [24, 30]. Nonetheless, we demonstrate substantial variation between geographical isolates and the evolutionary impact the host could have on the structure and antigenic variability of TSPs from geographically distinct parasite populations.

SjtTSP proteins exhibit notable antigenic propensity, with several large peaks in antigenicity predicted across the seven SjtTSPs, as previously reported for SmTSP-23 [33]. Findings here support both extracellular loop domains of SjtTSPs as highly antigenic, with the LEL containing the highest density of antibody binding sites, confirming the LEL domain of SjtTSPs as an attractive target for host immune effector molecules, such as IgG antibodies and B lymphocytes [23, 26, 30]. The LEL domains of variable SjtTSPs (TSPs-2, -8 and -23) displayed the greatest fluctuations in antigenicity over the whole protein, supporting the suggestion by Sealey et al. [33] of an association between amino acid variation and antigenic variation in TSP-23, which also extends to SjtTSP-2 and -8. Conversely, limited geographical antigenic variability of SjtTSP-1, -13 and -14 appears to result from the lack of amino acid variation. Invariable SjtTSP-25 antigenicity, despite considerable amino acid variation, however, suggests there is a more complex relationship between amino acid and antigenic variation than previously proposed. It was also predicted that antigenically variable TSPs would contain substantial differences in the length and location of antibody binding sites in the LEL domain [33]. Many sites within the LEL domain of SjtTSP-2 and -8 with low antigenicity were associated with a greater occurrence of polymorphic sites. SjtTSP-2 displayed evidence of fractured epitopes around the LEL hypervariable region, and SjtTSP-23 had shortened epitopes at the terminus of the LEL domain compared to the other SjtTSPs. Incidentally, host-derived selection pressures appear to drive a reduction in antigenicity and antibody binding, thereby rendering certain SjtTSPs less desirable to host immune components through the accumulation of mutations, highlighting a compelling mechanism of host immunomodulation and evasion by the parasite. Another

tactic of host immune evasion and regulation, shown to be facilitated using SjtTSP-23, is the inhibition and paralysis of host complement activation through preferential binding to the Fc domain of non-immune IgG [26, 69]. This strategy would require portions of SjtTSP-23-LEL to remain attractive targets for IgG, which may be reflected here as a large increase in antibody binding affinity at site 123 in the LEL domain.

Despite evidence of hypervariability and diversifying selection pressures acting on regions encoding SjtTSP-2-LEL, high levels of purifying selection were seen across the entire protein, particularly in the transmembrane helical regions. The resulting sequence conservation may be critical for protein localisation and membrane stability and is likely necessary to maintain function and structure. Primarily, selection is episodic, acting more frequently on individual sites than on large gene regions, suggesting that more frequent neutral or purifying selection may obscure phases of adaptive evolution [55]. Minimal local amino acid substitutions can be driven by adaptive evolution if the amino acid changes are biochemically advantageous, as supported by the major functional alterations highlighted by TreeSAAP, and predicted changes in LEL immunogenicity. Use of amino acid substitutions to explore the effect of selection on protein structure and function provided a high degree of resolution and is considered a more sensitive approach than  $\omega$  [63]. Evolution of antigenic variation, indicated here in certain SjtTSPs, is also demonstrated in the pathogen *Bordetella pertussis* (aetiological agent of whooping cough), which has seen a recent resurgence and epidemiological shift to infect older children and adults [70]. The evolution of a novel strain, adapted to express an antigen variant that escapes vaccine-induced immunity, may have caused this shift through sustained selection pressure from host immunity [71]. This evolutionary strategy of immune avoidance is also utilised by the malaria parasite (*Plasmodium* spp.), where antigenic variation, such as seen in PfEMP-1, is the leading mechanism used to evade detection by the host's immune system [72–74]. The parallels between schistosomes and other pathogenic organisms in their responses to host selection pressures, and mechanisms to drive antigenic variation, could inform us as to how antigenic variation can affect pathogenesis and parasite survivability inside the host in general.

Selection-driven changes in functional properties of TSP-LELs, such as changes in refractive index in SjtTSP-8, polar requirement across SjtTSP-23 and changes in isoelectric point in SjtTSP-2, alter protein folding and might interfere with SjtTSP interaction with host molecules, such as antibodies. Antigenic drift, whereby amino acid substitutions lead to changes in antigenicity over time, could also be indicative of balancing selection acting to

maintain a greater than expected genetic diversity of SjtTSPs through conservation of allelic variants in *S. japonicum* populations. Function-altering mutations within the LEL domain could have direct implications for SjtTSP immunogenicity and parasite survival within its definitive host. Although the implications of physicochemical change on antibody binding affinity are seldom investigated, changes in such properties could significantly impact antigen structure, function and immunogenicity. Host immunity differs greatly between species [75], and therefore two different host species could generate contrasting and variable selection pressures [10], leading to different host assemblages promoting the development of host-specific mutations in schistosome TSPs [30, 33, 76]. It could therefore be important to elucidate the variation in ACGs within and between different host species in future studies, not only to increase knowledge regarding the development of an effective anti-schistosome vaccine but also for understanding the varying disease pathology seen in humans infected with schistosomes [77].

Members of the TSP family have consistently demonstrated remarkable genetic diversity in schistosome species [30, 32, 33, 78], which is thought to be responsible, at least in part, for the variable protective efficacy demonstrated so far in TSP vaccine studies [19, 23, 24, 30, 35]. Vaccine development often involves using monomorphic laboratory isolates [79]. However, the geographic variation revealed here for SjtTSPs indicates that a diverse range of field-collected parasite isolates should be included in future vaccine trials as they better represent 'natural' parasite populations; this strategy would help develop more robust, efficacious vaccines. It has been demonstrated here, for the first time to our knowledge, that SjtTSPs with low genetic diversity and invariable antigenicity between provinces, such as SjtTSP-1, -13 and -14, could therefore prove more robust vaccine candidates. Several immune-exposed schistosome antigens, however, have demonstrated significant diversifying selection across parasite populations [78]—similar to that indicated here for certain SjtTSPs. This leads to a conundrum, as exposed antigens with an apparent propensity for diversification are likely polymorphic in natural, genetically heterogeneous parasite populations, and vaccines formulated against them would likely be ineffective. Non-variable antigens, on the other hand, may be internal/non-exposed, rendering them difficult for the immune system to recognise and mount an effective response against [78]. As little is known about the cellular localisation or function of SjtTSP-13, -14 and -25, further studies on these TSPs are warranted. Regardless, vaccine development based on a single polymorphic form of a TSP would likely offer low efficacy against natural parasite populations. It has recently been proposed that the

endeavour to develop a single-target vaccine has significant limitations, with evidence suggesting a multi-target, multivalent DNA or mRNA vaccine would provide the best protection [80, 81]. Furthermore, any vaccine would require deployment alongside, and integration within, existing control strategies to be suitably effective [17, 81].

## Conclusion

Despite significant reductions in infections across China, *S. japonicum* remains endemic in humans and animal reservoirs in several provincial regions [13], and the risk of re-emergence remains a legitimate concern [9]. It seems that the development of an effective anti-schistosome vaccine, intended to reach end-game elimination targets in China [8], has been hampered by the knowledge gap associated with the genetic diversity and antigenic variability of vaccine candidate antigens within and between parasite populations. To begin to address this, measuring the genetic diversity and evolutionary processes that effect vaccine targets, such as SjtTSPs, could be crucial in identifying robust vaccine targets as well as potential mechanisms employed by the parasite to evade the host immune response. As such, these findings are not only useful in evaluating the use of SjtTSPs as vaccine candidates, but they also provide an analytical framework with which to explore genetic and antigenic variability in other schistosome vaccine candidates. Findings reported here show that SjtTSPs' genetic and antigenic variability between geographically distinct parasite isolates could feasibly influence the effectiveness of the host immune response, impacting vaccine efficacy and representing a challenge for vaccine development against *S. japonicum*. It is therefore recommended that robust evaluation of vaccine target genetic and antigenic variability in natural parasite populations should be a prerequisite for any vaccine development campaign in addition to further immunological work experimentally validating alterations in TSP antigenicity and antibody binding affinity detected here.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13071-023-05706-3>.

**Additional file 1: Fig S1.** Map of China outlining the seven parasite populations (red) used in this study, which were sampled by Young et al. [34], and provincial boundaries (white). In relation to sampling locations reported by Young et al. [34]; Ya'an (Sichuan) represents Tianquan, Jiujiang (Jiangxi) represents Yongxia, Chizhou (Anhui) represents Guichi, and Jiaxing (Zhejiang) represents Jiashan. Sampling location names were changed to reflect the proper district-level classifications. Created using QGIS v3.26.3 [87].

**Additional file 2: Table S1.** Identification of individual nucleotide bases under selection in SjtTSP-2, SjtTSP-8 and SjtTSP-23 using MEME and FUBAR.

**Additional file 3: Fig S2.** Investigating site-specific functional alterations resulting from amino acid changes using PROVEAN. SjtSP-2 (A), SjtSP-8 (B), SjtSP-23 (C), SjtSP1 (D), SjtSP-13 (E), SjtSP-14 (F), SjtSP-25 (G). A PROVEAN score < -2.5, denoted by the back horizontal line, outlines an amino acid change that will induce a functional alteration at that site.

**Additional file 4: Fig. S3.** Identical predicted antigenicity profiles determined for tetraspanins (TSPs) of *Schistosoma japonicum* from Chinese provinces; SjtSP-1 (A), SjtSP-13 (B), SjtSP-14 (C), SjtSP-25 (D). SjtSP proteins are considered antigenic where the antigenicity score is > 1 (solid black line).

#### Acknowledgements

Not applicable.

#### Author contributions

All authors were involved in the design and conception of this study. Data acquisition was undertaken by DAJP and SPL, and analysis was carried out by DAJP, with support and supervision from SPL, AJW, AME and JPW. The original draft manuscript was written by DAJP and subsequently critiqued, revised and improved by SPL, AME, AJW and JPW. All authors read and approved the final manuscript.

#### Funding

This work was completed as part of a partially funded PhD studentship supported by Kingston University London GRF scheme.

#### Availability of data and materials

In silico study based on published *Schistosoma japonicum* genome data as described in the methods.

#### Declarations

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

Not applicable.

#### Author details

<sup>1</sup>Molecular Parasitology Laboratory, School of Life Sciences, Pharmacy and Chemistry, Kingston University London, Kingston Upon Thames, Penrhyn Rd, Surrey KT1 2EE, UK. <sup>2</sup>Department of Life Sciences, Natural History Museum, Cromwell Rd, South Kensington, London SW7 5BD, UK. <sup>3</sup>Department of Pathobiology and Population Sciences, Royal Veterinary College, University of London, Hawkshead Campus, Herts AL9 7TA, UK. <sup>4</sup>Centre for Epidemiology & Planetary Health, Department of Veterinary and Animal Sciences, Northern Faculty, Scotland's Rural College, An Lòchran, 10 Inverness Campus, Inverness IV2 5NA, UK.

Received: 13 November 2022 Accepted: 17 February 2023

Published online: 14 March 2023

#### References

- Rudge JW, Webster JP, Lu DB, Wang TP, Fang GR, Basáñez MG. Identifying host species driving transmission of *Schistosoma japonicum*, a multihost parasite system, in China. *Proc Natl Acad Sci USA*. 2013;110:11457–62.
- Gordon CA, Kurscheid J, Williams GM, Clements ACA, Li Y, Zhou XN, et al. Asian schistosomiasis: current status and prospects for control leading to elimination. *Trop Med Infect Dis*. 2019;4:40. <https://doi.org/10.3390/tropicalmed4010040>.
- Xu J, Li SZ, Zhang LJ, Bergquist R, Dang H, Wang Q, et al. Surveillance-based evidence: elimination of schistosomiasis as a public health problem in the Peoples' Republic of China. *Infect Dis Poverty*. 2020;9:63. <https://doi.org/10.1186/s40249-020-00676-5>.
- World Health Organisation. Ending the neglect to attain the Sustainable Development Goals: a road map for neglected tropical diseases 2021–2030. Geneva: World Health Organisation; 2021.
- Xianyi C, Liying W, Jiming C, Xiaonong Z, Jiang Z, Jiagang G, et al. Schistosomiasis control in China: the impact of a 10-year World Bank Loan Project (1992–2001). *Bull World Health Organ*. 2005;83:43–8.
- Lu DB, Rudge JW, Wang TP, Donnelly CA, Fang GR, Webster JP. Transmission of *Schistosoma japonicum* in marshland and hilly regions of China: parasite population genetic and sibship structure. *PLoS Negl Trop Dis*. 2010;4:e781. <https://doi.org/10.1371/journal.pntd.0000781>.
- Cao Z, Huang Y, Wang T. Schistosomiasis Japonica control in domestic animals: progress and experiences in China. *Front Microbiol*. 2017;14:2464. <https://doi.org/10.3389/fmicb.2017.02464>.
- Wang W, Bergquist R, King CH, Yang K. Elimination of schistosomiasis in China: current status and future prospects. *PLoS Negl Trop Dis*. 2021;15:e0009578. <https://doi.org/10.1371/journal.pntd.0009578>.
- Liang S, Yang C, Zhong B, Qiu D. Re-emerging schistosomiasis in hilly and mountainous areas of Sichuan, China. *Bull World Health Organ*. 2002;84:139–44.
- Wang TP, Shrivastava J, Johansen MV, Zhang SQ, Wang FF, Webster JP. Does multiple hosts mean multiple parasites? Population genetic structure of *Schistosoma japonicum* between definitive host species. *Int J Parasitol*. 2006;12:1317–25. <https://doi.org/10.1016/j.ijpara.2006.06.011>.
- Wang W, Dai JR, Liang YS. Apropos: factors impacting on progress towards elimination of transmission of *Schistosoma japonicum* in China. *Parasit Vectors*. 2014;7:408. <https://doi.org/10.1186/1756-3305-7-408>.
- Yang GJ, Bergquist R. Potential impact of climate change on schistosomiasis: a global assessment attempt. *Trop Med Infect Dis*. 2018;3:117. <https://doi.org/10.3390/tropicalmed3040117>.
- Collins C, Xu J, Tang S. Schistosomiasis control and the health system in P.R. China. *Infect Dis Poverty*. 2012;1:8. <https://doi.org/10.1186/2049-9957-1-8>.
- Webster JP, Molyneux DH, Hotez PJ, Fenwick A. The contribution of mass drug administration to global health: past, present, and future. *Philos Trans R Soc Lond B Biol Sci*. 2014;369:20130434. <https://doi.org/10.1098/rstb.2013.0434>.
- Wang W, Wang L, Liang YS. Susceptibility or resistance of praziquantel in human schistosomiasis: a review. *Parasitol Res*. 2012;111:1871–7. <https://doi.org/10.1007/s00436-012-3151-z>.
- McManus DP, Loukas A. Current status of vaccines for schistosomiasis. *Clin Microbiol Rev*. 2008;21:225. <https://doi.org/10.1128/CMR.00046-07>.
- Tebeje BM, Harvie M, You H, Loukas A, McManus DP. Schistosomiasis vaccines: where do we stand? *Parasit Vectors*. 2016;9:528. <https://doi.org/10.1186/s13071-016-1799-4>.
- Xiao SH, Sun J, Chen MG. Pharmacological and immunological effects of praziquantel against *Schistosoma japonicum*: a scoping review of experimental studies. *Infect Dis Poverty*. 2018;7:9. <https://doi.org/10.1186/s40249-018-0391-x>.
- Da'dara AA, Li YS, Xiong T, Zhou J, Williams GM, McManus DP, et al. DNA-based vaccines protect against zoonotic schistosomiasis in water buffalo. *Vaccine*. 2008;26:3617–25. <https://doi.org/10.1016/j.vaccine.2008.04.08>.
- Yuan C, Fu YJ, Li J, Yue YF, Cai LL, Xiao WJ, et al. *Schistosoma japonicum* efficient and rapid purification of the tetraspanin extracellular loop 2, a potential protective antigen against schistosomiasis in mammalian. *Exp Parasitol*. 2010;v126:456–61. <https://doi.org/10.1016/j.exppara.2010.05.018>.
- Lee J, Chong T, Newmark PA. The esophageal gland mediates host immune evasion by the human parasite *Schistosoma mansoni*. *Proc Natl Acad Sci*. 2020;117:19299–309. <https://doi.org/10.1073/pnas.2006553117>.
- Wilson RA, Jones MK. Fifty years of the schistosome tegument: discoveries, controversies, and outstanding questions. *Int J Parasitol*. 2021;51:1213–32.
- Tran MH, Pearson MS, Bethony JM, Smyth DJ, Jones MK, Duke M, et al. Tetraspanins on the surface of *Schistosoma mansoni* are protective antigens against schistosomiasis. *Nat Med*. 2006;12:835–40. <https://doi.org/10.1038/nm1430>.
- Cai P, Bu L, Wang J, Wang Z, Zhong X, Wang H. Molecular characterization of *Schistosoma japonicum* tegument protein tetraspanin-2: sequence variation and possible implications for immune evasion. *Biochem*

- Biophys Res Commun. 2008;372:197–202. <https://doi.org/10.1016/j.bbrc.2008.05.042>.
25. Jiang N, Cai P, Yin J, Hao L, Lu H, Wang X, et al. Characterization of antibody responses to the Sj23 antigen of *Schistosoma japonicum* after infection and immunization. *Acta Trop*. 2010;116:9–14. <https://doi.org/10.1016/j.actatropica.2010.04.015>.
  26. Wu C, Cai P, Chang Q, Hao L, Peng S, Sun X, et al. Mapping the binding between the tetraspanin molecule (Sjc23) of *Schistosoma japonicum* and human non-immune IgG. *PLoS ONE*. 2011;6:e19112. <https://doi.org/10.1371/journal.pone.0019112>.
  27. Tran MH, Freitas TC, Cooper L, Gaze S, Gatton ML, Jones MK, et al. Suppression of mRNAs encoding tegument tetraspanins from *Schistosoma mansoni* results in impaired tegument turnover. *PLoS Pathog*. 2010;6:e1000840. <https://doi.org/10.1371/journal.ppat.1000840>.
  28. Gobert GN, Tran MH, Moertel L, Mulvenna J, Jones MK, McManus DP, et al. Transcriptional changes in *Schistosoma mansoni* during early schistosomula development and in the presence of erythrocytes. *PLoS Negl Trop Dis*. 2010;4:e600. <https://doi.org/10.1371/journal.pntd.0000600>.
  29. Jiang Y, Xu X, Qing X, Pan W. Identification and characterization of six novel tetraspanins from *Schistosoma japonicum*. *Parasit Vectors*. 2011;4:190. <https://doi.org/10.1186/1756-3305-4-190>.
  30. Zhang W, Li J, Duke M, Jones MK, Kuang L, Zhang J, et al. Inconsistent protective efficacy and marked polymorphism limits the value of *Schistosoma japonicum* tetraspanin-2 as a vaccine target. *PLoS Negl Trop Dis*. 2011;5:e1166. <https://doi.org/10.1371/journal.pntd.0001166>.
  31. Hambrook JR, Hanington PC. Immune evasion strategies of schistosomes. *Front Immunol*. 2021;11:624178. <https://doi.org/10.3389/fimmu.2020.624178>.
  32. Wu W, Cai P, Chen Q, Wang H. Identification of novel antigens within the *Schistosoma japonicum* tetraspanin family based on molecular characterization. *Acta Trop*. 2011;117:216–24. <https://doi.org/10.1016/j.actatropica.2011.01.001>.
  33. Sealey KL, Kirk RS, Walker AJ, Rollinson D, Lawton SP. Adaptive radiation within the vaccine target tetraspanin-23 across nine *Schistosoma* species from Africa. *Int J Parasitol*. 2013;43:95–103. <https://doi.org/10.1016/j.ijpara.2012.11.007>.
  34. Young ND, Chan KG, Korhonen PK, Min Chong T, Ee R, Mohandas N, et al. Exploring molecular variation in *Schistosoma japonicum* in China. *Sci Rep*. 2015;5:17345. <https://doi.org/10.1038/srep17345>.
  35. Cupit PM, Steinauer ML, Tonnessen BW, Eric Agola L, Kinuthia JM, Mwangi IN, et al. Polymorphism associated with the *Schistosoma mansoni* tetraspanin-2 gene. *Int J Parasitol*. 2011;41:1249–52. <https://doi.org/10.1016/j.ijpara.2011.07.007>.
  36. Wu Z, Yu X, Xu JJ, Zheng Y, Cao A. Molecular cloning and characterization of a novel gene encoding tetraspanin of *Schistosoma japonicum*. *Helminthologia*. 2004;41:15–9.
  37. Howe KL, Bolt BJ, Shafie M, Kersey P, Berriman M. WormBase ParaSite—a comprehensive resource for helminth genomics. *Mol Biochem Parasitol*. 2017;215:2–10. <https://doi.org/10.1016/j.molbiopara.2016.11.005>.
  38. Fan J, Brindley PJ. Characterization of cDNAs encoding a new family of tetraspanins from schistosomes—the Sj25 family. *Gene*. 1998;219:1–8. [https://doi.org/10.1016/s0378-1119\(98\)00397-7](https://doi.org/10.1016/s0378-1119(98)00397-7).
  39. Birney E, Clamp M, Durbin R. GeneWise and genomewise. *Genome Res*. 2004;14:988–95. <https://doi.org/10.1101/gr.1865504>.
  40. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol*. 2001;305:567–80. <https://doi.org/10.1006/jmbi.2000.4315>.
  41. Letunic I, Doerks T, Bork P. SMART: recent updates, new developments and status in 2015. *Nucleic Acids Res*. 2015;43:D257–60. <https://doi.org/10.1093/nar/gku949>.
  42. Omasits U, Ahrens CH, Müller S, Wollscheid B. Protter: interactive protein feature visualization and integration with experimental proteomic data. *Bioinformatics*. 2014;30:884–6. <https://doi.org/10.1093/bioinformatics/btt607>.
  43. Edgar RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinform*. 2004;5:113. <https://doi.org/10.1186/1471-2105-5-113>.
  44. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol Biol Evol*. 2016;33:1870–4. <https://doi.org/10.1093/molbev/msw054>.
  45. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol*. 1987;4:406–25. <https://doi.org/10.1093/oxfordjournals.molbev.a040454>.
  46. Sinha S, Saho G, Rana S, Topno R, Ansari M, Dikhit MR, et al. Next generation sequence analysis of sequences from SRA file of HIV-1 envelope protein. *Int J Adv Res Publ*. 2018;6:1118–30.
  47. Charrier NP, Couton M, Voordouw MJ, Rais O, Durand-Hermouet A, Hervet C, et al. Whole body transcriptomes and new insights into the biology of the tick *Ixodes ricinus*. *Parasites Vectors*. 2018;11:364. <https://doi.org/10.1186/s13071-018-2932-3>.
  48. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. *BMC Bioinform*. 2009;10:421. <https://doi.org/10.1186/1471-2105-10-421>.
  49. Hall TA. BioEdit: a user friendly biological sequence alignment program editor and analysis program for Windows 95/98/NT. *Nucleic acids symp ser*. 1999;41:95–8.
  50. Huang X, Madan A. CAP3: a DNA sequence assembly program. *Genome Res*. 1999;9:868–77. <https://doi.org/10.1101/gr.9.9.868>.
  51. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990;215:403–10. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2).
  52. Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, et al. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol Biol Evol*. 2017;34:3299–302. <https://doi.org/10.1093/molbev/msx248>.
  53. Murrell B, Moola S, Mabona A, Weighill T, Sheward D, Kosakovsky Pond SL, et al. FUBAR: a fast, unconstrained bayesian approximation for inferring selection. *Mol Biol Evol*. 2013;30:1196–205. <https://doi.org/10.1093/molbev/mst030>.
  54. Kosakovsky Pond SL, Frost SD. Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Mol Biol Evol*. 2005;22:1208–22. <https://doi.org/10.1093/molbev/msi105>.
  55. Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K, Kosakovsky Pond SL. Detecting individual sites subject to episodic diversifying selection. *PLoS Genet*. 2012;8:e1002764. <https://doi.org/10.1371/journal.pgen.1002764>.
  56. Weaver S, Shank SD, Spielman SJ, Li M, Muse SV, Kosakovsky Pond SL. Datamonkey 2.0: a modern web application for characterizing selective and other evolutionary processes. *Mol Biol Evol*. 2018;35:773–7. <https://doi.org/10.1093/molbev/msx335>.
  57. Jespersen MC, Peters B, Nielsen M, Marcotili P. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Res*. 2017;45:W24–9. <https://doi.org/10.1093/nar/gkx346>.
  58. Zhang Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinform*. 2008;9:40. <https://doi.org/10.1186/1471-2105-9-40>.
  59. Schrödinger L, DeLano W. The PyMOL molecular graphics system, Version 2.0 Schrödinger, LLC.
  60. Holm L. DALI and the persistence of protein shape. *Protein Sci*. 2020;29:128–40. <https://doi.org/10.1002/pro.3749>.
  61. Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the functional effect of amino acid substitutions and indels. *PLoS ONE*. 2012;7:e46688. <https://doi.org/10.1371/journal.pone.0046688>.
  62. Woolley S, Johnson J, Smith MJ, Crandall KA, McClellan DA. TreeSAAP: selection on amino acid properties using phylogenetic trees. *Bioinformatics*. 2003;19:671–2. <https://doi.org/10.1093/bioinformatics/btg043>.
  63. McClellan DA, Palfreyman EJ, Smith MJ, Moss JL, Christensen RG, Sailsbury JK. Physicochemical evolution and molecular adaptation of the cetacean and artiodactyl cytochrome b proteins. *Mol Biol Evol*. 2005;22:437–55. <https://doi.org/10.1093/molbev/msi028>.
  64. McClellan DA, Ellison DD. Assessing and improving the accuracy of detecting protein adaptation with the TreeSAAP analytical software. *Int J Bioinform Res Appl*. 2010;6:120–33. <https://doi.org/10.1504/IJBRA.2010.032116>.
  65. Vita R, Mahajan S, Overton JA, Dhanda SK, Martini S, Cantrell JR, et al. The Immune Epitope Database (IEDB): 2018 update. *Nucleic Acids Res*. 2019;47:D339–43. <https://doi.org/10.1093/nar/gky1006>.
  66. Kolaskar AS, Tongaonkar PC. A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS Lett*. 1990;276:172–4. [https://doi.org/10.1016/0014-5793\(90\)80535-q](https://doi.org/10.1016/0014-5793(90)80535-q).

67. Pires DE, Ascher DB. mCSM-AB: a web server for predicting antibody-antigen affinity changes upon mutation with graph-based signatures. *Nucleic Acids Res.* 2016;44:469–73. <https://doi.org/10.1093/nar/gkw458>.
68. Garcia-España A, Chung PJ, Sarkar IN, Stiner E, Sun TT, DeSalle R. Appearance of new tetraspanin genes during vertebrate evolution. *Genomics.* 2008;91:326–34. <https://doi.org/10.1016/j.ygeno.2007.12.005>.
69. Wu C, Hou N, Piao X, Liu S, Cai P, Xiao Y, et al. Non-immune immunoglobulins shield *Schistosoma japonicum* from host immunorecognition. *Sci Rep.* 2015;5:13434. <https://doi.org/10.1038/srep13434>.
70. Burns DL, Meade BD, Messonnier NE. Pertussis resurgence: perspectives from the working group meeting on pertussis on the causes, possible paths forward, and gaps in our knowledge. *Int J Infect Dis.* 2014;209:32–5. <https://doi.org/10.1093/infdis/jit491>.
71. Preston A. The role of *B. pertussis* vaccine antigen gene variants in pertussis resurgence and possible consequences for vaccine development. *Hum Vaccin Immunother.* 2016;12:1274–6. <https://doi.org/10.1080/21645515.2015.1137402>.
72. McKenzie FE, Ferreira MU, Baird JK, Snounou G, Bossert WH. Meiotic recombination, cross-reactivity, and persistence in *Plasmodium falciparum*. *Evolution.* 2001;55:1299–307. <https://doi.org/10.1111/j.0014-3820.2001.tb00652.x>.
73. Smith JD, Chitnis CE, Craig AG, Roberts DJ, Hudson-Taylor DE, Peterson DS, et al. Switches in expression of *Plasmodium falciparum* var genes correlate with changes in antigenic and cytoadherent phenotypes of infected erythrocytes. *Cell.* 1995;82:101–10. [https://doi.org/10.1016/0092-8674\(95\)90056-x](https://doi.org/10.1016/0092-8674(95)90056-x).
74. Ferreira MU, da Silva NM, Wunderlich G. Antigenic diversity and immune evasion by malaria parasites. *Clin Diagn Lab Immunol.* 2004;11:987–95. <https://doi.org/10.1128/CDLI.11.6.987-995.2004>.
75. Bailey M, Christoforidou Z, Lewis MC. The evolutionary basis for differences between the immune systems of man, mouse, pig and ruminants. *Vet Immunol Immunopathol.* 2013;152:13–9. <https://doi.org/10.1016/j.vetimm.2012.09.022>.
76. He YX, Salafsky B, Ramaswamy K. Host–parasite relationships of *Schistosoma japonicum* in mammalian hosts. *Trends Parasitol.* 2001;17:320–4. [https://doi.org/10.1016/s1471-4922\(01\)01904-3](https://doi.org/10.1016/s1471-4922(01)01904-3).
77. Olveda DU, Olveda RM, McManus DP, Cai P, Chau TN, Lam AK, et al. The chronic enteropathogenic disease schistosomiasis. *Int J Infect Dis.* 2014;28:193–203. <https://doi.org/10.1016/j.ijid.2014.07.009>.
78. Philippesen GS, Wilson RA, DeMarco R. Accelerated evolution of schistosome genes coding for proteins located at the host–parasite interface. *Genome Biol Evol.* 2015;7:431–43.
79. Shrivastava J, Qian BZ, Mcvean G, Webster JP. An insight into the genetic variation of *Schistosoma japonicum* in mainland China using DNA microsatellite markers. *Mol Ecol.* 2005;14:839–49. <https://doi.org/10.1111/j.1365-294X.2005.02443.x>.
80. Copy Li XH, Vance GM, Cartwright J, Cao JP, Wilson RA, Castro-Borges W. Mapping the epitopes of *Schistosoma japonicum* esophageal gland proteins for incorporation into vaccine constructs. *PLoS ONE.* 2020;15:e0229542. <https://doi.org/10.1371/journal.pone.0229542>.
81. Molehin AJ, McManus DP, You H. Vaccines for human schistosomiasis: recent progress, new developments and future prospects. *Int J Mol Sci.* 2022;23:2255. <https://doi.org/10.3390/ijms23042255>.
82. Hammer O, Harper DAT, Ryan PD. PAST: paleontological statistics software package for education and data analysis. *Palaeontol Electron.* 2001;4:9.
83. Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, et al. Welcome to the Tidyverse. *J Open Source Softw.* 2019;4:1686. <https://doi.org/10.21105/joss.01686>.
84. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>. 2020;91.
85. Zhou Y, Zheng H, Chen Y, Zhang L, Wang K, Guo J, et al. The *Schistosoma japonicum* genome reveals features of host–parasite interplay. *Nature.* 2009;460:345–51. <https://doi.org/10.1038/nature08140>.
86. Liu F, Lu J, Hu W, Wang SY, Cui SJ, Chi M, et al. New perspectives on host–parasite interplay by comparative transcriptomic and proteomic analyses of *Schistosoma japonicum*. *PLoS Pathog.* 2006;2:e29. <https://doi.org/10.1371/journal.ppat.0020029>.
87. QGIS Association. QGIS Geographic Information System. 2022. Open-Source Geospatial Foundation Project. <https://www.qgis.org>

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

