

Министерство образования и науки Российской Федерации

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г. ЧЕРНЫШЕВСКОГО»

Кафедра медицинской физики

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

**МОЛЕКУЛЯРНОЕ МОДЕЛИРОВАНИЕ ОСНОВНЫХ
ИММУНОЛОГИЧЕСКИХ ХАРАКТЕРИСТИК БЕЛКА Pfra
КАК ОСНОВЫ ПЕПТИДНЫХ ВАКЦИН ПРОТИВ ЛИСТЕРИОЗА**
студента 2 курса 2221 группы

направления 03.04.02 «Физика», профиль «Медицинская физика»

код и наименование направления


института физики

наименование факультета, института

Лаврухин Максим Сергеевич


фамилия, имя, отчество

Научный руководитель
Профессор кафедры, д.м.н.,
профессор
должность, уч. степень, уч. звание


подпись, дата
3.06.23

В.А. Федорова
инициалы, фамилия

Зав. кафедрой:
д.ф.-м.н., профессор
должность, уч. степень, уч. звание


подпись, дата
3.06.23

АН.В. Скрипаль
инициалы, фамилия

Саратов 2023 г.

Введение

Методы классической вакцинологии по созданию вакцин существуют уже более 2 веков. Они хорошо отработаны и отлажены и главное доступны. Если инфекционный агент обладает низким показателем изменчивости нуклеотидной, и/или аминокислотной последовательности, с помощью классического метода можно с высокой долей вероятности обнаружить иммунологический агент способный индуцировать сильную и продолжительную протективную иммунную реакцию, что позволит создать на основе этого агента вакцинный препарат. Но эти методы имеют и свои ограничения: разработка препаратов подобным методом занимает много времени (15-20 лет), требует большой штат разнообразных специалистов; методы остаются несистемными и вакцины которые получены этим методами имеют относительную эффективность из-за эмпирического подхода при их создании; из-за несовершенства этих методов, все вакцинные препараты имеют поствакцинальные осложнения.

Одними из наиболее перспективных направлений молекулярной биологии и вакцинологии в частности являются методы создания вакцин нового поколения. Основной особенностью таких методов будет предиктивная обработка нуклеотидных и аминокислотных последовательностей инфекционных агентов, для последующего предиктивного анализа иммунологических характеристик эпитопных последовательностей. Достоинствами подобных вакцинных препаратов перед классическими методами будет: высокая скорость разработки вакцинных препаратов (1-5 лет), высокая эффективность вакцин, так как подобные методы будут учитывать видовые особенности организмов восприимчивых к инфекции, против которой и разрабатывается вакцина, и вместе с этим снижением, или полным отсутствием поствакцинальных осложнений.

Основными трудностями при разработке методов создания вакцин нового поколения, является отсутствие полного понимания всей картины иммунологических реакций происходящих в макроорганизме при

взаимодействии с инфекционным агентом, а так же необходимость мониторинга большого количества факторов, что порождает огромный поток данных требующий интерпретации. На основе существующих проблем нами предложен метод молекулярного моделирования основных иммунологических характеристик, который позволит прогнозировать эпитопные последовательности способные к индукции иммунной реакции восприимчивого животного, с минимизацией пост вакцинальных осложнений.

Все вышеуказанные данные подтверждают **актуальность** молекулярное моделирование основных иммунологических характеристик белка *Pfra* как основы пептидных вакцин против листериоза.

В связи с актуальностью поисков наиболее подходящего моделирования иммунологических характеристик **целью работы** является метода молекулярного моделирования иммунологических характеристик эпитопов на модели белка *Pfra* и кодирующего его гена *pfra* в штамме *Lm* АУФ для КРС в качестве целевых сельскохозяйственных животных.

Для достижения поставленной цели были поставлены следующие **задачи**:

1. Проанализировать состояние современной литературы, посвящённой исследованиям в области создания вакцин нового поколения, путём программного прогнозирования иммунологических характеристик белка *Pfra*;

2. Сравнить нуклеотидную последовательность гена *pfra* штамма *Lm* АУФ с гомологичной последовательностью гена *pfra* референтного полновирулентного штамма *Lm* EGD-e.

3. Программно спрогнозировать иммунологические характеристики для гена *pfra* для обоих штаммов.

4. Обработать полученные данные, выделить группы эпитопов, а так же провести поиск в этих группах эпитопов потенциально способных к провокации протективного иммунного ответа у КРС.

Новизна работы: проведено молекулярное моделирование иммунологических характеристик белка Pfra, как основы пептидных вакцин против листериоза.

1. Результаты

Из базы данных GENBANK были получены нуклеотидные и aa последовательности *pfra*_EGD и *pfra*_АУФ. Последовательности представлены в табл 2.

Таблица 2. Нуклеотидные и aa последовательности гена <i>pfra</i> штаммов Lm АУФ и EGD-е		
	АУФ	EGD-е
Нуклеотидная последовательность	atgtatgatcgattacagcgggtggaagaccggtatgatgaactaaatg agttattaagtgatccagatgtagtaccgatacccaaacgactacgcga ccttccaaagaacaatccggcaltaccgcgacagtcgaaacgtaccg cgaatacaaaaatgtaaacgacaataacgaaacaaaagaactcct aggagaaaaactagacgatgaaatgcgcgaatggccaaagaagaat tcgggaaactcaaaaagaanaaggcagacitagaagaacgactaaaa ctattactcgtgccaaaagatcctaatgatgacaaaaacgtiatttagaa atcccgagcagctgggtggagatgaagcagcttatttgcggcgatt tattccgatgtacagtaaatatgcggaatcagcggctggaaagtaga aatcatggacgcgaaccgactggattggcgggtacaagaatatt gccgatgatgaacggnaatgacgccttccgcaatgaaataggaac ggggcgcaccggttcaacgltacagaacagaaatcagggtgctc aatccatacatcaacagcgaacagtgccattttgccgaagcgggaaga agtggaaatcaggttgcatacaagaatccgtacagatatacgcac tccactggtccggtggacaagaatgcaataaacgatgtagcgtgac gttaacgcataatccgaccggaattgctgttcgatgcaagatgagcgt ctcagttgaaaaacaagataaagcagataaagatttgcgcgcacggt ttatgataaattgagcgtgagcctccgaggaatcagatccaaccgt aagtcagctgttggaacgggtgaccgctccgaacgtatccgactata actatccgaaaaccgcgtaccgaccaccggtatcggttaaacatfca aaagctagatacaatcatggaagcgaactcgatgaaatcattgacgc gctcatcctagaagatcaacaagtaaacggagcatttaaatgacgca aattag	atgtatgatcgattacagcgggtggaagaccggtatgatgaactaaatg aactactaagtatccagatgtagtaccgatacccaaacgactacgcga ccttccaaagaacaatccggcaltaccgcgacagtcgaaacgtaccg gaatacaaaaatgtaaacgacaataacgaaacaaaagaactcct ggagaaaaactagacgatgaaatgcgcgaatggccaaagaagaat cgcggaaactcaaaaagaanaaacagactgggaagaacgactaaaa tattactcgttccaaaagatcctaatgatgacaaaaacgtiatttagaa atcccgagcagctgggtggagacgaagcagcttatttgcggcgatt tccgtatgtacagcaaatatgcggaaatcagcggctggaaagtagaa attatggacgcaaatccaactggtattggcgggtacaagaatatttgc gatgatgaacggaaatgatgcttttccgaatgaaatgaaaacggg gccacccgtgtacaacgttccctgaaacagagtcagcggctcgaatl catacatcaacagcagctgtgcaattttccgaaagcgggaagaagtg gaaatcaggttgcatacaagaatccgtacagatatacgcacatcc ctggtccgggtggacaagaatgcaataaacgatgtagcgtgacgtt aaccatattccgaccggaattgctgttcgatgcaagatgagcgttcc agttgaaaaacaagataaagcagataaagatttgcgcgcacgtgtta tgataaattgagcgtgaggctgtgtaggaatcagatccaaccgtaag tcagctgttgaacgggtgaccgctccgaacgtatccgacttataact atccgaaaaccgcgtaacgcgaccgctatcggttaaacatfca agctagatacaatcatggaagcgaactcgatgaaatcattgacgc catcctagaagatcaacaagtaaacggagcatttaaatgacgcaaat tag
Длина	1077 нуклеотидов	
Аа последовательность	MYDRLQAVEDRYDELNELLSDPDVVSDPKR LRDLSKEQSGITATVETRYREYKKNVNEQINETK ELLGEKLDDEMREMAKEEFAELQKEKADLE ERLKLLLVPKDPNDKKNVLEIRGAAGGDEAA LFAGDLFRMYSKYAESRGWKVEIMDANPT GIGGYKEIJAMMNGNDAFSRMKYENGAHRV QRVPETESGGRHITSTATVAILPEAEVEIELH DKDIRTDTFASTGAGGQSVNTTMSAVRLTHI PTGIVVSMQDERSQLKNKDKAMKVLRARVY DKFEREFARYDANRKS AVGTGDRSERIRTY NYPQNRVTDHRIGLTIQKLDQIMEGKLEIHD ALILEDQTSKLEIHLNDAN	MYDRLQAVEDRYDELNELLSDPDVVSDPKR LRDLSKEQSGITATVETRYREYKKNVNEQINETK ELLGEKLDDEMREMAKEEFAELQKEKTDLFE RLKLLLVPKDPNDKKNVLEIRGAAGGDEAA LFAGDLFRMYSKYAESRGWKVEIMDANPTGI GGYKFIAMMNGNDAFSRMKYENGAHRVQR VPETESGGRHITSTATVAILPEAEVEIELHDK DIRTDTFASTGAGGQSVNTTMSAVRLTHIPTG IVVSMQDERSQLKNKDKAMKVLRARVYDKF EREAREFYDANRKS AVGTGDRSERIRTYNYP QNRVTDHRIGLTIQKLDQIMEGKLEIHDALIL EDQTSKLEIHLNDAN
Длина	358 аминокислот	

Было проведено сравнение нуклеотидных последовательностей *pfra*_АУФ и *pfra*_EGD-е с помощью инструмента Multalin. Результат сравнения представлен на Рис.2.

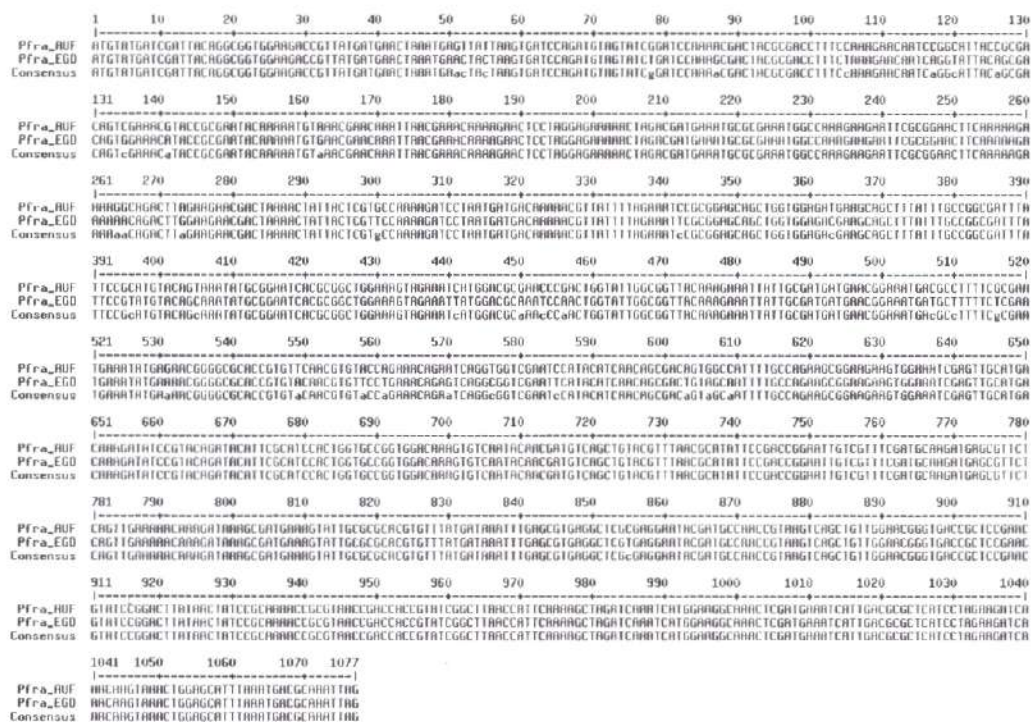


Рис 2. Сравнение нуклеотидных последовательностей генов *pfra*_АУФ и *pfra*_EGD-е.

На рис. 2. видно, что большая часть нуклеотидной последовательности этих 2 генов идентична. Из 1077 нуклеотидов 1038 оказались полностью идентичны, что составило 96,5% (1039/1077), но так же мы обнаружили различия в 38 нуклеотидов - 3,5% (38/1077)

Нуклеотидные замены (далее гапы) имели чёткую локализацию в первой половине гена (538 нуклеотидов) 26 гапов, что составляет 68,42% от общего числа гапов в этих последовательностях (26/38), тогда как вторая половина гена (539 нуклеотидов) демонстрировала наличие только 12 гапов - 31,57% (12/38). Это может говорить о том, что вторая часть гена *pfra* более консервативна - в ней накапливается больше мутаций, чем в первой части, которая явно демонстрировала определенную вариабельность.

Далее была произведена загрузка *aa* последовательностей Pfra_AУФ и Pfra_EGD на сервер IEDB. Для ускорения загрузки на сервер IEDB аллельных профилей коров, вместе с *aa* последовательностями белка Pfra для разных штаммов, были составлены аллельные файлы в виде столбца, содержащего 2 элемента в каждой строчке: название и номер аллели, далее через запятую указывается длина эпитопов предсказываемая для конкретной аллели. Из базы данных IEDB были взяты наборы аллельных профилей для модельного животного. Полученный аллельный файл был загружен на сервер IEDB вместе с *aa* последовательностью белка Pfra для каждого штамма Lm.

Стоит отметить, что, из-за серверных ограничений, загрузка и предсказание эпитопов для всех аллелей для каждого животного оказалась невозможной. Серверный буфер позволяет загружать информацию не больше, чем для 10 аллельных профилей с указанием длин для каждой аллели индивидуально, поэтому аллельные файлы были поделены на фрагменты, вмещающие число аллелей равное n , где n было ≤ 10 .

С помощью инструмента MHC-I Binding Predictions нами были предсказаны Т клеточные эпитопы МНС - I типа для таргетных животных. Данные были выведены в виде дата фреймов (Рис. 3.).

sequence_number	peptide	length	start	end	allele	peptide_index	core	icore	score	percentile
1	KLDEIIDAL	9	334	342	BoLA-5*003:01	6275	KLDEIIDAL	KLDEIIDAL	0.68537	0.01
2	KLDEIIDAL	9	334	342	BoLA-5*003:01	6974	KLDEIIDAL	KLDEIIDAL	0.68537	0.01
1	FEREAREEY	9	280	288	BoLA-1*020:01	16982	FEREAREEY	FEREAREEY	0.34737	0.01
2	FEREAREEY	9	280	288	BoLA-1*020:01	19418	FEREAREEY	FEREAREEY	0.34737	0.01
1	FEREAREEY	9	280	288	BoLA-1*020:01	16982	FEREAREEY	FEREAREEY	0.34737	0.01
2	FEREAREEY	9	280	288	BoLA-1*020:01	19418	FEREAREEY	FEREAREEY	0.34737	0.01
1	EEYDANRKS	10	286	295	BoLA-1*019:01	10030	EEYDARKSA	EEYDANRKS	-0.17702	0.01
2	EEYDANRKS	10	286	295	BoLA-1*019:01	12466	EEYDARKSA	EEYDANRKS	-0.17702	0.01
1	EEYDANRKS	10	286	295	BoLA-1*019:01	10030	EEYDARKSA	EEYDANRKS	-0.17702	0.01
2	EEYDANRKS	10	286	295	BoLA-1*019:01	12466	EEYDARKSA	EEYDANRKS	-0.17702	0.01
1	RYDELNELL	9	11	19	BoLA-3*065:01	60561	RYDELNELL	RYDELNELL	0.16908	0.01
2	RYDELNELL	9	11	19	BoLA-3*065:01	62997	RYDELNELL	RYDELNELL	0.16908	0.01
1	TMSAVRLTH	9	238	246	BoLA-1*009:01	2324	TMSAVRLTH	TMSAVRLTH	0.08791	0.01
2	TMSAVRLTH	9	238	246	BoLA-1*009:01	4760	TMSAVRLTH	TMSAVRLTH	0.08791	0.01
1	TMSAVRLTH	9	238	246	BoLA-1*009:01	2324	TMSAVRLTH	TMSAVRLTH	0.08791	0.01
2	TMSAVRLTH	9	238	246	BoLA-1*009:01	4760	TMSAVRLTH	TMSAVRLTH	0.08791	0.01
1	KLDEIIDAL	9	334	342	BoLA-5*072:01	10469	KLDEIIDAL	KLDEIIDAL	0.447572	0.02
2	KLDEIIDAL	9	334	342	BoLA-5*072:01	11168	KLDEIIDAL	KLDEIIDAL	0.447572	0.02
1	KLDEIIDAL	9	334	342	BoLA-1*067:01	65756	KLDEIIDAL	KLDEIIDAL	0.40441	0.02
2	KLDEIIDAL	9	334	342	BoLA-1*067:01	68192	KLDEIIDAL	KLDEIIDAL	0.40441	0.02
1	KLDEIIDAL	9	334	342	BoLA-1*067:01	65756	KLDEIIDAL	KLDEIIDAL	0.40441	0.02
2	KLDEIIDAL	9	334	342	BoLA-1*067:01	68192	KLDEIIDAL	KLDEIIDAL	0.40441	0.02
1	FEREAREEY	9	280	288	BoLA-1*042:01	51086	FEREAREEY	FEREAREEY	0.34737	0.02
2	FEREAREEY	9	280	288	BoLA-1*042:01	53522	FEREAREEY	FEREAREEY	0.34737	0.02
1	FEREAREEY	9	280	288	BoLA-1*042:01	51086	FEREAREEY	FEREAREEY	0.34737	0.02
2	FEREAREEY	9	280	288	BoLA-1*042:01	53522	FEREAREEY	FEREAREEY	0.34737	0.02
1	KLDEIIDAL	9	334	342	BoLA-5*064:01	9071	KLDEIIDAL	KLDEIIDAL	0.258212	0.02
2	KLDEIIDAL	9	334	342	BoLA-5*064:01	9770	KLDEIIDAL	KLDEIIDAL	0.258212	0.02

Рис. 3. Пример дата фрейма получаемого с использованием инструмента MHC-I Binding Predictions, предсказаны эпитопы KPC.

Каждый получившийся аллельный файл индивидуально загружали на сервер, вместе с *aa* последовательностью сначала Pfra_AУФ, затем Pfra_EGD-е, соответственно.

После загрузки полученные файлы с датафреймами содержащие эпитопные *aa* последовательности, были объединены используя скрипт M1, написанного нами на языке программирования Python.

Суть скрипта M1 состояла в соединении нескольких текстовых файлов в один. Для этого использовались модули phyton - fileinput и glob, отвечающие за последовательную загрузку файлов и их сшивание, соответственно. Пример кода для скрипта приставлен на Рис. 5.


```

import fileinput
import glob

file_list = glob.glob("*.txt")

with open('all.txt', 'w') as file:
    input_lines = fileinput.input(file_list)
    file.writelines(input_lines)

```

Рис. 4. Код скрипта M1, отвечающего за сшивание скачанных файлов с БД IEDB.

Затем, после форматирования получившегося сшитого файла, запускался скрипт M2, отвечающий за сортировку получившегося датафрейма с сортировкой по возрастанию. Метрикой для этого стал параметр «percentile_rank», отвечающий за иммунологические свойства эпитопа. С повышением значения этого параметра увеличивается способность эпитопов индуцировать иммунологическую реакцию иммунной системы реципиента.

После сортировки и использования библиотеки pandas был создан новый датафрейм, содержащий только эпитопы с параметрами percentile_rank ниже или равным единице. Это порог отсечения был выбран из-за рекомендаций IEDB, как оптимальный для этой работы. Это связано с тем, что 99,88% от предсказанных эпитопов оказались выше порога отсечения, причем увеличение их значений происходило логарифмически.

Далее полученный датафрейм был переформатирован так, чтобы убрать повторяющиеся эпитопы, так как один и тот же эпитоп может встречаться как в одном аллельном профиле, так и в нескольких. Для этой работы применялись эпитопы, наиболее часто встречающиеся в наибольшем кол-во аллельных профилей, доступных в БД IEDB, так как эпитопы, что имеют охват большего кол-ва аллельных профилей позволит предсказать более универсальные эпитопы. Для подсчёта повторений эпитопов была

использована функция «value_counts» из библиотеки pandas. Пример кода скрипта M2 указан на Рис. 5.

```
import pandas as pd
from tkinter import Tk

from tkinter.filedialog import askopenfilename
Tk().withdraw()
filename = askopenfilename()

df = pd.read_csv(filename, sep='\t')

g=(len(open(filename).readlines()))

df_rank = df[df['percentile_rank'] <= 1]

with open('cut_off.txt', "w") as f:
    print(df_rank.to_string(), file=f)

aga=df_rank[['peptide', 'percentile_rank']]
aga_a=aga.set_index('peptide')

with open('peptide+rank.txt', "w") as f:
    print(aga_a.to_string(), file=f)

with open('allele-unicum.txt', "w") as f:
    print((df_rank['peptide'].value_counts(ascending=False)).to_string(), file=f)
```

Рис. 5. Код скрипта M2, отвечающий за сортировку, 1 отсечение и подсчёт повторений эпитопных последовательностей.

В результате обработки первичных «сырых» данных были получены три панели эпитопов, кол-во которых на разных этапах отображены на Рис. 6.



Рис. 6. Сравнение кол-ва предсказанных эпитопов на разных этапах работы алгоритма.

Как видно на Рис. 6, всего было предсказано 295878 эпитопов для каждого гена Lm. После этого нами был применён первый порог отсечения, что сократило количество эпитопов до 2498 - 0,84% от изначального количества предсказанных эпитопов (2498/295878). Следующим этапом было применение второго порога отсечения, вследствие чего количество анализируемых эпитопов снизилось до 40 aa последовательностей, что составило 1,6% от набора эпитопов первого порога отсечения (40/2498) и всего 0,013% от первоначальной эпитопной выборки (40/295878).

Далее полученные датафреймы с эпитопами и их кол-вом повторений в каждом файле для каждого штамма Lm были перенесены в Excel для дальнейшего сравнительного анализа. С этой целью для каждого из 40 отобранных эпитопов был предсказан индекс иммуногенности, затем этот эпитопный набор был отсортирован по убыванию, именно по этому параметру.

Проведено сравнение предсказанных эпитопов для штамма АУФ и EGD-е. Результаты сравнения представлены на Рис. 7.

Peptide	Length	score	Иммуноген/н антиген/е иммуногенный score	antigen/n on-antigen	allergenicity	Toxicity	Штамм
TSTATVAL	9	0.23941	Immunogen	100% Probable ANTIGEN	PROBABLE NON-ALLERGEN	Non-Toxin	AYF+EG D-e
KEKIDLEERL	10	0.18883	Immunogen	66% Probable ANTIGEN	PROBABLE NON-ALLERGEN	Non-Toxin	EGD-e
KIDDEMRM	9	0.0639	Immunogen	66% Probable ANTIGEN	PROBABLE NON-ALLERGEN	Non-Toxin	AYF+EG D-e
KEQSGHATV	10	0.05077	Immunogen	66% Probable ANTIGEN	PROBABLE NON-ALLERGEN	Non-Toxin	AYF+EG D-e
EQNETKEL	9	0.04733	Immunogen	100% Probable ANTIGEN	PROBABLE NON-ALLERGEN	Non-Toxin	AYF+EG D-e
SKYAESBQW	9	0.04342	Immunogen	66% Probable ANTIGEN	PROBABLE NON-ALLERGEN	Non-Toxin	AYF+EG D-e
KLDEIDAL	9	0.40441	Immunogen	100% Probable ANTIGEN	PROBABLE ALLERGEN	Non-Toxin	AYF+EG D-e
AALPEALEV	9	0.26257	Immunogen	100% Probable ANTIGEN	PROBABLE ALLERGEN	Non-Toxin	AYF+EG D-e
IHIPTGIVV	9	0.23818	Immunogen	66% Probable ANTIGEN	PROBABLE ALLERGEN	Non-Toxin	AYF+EG D-e
EQSGHATV	9	0.20224	Immunogen	66% Probable ANTIGEN	PROBABLE ALLERGEN	Non-Toxin	AYF+EG D-e
KEKADLEERL	10	0.18916	Immunogen	100% Probable ANTIGEN	PROBABLE ALLERGEN	Non-Toxin	AYF+EG D-e
ADLEERLKL	9	0.10801	Immunogen	66% Probable ANTIGEN	PROBABLE ALLERGEN	Non-Toxin	AYF+EG D-e
TDLEERLKL	9	0.10801	Immunogen	66% Probable ANTIGEN	PROBABLE ALLERGEN	Non-Toxin	EGD-e
FEREAREEY	9	0.34737	Immunogen	66% Probable NON-ANTIGEN	PROBABLE ALLERGEN	Toxin	AYF+EG D-e
ATVETYREY	9	0.25065	Immunogen	66% Probable NON-ANTIGEN	PROBABLE ALLERGEN	Non-Toxin	AYF+EG D-e
SERERTYNY	9	0.23076	Immunogen	100% Probable NON-ANTIGEN	PROBABLE ALLERGEN	Non-Toxin	AYF+EG D-e
RLIHIPTGI	9	0.21687	Immunogen	66% Probable NON-ANTIGEN	PROBABLE ALLERGEN	Non-Toxin	AYF+EG D-e
AVEDRYDEL	9	0.17806	Immunogen	100% Probable NON-ANTIGEN	PROBABLE ALLERGEN	Non-Toxin	AYF+EG D-e
RYDEINELL	9	0.16906	Immunogen	66% Probable NON-ANTIGEN	PROBABLE ALLERGEN	Non-Toxin	AYF+EG D-e

Рис.7. Сравнение эпитопов, предсказанных для штаммов Lm EGD-е и АУФ с помощью ПО Microsoft Excel. Зелёным цветом выделены потенциально иммунопротективные эпитопы, жёлтым

- эпитопы, способные вызывать как иммунный ответ, так и аллергическую реакцию (абберантный иммунный ответ), оранжевым - все остальные характеристики. Синим отмечены эпитопы, способные индуцировать протективный иммунный ответ, которые встречаются только в штамме EGD-e, красным - способные индуцировать абберантный иммунный ответ, найденные только в штамме АУФ.

Как видно на Рис. 7. сформированные группы оказались не одинаковы по размеру. К первой группе относится 6 эпитопов, что составило 15% от числа отобранных (7/40) и 0,002% от общего числа предсказанных эпитопов (7/295878). Важно отметить, что эпитоп «KEKTDLEERL» имеет необходимые параметры, чтобы считать его потенциально иммунопротективным, и был предсказан только для полновирулентного штамма EGD-e, что может говорить об утрате этого эпитопа в слабовирулентном штамме АУФ.

Ко второй группе относилось 7 эпитопов - 18,75% (6/32), или 17,5% (7/40) и 0,0023% (7/295878), соответственно. Важно отметить, что к этой группе были отнесены 2 эпитопа, которые являются специфичными для штамма АУФ и не встречаются в референтном полновирулентном штамма.

Большее количество эпитопов было включено в третью группу - 10,8% (27/40).

Таким образом, нами было выявлено 3 группы эпитопов:

1.Потенциально иммунопротективные (далее просто иммунопротективные) – эпитопы, которые потенциально обладают иммуногенными, антигенными, не аллергенными свойствами и не являются токсичными для организма - эпитопы с 1 по 6 включительно;

2.Потенциально абберантные (далее абберантные) - эпитопы, обладающие иммуногенностью и антигенностью, но при этом потенциальные аллергены. Это эпитопы способны вызывать выраженную аллергическую реакцию вместо протективного иммунного ответа. Эпитопы с 7 по 13 включительно;

3. Все остальные комбинации - в эту группу мы объединили все эпитопы, которые частично обладали искомыми свойствами, но отсутствие одного или нескольких из них не позволяет им вызвать выраженный проективный иммунный ответ организма. Эпитопы с №>14;

Заключение

Подводя итоги данной дипломной работы, можно сказать, что поставленная цель, а именно – *молекулярное моделирование иммунологических характеристик белка Pfra, как основы пептидных вакцин против листериоза.* – была достигнута, в результате решения следующих задач:

1. Было проанализировано состояние современной литературы, посвящённой исследованиям в области создания вакцин нового поколения;
2. Было проведено сравнение нуклеотидной последовательности гена *pfra* штамма *Lm АУФ* с гомологичной последовательностью гена *pfra* референтного полновирулентного штамма *Lm EGD-e*;
3. Было проведено программное прогнозирование иммунологических характеристик гена *pfra* для обоих штаммов;
4. Была проведена обработка полученных данных, выделены группы эпитопов, а также поиск в этих группах эпитопов, потенциально способных к провокации протективного иммунного ответа у КРС.

В ходе выполнения магистерской работы установлено, что разработанный метод молекулярное моделирование иммунологических характеристик белка Pfra, как основы пептидных вакцин против листериоза. Полученные результаты согласуются с теоретическими расчётами и позволяют ускорить процесс поиска вакцинных таргетов.

Список использованной литературы

1. Кичемазова Н.В., Хижнякова М.А., Ляпина А.М., Колосова А.А., Федорова В.А. ВАКЦИНОПРОФИЛАКТИКА ЛИСТЕРИОЗА СЕЛЬСКОХОЗЯЙСТВЕННЫХ ЖИВОТНЫХ УДК 619:616.9: 615.371:579.62: 636.03
2. Шамшева О.В., Учайкин В.Ф., Медуницын Н.В. Клиническая вакцинология. Москва : ГЭОТАР-Медиа, 2016. 576 с.
3. Зверев В.В. Вакцины и вакцинация : национальное руководство / под ред. В.В. Зверева, Р.М. Хаитова. Москва : ГЭОТАР-Медиа, 2014. 640 с.
4. Мякинькова Л.Л., Букач О.В., Логунова А.В. Современные проблемы, вызовы и перспективные направления в области вакцинологии // Инноватика и экспертиза: научные труды. 2015. Т. 14, № 1. С. 96-109.
5. Doolittle R.F., Hunkapiller M.W., Hood L.E. et al. Simian sarcoma virus oncogene, v-sis, is derived from the gene (or genes) encoding a platelet-derived growth factor//Science. 1983. V. 221. P. 275-277.
6. Гельфанд М.С., Миронов А.А. Вычислительная биология на рубеже десятилетий // Молекулярная биология. 1999. Т. 33. С. 969-984.
7. Vitreschak A.A., Rodionov DA., Mironov A.A., Gelfand M.S. Regulation of riboflavin biosynthesis and transport genes in bacteria by transcriptional and transia-tional attenuation // Nucleic Acids Research. 2002. Т. 30. P. 3141-3151.
8. Yershov G., Barsky V., Belgovski A. et al. DNA analysis and diagnostics on oligonucleotide microchips // Proc. Natl. Acad. Sci. USA. 1996. V. 93. P. 4913-4918.
9. Nuritdinov R.N., Artamonova I.I., Mironov AA., Gelfand M.S. Low conservation of alternative splicing patterns in the human and mouse genomes // Hum. Mol. Genet. 2003. V. 12. P. 1313-1320.

10.Thorisson G.A., Stein L.D. The SNP Consortium web-site: past, present and future // *Nucleic Acids Res.* 2003. V. 31. P. 124-127.

11.Базовые понятия аллергологии (часть 2) // cyberleninka URL: <https://cyberleninka.ru/article/n/bazovye-ponyatiya-allergologii-chast-2> (дата обращения: 04.05.2023).

12. Чеснокова С.М., Чугай Н.В. Биологические методы оценки качества объектов окружающей среды Учебное пособие В двух частях Часть 2. Методы биотестирования

13. Giesker K, Hensel M. Types of Bacterial Vaccines. *Science Direct.* 2014. <https://www.sciencedirect.com/topics/medicine-and-dentistry/bacterial-vaccine>

14. David E, Bloom SB. Antimicrobial resistance and the role of vaccines. *PNAS.* 2018.

15.Oxford Vaccine Group. Types of vaccines. 2020. <https://vk.ovg.ox.ac.uk/vk/types-of-vaccine>

16.WHO. Vaccine quality. Retrieved from Immunization standards. 2020. https://www.who.int/immunization_standards/vaccine_quality/en/

17.OIE. OIE terrestrial manual. OIE. 2018.

18.Creative Biolabs. Subunit Vaccines. 2020. <https://www.creativebiolabs.com/vaccine/subunit-vaccine-design.htm>

19.Lidder P, Sonnino A. Biotechnologies for the Management of Genetic Resources for Food and Agriculture. *Advances in Fenetics.* 2016.

20.National Institute of Allergy and Infectious Diseases. Vaccine Types. 2019. <https://www.niaid.nih.gov/research/vaccine-types>

21. Ministry Health System. Types of vaccines. 2020. <https://www.health.mil/Military-Health-Topics/Health->

Readiness/ImmunizationHealthcare/Clinical-Consultation-Services/Types-of-Vaccines

22.The immunization advisory center. Types of vaccines. 2020. <https://www.immune.org.nz/vaccines/vaccine-development/types-vaccines>

23.WHO. WHO finds India's vaccine regulatory authority compliant with international standards. 2020. https://www.who.int/medicines/regulation/indiaauthority_reg_compliant-int-standards/en/

24.Endert P. Antigen Processing (2nd ed.). New York: Springer Science+Business Media, LLC. 2019.

25.Fox CB.. Vaccine adjuvants. New York: Humana Press. 2017.

26.Ferran MC, Skuse GR. Recombinant Virus Vaccines. New York: Springer Science+Business Media LLC. 2017.

27.CHOP. Making Vaccines: Process of Vaccine Development. 2020. <https://www.chop.edu/centers-programs/vaccine-education-center/making-vaccines/process-vaccine-development>

28. UK Vaccine Network. (n.d.). Welcome to the Vaccine Development Process Map. 2020. <http://vaccinedevelopment.org.uk/>

29.Irini A Doytchinova and Darren R Flower. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinformatics. 2007 8:4.

30.Irini A Doytchinova and Darren R Flower. Identifying candidate subunit vaccines using an alignment-independent method based on principal amino acid properties. Vaccine. 2007 25:856-866.

31.Irini A Doytchinova and Darren R Flower. Bioinformatic Approach for Identifying Parasite and Fungal Candidate Subunit Vaccines. *Open Vaccines Journal*, 2008 1:22-26.

32.Sudheer Gupta, Pallavi Kapoor, Kumardeep Chaudhary, Ankur Gautam, Rahul Kumar, Gajendra PS Raghava In Silico Approach for Predicting Toxicity of Peptides and Proteins // *PloS one*. - 2013. - №9. - C. e73957.

33.Nielsen M., Andreatta M. NetMHCpan-3.0; improved prediction of binding to MHC class I molecules integrating information from multiple receptor and peptide length datasets. *Genome Med.* 2016; 8:33.

34.Karosiene E., Rasmussen M., Blicher T., Lund O., Buus S., Nielsen M. NetMHCIIpan-3.0, a common pan-specific MHC class II prediction method including all three human MHC class II isotypes, HLA-DR, HLA-DP and HLA-DQ. *Immunogenetics.* 2013; 65:711–724.

35.O'Donnell T.J., Rubinsteyn A., Bonsack M., Riemer A.B., Laserson U., Hammerbacher J. MHCflurry: open-source Class I MHC binding affinity prediction. *Cell Syst.* 2018; 7:129–132.

36.Kim Y., Sidney J., Pinilla C., Sette A., Peters B.. Derivation of an amino acid similarity matrix for peptide: MHC binding and its application as a Bayesian prior. *BMC Bioinformatics.* 2009; 10:394.

37.Nielsen M., Andreatta M. NetMHCpan-3.0; improved prediction of binding to MHC class I molecules integrating information from multiple receptor and peptide length datasets. *Genome Med.* 2016; 8:33.

38.Karosiene E., Rasmussen M., Blicher T., Lund O., Buus S., Nielsen M. NetMHCIIpan-3.0, a common pan-specific MHC class II prediction method including all three human MHC class II isotypes, HLA-DR, HLA-DP and HLA-DQ. *Immunogenetics.* 2013; 65:711–724.

39.O'Donnell T.J., Rubinsteyn A., Bonsack M., Riemer A.B., Laserson U., Hammerbacher J.. MHCflurry: open-source Class I MHC binding affinity prediction. *Cell Syst.* 2018; 7:129–132.

40.Kim Y., Sidney J., Pinilla C., Sette A., Peters B.. Derivation of an amino acid similarity matrix for peptide: MHC binding and its application as a Bayesian prior. *BMC Bioinformatics.* 2009; 10:394.

41.Nielsen M., Connelley T., Ternette N.. Improved prediction of bovine leucocyte antigens (BoLA) presented ligands by use of mass-spectrometry-determined ligand and in vitro binding data. *J. Proteome Res.* 2018; 17:559–567.

42.Alvarez B., Reynisson B., Barra C., Buus S., Ternette N., Connelley T., Andreatta M., Nielsen M.. NNAlign_MA; MHC peptidome deconvolution for accurate mhc binding motif characterization and improved t-cell epitope predictions. *Mol. Cell Proteomics.* 2019; 18:2459–2477.

43.Bassani-Sternberg M., Chong C., Guillaume P., Solleder M., Pak H., Gannon P.O., Kandalafi L.E., Coukos G., Gfeller D.. Deciphering HLA-I motifs across HLA peptidomes improves neo-antigen predictions and identifies allosteric regulating HLA specificity. *PLoS Comput. Biol.* 2017; 13:e1005725.

44.Jurtz V., Paul S., Andreatta M., Marcatili P., Peters B., Nielsen M.. NetMHCpan-4.0: improved peptide-MHC class I interaction predictions integrating eluted ligand and peptide binding affinity data. *J. Immunol.* 2017; 199:3360–3368.

45.Barra C., Alvarez B., Paul S., Sette A., Peters B., Andreatta M., Buus S., Nielsen M.. Footprints of antigen processing boost MHC class II natural ligand predictions. *Genome Med.* 2018; 10:84.

46.Языковые разногласия: как Oracle и Google судятся из-за Java.[Электронный ресурс]: <https://www.pvsm.ru/java/197831> (дата обращения: 17.04.2023);

47. Google Fuchsia. [Электронный ресурс]:
https://ru.wikipedia.org/wiki/Google_Fuchsia (дата обращения: 17.04.2023);

48. Новую ОС Fuchsia от Google скомпилировали и изучили. [Электронный ресурс]: <https://habr.com/ru/post/396891/> (дата обращения: 17.04.2023);

49. Расцвет фуксии. Что мы знаем о Fuchsia OS от Google? [Электронный ресурс]: <https://tproger.ru/translations/google-fuchsia-osreview/> (дата обращения: 17.04.2023).

