



OPEN ACCESS

EDITED BY

Camila Koutsodontis Cerqueira Cézar,
São Paulo State University, Brazil

REVIEWED BY

Nattawooti Sthitmatee,
Chiang Mai University, Thailand
David Smith,
Moredun Research Institute,
United Kingdom

*CORRESPONDENCE

Martha Betson
m.betson@surrey.ac.uk

SPECIALTY SECTION

This article was submitted to
Parasitology,
a section of the journal
Frontiers in Veterinary Science

RECEIVED 08 August 2022

ACCEPTED 12 October 2022

PUBLISHED 01 November 2022

CITATION

Evangelista FMD, van Vliet AHM,
Lawton SP and Betson M (2022) A
reverse vaccinology approach
identifies putative vaccination targets
in the zoonotic nematode *Ascaris*.
Front. Vet. Sci. 9:1014198.
doi: 10.3389/fvets.2022.1014198

COPYRIGHT

© 2022 Evangelista, van Vliet, Lawton
and Betson. This is an open-access
article distributed under the terms of
the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution
or reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

A reverse vaccinology approach identifies putative vaccination targets in the zoonotic nematode *Ascaris*

Francisco M. D. Evangelista¹, Arnoud H. M. van Vliet¹,
Scott P. Lawton² and Martha Betson^{1*}

¹School of Veterinary Medicine, Faculty of Health and Medical Sciences, University of Surrey, Guildford, United Kingdom, ²Centre for Epidemiology and Planetary Health, Department of Veterinary and Animal Sciences, Northern Faculty, Scotland's Rural University College (SRUC), An Lóchran, Inverness, United Kingdom

Ascariasis is the most prevalent helminthic disease affecting both humans and pigs and is caused by the roundworms *Ascaris lumbricoides* and *Ascaris suum*. While preventive chemotherapy continues to be the most common control method, recent reports of anthelmintic resistance highlight the need for development of a vaccine against ascariasis. The aim of this study was to use a reverse vaccinology approach to identify potential vaccine candidates for *Ascaris*. Three *Ascaris* proteomes predicted from whole-genome sequences were analyzed. Candidate proteins were identified using open-access bioinformatic tools (e.g., Vacceed, VaxiJen, Bepipred 2.0) which test for different characteristics such as sub-cellular location, T-cell and B-cell molecular binding, antigenicity, allergenicity and phylogenetic relationship with other nematode proteins. From over 100,000 protein sequences analyzed, four transmembrane proteins were predicted to be non-allergen antigens and potential vaccine candidates. The four proteins are a Piezo protein, two voltage-dependent calcium channels and a protocadherin-like protein, are all expressed in either the muscle or ovaries of both *Ascaris* species, and all contained high affinity epitopes for T-cells and B-cells. The use of a reverse vaccinology approach allowed the prediction of four new potential vaccination targets against ascariasis in humans and pigs. These targets can now be further tested in *in vitro* and *in vivo* assays to prove efficacy in both pigs and humans.

KEYWORDS

Ascaris lumbricoides, *Ascaris suum*, Neglected Tropical Diseases (NTDs), nematodes, zoonoses, vaccine

Introduction

The giant roundworm *Ascaris* is the most prevalent soil-transmitted helminth (STH) infection in humans, being responsible for 0.861 million disability-adjusted life years (DALYs) worldwide (1). In parallel, *Ascaris* infections remain an issue on many pig farms worldwide. This parasite is especially important in farms with lower levels of biosecurity, such as organic farms, and/or in lower to medium income countries, where backyard

farming is common (2, 3). Being a zoonotic disease, the close contact between humans and pigs in many highly endemic areas increases the risk *Ascaris* transmission among both hosts (3, 4).

Preventive chemotherapy, bolstered by implementation of water, sanitation and hygiene (WASH) protocols, continues to be the mainstay of ascariasis control in humans, as advocated in the World Health Organization (WHO) roadmap for the control of Neglected Tropical Diseases (NTDs) (5). Reduced efficacy and resistance to benzimidazole drugs have already been reported in human *Ascaris* infections (6, 7), and this directly affects the implementation of control protocols in endemic areas, necessitating the development of effective vaccines and vaccination protocols. Vaccination against *Ascaris* has had some degree of efficacy in mouse and pig animal models, but no candidate vaccine has undergone human clinical trials yet (8). The proteins used in the past vaccination assays were either secreted proteins or those retrieved from crude extracts of adult worms. Recently, vaccination with a chimeric protein led to up to 73% larval burden reduction in mice, in contrast to a 99.8% reduction when mice were repeatedly infected with *Ascaris* sterile eggs (9–12). These results highlight that there are other antigens that could be tested. These antigens could then be incorporated in a multi-component vaccine along with the already known vaccination targets to stimulate a more complete immune response.

Several annotated genomes are now available for *Ascaris lumbricoides* and *Ascaris suum*, which enable the search for putative vaccine candidates using a reverse vaccinology analysis. A reverse vaccinology approach combines genome information and bioinformatic tools for identification of vaccine candidates, and has been successfully applied to other nematodes before, such as *Toxocara canis* and *Trichuris muris* (13, 14). Such methodology allows researchers to uncover vaccine targets from predicted proteomes by assessing if proteins have useful characteristics. Different analyses, such as protein sub-cellular location and the prediction of B-cell and T-cell epitopes, are examples of steps used in these approaches that help select proteins for further testing. This is especially important when funding is limited, and vaccination trials need to be focused. The aim of this study was to apply an *in silico* methodology to analyse protein sequences predicted from three *Ascaris* annotated proteomes to identify potential new vaccination targets that could be used in vaccination assays against *Ascaris*.

Materials and methods

Data retrieval

The annotated proteomes for three assembled genomes of *Ascaris* spp. were retrieved from the WormBase ParaSite database (15). The *A. lumbricoides* proteome from BioProject PRJEB4950 (16) has 23,604 protein sequences. The *A.*

suum proteomes from BioProject PRJNA80881 (17) and BioProject PRJNA62057 (18, 19) have 18,542 and 57,968 protein sequences, respectively.

Protein subcellular location analysis

The retrieved proteomes were first visualized and analyzed with BioEdit v7.2.5 (20). A total of 2,604 protein sequences that included unknown amino acids (aa) (indicated with the symbol “X”) were excluded from further analysis as they tend to be a consequence of poor annotations and most of the bioinformatics tools used do not analyse protein sequences with unknown amino acids.

The framework Vacceed v2.1 (21) was used to identify potential vaccine candidates. The tools employed in Vacceed analysis were: WoLF PSORT v0.2 (22), DeepLoc v1.0 (23), SignalP v5.0 (24), TargetP v2.0 (25), TMHMM v2.0 (26) and Phobius v1.01 (27). Proteins were ranked with scores between 0 and 1, from low immunogenicity (final score = 0) to high immunogenicity potential (final score = 1). Higher scores were given to proteins which were predicted to be secreted, signal or transmembrane peptides. The proteins sequences with a Vacceed score of ≥ 0.750 were retrieved (28) and later tested for epitope binding to CD4+ T helper (Th) cells.

As Vacceed makes use of different bioinformatic tools, there was the need to check how much each tool influences the final protein scores. After retrieving the Vacceed scores from runs including all programs, the process was repeated six times excluding one program per run (i.e., first run with all the bioinformatic tools except TMHMM, second with all the bioinformatic tools except DeepLoc, and so on) (28). The proteins scores were then compared between the different assessments through a Pearson correlation test using R v3.6.3 (29), RStudio v1.2.5033 (30) [with the package *corrplot* (31)] and IBM SPSS Statistics v26.

CD4+ Th cell binding predictions

The protein sequences retrieved with Vacceed were submitted to the standalone version of the Immune Epitope Database (IEDB) Major Histocompatibility Complex class II (MHC-II) binding predictor v2.22.3 (MHCII-IEDB, available at: <http://tools.iedb.org/mhcii/>). This tool employs neural networks trained on IEDB experimentally validated epitopes to predict and quantify the binding affinity between a given peptide/antigen epitope and a selected MHC-II molecule recognized by CD4+ Th cells. The IEDB-recommended 2.22 prediction method was used, comprising the Consensus approach, NN-align, SMM-align, CombLib and Sturniolo methods, or NetMHCIIpan, if any of the previous methods was not available for the selected MHC-II allele (32). This method

was used against the 27 human leukocyte antigen (HLA) allele reference set, covering around 99% of the world's human population (33). The default peptide epitope length of 15 aa with a core of 9 aa was selected. The protein sequences which had epitopes ranked from zero to one (from the maximum of 100) and were simultaneously predicted to bind to all the 27 alleles in the used reference set were selected for further testing. These MHC-II binding predictions were only made for human alleles due to the lack of *in silico* tools that make binding predictions for swine MHC-II alleles.

Allergenicity, antigenicity, and function predictions

The protein sequences retrieved with MHCII-IEDB were tested for antigenicity and allergenicity. Protein antigenicity was analyzed using IEDB Kolaskar and Tongaonkar antigenicity scale and VaxiJen 2.0 (34, 35). The IEDB Kolaskar and Tongaonkar antigenicity scale method, available at <http://tools.iedb.org/bcell/>, was used with default setting and threshold of 1.000 and the VaxiJen 2.0 tool, server accessible at <http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html>, was used with the default threshold of 0.5. Allergenicity was evaluated using AllerTop 2.0 and AllergenFP (36, 37). The AllerTop 2.0 server, <https://www.ddg-pharmfac.net/AllerTOP/index.html>, and the AllergenFP tool, <http://www.ddg-pharmfac.net/AllergenFP/index.html>, analyse protein sequences and compare them to a training set of 2,427 known allergens and 2,427 non-allergens using two different algorithms for the better recognition of allergens and non-allergens, respectively, for each tool. Proteins that were classified as both potential antigens and non-allergens were considered as good vaccination targets.

The function of the selected proteins, their protein family and tissue location in both *A. lumbricoides* and *A. suum* were analyzed based on the most recent genomic and transcriptomic studies (4, 19). These assays were able to map the location where all the genes and, therefore, proteins were transcribed in both *Ascaris* species as well as in what stages of the parasite lifecycle. This information is important as it allows us to infer what roles these proteins could play in the *Ascaris* lifecycle and how useful they would be as components in a vaccine.

CD4+ Th cell binding epitope selection

To further streamline the epitope selection process, we integrated the information gathered using the Phobius and MHCII-IEDB tools to select extracellular CD4+ Th binding epitopes. We used Phobius to assess the extracellular domains of the predicted proteins and identify if the previously predicted epitopes were present in those areas. Proteins that lacked the full 15 aa epitopes in these areas were disregarded. For each

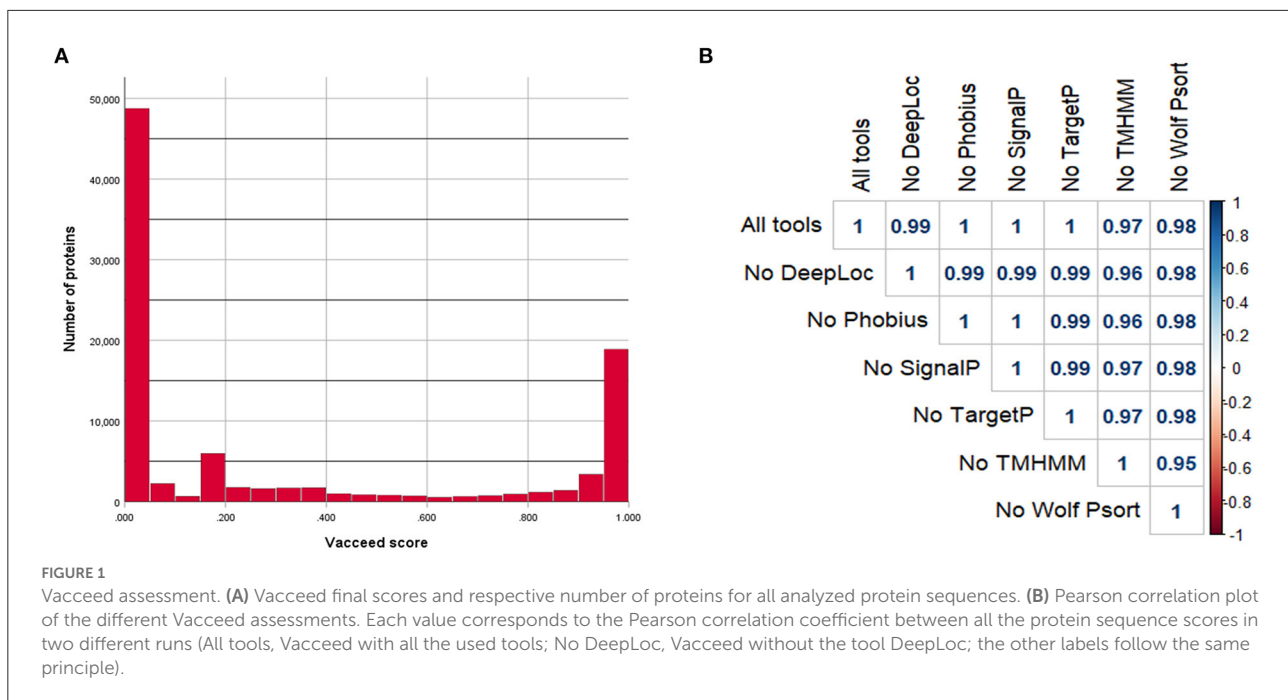
protein, the two non-redundant epitopes (without overlapping aa) predicted to bind to the largest combined number of unique alleles in MHCII-IEDB tool analysis were retrieved. These epitopes were submitted to a BLASTp (Basic Local Alignment Search Tool protein) search (available at: <https://blast.ncbi.nlm.nih.gov/Blast.cgi>) to check for identity in humans and pigs. Epitopes with 100% identity to human or pig epitopes were discarded. The BLASTp analysis was conducted with a threshold of $e\text{-value}=0.05$. Each epitope was also submitted to Allertop 2.0 (36) to confirm they were not potential allergens. The best expressed protein transcript for each vaccine target, as found in the most recent genomic studies (4, 19), was used to predict the presence of glycosylation sites using NetNGlyc 1.0 (38). Only epitopes that were not in direct contact with glycosylation sites were selected. These transcripts were also submitted to Protter to draw the two-dimensional (2D) protein structure using the information retrieved with Phobius (27, 39).

B-cell linear epitope prediction and selection

The presence of B-cell linear epitopes was assessed using BepiPred 2.0 webservice (available at: <http://www.cbs.dtu.dk/services/BepiPred/>) (40). Using the same methodology as in the previous step, proteins that had epitopes found in extracellular domains, exposed to the host immunological system by not being in predicted glycosylation sites, predicted to be non-allergens, and present in most of the protein transcripts of the different genes were considered good vaccination targets. The two highest scored non-allergen epitopes, with a size between 8 and 40 aa, were retrieved for each protein.

Protein phylogenetic analysis

To assess the relationship between the predicted vaccine targets and their orthologs in other nematodes, a phylogenetic analysis of the predicted vaccine targets was performed including predicted orthologs/similar proteins present in annotated genomes of other parasitic nematodes and *Caenorhabditis elegans* available in WormBase ParaSite database (15). Proteins were selected according to the list of orthologs provided by WormBase and, if no ortholog was present for a given nematode, a BlastP analysis was carried out and proteins were retrieved according to the combination of query coverage, identity percentage and e-value. The protein sequences were aligned using MUSCLE (<http://www.ebi.ac.uk/Tools/muscle/index.html>), the phylogenetic analysis was performed with Maximum-likelihood method using the JTT + G substitution model in MEGA-X (41) and the predicted trees were visualized in iTOL (42). Nodal support was tested using bootstrap values, which were calculated with 500 replicates.



Results

Initial protein selection

From the 100,114 protein sequences extracted from the combined *A. suum* and *A. lumbricoides* proteomes, 97,510 were analyzed with the Vacceed framework for protein subcellular location. A total of 28,085 proteins were selected for further analysis (Figure 1), from which 5,984 protein sequences were from *A. lumbricoides*, and 5,004 and 17,097 were *A. suum* protein sequences from the two *A. suum* proteomes, BioProject PRJNA80881 and BioProject PRJNA62057, respectively. The BioProject PRJNA62057 has a higher number of proteins in its proteome due to the way it was assembled, where all the expressed transcripts were kept instead of being merged into a single protein sequence. The Vacceed final scores for each protein in the three proteomes are provided in Supplementary Table 1.

The correlation assessment of the Vacceed framework demonstrated a high correlation between the results of the different Vacceed runs (Figure 1). The removal of TMHMM from the framework resulted in a slightly lowered correlation score when compared with the original Vacceed run ($r = 0.97$).

Selection of vaccine targets and identification of epitopes

Before the CD4+ Th cell binding prediction was carried out, a further 2,906 duplicate sequences were identified and

removed from the *A. suum* PRJNA62057 dataset. A total of 25,179 protein sequences were then analyzed for epitope binding to MHC-II alleles. When analyzing the final output, 32 protein sequences had epitopes predicted to bind to all the 27 MHC-II alleles in the reference set with a rank between 0.01 and 1. From this set of 32 protein sequences, 9 proteins belonged to the *A. lumbricoides* proteome, while the other 23 proteins were from the two *A. suum* proteomes, 3 from BioProject PRJNA80881 and 20 BioProject PRJNA62057. Only three of these protein sequences were predicted to be secreted proteins. A table detailing the WormBase protein identifiers and protein aa length for these 32 candidate proteins is provided in Supplementary Table 2.

The analysis with AllerTop 2.0 and AllergenFP did not predict any allergens in the 32 protein sequences. However, antigen prediction tools classified four proteins sequences as non-antigenic, and these proteins were disregarded. The remaining protein sequences that had similar results between their respective orthologs in *A. lumbricoides* and *A. suum* were grouped into five clusters, represented by distinct genes in all three *Ascaris* genomes. The five genes and their respective transcripts are predicted to be responsible for the transcription of four membrane transporters and one cell adhesion protein.

Four genes and respective proteins were selected as promising vaccination targets based on epitope exposure to the host immunological system: “Voltage-dependent T-type calcium channel subunit alpha” [ATtype (WormBase gene identifiers: *GS_24322*, *ALUE_0000418301* and *AgB13X_g094*)], “Piezo-type mechanosensitive ion channel component” [APiezo (WormBase gene identifiers: *GS_03113*, *ALUE_0000666901*

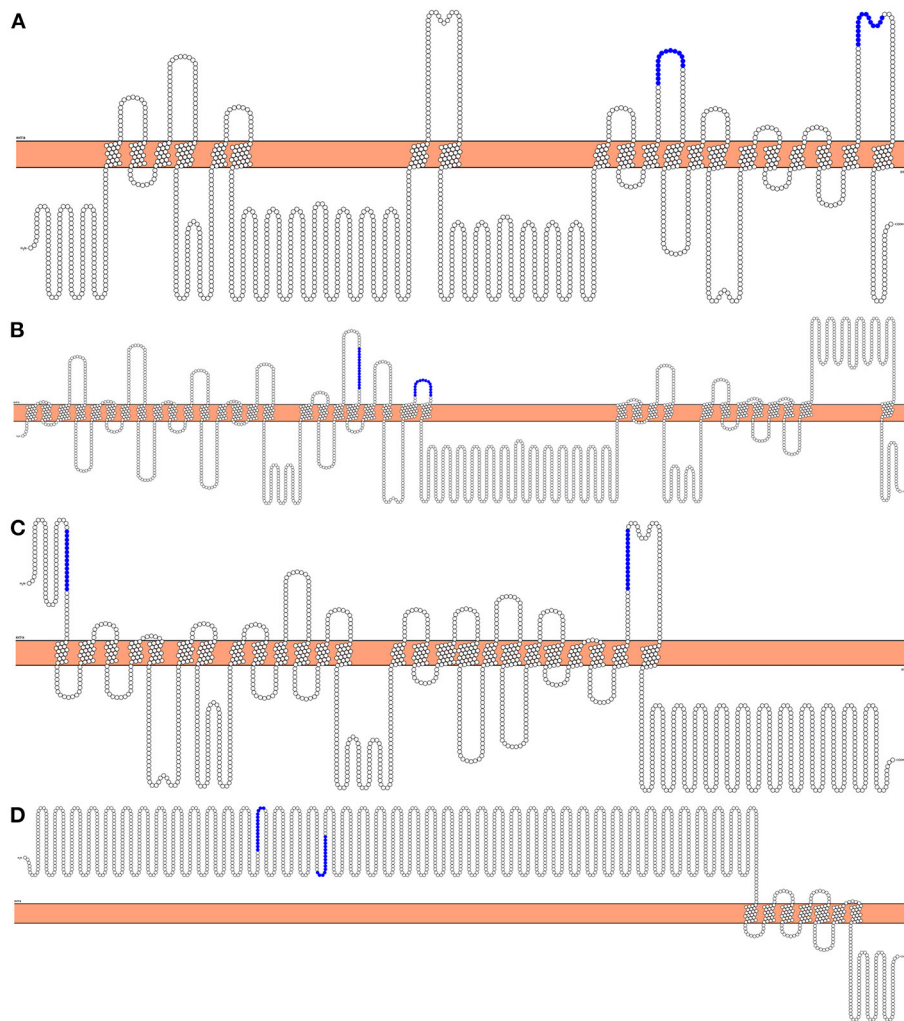


FIGURE 2

2D representation of the topology of predicted vaccination targets using Phobius predictions. The horizontal bar in the middle represents the membrane, above it is the extracellular domain, and below is the intracellular domain. Highlighted are the Cd4+ Th epitopes chosen to be incorporated in future vaccination assays. **(A)** “*AgB13X_g094_t05*” as the best expressed protein transcript for ATtype. **(B)** “*AgR007_g063_t01*” as the best expressed transcript for APiezo. **(C)** “*AgR007_g282_t05*” as the best expressed protein transcript for ALtype. **(D)** “*AgB13X_g096_t06*” as the best expressed protein transcript for AProto.

and *AgR007_g063*], “Voltage-dependent L-type calcium channel subunit alpha” [ALtype (WormBase gene identifiers: *GS_04697*, *ALUE_0001482301* and *AgR007_g282*)], and “Protocadherin-like” [AProto (WormBase gene identifiers: *GS_06422*, *ALUE_0000418601* and *AgB13X_g096*)]. According to the most recent transcriptomic data, both ATtype and AProto are highly expressed in the ovaries while APiezo and ALtype are highly expressed in the muscle of adults (4, 19). These proteins were predicted to have both CD4+ Th cell and B cell binding epitopes in extracellular areas (Figure 2). For each protein, two non-allergen CD4+ Th cell binding

epitopes were selected based on the results of the Allertop 2.0, MHCII-IEDB tool and their presence in the extracellular areas of the protein (Table 1). Two B cell epitopes were also selected for each target. Supplementary Table 3 lists the epitopes found for each predicted vaccine target using MHCII-IEDB that scored between 0 and 1 and the MHC-II alleles they were predicted to bind to (only epitopes that were predicted to bind to two or more MHC-II alleles were retrieved). None of the selected epitopes were in the direct vicinity of predicted glycosylation sites (Supplementary material 4). A workflow diagram of the analysis can be seen in Figure 3.

TABLE 1 Proteins and respective epitopes identified as potential vaccination targets against *Ascaris* spp.

Protein function	Protein Family	Species	WormBase Protein transcript	Selected CD4+ Th cell binding epitopes (aa)	Selected B-cell epitopes (aa)
Membrane transporter	Piezo Family	<i>A. suum</i>	GS_03113	NCLKYFANFFFYRFG	LLSVHLKNDDDSIEST
			AgR007_g063_t01	SLFLRPMRVALALLN	VDPSFDPVIPKEEVI
			AgR007_g063_t02		
			AgR007_g063_t03		
			AgR007_g063_t04		
			AgR007_g063_t06		
		<i>A. lumbricoides</i>	ALUE_0000666901		
	Voltage-dependent calcium channel, T-type, alpha-1 subunit	<i>A. suum</i>	GS_24322	LRLLRALRPLRVINR	DATGVDMQPVENYN
			AgB13X_g094_t03	FKNFGMAFLTLFRIA	SIPPKSVER
			AgB13X_g094_t04		
AgB13X_g094_t05					
	<i>A. lumbricoides</i>	ALUE_0000418301			
Voltage-dependent calcium channel, L-type, alpha-1 subunit	<i>A. suum</i>	GS_04697	NNNFHTFPAAILVLF	ALNDETHIHRNNN	
		AgR007_g282_t01	ERSLLCLTLNPLRK	SNEEDRGPVYNA	
		AgR007_g282_t02			
		AgR007_g282_t14			
	<i>A. lumbricoides</i>	ALUE_0001482301			
Cell adhesion	No predicted Family.	<i>A. suum</i>	GS_06422	HTFRRFITAISLLDR	LSQSDHHILPRFANFVDDRTESLRSVTIQLLCSLR
	Protocadherin-like proteins		AgB13X_g096_t02	NQEGVVHILSRKIFD	RQQFTLTFPYFSDGKFK
		<i>A. lumbricoides</i>	ALUE_0000418601		

aa, Amino acid; *A. suum*, *Ascaris suum*; *A. lumbricoides*, *Ascaris lumbricoides*.

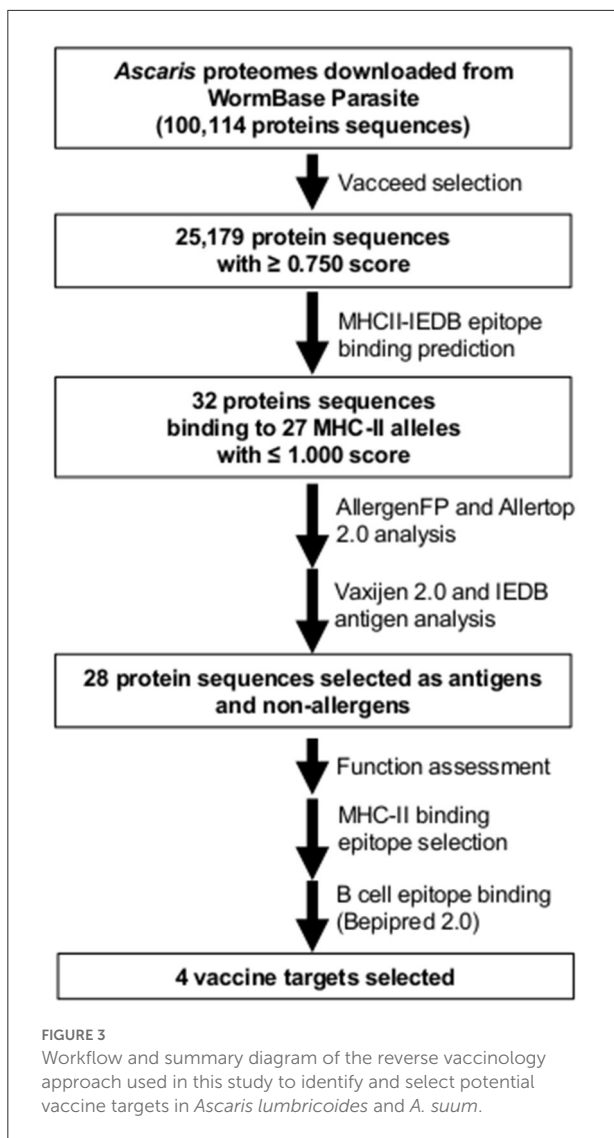
Phylogenetic relationships of vaccine targets identified across nematode species

To analyse the potential relationship between the predicted vaccine targets and proteins in other nematodes, orthologs of the selected vaccine targets were retrieved from other nematode proteomes present in WormBase Parasite (Supplementary Table 5), aligned and used for the generation of a phylogenetic tree. There is a general clustering of the proteins within each nematode clade. The exception occurred in the APiezo orthologs where clade III nematodes were separated in two different groups, with the Ascaridomorpha nematodes (e.g., *Ascaris*, *Toxocara* and *Parascaris*) being more closely related to the nematodes in Clade IV and the Spiruromorpha nematodes (e.g., *Onchocerca*, *Dirofilaria*, *Loa*, *Brugia* and *Wuchereria*) being closer to the Clade V nematodes. These relationships can be seen in Figure 4.

Discussion

In this study we have used bioinformatic approaches to identify four different *Ascaris* proteins that could be included

in multi-epitope vaccines against human and pig *Ascaris* infections. These genes are highly expressed in two distinct regions of the parasite: the ovaries and early egg stages, in the case of ATtype and Aproto, and in the muscle, the case for APiezo and ALtype (4, 19). While both ATtype and ALtype are predicted to be calcium channels that promote calcium import with muscle and smooth muscle contraction through a voltage mechanism, APiezo is a mechanosensitive calcium channel that in *C. elegans* was shown to affect reproductive tissue development and malfunction (43). One interesting relationship shown in Figure 4B is how APiezo appears to be more closely related between blood feeding parasites, such as *Necator americanus* and filarial parasites such as *Loa loa*, suggesting a role in adaptation to contact with blood. The AProto protein is the most unique as to our knowledge, as protocadherins have not been reported before in nematodes. The structural domains appear to be more closely related to that of Flamingo/Stan cadherins due to the presence of both laminin-G and EGF-like receptors with seven transmembrane domains (with these last highlighted in Figure 2D) (4, 44). Being highly transcribed in the ovaries and early egg stages, while predicted to be responsible for homophilic cell adhesion, AProto could have a role in early oocyte development. Although present in regions usually considered difficult recognize by



the host immune system, such as the parasite's muscle and ovaries, an IgG immunoblot assay showed that this is possible, as proteins highly transcribed in the muscle, ovaries and intestines of the parasite were recognized by serum from pigs infected with *Ascaris* (45). It is also interesting to highlight the close relationship between the predicted vaccine targets and the orthologs in other nematodes, suggesting that these orthologs might also be useful in the control of the respective species.

Previous attempts to generate a vaccine against ascariasis used crude extracts, recombinant proteins and, more recently, chimeric proteins (9–12). The highest lung larvae burden reduction achieved using a multi-epitope or recombinant protein vaccine was 73.5%, when using a chimeric protein containing B-cell epitopes of the As14, As16 and As37 *Ascaris* proteins (12). In comparison, vaccination against trichuriasis

in mice resulted in up to 97% reduction of adult nematodes (46). The vaccine against *Trichuris* is based on recombinant proteins and showed efficacies vastly superior to similar vaccines against ascariasis. With vaccine development against ascariasis lagging behind other parasitic diseases, there is a need to discover other antigens to be tested as vaccine candidates. The reverse vaccinology approach used in our study is based on genomic and proteomic data to predict which proteins may be usable as vaccine candidates prior to new *in vivo* studies. This methodology allows researchers to focus down vaccination assays to a smaller set of proteins, effectively reducing costs and time. This reverse vaccinology approach has been successfully applied to identify vaccination targets for other helminths, for example *T. canis* (14), *T. muris* (13), *Echinococcus granulosus* (47), and *Schistosoma mansoni* (48, 49).

Our reverse vaccinology analysis used all the proteins predicted in the three *Ascaris* proteomes. Previous *in silico* studies on vaccine candidate prediction in *Ascaris* focused exclusively on secreted proteins (50). The workflow applied in this study allows the testing of all the proteins predicted from a genome analysis, without automatically excluding non-secreted proteins. In a recent study, an *A. lumbricoides* multi-epitope vaccine candidate was developed using *in silico* methodology and proteins were selected based on their binding to the HLA-DRB1*07:01 and HLA-DRB1*15:01 MHC-II alleles (51). Although these MHC-II alleles are known to recognize *Ascaris* antigens in humans, they only cover up to 15% of the human population in areas where human ascariasis is endemic (50). This might prove detrimental in *in vivo* studies that cover population that do not have these MHC-II alleles, limiting its usefulness.

Each of the three genomes included in the analysis were predicted to have over 5,000 proteins that could be further investigated as good vaccine targets according to Vaxceed, which corresponds to 25–29% of all the proteins present. As the number of secreted proteins in *Ascaris* is predicted to be 254 proteins (50), this suggests that the number of potential vaccination targets might be vastly superior to the ones that are usually investigated in these species and other helminths. The final candidates are all predicted to be non-secreted proteins. This contrasts with most of the previously studied vaccine candidates, except for the muscle membrane-bound As37 protein (52). As37 was not predicted to be a good target in Vaxceed (with scores of 0.001 in all three genomes), showing some limitations to the method we used. However, our predictions and the protection achieved with the As37 recombinant protein support the idea that only targeting secreted proteins can be detrimental to the selection of good vaccination targets in *Ascaris*. Recent work in *Toxocara canis*, a parasite of the same family as *Ascaris*, showed that membrane proteins are capable to induce protection in a mouse model,

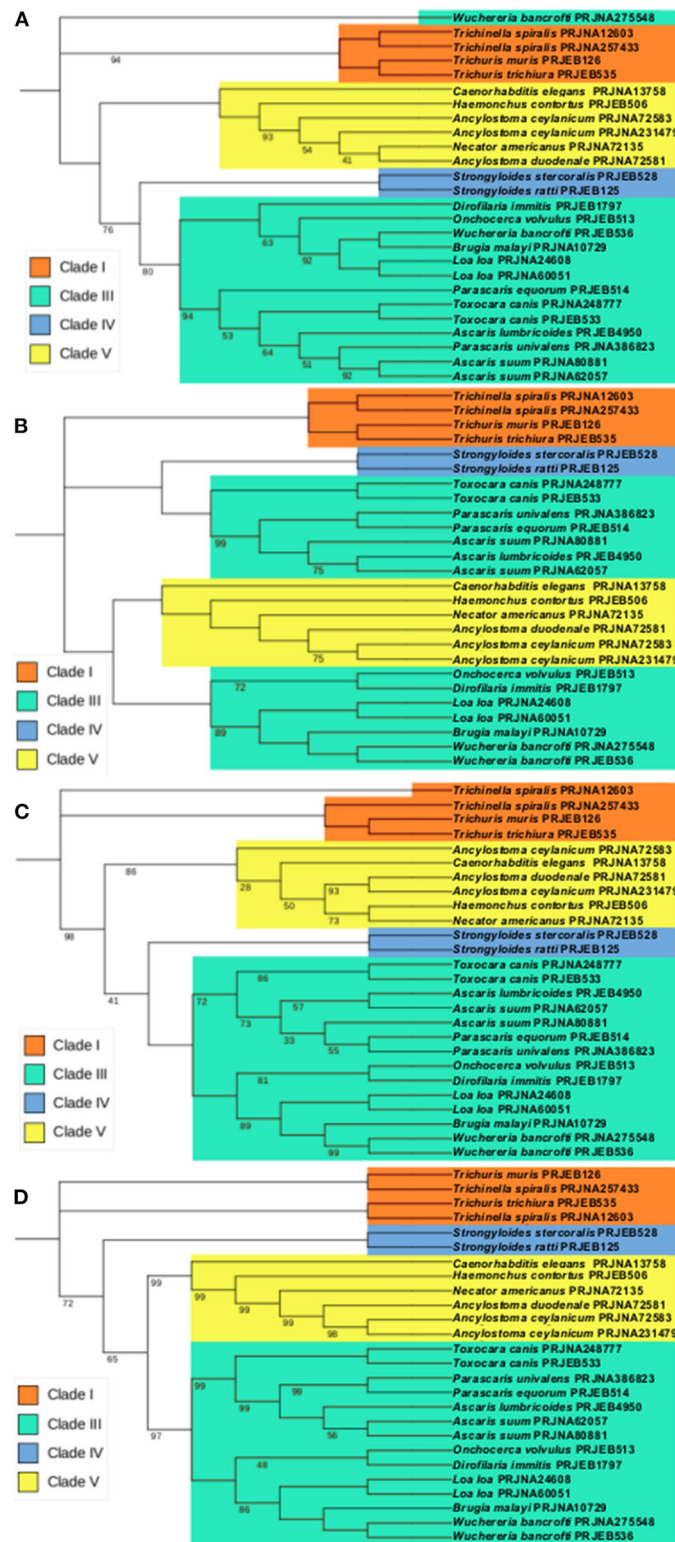


FIGURE 4

Maximum likelihood phylogenies using the predicted vaccine targets and orthologs found in other nematodes. The trees were inferred using bootstrap values with 500 replicates. Values on the nodes represent the percentage of bootstrap support values and nodes without bootstrap values were supported by 100% of the 500 replicates. Each sequence is identified by the nematode species followed by the respective BioProject (proteins transcripts used in these analyses are individualized in Supplementary Table 4). (A) Phylogenetic tree of ATtype orthologs. (B) Phylogenetic tree of APiezo orthologs. (C) Phylogenetic tree of ALtype orthologs. (D) Phylogenetic tree of AProto orthologs.

reinforcing the idea that it is an error to disregard these proteins when looking for vaccination targets in nematodes (14).

The underlying host's immune responses against *Ascaris* are still only partially understood (53). This is a disadvantage when selecting the right tools to predict which proteins could be useful as vaccination targets. MHC-II molecules appear to have a prominent role in the control of nematode infections in mammals, including *A. lumbricoides* infections in humans (53, 54). This role makes the discovery of epitopes that bind to these molecules a priority for the design of multi-epitope/subunit-based vaccines. The MHCII-IEDB tool was chosen for this purpose as it was used in selecting epitopes for vaccination assays against other nematodes and is one of the most accurate tools available (13, 50). A reference set of 27 different alleles was chosen due to the fact that heterogeneity throughout the human population leads to different immune responses (50). Thus, we wanted to select proteins that would be able to induce a helpful immune reaction in a large portion of the population, and not just focus on one specific allele. Unfortunately, this biases toward selecting larger proteins as they contain a larger number of epitopes. These predictions were made only for human alleles due to unavailability of similar open-access bioinformatic tools for swine MHC-II alleles. This is a limitation of the reverse vaccinology approach to discovering vaccine targets in pigs. Bepipred 2.0 predicted the presence of B cell binding epitopes in same proteins which would enhance the host immune response against these proteins in their native form in the parasite. Another feature that needs to be considered when selecting vaccination targets and respective epitopes for a multi-epitope vaccine is the presence of glycosylation sites in the native antigen. In general, glycosylation is known to play a key role in regulating T-cell activity and highly glycosylated areas in proteins will downregulate the activity of these cells (55, 56). Selecting epitopes that are not predicted to be in these glycosylation sites would then improve the chance of the host reacting to the native antigen after receiving the multi-epitope vaccine (56). However, in *Haemonchus contortus* vaccination assays, glycosylation was fundamental for the recognition and protection induced by some antigens (57). Glycosylation in *Ascaris* proteins and how it affects the host is still not entirely understood and requires more work, but it is known to be a strategy that the parasite has developed to modulate the host immune response and make it less responsive (58, 59). Another consideration is how difficult or impossible it is to mimic the protein glycosylation that happens in nematodes using the most common recombinant protein production systems, such as the ones using *Escherichia coli* (60). To replicate glycosylation the use of other systems (e.g., mammalian cell lines) is required that tend to be less productive and harder to maintain than bacterial systems (61).

Whilst the proteins we selected have been predicted to be useful and have similarities to other proteins tested in vaccination assays against other nematodes, there is still the

need to test them *in vitro* and *in vivo*. The next step should be to confirm *in vitro* that both humans and pigs are able to recognize these proteins targets and their respective epitopes. Should these proteins prove useful in stimulating an immune reaction in the host, we propose that the proteins and respective epitopes identified in this work should be incorporated into a multi-epitope vaccine which, ideally, would include CD4+ Th cell and B-cell binding epitopes from other proteins, such as As14, As16 and As37. Ideally, such a vaccine should be tested using a pig model to assess its potential effect on both larvae and adult *Ascaris*, impossible to assess using a mouse model. This is more relevant when testing for the utility of incorporating ATtype and AProto epitopes due to protein higher expression in the ovaries of adult female *Ascaris* and the released eggs (4, 19). Should these proteins not prove useful, a protocol optimization should be done to identify new targets so that, in the future, it might also be applied to other species that are in need of further studies.

In conclusion, this study highlights the role that reverse vaccinology and *in silico* methodology can play in identification of vaccine candidates for parasitic diseases. The genome-wide approach, without bias toward secreted proteins, led to the prediction of four novel candidates that were not identified in previous studies but show the promise of promoting a useful immune response in vaccination assays against *A. lumbricoides* and *A. suum*. These proteins should now be tested *in vitro* and in combination with already known vaccine targets. Ultimately, the findings of this study will support the future development of a vaccine against both ascariasis in humans and pigs, thus promoting the health of both populations by reducing the need to use mass-drug administration and decreasing the risk of anthelmintic resistance appearing.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary materials, further inquiries can be directed to the corresponding author.

Author contributions

Concept and design, data interpretation, and critical review and editing of the manuscript: FE, AV, SL, and MB. Data acquisition and analysis and manuscript draft: FE. All authors contributed to the article and approved the submitted version.

Funding

This study was funded by a Doctoral College Studentship Award from the University of Surrey awarded to FE.

Acknowledgments

The authors would like to thank Dr. Stephen Goodswen for assistance and troubleshooting with the Vacceed tool and Dr. Pritesh Taylor for help with computer resources. The present work is available as a preprint in Evangelista et al. (62). This work was presented as a lecture at the 74th AVTRW Annual Conference (Online/Zoom, hosted from Glasgow, 14th-15th September 2020), BAVP Early Career Parasitologists meeting (Online/Zoom, 29th-30th April 2021) and LE STUDIUM Conference (Online/Zoom, 29th November to 1st December 2021).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher,

the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fvets.2022.1014198/full#supplementary-material>

SUPPLEMENTARY TABLE 1

The Vacceed final scores for each proteome in the three *Ascaris* proteomes.

SUPPLEMENTARY TABLE 2

Protein sequences predicted to have epitopes that bind to all 27 MHC-II alleles, used as a reference set, with a score between 0.01 and 1 in the MHCII-IEDB tool.

SUPPLEMENTARY TABLE 3

The epitopes found for each predicted vaccine target using MHCII-IEDB that scored between 0 and 1 and the MHC-II alleles they were predicted to bind to (only epitopes that were predicted to bind to two or more MHC-II alleles were retrieved).

SUPPLEMENTARY TABLE 4

Glycosylation assessment.

SUPPLEMENTARY TABLE 5

List of orthologs used in the phylogenetic analysis.

References

- Kyu HH, Abate D, Abate KH, Abay SM, Abbafati C, Abbasi N, et al. Global, regional, and national disability-adjusted life-years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet*. (2018) 392:1859-922. doi: 10.1016/S0140-6736(18)32335-3
- Katakam KK, Thamsborg SM, Dalsgaard A, Kyvsgaard NC, Mejer H. Environmental contamination and transmission of *Ascaris suum* in Danish organic pig farms. *Parasit Vectors*. (2016) 9:80. doi: 10.1186/s13071-016-1349-0
- Dold C, Holland CV. *Ascaris* and ascariasis. *Microbes Infect*. (2011) 13:632-7. doi: 10.1016/j.micinf.2010.09.012
- Easton A, Gao S, Lawton SP, Bennuru S, Khan A, Dahlstrom E, et al. Molecular evidence of hybridization between pig and human *Ascaris* indicates an interbred species complex infecting humans. *Elife*. (2020) 9:e61562. doi: 10.7554/eLife.61562
- World Health Organization. *Ending the neglect to attain the Sustainable Development Goals - A Road Map for Neglected Tropical Diseases 2021-2030*. Geneva: World Health Organization (2020). Licence: CC BY-NC-SA 3.0 IGO.
- Furtado LFF, da Silva Medeiros C, Zuccherato LW, Alves WP, de Oliveira VNGM, da Silva VJ, et al. First identification of the benzimidazole resistance-associated F200Y SNP in the beta-tubulin gene in *Ascaris lumbricoides*. *PLoS ONE*. (2019) 14:e0224108. doi: 10.1371/journal.pone.0224108
- Krücken J, Fraundorfer K, Mugisha JC, Ramünke S, Siftt KC, Geus D, et al. Reduced efficacy of albendazole against *Ascaris lumbricoides* in Rwandan schoolchildren. *Int J Parasitol Drugs Drug Resist*. (2017) 7:262-71. doi: 10.1016/j.ijpddr.2017.06.001
- Hotez PJ, Strych U, Lustigman S, Bottazzi ME. Human anthelmintic vaccines: Rationale and challenges. *Vaccine*. (2016) 34:3549-55. doi: 10.1016/j.vaccine.2016.03.112
- Gazzinelli-Guimarães AC, Gazzinelli-Guimarães PH, Nogueira DS, Oliveira FMS, Barbosa FS, Amorim CCO, et al. IgG induced by vaccination with *Ascaris suum* extracts is protective against infection. *Front Immunol*. (2018) 9:25-35. doi: 10.3389/fimmu.2018.02535
- Tsuji N, Suzuki K, Kasuga-Aoki H, Matsumoto Y, Arakawa T, Ishiwata K, et al. Intranasal immunization with recombinant *Ascaris suum* 14-kilodalton antigen coupled with cholera toxin B subunit induces protective immunity to *A. suum* infection in mice. *Infect Immun*. (2001) 69:7285-92. doi: 10.1128/IAI.69.12.7285-7292.2001
- Tsuji N, Miyoshi T, Islam MK, Isobe T, Yoshihara S, Arakawa T, et al. Recombinant *Ascaris* 16-Kilodalton protein-induced protection against *Ascaris suum* larval migration after intranasal vaccination in pigs. *J Infect Dis*. (2004) 190:1812-20. doi: 10.1086/425074
- de Castro JC, de Almeida LV, Cardoso MS, Oliveira FMS, Nogueira DS, Reis-Cunha JL, et al. Vaccination with chimeric protein induces protection in murine model against ascariasis. *Vaccine*. (2021) 39:394-401. doi: 10.1016/j.vaccine.2020.11.046
- Zawawi A, Forman R, Smith H, Mair I, Jibril M, Albaqshi MH, et al. In silico design of a T-cell epitope vaccine candidate for parasitic helminth infection. *PLoS Pathog*. (2020) 16:e1008243. doi: 10.1371/journal.ppat.1008243
- Salazar Garcés LF, Santiago LF, de Oliveira Santos SP, Jaramillo Hernández DA, da Silva MB, Dos Santos Alves V, et al. Immunogenicity and protection induced by recombinant *Toxocara canis* proteins in a murine model of toxocarasis. *Vaccine*. (2020) 38:4762-72. doi: 10.1016/j.vaccine.2020.04.072
- Howe KL, Bolt BJ, Shafie M, Kersey P, Berriman M. WormBase ParaSite - a comprehensive resource for helminth genomics. *Mol Biochem Parasitol*. (2017) 215:2-10. doi: 10.1016/j.molbiopara.2016.11.005
- International Helminth Genomes Consortium. Comparative genomics of the major parasitic worms. *Nat Genet*. (2019) 51:163-74. doi: 10.1038/s41588-018-0262-1
- Jex AR, Liu S, Li B, Young ND, Hall RS, Li Y, et al. *Ascaris suum* draft genome. *Nature*. (2011) 479:529-33. doi: 10.1038/nature10553
- Wang J, Mitreva M, Berriman M, Thorne A, Magrini V, Koutsovoulos G, et al. Silencing of germline-expressed genes by DNA elimination in somatic cells. *Dev Cell*. (2012) 23:1072-80. doi: 10.1016/j.devcel.2012.09.020

19. Wang J, Gao S, Mostovoy Y, Kang Y, Zagoskin M, Sun Y, et al. Comparative genome analysis of programmed DNA elimination in nematodes. *Genome Res.* (2017) 27:2001–14. doi: 10.1101/gr.225730.117
20. Hall TA. BIOEDIT a user-friendly biological sequence alignment editor and analysis program for windows 95/98/ NT. *Nucleic Acids Symp Ser.* (1999) 41:95–8.
21. Goodswen SJ, Kennedy PJ, Ellis JT. Vacceed: a high-throughput in silico vaccine candidate discovery pipeline for eukaryotic pathogens based on reverse vaccinology. *Bioinformatics.* (2014) 30:2381–3. doi: 10.1093/bioinformatics/btu300
22. Horton P, Park K-J, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, et al. WoLF PSORT: protein localization predictor. *Nucleic Acids Res.* (2007) 35:W585–7. doi: 10.1093/nar/gkm259
23. Almagro Armenteros JJ, Sønderby CK, Sønderby SK, Nielsen H, Winther O. DeepLoc: prediction of protein subcellular localization using deep learning. *Bioinformatics.* (2017) 33:3387–95. doi: 10.1093/bioinformatics/btx431
24. Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, et al. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol.* (2019) 37:420–3. doi: 10.1038/s41587-019-0036-z
25. Almagro Armenteros JJ, Salvatore M, Emanuelsson O, Winther O, von Heijne G, Elofsson A, et al. Detecting sequence signals in targeting peptides using deep learning. *Life Sci Alliance.* (2019) 2:e201900429. doi: 10.26508/lsa.201900429
26. Krogh A, Larsson B, Von Heijne G, Sonnhammer ELL. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *J Mol Biol.* (2001) 305:567–80. doi: 10.1006/jmbi.2000.4315
27. Käll L, Krogh A, Sonnhammer ELL. A combined transmembrane topology and signal peptide prediction method. *J Mol Biol.* (2004) 338:1027–36. doi: 10.1016/j.jmb.2004.03.016
28. Palmieri N, Shrestha A, Ruttkowski B, Beck T, Vogl C, Tomley F, et al. The genome of the protozoan parasite *Cystoisospora suis* and a reverse vaccinology approach to identify vaccine candidates. *Int J Parasitol.* (2017) 47:189–202. doi: 10.1016/j.ijpara.2016.11.007
29. R Core Team. *R: A Language and Environment for Statistical Computing.* Vienna (2020).
30. RStudio Team. *RStudio: Integrated Development Environment for R.* Boston, MA (2019).
31. Wei T, Simko V. *R package “corrplot”: Visualization of a Correlation Matrix* (2017).
32. Fleri W, Paul S, Dhanda SK, Mahajan S, Xu X, Peters B, et al. The immune epitope database and analysis resource in epitope discovery and synthetic vaccine design. *Front Immunol.* (2017) 8:278. doi: 10.3389/fimmu.2017.00278
33. Greenbaum J, Sidney J, Chung J, Brander C, Peters B, Sette A. Functional classification of class II human leukocyte antigen (HLA) molecules reveals seven different supertypes and a surprising degree of repertoire sharing across supertypes. *Immunogenetics.* (2011) 63:325–35. doi: 10.1007/s00251-011-0513-0
34. Kolaskar AS, Tongaonkar PC. A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS Lett.* (1990) 276:172–4. doi: 10.1016/0014-5793(90)80535-Q
35. Doytchinova IA, Flower DR. Identifying candidate subunit vaccines using an alignment-independent method based on principal amino acid properties. *Vaccine.* (2007) 25:856–66. doi: 10.1016/j.vaccine.2006.09.032
36. Dimitrov I, Bangov I, Flower DR, Doytchinova I. AllerTOP v. 2—a server for in silico prediction of allergens. *J Mol Model.* (2014) 20:2278. doi: 10.1007/s00894-014-2278-5
37. Dimitrov I, Naneva L, Doytchinova I, Bangov I. AllergenFP: allergenicity prediction by descriptor fingerprints. *Bioinformatics.* (2013) 30:846–51. doi: 10.1093/bioinformatics/btt619
38. Gupta R, Brunak S. Prediction of glycosylation across the human proteome and the correlation to protein function. *Pac Symp Biocomput.* (2002) 310–22.
39. Omasits U, Ahrens CH, Müller S, Wollscheid B, Protter: interactive protein feature visualization and integration with experimental proteomic data. *Bioinformatics.* (2014) 30:884–6. doi: 10.1093/bioinformatics/btt607
40. Jespersen MC, Peters B, Nielsen M, Marcattili P. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Res.* (2017) 45:W24–9. doi: 10.1093/nar/gkx346
41. Kumar S, Stecher G, Li M, Nkay C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol.* (2018) 35:1547–9. doi: 10.1093/molbev/msy096
42. Letunic I, Bork P. Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* (2021) 49:W293–6. doi: 10.1093/nar/gkab301
43. Bai X, Bouffard J, Lord A, Brugman K, Sternberg PW, Cram EJ, et al. *Caenorhabditis elegans* PIEZO channel coordinates multiple reproductive tissues to govern ovulation. *Elife.* (2020) 9:e53603. doi: 10.7554/eLife.53603
44. Hardin J, Lynch A, Loveless T, Pettitt J. Chapter eleven - cadherins and their partners in the nematode worm *Caenorhabditis elegans*. In: van Roy FBT-P, editor. *The Molecular Biology of Cadherins.* Academic Press (2013). p. 239–62.
45. González-Miguel J, Morchón R, Gussoni S, Bossetti E, Hormaeche M, Kramer LH, et al. Immunoproteomic approach for identification of *Ascaris suum* proteins recognized by pigs with porcine ascariasis. *Vet Parasitol.* (2014) 203:343–8. doi: 10.1016/j.vetpar.2014.03.031
46. Gomez-Samblas M, García-Rodríguez JJ, Trelis M, Bernal D, Lopez-Jaramillo FJ, Santoyo-Gonzalez F, et al. Self-adjuncting C18 lipid vinyl sulfone-PP2A vaccine: study of the induced immunomodulation against *Trichuris muris* infection. *Open Biol.* (2017) 7:170031. doi: 10.1098/rsob.170031
47. Gan W, Zhao G, Xu H, Wu W, Du W, Huang J, et al. Reverse vaccinology approach identify an *Echinococcus granulosus* tegumental membrane protein enolase as vaccine candidate. *Parasitol Res.* (2010) 106:873–82. doi: 10.1007/s00436-010-1729-x
48. de Souza C, Lopes MD, De Oliveira FM, Passos MJF, Ferreira LCG, Faria BF, et al. Rational selection of immunodominant and preserved epitope Sm043300e from *Schistosoma mansoni* and design of a chimeric molecule for biotechnological purposes. *Mol Immunol.* (2018) 93:133–43. doi: 10.1016/j.molimm.2017.11.019
49. Oliveira FM, Coelho IE V, Lopes MD, Taranto AG, Junior MC, Santos LLD, et al. The use of reverse vaccinology and molecular modeling associated with cell proliferation stimulation approach to select promiscuous epitopes from *Schistosoma mansoni*. *Appl Biochem Biotechnol.* (2016) 179:1023–40. doi: 10.1007/s12010-016-2048-1
50. Ebner F, Morrison E, Bertazzon M, Midha A, Hartmann S, Freund C, et al. CD4+ Th immunogenicity of the *Ascaris* spp. secreted products. *NPJ Vaccines.* (2020) 5:25. doi: 10.1038/s41541-020-0171-z
51. Kaur R, Arora N, Rawat SS, Keshri AK, Singh N, Show SK, et al. Immunoinformatics driven construction of multi-epitope vaccine candidate against *Ascaris lumbricooides* using its entire immunogenic epitopes. *Expert Rev Vaccines.* (2021) 3:1–13. doi: 10.1080/14760584.2021.1974298
52. Versteeg L, Wei J, Liu Z, Keegan B, Fujiwara RT, Jones KM, et al. Protective immunity elicited by the nematode-conserved A37 recombinant protein against *Ascaris suum* infection. *PLoS Negl Trop Dis.* (2020) 14:e0008057. doi: 10.1371/journal.pntd.0008057
53. Zawawi A, Else KJ. Soil-transmitted helminth vaccines: are we getting closer? *Front Immunol.* (2020) 11:576748. doi: 10.3389/fimmu.2020.576748
54. Garamszegi LZ, Nunn CL. Parasite-mediated evolution of the functional part of the MHC in primates. *J Evol Biol.* (2011) 24:184–95. doi: 10.1111/j.1420-9101.2010.02156.x
55. Wolfert MA, Boons G-J. Adaptive immune activation: glycosylation does matter. *Nat Chem Biol.* (2013) 9:776–84. doi: 10.1038/nchembio.1403
56. Pereira MS, Alves I, Vicente M, Campar A, Silva MC, Padrão NA, et al. Glycans as key checkpoints of T cell activity and function. *Front Immunol.* (2018) 9:2754. doi: 10.3389/fimmu.2018.02754
57. van Stijn CMW, van den Broek M, Vervelde L, Alvarez RA, Cummings RD, Tefsen B, et al. Vaccination-induced IgG response to Gal α 1-3GalNAc glycan epitopes in lambs protected against *Haemonchus contortus* challenge infection. *Int J Parasitol.* (2010) 40:215–22. doi: 10.1016/j.ijpara.2009.07.009
58. van Die I, Cummings RD. Glycan gimmickry by parasitic helminths: a strategy for modulating the host immune response? *Glycobiology.* (2010) 20:2–12. doi: 10.1093/glycob/cwp140
59. Favoretto BC, Casabuono AAC, Portes-Junior JA, Jacysyn JF, Couto AS, Faquim-Mauro EL. High molecular weight components containing N-linked oligosaccharides of *ascaris suum* extract inhibit the dendritic cells activation through DC-SIGN and MR. *Mol Immunol.* (2017) 87:33–46. doi: 10.1016/j.molimm.2017.03.015
60. Bhatwa A, Wang W, Hassan YI, Abraham N, Li XZ, Zhou T. Challenges associated with the formation of recombinant protein inclusion bodies in *escherichia coli* and strategies to address them for industrial applications. *Front Bioeng Biotechnol.* (2021) 9:65. doi: 10.3389/fbioe.2021.630551. Epub ahead of print.
61. Tripathi NK, Shrivastava A. Recent developments in bioprocessing of recombinant proteins: expression hosts and process development. *Front Bioeng Biotechnol.* (2019) 7:420. doi: 10.3389/fbioe.2019.00420
62. Evangelista FMD, van Vliet AHM, Lawton SP, Betson M. A reverse vaccinology approach identifies putative vaccination targets in the zoonotic nematode *Ascaris*. *bioRxiv.* (2022). doi: 10.1101/2022.04.27.489657