

In silico analysis for identification of tick phagotopes selected by phage-displayed libraries

Análises in silico na identificação de fagotopos de carrapatos selecionados por bibliotecas de *phage display*

Carlos Roberto Prudencio^{1*}; Rafael Nascimento¹; Moacir Marchiori Filho²; Andrea de Oliveira Marques Marra²; Guilherme Rocha Lino de Souza³; Juliana Franco Almeida¹; Rone Cardoso¹; Matias Pablo Juan Szabó⁴; Luiz Ricardo Goulart¹

¹Laboratório de Bionanotecnologia, Instituto de Genética e Bioquímica, Universidade Federal de Uberlândia – UFU

²Vallée S/A., São Paulo - SP, Brasil

³Departamento de Produção Vegetal Alto Universitário, Centro Agropecuário – CCA, Universidade Federal do Espírito Santo – UFES

⁴Faculdade de Medicina Veterinária, Universidade Federal de Uberlândia – UFU

Received May 31, 2008

Accepted February 27, 2009

Abstract

Phage display techniques have been widely employed to map epitope structures which have served as the basis for developing molecular vaccines. We have applied this technique to map specific epitopes of *Rhipicephalus (Boophilus) microplus*. In the present study, we have identified the potential immunogens using a process in which the selected phage clones were analyzed through bioinformatics, prior to final field tests. The present study demonstrates the feasibility of identifying important *R. (B.) microplus* phagotopes for vaccine development through screening of phage-displayed random peptide libraries and bioinformatics tools.

Keywords: bioinformatics, phagotopes, phage display, tick, vaccines.

Resumo

Técnicas de *phage display* têm sido amplamente empregadas para o mapeamento de epítomos os quais tem servido como base para o desenvolvimento de vacinas moleculares. Esta técnica foi aplicada no mapeamento de epítomos do *Rhipicephalus (Boophilus) microplus*. Neste estudo, potenciais imunógenos foram identificados pela adoção de um processo em que os clones de fagos foram analisados por bioinformática, previamente à realização dos testes. Os resultados demonstraram a possibilidade da identificação de importantes mimetopos do *R. (B.) microplus* para o desenvolvimento de vacinas através da seleção de bibliotecas de phage display associada à análise de bioinformática.

Palavras-chave: bioinformática, fagotopos, *phage display*, carrapatos, vacinas.

Introduction

The need for high-throughput bioinformatic methods to characterize gene function is being driven by the generation of sequences at a rate far beyond our ability to carry out experimental functional analyses. An important clue to discovering gene function is the identification of binding partners. Phage display technology is a widely used tool for identifying either protein- or small molecule-binding (MANDAVA, 2004). This methodology is now well established as an important experimental approach in designing new reagents for diagnosis of diseases and development of novel vaccines (WANG, 2007).

This technique has been widely employed to map epitope structures which have served as the basis for developing molecular

vaccines. The use of peptide epitope mimics (i.e., phagotopes or mimotopes) as vaccines has become a promising strategy. The peptide epitopes may serve as useful tools for the development of effective vaccines for the design of recombinant multi-epitope antigens, peptide-based antigens, or peptide-displaying phages themselves used for immunization (BENHAR, 2001; IRVING, 2001).

We tested six phage-displayed random peptides that were selected with seven different strategies and different stringency conditions against the purified hyperimmune serum of chickens (IgY) immunized with total proteins of larvae and adults of *R. (B.) microplus* (unpublished data). In the present study, we have applied this technique to map specific motifs of the tick *R. (B.) microplus* and to validate the peptides as potential immunogens. We adopted a multi-process of similarity analysis between the recombinant peptides and proteins of *R. (B.) microplus* deposited in the GenBank.

*Corresponding author: Carlos Roberto Prudencio

Laboratório de Genética Molecular, Instituto de Genética e Bioquímica, Universidade Federal de Uberlândia – UFU, Campus Umuarama, Bloco 2E, Sala 24, CEP 38400-902 Uberlândia - MG, Brasil; e-mail: crprudencio@gmail.com

Supported by: FINEP, CNPq, CAPES, Vallée S/A, ImunoScan.

Material and Methods

Peptide sequences obtained by phage DNA sequencing were analyzed with available bioinformatics programs (<http://relic.bio.anl.gov/programs.aspx>). FASTAskan was used to rank the proteins in descending order such that the proteins most likely to bind to a ligand would be clustered at the top. Scores were generated by calculating the similarity between each peptide sequence as compared to the length of the protein sequence of the *R. (B.) microplus*. The similarity threshold was selected both empirically and experimentally and corresponded to approximately three identities and one similarity for every five amino acids.

Next, the similarity between peptides and proteins related to the bovine *R. (B.) microplus* tick and deposited in the GenBank was investigated. Analysis was done by using the function "Search for Short Nearly Exact Matches" through the PAM30 matrix of the BLAST software program. Amino acids similar in four or more identical AA within a peptide sequence were determined as a similar motif of the corresponding protein sequences.

A search for similarity between peptides and predicted protein based on ESTs of the *R. (B.) microplus* was carried out in GenBank. For this we used the program BLAST defined as: "\$blastall -p tblastn -i peptides -d ests -e 100 -b 1 -F f -o result."

Results

The selected phage clones were sequenced, translated, and analyzed using bioinformatics, and generated 107 different peptides in a total of 281 sequences. Some selected phagotopes showed excellent matches with linear sequences of known proteins of *R. (B.) microplus*. Peptides that did not present significant matches with known proteins, but shared extensive homology among each other, were clustered and classified as conformational epitopes of *R. (B.) microplus*, or considered as phagotopes of non-proteic antigens.

A list of proteins classified in decreasing order of similarities to peptides was generated. Tables 1 and 2 present the *R. (B.) microplus* proteins described in GenBank, classified in accordance with the similarity to phagotopes. In Table 1, the subpopulations of selected peptides were divided in accordance with the seven selection strategies employed and were classified in decreasing order, with only the protein with the highest index of similarity for each selection strategy being represented. Table 2 shows only the 10 proteins having the highest indices of similarity. The proteins showing the highest matches in decreasing order were the precursor protein of GP80 putative secreted protein and the notch-like protein.

In Table 3, alignments were classified in decreasing order of frequencies of peptide phagotopes. The first column illustrates the ranking, in decreasing order, of proteins deposited in GenBank that contain motifs similar to the selected peptides. The third column represents the relation of the number of peptides similar to proteins (mimotopes) and the fourth column represents the number of alignments. A total of 66 different proteins referring to *R. (B.) microplus* were aligned. Three proteins con-

Table 1. Subpopulations of proteins aligned with the peptides were divided in accordance with seven adopted selection strategies and classified in decreasing order (FastaScan).

Selection	Score	Proteins (GenBank)
S1	233.33	BmTI-A
S2	178.57	Cystatin
S3	259.26	Serotonin receptor
S4	650.00	Sodium channel
S5	470.00	Carrapatin inhibiting protease
S6	666.67	BM86
S7	657.14	GP80 precursor

Table 2. Alignments classified in decreasing order of frequencies of peptide phagotopes (FastaScan).

Class	Score	Proteins (GeneBank)
1	107.92	GP80 precursor
2	69.31	putative secreted protein
3	50.50	notch-like protein
4	47.52	putative sodium channel
5	44.55	leucokinin-like peptide receptor
6	44.55	salivary gland metalloprotease
7	44.55	yolk cathepsin
8	43.56	G-protein coupled receptor
9	41.58	cell surface glycoprotein BM86
10	41.58	glucose 6-phosphate dehydrogenase isoform C

Table 3. Alignments classified in decreasing order of frequencies of peptide phagotopes (BLAST).

Number	Proteins	Frequency	Match
1	Salivary gland metalloproteases	28	35
2	Notch-like protein	27	43
3	GP80 Precursor	27	36
4	Esterase	17	19
5	Membrane B antigen	14	16
6	Calreticulin	14	14
7	Sodium channel	12	13
8	Reverse transcriptase-type protein	11	12
	P450 CYP319A1	11	12
9	Paramyosin	11	11
10	Acetylcholinesterase	10	12

taining more protein motifs similar to peptides are highlighted, with the most frequent proteins being salivary gland metalloproteases, notch-like proteins, and GP80 precursor. The proteins showing the highest matches in decreasing order were the notch-like protein, the precursor protein of GP80 protein, and the salivary gland metalloproteases.

Table 4 shows only the eight peptides having the highest indices of similarity with the predicted proteins to *R. (B.) microplus* ESTs. The function of these proteins has not been identified yet.

Discussion

This strategy was established to discriminate and reduce the number of clones, to be specific to the target. The stringent selection processes coupled with bioinformatics assays allowed us to narrow down the number of peptides that should be tested directly as immunogens and at the same time indicate the native antigens to be tested as new targets.

After the selection experiments, our preliminary results suggest that the epitope-displaying M13 phages could be used as immunogens to mapping new *R. (B.) microplus* antigens according to three findings. First, the peptide epitope sequences sharing the consensus sequence frequently correlate with many secreted proteins, indicating some of the most likely antigens directly involved. Second, alignment revealed similarities of motifs with several antigenic proteins and also the proteins having important cell functions. Third, the clones were shown to present high similarity with the tick GP80 precursor. A parallel experiment using antibody libraries from these same chickens and the scFv selected with the larval antigens could recognize a GP80 precursor that was determined by N-terminal sequencing (SOUZA, 2007). Finally, taken together, the data strongly suggest the specificity of mapping possible native tick immunogens and it may be used to build a *R. (B.) microplus* epitope bank (epitome) in the future to select and construct a multi-epitope vaccine.

The results also indicate the potential of using this methodology as an alternative approach for finding and validating new proteins. This is specially useful for *R. (B.) microplus*, whose novel antigenic proteins are of great necessity for the control of this parasitism and for pointing new directions for the development of a successful vaccine. Thus, the possibility to develop new recombinant subunit vaccines produced by using biotechnology holds great promise for the future of veterinary vaccines that apply genomic tools as new data become available.

By using bioinformatics analysis methods, phage clones and *R. (B.) microplus* proteins were determined according to a sequence of linear amino acids, which are the constituent parts of these fusion peptides. Thus, we were able to demonstrate the antigenic profile of *R. (B.) microplus* by selecting phage display libraries. Our analyses revealed the potential of this methodology for the discovery of new vaccine candidates.

Conclusion

The present study demonstrates the feasibility of identifying important *R. (B.) microplus* phagotopes and native proteins for vaccine development through the screening of phage-displayed random peptide libraries and bioinformatics tools. Future work will be done to identify and characterize additional phagotopes, and to optimize the validation process for the best vaccination results.

Table 4. Representation of similarities between the recombinant peptides and proteins predicted to ESTs of *Rhipicephalus (Boophilus) microplus* deposited in GenBank.

Sequence (Fago)	Accession (Gb)	Clone (mRNA)
Sequence ID N° 4	CV447142.1	BEAFD73
Sequence ID N° 22	CV442976.1	BEAEP08
Sequence ID N° 24	CV443271.1	BEAEQ76
Sequence ID N° 42	CK192020.1	BEAD714
Sequence ID N° 75	CV455770.1	BEAGS85
Sequence ID N° 79	CV445416.1	BEAF345
Sequence ID N° 81	CK185650.1	BEAC593
Sequence ID N° 93	CK183322.1	BEABS50

Acknowledgements

This work was supported by grants from Financiadora de Estudos e Projetos (FINEP), Vallée S/A, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG). The authors would like to thank the excellent assistance of Dra. Aline Aparecida Rezende Rodrigues and Dr. José Miguel Ortega for his valuable assistance for the in silico analyses.

References

- BENHAR, I. Biotechnological applications of phage and cell display. **Biotechnology advances**, v. 19, n. 1, p. 1-33, 2001.
- IRVING, M. B.; PAN, O.; SCOTT, J. K. Random-peptide libraries and antigen-fragment libraries for epitope mapping and the development of vaccines and diagnostics. **Current Opinion in Chemical Biology**, v. 5, n. 3, p. 314-324, 2001.
- MANDAVA, S. et al. RELIC-a bioinformatics server for combinatorial peptide analysis and identification of protein-ligand interaction sites. **Proteomics**, v. 4, n. 5, p. 1439-1460, 2004.
- SOUZA, G. R. L. **Construção de uma biblioteca de anticorpos (scFv) apresentada em fagos para seleção, análise e caracterização de antígenos do carrapato bovino (Boophilus microplus)**. Uberlândia, 2007. 83f. Tese (Doutorado) - Instituto de Genética e Bioquímica, Universidade Federal de Uberlândia.
- WANG, Y. S. et al. Development of a multi-mimotope peptide as a vaccine immunogen for infectious bursal disease virus. **Vaccine**, v. 25, n. 22, p. 4447-4455, 2007.