COMPARATIVE ANALYSIS TO DETERMINE THE EFFICACY OF PEPTIDE VACCINE AGAINST DIFFERENT STRAINS OF SARS-COV-2

A DISSERTATION

SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE AWARD OF THE DEGREE

OF

MASTER OF SCIENCE

IN

BIOTECHNOLOGY

Submitted by:

VANSHIKA DUREJA

2K20/MSCBIO/34

Under the supervision of

Dr. Asmita Das

DEPARTMENT OF BIOTECHNOLOGY

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

MAY, 2022

DELHI TECHNOLOGIAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

CANDIDATE'S DECLARATION

I, Vanshika Dureja, Roll No.- 2k20/MSCBIO/34 student of M.sc BIOTECHNOLOGY, hereby declare that the project Dissertation titled "name of project" which is submitted by me to the Department of Biotechnology, Delhi Technological University, Delhi in partial fulfilment of the requirement for the award of the degree of Master of Science, is original and not copied from any source without proper citation. This work has not previously formed the basis for the award of any Degree, Diploma Associateship, Fellowship or other similar title or recognition.

Place: Delhi VANSHIKA DUREJA

Date:

DEPARTMENT OF BIOTECHNOLOGY

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

CERTIFICATE

I hereby certify that the Project Dissertation Titled "**Comparative analysis to determine the efficacy of peptide vaccine against different strains of SARS-CoV-2**" which is submitted by **Vanshika Dureja**, Roll No. – 2k20/MSCBIO/34 (Department of Biotechnology), Delhi Technological University, Delhi in partial fulfilment of the requirement for the award of the degree of Master of Science, is a record of the project work carried out by the students under my supervision. To the best of my knowledge this work has not been submitted in part or full for any Degree or Diploma to this University or elsewhere.

Place: Delhi (Dr. ASMITA Das)

Date: **SUPERVISOR**

ACKNOWLEDGEMENT

Firstly, I would like to express my sincere gratitude to my supervisor Dr. Asmita Das for her continuous support during my dissertation, for her patience, motivation and scientific discussions without which this research could not have been completed. I would like to thank Mr. Sunil, for his guidance during entire project, despite the length sof this project he never hesitated to answer my questions and help me throughout my dissertation. I would also like to thank all my lab members Mr Mohd Tauheed Rayeen, Mr Shayon Mahalonobis and Ms Priyanka for their constant support and good time. I would also like to give special thanks to my parents and family as a whole for their continuous support and understanding when undertaking my research and writing my project. Your prayer for me was what sustained me so far. Finally, I would like to thank God, for letting me through all the difficulties. I have experienced your guidance day by day. You are the one who let me finish my degree.

Vanshika Dureja

ABSTRACT

SARS CoV-2 is a global threat these days and has been declared as a pandemic by WHO. With the passage of time from 2019 to 2021 different variants of the SARS CoV2 have been reported due to mutations in the S glycoprotein of the SARS CoV-2. Vaccines are being developed by different institutes and organization to prevent/ inhibit the further destruction by the coronavirus. Out of the vaccines synthesised using various approaches the protein subunit vaccines are easy to manufacture and do not trigger other immunologic reactions. Since the s glycoprotein of the COVID-19 is the most immunologic component amongst the complete virion and is an essential site for vaccine design. Due to this reason in our study we focused on the NVX-CoV-2373 which a peptide based sub unit vaccines with sequence similar to the S Glycoprotein of the virus except the dual mutations and it alters the binding of surface glycoprotein of virus to hACE2 of host and produces both humoral and cellular immunity in host. But this vaccine has not been authorized by all the countries, so here, we have used an in silico approach to determine the efficacy of the vaccine against few strains of SARS-

CoV-2 by identifying the similarities between the S glycoprotein sequence of the vaccine and the variants by performing Blast which is followed by the linear B cell epitope prediction using Bepipred 2.0 and the results obtained were analysed for similarity. It was found that the maximum number of epitopes of the Lambda variants were identical to the NVX-Cov-2373 vaccine.

CONTENTS

Title page **01**

Candidates declaration **02**

Certificate **03**

Acknowledgement **04**

Abstract **05**

Contents **07**

List of Figures **10**

List of Tables **12**

List of Symbols/abbreviations **14**

Chapter 1 INTRODUCTION **15**

1.1 About the novel SARS-CoV-2 **15**

1.2 Symptoms of SARS-CoV-2 **16**

1.3 Structure and Features of SARS- CoV-2 **17**

1.4 Spike protein and its role in pathogenesis **18**

1.5 Variants of SARS-CoV-2 **19**

1.5.1 Variants of Concern **20**

1.5.2 Variants of Interest **22**

1.6 Need of vaccine and significance of Spike glycoprotein in vaccine development **23**

1.7 Immune responses to SARS CoV-2 **24**

1.7.1 Innate immune responses by SARS-CoV-2. **25**

1.7.2 Humoral immune responses by SARS-CoV-2. **26**

1.8 Limitations associated with production of vaccine **26**

1.9 Epitope prediction and its significance for vaccine efficacy and design. **27**

1.9.1 Epitope prediction. **28**

Chapter 2 LITERATURE REVIEW 30

2.1 Vaccines for COVID-19 **30**

2.1.1 The Whole Virus Approach **31**

2.1.2 Nucleic Acid Approach **31**

2.1.3 Subunit Approach **32**

2.2 NVX-CoV2373 Vaccine **32**

2.3 B cell epitope prediction **35**

2.3.1 Prediction of Linear B cell Epitopes. **37**

2.3.2 Prediction of Conformational B cell Epitopes **38**

Chapter 3 MATERIALS AND METHODOLGY 41

3.1 Tools used **41**

3.1.1 BLAST **41**

3.1.2 Bepipred **42**

3.2 Methodolgy **45**

3.2.1 Retrieving data from NCBI **45**

3.2.2 Retrieving sequence data of NVX-CoV2373 **46**

3.2.3 Performing BLAST **46**

3.2.4 Linear B cell epitope Prediction **47**

Chapter 4 RESULTS AND DISCUSSION 48

4.1 Results **48**

4.1.1 Sequences of SARS- CoV-2 variants **48**

4.1.2 BLAST Results **49**

4.1.3 Bepipred Results **51**

4.2 Discussion **53**

Appendices **54**

References **69**

List of Publication **77**

LIST OF FIGURES

FIGURE 1.1 Timeline of SARS-CoV-2

FIGURE1.2 This Diagram depicts the features of COVID 19.

FIGURE 1.3 Diagrammatic representation of strains of variants of SARS CoV-2.

FIGURE 1.4 Present in the Alpha variant along with its properties.

FIGURE 1.5 Diagrammatic representation of the types of immunity and the cells involved.

FIGURE 2.1 Showing various techniques for designing COVID-19 vaccine.

FIGURE 2.2 Steps involved in synthesis of Novavax vaccine.

FIGURE 2.3 Diagrammatic representation of different B-cell epitopes and their prediction methods.

Figure 3.1 Schematic representation of Various types of BLAST.

Figure 3.2- Schematic representation of the overall methodology involved.

Figure 4.1- Blast Tree view of the provided sequences of Vaccine and SARS CoV2 strain

Figure 4.2 - Blastp results for alignment of multiple sequence of SARS CoV-2 against the vaccine sequence.

LIST OF TABLES

Table 2.1 List of few vaccines available for SARS CoV-2 along with their type, company and target antigen

Table 2.2 List of tools for linear epitope prediction

Table 2.3 List of tools for conformational epitope prediction.

Table 3.1 Shows the relationship between selected thresholds and the sensitivity/specificity of the prediction method, calculated on basis of the epitope/non-epitope predictions.

Table 3.2 It shows the relationship between selected thresholds and the sensitivity/specificity of the prediction methods.

Table 3.3 Differences between Bepipred 1.0 and 2.0

Table 4.1 List of variant of SARS CoV-2 under study along with their accession number obtained.

Table 4.2 Bepipred result of vaccine sequence and different covid strains, here the predicted linear epitopes of vaccine are mentioned and the cells are filled with a distinct

color, the epitope number of the variants that match the epitopes of vaccine are also colored with the similar color.

LIST OF SYMBOLS/ABBREVIATIONS

1) ML- Machine learning
2) MHC- Major Histocompatibility Complex
3) BCR- B cell Receptors
4) CEP- Conformational Epitope Prediction
5) hACE2- Human Angiotensin Converting Enzyme 2
6) WHO – World Health Organization
7) CDC- Centre for Disease Control
8) SARS- Severe Acute Respiratory Syndrome
9) VOC- Variant of Concern
10) VOI- Variant of Interest
11) MMR- Measles, Mumps and Rubella
12) IFN- Interferon
13) BLAST- Basic Local Alignment Search Tool
14) ESTs- Expressed Sequence Tags
15) NCBI- National Centre for Biotechnology Information
16) IEDB- The Immune Epitope Database

CHAPTER 1

INTRODUCTION

1.1 About the Novel SARS-CoV-2

Coronaviruses (CoVs) are a class of viruses that majorly give rise to the infections in the human respiratory tract. Many genera of CoVs infect humans but SARS CoV-2 is the most pathogenic amongst all and it leads to critical diseases of the respiratory tract and can cause death of the infected individuals as well. In December 2019, a novel form of SARS CoV-2 was seen in Wuhan, China which was named COVID-19/Coronavirus Disease 2019 by WHO [1]. The COVID-19 pandemic converted into a very critical worldwide health emergency and one of the highest threats for the population [2] along with being a global health emergency various fields of life such as economic, social, environmental and cultural are affected by this virus. On 11th of March 2020, WHO (World Health Organization) revealed it as a worldwide rife. When compared with SARS CoV and MERS-CoV, SARS-CoV-2 is extremely infectious and transmissible having a reproductive number around 2.2 (i.e. COVID-19 infection in one of the individual can further lead to an average of 2.2 new infections) [3]. Furthermore, it has created many problems in its containment process because of its tendency to get transmitted from patients without symptoms [4]. The infection by SARS CoV-2 reached

around 1.5 million and resulted in 96,000 deaths by April 2020. Three waves of COVID—19 are reported till date. First in December, 2019 the second one in July 2020 and third one in December 2021.

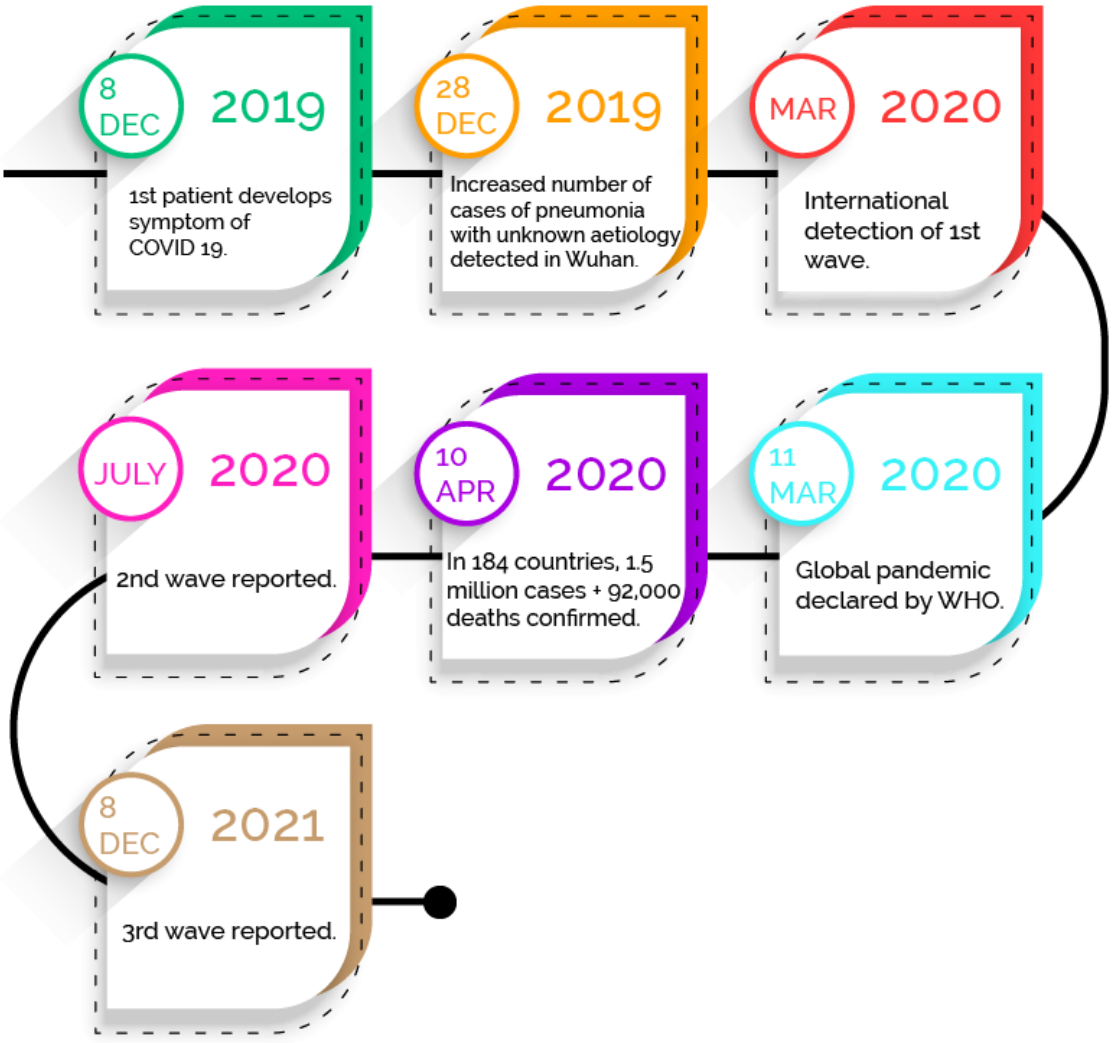


Figure1.1 Timeline of SARS-CoV 2

1.2 **Symptoms of SARS CoV-2**

The medical indications of COVID-19 can differ from symptom less and mild flu-like symptoms to critical respiratory distress syndrome and demise [1]. The rise in number of infected individuals/cases, the symptoms observed for coronavirus were variable, few of the common symptoms are continued fatigue and tiredness, join pain, insomnia, increased heart rate, dizziness, difficulty in breathing and others include permanent

pulmonary damage, myalgia and neurological degeneration. The severity of symptoms and the increased death rate are commonly reported due to hypertension, diabetes, obesity in patients suffering from COVID-19 infection [5,6], since three waves of the COVID-19 pandemic were seen the symptoms of all the three waves were found to be variable, symptoms observed in first wave of coronavirus were fever, cough, chest and muscle pain, confusion, sore throat, anosmia, dyspnoea, ageusia and headache and those for the second wave were cold, pneumonia, fever and dyspnea. However, the symptoms observed in the third wave were flu like with neurological effect.

1.3 Structure and features of SARS-CoV-2

It is a virus with envelope and is spherical in shape with 80-160nm diameter with a large fragmented +ve sense ssRNA with a genome size of 29.9kb, which is an intermediate value between the size of SARS CoV (29.9kb) and middle east respiratory syndrome coronavirus(30.1kb).[7,8,9]. Two third of the genome of coronavirus is spanned by 2 ORFs: i.e. ORF1a and ORF1b present at the 5’ terminus [9]. Both these ORFs are further translated to form non-structural proteins that are 16 in number (Nsps) which are responsible for the growth and multiplication of virus. The 3’ terminal of this virus code for 4 structural proteins such as (S), (E), (M), and (N) proteins. The S protein mentioned above signifies the name of virus i.e. is “corona” as it gives a crown like structure and is the most immunogenic part of the virus and thus is essential targets used by antibodies to neutralise and prevent viral infection [10], while M protein is abundant protein and tells about the shape of the virion because it is the largest protein and is essential for release of newly formed viral particles form host after infection. Viral infection and its replication are is maintained by E protein and all these three proteins that is S, M, and E are accountable for forming the viral coat, on the other hand N protein maintains the integrity of viral genome (RNA) within the envelope by binding to it (it creates a capsid to load the genomic RNA). When virus undergoes self-assembly (after infection) M protein coordinate with all the other proteins to form a complete virion (11,12,13). The 3’ terminus also encodes for some accessory proteins that are specific for a genus and helps virus from the effect of immune system of host or by increasing its virulence.

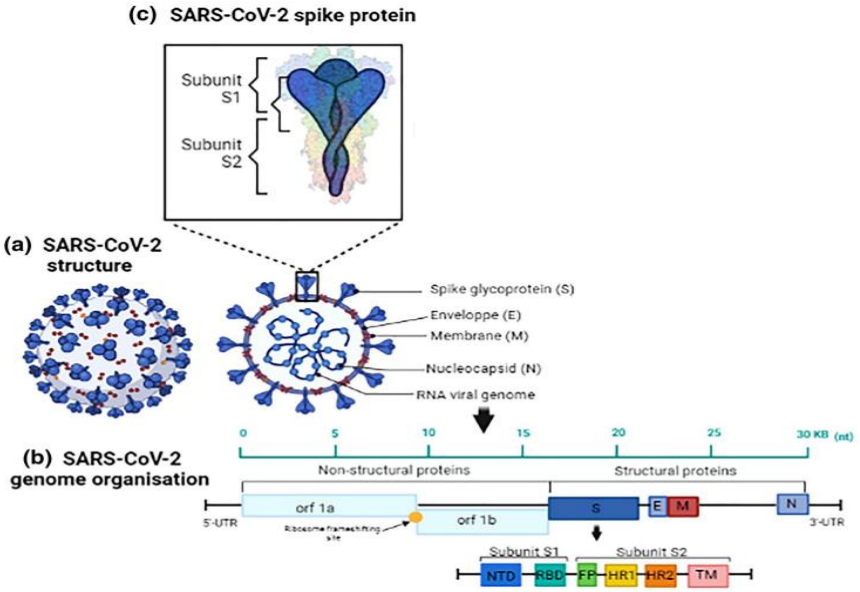


Figure 1.2 This Diagram depicts the features of COVID 19.

Ikbel Hadj Hassine (2021), Covid-19 vaccines and variants of concern: A review, Wiley

1.4 Spike protein and its role in pathogenesis

Spike protein or S protein in SARS CoV-2 is around 1273 amino acid long, and has an RRAR furin cleavage site that forms the two subunits S1(cap) and S2 (stalk) at the endosome. The S1 subunit has receptor binding domains (RBD) for binding to receptors as well as N terminal domains. While S2 subunit (non RBD containing) has an internal membrane fusion protein, HR1 and HR2 (heptapeptide repeat sequence), transmembrane domain and a membrane proximal external region. Filament of S2 subunit helps in the viral genome entry in host cell by fusing with the host membrane [13,14,15]. Assembly of three S1/S2 in a non-covalent manner leads to an active and functional S glycoprotein. The trimeric spike protein is an essential surface glycoprotein that associates with (hACE2) i.e. human angiotensin converting enzyme for viral priming, insertion and release (by undergoes various rearrangements in its structure thereby exposing the hydrophobic fusion peptides). Therefore, one of the most relevant step in development of infection and can be targeted for development of vaccine. [16,17]

1.5 Variants of SARSCoV-2

Due to the presence of error prone viral polymerase and speedy replication viruses such as SARS CoV-2 gain mutations [18]. Viruses like SARS CoV-2 acclimatize to immune response in various species and tissues leading to mutations majority of these mutations do not alter the infectivity or phenotype of virus [19]. Around the globe, a large variety of strains of SARS CoV-2 exist, but majority of them does not alter the mechanism of action of virus. Most of the anomalies in different mutants are located in S gene of SARS CoV-2. Variants are produced from Natural selection and these new variants show better survival due to an efficient replication/ penetration and transmissibility between humans. Furthermore, an important role in development of novel variants is played by the frequency of random events leading to development of more transmissible forms rather than the more pathogenic ones.[20]. Variants are classified in two categories based on the transmissibility of variant and how critical the variant is ? and these categories are 1) variants of concern (VOC) and 2) variants of interest (VOI).

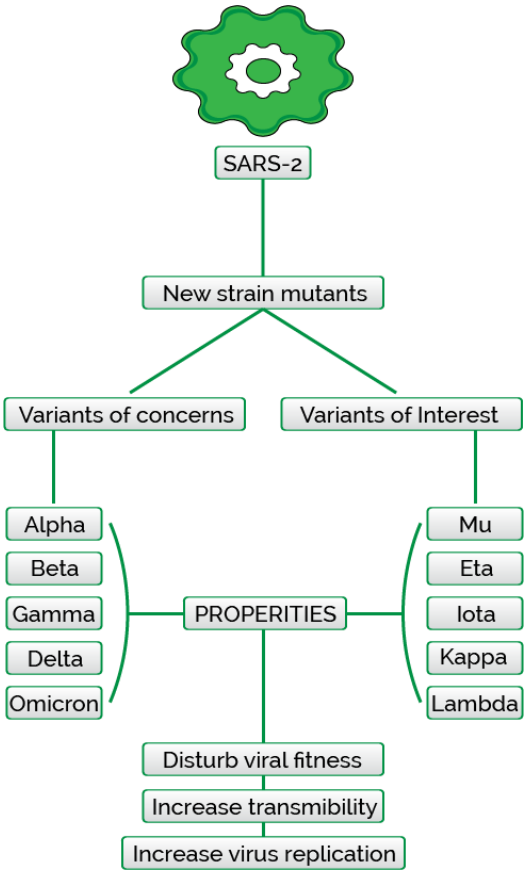


Figure 1.3 Diagrammatic representation of strains of variants of SARS CoV-2

1.5.1 Variants of Concern

They are those variants of virus who have mutations in the spike proteins (most immunogenic protein of the virus) and therefore are of major concerns as these spike protein sequences are used in vaccine design and any alteration in these sequences will alter the efficacy of vaccine. Few VOC are Alpha, beta, gamma, omicron and delta variants.

1) Alpha Variant (B.1.1.7)

U.K. was the initial place where this variant was first discovered in around September 2020 and later it spread in Europe first and then in U.S where it remained dominant. On comparison with the native COVID strain of Wuhan, the alpha variant is around 50% (46.6%) extra infectious, dangerous and cause a higher number of deaths and hospitalizations.

Alpha variant has around 10 mutations in the S protein which have caused all the structural alterations. In this strains N501Y mutation is reported in the RBD domain of S glycoprotein, D614G and P681H mutations were seen in the spike proteins that lead to increased binding to the ACE2 receptors, a comparatively higher deposition of the variants could be seen in pharynx and nasal cavities leading to higher transmission rate were observed due to these above mentioned mutations [21-24].

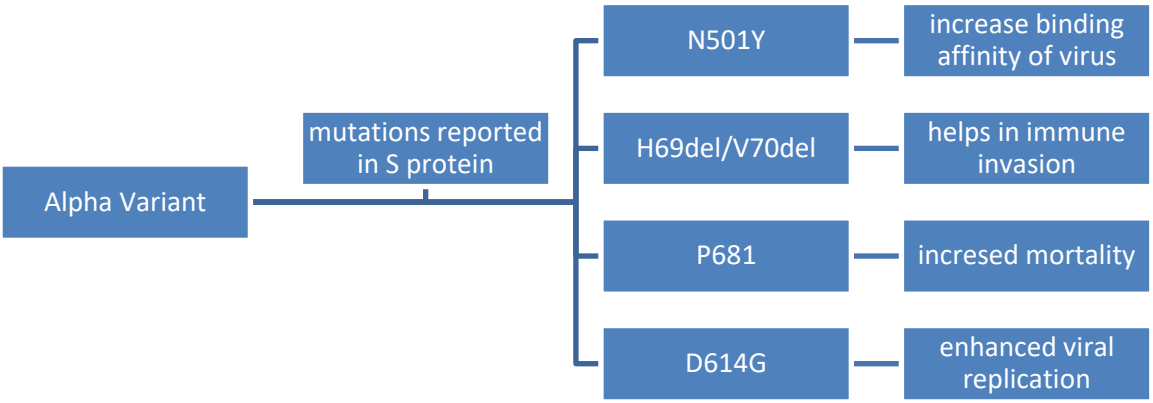


Figure 1.4 Mutations present in the Alpha variant along with its properties.

2) **Beta variant (B.1.351, B.1.351.2, B.1.351.3)**

In South Africa Beta variant was initially discovered and is found to be 50% more transmissible tan the original strain. And has better tendency to escape from the neutralizing antibody. Due to the above mentioned reasons it is challenging to treat the infections caused by Beta variants as there are more chances of reinfection in contrast to the native strain. Three subgroups with 12 possible mutations are available for the beta variant, few of these mutations are K417N, E484K, N501Y and D614G which are reported in spike protein sequences, these mutations has led to an escalated transmissibility and severity of infection (decreased binding to mRNA vaccine induced and monoclonal antibody induced antibodies) [25,26,27].

3) Gamma variant (P.1, P.1.1, P.1.2)

In November 2020, it was initially found in Japan and then later in U.S. in 2021. As in beta variant, spike protein alterations of gamma variant also helps escaping neutralizing antibodies, so the patients with prior COVI-19 infections or the immunized individuals can also experience this infection. In this strain, four mutations are seen (K417T, E484K, N501Y, D614G) out of which former three are present RBD this variant is included in VOC. This variant has in total 17 mutations majority of which that is 11 mutations are in the S protein. These mutations are related to improved transmissibility and higher capacity to hide from the immune system by escaping neutralizing antibodies [25,24,28].

4) Delta variant (B.1.617.2, AY.1, AY.2, AY.3)

In India, it was initially discovered in October 2020. This variant acquired more importance once it was identified in the U.S. They also have mutations reported in the S protein as observed in the case of both β and γ variants that prevent antibody neutralization. multiple mutations that are present are L452R, E484Q, D614G and P681R that makes this variant distinct from others is that it becomes very efficient while adhering to the host cells, therefore, lowers the vaccine efficacy and is much more contagious [36,37].

5) **Omicron (B.1.1.529)**

On November 9, 2021 omicron was seen in South Africa. This variant comprise of 3 lineages and one sub lineage. Omicron has a very high rate of mutations > 60 substitutions, deletions and insertions, including 39 on the spike region of the virus out of 39, 15 mutations are there in RBD, which influences the spread and its response to different treatment measure and vaccinations [29.30].

1.5.2 Variants of interest (VOIs)

They are known to be even more dangerous, due to their liberation from the immune system, and counteracts to treatment methods that are available, or are challenging determine. Based on the rates of infections VOIs are "of interest" to few countries. Some of the variants categorised as VOIs are mu, eta, kappa and lambda variants.

1) Mu variant (B.1.621 and B.1.621.1)

In Colombia in January 2021, this variant was sequenced initially and is been reported in more than 39 countries that include U.S. Mu is not categorised as a VOI by CDC but is categorised by WHO due to sudden outbreaks in Europe. Few research had indicated that alterations in the S protein of SARS CoV-2 help mu variant to escape from immune system.

2) Eta variant (B.1.525)

It is designated as VOI by WHO in March 2021 but was initially discovered in the U.K. and Nigeria in December 2020. It also has mutations in the spike protein which makes it more resistant to therapies, as reported by the CDC.

3) Iota variant (B.1.526)

In November 2020, it was initially detected in New York and was listed as a VOI by WHO in March 2021. With the onset of pandemic, it has leads to 3% of all COVID-19 infections. Studies have suggested that vaccines and treatments are not effective for this variant making it uncertain.

4) Kappa variant (B.1.617.1)

Initially witnessed in India in October 2020 throughout a similar schedule as that the delta variant and was declared as a VOI in April 2021 by WHO. It makes the antibody therapy ineffective due to the presence of spike protein mutations.

5) Lambda variant (B.1.1.1.37, alias C.37)

In December 2020, it was initially announced in Peru. In June 2021 it accounted for 10% of COVID-19 cases and was declared as a VOI by WHO.

1.6 Need of vaccine and the significance of Spike glycoprotein in vaccine development

In past two decades, our knowledge about the pathogenic CoVs has been rising but no preventive effective vaccines are discovered. There is urgent need of an effective vaccine to prevent the quick spread and mortality caused by COVID-19. The S protein is of great significance for virus-cell receptor binding and virus-cell membrane fusion, depicting that it is a potential target for CoV vaccine design [31]. It has been reported in few studies that antibodies produced against S protein are immune-dominant and have a long term effect [32, 33]. Many vaccines against S proteins have produced protective responses. Other than S protein other structural proteins are also evaluated for vaccine development but those targeting N protein do not induce neutralizing antibodies because N protein are not exposed on the surface of Coronaviruses [26]. But N protein has advantage that it is more conserved amongst the CoV species that the S protein that undergoes limitless mutations thereby making it an essential target for T cell inducing vaccine development that can be universal [26]. Vaccines based on M protein generates a large amount of antibody responses in vaccinated animals [34] but does not provide protective immunity.

1.7 Immune responses to SARS CoV-2

The immune system especially in humans is extremely complicated and perform its action at various levels. A distinct immune system can be seen in different individuals and immune challenges affect them differently. It is usually classified into two types: innate and adaptive immune system. Innate immunity includes non-specific defence mechanisms that either acts right away or hours after the arrival of microorganisms into the body therefore, it is quick, old and conserved and acts as a first line defence [35][36]. All multicellular organisms show few forms of innate immunity, whereas adaptive immunity is highly specific and exists only in vertebrates. The adaptive immune system, in reality is capable of individually identifying and destroying invading pathogens. Furthermore, the adaptive immune system will remember to fight the pathogens, obtain pathogen-specific durable protective memory, and allow further strikes every time a pathogen is encountered [37]. This immunity is orchestrated by a system of highly specialised cells that transmit information via chemical interactions that occur on the surface of cells along with complicated assembly of molecules such as cytokines and chemokines that are known for intercellular communication.[38] However, both of these two immune mechanisms work jointly, and the initiation of adaptive immunity depends on the previous stimulation of the non-specific immune response [37]. Acquired immunity is expressed by B and T lymphocytes that recognise the antigen (as they bind to specific receptors on these lymphocytes) rather than recognising the entire pathogen. The recognition patterns of B and T cells vary highly, i.e., antigens exposed to solvents are recognised by B lymphocytes via binding to the B cell receptors that contain the membrane bound Ig that on activation leads to differentiation followed by secretion of soluble immunoglobulins or antibodies. It also helps in regulating the humoral adaptive responses by performing various functions such as neutralising toxins and pathogens and labelling these pathogens for destruction or elimination [37].

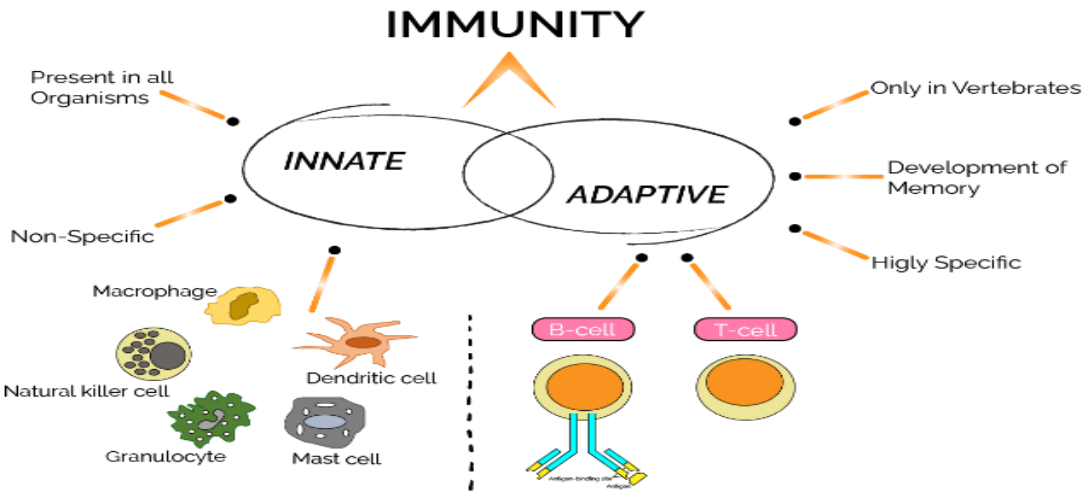


Figure. 1.5 Diagrammatic representation of the types of immunity and the cells involved.

The immune system's hierarchical and combinatorial properties contribute to its complexity. As a result, a massive quantity of details and facts is developed about the immune system. Immunological research is required in order to deal with this complication. For a long time, immunologists have used high throughput experimental approaches, which have resulted in a massive amount of functional, clinical, and epidemiological data. As a result, new computational techniques are being developed. It is necessary to develop methods for storing and analysing these data. This results in Immunoinformatics is a field that deals with the study of the immune system.[39] Therefore, it is important to understand the importance of immune reactions to infections by SARS-CoV-2 for the development of vaccine. As discussed before, SARS-CoV-2 constitutes of four structural proteins i.e S,N, M and E. S glycoprotein as mentioned protrudes from the viral surface and has 2 subunits one for viral binding and other viral and host cell membrane fusion. Antibodies that are produced against the 'S' protein inhibit the entry of virus inside host, on the other hand antibodies against other proteins activate the killing of virus [40].

1.7.1 Innate immune response by SARS Cov-2

The knowledge about the innate immune response shown by COVID-19 is very limited but the general RNA viruses are known to initiate the immune response by recognition by TLR and PPRs. This recognition in return activates a series of events leading to

production of Cytokines such as TNF and ILs and these molecules jointly regulate the adaptive immune response. If this response is present early and effectively, IFN-1 can inhibit the infection caused by COVID-19 but this virus can inhibit the activation and signalling of IFN.

1.7.2 Humoral response by SARS CoV-2

IgM, IgG and IgA antibodies are found to be developed in some COVID-19 patients against the 'N' and 'S' proteins, amongst both the proteins neutralizing antibodies against S protein are important for effective immune response against SARS-CoV-2 as the RBD and few other regions of S glycoprotein behave as epitopes leading to Neutralizing antibody production.

1.8 Limitations associated with production of Vaccines

1) The exact innate immune response shown against SARS CoV-2 is very limited which restricts the development of vaccine due to in complete knowledge.
2) The significance of T- cell immunity in COVID-19 or related infection is not well understood.
3) **Route of Administration**- the route by which a vaccine is administered is essential for effectiveness of vaccine. In a study on the mice model given a intramuscular vaccine dose, it was found that adenovirus vectored vaccine produces a systemic humoral and cell mediated response, it does not prevented or provided immunity while one intranasal dose of same vaccine induced increased levels of Nabs, T cell responses and promotes systemic and mucosal IgA and almost prevented SARS CoV-2 completely. But the current trials are not considering the nasal route of administration [41].
4) **The type and number of antigen**- for production of Nabs targeting of S glycoprotein is essential but targeting other antigens for vaccine development can also be advantageous for example the 'N' protein is more conserved than the 'S' protein and vaccine against the 'N' protein can therefore, be able to provide an enduring immunity against SARS-CoV-2. Another limitation while vaccine development is the no. of antigens to be introduced in the candidates as although SARS-CoV-2 is a single stranded virus but is more prone to mutations and even smaller mutations can lead to formation of different lineages [42].

5) **Duration of immunity and need for boosting-** as in common flu infection the antibodies formed in host are not active for longer than few weeks to months therefore common cold and flu are catched easily. Similar is the case for SARS-CoV-2 with the upcoming mutations it is difficult for body to develop a long term immunity against this disease by vaccine administration. Moreover, the development of antibodies in the infected individuals is also delayed (i.e not easily prominent to be detected by assays) which adds another complexity to the process of vaccine design [43].
6) **Endpoint for assessing efficacy**- it a big doubt that what should be the ideal 'primary endpoint' that is for ideal vaccine to be effective what do we expect it to do, to provide complete protection from infection or just attenuation of severity of disease which is further challenged by the limited knowledge about the transmission rate and incidence dynamics of the virus.
7) **Dosing issues**- it is important to identify the correct dose of vaccine to prevent it from failing the phase 3 trials by managing the correct efficacy and safety. Further, the elderly people have sub optimized immunity than the younger generation therefore, they might require higher doses.
8) **Fears and concern about COVID-19 vaccines**- myths and wrong information is spread by media addressing public about the side effects thereby creating resistance to vaccine uptake [44].

1.9 Epitope prediction and its significance for vaccine efficacy and design

Immunoinformatics basically deals with the in-silico modelling and data analysis of immunology. The field of immunoinformatics has different areas of computational research, it majorly relies on the research and development of algorithms for recognising probable B-cell and T-cell epitopes. The epitopes are useful in chemical synthesis of peptide vaccines. To make the target B-cell epitope immunogenic, it coalesces with the T-cell epitope. So, this method of recognising the potential antigen epitopes is called "reverse vaccinology." B cell epitopes is the segment of the antigen that attaches to the antibody, these epitopes constitute the solvent-exposed region which are being identified by the B cells. Since the majority of antigens are proteins, they are therefore subjected to epitope prediction. Since vaccination is very important for a healthy population worldwide but for new emerging diseases such as SARSCoV-2

which have a complex lifecycle and variable antigens designing a vaccine is difficult. With the introduction to immunoinformatics and computation sciences in biology new methods are proposed for vaccine design and its efficacy detection as a normal experimental method for a vaccine development and release takes about 10-11 years and waiting for long in this pandemic is not possible. With the availability of genome data in the databases it analyses can be done for predicting epitopes of the pathogens making the process less labour intensive and quick [45-47].By looking for a protein sequence in pathogen of interest by computational methods we can discover the epitopes that can be targeted for designing vaccines. Being pathogens specific and unique, these predicted epitopes represent ready candidates in vaccine construction.

1.9.1 EPITOPE PREDICTION

The antigen's potential to attach to a distinct complementary antibody is termed as the antigenicity of an antigen while immunogenicity is production of immune response. The purpose of predicting epitope is to device a molecule that could substitute the position of an antigen in synthesising or detecting antibodies. Once this molecule is devised it can be produced or cloned into a vector.[35] Generally, in predicting epitopes, the primary goal emphasizes the interaction of MHC molecule to the peptide antigen. The experimental technique is found to be complex and time-consuming. Consequently, many computer-based methods are continually being developed and are being used to determine the epitope. The series of methods comprises of array-controlled methods, QSAR analysis, recognition of structural binding motifs, threading of proteins, homology modelling and ML tools. In former times, computer technology could only identify the characteristics of the sequence. However, new improved algorithms and techniques are being developed continuously to boost prediction accuracy. For practical reasons, the identification of epitopes in antigens is very interesting, including, understanding the root cause of the disease, immune monitoring, the development of diagnostic assays, and the epitope-based vaccine construction. The B cell epitopes has various application, including analysing the antigen-antibody complex3D structure, using a peptide library to detect antibody binding, or using a specific diagnostic test [48][49]. While, polymer MHC, lymphatic proliferation, or ELISPOT assays are used for the experimental T cell epitope analysis. Traditional epitope prediction is entirely based on experimental technology, which is expensive and time-consuming. As a result,

epitope forecasting methods are devised to enhances identification of epitope and also reducing the load of associated tests.

CHAPTER 2

LITERATURE REVIEW

2.1 VACCINES for COVID-19:

Three essential vaccine technology are exploited for the development of successful and harmless vaccine against COVID-19. These platforms are listed below.

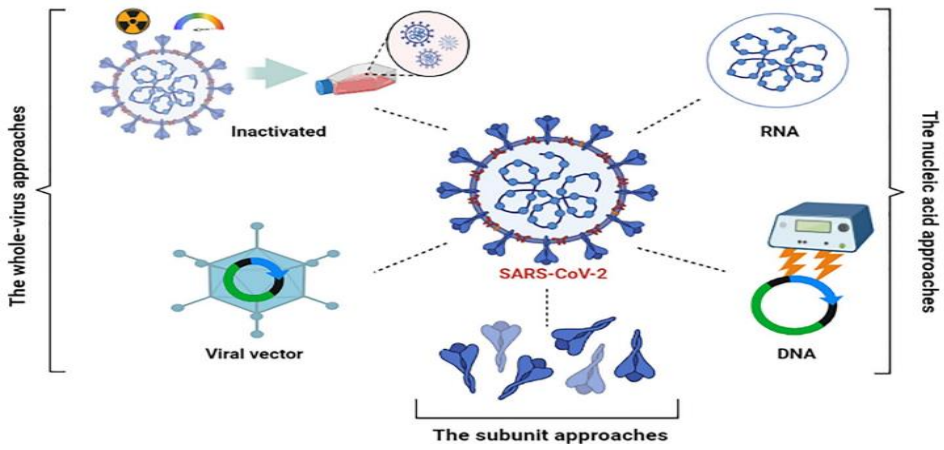


Figure 2.1- showing various techniques for designing COVID-19 vaccine

2.1.1 The whole virus approach– In this approach the complete virus or virion is used in vaccine synthesis, this platform is further classified into three forms which are as follows:

2.1.1.1 Inactivated Vaccines- It contains the complete virus which has been inactivated/ killed by exposure to U.V. radiations, chemicals or heat. The development of such vaccines require specific facilties for growth and culturing of virus safely, these vaccine are mostly used for flu and polio [50, 51]. They develop humoral response and are beneficial as they do not contain any live component. **CoronaVac, Covaxin,WIBP-CorV are examples of some common inactivated COVID-19 vaccines that are developed.**

2.1.1.2 Live attenuated Vaccines- These types of vaccines constitute a live, weak form of virus. MMR vaccine, the chickenpox vaccine and the shingles vaccines are all live attenuated vaccines. They resemble the weakened form of vaccines but might pose a risk when consumed/introduced into immune compromised individuals.

2.1.1.3 Viral vector vaccines- They include a safer viral form to deliver certain protein of pathogenic virus in the host to develop immunity without causing disease. these are developed quickly for example the Ebola virus vaccine [52, 53]. These vaccines generate both humoral and cellular responses and are essential as their large scale manufacturing is easy and are highly immunogenic. AZD1222 (chimpanzee adenoviral vector), As26.COV2S and Sputnik V (human adenoviral vector) are common viral vector vaccines for COVID-19.

2.1.2 Nucleic acid approach- In this approach a part of the genetic constituent of virus/ nucleic acid of virus is introduced in the vaccine. Set of instructions such as DNA or mRNA are introduced into a host cell by vaccine in this technique, this introduced instruction helps in protein synthesis which is specific and to which our immune system has to respond. These vaccines also generate the dual response and are beneficial as they are easy to manufacture and developed and are in least risk in triggering disease. mRNA-1273, and BNT162b2 are examples of few mRNA based covid-19 vaccines.

2.1.3 Subunit approach- In this approach rather than the complete virus being introduced into the host only a specific portion or the subunit of virus is introduced so that it can be recognized by host's immune system to develop immunity.it utilises purified forms of antigen [54]. They are also known to develop both humoral and cellular responses and are of greater significance when compared to other approaches as there is no risk with use as no live components are involved further the subunit vaccines are comparatively stable than other forms thereby, making their use more significant. EpiVacCorona and Novavax or NVX-CoV2373 are examples of subunit vaccines that aim on SARS CoV-2 by focussing on S protein.

2.2 Novavax NVX-CoV2373 Covid-19 vaccine:

It is a peptide subunit vaccine that was developed by the partnership of the American Biotechnology Novavax and the Coalition for Epidemic Preparedness Innovations Foundation. The drug Bank Accession Number of the vaccine is DB15810. This vaccine is of great significance as the S glycoprotein of the SARS CoV-2 collaborates with angiotensin-converting enzyme 2 (hACE2) explicited on the surface of the host cell to produce infection. The antibodies that are directed against the S protein can inhibit its interaction with hACE2 leading to viral neutralization. This vaccine consist of a recombinant mutant S protein that is codon optimised for Baculovirus mediated expression in Sf9cells that are attached to Matrix-M adjuvant, leading to formation of stable nanoparticles that can generate both cellular and humoral immune responses in Humans and animals [55-57]. The Novavax vaccine has the same utilises the same S glycoprotein sequence of SARS CoV-2 i.e. 1273amino acid long, along with dual mutations that are 682-RRAR-685 t 682-QQAQ-685 and two proline substitution at residues K986P and V987P. This vaccine is approved by Health Canada. The NVX-CoV2373 is introduced via intramuscular route.

Steps involved in synthesis of NVX-CoV2373 Vaccine are as follows:

1) Engineering of the full length, stabilized spike gene into Baculovirus.
2) The recombinant Baculovirus formed infects moth cells in the *S. frugiperda* (Sf9) expression system.
3) After infection, the spike proteins enters into the nucleus of Sf9 cells.

4) On entering the nucleus the spike DNA is transcribed into mRNA within the nucleus.
5) Sf9 cells then produce spike proteins on their translation and glycosylation in their native form that is trimer conformation.
6) Once formed, the spike protein trimmers are harvested and the vaccine nanoparticles are assembled around a Polysorbate 80 core (PS80).
7) The assembled nanoparticles are mixed with Matrix-M^{TM} adjuvant thereby creating a ready to used NVX-CoV2373 vaccine.

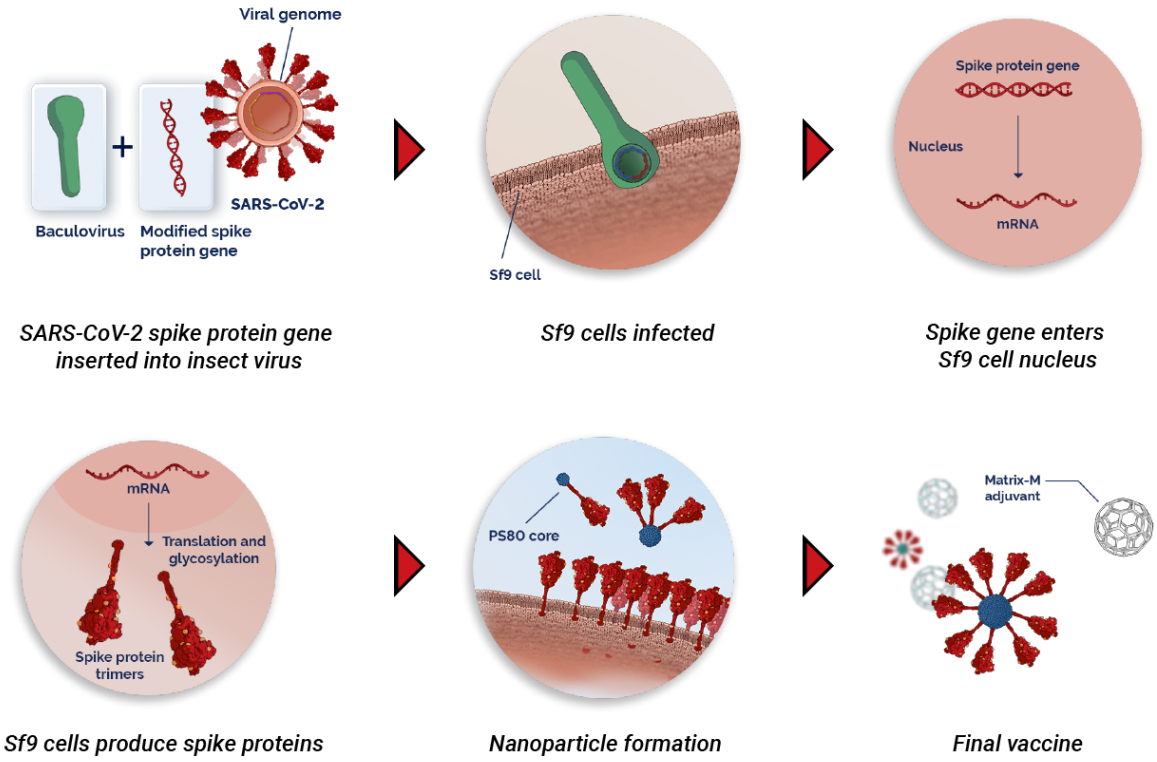


Figure 2.2 Steps involved in synthesis of Novavax vaccine.

Mechanism of action:

As the S protein of SARS-CoV-2 is of great importance in causing infection, active immunization with its variant is an essential method to regulate infection. The NVX-CoV2373 on immunization binds with hACE2 and exhibit higher anti-S IgG titers, increased levels of IFN- gamma and interlukins-4,5 and decreased viral load due to blocking of interactions.

Table 2.1 List of few vaccines available for SARS CoV-2 along with their type, company and target antigen

S.No.	Name of Vaccine	Type of Vaccine	Company	Target Antigen	Side effects
1	Covaxin	Inactivated virus	Bharat Biotech	Whole Virus	Pain, fever, headache, fatigue
2	BBIP-CorV	Inactivated virus	Sinopharm	Whole virus	Pain, fever, headache, fatigue
3	Ad26.COV2.S	Adenoviral vector	Janssen and Johnson & Johnson	S protein	Muscle pain, nausea, headache
4	Sputnik V	Adenoviral vector	Gamaleya	S protein	Hyperthermia, asthemia, muscle and joint pain
5	mRNA-1273	mRNA	Moderna	S protein	Pain, fever, headache, fatigue
6	NVX-CoV-2373	Protein subunit	Novavax	S protein	Muscle Pain, tenderness, fatigue

Although many platforms are being accessed for vaccine development but due to quick transmission and spread without symptoms, it can be stated that a global vaccine with higher coverage is required. Out of the given approaches limitations are associated with different techniques for vaccine productions such as:

1) Epitope alteration I case of inactivated vaccines.
2) Risk of residual virulence especially in immune compromised individuals for live attenuated vaccines.
3) For viral vector based vaccines are difficult to manufacture and there are chances of genome integration.
4) While genome based vaccines produce lower immunogenicity, difficult to administer requirement of lower room temperature for vaccine storage and transportation
5) There are risk for immune responses induced by RNA in RNA based vaccines.

Due to the above mentioned reasons Subunit vaccines are important vaccine candidates but it becomes difficult to try and analyse the vaccine against different variants and strains experimentally. As it is more time consuming, and expensive. Therefore, epitope predictions using Insilco tools can be done for various strains and vaccines to check the epitopic similarity and hence efficacy between vaccine and pathogen.

2.3 B-CELL EPITOPE PREDICTION

B cells are antigenic factors which are seen on the exterior facet of the pathogen or infectious agent and are known to interact with the hydrophobic binding site of B cell receptors (BCR) and have six hyper-variable loops of different length and constitution of amino acid. These prediction helps in recognising the B cell epitopes by exchanging the antigens required for the production of antibodies as well as their function-based studies. Since, any of the free or solvent-exposed areas in the antigen are the sites for antibody identification [58]. B cell epitopes are categorised into 2 classes.

- Linear or continuous epitopes consist of consecutive or successive residues and contribute a minority of the native antigens.
- Conformational or discontinuous epitopes consist of scattered solvent-exposed regions that may or may not be sequential.

Predominantly many B cell epitopes are discontinuous, as protein folding brings the far-off residues into close proximity due to protein folding [59]. Antibodies that are known to recognise linear or continuous epitopes have the potential to identify and recognise antigens that are denatured, which causes depletion of conformational B cell

recognition. Many approaches are accessible to investigate B cell epitopes, but they are unsuitable for genomic scale study due to their high cost and cumbersome nature [60]. Therefore, advanced computational tools are being designed for studying or estimating the B cell epitopes effectively, as they are quick, economical, and extensible (scalable) [61].

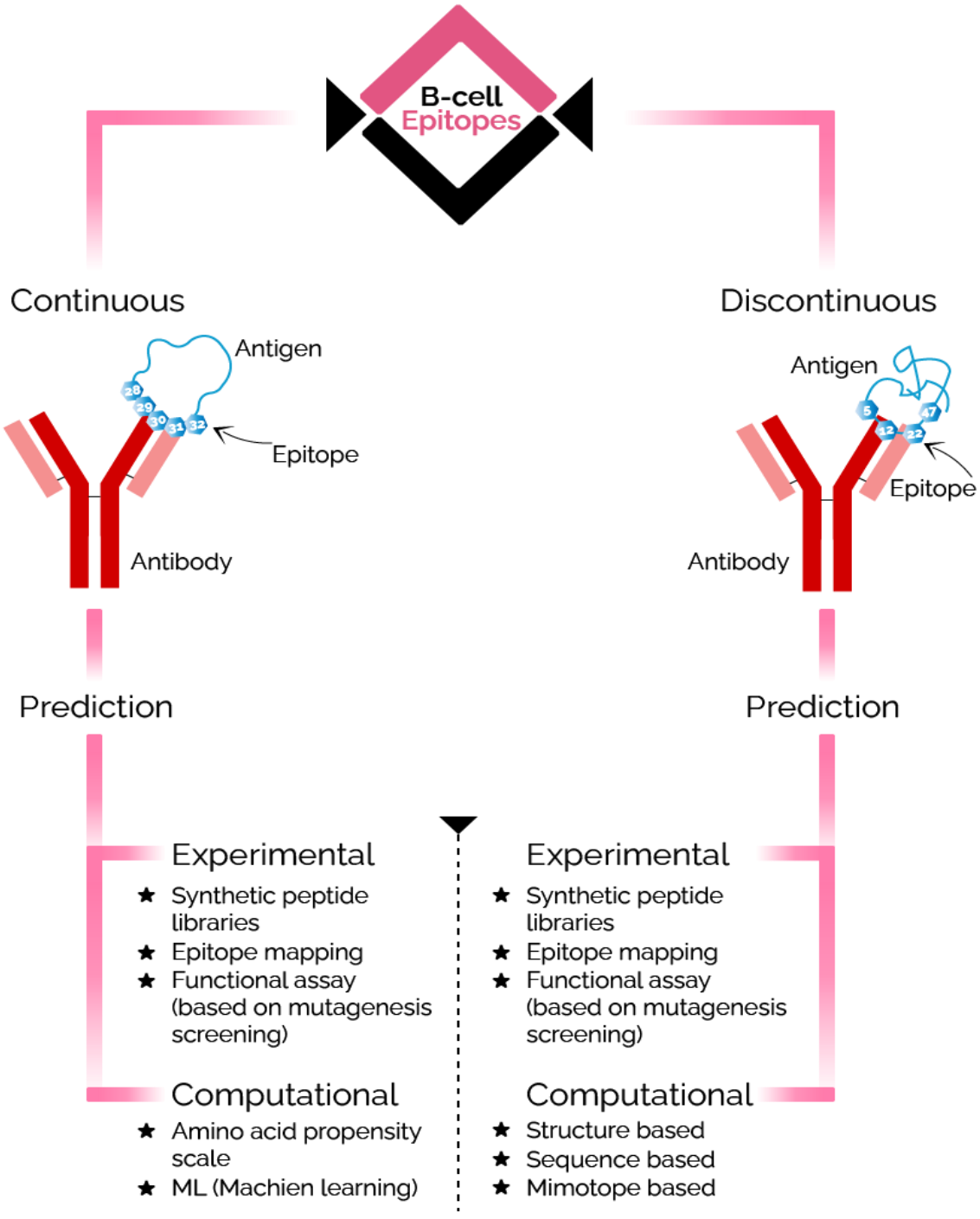


Figure. 2.3 Diagrammatic representation of different B-cell epitopes and their prediction methods.

2.3.1 Prediction of linear B cell epitopes: Although the linear B cell epitopes constitute the minority, they have gained much attention since they can substitute antigens for administration and antibody generation. Antibody-antigen binding in relation to Linear B cell epitope is dependent on the configuration, due to which this prediction is more complicated [61, 62].

- **Various methods for prediction are as follows:**

1. Methods based on a propensity scale or amino acid analysis Propensity scale methods: It is one of the traditional methods of identifying the most probable antigenic sequences. It computes the score for amino acid residues in the specified peptide considering the propensity values and the capacity of the amino acids that contribute in B cell epitope that can be allocated by the propensity score [63]. For instance, see Pellet J et al., 1993. The ones developed by Parker et al., Pekkequer et al., and Emini et al. use hydrophilicity turns, [64, 65] pliability, and solvent availability propensity scales, while those developed by Hopp and Woods [66. 67] use only the Levitt hydrophilicity scale. Propensity scale method assumes that hydrophilic areas on peptides surfaces are antigenic [65]. For example, tools such as Bepitope and PREDITOPE predicts epitopes based on turns in protein and the propensity scale values (more than 30) respectively.

2. **Machine learning methods:** Due to the inefficient performance of the above method, the machine based methods were developed whereby the ML (Machine learning) [68] algorithm initially translates the epitopes into feature vectors to look at the specified properties as given by propensity scales and then is trained to differentiate experimental B cell epitopes from the region which do not have any epitope [69].

TABLE 2.2 LIST OF TOOLS FOR LINEAR EPITOPE PREDICTION

TABLE

Methods	Tools	Server	About the tool
Linear B cell epitope			
Propensity scale methods	PEOPLE	http://www.iedb.org/	It uses a multipara metric algorithm based on hydrophilicity, accessibility, flexibility, and secondary structure properties of the amino acids along with the assessment of β-turns.
Machine learning methods	BepiPred (DT)	http://www.cbs.dtu.dk/services/BepiPred/	It is based on random forests trained on B-cell epitopes obtained from 3D-structures of antigen-antibody complexes.
	ABCpred (ANN)	http://www.imtech.res.in/raghava/abcpred/	It is a SVM-based model trained on anchoring pair composition.
	BCPREDS (SVM)	http://ailab.ist.psu.edu/bcpred/	Trained using various string kernels that eliminate the need for representing the sequence into length-fixed feature vectors.
	SVMtrip (SVM)	http://sysbio.unl.edu/SVMTriP/prediction.php	Trained on length-fixed tripeptide composition vectors.

2.3.2 Prediction of conformational B cell epitopes: although they constitute a majority, that is the percentage of conformational B cell epitopes is 90 in comparison to the linear B cell epitope, its estimation struggles for a few reasons:

1 It requires prior knowledge of protein 3–D structure, which is available only for a few proteins [70].
2. Any alteration in the folding of protein might change the number of epitopes.[71]
3. Separating discontinuous B cell epitope from its protein for specific antibody development requires an appropriate scaffold for epitope grafting, which makes the process time-consuming and complicated. Despite the above-mentioned factors, the

prediction helps in structure-function studies that are based on the antigen-antibody interaction.

- **Various methods for prediction are as follows:**

1. **Sequence-based Methods:** These methods do not require prior information about the structure of the target antigen. It assigns a score to the antigen and is an essential tools in conformational B cell prediction. While the machine learning methods also utilise sequence-based classifiers for a trustable prediction.

2. **Structure-based method:** The structure-based epitope prediction method considers the three dimensional protein structure [67], and it happens to be the most definite experimental method to identify epitopes. It takes into account the resolution of the antigen-antibody complex structure inferred from X-ray crystallography [68, 69]. There are various structure-based B-cell epitope predictors like CEP, Disco Tope, ElliPro, PEPITO, SEPPA, EPITOPIA, or EPIPRED. The CEP (Conformational Epitope Predictor) [70] was the first one, followed by the Disco Tope which was formed by Andersen et al. [71] operating on the protein structure statistics and spatial properties along with the surface accessibility of the amino acid [72]. The structure-based method is used less because of its high cost and the growing complexity of 3-D protein structures. This method is preferred over the sequence-based method because for training it uses a small-scale dataset and can identify the peptides for alleles that have been trailed before or in which the sequence-based method has failed [73]. Structure-based epitope prediction is the only method that predicts discontinuous epitopes.

3. **Mimotope-based prediction**: It is based on a combined approach for mapping epitopes,[74] by random peptide library preparation that are observed against the required antibody creating a group called Mimotopes, which copies the features and organization and not the sequence of the real epitopes [75, 76].

TABLE 2.3 LIST OF TOOLS FOR CONFORMATIONAL B – CELL EPITOPE PREDICTION

Conformational B-cell epitope			
Sequence-based prediction methods	CBTOPE (SVM)	http://www.imtech.res.in/raghava/cbtope/submit.php	SVM-based model trained on the physiochemical and sequence featured of conformational B-cell epitopes.
VStructure-based prediction methods	CEP	http://bioinfo.ernet.in/cep.htm	Relies entirely on predicting patches of solvent-exposed residues.
	DiscoTope	http://tools.iedb.org/discotope	In addition to solvent accessibility, it also considers amino acid statistics and spatial information to predict conformational B-cell epitopes.
	ElliPro	http://tools.iedb.org/ellipro/	Identify protruding regions in antigen surfaces.
	PEPITO	http://pepito.proteomics.ics.uci.edu/	Combine single physicochemical properties of amino acids and geometrical structure property
	SEPPA	http://lifecenter.sgst.cn/seppa/	Combine single physicochemical properties of amino acids and geometrical structure property
	EPITOPIA	http://epitopia.tau.ac.il/	Based on naïve Bayes classifiers and support vector regressions.
	EPSVR	http://sysbio.unl.edu/EPSVR/	Based on naïve Bayes classifiers and support vector regressions.
	EPIPRED	http://opig.stats.ox.ac.uk/webapps/sabdabsabpred/EpiPred.php	It uses a docking-like approach to match up antibody and antigen structures, thus identifying epitope regions on the antigen.
	PEASE	http://www.ofranlab.org/PEASE	It utilizes the sequence of the antibody and the 3D-structure of the antigen
Mimotope analysis -based prediction	PEPITOPE	http://pepitope.tau.ac.il/	It predict epitopes based on peptides extracted from a phage display library, or to align a linear peptide sequence onto a three dimensional protein structure.
	EpiSearch	http://curie.utmb.edu/episearch.html	It predicts conformational epitopes on antigen protein using peptides selected from phage display experiments.

CHAPTER 3

MATERIAL AND METHODOLOGY

Since the epitope prediction and analysis is done in silico using software's therefore, no such materials are required as such for analysis except the software's.

3.1 Tools Used:

3.1.1 BLAST

It stands for Basic Local alignment search tool it helps in finding out the areas of similarity between sequences. The sequences used can be either nucleotide or protein sequences. It helps to determine the statistical significance of matches and can even detect the evolutionary relationships between the sequences. BLAST are of several types:

1) BLASTn- it is a nucleotide BLAST, it analyse one or more nucleotide query sequence to a subject nucleotide sequence. Used mainly in determining the evolutionary relationship amongst various organisms.
2) BLASTx- it looks for translated nucleotide sequences against protein sequences i.e. it analyses a query of sequence of nucleotide that is translated in all six frames of

reading against a database of protein sequence and is usually the initial test that is performed for newly obtained sequences.

3) tBLASTn- it analyses the query protein sequence against a six frame translation of nucleotide sequence. And is essential in locating the protein coding regions that are homologous in the unannotated nucleotide sequences such as ESTs and HTG.
4) BLASTp- it stands for protein BLAST and analyses a single or many protein sequences (query) to another subject protein sequence. Used mainly for protein identification.

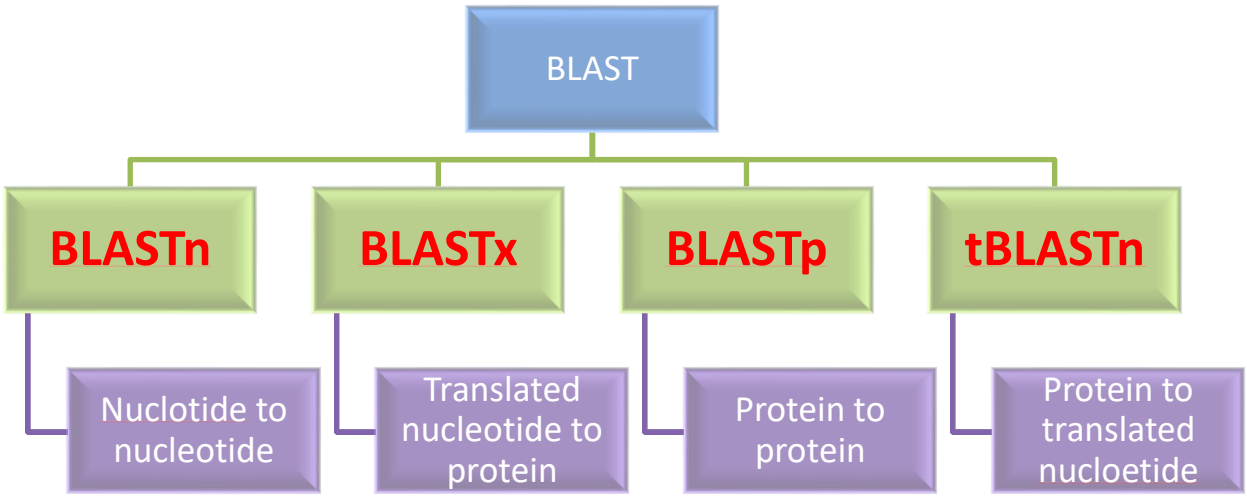


Figure 3.1 Schematic representation of Various types of BLAST

3.1.2 Bepipred:

As described previously B cell epitopes are protein fragment that bind to antibodies, a large number of these epitopes are made up of variable segments of the peptide chains that are carried together by the protein turn over, but for a minority around 10% of the epitopes, the related antibodies are reactive against a linear peptide [77] (constitute a single polypeptide stretch).

The traditional method for linear B cell epitope prediction is the Propensity scale method, where looking at the physio chemical properties, it provides each amino acid with a propensity value. And a running average window was discovered to decrease the changes in the prediction values [78]

For the B cell epitope prediction a variety of structure tools are available but they need prior information regarding the antigen structure [79-82] but this situation is not possible for all the majority of cases as the structural information is not available for all the antigens/ proteins. Therefore, the left out option always is analysing the sequence based information but prediction through such information is not accurate. There is a need of reliable prediction tools in numerous healthcare and biotechnological applications such as in antibody development, vaccine design, efficacy testing and understanding of the immune system [83-85].

Therefore, Bepipred-1.0 was introduced which is a linear sequence based tool for B cell epitope prediction. It is an innovative approach for linear B cell epitopes prediction, and is better than random predictions and the variety of propensity scales being tested. This tool is an amalgamation method based on the prediction of Hidden Markov Model and the propensity scale by Parker et.al [86]. The Bepipred-1.0 server and data sets are available publicly at http://www.cbs.dtu.dk/services/BepiPred. On providing input sequence to the software, linear epitopes are predicted and are represented in both graphical and tabular format. The residues with a score value above the threshold (i.e. default value of 0.35) are forecasted as an epitope part and are displayed in yellow on the graph (where Y-axes shows the residue scores and X-axes depicts the residue positions in the sequence) and is shown with "E" in the results.

Table 3.1 it shows the correlation between selected thresholds and the sensitivity/specificity of the prediction method.

Threshold	Sensitivity	Specificity
-0.20	0.75	0.50
0.20	0.56	0.68
0.35	0.49	0.75
0.90	0.25	0.91
1.30	0.13	0.96

But this tool shows lower sensitivity to higher thresholds and any reduction in threshold will decrease the specificity so Bepipred-2.0 was launched and is skilled entirely on epitope data obtained from the crystal structures, which is assumed to be prime and indeed leads to considerably accelerated predictive performance when compared to other reachable tools. To use Bepipred-2.0 the protein sequence of interest has to be uploaded in the FASTA format to this interface http://www.cbs.dtu.dk/services/BepiPred/.

The BepiPred-2.0 server takes into account the protein sequence to give the B-cell epitopes, it uses a Random Forest algorithm that is instructed to determine the amino acid that are on epitopes and non-epitopes by looking at the crystal structures. After the above process a step wise smoothing is done and finally the residues predicted with value above the limiting value (default value is 0.5) are calculated as an epitopic part and is shown with yellow on the graph and pronounced with "E" in results.

Table 3.2 it shows the correlation between threshold that is selected and the sensitivity/specificity of the prediction method.

Threshold	Sensitivity	Specificity
0	1	0
0.05	1	0
0.10	1	0
0.15	1	0
0.20	1	0.00019
0.25	0.99743	0.00419
0.30	0.98995	0.0276
0.35	0.97212	0.07036
0.40	0.93605	0.15606
0.45	0.82607	0.3307
0.50	0.58564	0.57158
0.55	0.29159	0.81655
0.60	0.09559	0.95116
0.65	0.01969	0.99272
0.70	0.00182	0.99954
0.75	0	1
0.80	0	1
0.85	0	1
0.90	0	1
0.95	0	1
1	0	1

In reality, software or tool users are interested only in inspecting the top-scoring predictions, in order to prioritize the data to be checked experimentally. On comparing the AUC values for Bepipred -1.0 and 2.0 the AUC scores for 2.0 is comparatively higher that is around 0.548.

Bpipred-1.0	Bepipred-2.0
• http://www.cbs.dtu.dk/services/BepiPred	• http://www.cbs.dtu.dk/services/BepiPred
• Not trained on the crystal structure data	• Instructed entirely on crystal structure epitope data.
• AUC score 0.548.	• AUC score 0.574.
• Combination of HMM and propensity scale method	• Uses Random Forest Algorithm
• Less predictive power	• More predictive power
• Less sensitive	• More sensitive and specific

Table 3.3 differences between Bepipred 1.0 and 2.0

3.2 Methodology:

3.2.1 Retrieving Data from NCBI

NCBI virus is a public server that contains information about the viral sequence data from RefSeq, GenBank and other NCBI repositories. It has a quick access to SARS CoV-2 data such as Refseq, Genbank, and Nucleotide and protein sequence information.

❖ Steps to retrieve data in NCBI virus:

- Opening NCBI virus in Toolbar
- Selecting novel SARS CoV-2 protein sequences from quick access data.
- A new web page will appear where data is filtered by selecting surface glycoproteins from the protein filter and specific pangolin series from pangolin filter.
- Obtained results in the table were looked for the sequence of interest and their accession numbers were selected.
- Using the specific accession number, surface glycoprotein sequences for different strains of SARS CoV-2 were retrieved. (i.e. Alpha, Omicron, Lambda, Eta, Gamma variants)

3.2.2 Retrieving Sequence Data of Novavax Vaccine (NVX-CoV2373)

- SARS CoV-2 S glycoprotein sequence was obtained from NCBI.
- Two mutations were performed in the sequence to obtain the sequence of the desired vaccine:

1) First, mutation was performed at the furin cleavage site 682-RRAR-685 to 682-QQAQ-685,
2) Second, is 2 proline substitutions were performed at residues K986P and V987P respectively.

3.2.3 Performing BLASTp

- **Protein BLAST: Align two or more sequences using BLAST (nih.gov), this site was accessed.**
- **Multiple sequence alignment were choosed and in the query space the sequence of Peptide Vaccine (Novavax) was copied**
- While in the subject sequence the retrieved sequence of the above mentioned SARS CoV 2 strains were added one after the other.

- And BLASTp was performed to determine the percentage identity between the strain and vaccine sequences.

3.2.4 Linear B cell Epitope Prediction:

- IEDB analysis resources were used through the link (Antibody Epitope Prediction (iedb.org)).
- Bepipred linear B cell prediction 2.0 was selected form the below options.
- Sequences retrieved for the peptide vaccine as well as covid strains were pasted individually in the space provided.
- The results of the prediction obtained for the vaccines were compared with each of the covid strains.

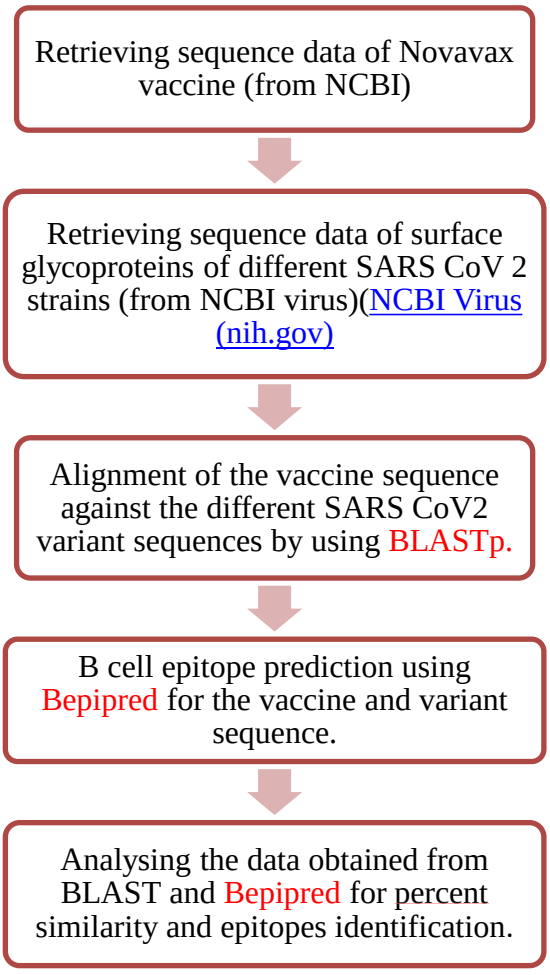


Figure 3.2- Schematic representation of the overall methodology involved.

CHAPTER 4

RESULT AND DISCUSSION

4.1 Results

4.1.1 SEQUENCES OF SARS CoV-2 VARIANTS

Since SARS CoV-2 is known to have a variety of variants but here, data of only 5 of the variants i.e. Alpha, Gamma, Lambda, ETA and Omicron are obtained from NCBI virus. Sequence information of the variants are attached in the Appendix1 below. While a tabular representation of the variants under study, their accession number obtained the Pangolin lineage they belong to, the country the variant was first reported and geographical location of the variant are listed in the table below**:**

Table 4.1 list of variants of SARS CoV-2 under study along with their accession number obtained.

SN O.	COVID VARIANT	PANGOLIN SERIES	ACESSION NUMBER	GEOGRAPHICAL LOCATION	FIRST REPORTED IN
1	ALPHA VARAINT	B.1.1.7	UPI04166	U.S.A., U.K.	U.K.
2	OMICRON VARIANT	B.1.1.529	UOL24929	U.S.A.	SOUTHERN AFRICA
3	GAMMA VARIANT	P.1	UPI04453	U.S.A., JAPAN	JAPAN (TOKYO)
4	LAMBDA VARIANT	C.37	UPI04200	U.S.A. (PERU)	PERU
5	ETA VARIANT	B.1.525	UPD40192	FRANCE	U.K., NIGERIA

4.1.2 BLASTp Results

While performing BLASTp result obtained has few parameters such as the maximum score, total score, E value (it is basically the likelihood that 2 sequences are similar by chance) percentage identity (it determines the similarity between the aligned and query sequence), accession length and query number (which is a unique identifier).

The results for the alignment shoe a similar e value of 0.0 for all the sequences under study while the percentage identity was different i.e a value of around 98.82% which was the highest amongst all was obtained for the alpha variant, followed by value of

about 98.66% for both eta and Gamma Variant and a least value of 98.59 % for lambda variant.

On obtaining the distance tree of the result, the peptide vaccine was found to be more evolutionary linked to lambda variant than the alpha, gamma, eta and omicron variants.

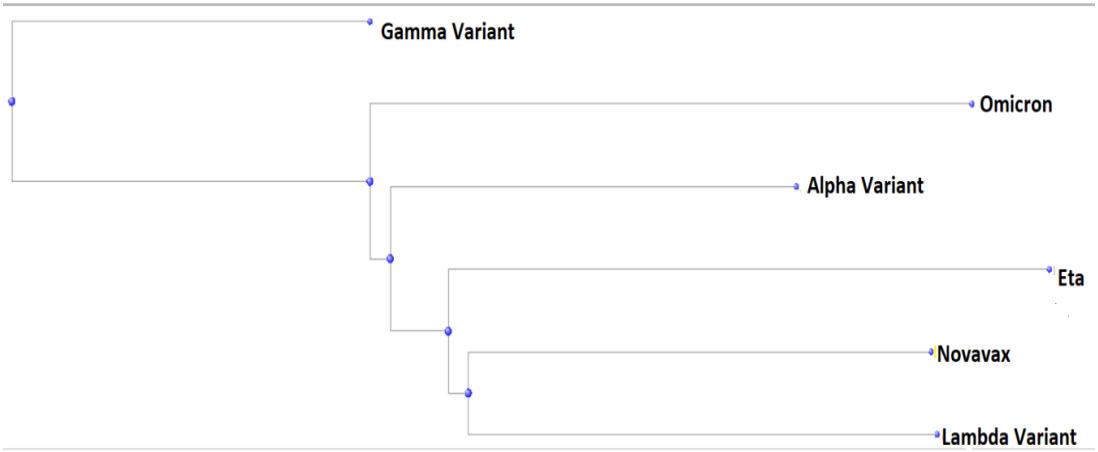


Figure 4.1- Blast Tree view of the provided sequences of Vaccine and SARS CoV2 strain

Sequences producing significant alignments Download Select columns Show 100

select all *0 sequences selected* Graphics Distance tree of results Multiple alignment MSA Viewer

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
Gamma Variant		2598	2598	100%	0.0	98.66%	1273	Query_10748
Alpha Variant		2591	2591	100%	0.0	98.82%	1270	Query_10747
Omicron		2591	2591	100%	0.0	98.59%	1270	Query_10750
Eta Variant		2589	2589	100%	0.0	98.66%	1270	Query_10751
Lambda Variant		2586	2586	100%	0.0	98.51%	1266	Query_10749

Figure 4.2 - Blastp results for alignment of multiple sequence of SARS CoV 2 against the vaccine sequence.

4.1.3 Bepipred Results

The linear epitopes for the peptide NOVAVAX vaccine and different SARS CoV2 variants were predicted using Bepipred 2.0. 35 linear epitopes were predicted for the NOVAVAX while the number of epitopes for the variants of SARS CoV 2 were 35 for the alpha and variant and 36 and 29 for the lambda and gamma variants respectively. Out of the 35 peptides of alpha variants 9 of the predicted epitopes were similar to that of the vaccine, same is the case for both gamma and Omicron variant as well. While a highest of around 11 linear epitopes of the lambda variant were found similar to the epitopes predicted for vaccine Likewise, 10 epitopes of the eta variant were found similar to the NOVAVAX epitopes. Therefore, with these results it can be interpreted that the vaccine might show more efficacy towards the lambda variant of coronavirus followed by the eta variant.

TABLE 4.2 Bepipred result of vaccine sequence and different covid strains, here the predicted linear epitopes of vaccine are mentioned and the cells are filled with a distinct color, the epitope number of the variants that match the epitopes of vaccine are also colored with the similar color.

predicted peptide no.	Predicted Peptides of NOVAVAX VACCINE	Alpha Variant	Gamma Variant	Lambda variant	Omicron	eta variant
1	SQCVNLTTRTQLPPAYTNSFTRGVY					
2	FSNVTWFHAIHVSGTNGTKRFDN					
3	KS					
4	DPFLGVYYHKNNKSWME					
5	MDLEGKQGNFKNL					
6	KHTPINLVRDLPQGFS					
7	TPGDSSSGWTA					
8	DPL					
9	FTVE					
10	YQTSNFRVQP					
11	PNITNLCPFGEVFNATRFASVYAWNRKRISNCVA					
12	LYNSASFSTFKCYGVSPTKLNDLCFT					
13	GDEVRQIAPGQTGKIADY					

14	YKL					
15	NLDSKVGGNYN					
16	FRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTN					
17	ELLHAPATVCGPKKSTNLVKN					
18	SNKKFLP					
19	QTLE					
20	TNTSN					
21	NCTEVPVAIHADQLTPTWRVYSTGSNVF					
22	VNNSYECDIP					
23	ASYQTQTNSPQQAQSVASQSIIAYTMSLGAENSVAYSNN					
24	E					
25	DKNTQ					
26	KQIYKTPPIKDFGGF					
27	PDPSKPSKR					
28	LADAGFIKQYGDCLGD					
29	DPPEAEVQI					
30	GQSKRVDFC					
31	RNFYEPQIITTD					
32	VNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGI					
33	LGKY					
34	SCCKFDEDDSEPVLKG					
35	K					
36	_					
37	_					

4.2 DISCUSSIONS

With the ongoing COVID-19 pandemic it is important to provide vaccination to all individuals as it is important for global health. The actual process of vaccine development that is from the idea to experimentation, results, clinical trials, approval to manufacturing and consumption is a long process that takes 10- 11 years but in the pandemic situation there is shortage of time and resource. Although many vaccines are being developed by different institutions against SAR-CoV-2 but to check the efficacy of vaccine against different strains of SARS CoV-2 experimentally is time consuming and labour intensive. Therefore, we have focussed on the in silico approach for analysing the efficacy of different variants of virus against the NOVAVAX peptide subunit vaccine as these vaccines are easy to manufacture and there are no associated risk of virulence and are safe to use. The linear epitopes of vaccine are compared to those predicted of the variants of the virus and it was observed that the maximum number of predicted epitopes of vaccine were found similar to the Lambda variant followed by the Eta variant of COVID-19. Hence, it can be proposed that out of all studies variants of COVID-19, NOVAVAX vaccine can have more efficacy against the Lambda and Eta variants due to greater similarity amongst predicted epitopes. Since both the Lambda and Eta variants are geographically located in the U.S.A mainly in Southern America (Chile and Peru) and U.K. respectively. But in these countries this vaccine is not approved yet and is still under 1st clinical trial. Therefore, on approval in the U.S.A. and U.K., Novavax vaccine could work wonders in these geographical areas.

APPENDICES

Appendix 1

SEQUENCE OF NOVAVAX VACCINE

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFR
SSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIR
GWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVY
SSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQ
GFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRT
LKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITN
LCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCF
TNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYN
YLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPY
RVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFG
RDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAI
HADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPQ
QAQSVASQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTM
YICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFG
GFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFN
GLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQN
VLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGA
ISSVLNDILSRLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMS
ECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAH
FPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELD
SFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELG
KYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSE
PVLKGVKLHYT

S Glycoprotein sequences of different variants of SARS CoV-2 (under study)

1) ALPHA VARIANT (UPI04166.1)

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAISG
TNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLG
VYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLV
RDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENG
TITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAW
NRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNY
KLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQ
SYGFQPTYGVGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQ
QFGRDIDDTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTW
RVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSHRRARSVASQSIIAYTMSLGAEN
SVAYSNNSIAIPINFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVE
QDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGD
IAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQ
NVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILAR
LDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFP
QSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTHNTFVS
GNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNL
NESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEP
VLKGVKLHYT

2) GAMMA VARIANT (P.1)

MFVFLVLLPLVSSQCVNFTNRTQLPSAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHV
SGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNYPF
LGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLSEFVFKNIDGYFKIYSKHTPI
NLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYN
ENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASV
YAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGTIAD
YNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVKGFNCYF
PLQSYGFQPTYGVGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFL
PFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLT
PTWRVYSTGSNVFQTRAGCLIGAEYVNNSYECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMSLG
AENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGI
AVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDC
LGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIG
VTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDI
LSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAAIKMSECVLGQSKRVDFCGKGYHLM
SFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNT
FVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASFVNIQKEIDRLNEVA
KNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDD
SEPVLKGVKLHYT

3) ETA VARIANT (UPD40192.1)

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTRDLFLPFFSNVTWFHVISG
TNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLG
VYHKNNNSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLV
RYLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENG
TITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAW
NRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNY
KLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVKGFNCYFPLQ
SYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQ
QFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTW

RVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTHTNSPRRARSVASQSIIAYTMSLGAEN
SVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVE
QDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGD
IAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTLGAGAALQIPFAMQMAYRFNGIGVTQ
NVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSR
LDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFP
QSAPHGVVFLHVTYVPAQEKNFTTAPAICHGGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVS
GNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNL
NESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEP
VLKGVKLHYT

4) LAMBDA VARIANT (UPI04200.1)

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHV
SGTNVIKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPF
LGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPI
NLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHNSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITD
AVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKR
ISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPD
DFTGCVIAWNSNNLDSKVGGNYNYQYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYSPLQSYGF
QPTNGVGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGR
DIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTWRVYS
TGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMSLGAENSVAY
SNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKN
TQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAAR
DLICAQKFNGLNVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLY
ENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKV
EAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAP
HGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCD
VVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESL
IDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKG
VKLHYT

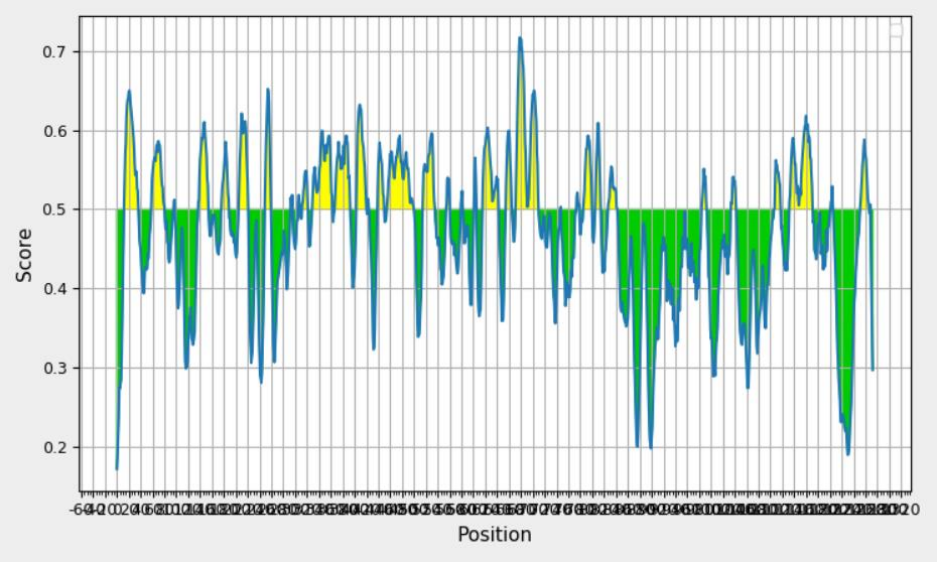
5) OMICRON (B1.1.529)

MFVFLVLLPLVSSQCVNLITRTQSYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSGT
NGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLDV
YYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLG
RDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENG
TITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAW
NRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNY
KLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQ
SYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQ
QFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTW
RVYSTGSNVFQTRAGCLIGAEYVNNSYECDIPIGAGICASYQTQTKSHRRARSVASQSIIAYTMSLGAEN
SVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVE
QDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGD
IAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQ
NVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNHNAQALNTLVKQLSSKFGAISSVLNDILSR
LDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFP
QSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVS
GNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNL
NESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEP
VLKGVKLHYT

Appendix 2

Bepipred results

1) Graphical and tabular results of predicted peptides for NOVAVAX vaccine



Predicted peptides:

No.	Start	End	Peptide	Length
1	13	37	SQCVNLTTRTQLPPAYTNSFTRGVY	25
2	59	81	FSNVTWFHAIHVSGTNGTKRFDN	23
3	97	98	KS	2
4	138	154	DPFLGVYYHKNNKSWME	17
5	177	189	MDLEGKQGNFKNL	13
6	206	221	KHTPINLVRDLPQGFS	16
7	250	260	TPGDSSSGWTA	11
8	294	296	DPL	3
9	306	309	FTVE	4
10	313	322	YQTSNFRVQP	10
11	330	363	PNITNLCPFGEVFNATRFASVYAWNRKRISNCVA	34
12	368	393	LYNSASFSTFKCYGVSPTKLNDLCFT	26
13	404	421	GDEVRQIAPGQTGKIADY	18
14	423	425	YKL	3
15	440	450	NLDSKVGGNYN	11
16	456	501	FRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTN	46
17	516	536	ELLHAPATVCGPKKSTNLVKN	21
18	555	561	SNKKFLP	7
19	580	583	QTLE	4

20	602	606	TNTSN	5
21	616	643	NCTEVPVAIHADQLTPTWRVYSTGSNVF	28
22	656	665	VNNSYECDIP	10
23	672	710	ASYQTQTNSPQQAQSVASQSIIAYTMSLGAENSVAYSNN	39
24	748	748	E	1
25	775	779	DKNTQ	5
26	786	800	KQIYKTPPIKDFGGF	15
27	807	815	PDPSKPSKR	9
28	828	843	LADAGFIKQYGDCLGD	16
29	985	993	DPPEAEVQI	9
30	1035	1043	GQSKRVDFC	9
31	1107	1118	RNFYEPQIITTD	12
32	1133	1172	VNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGI	40
33	1203	1206	LGKY	4
34	1252	1267	SCCKFDEDDSEPVLKG	16
35	1269	1269	K	1

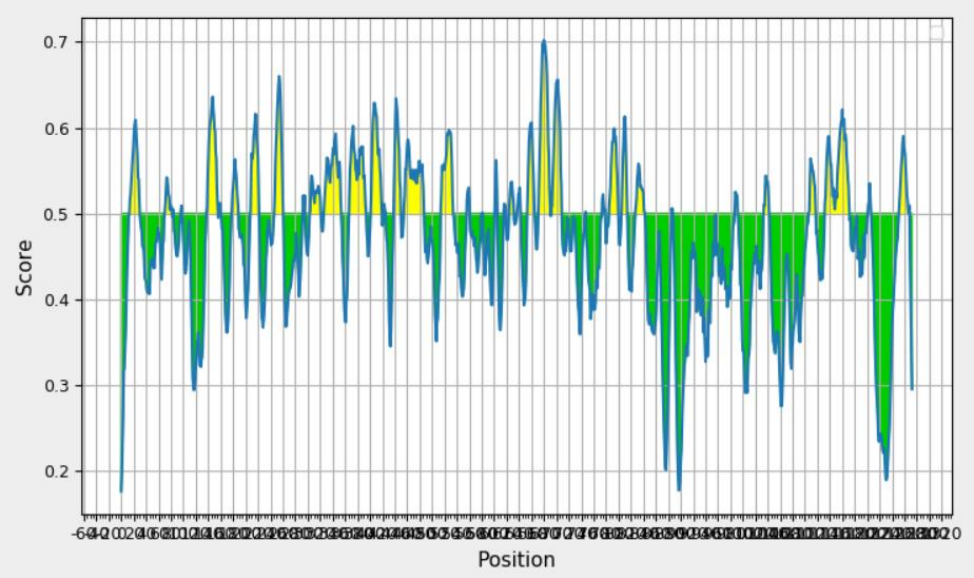
2) Graphical and tabular results of predicted peptides for Alpha variant of COVID-19

Predicted peptides:

No.	Start	End	Peptide	Length
1	13	37	SQCVNLTTRTQLPPAYTNSFTRGVY	25
2	55	63	FLPFFSNVT	9
3	67	82	AISGTNGTKRFDNPVL	16
4	94	97	EKSN	4
5	137	162	PFLGVYHKNNKSWMESEFRVYSSANN	26
6	173	188	LMDLEGKQGNFKNLRE	16
7	205	219	TPINLVRDLPQGFSA	15
8	246	257	LTPGDSSSGWTA	12
9	291	293	DPL	3
10	303	320	FTVEKGIYQTSNFRVQPT	18
11	326	351	FPNITNLCPFGEVFNATRFASVYAWN	26
12	363	389	SVLYNSASFSTFKCYGVSPTKLNDLCF	27
13	399	416	IRGDEVRQIAPGQTGKIA	18
14	420	422	YKL	3
15	436	448	NNLDSKVGGNYNY	13
16	455	500	KSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTYGV	46
17	513	532	ELLHAPATVCGPKKSTNLVK	20
18	552	559	SNKKFLPF	8
19	579	579	L	1
20	600	602	NTS	3

21	613	642	NCTEVPVAIHADQLTPTWRVYSTGSNVFQT	30
22	653	662	VNNSYECDIP	10
23	669	687	ASYQTQTNSHRRARSVASQ	19
24	692	707	YTMSLGAENSVAYSNN	16
25	745	745	E	1
26	770	776	EQDKNTQ	7
27	783	797	KQIYKTPPIKDFGGF	15
28	804	812	PDPSKPSKR	9
29	825	840	LADAGFIKQYGDCLGD	16
30	985	989	EAEVQ	5
31	1031	1040	LGQSKRVDFC	10
32	1104	1115	RNFYEPQIITTH	12
33	1130	1169	VNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGI	40
34	1199	1203	ELGKY	5
35	1250	1266	CCKFDEDDSEPVLKGVK	17

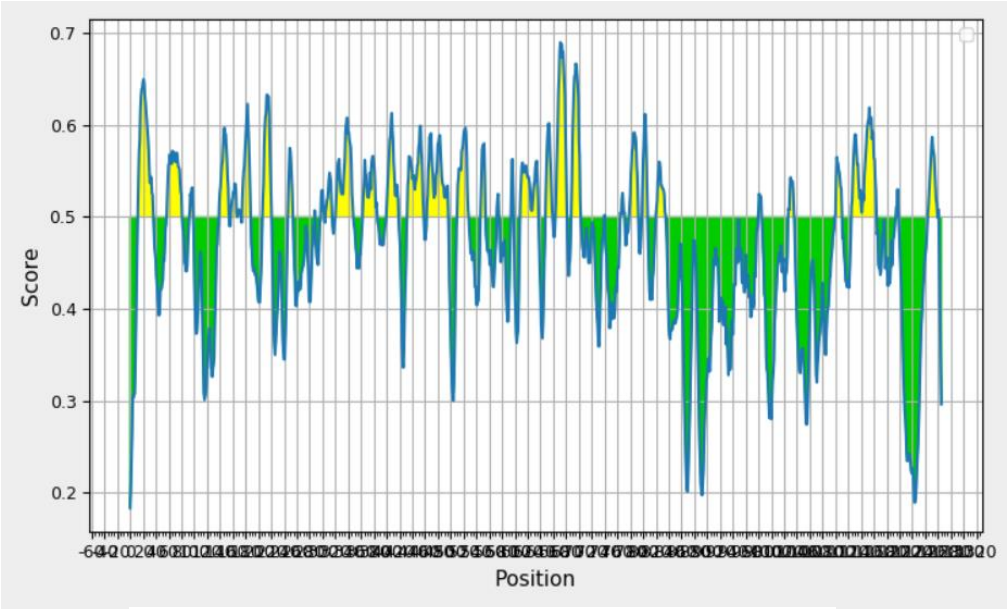
3) Graphical and tabular results of predicted peptides for Gamma variant of COVID-19



Predicted peptides:

No.	Start	End	Peptide	Length
1	14	30	QCVNFTNRTQLPSAYTN	17
2	71	82	SGTNGTKRFDNP	12
3	84	84	L	1
4	97	98	KS	2
5	138	156	YPFLGVYYHKNNKSWMESE	19
6	158	161	RVYS	4
7	179	189	LEGKQGNFKNL	11
8	209	222	PINLVRDLPQGFSA	14
9	247	261	SYLTPGDSSSGWTAG	15
10	293	296	LDPL	4
11	305	321	SFTVEKGIYQTSNFRVQ	17
12	327	354	VRFPNITNLCPFGEVFNATRFASVYAWN	28
13	368	395	LYNSASFSTFKCYGVSPTKLNDLCFTNV	28
14	403	420	RGDEVRQIAPGQTGTIAD	18
15	423	426	YKLP	4
16	439	450	NNLDSKVGGNYN	12
17	456	488	FRKSNLKPFERDISTEIYQAGSTPCNGVKGFNC	33
18	515	534	FELLHAPATVCGPKKSTNLV	20
19	557	561	KKFLP	5
20	582	582	L	1
21	602	606	TNTSN	5
22	617	619	CTE	3
23	625	632	HADQLTPT	8
24	638	643	TGSNVF	6
25	655	665	YVNNSYECDIP	11
26	672	691	ASYQTQTNSPRRARSVASQS	20
27	693	709	IAYTMSLGAENSVAYSN	17
28	748	748	E	1
29	775	779	DKNTQ	5
30	786	800	KQIYKTPPIKDFGGF	15
31	806	815	LPDPSKPSKR	10
32	828	843	LADAGFIKQYGDCLGD	16
33	887	888	TF	2
34	988	992	EAEVQ	5
35	1035	1043	GQSKRVDFC	9
36	1107	1118	RNFYEPQIITTD	12
37	1133	1172	VNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGI	40
38	1203	1207	LGKYE	5
39	1252	1269	SCCKFDEDDSEPVLKGVK	18

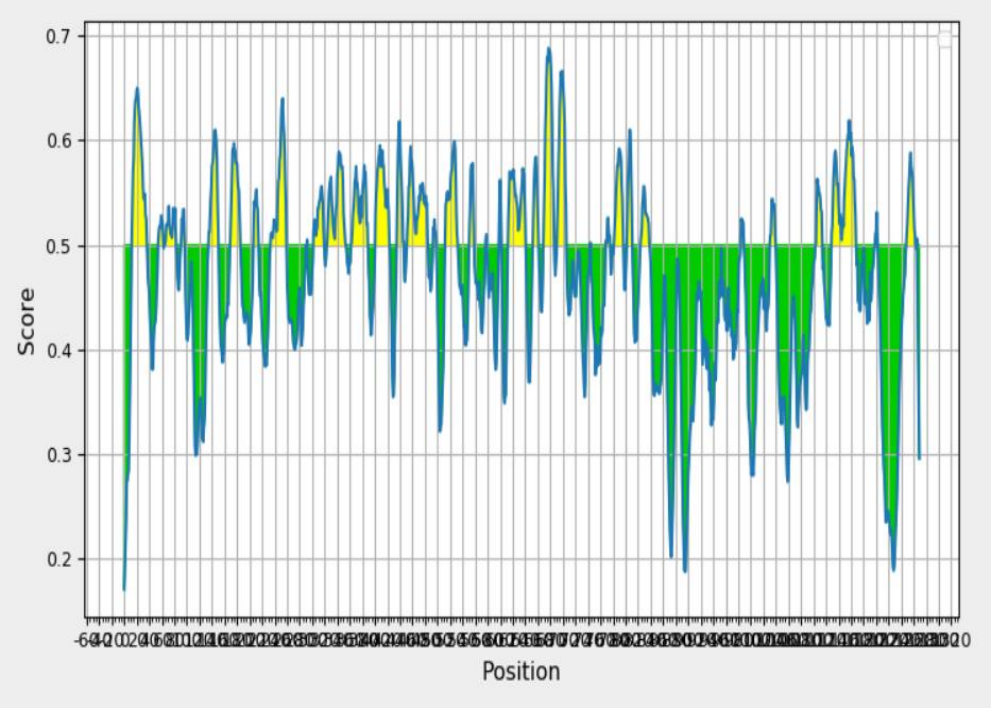
4) Graphical and tabular results of predicted peptides for Lambda variant of COVID-19



Predicted peptides:

No.	Start	End	Peptide	Length
1	13	37	SQCVNLTTRTQLPPAYTNSFTRGVY	25
2	58	82	FFSNVTWFHAIHVSGTNVIKRFDNP	25
3	94	99	STEKSN	6
4	140	155	FLGVYYHKNNKSWMES	16
5	158	174	RVYSSANNCTFEYVSQP	17
6	176	188	LMDLEGKQGNFKN	13
7	208	221	TPINLVRDLPQGFS	14
8	247	254	SSSGWTAG	8
9	289	289	L	1
10	299	304	FTVEKG	6
11	306	315	YQTSNFRVQP	10
12	321	349	RFPNITNLCPFGEVFNATRFASVYAWNRK	29
13	364	386	SASFSTFKCYGVSPTKLNDLCFT	23
14	401	419	RQIAPGQTGKIADYNYKLP	19
15	433	458	NLDSKVGGNYNYQYRLFRKSNLKPFE	26
16	464	497	EIYQAGSTPCNGVEGFNCYSPLQSYGFQPTNGVG	34
17	511	528	LHAPATVCGPKKSTNLVK	18
18	548	555	SNKKFLPF	8
19	573	576	QTLE	4
20	595	599	TNTSN	5
21	610	637	CTEVPVAIHADQLTPTWRVYSTGSNVFQ	28
22	649	659	VNNSYECDIPI	11
23	665	682	ASYQTQTNSPRRARSVAS	18
24	689	703	TMSLGAENSVAYSNN	15
25	741	741	E	1
26	766	772	EQDKNTQ	7
27	779	793	KQIYKTPPIKDFGGF	15
28	799	808	LPDPSKPSKR	10
29	821	836	LADAGFIKQYGDCLGD	16
30	981	985	EAEVQ	5
31	1028	1036	GQSKRVDFC	9
32	1100	1111	RNFYEPQIITTD	12
33	1126	1165	VNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGI	40
34	1195	1199	ELGKY	5
35	1245	1260	SCCKFDEDDSEPVLKG	16
36	1262	1262	K	1

5) Graphical and tabular results of predicted peptides for ETA variant of COVID-19

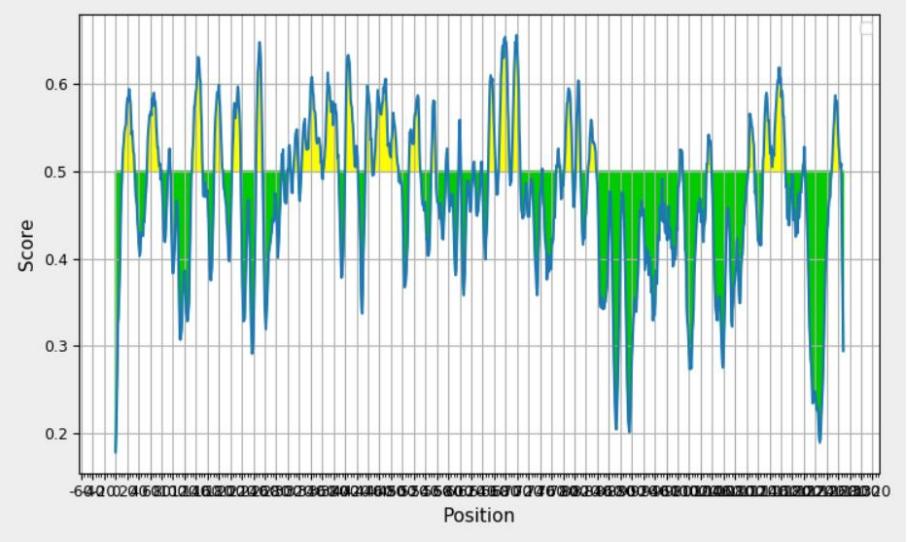


Predicted peptides:

No.	Start	End	Peptide	Length
1	13	37	SQCVNLTTRTQLPPAYTNSFTRGVY	25
2	55	63	FLPFFSNVT	9
3	65	83	FHVISGTNGTKRFDNPVLP	19
4	92	97	STEKSN	6
5	136	151	DPFLGVYHKNNNSWME	16
6	171	186	PFLMDLEGKQGNFKNL	16
7	208	215	NLVRYLPQ	8
8	235	259	FQTLLALHRSYLTPGDSSSGWTAGA	25
9	291	291	D	1
10	293	293	L	1
11	303	320	FTVEKGIYQTSNFRVQPT	18
12	325	355	RFPNITNLCPFGEVFNATRFASVYAWNRKRI	31
13	363	389	SVLYNSASFSTFKCYGVSPTKLNDLCF	27
14	402	424	DEVRQIAPGQTGKIADYNYKLPD	23
15	436	447	NNLDSKVGGNYN	12
16	453	485	FRKSNLKPFERDISTEIYQAGSTPCNGVKGFNC	33
17	494	499	FQPTNG	6
18	513	533	ELLHAPATVCGPKKSTNLVKN	21
19	552	559	SNKKFLPF	8
20	578	579	TL	2
21	599	603	TNTSN	5
22	614	641	CTEVPVAIHADQLTPTWRVYSTGSNVFQ	28

23	653	662	VNNSYECDIP	10
24	670	687	SYQTHTNSPRRARSVASQ	18
25	692	706	YTMSLGAENSVAYSN	15
26	745	745	E	1
27	770	776	EQDKNTQ	7
28	783	797	KQIYKTPPIKDFGGF	15
29	804	811	PDPSKPSK	8
30	825	839	LADAGFIKQYGDCLG	15
31	985	989	EAEVQ	5
32	1032	1040	GQSKRVDFC	9
33	1105	1115	NFYEPQIITTD	11
34	1130	1169	VNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGI	40
35	1199	1203	ELGKY	5
36	1249	1264	SCCKFDEDDSEPVLKG	16
37	1266	1266	K	1

6) Graphical and tabular results of predicted peptides for Omicron variant of COVID-19



Predicted peptides:

No.	Start	End	Peptide	Length
1	13	34	SQCVNLITRTQSYTNSFTRGVY	22
2	56	79	FSNVTWFHAIHVSGTNGTKRFDNP	24
3	81	81	L	1
4	93	96	EKSN	4
5	135	152	DPFLDVYYHKNNKSWMES	18
6	173	188	LMDLEGKQGNFKNLRE	16
7	204	219	HTPINLGRDLPQGFSA	16
8	246	257	LTPGDSSSGWTA	12
9	290	293	LDPL	4
10	301	306	KSFTVE	6
11	312	319	TSNFRVQP	8
12	325	360	RFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVA	36
13	363	389	SVLYNSASFSTFKCYGVSPTKLNDLCF	27
14	401	423	GDEVRQIAPGQTGKIADYNYKLP	23
15	437	448	NLDSKVGGNYNY	12
16	452	494	LFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGF	43
17	513	532	ELLHAPATVCGPKKSTNLVK	20
18	552	559	SNKKFLPF	8
19	577	580	QTLE	4
20	599	603	TNTSN	5
21	624	627	DQLT	4
22	639	639	V	1
23	651	663	EYVNNSYECDIPI	13
24	669	688	ASYQTQTKSHRRARSVASQS	20
25	692	706	YTMSLGAENSVAYSN	15
26	745	745	E	1
27	770	776	EQDKNTQ	7
28	783	797	KQIYKTPPIKDFGGF	15
29	804	812	PDPSKPSKR	9
30	825	839	LADAGFIKQYGDCLG	15
31	985	989	EAEVQ	5
32	1032	1040	GQSKRVDFC	9
33	1104	1115	RNFYEPQIITTD	12
34	1130	1169	VNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGI	40
35	1200	1203	LGKY	4
36	1250	1266	CCKFDEDDSEPVLKGVK	17

7) Comparative representation of total epitopes predicted and those similar to the vaccine sequences

SEQUENCE	Total linear epitopes predicted by the tool	Number of epitopes matching the Vaccine epitopes
Novavax Vaccine	35	35
Alpha Variant	35	9
Lambda Variant	36	11
Gamma Variant	39	9
Eta Variant	37	10
Omicron Variant	36	9

8) The predicted epitope of COVID-19 variants that matches the vaccine epitopes.

Novavax predicted epitope number	Alpha variant predicted epitope number
1	1
8	9
14	14
22	22
24	25
26	27
27	28
28	29
32	33

Novavax predicted epitope number	Lambda variant predicted epitope number
1	1
10	11
19	19
20	20
24	25
26	27
28	29
31	32
32	33
34	35
35	36

Novavax predicted epitope number	Gamma variant predicted epitope number
3	4
20	21
24	28
25	29
26	30
28	32
30	35
31	36
32	37

Novavax predicted epitope number	Eta variant predicted epitope number
1	1
17	18
20	21
22	23
24	26
26	28
30	32
32	34
34	36
35	37

Novavax predicted epitope number	Omicron variant predicted epitope number
19	19
20	20
24	26
26	28
27	29
30	32
31	33
32	34
33	35

REFERENCES

1) Yen-Der Li, Wei-Yu Chi, Jun-Han Su , Louise Ferrall , Chien-Fu Hung and T.-C. Wu (2020), Coronavirus vaccine development: from SARS to MERS to CVID-19, Journal of Biomedical Sciences, (2020) 27:104 https://doi.org/10.1186/s12929-020-00695-2.
2) Y. Wu, W. Jing, J. Liu, Q. Ma, J. Yuan, Y. Wang, M. Du, M. Liu, Effects of temperature and humidity on the daily new cases and new deaths of COVID-19 in 166 countries, Science of the Total Environment 729 (2020) 1–7.
3) Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, Ren R, Leung KSM, Lau EHY, Wong JY, Xing X, Xiang N, Wu Y, Li C, Chen Q, Li D, Liu T, Zhao J, Liu M, Tu W, Chen C, Jin L, Yang R, Wang Q, Zhou S, Wang R, Liu H, Luo Y, Liu Y, Shao G, Li H, Tao Z, Yang Y, Deng Z, Liu B, Ma Z, Zhang Y, Shi G, Lam TTY, Wu JT, Gao GF, Cowling BJ, Yang B, Leung GM, Feng Z. Early transmission dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. N Engl J Med. 2020;382(13):1199–207.
4) Gandhi M, Yokoe DS, Havlir DV. Asymptomatic transmission, the Achilles' Heel of current strategies to control covid-19. N Engl J Med. 2020;382(22):2158–60.
5) JE. de la Pena, ˜ R.A. Rascon-Pacheco, ´ E. Gonzalez-Figueroa, ´ J.E. Fern´ andez-Garate, ´ O.S. Medina-Gomez, ´ P. Borja-Bustamante, J.A. Santill´ an-Oropeza, V.H. BorjaAburto, I.de J. Ascencio-Montiel, Hypertension, Diabetes and Obesity, Major Risk Factors for Death in Patients with COVID-19 in Mexico, Arch. Med. Res. 52 (2021) 443–449, https://doi.org/10.1016/j.arcmed.2020.12.002.
6) S. Garg, L. Kim, M. Whitaker, A. O'Halloran, C. Cummings, R. Holstein, M. Prill, S. J. Chai, P.D. Kirley, N.B. Alden, B. Kawasaki, K. Yousey-Hindes, L. Niccolai, E. J. Anderson, K.P. Openo, A.

Weigel, M.L. Monroe, P. Ryan, J. Henderson, S. Kim, K. Como-Sabetti, R. Lynfield, D. Sosin, S. Torres, A. Muse, N.M. Bennett, L. Billing, M. Sutton, N. West, W. Schaffner, H.K. Talbot, C. Aquino, A. George, A. Budd, L. Brammer, G. Langley, A.J. Hall, A. Fry, Hospitalization rates and characteristics of patients hospitalized with laboratory-confirmed coronavirus disease 2019 — COVID-NET, 14 states, march 1–30, MMWR Morb. Mortal. Wkly Rep. 69 (2020) (2020) 458–464, https://doi.org/10.15585/mmwr.mm6915e3.

7) Masters PS. The molecular biology of coronaviruses. Adv Virus Res. 2006;66:193–292.
8) Stadler K, Masignani V, Eickmann M, Becker S, Abrignani S, Klenk HD, Rappuoli R. SARS–beginning to understand a new virus. Nat Rev Micro- biol. 2003;1(3):209–18.
9) Enjuanes L, Zuniga S, Castano-Rodriguez C, Gutierrez-Alvarez J, Canton J, Sola I. Molecular Basis of Coronavirus Virulence and Vaccine Develop- ment. Adv Virus Res. 2016;96:245–86.
10) Goldsmith CS, Miller SE, Martines RB, Bullock HA, Zaki SR. Electron microscopy of SARS-CoV-2: a challenging task. Lancet. 2020;395(10238):e99. https://doi.org/10.1016/S0140-6736.
11) 31188-0 20. Mishra SK, Tripathi T. One year update on the COVID-19 pandemic: where are we now? Acta Trop. 2021;214:105778. https://doi.org/10.1016/j.actatropica.2020.105778.
12) De Wit E, Van Doremalen N, Falzarano D, Munster VJ. SARS and MERS: recent insights into emerging coronaviruses. Nat Rev Microbiol. 2016;14(8):523-534. https://doi.org/10.1038/nrmicro.2016.81.
13) Cui J, Li F, Shi ZL. Origin and evolution of pathogenic coronaviruses. Nat Rev Microbiol. 2019;17(3):181-192. https://doi.org/10. 1038/s41579-018-0118-9.
14) Li F. Structure, function, and evolution of coronavirus spike proteins. Annu Rev Virol. 2016;3:237-261. https://doi.org/10.1146/ annurev-virology-110615-042301
15) Chen Y, Guo Y, Pan Y, Zhao ZJ. Structure analysis of the receptor binding of 2019-nCoV. Biochem Biophys Res Commun. 2020;525(1): 135-140. https://doi.org/10.1016/j.bbrc.2020.02.071
16) Sharma O, Sultan AA, Ding H, Triggle CR. A review of the progress and challenges of developing a vaccine for COVID-19. Front Immunol. 2020;11:585354. https://doi.org/10.3389/fimmu.2020. 585354.
17) Tortorici MA, Veesler D. Structural insights into coronavirus entry. Adv Virus Res. 2019;105:93-116. https://doi.org/10.1016/bs.aivir. 2019.08.002.
18) Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, Si HR, Zhu Y, Li B, Huang CL, et al: A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature 579: 270-273, 2020
19) Kistler K, Huddleston J and Bedford T: Rapid and parallel adaptive mutations in spike S1 drive clade success in SARS-CoV-2. Cell Host Microbe 30: 545-555.e4, 2022.

20) Turnbull CD, Porter BML, Evans SB, Smith O, Lardner R, Hallifax R, Bettinson HV, Talbot NP, Bafadhel M, Rahman NM, et al: Improved COVID-19 outcomes in a large non-invasive respiratory support cohort despite emergence of the alpha variant. BMJ open Respir Res 8: e001044, 2021.

21) Greaney AJ, Loes AN, Crawford KHD, et al. Comprehensive mapping of mutations in the SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal human plasma antibodies. Cell Host Microbe. 2021;29(3):463-476. https://doi.org/10.1016/J.CHOM. 2021.02.003

22) Tu H, Avenarius MR, Kubatko L, et al. Distinct patterns of emergence of SARS-CoV-2 spike variants including N501Y in clinical samples in Columbus Ohio. bioRxiv. https://doi.org/10.1101/2021. 01.12.426407 36) Leung K, Shum MH, Leung GM, et al. Early transmissibility assessment of the N501Y mutant strains of SARS-CoV-2 in the Severe acute respiratory syndrome coronavirus 2 501Y Variant 1 and Variant 2. Eurosurveillance. 26(1):2002106. https://doi.org/10. 2807/1560- 7917.ES.2020.26.1.2002106

23) Conti P, Conti P, Caraffa A, et al. The British variant of the new coronavirus-19 (SARS-CoV-2) should not create a vaccine problem. J Biol Regul Homeost Agents. 2021;35:1-4.

24) Fratev, F. The N501Y and K417N mutations in the spike protein of SARS-CoV-2 alter the interactions with both hACE2 and human derived antibody: a free energy of perturbation study. BioRxiv. https://doi.org/10.1101/2020.12.23.424283.

25) Khan A, Khan T, Ali S, Aftab S, Wang Y, Qiankun W, Khan M, Suleman M, Ali S, Heng W, et al: SARS-CoV-2 new variants: Characteristic features and impact on the efficacy of different vaccines. Biomed Pharmacother 143: 112176. 2021.

26) Wang Z, Schmidt F, Weisblum Y, Muecksch F, Barnes CO, Finkin S, Schaefer-Babajew D, Cipolla M, Gaebler C, Lieberman JA, et al: mRNA vaccine-elicited antibodies to SARS-CoV-2 and circulating variants. Nature 592: 616-622, 2021.

27) Moss DL and Rappaport J: SARS-CoV-2 beta variant substitu- tions alter spike glycoprotein receptor binding domain structure and stability. J Biol Chem 297: 101371, 2021.

28) Lasek-Nesselquist E, Pata J, Schneider E, George KS. A tale of three SARS-CoV-2 variants with independently acquired P681H mutations in New York State. medRxiv. https://doi.org/10.1101/2021. 03.10.21253285.

29) Kim S, Nguyen TT, Taitt AS, Jhun H, Park HY, Kim SH, Kim YG, Song EY, Lee Y, Yum H, et al: SARSCoV-2 omicron mutation is faster than the chase: Multiple muta- tions on spike/ACE2 interaction residues. Immune Netw 21: e38, 2021.

30) Meo SA, Meo AS, Al-Jassir FF and Klonoff DC: Omicron SARS-CoV-2 new variant: Global prevalence and biological and clinical characteristics. Eur Rev Med Pharmacol Sci 25: 8012-8018, 2021.

31) Du L, He Y, Zhou Y, Liu S, Zheng BJ, Jiang S. The spike protein of SARS-CoV–a target for vaccine and therapeutic development. Nat Rev Microbiol. 2009;7(3):226–36.

32) Cao Z, Liu L, Du L, Zhang C, Jiang S, Li T, He Y. Potent and persistent antibody responses against the receptor-binding domain of SARS-CoV spike protein in recovered patients. Virol J. 2010;7:299.

33) Zhong X, Yang H, Guo ZF, Sin WY, Chen W, Xu J, Fu L, Wu J, Mak CK, Cheng CS, Yang Y, Cao S, Wong TY, Lai ST, Xie Y, Guo Z. B-cell responses in patients who have recovered from severe acute respiratory syn- drome target a dominant site in the S2 domain of the surface spike glycoprotein. J Virol. 2005;79(6):3401–8.

34) He Y, Zhou Y, Siddiqui P, Niu J, Jiang S. Identifcation of immunodomi- nant epitopes on the membrane protein of the severe acute respiratory syndrome-associated coronavirus. J Clin Microbiol. 2005;43(8):3718–26.

35) Kimbrell DA, Beutler B (2001) The evolution and genetics of innate immunity. Nat Rev Genet 2:256–267.

36) Rajat K. De and Namrata Tomar (eds.), (2014), Immunoinformatics, Methods in Molecular Biology, vol. 1184, DOI 10.1007/978-1-4939-1115-8_3.

37) W. E. Paul,(2012) Fundamental Immunology, Lippincott Williams & Wilkins.

38) Thomas K, Goldsby J, Osborne RA, Barbara A, Kuby J (2006) Kuby immunology, 6th edn. Freeman and Co., WH.

39) Gardy JL, Lynn DJ, Brinkman FSL, Rew H (2009) Enabling a systems biology approach to immunology: focus on innate immunity. Trends Immunol 30:249–262.

40) Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT, Veesler D. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. Cell. 2020;181:281–92. doi:10.1016/j.cell.2020.02.058.

41) Hassan AO, Kafai NM, Dmitriev IP, Fox JM, Smith BK, Harvey IB, Chen RE, Winkler ES, Wessel AW, Case JB, et al. A Single-Dose Intranasal ChAd Vaccine Protects Upper and Lower Respiratory Tracts against SARS-CoV-2. Cell. 2020 Oct 1;183(1):169–184.e13. Epub 2020 Aug 19. PMID: 32931734; PMCID: PMC7437481. doi:10.1016/j.cell.2020.08.026.

42) Agnihothram S, Gopal R, Yount BL, Donaldson EF, Menachery VD, Graham RL, Scobey TD, Gralinski LE, Denison MR, Zambon M, et al. Evaluation of serologic and antigenic relationships between Middle Eastern respiratory syndrome coronavirus and other coronaviruses to

develop vaccine platforms for the rapid response to emerging coronaviruses. J Infect Dis. 2013;209:995–1006. doi:10.1093/infdis/jit609.

43) Lin Q, Zhu L, Ni Z, Meng H, You L. Duration of serum neutralizing antibodies for SARS-CoV-2: lessons from SARS-CoV infection [published online ahead of print, 2020 Mar 25]. J Microbiol Immunol Infect. 2020:S1684-1182(20)30075-X. doi:10.1016/j. jmii.2020.03.015.

44) French J, Deshpande S, Evans W, Obregon R. Key Guidelines in Developing a Pre-Emptive COVID-19 Vaccination Uptake Promotion Strategy. Int J Environ Res Public Health. 2020;17 (16)):E5893. doi:10.3390/ijerph17165893.

45) Urrutia-Baca VH, Gomez-Flores R, De La Garza-ramos MA, TamezGuerra P, Lucio-Sauceda DG, Rodríguez-Padilla MC. Immunoinformatics approach to design a novel epitope-based oral vaccine against Helicobacter pylori. J Comput Biol. 2019;(2019 (May):23. doi:10.1089/cmb.2019.0062

46) Ali SA, Almofti YA, Abd-Elrahman KA. Immunoinformatics approach for multiepitopes vaccine prediction against glycoprotein B of avian infectious laryngotracheitis virus. Adv Bioinformatics. 2019;2019(1270485):1–23. doi:10.1155/2019/1270485.

47) . Kaliamurthi S, Selvaraj G, Kaushik AC, Gu KR, Wei DQ. Designing of CD8+ and CD8+ - overlapped CD4+ epitope vaccine by targeting late and early proteins of human papillomavirus. Biologics. 2018;12:107–125. doi:10.2147/BTT.S177901.

48) M. H. Van Regenmortel, (2009),“What is a B-cell epitope?,” Methods in Molecular Biology, vol. 524, pp. 3–20.

49) J. Ponomarenko and M. Van Regenmortel, (2009),“B-cell epitope prediction,” in Structural Bioinformatics, pp. 849–879, John Wiley & Sons, Inc.

50) Hopp T, Woods K: Prediction of protein antigenic determinants from amino acid sequence. Proc Natl Acad Sci U S A 1981, 78(6):3824-3828.

51) Klausen,M.S., Anderson,M.V., Jespersen,M.C., Nielsen,M. and Marcatili,P. (2015) LYRA, a webserver for lymphocyte receptor structural modeling. Nucleic Acids Res., 43, W349–W355.

52) Olimpieri,P.P., Chailyan,A., Tramontano,A. and Marcatili,P. (2013) Prediction of site-specific interactions in antibody-antigen complexes: the proABC method and server. Bioinformatics, 29, 2285–2291.

53) Sela-Culang,I., Benhnia,M.R.E.I., Matho,M.H., Kaever,T., Maybeno,M., Schlossman,A., Nimrod,G., Li,S., Xiang,Y., Zajonc,D. et al. (2014) Using a combined computational-experimental approach to predict antibody-specific B cell epitopes. Structure, 22, 646–657.

54) Sircar,A. and Gray,J.J. (2010) SnugDock: paratope structural optimization during antibody-antigen docking compensates for errors in antibody homology models. PLoS Comput. Biol., 6, e1000644.

55) Kringelum,J.V., Lundegaard,C., Lund,O. and Nielsen,M. (2012) Reliable B cell epitope predictions: impacts of method development and improved benchmarking. PLoS Comput. Biol., 8, e1002829.

56) Shirai,H., Prades,C., Vita,R., Marcatili,P., Popovic,B., Xu,J., Overington,J.P., Hirayama,K., Soga,S., Tsunoyama,K. et al. (2014) Antibody informatics for drug discovery. Biochim. Biophys. Acta, 1844, 2002–2015.

57) El-Manzalawy,Y., Dobbs,D. and Honavar,V.G. (2017) In silico prediction of linear B-cell epitopes on proteins. Methods Mol. Biol., 1484, 255–264.

58) Saha S, Raghava G:((2006), Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. Proteins, 65:40-48.

59) Walter G:(1986) Production and use of antibodies against synthetic peptides. J. Immunol. Methods, 88:149-61.

60) Namrata Tomar, Rajat K.De (2014), Immunoinformatics: a brief review, chapter 3.

61) Korber B, LaBute M, Yusim K: (2006), Immunoinformatics comes of age. PLoS Comput. Biol, 2:e71.

62) Flower D: . Immunoinformatics: predicting immunogenicity in silico Quantum distributor, 1 2007.

63) Karplus P, Schulz G: Prediction of chain flexibility in proteins: a tool for the selection of peptide antigen. Naturwiss 1985, 72:21-213.

64) Parker J, Guo HRD: New hydrophilicity scale derived from highperformance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and x-ray-derived accessible sites. Biochemistry 1986, 25:5425-5432.

65) Emini E, Hughes J, Perlow D, Boger J: Induction of hepatitis A virusneutralizing antibody by a virus-specific synthetic peptide. J. Virol. 1985, 55:836-839.

66) Levitt M: A simplified representation of protein conformations for rapid simulation of protein folding. Journal of molecular biology 1976, 104:59.

67) Pellequer J, Westhof E, Van Regenmortel M: Correlation between the location of antigenic sites and the prediction of turns in proteins. Immunol. Lett. 1993, 36:83-99

68) Blythe M, Flower D: Benchmarking B cell epitope prediction: underperformance of existing methods. Protein Science: A Publication of the Protein Society 2005, 14:246

69) Jose L. Sanchez-Trincado, Marta Gomez-Perosanz, and Pedro A. Reche,(2017), Fundamentals and Methods for T- and B-Cell Epitope Prediction, Journal of Immunology Research Volume 2017, Article ID 2680160, 14 pages https://doi.org/10.1155/2017/2680160.

70) M. Levitt, "Nature of the protein universe," Proceedings of the National Academy of Sciences of the United States of America, vol. 106, no. 27, pp. 11079–11084, 2009..

71) Pomes A (2010) Relevant B cell epitopes in allergic disease. Int Arch Allergy Immunol 152:1–11.

72) Y. Wu, W. Jing, J. Liu, Q. Ma, J. Yuan, Y. Wang, M. Du, M. Liu, Effects of temperature and humidity on the daily new cases and new deaths of COVID-19 in 166 countries, Science of the Total Environment 729 (2020) 1–7.

73) J.E. de la Pena, ˜ R.A. Rascon-Pacheco, ´ E. Gonzalez-Figueroa, ´ J.E. Fern´ andez-Garate, ´ O.S. Medina-Gomez, ´ P. Borja-Bustamante, J.A. Santill´ an-Oropeza, V.H. BorjaAburto, I.de J. Ascencio-Montiel, Hypertension, Diabetes and Obesity, Major Risk Factors for Death in Patients with COVID-19 in Mexico, Arch. Med. Res. 52 (2021) 443–449, https://doi.org/10.1016/j.arcmed.2020.12.002.

74) S. Garg, L. Kim, M. Whitaker, A. O'Halloran, C. Cummings, R. Holstein, M. Prill, S. J. Chai, P.D. Kirley, N.B. Alden, B. Kawasaki, K. Yousey-Hindes, L. Niccolai, E. J. Anderson, K.P. Openo, A. Weigel, M.L. Monroe, P. Ryan, J. Henderson, S. Kim, K. Como-Sabetti, R. Lynfield, D. Sosin, S. Torres, A. Muse, N.M. Bennett, L. Billing, M. Sutton, N. West, W. Schaffner, H.K. Talbot, C. Aquino, A. George, A. Budd, L. Brammer, G. Langley, A.J. Hall, A. Fry, Hospitalization rates and characteristics of patients hospitalized with laboratory-confirmed coronavirus disease 2019 — COVID-NET, 14 states, march 1–30, MMWR Morb. Mortal. Wkly Rep. 69 (2020) (2020) 458–464, https://doi.org/10.15585/mmwr.mm6915e3.

75) Goldsmith CS, Miller SE, Martines RB, Bullock HA, Zaki SR. Electron microscopy of SARS-CoV-2: a challenging task. Lancet. 2020;395(10238):e99. https://doi.org/10.1016/S0140-6736

76) 31188-0 20. Mishra SK, Tripathi T. One year update on the COVID-19 pandemic: where are we now? Acta Trop. 2021;214:105778. https://doi.org/10.1016/j.actatropica.2020.105778.

77) Klausen,M.S., Anderson,M.V., Jespersen,M.C., Nielsen,M. and Marcatili,P. (2015) LYRA, a webserver for lymphocyte receptor structural modeling. Nucleic Acids Res., 43, W349–W355.

78) Olimpieri,P.P., Chailyan,A., Tramontano,A. and Marcatili,P. (2013) Prediction of site-specific interactions in antibody-antigen complexes: the proABC method and server. Bioinformatics, 29, 2285–2291.

79) Sela-Culang,I., Benhnia,M.R.E.I., Matho,M.H., Kaever,T., Maybeno,M., Schlossman,A., Nimrod,G., Li,S., Xiang,Y., Zajonc,D. et al. (2014) Using a combined computational-experimental approach to predict antibody-specific B cell epitopes. Structure, 22, 646–657.

80) Sircar,A. and Gray,J.J. (2010) SnugDock: paratope structural optimization during antibody-antigen docking compensates for errors in antibody homology models. PLoS Comput. Biol., 6, e1000644.

81) Kringelum,J.V., Lundegaard,C., Lund,O. and Nielsen,M. (2012) Reliable B cell epitope predictions: impacts of method development and improved benchmarking. PLoS Comput. Biol., 8, e1002829.

82) Shirai,H., Prades,C., Vita,R., Marcatili,P., Popovic,B., Xu,J., Overington,J.P., Hirayama,K., Soga,S., Tsunoyama,K. et al. (2014) Antibody informatics for drug discovery. Biochim. Biophys. Acta, 1844, 2002–2015.

83) El-Manzalawy,Y., Dobbs,D. and Honavar,V.G. (2017) In silico prediction of linear B-cell epitopes on proteins. Methods Mol. Biol., 1484, 255–264.

84) El-Manzalawy,Y. and Honavar,V. (2010) Recent advances in B-cell epitope prediction methods. Immunome Res., 6, S2.

85) Parker J, Guo D, Hodges R: New hydrophilicity scale derived from High-Performance Liquid Chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. Biochemistry 1986, 25:5425-5432.

86) Jens Erik Pontoppidan Larsen, Ole Lund and Morten Nielsen. Improved method for predicting linear B-cell epitopes. Immunome Res. 2006; 2: 2.

LIST OF PUBLICATIONS

Paper ID:GS-ICMPH-DELHI-210422-3433

GLOBAL Society For research And Development

ICMPH 2022 New Delhi | 24th April 2022

International Conference on Medical, Pharmaceutical and Health Sciences

New Delhi, India | 24th April, 2022

Organized by

GLOBAL SOCIETY FOR RESEARCH AND DEVELOPMENT

Certificate

— OF PRESENTATION —

This is to certify that ***Vanshika Dureja*** *has presented a paper entitled "A Comparative Study of the Structural Basis of B-Cell Epitope Prediction Tool (ElliPro & DiscoTope)" at the International Conference on Medical, Pharmaceutical and Health Sciences (ICMPH) held in New Delhi, India on 24th April, 2022.*

GLOBAL SOCIETY FOR RESEARCH AND DEVELOPMENT GSRD

Associated with UGC Care Scopus

D. Das
Conference Coordinator
GSRD
www.gsrd.co | info.gsrd@gmail.com

Chairman
GLOBAL SOCIETY FOR RESEARCH AND DEVELOPMENT

www.gsrd.co

International Conference on Medical, Pharmaceutical and Health Sciences (ICMPH)
21st - 22nd April, 2022 at New Delhi, India

EVENT ACCEPTANCE LETTER

Dear MOHD TAUHEED RAYEEN, VANSHIKA DUREJA, ASMITA DAS

We are happy to inform you that your PAPER has been selected for **ICMPH** on **21st - 22nd April, 2022 at New Delhi, India** after peer review process which will be organized by **GSRD** and in association with **PET** for presentation (oral presentation/ poster presentation) at the Conference. Registered paper/Abstract will get Conference Proceeding having ISBN *(International Standard Book Number)* and certificates of presentation.

Paper Title: *A Comparative Study of the Structural Basis of B-Cell Epitope Prediction Tool (ElliPro & DiscoTope)*

Author's Name: MOHD TAUHEED RAYEEN, VANSHIKA DUREJA, ASMITA DAS

Paper ID: GS-ICMPH-DELHI-210422-3433

Kindly confirm your Registration and Event Participation by following links.

Official Page of event: http://gsrd.co/Conference2022/4/NewDelhi/ICMPH/

For Registration guidelines: http://gsrd.co/Conference2022/4/NewDelhi/ICMPH/registration.php

Register now Online by clicking below(By Using Credit Card/Debt Card/Net Banking)
http://paymentnow.in/
Or Bank Details(For Offline payment)
ACCOUNT NAME : Institute For Technology and Research **ACCOUNT TYPE :** Current Account **BANK NAME :** State Bank of India **ACCOUNT NO :** 32764752761 **IFSC CODE :** SBIN0010927 **SWIFT CODE :** SBININBB270 (For foreign MONEY transfer) **ADDRESS :** Khandagiri,Bhubaneswar,Odisha-751030

* After the payment via credit card or debit card Kindly mail us the transaction details along with the scan copy of the Identity proof of the card holder.
** For any query related to payment you can mail us to- *info.gsrd@gmail.com*

The registration fee received by GSRD is not refundable but transferrable. Registration fee includes charges for conference participation only (Day 01). Arrangements and costs of visa, travelling and accommodation are not the responsibility of our organization; they will be borne by the author himself/herself.

Last date of Registration	***05th Apr 2022*** (Kindly confirm your registration process before this date)

Note: *Kindly send us the details regarding payment and Registration form to the official mail Id of the Event before last date of registration.*

All Selected and registered papers will also be forwarded for publication in any one of the following International Journals after the conference.

List A Journals

- International Journal of Mechanical and Production Engineering(IJMPE)
Impact Factor: 4.022
- International Journal of Electrical, Electronics and Data Communication(IJEEDC)
Impact Factor: 3.46
- International Journal of Advance Computational Engineering and Networking(IJACEN)
Impact Factor: 3.2
- International Journal of Soft Computing And Artificial Intelligence(IJSCAI)
Impact Factor: 1.95
- International Journal of Management and Applied Science (IJMAS)
Impact Factor: 4.1
- International Journal of Advances in Mechanical and Civil Engineering(IJAMCE)
Impact Factor: 3.46
- International Journal of Industrial Electronics and Electrical Engineering(IJIEEE)
Impact Factor: 3.2
- International Journal of Advances in Science, Engineering and Technology(IJASEAT)
Impact Factor: 3.15
- International Journal of Advances in Computer Science and Cloud Computing(IJACSCC)
Impact Factor: 2.05
- International Journal of Advances in Electronics and Computer Science (IJAECS)
Impact Factor: 2.68

List B Journals

- International Academy of Science, Engineering and Technology (IASET) (Scopus Indexed)
- International Journal of Engineering & Technology(Scopus Indexed)
- International Journal of Engineering and Technology Innovation(Scopus Indexed)
- International Journal of Economic Sciences(WEB OF SCIENCE Indexing)
- Archives of Business Research(ABR)

Important Information

The GSRD can submit the extended version registered conferences papers to the above International journals (List-A and List-B) with the written request form the author of the papers.

Cost of Publication in List-A Journal is Free (Only online Publication) for all **GSRD Registered Papers** and authors **need not to pay** if it's got selected for the journal for online publication.

Cost of Publication in List-B Journal is 250 USD to 350 USD per paper which will not be included in the GSRD conference registration fees.

The tentative time period for Publication:

For List-A Journal the duration of publication is 30 to 90 days after the day of the Conference.

For List-B Journal the duration of publication is 30 to 120 days after the day of the conference (may exceed 6 months).

Note: This is the officially letter for your paper Acceptance only. You need to confirm your registration and event participation by paying the registration fees before the last date of Registration.

Important Note: *Above mentioned conference dates will be adjust to single day conference based on the availability of space in the hotel and numbers of participants which will reassured to the participants a week before the conference date. There will be No refund for making room reservations or flight tickets before the final confirmations.*

Thank You!!!

Conference Coordinator,

Global Society for Research and Development-GSRD

Mob: +91-9777755483 (Call/ Whatsapp)

Mail-id- info.gsrd@gmail.com

Official website: www.gsrd.co

turnitin

Similarity Report ID: oid:27535:16468344

PAPER NAME

draft 2 (1).pdf

WORD COUNT	CHARACTER COUNT
7524 Words	**40241 Characters**
PAGE COUNT	FILE SIZE
32 Pages	**1.8MB**
SUBMISSION DATE	REPORT DATE
May 2, 2022 11:21 PM GMT+5:30	**May 2, 2022 11:22 PM GMT+5:30**

9% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

- 5% Internet database
- 6% Publications database
- Crossref database
- Crossref Posted Content database
- 5% Submitted Works database

Excluded from Similarity Report

- Bibliographic material

Summary

turnitin

Similarity Report ID: oid:27535:16468344

9% Overall Similarity

Top sources found in the following databases:

- 5% Internet database
- Crossref database
- 5% Submitted Works database
- 6% Publications database
- Crossref Posted Content database

TOP SOURCES

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	abc7news.com Internet	<1%
2	ncbi.nlm.nih.gov Internet	<1%
3	Ikbel Hadj Hassine. "Covid-19 vaccines and variants of concern: A revi... Crossref	<1%
4	neueve.com Internet	<1%
5	"Immunoinformatics", Springer Science and Business Media LLC, 2020 Crossref	<1%
6	Juan Mosqueda, Diego Josimar Hernández-Silva, Andrea Romero-Mal... Crossref	<1%
7	Mayo Clinic College of Medicine on 2017-01-10 Submitted works	<1%
8	"Vaccine Design", Springer Science and Business Media LLC, 2022 Crossref	<1%

Sources overview

turnitin **Similarity Report ID:** oid:27535:16468344

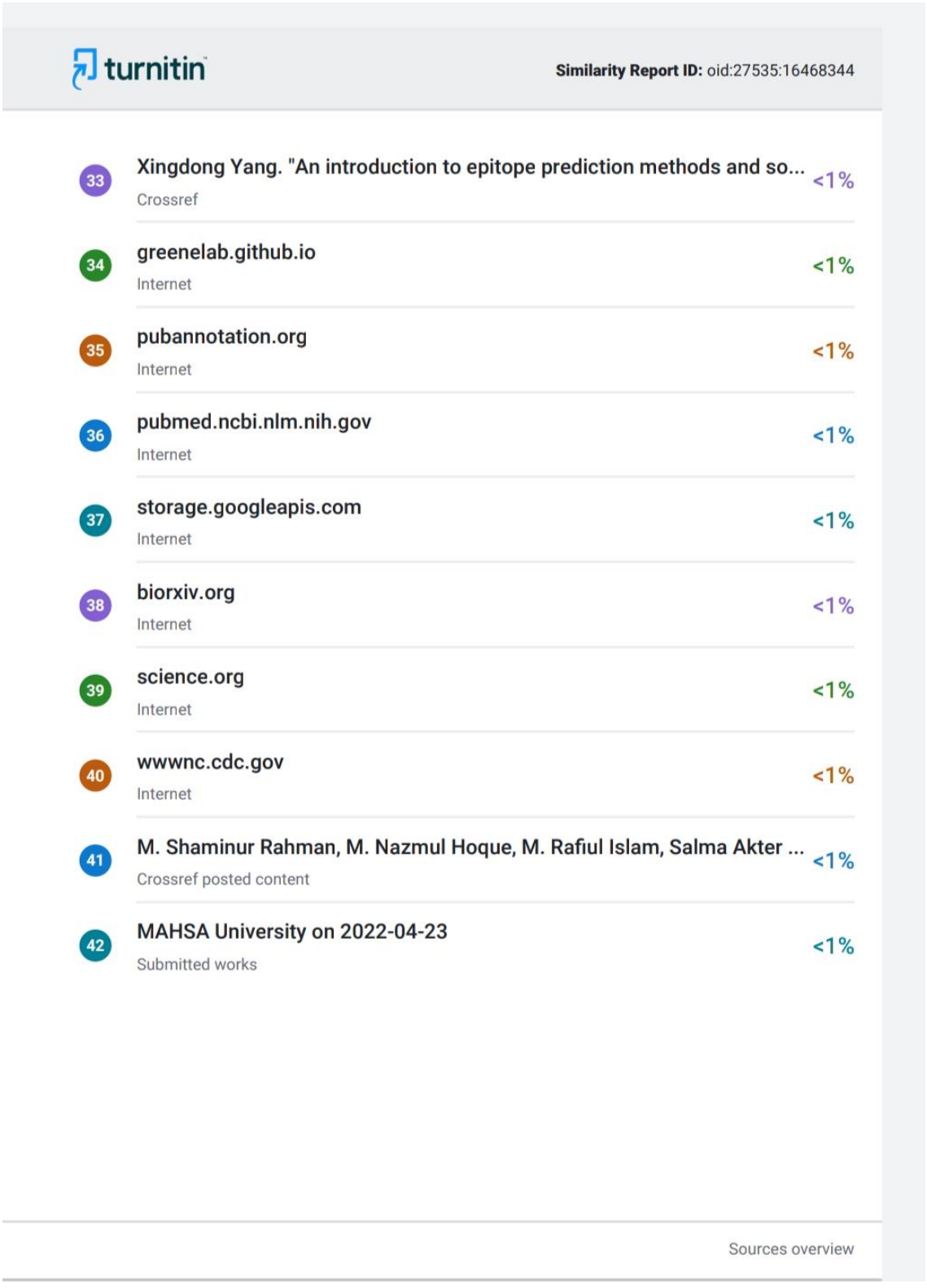
9	mdpi.com Internet	<1%
10	RMIT University on 2018-10-05 Submitted works	<1%
11	nature.com Internet	<1%
12	researchsquare.com Internet	<1%
13	University of Hertfordshire on 2021-01-27 Submitted works	<1%
14	"Immunoinformatics", Springer Science and Business Media LLC, 2014 Crossref	<1%
15	Yi Zhang, Haocheng Zhang, Wenhong Zhang. "SARS-CoV-2 variants, im... Crossref	<1%
16	University of Essex on 2021-04-13 Submitted works	<1%
17	University of Kent at Canterbury on 2017-03-02 Submitted works	<1%
18	hindawi.com Internet	<1%
19	J. A. Tenreiro Machado, J. M. Rocha-Neves, Filipe Azevedo, J. P. Andr... Crossref	<1%
20	parasitesandvectors.biomedcentral.com Internet	<1%

Sources overview

turnitin **Similarity Report ID:** oid:27535:16468344

21	Kosmas A. Galanis, Katerina C. Nastou, Nikos C. Papandreou, Georgios... Crossref posted content	<1%
22	Nottingham Trent University on 2020-09-02 Submitted works	<1%
23	coursehero.com Internet	<1%
24	Enjuanes, L.. "Vaccines to prevent severe acute respiratory syndrome ... Crossref	<1%
25	Vasco Azevedo, Vinicius Abreu, Sintia Almeida, Anderson Santos et al. ... Crossref	<1%
26	iris.paho.org Internet	<1%
27	Manisha Pritam, Garima Singh, Rajnish Kumar, Satarudra Prakash Sing... Crossref	<1%
28	Pei-Hua Wang, Ya-Qian Li, Yuan-Qing Pan, Yan-Yan Guo, Fan Guo, Rui-... Crossref	<1%
29	RMIT University on 2021-08-16 Submitted works	<1%
30	University of Adelaide on 2020-06-04 Submitted works	<1%
31	University of Hertfordshire on 2021-04-30 Submitted works	<1%
32	University of Kent at Canterbury on 2020-08-20 Submitted works	<1%

Sources overview



turnitin

Similarity Report ID: oid:27535:16468344

33	Xingdong Yang. "An introduction to epitope prediction methods and so... Crossref	<1%
34	greenelab.github.io Internet	<1%
35	pubannotation.org Internet	<1%
36	pubmed.ncbi.nlm.nih.gov Internet	<1%
37	storage.googleapis.com Internet	<1%
38	biorxiv.org Internet	<1%
39	science.org Internet	<1%
40	wwwnc.cdc.gov Internet	<1%
41	M. Shaminur Rahman, M. Nazmul Hoque, M. Rafiul Islam, Salma Akter ... Crossref posted content	<1%
42	MAHSA University on 2022-04-23 Submitted works	<1%

Sources overview

DELHI TECHNOLOGIAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

CANDIDATE'S DECLARATION

I, Vanshika Dureja, Roll No.- 2k20/MSCBIO/34 student of M.sc BIOTECHNOLOGY, hereby declare that the project Dissertation titled "name of project" which is submitted by me to the Department of Biotechnology, Delhi Technological University, Delhi in partial fulfilment of the requirement for the award of the degree of Master of Science, is original and not copied from any source without proper citation. This work has not previously formed the basis for the award of any Degree, Diploma Associateship, Fellowship or other similar title or recognition.

Vanshika Dureja

VANSHIKA DUREJA

Place: Delhi

Date: 5, May, 2022

DEPARTMENT OF BIOTECHNOLOGY

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

CERTIFICATE

I hereby certify that the Project Dissertation Titled "**Comparative analysis to determine the efficacy of peptide vaccine against different strains of SARS-CoV-2**" which is submitted by **Vanshika Dureja**, Roll No. – 2k20/MSCBIO/34 (Department of Biotechnology), Delhi Technological University, Delhi in partial fulfilment of the requirement for the award of the degree of Master of Science, is a record of the project work carried out by the students under my supervision. To the best of my knowledge this work has not been submitted in part or full for any Degree or Diploma to this University or elsewhere.

Place: Delhi

Date: 4 May,2022

5/5/22

Dr. Asmita Das
(SUPERVISOR)
Assistant professor
Department of Biotechnology
Delhi Technological University

05/05/2022

Prof. Pravir Kumar
Head of Department
Department of Biotechnology
Delhi Technological University

Head of the Department
Department of Biotechnology
Delhi Technological University
(Formerly College of Engg.)
Bawana Road, Delhi-110042

ACKNOWLEDGEMENT

Firstly, I would like to express my sincere gratitude to my supervisor Dr. Asmita Das for her continuous support during my dissertation, for her patience, motivation and scientific discussions without which this research could not have been completed. I would like to thank Mr. Sunil, for his guidance during entire project, despite the length sof this project he never hesitated to answer my questions and help me throughout my dissertation. I would also like to thank all my lab members Mr Mohd Tauheed Rayeen, Mr Shayon Mahalonobis and Ms Priyanka for their constant support and good time. I would also like to give special thanks to my parents and family as a whole for their continuous support and understanding when undertaking my research and writing my project. Your prayer for me was what sustained me so far. Finally, I would like to thank God, for letting me through all the difficulties. I have experienced your guidance day by day. You are the one who let me finish my degree.

Vanshika Dureja