doi 10.18699/bgrs2024-12.2-10

Mathematical modeling of the promising composition of the next generation safe epitope vaccine against porcine salmonellosis

Lavrukhin M., Kichemazova N., Oglodina D., Larionova O., Zaitsev S., Feodorova V.*

Saratov State University of Genetics, Biotechnology and Engineering named after N.I. Vavilov, Saratov, Russia

** feodorovav@mail.ru*

Key words: *Salmonella*; immunity; *in silico* analysis; epitope; vaccine

Motivation and Aim: Currently, the methods of artificial neural network are considered to be one of the promising approaches for development of a new generation effective safe vaccines. The main advantages of these methods are a better prediction rate, speed, data efficiency and high rank of selection of MHC molecules which are capable to elicit successfully the protective immune response in target hosts. In fact, a rational design of either single- or multi-peptide vaccines predicted *in silico* could be critical to induce enhanced immunity against actual infection diseases in specific mammalian species. Moreover, these vaccines are absolutely safe because possess no live microbial material that highlights their great potency for both animal health and welfare, and overall, for global agricultural economy. *Salmonella enterica* subsp. *enterica* serovar Choleraesuis (*S. choleraesuis*), the causative agent of salmonellosis, is known as a swine-adapted foodborne pathogen associated to invasive infections in both animals and humans. Nowadays, *Salmonella* infection in humans is considered as a global health burden. Contaminated animal-derived foodstuffs, especially pork products are the main source for human salmonellosis. Overall, *Salmonella* infection in swine and other livestock could lead to marked economic losses in national pork industry and global agricultural sector as well.

Several live whole cell (LWC) and killed whole cell (KWC) vaccines for immunization against porcine salmonellosis are now available in Russian Federation and worldwide. However, all of them need to improve their level of immunogenicity and safety as well. The main goal of this study was to predict *in silico* potential ligands for CD8+ T-cell epitopes of MHC class I in pig model using freely available machine leaning algorithms.

Methods and Algorithms: The IEDB public database "T Cell Prediction – Class I" was used as the main tool for the current study. This tool is based on the selective simultaneous use of three different neural networks, such as: (i) conventional sparse encoding [1]; (ii) BLOSUM encoding [2]; and (iii) hidden Markov model encoding [3]. The NetMHCpan 4.1 EL was applied as the main predictive method by IEDB (recommended epitope prediction – 2023.09). Both nucleotide and amino acid sequences for 14 virulence genes (*ipfB*, *sopB*, *ompA*, *pipB*, *pipB2*, *sifA*, *sinH*, *sopD*, *sopD2*, s*opE2*, *sseB*, *spvB*, *spvC* and *spvR*) of the *S. choleraesuis* strain TC-177 were obtained from NCBI GenBank database (https://www.ncbi.nlm.nih.gov/nucleotide/). The relevant epitope data sets for MHC class I were predicted for all 44 available pig alleles (access date 02/24/2024). The panel of potentially protective epitopes was selected based on

3 main parameters, namely antigenicity, allergenicity and immunogenicity using the DDG (https://www.ddg-pharmfac.net/ddg/index.html) and IEDB database software (https://www.iedb.org), respectively, with default settings.

Results: Initially, in range 60,000–400,000 epitopes were predicted for each protein. Further, two consecutive cut-offs resulted for individual protein from 14 to 88 epitopes (0.007–0.044 %, i. e. lesser then 1 %) which were combined into 3 groups based on their desired immunological characteristics, such as: (i) the epitopes with marked antigenicity, immunogenicity and non-allergenic activity (n = in the range from 1 to 20), which may be involved in a protective immune response, and can become the basis for a new generation vaccine; (ii) the epitopes with marked antigenicity, immunogenicity and allergenic activity (n = in the range 1–13); and (iii) the largest group, including up to 90 % of epitopes (n = in the range 11–64) which demonstrated no required characteristics resulting no protective aberrant immune response in target animals. Importantly, the majority of potentially protective epitopes were associated with only 5 proteins encoded by such genes as *pipB*, *sinH*, *sopD*, *spvB* and *spvC*. In contrast, no epitopes with relevant protective activity were predicted for the proteins expressed by the remaining 9 genes (*ipfB*, *sopB*, *ompA*, *pipB2*, *sifA*, *sopD2*, *sopE2*, *sseB* and *spvR*).

Conclusion: The data obtained could be perspective for development of multi-epitope peptide vaccine against porcine salmonellosis.

Funding: The study is supported by Russian Science Foundation (No. 22-16-00165).

References

1. Baldi P., Brunak S. Bioiformatics: The Machine Learning Approach. A Bradford Book. Publisher: The MIT Press Cambridge, Massachusetts London, England, 2001
2. Henikoff S., Henikoff J.G. Amino acid substitution matrices from protein blocks. *Proc Natl Acad Sci USA*. 1992;89(22):10915-10919. doi 10.1073/pnas.89.22.10915
3. Nielsen M., Lundegaard C., Worning P., Hvid C.S., Lamberth K., Buus S., Brunak S., Lund O. Improved prediction of MHC class I and class II epitopes using a novel Gibbs sampling approach. *Bioinformatics.* 2004;20(9):1388-1397. doi 10.1093/bioinformatics/bth100