MASTER OF SCIENCE (BIOTECHNOLOGY) [MOHD TAUHEED RAYEEN] 2022

PREDICTING THE EFFECTIVENESS OF NOVOVAX (NVXCoV2373) AGAINST THE SARS-CoV-2 MUTATIONS: AN IN SILICO ANALYSIS USING LINEAR B-CELL EPITOPE PREDICTION

A DISSERTATION

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF

MASTER OF SCIENCE

IN

BIOTECHNOLOGY

Submitted by:

Mohd Tauheed Rayeen

2K20/MSCBIO/15

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 TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

CANDIDATE'S DECLARATION

I, Mohd Tauheed Rayeen, 2K20/MSCBIO/15 hereby certify that the work which I presented in the Major Project entitled 'Predicting the effectiveness of NOVOVAX (NVXCoV2373) against the SARS-CoV-2 mutations: An In silico study using linear B-cell epitope prediction' in fulfillment of the requirement for the award of the Degree of Masters of Science in Biotechnology and submitted to the Department of Biotechnology, Delhi Technological University, Delhi is an authentic record of my own, carried out during a period from 7-Jan-2022 to 4-May-2022, under the supervision of Dr Asmita Das.

The matter presented in this thesis has not been submitted by me for the award of any other degree of this or any other University. The work has been communicated in a GSRD conference with Scopus indexed proceedings. The details of which are given below:

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

Names of Authors: **Mohd Tauheed Rayeen**, Vanshika Dureja and Asmita Das

Name of the Conference: International Conference on Medical, Pharmaceutical and Health Sciences (ICMPH)

Conference date with Venue: 24 April, 2022, The Suncourt Hotel Yatri, karol bagh, Delhi

Registration for the conference: Completed

Status of the Paper (Accepted/Published/Communicated): Accepted

Date of Paper Communication: 18 March, 2022

Date of Paper Acceptance: 21 March, 2022

MOHD TAUHEED RAYEEN

2K20/MSCBIO/15

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

CERTIFICATE

I hereby certify that the Project dissertation titled **'Predicting the effectiveness of NOVOVAX (NVXCoV2373) against the SARS-CoV-2 mutations: An In silico study using linear B-cell epitope prediction'** which is submitted by **Mohd Tauheed Rayeen**, **2K20/MSCBIO/15**, Department of Biotechnology, Delhi Technological University, Delhi in partial fulfillment of the requirement for the award of the degree of Master of Science, is a record for the project work carried out by the student 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

Dr. Asmita Das

(SUPERVISOR)

Assistant professor

Department of Biotechnology

Delhi Technological University

Prof. Pravir kumar

Head of Department

Department of Biotechnology

Delhi Technological University

ACKNOWLEDGEMENT

It is my privilege to express my profound sense of gratitude and indebtedness to my mentor **Dr. Asmita Das**, Assistant Professor in the Department of Biotechnology, Delhi Technological University for her valuable guidance and consistent encouragement during the progress of the project work. The dissertation wouldn't be completed within a short period without her insightful suggestions and support.

I also take the opportunity to acknowledge the contribution of **Prof. Pravir kumar**, Head of Department of Biotechnology, Delhi Technological University for allowing us to use the department facilities and his full support and assistance during the development of project. I would also not like to miss the opportunity to acknowledge the contribution of all faculty members of the department for their cooperation and assistance during the development of project. I am highly thankful to Mr. Chhail Bihari and Mr. Jitendra Singh for their support.

I am equally grateful and wish to express my wholehearted thanks to respected lab seniors Mr. Sunil, Ms. Ritu Dhankas and Ms. Shweta for their kind support and help in the course of my research work. I would also wish to express my gratitude and affection to my family for their constant love and support.

Mohd Tauheed Rayeen

2K20/MSCBIO/15

ABSTRACT

The current COVID-19 pandemic has resulted in almost 45 lakh cases of infection in people with infected patients in over 200 countries, with an overall death rate of nearly 7%, resulting in partial or complete isolation of various countries. Generally, in predicting epitopes, the primary goal emphasizes the interaction of MHC molecule to thepeptide antigen. For practical reasons, the identification of epitopes in antigens is very interesting, including, understanding the root cause of the disease, immunemonitoring, 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. In this paper we aim to find out the NOVOVAX vaccine efficiency against the different covid strains namely Beta, Delta, Mu, Iota, Kappa. Novovax is a protein based vaccine which is being developed by Novavax (Gaithersburg, USA) and Covovax is the brand name of the same vaccine which is being evaluated in India. Therefore bepipred which a newl tool for estimating linear B-cell epitopes was used as it utilises the sequence as input and predicts the positive epitopes. The bepipred result of all the prediction obtained from the vaccine and from the different strains were computed. The result show's that kappa variants has maximum i.e. 12 similar peptide sequence with the vaccine and therefore NOVOVAX will be more effective against the kappa strain of corona virus, the other 4 different strains have lesser matching petide i.e mu has 10 matching epitopes, delta has 8, iota has 7 similar peptide chains and the beta variant has only 5 similar epitopic peptide sequence. As a result, just one method to prevent the outbreak is to broadly supply safe and effective vaccines for spreading strains over the globe.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER-1 (INTRODUCTION)

GENERAL INTRODUCTION

In general, the immune system is divided into two types: innate and adaptive immune systems. Innate immunity includes non-specific defence mechanisms that either acts right away or hours after the arrival of microorganisms into the body. 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 individuallyidentifying 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 [1]. However, both of these two immunemechanisms work jointly, and the initiation of adaptive immunity depends on the previous stimulation of the non-specific immune response[1]. Acquired immunity is expressed by B and T lymphocytes that recognise the antigen (as they bindto 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 adaptiveresponses by performing various functions such as neutralising toxins and pathogens and labelling these pathogens for destruction or elimination [1].

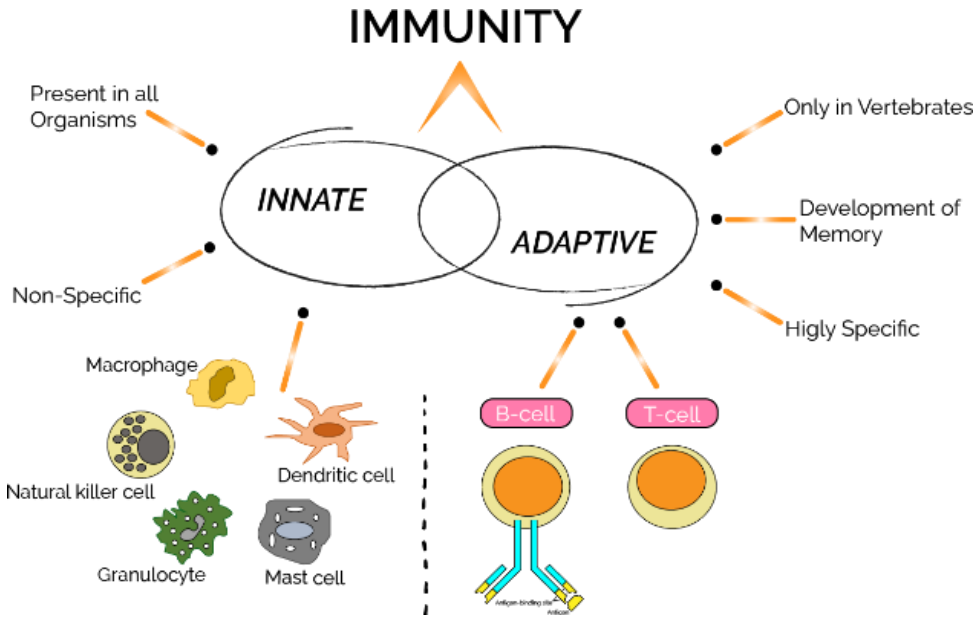


Figure 1- Immunity and its types.

The current pandemic produced by the COVID-19 has caused in almost 45 lakh identified cases in person with infected patients in more than 200 nations, the overall fatality rate of nearly 7%, resulting

in partially or completely confinement of different nations.[28],[29] This extraordinary crisis has wreaked havoc on the global economy, putting strain on medical management across both industrialized and progressing economies.[30] Although quantitative reverse transcription polymerase chain reaction remains the only and best method for early identification of covid-19 ailing people, false negative diagnoses because of poor and inadequate viral genetic material at the site of observation have hampered the use of such molecular-based techniques.[30] Although, in light of the present pandemic, fast diagnostics to diagnose acute infections are critical.

Due to their molecular shape and phylogenetic relationship, they are 7 types classified into four categories: alpha, Beta, gamma, and detla. Coronaviruses have a ss-RNA genetic material of 26–32 kb and are members of the Coronaviridae virus family. RNA viruses have a million-fold much larger alteration rate than their hosts, that is associated to changes in phenotypic plasticity & severity, the majority of which are regarded to be favourable to viral adaptability.[31],[32] The associated mutations connected to the spike protein, which is required for adhesion and viral admission in host by the ACE2 receptor, play a significant role in spike protein docking and are the main reason for its global dissemination.

Generally, in predicting epitopes, the primary goal emphasizes the interaction of MHC molecule to thepeptide 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 method comprises of array-controlled methods, QSAR analysis, recognition of structural binding motifs, threading of proteins, homology modelling and MLtools. In former times, computer technology could only identify the characteristics of the sequence.However, new improved algorithms and techniquesare being developed continuously to boostprediction accuracy.

For practical reasons, the identification of epitopes in antigens is very interesting, including, understanding the root cause of the disease, immunemonitoring, 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 [2][3]. While, polymer MHC, lymphatic proliferation, or ELISPOT assays are utilised for experimental T cell epitope analysis. Traditional tools were entirely based on experimental technology, which isexpensive and time-consuming. As a result epitope forecasting algorithms were devised just to enhances the identification of probable epitope and also reducing the load of associated tests. Here, we will first discuss the tools related to B cells that recognise the antigenin order to better understand the concept of epitope prediction. Following that, we'll go over the most essential forecasting methodologies and tools in detail, giving specific emphasis to their foundations and potential.

CHAPTER -2 (REVIEW OF LITERATURE)

B-CELL EPITOPE PREDICTION

These are antigenic determinants which areseen 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 [4]. These predictionhelps 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 theantigen are the sites for antibody identification [5]. These 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 [6]. Antibodies that are known to recognise linear orcontinuous epitopes have the potential to identify and recognise denatured antigens, while a denaturedantigen result in a loss of conformational B cell recognition.

The study of B cell epitopes can be done in a variety of ways., but they are unsuitable for genomic scale study due to their highcost and cumbersome nature [7]. Therefore, advanced computational tools are being designed for studying or estimating theese epitopes effectively, being quick, economical, and extensible (scalable) [8].

Prediction of linear B cell epitopes: Though these 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 [8],[9].

2.1. Various methods for prediction are asfollows:

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 calculates the amino acid residue score in the specifiedpeptide considering the propensity values and the capacity of the peptide to be a part of a B cell epitope that

can be allocated by the propensity score[10]. For instance, see Pellet J et al., 1993. The ones developed by Parker, Pekkequer, and Emini use hydrophilicityturns, [11],[12] pliability, and solvent avalibility propensity scales, while those developed by Hopp and Woods [13],[14] use only the Levitt hydrophilicity scale. Propensity scale method assumes that hydrophilic areas on peptides surfaces are antigenic[12].

2. **Machine learning methods:**Due to the inefficient conduct of the above system, the machine based methods were developed whereby the ML (Machine learning) [15]algorithm initially translates the epitopes into feature vectors to look at the specified properties as given by propensity scales and then is trained todifferentiate experimental B cell epitopes from the region which do not have any epitope [16]

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.

Table 1 linear B-cell epitopes prediction

Prediction of conformational B cell epitopes: although they constitute greater number, that is more than 90% are conformational in comparison to the linear B cell epitope, its estimation struggles for a few reason

1. It requires prior knowledge of protein 3–Dstructure, which is available only for a fewproteins [17].
2. 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.

2.2. 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 assign a score to the antigen and is an essential toolsin conformational B cell prediction. Whilethe machine learning methods also utilise sequence-based classifiers for a trustable prediction.
2. **Structure-based method:** The structure- based epitope prediction method considersthe three dimensional protein structure[18],and it happens to be the most definite experimental method to identify epitopes. It takes into account the resolution for antigen-antibody complex structure inferred by X-ray crystallography [19],[20]. There are various structure-based B-cell epitope predictors like CEP, Disco Tope, ElliPro, PEPITO, SEPPA, EPITOPIA, or EPIPRED. The CEP(Conformational Epitope Predictor) [21] was the first one, followed by the Disco Tope which was formed by Andersen et al. [22] operating on the protein structure statistics and spatial properties along with the surface accessibility of the amino acid [23]. The structure-based method is used less because of its high cost and the growing complexity of 3-D proteinstructures. 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 [24]. Structure-based epitope prediction is theonly method that predicts discontinuousepitopes.
3. **Mimotope-based prediction**: it is based on a combined approach for mapping epitopes,[25] 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 [26], [27].

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.

Table 2 Non linear B- cell epitope prediction

COVID-19 VARIENTS AND ITS EVOLUTION

The SARS-CoV-2 epidemic has affected nearly every country on the planet; unexpectedly, vaccine development has kept pace with the escalating number of patients.[33],[34] Many vaccines have been approved for use, and governments around the world are immunising their citizens as quickly as possible, knowing that this is the only way to stop the COVID-19 epidemic. Pfizer-BioNTech vaccines and Moderna created most of these authorised vaccination, which are generally spiking protein or mRNA-based. Other adenovirus vector-based vaccines are being researched by Oxford AstraZeneca, Cansino, Johnson & Johnson, and others[34],[35].[36]. On the global endeavour on SAIDP, the WGS has provided roughly 36k covid sequences for speedy vaccine making. The sequence supply a definite path for investigators to follow as they tracked the evolution of various lineages across the world.[37] Many research have discovered a relationship between corona virus genetic changes and the participants' immune responsiveness. In this circumstance, do we consider whether the alterations that happened had any effect on vaccination efficacy? Mumps, rubella, measles, rotavirus, sabin oral poliovirus, influenza, hepatitis A, rabies, and yellow fever vaccines, among many others, deliver the entire virus inactivated or live-attenuated, eliciting a polyclonal B-cell reaction against many autoantigens.[38],[39] Due to the variability of host defense and cell-mediated resistance, no vaccine escaping variations for these pathogens have been discovered. As a result, the vaccine may be ineffective to prevent the emergence of new mutant strains.

Figure 2 variants of covid-19

According to reports, the antigen utilised in currently marketed covid vaccines is minimal, including roughly 1270 amino acids (AAs). RBD and NTD are the only antigens that elicit an antibody

response.[40] Drift in the antigens could result in mutations for a variety of reasons, including naturally occurring infection, selective breeding, or the vaccine itself. Antibody neutralisation will be lost as a result of mutations in an antigen's sequence. As a result, mabs against a specific vaccine sequence resulted in a reduction in efficency.[41]

As COVID-19 expanded throughout China, new strains of the virus emerged from throughout the world, notably B.1.1.7 from the U.k. and B.1.351 from South Africa. When it came to population immunity, the South African strain had a high rate of transmission among individuals, which may have facilitated development and dispersion. Modifications in the RBD of S-protein were found in these variations, leading to a high rate of transmission among people.[42] The new virus spread at a 40–70% quicker rate than the first. The viruses in the South African strain has two additional modifications in the S-protein that enabled it to elude mitigating antibodies. The next sequence of mutation in the novel strain P.1lineage was detected in Brazil. In 2021, the UK's NERVTAG released an article reporting the outcome of various prvious B.1.17 outcomes.[42],[43] This UK variety was discovered in the England in october 2020, and the transmission rate was outlined to be extremely high. It spread to tens of other countries due to the high transmission rate, and it is still expanding. The B.1.17 has 17 mutations, eight are in the spike envelope protein. The fundamental issue with all the changes is that the antisera dosage that have been licenced in the United Kingdom are built on spike protein, which only has effectiveness implications. According to NERVTAG, the B.1.1.7 virus could have a higher death rate than the non-mutated virus.[43]

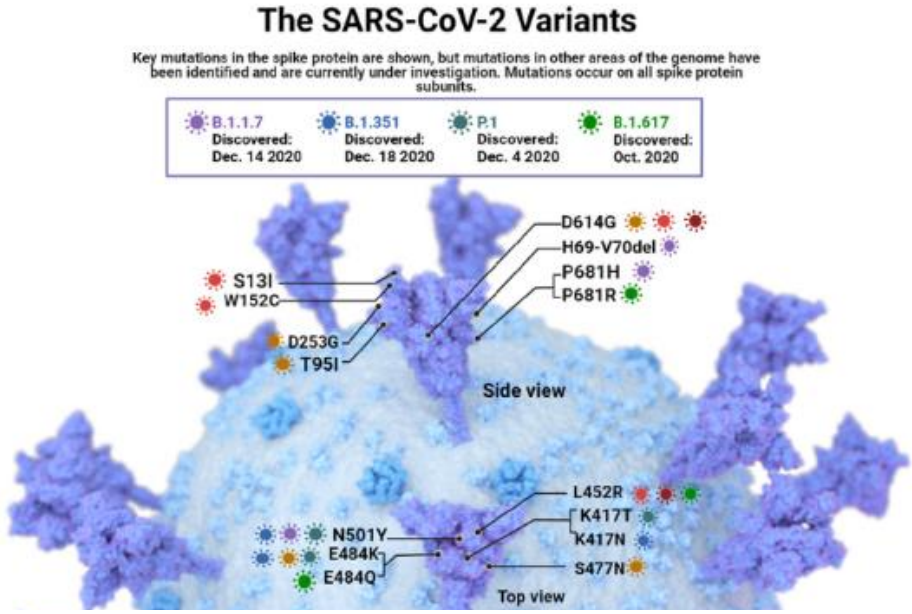


Figure 3 SARS covid variants lineage

The other covid type is P.1, which has been circulating in Brazil since the middle of 2020. Because this strain was so contagious, Brazil became the third in the world country to be affected. The enormous outbreak of infections that hit Manaus has wreaked havoc on the Brazilian healthcare

system.[44] The other strain is B.1.351, which was discovered in Southern Africa in 2020. The SA variation has resulted in a modest decrease in neutralising antibody production. Variations of covid are appearing, and the pathogenicity rate is rising. Knowing how the virus morphs is crucial, to be sure. [45]Creating different vaccine versions in the is silico and further evaluating them for approved vaccines which can offer data on vaccine performance, pathogenicity rate, antibody titers, and a range of many different factors. On either side, the novel formed G2K-UK Virus Consortium has assumed lead in determining in what way numerous known strains will effect the epidemic. It's a joint effort with 11 other United kingdom institutions that will operate alongside its genome sequencing.[42]

The UK has formulated 4 million covid genomic sequences to find the effect of distinct versions on vaccine efficiency, etiology, illness severity, as well as other aspects. The major objective is to look into the effect of polymorphisms on the rate of infection between individuals, vaccine efficacy, and various therapies using a variety of cell-based and animal-based systems in scientific investigations to figure out how variations affect commercially available vaccines.[44]
It's crucial to find out where the virus is altering because it's been mutating since it first arrived in Wuhan, China, and folks are anxiously hoping for its vaccination. Although vaccine evasion is a serious concern, Hibberd points out is that contemporary variations in transmission rates and vaccination receptivity have emerged.[46] The goal of the cooperative investigation on Severe acute respiratory variants with and without modifications will be to determine how each mutation affects viral activity. Our knowledge of vaccinations will be radically altered once observation - based and clinical data on variations and changes are collected.[44],[46]

COVID WAVES

COVID-19 waves are easily understood by imagining sea waves. Every round represents one "wave" of covid-18, in which the rate of infection rises and then falls. COVID-19 was initially detected in China around a year and a half ago. [47]It had spread all around the world in less than a week. The pandemic had a substantial impact on several cultures' economics, resulting in a steep decline in the production of a variety of commodities. Owing to reduced infection counts throughout the summer months, several elements of society felt the pandemic was over, disregarding the early disclosure of the approaching second wave.[48] However, mutation emergence, raising the risk of covid, and the percentage of patients hospitalized to critical care units (ICUs) grew as a consequence.

A. FIRST WAVE

The first wave struck damage on nearly every area of the world, albeit due to periodic shifts; the south part of the globe was impacted later, but when hit was no less brutal. The 2nd wave

happened, and it was significantly more powerful than the 1st, as some analysts estimated. The COVID-19's initial wave impacted a number of countries, but the narrative does not stop there; the second wave could be far more lethal.[49]

B. **SECOND WAVE**

The second wave has a medical definition. "A viral disease that can appear during a pandemic." The virus attacks a specific group of people first. Disease appear to be diminishing. Then illnesses spread to a different segment of the people, leading in a 2nd batch of infections." In April 2020, a Lancet article warned of the possibility of a second wave of the COVID-19 pandemic. Furthermore, " A second-wave pandemic is on the horizon, posing a significant threat to the community in relation to human mortality and a devastating economic crisis," as per a research reported in Nature. The new outbreak of COVID-19 has had a significant impact on India and other countries, and it poses a greater threat to COVID-19 than first wave. India has the most dangerous or severe COVID-19 cases in the world, with 8944 cases as of May 17, 2021.[49] The gradient of the cases then began to decline again. Experts fear that the story does not finish here, and that the third wave, which has arisen with a new mutation, may be more hazardous than the previous waves.[50]

C. **THIRD WAVE**

Since the first week of May, the weekly account of Coronavirus infections in the countries appears to be decreasing. The cases have reduced to 0.26 million per day, down from the altitude of 0.4 million. The infectivity rate of current cases has also decreased, with by overall number falling from 3.7 million to approximately 3.2 million. If the declining pattern remains the same, cases might reach the Feb 2020 level till the starting of august, according to the Indian Express report.[51] Following an increase in the number of cases and deaths in the second batch, some experts believe the third COVID-19 wave will be even more lethal.

D. FOURTH WAVE

South Africa ministry of heath and development , announced on December 3, 2021 that the country is having the fourth outbreak of COVID-19. Between November and December, the number of total cases in South Africa surged by 13 times. Germany also recorded 50k infection recorded per day, the peak of daily cases total ever since this COVID-19 pandemic began. According to the report, the huge proportion of unvaccinated citizens is the cause of COVID-19's 4th wave spike in Germany.[51] In addition, Russia is dealing with the fourth spike of COVID-19, which is causing 40,000 cases per day. Furthermore, the 4th wave is predicted to impact all countries by a large no. of infection. A step is to raise vaccination rates, which could be one rationale for fighting this current Pandemic.[52]

VACCINES

A vaccination is a substance which triggers an immune response against a specific infection-causing virus by revealing the body's adaptive response to specific antigens.[53] Immune responses to antigens, such as cell-mediated and auto immune response, are then developed by the adaptive immune system. These reactions safeguard the person from infection by that virus in the future. Vaccines are normally regarded preventive agents, with vaccination being the greatest option for preventing the spread of a pandemic, especially when no specific treatment drugs are available. Vaccine development, however, entails a wide spectrum of analysis and studies, which can take months or even years for an individual vaccine.[54],[55] Given that COVID-19 is propagating like wildfire over the world, producing the vaccine in the midst of the pandemic was a major challenge. Certain vaccinations have previously been approved to protect mankind from the SARS-CoV-2 virus, based on the concepts of building a vaccine for covid and other infections.[56] WHO recently approved the use of ChAdOx1-S, a viral vector vaccine prepared by AstraZeneca-SKBio (Korea) and Serum Institute Of India, it has shown efficieny of 63 percent.[57] COVAXIN is a vaccine produced by Bharat Biotech in conjunction with the ICMR and the National Institute of Virology in India.[59] This vaccination is administered in a number of countries, including Sri Lanka, Myanmar, Bahrain, and Oman. Aside from the vaccines indicated above, certain vaccines have also been licenced in other countries. Comirnaty (BNT162b2) is a nucleoside-modified m-RNA-based vaccine manufactured by Pfizer/BioNTech in the United States and Germany.[58],[61] The vaccine's effectiveness was assessed to be between 91.2 percent and 100 percent after testing in multiple subgroups. This vaccine has a 6-month shelf life at -90°C to -60°C and is stable for up to 5 days after being removed from the freezer and stored at 2°C to 8°C. National Regulatory Authorities in various countries have granted 19 vaccinations based on EUA against corona virus; though, WHO only approved 6 vaccines out of those 19. Pfizer/BioNTech BNT162b2 vaccine, Moderna's mRNA-1273 vaccine, Johnson & Johnson's Ad26.COV2.S vaccine, AstraZeneca's AZD1222, Sinopharm COVID-19 vaccine, and Sinovac COVID-19 vaccine are among the WHO-recommended vaccinations.[60],[61]

PEPTIDE BASED VACCINES

Peptide subunit vaccines and virus-like particles are two forms of protein-based vaccinations[62],[63]. As an immune-stimulating antigen, peptide vaccination use a viral peptide chain that is particular to the infectious virus or a component of a peptide chain (epitope). [64],[66],[67]Recombinant technology can be used to make these proteins. Adjuvants are required for better immunogenicity in these protein subunit vaccines, and repeated doses must be given.[65],[66] 3 peptide subunit-based vaccination types are in research and observation, while 51

are in clinical trail testing, according to the WHO (7th July 2020). The corona virus spike protein or RBD of spike protein is used in all subunit vaccines in clinical trials.

Virus like particles are multi-peptide complexes that are identical to existing virus particles but lack genetic code, making them incapable of replication.[67] Because VLPs' protein structure is similar to that of the parent virus particle, they can elicit significant immune responses.[69],[72] As a result, when these substances are introduced into the body, they might trigger powerful immunological reactions.[68],[69] These vaccine candidates are safe since they contain no genetic material and hence cannot cause disease; nonetheless, they are laborious to produce. Medicago Inc. is working on a virus like particle vaccine that is currently in clinical trail of phase 1 studies the vaccine uses plant-based VLPs.[70],[71],[73],[74]

NOVAVAX (NVXCoV2373) COVID-19 VACCINE

Novovax is a peptide based vaccine which is being developed by Novavax (Gaithersburg, USA). This protein subunit vaccine has a regime of 2 dosage which is separated by 3 weeks apart. The booster dose can be given at any day after the completion of 3 week or more from the preliminary dosage. The vaccine is given in the United States of America. SARS-CoV-2 rS is another name for this vaccination.[75] Novavax and CEPI developed the Matrix M1 vaccine, which is a combination nanoparticle-associated adjuvant vaccination. Covovax is the brand name under which it is now being evaluated in India. It produces a greater neutralising Ab titer when given with vaccine than corona virus convalescent serum.[76] Many of the competitior ' CD4+ T-cell responses were of the Th1 type. On March 12, 2021, Novavax announced that their vaccine applicant was 96.4 percent effective against the parent strain and 86 percent effective against Lineage B.1.1.7, respectively. It was reported to be 55 percent efficient against the 501.V2 strain in persons without HIV/AIDS. It was also 100 percent effective at preventing serious illness.[75]

The above-mentioned vaccination is both safe and effective in the prevention of COVID-19. All of these vaccinations have received approval in one or more countries. In addition, they are undergoing clinical testing in a number of nations. Vaccination appears being the only way to halt the progress of a global disease. Concerns have been expressed about covid fluctuation after a short length of time. As South Africa has previously witnessed, several other escape variants could arise in the future, leading in a terrible pandemic scenario.[76] As viral propagation rises, the likelihood of Severe acute respiratory mutations keep increasing. As a consequence, there is only one method to terminate the epidemic: widespread distribution of safe and effective vaccines targeting circulating strains. Currently, all strong countries are rushing to immunise their netizens as quickly as possible. They put themselves at risk of developing a new variant that vaccinations might not be able to protect against.[76] It can be a compulsion to develop & circulate new vaccines on a regular basis to tackle

new SARS-CoV-2 variants. To achieve herd immunity, more vaccination coverage is required.

BETA VARIENT

E484K mutations are among the mutations found in the beta version from South Africa, which was discovered in November 2020. In this lineage, there were a sum of 12 variants and one deletion. This variant is more likely to infect young people with comorbidities, and it produces more serious disease than other mutations in same settings. The novel strains from the UK and South Africa seem to be more infectious; even so, genetic variations in the UK versions are unlikely to obstruct the performance of the developed vaccines, whereas mutations in the South African variants (501. V2) may obstruct it to some extent, particularly due to K417N and E484K mutations. In this sense, predicting which of the recently disclosed SARS-CoV-2 strains is more fatal is difficult due to a lack of empirical data. While the hospitalisation of patients is still being investigated, the transmissibility of this variation has indeed been found to raise up to 52 percent. Furthermore, the mortality rate associated with this variant's infection is higher.[77] Because the T cell response elicited by the D614G mutation remains effective against it, the likelihood of reinfection with this mutation has apparently decreased. More crucially, the vaccine's efficacy has been lowered for many people. Furthermore, the E484K mutation causes conformational changes in the flexible loop region of S RBD, which play a vital role in inflammatory processes, viral receptor binding strength, and pathogenicity.[77]

DELTA VARIENT

The delta variant has huge impact in Indian subcontinent and has caused the maximum number of hospitalization and deceased cases in India. The double mutant's official designation is B.1.617, and it has 13 alterations, seven of those are in the spike protein. The L452R and E484Q mutations appeared together for the first time in this strain.[78] The E484Q mutation is notable because it is similar to the E484K mutations previously discovered in SVV B.1.351 and BVV P.1.[79]

The term "triple mutant" refers to a COVID variant created by combining three distinct mutations.[81] The variation was originally discovered in Maharashtra in October, and then in another Indian state. It arose from the COVID double mutant.[79] The spike peptide and alterations E484K and D618G are distinguished by the loss of H146 and Y145. According to the CDC, WHO designated this strain/variant as "VOC" and "VOI."[79],[80]

MU VARIENT

In the midst of the Delta version's summer rise in COVID-19 cases, illnesses, and deaths, several specialists cautioned that another possible danger, the Mu strain of the coronavirus, was beginning to appear. Mu gained attention because of a mutation that let it to defy vaccine immunity, causing a new wave of anxiety among vaccine recipients. But then something unusual happened: the predominance of Mu in circulation in the United States plummeted dramatically in a couple of

weeks. Despite the risk from Mu, the Delta variant's high transmissibility caused it to remain the dominant covid strain, according to some specialists. The Mu variety was initially detected in the United States in March - April, with only a few cases reported. Then, in July, Mu cases began to decline at a similar rate, dropping to 2% on July 9 and 1% on July 22, before dropping to fewer than 0.5 percent in August. By September, the number of Mu cases had dropped to single digits. The most recent came on September 20.

IOTA VARIENT

One of the forms of the virus that causes COVID-19, is the Iota variant, also known as lineage B.1.526. It was initially discovered in November 2020 in New York City. The E484K spike alteration, which can sometimes enable the virus escape antibodies, and the S477N genetic variation, which may assist the virus attach more closely to human cells, have both appeared in the variation. By February 2021, it had quickly expanded throughout the New York area, accounting for almost one out of every four viral sequences. The variation had been discovered in at least fifty U.S. states and 18 nationalities by April 11, 2021. B.1.526 has been classified Iota variant using the World Health Organization's simplified nomenclature approach, and is deemed a variant of interest (VOI), but still not a variant of concern. Until June 2021, approximately 45 thousand cases were detected in the United States. Even by end of July 2021, the Iota variety had lost its dominance in the United States, and the Delta variation had taken its place.

KAPPA VARIENT

SARS-CoV-2, the causative agent of COVID-19, has a variety known as the Kappa variant. It is one of Pango lineage B.1.617's three sublineages. The SARS-CoV-2 Kappa form, also called as lineage B.1.617.1, was discovered in India for the first time in December 2020. The Kappa mutant accounted for more than half of the genomes submitted from India at the end of March 2021. Public Health England categorized it as a Variant Under Investigation on April 1, 2021. The Delta and Kappa variations are actually siblings, direct descendants of a variety known as the double mutant, or B.1.617, in the past. Previously, the primary coronavirus varieties were identified by the names of the nation where they were originally discovered. As a result, they were dubbed "UK variant," "South Africa variation," and "Brazil variant," with the double mutant B.1.617 being dubbed the "India variant." The WHO named these major variants after symbols of the Greek alphabet at the end of May to stop the association with specific countries, which had been leading to name-calling and guilt games.

CHAPTER-3 (MATERIALS & METHODOLOGY)

TOOLS USED

A. BLAST

The Basic Local Alignment Search Tool (BLAST) is a programme that searches for regions of similarity between sequences. The software compares nucleotide or protein sequences to see whether there are any clinically important matches. BLAST can be used to discover genetic family members as well as deduce physiological and phylogenetic relationships between sequences. The BLAST is bascially a measure based on accurate mutation scores. It closely resembles the results that a dynamic programming technique for optimising this metric would produce. The method is more than an order of magnitude more precise than previous heuristic algorithms for detecting weak but biologically important sequence similarities.

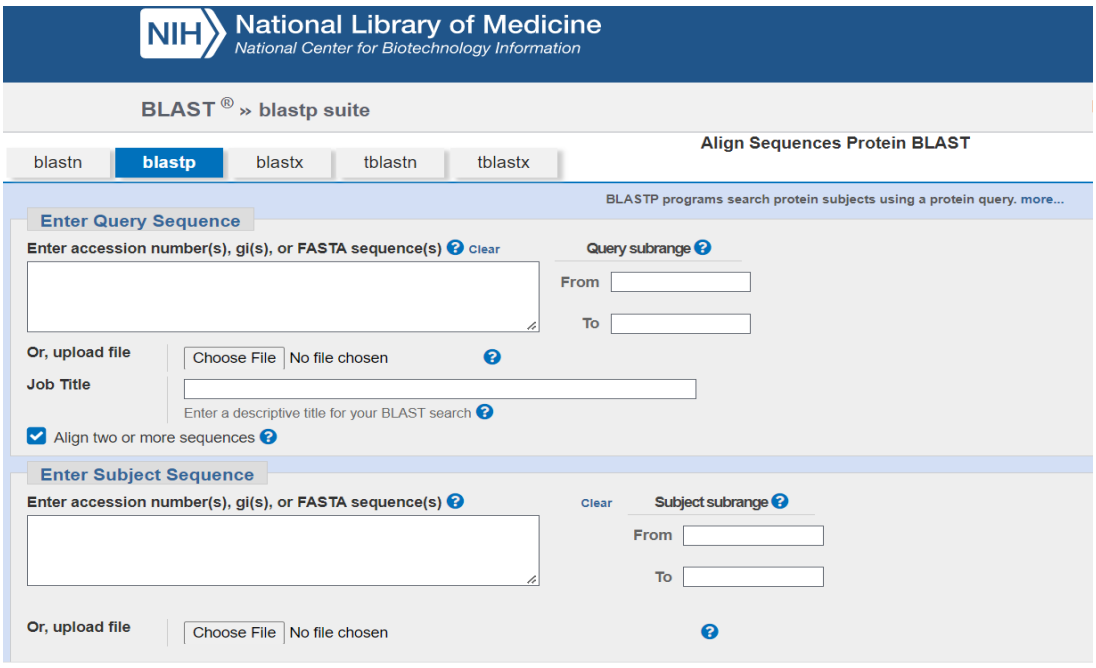


Figure 4 BLAST window

The BLAST method allows for the creation of incredibly rapid database search systems that can also be mathematically analysed. Alternative implementations and variations of the core principle, such as the ones discussed above, can adapt the method to diverse circumstances. BLAST can be a useful tool for molecular biologists, especially when sequence databases grow in size.

Here the reason while doing this study I opted to do the BLAST first was to check that how much similarity is present among the sequences of the vaccine candidate and all the other 5

different varients of COVID 19 surface glycoprotein. The sequence of the vaccine surface glycoprotien was being modified from the original surface glycoprotein to get the desired vaccine sequence of NOVOVAX and rest the 5 different variant namely beta, kappa, delta, mu, iota's surface glycoprotein sequences were retrieved from NIH site.

B. BEPIPRED

The traditional method for linear B cell epitope prediction is the Propensity scale method, where looking at the physio chemical properties, a propensity value to each amino acid is provided. And a running average window was discovered to decrease the changes in the prediction values[82]

For the B cell epitope prediction a variety of structure tools are available but they require prior knowledge of the structure of antigen[83],[84],[85],[86],[88] 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 many clinical and biotechnological implementations such as in antibody construction, vaccine development, efficacy testing and understanding of the immune system.[88],[89],[90].

Therefore, Bepipred was introduced which is a tool for linear B cell epitope prediction. It is a novel method for predicting epitopes and gives out the positive epitopic peptide chains, and is comparatively good than the other random predictions and the variety of propensity scales being tested. This tool is a hybrid method based on the Hidden Markov Model prediction and Parker et al propensity's scale (N). The BepiPred-1.0 server and data sets can be found 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. Residues with scores greater than the threshold (standard set of 0.35) are expected to be epitopes and are coloured yellow on the graph and indicated with a "E" in the output table. The chosen threshold has no effect on the values of the scores.

BepiPred 2.0 employs a random algorithm that was based on epitopes discovered in antibody-antigen protein sequences. On both epitope data collected from solved 3D structures and a big data compilation of linear b-cell epitopes retrieved on the IEDB database, this new method outperformed previously known methods for sequence-based epitope prediction. The method presents results in a user-friendly and instructive manner that is suitable for both computer savvy and non-technical people. We believe that BepiPred-2.0 will be a useful tool

for computational and immuno research.

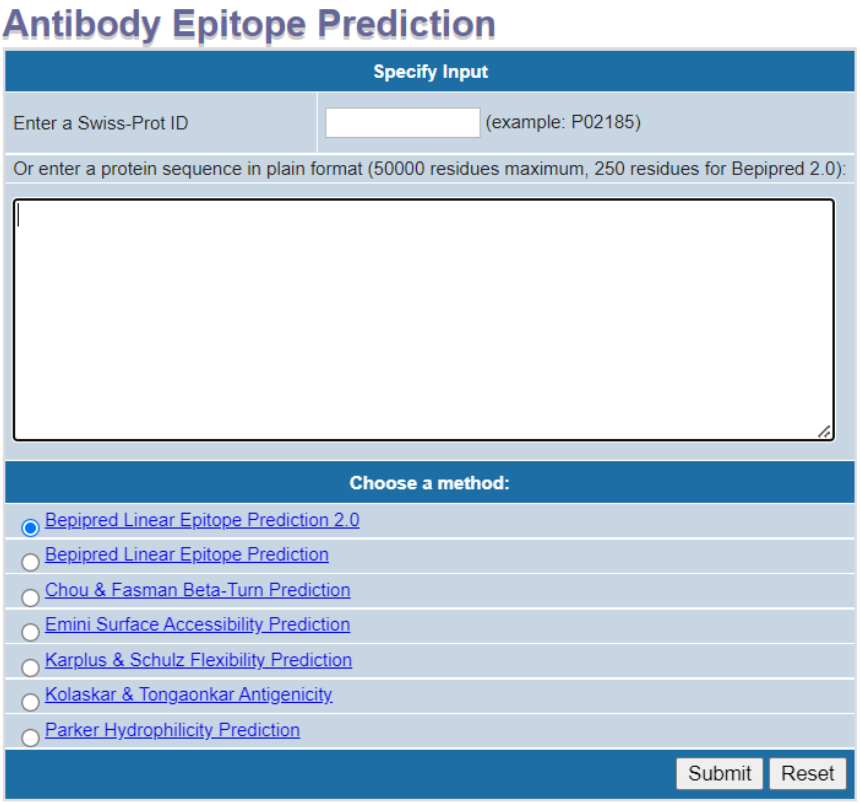


Figure 5 BepiPred Input window

So the idea of using bepipred in our research work is that bepipred gives the result based on the sequence not on the structure and in our work we got only the protein sequences for all the covid varients from the NIH site hence they could easily examined and the possible peptide were match from the vaccine result to all the different varients possible linear epitopic peptides.

METHODOLOGY

1. **Retrieving Data from NCBI (for varients)**

   NCBI virus is a community portal for viral sequence data from RefSeq. GenBank, and other NCBI repositories. It has a quick access to SARS CoY 2 data such as Refseq. Genbank, and Nucleotide and protein sequence information.

Steps to retrieve data in NCBI virus:

1. Opening NCBI virus in Toolbar
2. Selecting novel SARS CoV2 protein sequences from quick access data.
3. 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.

4. Obtained results in the table were looked for the sequence of interest and their accession numbers were selected.
5. Using the specific accession number, surface glycoprotein sequences for different strains of SARS CoV 2 were retrieved.(beta, delta, mu, iota & kappa variants)

2. Retrieving Sequence Data of Novavax Vaccine (NVX-CoV2373)

SARS CoV2 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. BLAST

So we will first do the BLAST of the vaccine sequence vs the different strains of covid18

Step 1: Open the BLAST application.

The user must select the BLAST programme type from the database, such as BLASTp, BLASTn, BLASTx, tBLASTn, tBLASTx. We'll use BLASTp here because we have protein sequences.

Step 2: Type a query sequence into the box (here our query sequence will always be of the NOVOVAX vaccine sequence)

Enter a query sequence for similarity search by entering it into the input box or attaching a FASTA file containing the sequence. For all BLAST programmes, this stage is the same. The accession number, gi number, or maybe even a pure FASTA sequence can be provided by the user.

Step 3: Click on align two or more sequences

This will help to find the similarity between the two input sequences, its match, mismatch & gap penalty and the over all similarity.

Step 4: Type a subject sequence into the box (here the subject sequence will be of the different covid strains.

Enter a subject sequence for similarity search by entering it into the input box or attaching a FASTA file containing the sequence. For all BLAST programmes, this stage is the same. The accession number, gi number, or maybe even a pure FASTA sequence can be provided by the user.

Step 5: Run BLAST.

By clicking the BLAST button at the bottom of the page, you can submit your BLAST programme.

Step 6: Repeat the same step with different strains.

So we need to repeat the same step for getting the BLAST of all the five varients of covid-19 namely Beta, delta, mu, iota, & kappa in the subject sequence the query sequence will remain the same.

Step 7: Final result

The result of the BLAST is then put in a tabular form and will further be analysed with the Bepipred results.

4. BEPIPRED

To predict the possible linear B cell epitopic peptides of the vaccine and the different strains we will use bepipred.

Step 1: Open the IEDB application

Click on the link http://tools.immuneepitope.org/bcell/

Step 2: Enter a protein sequence in plain format and click on the bepipred linear epitope prediction 2.0

Step 3: Add input first from the sequence retrived from the peptide vaccine surface glycoprotein individually

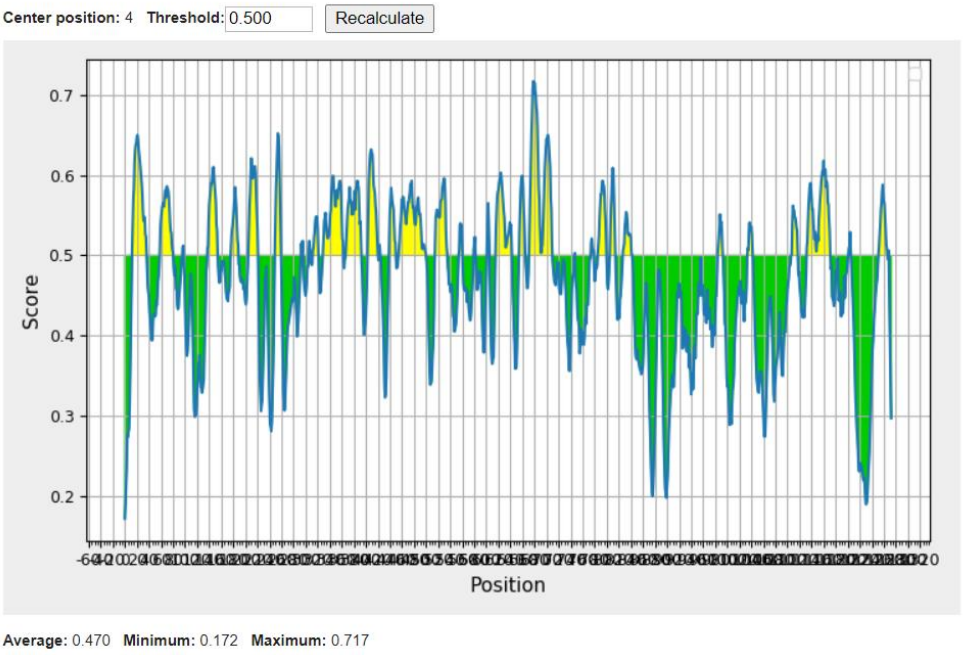


Figure 6 BepiPred result graph

Step 4: then further do the same step with all the covid strains(beta, delta, iota, mu,& kappa)

Step 5: the result of all the prediction obtained from the vaccine and from the different strains were computed against each covid strain

CHAPTER-4 (RESULT)

BLAST RESULT

The result of BLAST between the vaccine and the beta variant of covid was near about 98% , vaccine and the delta variant was about 99% , the vaccine and the mu strain of covid was about 99%, the iota variant vs the vaccine was about 98% and finally the kappa variant vs the novovax vaccine was about 99%.

BLAST RESULT	BETA	DELTA	MU	IOTA	KAPPA
NOVOVAX	**98%**	**99%**	**99%**	**98%**	**99%**

Table 3 BLAST RESULT

BEPIPRED RESULT

BETA VARIENT RESULT

The result of the bepipred for the vaccine vs the beta variant showed the matching of 5 similar peptide chain only with the longest peptide length of 15 amino acid.

S.NO.	VACCINE MATCHING PEPTIDE	BETA VARIENT MATCHING PEPTIDE	LENGTH
1.	QTLE	QTLE	4
2.	VNNSYECDIP	VNNSYECDIP	10
3.	E	E	1
4.	KQIYKTPPIKDFGGF	KQIYKTPPIKDFGGF	15
5.	GQSKRVDFC	GQSKRVDFC	9

Table 4 BETA VARIENT RESULT

DELTA VARIENT RESULT

The delta variant against the NOVOVAX vaccine showed almost 8 similar peptides chains with the longest length of 40 amino acid. These peptides chain are the probable linear B-cell epitopic sites, where the possibility of getting the epitopes is very high. Hence, we have 8 similar peptide chains.

S.NO.	VACCINE MATCHING PEPTIDE	DELTA VARIENT MATCHING PEPTIDE	LENGTH
1	DPL	DLP	3
2	YKL	YKL	3
3	QTLE	QTLE	4
4	TNTSN	TNTSN	5
5	E	E	1
6	KQIYKTPPIKDFGGF	KQIYKTPPIKDFGGF	15
7	RNFYEPOIITTD	RNFYEPOIITTD	12
8.	VNNTVYDPLQPELDSFKEELDKYF KNNHTSPDVDLGDISGI	VNNTVYDPLQPELDSFKEELDKYF KNNHTSPDVDLGDISGI	40

Table 5DELTA VARIENT RESULT

MU VARIENTS RESULT

The MU variant has 10 similar peptide chains with 4 more than 10amino acid long chain of positive epitopic sites.

S.NO.	VACCINE MATCHING PEPTIDE	MU VARIENT MATCHING PEPTIDE	LENGTH
1.	SQCVNLTTRTQLPPAYTNSFTRGV Y	SQCVNLTTRTQLPPAYTNSFTRGV Y	25
2.	DLP	DLP	3
3.	YKL	YKL	3
4.	SNKKFLP	SNKKFLP	
5.	TNTSN	TNTSN	5
6.	E	E	1
7.	DKNTQ	DKNTQ	5
8.	KQIYKTPPIKDFGGF	KQIYKTPPIKDFGGF	15
9.	VNNTVYDPLQPELDSFKEELDKYK NHTSPDVDLGDISGI	VNNTVYDPLQPELDSFKEELDKYK NHTSPDVDLGDISGI	40
10.	SCCKFDEDDSEPVLKG	SCCKFDEDDSEPVLKG	16

IOTA VARIENT RESULT

The IOTA variant has 7 similar peptide chain with the vaccine and has only two peptide which are more than 10 amino acid longer in length.

S.NO.	VACCINE MATCHING PEPTIDE	IOTA VARIENT MATCHING PEPTIDE	LENGTH
1.	TNTSN	TNTSN	5
2.	E	E	1
3.	DKNQT	DKNQT	5
4.	RNFYEPOIITTD	RNFYEPOIITTD	12
5.	VNNTVYDPLQPELDSFKEELDKYF KNHTSPDVDLGDISGI	VNNTVYDPLQPELDSFKEELDKYF KNHTSPDVDLGDISGI	40
6.	LGKY	LGKY	4
7.	K	K	1

Table 6 IOTA VARIENT RESULT

KAPPA VARIENT RESULT

The Kappa variant has the maximum matching peptide chain i.e. it has 12 similar positive epitopic site with the NOVOVAX vaccine, hence the vaccine will be more effective against the strain.

S.NO.	VACCINE MATCHING PEPTIDE	KAPPA VARIENT MATCHING PEPTIDE	LENGTH
1	SQCVNLTTRTQLPPAYTNSFTRG VY	SQCVNLTTRTQLPPAYTNSFTRG VY	25
2	FTVE	FTVE	4
3	YQTSNFRVQP	YQTSNFRVQP	10
4	LYNSASFSTFKCYGVSPTKLNDL CFT	LYNSASFSTFKCYGVSPTKLNDL CFT	26
5	YKL	YKL	3
6	SNKKFLP	SNKKFLP	7
7	QTLE	QTLE	4
8	VNNSYECDIP	VNNSYECDIP	10
9	E	E	1
10	KQIYKTPPIKDFGGF	KQIYKTPPIKDFGGF	15
11	GQSKRVDFC	GQSKRVDFC	9
12	VNNTVYDPLQPELDSFKEELDK YFKNHTSPDVDLGDISGI	VNNTVYDPLQPELDSFKEELDK YFKNHTSPDVDLGDISGI	40

Table 7 KAPPA VARIENT RESULT

CHAPTER-5 (CONCLUSION & DISCUSSION)

We frequently overlook the fact that the Covid-19 is still a pandemic, with billions of people worldwide having still to receive a single dose of vaccination. So there is need to increase the vaccination drive worldwide with save and effective vaccination against the new variant of covid-19. Only 8.1 percent of people in low-income nations have received nearly one dose, whereas in developed countries, about 60–80 percent of the population are fully vaccinated.

In our study we find out that BLAST of the NOVOVAX vaccine when compared to all the different strains of covid-19 were near about similar with Beta and Iota having 98% similarity with the vaccine candidate and Mu, Delta, Kappa having 99% similar sequence. Though there's not much difference in BLAST among the five variants so, we decided to find out the possible linear B-cell epitopes of the vaccine and it was then compared to strains epitopic peptide sequence and the result were tabulated. The result show's that kappa variants has maximum similar peptide sequence with the vaccine and therefore NOVOVAX will be more effective against the kappa strain of corona virus, the next strain having maximum matching peptide sequence is MU. The MU strain has 10 similar epitopic site so NOVOVAX will have good efficacy against this strain too. The delta variant which is the variant of concern in India has 8 similar peptide and has the maximum number of hospitalization in India so NOVOVAX can also be administered to the patient but the efficiency is still under review. The beta variant has 5 similar peptide sequence and the iota variant has 7 similar epitopic site as compared to the vaccine simultaneously. Though blast has similar result but bepipred has different findings as of which we find our probable result.

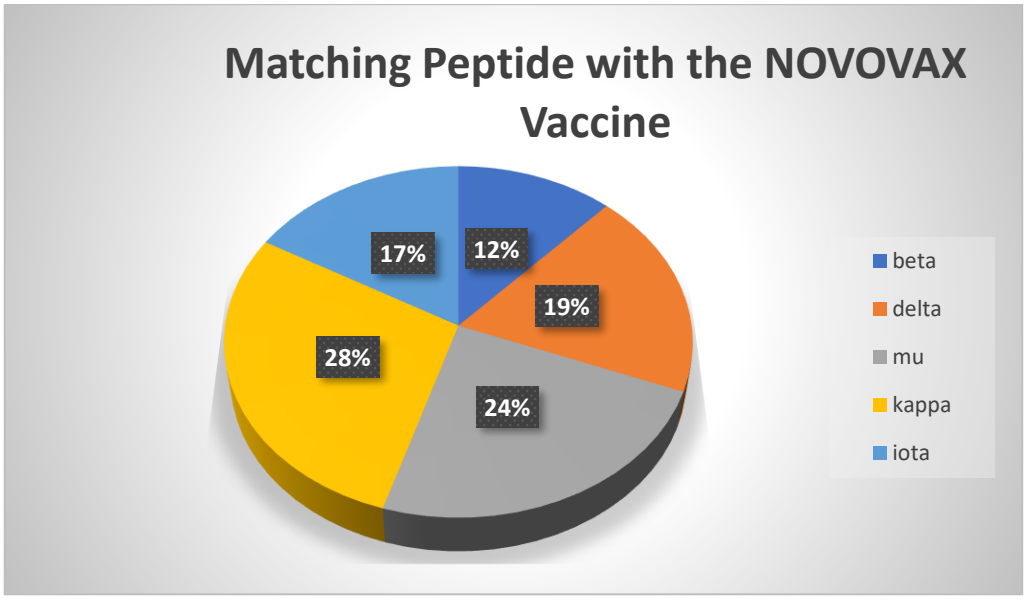


Figure 7 FINAL RESULT OF VACCINE EFFICIENCY

Owing to geographical location the use of NOVOVAX will more be effective in the United Kingdom, United States and the Indian subcontinent as the kappa & delta variant has maximum cases in India, US & UK with cases more than 1 million in these regions. The NOVOVAX is given to the citizen in the United States.

Based on our findings, vaccinations are extremely safe and effective against COVID 19, particularly in terms of reducing hospitalisation and severe disease (eg: death and hospitalization). The Kappa and Delta variants are effective against the NOVOVAX vaccine, although Delta variants are more likely to reinfect or evade the immune response elicited by COVID 19 infection or following immunisation.

REFERENCES

1. W. E. Paul, Fundamental Immunology,Lippincott Williams & Wilkins, 2012.
2. M. H. Van Regenmortel, “What is a B-cellepitope?,” Methods in Molecular Biology, vol. 524, pp. 3–20, 2009.
3. J. Ponomarenko and M. Van Regenmortel,“B-cell epitope prediction,” in Structural Bioinformatics, pp. 849–879, John Wiley & Sons, Inc, 2009.
4. Greenbaum JA, Andersen PH, Blythe M etal (2007) Towards a consensus on datasets and evaluation metrics for developing B- cell epitope prediction tools. J Mol Recognit 20: 75–82.
5. Saha S, Raghava G: Prediction of continuous B-cell epitopes in an antigenusing recurrent neural network. Proteins2006, 65:40-48.
6. Walter G: Production and use of antibodies against synthetic peptides. J.Immunol. Methods 1986, 88:149-61.
7. Namrata Tomar, Rajat K.De (2014),Immunoinformatics: a brief review, chapter 3.
8. Korber B, LaBute M, Yusim K: Immunoinformatics comes of age. PLoSComput. Biol. 2006, 2:e71.
9. Flower D: . Immunoinformatics: predicting immunogenicity in silicoQuantum distributor, 1 2007.
10. Karplus P, Schulz G: Prediction of chainflexibility in proteins: a tool for the selection of peptide antigen. Naturwiss 1985, 72:21-213.
11. 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.
12. Emini E, Hughes J, Perlow D, Boger J: Induction of hepatitis A virusneutralizingantibody by a virus-specific synthetic peptide. J. Virol. 1985, 55:836-839.
13. Levitt M: A simplified representation of protein conformations for rapid simulationof protein folding. Journal of molecular biology 1976, 104:59.
14. 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
15. Blythe M, Flower D: Benchmarking B cellepitope prediction: underperformance of existing methods. Protein Science: A Publication of the Protein Society 2005, 14:246

16. 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.
17. M. Levitt, "Nature of the protein universe," Proceedings of the National Academy of Sciences of the United Statesof America, vol. 106, no. 27, pp. 11079–11084, 2009..
18. Peters B, Sidney J, Bourne P, Bui HH, Buus S, Doh G, Fleri W, Kronenberg M, Kubo R, Lund O, Nemazee D, Ponomarenko JV, Sathiamurthy M, Schoenberger SP, Stewart S, Surko P, Way S, Wilson S, Sette A: The design andimplementation of the immune epitope database and analysis resource. *Immunogenetics* 2005, 57(5):326-336.
19. Mirza O, Henriksen A, Ipsen H, Larsen J, Wissenbach M, Spangfort M, Gajhede M: Dominant epitopes and allergic cross- reactivity: complex formation between a Fab fragment of a monoclonal murine IgGantibody and the major allergen from birch pollen Bet v 1. Journal of immunology (Baltimore, Md.: 1950) 2000,165:331.
20. Fleury D, Daniels R, Skehel J, KnossowM, Bizebard T: Structural evidence for recognition of a single epitope by two distinct antibodies. Proteins 2000, 40(4):572.
21. Kulkarni-Kale U, Bhosle S, Kolaskar A: CEP: a conformational epitope prediction server. Nucleic Acids Res. 2005, 33:W168.
22. Haste Andersen P, Nielsen M, Lund O: Prediction of residues in discontinuous B-cell epitopes using protein 3D structures. Protein Sci. 2006, 15:2558.
23. El-Manzalawy Y, Honavar V. Recent advances in B-cell epitope prediction methods. Immunome Res. 2010 Nov 3;6 Suppl 2(Suppl 2):S2. doi: 10.1186/1745- 7580-6-S2-S2. PMID: 21067544; PMCID:PMC2981878.
24. Yang X, Yu X. An introduction to epitopeprediction methods and software. Rev Med Virol. 2009 Mar;19(2):77-96. doi: 10.1002/rmv.602. PMID: 19101924.
25. Pizzi E, Cortese R, Tramontano A: Mapping epitopes on protein surfaces.

Biopolymers 1995, 36(5):675-680.

26. Halperin I, Wolfson H, Nussinov R: SiteLight: binding-site prediction using phage display libraries. Protein Science: APublication of the Protein Society 2003, 12(7):1344-1359.
27. Moreau V, Granier C, Villard S, Laune D,Molina F: Discontinuous epitope prediction based on mimotope analysis. Bioinformatics 2006, 22(9):1088-1095
28. Cui J, Li F, Shi ZL. Origin and evolution of pathogenic coronaviruses. Nat RevMicrobiol 2019;17:181–92, http://dx.doi.org/10.1038/s41579-018-0118-9.
29. Malik JA, Maqbool M. COVID-19: an overview of current scenario. CELLMED2020;10:21.1–8, http://dx.doi.org/10.5667/CELLMED.2020.0021.
30. Malik JA, Mulla AH, Farooqi T, Pottoo FH, Anwar S, Rengasamy KRR. Targetsand strategies for vaccine development against SARS-CoV-2. Biomed Pharma-cother 2021;137:111254, http://dx.doi.org/10.1016/J.BIOPHA.2021.111254.
31. Weekly epidemiological update on COVID-19 — 4 May 2021. n.d. https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19—4-may-2021. [Accessed 10 May 2021].
32. Coronavirus cases worldwide by country | Statista. n.d. https://www.statista.com/statistics/1043366/novel-coronavirus-2019ncov-cases-worldwide-by-country/. [Accessed 10 May 2021].
33. Fontanet A, Autran B, Lina B, Kieny MP, Karim SSA, Sridhar D. SARS-CoV-2variants and ending the COVID-19 pandemic. Lancet 2021;397:952–4, http://dx.doi.org/10.1016/S0140-6736(21)00370-6.
34. Gu H, Chen Q, Yang G, He L, Fan H, Deng YQ, et al. Adaptation of SARS-CoV-2in BALB/c mice for testing vaccine efficacy. Science 2020;369, http://dx.doi.org/10.1126/science.abc4730.
35. Volz E, Mishra S, Chand M, Barrett JC, Johnson R, Hopkins S, et al. Transmissionof SARS-CoV-2 lineage B.1.1.7 in England: insights from linking epidemiolog-ical and genetic data. MedRxiv 2021, http://dx.doi.org/10.1101/2020.12.30.20249034, 2020.12.30.20249034.
36. Genomic characterisation of an emergent SARS-CoV-2 lineage inManaus: preliminary findings – SARS-CoV-2 coronavirus/nCoV-2019genomic epidemiology – virological. n.d. https://virological.org/t/genomic-

characterisation-of-an-emergent-sars-cov-2-lineage-in-manaus-preliminary-findings/586. [Accessed 16 June 2021].

37. Burki T. Understanding variants of SARS-CoV-2. Lancet (London, England)2021;397:462, http://dx.doi.org/10.1016/S0140-6736(21)00298-1.
38. Williams TC, Burgers WA. SARS-CoV-2 evolution and vaccines: cause forconcern? Lancet Respir Med 2021;9:333–5, http://dx.doi.org/10.1016/S2213-2600(21)00075-8.
39. Korber B, Fischer WM, Gnanakaran S, Yoon H, Theiler J, Abfalterer W, et al.Tracking changes in SARS-CoV-2 spike: evidence that D614G increases infec-tivity of the COVID-19 virus. Cell 2020;182, http://dx.doi.org/10.1016/j.cell.2020.06.043, 812–827.e19.
40. Williams TC, Burgers WA. SARS-CoV-2 evolution and vaccines: cause forconcern? Lancet Respir Med 2021;9:333–5, http://dx.doi.org/10.1016/S2213-2600(21)00075-8.
41. Williams TC, Burgers WA. SARS-CoV-2 evolution and vaccines: cause forconcern? Lancet Respir Med 2021;9:333–5, http://dx.doi.org/10.1016/S2213-2600(21)00075-8.
42. Eguia RT, Crawford KHD, Stevens-Ayers T, Kelnhofer-Millevolte L, GreningerAL, Englund JA, et al. A human coronavirus evolves antigenically to escapeantibody immunity. PLoS Pathog 2021;17, http://dx.doi.org/10.1371/journal.ppat.1009453.
43. Romanò L, Paladini S, Galli C, Raimondo G, Pollicino T, Zanetti AR. Hepatitis Bvaccination: are escape mutant viruses a matter of concern? Hum VaccinesImmunother 2015;11:53–7, http://dx.doi.org/10.4161/hv.34306. Landes Bio-science.
44. Choi B, Choudhary MC, Regan J, Sparks JA, Padera RF, Qiu X, et al. Persistenceand evolution of SARS-CoV-2 in an immunocompromised host. N Engl J Med2020;383:2291–3, http://dx.doi.org/10.1056/nejmc2031364.
45. Mas V, Nair H, Campbell H, Melero JA, Williams TC. Antigenic and sequencevariability of the human respiratory syncytial virus F glycoprotein comparedto related viruses in a comprehensive dataset. Vaccine 2018;36:6660–73,http://dx.doi.org/10.1016/j.vaccine.2018.09.056.
46. Weisblum Y, Schmidt F, Zhang F, DaSilva J, Poston D, Lorenzi JCC, et al. Escapefrom neutralizing antibodies 1 by SARS-CoV-2 spike protein variants.

eLife2020;9:1,

47. Zhang SX, Arroyo Marioli F, Gao R. A second wave? What do people meanby COVID waves? — a working definition of epidemic waves. MedRxiv2021;2021(14):3775–82, 2021.02.21.21252147.
48. Hu J, Li C, Wang S, Li T, Zhang H. Genetic variants are identified toincrease risk of COVID-19 related mortality from UK Biobank data. MedRxiv2020;2020:11–21, http://dx.doi.org/10.1101/2020.11.05.20226761.
49. Xu S, Li Y. Beware of the second wave of COVID-19. Lancet 2020;395:1321–2,http://dx.doi.org/10.1016/S0140-6736(20)30845-X.[130] Cacciapaglia G, Cot C, Sannino F. Second wave COVID-19 pandemics in Europe:a temporal playbook. Sci Rep 2020;10, http://dx.doi.org/10.1038/s41598-020-72611-5.
50. Critical coronavirus (COVID-19) cases most impacted countries worldwide2021 | Statista. n.d. https://www.statista.com/statistics/1109971/serious-critical-coronavirus-2019ncov-cases-worldwide-most-impacted-countries/.[Accessed 20 May 2021].
51. Omicron: South Africa health minister declares 4th wave of COVID-19. n.d.https://www.downtoearth.org.in/news/africa/omicron-south-africa-health-minister-declares-4th-wave-of-covid-19-80556. [Accessed 10 December2021].
52. Unvaccinated people lead Covid 4th wave in Germany: report | WorldNews — Hindustan Times. n.d. https://www.hindustantimes.com/world-news/unvaccinated-people-lead-covid-4th-wave-in-germany-report-101636690003241.html. [Accessed 10 December 2021].
53. A.S. Clem, Fundamentals of vaccine immunology, in: J. Glob. Infect. Dis., Wolters Kluwer -- Medknow Publications, 2011: pp. 73–78. https://doi.org/10.4103/0974-777X.77299.
54. A.S. Monto, Vaccines and antiviral drugs in pandemic preparedness, Emerg. Infect. Dis. 12 (2006) 55–60. https://doi.org/10.3201/eid1201.051068.
55. WHO lists two additional COVID-19 vaccines for emergency use and COVAX rollout, (n.d.).
56. COVAXIN - India‘s First Indigenous Covid-19 Vaccine | Bharat Biotech, (n.d.).

57. comirnaty, (n.d.).
58. COVID-19 Vaccines Advice, (n.d.).
59. WHO | SARS-CoV-2 Variants, WHO. (2021).
60. Accelerating a safe and effective COVID-19 vaccine, (2020).
61. E. Callaway, The race for coronavirus vaccines: a graphical guide, Nature. 580 (2020) 576–577. https://doi.org/10.1038/d41586-020-01221-y.
62. R.M. Bill, Recombinant protein subunit vaccine synthesis in microbes: A role for yeast?, J. Pharm. Pharmacol. 67 (2015) 319–328. https://doi.org/10.1111/jphp.12353.
63. A. Jain, G. Sharma, G. Ghoshal, P. Kesharwani, B. Singh, U.S. Shivhare, O.P. Katare, Lycopene loaded whey protein isolate nanoparticles: An innovative endeavor forenhanced bioavailability of lycopene and anti-cancer activity, Int. J. Pharm. 546 (2018) 97–105. https://doi.org/10.1016/J.IJPHARM.2018.04.061.
64. A. Jain, K. Jain, P. Kesharwani, N.K. Jain, Low density lipoproteins mediated nanoplatforms for cancer targeting, J. Nanoparticle Res. 15 (2013). https://doi.org/10.1007/s11051-013-1888-7.
65. A. Tiwari, P. Kesharwani, V. Gajbhiye, N.K. Jain, Synthesis and characterization of dendro-PLGA nanoconjugate for protein stabilization, Colloids Surfaces B Biointerfaces. 134 (2015). https://doi.org/10.1016/j.colsurfb.2015.06.064.
66. F. Zeeshan, M. Tabbassum, P. Kesharwani, Investigation on Secondary Structure Alterations of Protein Drugs as an Indicator of Their Biological Activity Upon Thermal Exposure, Protein J. 38 (2019) 551–564. https://doi.org/10.1007/s10930-019-09837-4.
67. S.G. Reed, M.T. Orr, C.B. Fox, Key roles of adjuvants in modern vaccines, Nat. Med.19 (2013) 1597–1608. https://doi.org/10.1038/nm.3409.
68. D.J. Marciani, Vaccine adjuvants: Role and mechanisms of action in vaccine immunogenicity, Drug Discov. Today. 8 (2003) 934–943. https://doi.org/10.1016/S1359-6446(03)02864-2.
69. Novavax to Present COVID-19 Vaccine Candidate Progress at 2nd ISV COVID-19 Vaccines Virtual Congress | Novavax Inc. - IR Site, (n.d.).
70. Clover Biopharmaceuticals Initiates Phase 1 Clinical Trial for COVID-19 Vaccine Candidate | Business Wire, (n.d.).

71. China‘s Zhifei Finishes Covid-19 Vaccine‘s Phase II Clinical Trial in Three Months inRace to the Market, (n.d.).
72. B. Gorain, S.K. Bhattamishra, H. Choudhury, U. Nandi, M. Pandey, Overexpressed Receptors and Proteins in Lung Cancer, Nanotechnology-Based Target. Drug Deliv. Syst. Lung Cancer. (2019) 39–75. https://doi.org/10.1016/B978-0-12-815720-6.00003-4.
73. A. Singh, S.K. Maharana, R. Shukla, P. Kesharwani, Nanotherapeutics approaches for targeting alpha synuclien protein in the management of Parkinson disease, Process Biochem. 110 (2021) 181–194. https://doi.org/10.1016/J.PROCBIO.2021.08.008.
74. A. Roldão, M.C.M. Mellado, L.R. Castilho, M.J.T. Carrondo, P.M. Alves, Virus-like particles in vaccine development, Expert Rev. Vaccines. 9 (2010) 1149–1176. https://doi.org/10.1586/erv.10.115.
75. Keech C, Albert G, Cho I, Robertson A, Reed P, Neal S, et al. Phase 1–2 trial ofa SARS-CoV-2 recombinant spike protein nanoparticle vaccine. N Engl J Med2020;383:2320–32, http://dx.doi.org/10.1056/nejmoa2026920.
76. Novavax vaccine 96% effective against original coronavirus, 86% vs Britishvariant in UK trial | Reuters. n.d. https://www.reuters.com/article/uk-health-coronavirus-vaccines-novavax-idUSKBN2B32ZI. [Accessed 17 May 2021]
77. Inside the B.1.1.7 coronavirus variant — The New York Times. n.d. https:// www.nytimes.com/interactive/2021/health/coronavirus-mutations-B117-variant.html. [Accessed 21 May 2021].
78. This “double mutant” variant is adding fuel to India’s COVID-19 cri-sis; n.d. https://www.nationalgeographic.com/science/article/this-double- mutant-variant-is-adding-fuel-to-indias-covid-19-crisis. [Accessed 22 May2021].
79. What is the triple mutant coronavirus variant in India? n.d. https://www.womenshealthmag.com/health/a36232744/what-is-triple-mutant-variant-coronavirus/. [Accessed 13 May 2021].
80. Prakash SJ, Mishra AP, Samal KC. Triple mutant Bengal strain (B.1.618)research of coronavirus and the worst COVID outbreak in India. BioticaResearch Today; 2021. p. 1–5.
81. Cherian S, Potdar V, Jadhav S, Yadav P, Gupta N. Convergent evolution ofSARS-CoV-2 spike mutations, L452R, E484Q and P681R, in the second

waveof COVID-19 in Maharashtra, India; 2021.

82. E) 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.
83. F) 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.
84. G) 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.
85. H) 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.
86. I) 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.
87. J) 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.
88. K) 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.
89. L) El-Manzalawy,Y. and Honavar,V. (2010) Recent advances in B-cell epitope prediction methods. Immunome Res., 6, S2.
90. M) 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.

APPENDIX

1. NOVOVAX VACCINE SEQUENCE & PREDICTED EPITOPES

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQD
LLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGT
TLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVY
SSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLR
DLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYV
GYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRV
QPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSAS
FSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLP
DDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTP
CNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKKST
NLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEI
LDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYS
TGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPQQAQSVASQ
SIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGD
STECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFG
GFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICA
QKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAY
RFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNA
QALNTLVKQLSSNFGAISSVLNDILSRLDPPEAEVQIDRLITGRLQSLQTYVTQ
QLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVF
LHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQII
TTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLG
DISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFI
AGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT

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. BETA VARIENT SEQUENCE & PREDICTED EPITOPES

MFVFLVLLPLVSSQCVNFTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQD
LFLPFFSNVTWFHAIHVSGTNGTKRFANPVLPFNDGVYFASTEKSNIIRGWIFG
TTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRV
YSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINL
VRGLPQGFSALEPLVDLPIGINITRFQTLLALHISYLTPGDSSSGWTAGAAAYY
VGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFR
VQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSA
SFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGNIADYNYKL
PDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGST
PCNGVKGFNCYFPLQSYGFQPTYGVGYQPYRVVVLSFELLHAPATVCGPKKS
TNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTL
EILDITPCSFGGVSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTWRV
YSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPGSASSVAS
QSIIAYTMSLGVENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICG
DSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDF
GGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLIC
AQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMA
YRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQN
AQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTYVT
QQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVV
FLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQ
IITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLG
DISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQGSGYIPEAPRDGQ
AYVRKDGEWVLLSTFLGRSLEVLFQGPGHHHHHHHHSAWSHPQFEKGGGSG
GGGSGGSAWSHPQFEK

No.	Start	End	Peptide	Length
1	16	37	VNFTTRTQLPPAYTNSFTRGVY	22
2	70	76	VSGTNGT	7
3	110	110	L	1
4	138	153	DPFLGVYYHKNNKSWM	16
5	179	181	LEG	3
6	210	222	INLVRGLPQGFSA	13
7	249	260	LTPGDSSSGWTA	12
8	293	296	LDPL	4
9	306	322	FTVEKGIYQTSNFRVQP	17
10	331	354	NITNLCPFGEVFNATRFASVYAWN	24
11	368	395	LYNSASFSTFKCYGVSPTKLNDLCFTNV	28
12	404	419	GDEVRQIAPGQTGNIA	16
13	424	426	KLP	3
14	438	451	SNNLDSKVGGNYNY	14
15	453	488	YRLFRKSNLKPFERDISTEIYQAGSTPCNGVKGFNC	36
16	515	535	FELLHAPATVCGPKKSTNLVK	21
17	557	561	KKFLP	5
18	580	583	QTLE	4
19	603	605	NTS	3
20	617	619	CTE	3
21	625	644	HADQLTPTWRVYSTGSNVFQ	20
22	656	665	VNNSYECDIP	10
23	673	710	SYQTQTNSPGSASSVASQSIIAYTMSLGVENSVAYSNN	38
24	748	748	E	1
25	773	779	EQDKNTQ	7
26	786	800	KQIYKTPPIKDFGGF	15
27	807	814	PDPSKPSK	8
28	828	842	LADAGFIKQYGDCLG	15
29	887	887	T	1
30	987	993	VEAEVQI	7
31	1035	1043	GQSKRVDFC	9
32	1109	1118	FYEPQIITTD	10
33	1133	1179	VNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNI	47
34	1181	1182	KE	2
35	1184	1184	D	1
36	1205	1225	KYEQGSGYIPEAPRDGQAYVR	21
37	1247	1284	GPGHHHHHHHHSAWSHPQFEKGGGSGGGGSGGSAWSHP	38

3. DELTA VARIENT SEQUENCE & PREDICTED EPITOPES

MFVFLVLLPLVSSQCVNLRTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQD
LFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASIEKSNIIRGWIFG
TTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLVYYHKNNKSWMESGVYSS
ANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRD
LPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVG
YLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQ
PTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFS
TFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPD
DFTGCVIAWNSNNLDSKVGGNYNYRYRLFRKSNLKPFERDISTEIYQAGSKPC
NGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKKSTN
LVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEIL
DITPCSFGGVSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTWRVYST
GSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSRRRARSVASQSI
IAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDST
ECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGF
NFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQK
FNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRF
NGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQNVVNQNAQA
LNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQL
IRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLH
VTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITT
DNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDIS
GINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAG
LIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT

No.	Start	End	Peptide	Length
1	14	37	QCVNLRTRTQLPPAYTNSFTRGVY	24
2	70	88	VSGTNGTKRFDNPVLPFND	19
3	96	100	EKSNI	5
4	138	155	DPFLVYYHKNNKSWMESG	18
5	173	186	LMDLEGKQGNFKNL	14
6	207	217	INLVRDLPQGF	11
7	233	236	TRFQ	4
8	243	255	RSYLTPGDSSSGW	13
9	291	293	DPL	3
10	303	320	FTVEKGIYQTSNFRVQPT	18
11	326	353	FPNITNLCPFGEVFNATRFASVYAWNRK	28
12	356	356	S	1
13	360	360	A	1
14	362	390	YSVLYNSASFSTFKCYGVSPTKLNDLCFT	29
15	399	417	IRGDEVRQIAPGQTGKIAD	19
16	420	422	YKL	3
17	438	463	LDSKVGGNYNYRYRLFRKSNLKPFER	26
18	471	499	QAGSKPCNGVEGFNCYFPLQSYGFQPTNG	29
19	513	532	ELLHAPATVCGPKKSTNLVK	20
20	552	559	SNKKFLPF	8
21	577	580	QTLE	4
22	599	603	TNTSN	5
23	614	625	CTEVPVAIHADQ	12
24	633	643	YSTGSNVFQTR	11
25	653	662	VNNSYECDIP	10
26	670	687	SYQTQTNSRRRARSVASQ	18
27	692	707	YTMSLGAENSVAYSNN	16
28	745	745	E	1
29	770	776	EQDKNTQ	7
30	783	797	KQIYKTPPIKDFGGF	15
31	804	811	PDPSKPSK	8
32	825	839	LADAGFIKQYGDCLG	15
33	954	954	Q	1
34	984	990	VEAEVQI	7
35	1031	1040	LGQSKRVDFC	10
36	1104	1115	RNFYEPQIITTD	12
37	1130	1169	VNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGI	40
38	1199	1203	ELGKY	5
39	1249	1266	SCCKFDEDDSEPVLKGVK	18

4. MU VARIENT SEQUENCE & PREDICTED EPITOPES

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQD
LFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASIEKSNIIRGWIFG
TTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVTSNHKNNKSWMESEFR
VYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPIN
LVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAY
YVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNF
RVQPTESIVRFPNITNLCPFGEVFNATKFASVYAWNRKRISNCVADYSVLYNS
ASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYK
LPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGS
TPCNGVKGFNCYFPLQSYGFQPTYGVGYQPYRVVVLSFELLHAPATVCGPKK
STNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQT
LEILDITPCSFGGVSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTWR
VYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSHRRARSV
ASQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYI
CGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIK
DFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLI
CAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQM
AYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQNVVNQ
NAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTYV
TQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGV
VFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEP
QIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDL
GDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLG
FIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHY
T

Predicted peptides:

No.	Start	End	Peptide	Length
1	13	37	SQCVNLTTRTQLPPAYTNSFTRGVY	25
2	55	80	FLPFFSNVTWFHAIHVSGTNGTKRFD	26
3	97	100	KSNI	4
4	138	156	DPFLGVTSNHKNNKSWMES	19
5	179	192	DLEGKQGNFKNLRE	14
6	210	223	PINLVRDLPQGFSA	14
7	250	260	LTPGDSSSGWT	11
8	295	297	DPL	3
9	307	311	FTVEK	5
10	314	322	YQTSNFRVQ	9
11	331	355	PNITNLCPFGEVFNATKFASVYAWN	25
12	369	396	LYNSASFSTFKCYGVSPTKLNDLCFTNV	28
13	403	422	IRGDEVRQIAPGQTGKIADY	20
14	424	426	YKL	3
15	442	451	LDSKVGGNYN	10
16	457	489	FRKSNLKPFERDISTEIYQAGSTPCNGVKGFNC	33
17	498	504	FQPTYGV	7
18	516	536	FELLHAPATVCGPKKSTNLVK	21
19	556	562	SNKKFLP	7
20	582	583	TL	2
21	603	607	TNTSN	5
22	618	645	CTEVPVAIHADQLTPTWRVYSTGSNVFQ	28
23	657	668	VNNSYECDIPIG	12
24	673	690	ASYQTQTNSHRRARSVAS	18
25	697	711	TMSLGAENSVAYSNN	15
26	749	749	E	1
27	774	774	E	1
28	776	780	DKNTQ	5
29	787	801	KQIYKTPPIKDFGGF	15
30	807	815	LPDPSKPSK	9
31	829	843	LADAGFIKQYGDCLG	15
32	958	958	Q	1
33	988	994	VEAEVQI	7
34	1035	1044	LGQSKRVDFC	10
35	1109	1119	NFYEPQIITTD	11
36	1134	1173	VNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGI	40
37	1203	1207	ELGKY	5
38	1253	1268	SCCKFDEDDSEPVLKG	16
39	1270	1270	K	1

5. IOTA VARIENT SEQUENCE & PREDICTED EPITOPES

MFVFFVLLPLVSSQCVNFTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQD
LFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASIEKSNIIRGWIFG
TTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRV
YSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINL
VRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGGSSSGWTAGAAAY
YVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNF
RVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNS
ASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYK
LPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAG
NTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPK
KSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQ
TLEILDITPCSFGGVSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTW
RVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARS
VASQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMY
ICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPI
KDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARD
LICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQ
MAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVN
QNARALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTY
VTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHG
VVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYE
PQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVD
LGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWL
GFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLH
YT

Predicted peptides:

No.	Start	End	Peptide	Length
1	14	34	QCVNFTTRTQLPPAYTNSFTR	21
2	60	62	SNV	3
3	70	79	VSGTNGTKRF	10
4	140	164	FLGVYYHKNNKSWMESEFRVYSSAN	25
5	183	189	QGNFKNL	7
6	210	221	INLVRDLPQGFS	12
7	242	262	LALHRSYLTPGGSSSGWTAGA	21
8	293	296	LDPL	4
9	306	321	FTVEKGIYQTSNFRVQ	16
10	330	354	PNITNLCPFGEVFNATRFASVYAWN	25
11	366	392	SVLYNSASFSTFKCYGVSPTKLNDLCF	27
12	413	419	GQTGKIA	7
13	439	446	NNLDSKVG	8
14	459	466	SNLKPFER	8
15	477	486	NTPCNGVEGF	10
16	494	503	SYGFQPTNGV	10
17	515	535	FELLHAPATVCGPKKSTNLVK	21
18	555	562	SNKKFLPF	8
19	581	582	TL	2
20	602	606	TNTSN	5
21	617	627	CTEVPVAIHAD	11
22	637	643	STGSNVF	7
23	656	667	VNNSYECDIPIG	12
24	672	689	ASYQTQTNSPRRARSVAS	18
25	696	711	TMSLGAENSVAYSNNS	16
26	748	748	E	1
27	775	779	DKNTQ	5
28	785	800	VKQIYKTPPIKDFGGF	16
29	806	816	LPDPSKPSKRS	11
30	828	842	LADAGFIKQYGDCLG	15
31	988	992	EAEVQ	5
32	1035	1044	GQSKRVDFCG	10
33	1107	1118	RNFYEPQIITTD	12
34	1133	1172	VNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGI	40
35	1203	1206	LGKY	4
36	1253	1267	CCKFDEDDSEPVLKG	15
37	1269	1269	K	1

6. KAPPA VARIENT SEQUENCE & PREDICTED EPITOPES

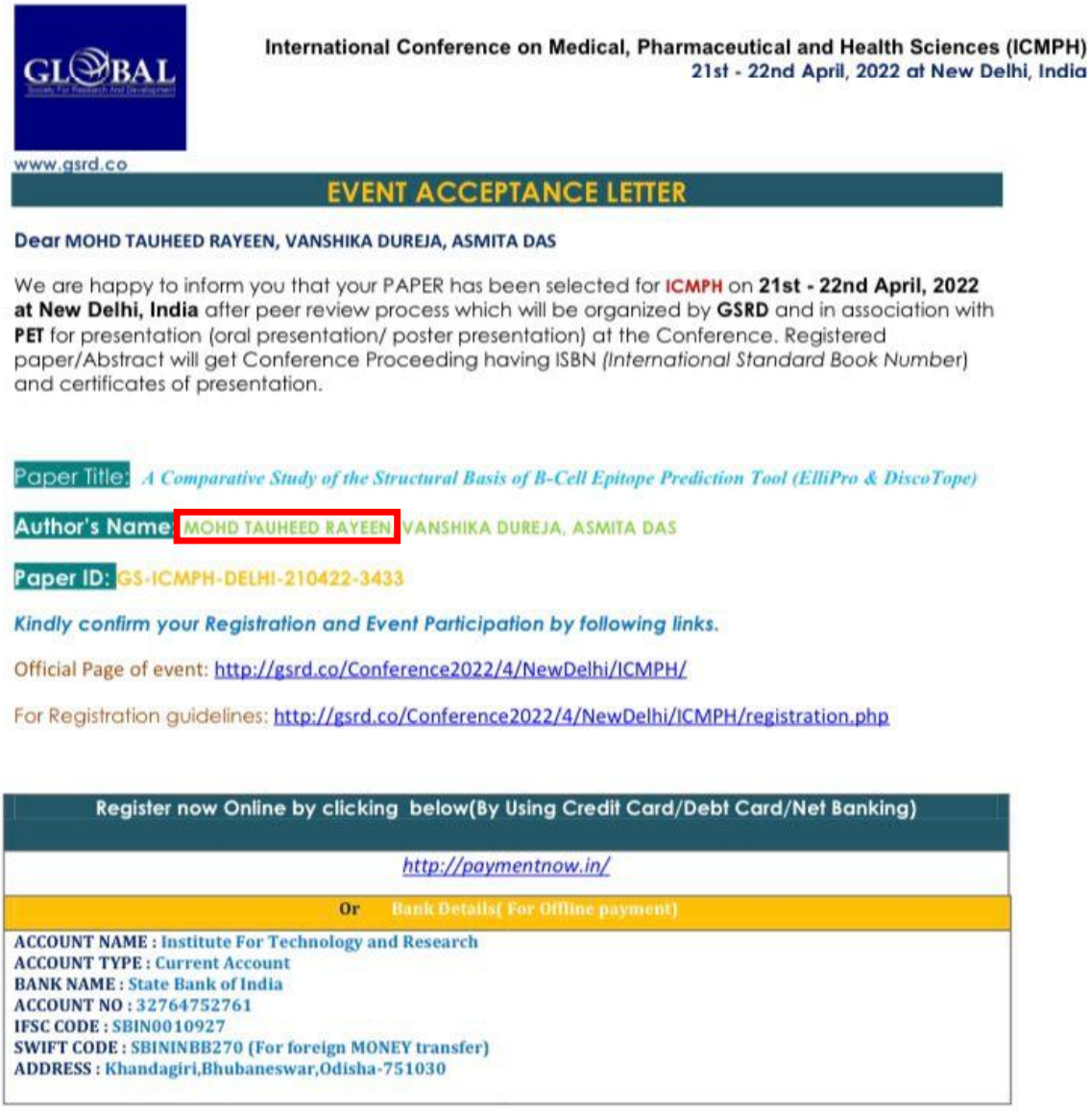
MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQD
LFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASIEKSNIIRGWIFG
TTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLDVYYHKNNKSWMKSEFRV
YSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINL
VRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAY
YVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNF
RVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNS
ASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYK
LPDDFTGCVIAWNSNNLDSKVGGNYNYRYRLFRKSNLKPFERDISTEIYQAGS
TPCNGVQGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKK
STNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQT
LEILDITPCSFGGVSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTWR
VYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSRRRARSV
ASQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYI
CGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIK
DFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLI
CAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQM
AYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQ
NAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTYV
TQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGV
VFLHVTYVPAHEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEP
QIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDL
GDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLG
FIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHY
T

Predicted peptides:

No.	Start	End	Peptide	Length
1	13	37	SQCVNLTTRTQLPPAYTNSFTRGVY	25
2	59	78	FSNVTWFHAIHVSGTNGTKR	20
3	96	100	EKSNI	5
4	140	155	FLDVYYHKNNKSWMKS	16
5	177	191	MDLEGKQGNFKNLRE	15
6	208	222	TPINLVRDLPQGFSA	15
7	251	260	PGDSSSGWTA	10
8	293	296	LDPL	4
9	306	309	FTVE	4
10	313	322	YQTSNFRVQP	10
11	329	355	FPNITNLCPFGEVFNATRFASVYAWNR	27
12	368	393	LYNSASFSTFKCYGVSPTKLNDLCFT	26
13	403	421	RGDEVRQIAPGQTGKIADY	19
14	423	425	YKL	3
15	440	451	NLDSKVGGNYNY	12
16	453	502	YRLFRKSNLKPFERDISTEIYQAGSTPCNGVQGFNCYFPLQSYGFQPTNG	50
17	515	536	FELLHAPATVCGPKKSTNLVKN	22
18	555	561	SNKKFLP	7
19	580	583	QTLE	4
20	603	605	NTS	3
21	617	627	CTEVPVAIHAD	11
22	637	645	STGSNVFQT	9
23	656	665	VNNSYECDIP	10
24	673	690	SYQTQTNSRRRARSVASQ	18
25	695	709	YTMSLGAENSVAYSN	15
26	748	748	E	1
27	773	779	EQDKNTQ	7
28	786	800	KQIYKTPPIKDFGGF	15
29	807	814	PDPSKPSK	8
29	807	814	PDPSKPSK	8
30	828	842	LADAGFIKQYGDCLG	15
31	988	992	EAEVQ	5
32	1035	1043	GQSKRVDFC	9
33	1108	1118	NFYEPQIITTD	11
34	1133	1172	VNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGI	40
35	1202	1206	ELGKY	5
36	1253	1269	CCKFDEDDSEPVLKGVK	17

LIST OF PUBLICATIONS

- Rayeen M. T., Dureja V. and Das A. A Comparative Study of the Structural Basis of B-Cell Epitope Prediction Tool (ElliPro & DiscoTope). Submitted In International Conference on Medical, Pharmaceutical and Health Sciences(21-22 April), New Delhi**.**



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

www.gsrd.co

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

CONFERENCE CERTIFICATE

PAPER NAME

FINAL PLAG CHECK (3).pdf

WORD COUNT

7457 Words

CHARACTER COUNT

40433 Characters

PAGE COUNT

28 Pages

FILE SIZE

1.2MB

SUBMISSION DATE

May 2, 2022 11:55 PM GMT+5:30

REPORT DATE

May 2, 2022 11:56 PM GMT+5:30

8% Overall Similarity

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

- 3% Internet database
- Crossref database
- 4% Submitted Works database
- 5% Publications database
- Crossref Posted Content database

Excluded from Similarity Report

- Bibliographic material

8% Overall Similarity

Top sources found in the following databases:

- 3% Internet database
- Crossref database
- 4% Submitted Works database
- 5% 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	**Far Eastern University on 2022-04-08** Submitted works	1%
2	**Abbas Khan, Taimoor Khan, Shughla Ali, Summiya Aftab et al. "SARS-C...** Crossref	<1%
3	**Ravi Bandaru, Smruti Rekha Rout, Omkar Kamble, Sangram K. Samal et...** Crossref	<1%
4	**ncbi.nlm.nih.gov** Internet	<1%
5	**RMIT University on 2021-08-16** Submitted works	<1%
6	**University of Wolverhampton on 2021-12-17** Submitted works	<1%
7	**University of Adelaide on 2022-02-18** Submitted works	<1%
8	**Yasser EL-Manzalawy. "Recent advances in B-cell epitope prediction m...** Crossref	<1%

9	Jonaid Ahmad Malik, Sakeel Ahmed, Aroosa Mir, Mrunal Shinde, Onur ... Internet	<1%
10	Jonaid Ahmad Malik, Sakeel Ahmed, Aroosa Mir, Mrunal Shinde, Onur ... Crossref	<1%
11	6x8nt0.blogspot.com Internet	<1%
12	"Immunoinformatics", Springer Science and Business Media LLC, 2020 Crossref	<1%
13	Charles Sturt University on 2012-09-19 Submitted works	<1%
14	Nanyang Technological University, Singapore on 2012-12-10 Submitted works	<1%
15	medrxiv.org Internet	<1%
16	economictimes.indiatimes.com Internet	<1%
17	Longbo Hu, Yuhua Xu, Liping Wu, Jin Feng et al. "The E484K Substituti... Crossref	<1%
18	doktori.bibl.u-szeged.hu Internet	<1%
19	International Medical University on 2020-11-20 Submitted works	<1%
20	health.govt.nz Internet	<1%

21	greenelab.github.io Internet	<1%
22	Amrita Vishwa Vidyapeetham on 2021-06-14 Submitted works	<1%
23	University of Wollongong on 2021-05-31 Submitted works	<1%
24	en.wikipedia.org Internet	<1%
25	biorxiv.org Internet	<1%
26	University of Hertfordshire on 2022-04-06 Submitted works	<1%

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

CANDIDATE'S DECLARATION

I, Mohd Tauheed Rayeen, 2K20/MSCBIO/15 hereby certify that the work which I presented in the Major Project entitled 'Predicting the effectiveness of NOVOVAX (NVXCoV2373) against the SARS-CoV-2 mutations: An In silico study using linear B-cell epitope prediction' in fulfillment of the requirement for the award of the Degree of Masters of Science in Biotechnology and submitted to the Department of Biotechnology, Delhi Technological University, Delhi is an authentic record of my own, carried out during a period from 7-Jan-2022 to 4-May-2022, under the supervision of Dr Asmita Das.

The matter presented in this thesis has not been submitted by me for the award of any other degree of this or any other University. The work has been communicated in a GSRD conference with Scopus indexed proceedings. The details of which are given below:

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

Names of Authors: Mohd Tauheed Rayeen, Vanshika Dureja and Asmita Das

Name of the Conference: International Conference on Medical, Pharmaceutical and Health Sciences (ICMPH)

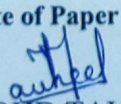
Conference date with Venue: 24 April, 2022, The Suncourt Hotel Yatri, karol bagh, Delhi

Registration for the conference: Completed

Status of the Paper (Accepted/Published/Communicated): Accepted

Date of Paper Communication: 18 March, 2022

Date of Paper Acceptance: 21 March, 2022



MOHD TAUHEED RAYEEN

2K20/MSCBIO/15

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

CERTIFICATE

I hereby certify that the Project dissertation titled **'Predicting the effectiveness of NOVOVAX (NVXCoV2373) against the SARS-CoV-2 mutations: An In silico study using linear B-cell epitope prediction'** which is submitted by **Mohd Tauheed Rayeen, 2K20/MSCBIO/15,** Department of Biotechnology, Delhi Technological University, Delhi in partial fulfillment of the requirement for the award of the degree of Master of Science, is a record for the project work carried out by the student 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

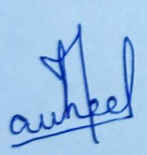
Head of the Department
Department of Biotechnology
Delhi Technological University
(Formerly Delhi College of Engg.)
Bawana Road, Delhi-110042

ACKNOWLEDGEMENT

It is my privilege to express my profound sense of gratitude and indebtedness to my mentor **Dr. Asmita Das**, Assistant Professor in the Department of Biotechnology, Delhi Technological University for her valuable guidance and consistent encouragement during the progress of the project work. The dissertation wouldn't be completed within a short period without her insightful suggestions and support.

I also take the opportunity to acknowledge the contribution of **Prof. Pravir kumar**, Head of Department of Biotechnology, Delhi Technological University for allowing us to use the department facilities and his full support and assistance during the development of project. I would also not like to miss the opportunity to acknowledge the contribution of all faculty members of the department for their cooperation and assistance during the development of project. I am highly thankful to Mr. Chhail Bihari and Mr. Jitendra Singh for their support.

I am equally grateful and wish to express my wholehearted thanks to respected lab seniors Mr. Sunil, Ms. Ritu Dhankas and Ms. Shweta for their kind support and help in the course of my research work. I would also wish to express my gratitude and affection to my family for their constant love and support.



Mohd Tauheed Rayeen

2K20/MSCBIO/15