EPITOPE PREDICTION AS THERAPEUTIC VACCINE FOR POLIO

A DISSERTATION

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE AWARD OF THE DEGREE

OF

MASTER OF TECHNOLOGY

IN

BIOINFORMATICS

Submitted by:

Saumya Verma

2K17/BIO/07

Under the Supervision of

Dr. ASMITA DAS

DEPARTMENT OF BIOTECHNOLOGY

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

CANDIDATE'S DECLARATION

I, SaumyaVerma, Roll No. 2K17/BIO/07 student of M.Tech (Bioinformatics), hereby declare that the project Dissertation titled **"Epitope prediction as therapeutic vaccine for polio"** which is submitted by me to the Department of Biotechnology, Delhi Technological University, Delhi in partial fulfillment of the requirement for the award of the degree of Master of Technology, is original and not copied from any source without proper citation. This work has not previously formed the basis for the award of any Degree, Diploma Associate ship, Fellowship or other similar title or recognition.

Place: Delhi SAUMYA VERMA

Date:

DEPARTMENT OF BIOTECHNOLOGY

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

CERTIFICATE

I hereby certify that the project dissertation titled "**Epitope prediction as therapeutic vaccine for polio**" which is submitted by SaumyaVerma, Roll No. 2K17/BIO/07, Department of Biotechnology, Delhi Technological University, Delhi in partial fulfillment of the requirement for the award of the degree of Master in Technology (Bioinformatics), is a record of a 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 ant Degree or Diploma to this University or elsewhere.

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

Prof. Jai Gopal Sharma
Head of Department
Department of Biotechnology
Delhi Technological University

ABSTRACT

The vaccines of polio are being used to halt poliomyelitis. They are of two types: inactivated poliovirus which is taken by injection and weakened poliovirus which is taken by mouth. To study the mechanism of the reversion of genes in live poliovirus in humans, the study of molecular evolution has to be done. The results came out from the recent vaccine try out that the rates of the substitution for the VP1 are higher for the isolates of sabin which is compared with the wild type. In polio transmission patterns geographical variation is important. Major Histocompatibility Complex class I binding prediction tool and CD8+ immunogenicity prediction tool are being used for the prediction of class I epitopes. Major Histocompatibility Complex class II binding prediction tool and CD4+ immunogenicity prediction tool are being used for the prediction of class II epitopes. Population coverage of the epitopes of MHC I and MHC II binding peptide was preconcieved in Indian population. Many population coverage epitopes were selected for each serotype. Epitopes were selected for VP1, VP2, VP3, VP4. A combinatorial epitope prediction would be used as a therapeutic vaccine for polio.

ACKNOWLEDGEMENT

My sincere thanks to my project guide **Dr.Asmita Das,** Assistant Professor, DTU, Delhi who has given me a chance to work on this project under her guidence. I would like to convey my heartfelt thanks and appreciation to her for providing me guidance and tips during my project work.

I would like to express my deep sense of appreciation to Prof. Jai Gopal Sharma, Head of the Department, Department of Biotechnology, DTU, who has given me a chance to study and work in the reputed institute.

I would also like to thank my seniors, Ph.D scholars, classmates, faculty members and the staff members for their direction which has helped me to achieve positive result in my work.

SAUMYA VERMA

(2K17/BIO/07)

Table of Contents

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Eradication of polio are the success stories in medicine and health of public and teaches us much about how to combat infectious diseases. Poliovirus is used as a model virus because a large body of research data exists on the physical, chemical, and biological properties of the virus, vaccination is available, and its culture is easy as compared to other viruses. The poliomyelitis which is originated from the Greek word, “polio” which means “grey” and “myelon” which means “marrow.” It is an infectious disease which is caused by poliovirus which is a member of the Enterovirus genus, which belongs to the Picornaviridae family.1 Poliomyelitis which is an exclusive disease of human transmitted from a patient or which is free from symptom carrier through the fecal-oral route. Manifestations which are differ ranging from most common form to the most severe forms of debilitating paralysis. Instant medical revolution came in the form of polio vaccines which was introduced in 1950 by Sabin and Stalk. The number of cases in polio keep on decreasing from nearly 58 000 to just 5600 in a year time. Further downfall in number of cases was seen after the second wave of mass immunization. In 1961, only 161 cases were listed. The case which was last detected of paralytic polio by endemic transmission were recorded in 1979 in the midwest United States. In 1988, World Health Organization (WHO) had passed a resolution ,to free the world from wild-type polio by the end of 2000 in the way of Global Polio Eradication Initiative. In 1994, the region in WHO of Americas which were certified polio- free and followed by the Western Pacific Region of WHO in 2000. World Health Organization of European Regions were declared free from polio of all the 3 types of wild poliovirus (types 1, 2, and 3) in June 2002.

Poliovirus is having the diameter of 25 to 30 nm. Its outer coat or capsid is composed of 60 protomers each made of 4 virion proteins VP1, VP2, VP3, and VP4 arranged in icosahedral symmetry. All the 4 virions are made of 8 strands of protein arranged in β sheet array forming a β barrel. Due to the intermingling of various proteins, loops are created, which serve as antigenic sites for combination with corresponding antibodies. Three serotypes of poliovirus have been recognized as types 1, 2, and 3. For type 1 prototype strains will be Brunhilde and Mahoney strains , Lansing and MEFI of type 2, and Leon and Saukett of type 3. The poliovirus enters the oropharynx and multiplies locally in the tonsils, lymph

nodes of the neck, and then subsequently in Peyer patches and small intestine. The incubation period ranges from 2 to 35 days. There is also a hypothesis to suggest that sometimes, virus enters the blood stream and then secondarily invades the tonsils.6 After 3 to 5 days, the virus is shed in stool and also can be recovered from the throat swabs of exposed patients. This period may be entirely asymptomatic or mild viremic symptoms may be seen. Self-limiting episodes of gastroenteritis, respiratory tract infection, and influenza-like illness can occur. The viremia may subside due to the appearance of antibodies or spread to the central nervous system (CNS) via bloodstream.

The work in this project which is trying to show is to stop the reversal in polio. The polio which has been eradicated in the previous years but some viruses remain left and not been eradicated. These viruses start replicating and the polio is back in some areas. So with the help of computational biology by using the various tools a vaccine can be designed which will target that particular target on which the replication of polio takes place. To determine the flowchart of genetic reversion of live poliovirus vaccine in the humans. Here observations are taken from a trail of recent vaccine that the substitution rates of VP1 have been increased for the sabin like isolates in respect to the rate for the wild type. Here the four serotypes of polio virus is taken and the sequences are retrieved from the UNIPROT which is run in the FASTA format. By using the IEDB prediction tool Tepitool which is used for the prediction of MHC I binding peptide and the set of alleles were selected which is based on their population frequencies where we use the most abundant set of alleles of HLA class I which covers almost 99% of the population in the world. For all the alleles immunogenecity prediction score is predicted with the help of immunogenecity prediction tool. This tool uses the amino acid properties as well as their position within the peptide so that it can predict the immunogenecity of a peptide complex. Prediction score is calculated for all the MHC I binding peptide and select only those peptides whose prediction score is more than zero and last but not the least population coverage of selected peptide are taken and will select only those peptides which will cover most of the population. This tool will calculate the related fraction of individuals which is predicted to respond to a given epitope set on the basis of HLA genotypic frequencies and on the basis of binding of HLA. Same process will take place for MHC II also and also for combination of both MHC I and MHC II. In the combination, HLA alleles are selected so that the population coverage of at least >80% is achieved. So different combination of epitopes are tested for the achievement of highest population coverage.

2. Review of literature

2.1 Polio

Poliovirus, the causative agent of polio (also known as poliomyelitis), is a member virus of Enterovirus C, in the family of Picornaviridae.[1]

Poliovirus is composed of an RNA genome and a protein capsid. The genome is a single-stranded positive-sense RNA genome that is about 7500 nucleotides long.[2] The viral particle is about 30 nm in diameter with icosahedral symmetry. Because of its short genome and its simple composition—only RNA and a nonenveloped icosahedral protein coat that encapsulates it, poliovirus is widely regarded as the simplest significant virus.[3]

Poliovirus was first isolated in 1909 by Karl Landsteiner and Erwin Popper.[4] In 1981, the poliovirus genome was published by two different teams of researchers: by Vincent Racaniello and David Baltimore at MIT[5] and by Naomi Kitamura and EckardWimmer at Stony Brook University.[6] Poliovirus is one of the most well-characterized viruses, and has become a useful model system for understanding the biology of RNA viruses. Perhaps the best-characterized model for this link between assembly and cell entry is the influenza A virus hemaglutinin [reviewed in (100, 117)]. In influenza A, the hemaglutinin glycoprotein spike provides the site for binding to its cellular receptor (sialic acid residues on glycoproteins and glycolipids on the cell surface) and serves as the fusion protein. The hemagglutinin is first synthesized as a precursor HA0 and is cleaved late in assembly to form two chains, HA1 and HA2. The newly generated N terminus of HA2 begins with a string of nonpolar amino acids, “the fusion peptide,” that ultimately facilitates fusion of the viral envelope with the cell membrane. When the virus binds to its cellular receptor, it is taken up into endosomes. Upon acidification of the endosomes, the hemagglutinin undergoes a massive conformational change that results in the exposure of the fusion peptide, its insertion into the membrane of the endosome, the fusion of the viral envelope with the endosome, and the release of the nucleocapsid into the cytoplasm. Thus, for influenza the receptor plays a single role, namely to concentrate virus at the surface of susceptible cells, and the trigger that releases the hemagglutinin from its metastable state is acidification of the endosome. In other

enveloped viruses the receptor also serves as the trigger that induces the conformational changes required for entry, in which case the virus envelope may fuse directly with the plasma membrane.

The HA of influenza A is the only protein where both the metastable and fusogenic forms of the glycoprotein have been characterized structurally. However, similar conformational changes have been proposed for the envelope glycoproteins of a number of other viruses based on structures of analogues of the fusogenic form (5, 20, 32, 61, 71, 72, 113, 114, 123). Using poliovirus as an example, we demonstrate that similar mechanisms in which receptor binding releases virus from a metastable state and exposes hydrophobic sequences may also occur in nonenveloped viruses.

2.2 Replication cycle

Poliovirus infects human cells by binding to an immunoglobulin -like receptor, CD155 (also known as the poliovirus receptor or PVR) on the cell surface. Interaction of poliovirus and CD155 facilitates an irreversible conformational change of the viral particle necessary for viral entry. Attached to the host cell membrane, entry of the viral nucleic acid was thought to occur one of two ways: via the formation of a pore in the plasma membrane through which the RNA is then “injected” into the host cell cytoplasm, or that the virus is taken up by endocytosis. Recent experimental evidence supports the latter hypothesis and suggests that poliovirus binds to CD155 and is taken up by endocytosis. Immediately after internalization of the particle, the viral RNA is released.

Poliovirus is a positive-stranded RNA virus. Thus, the genome enclosed within the viral particle can be used as messenger RNAand immediately translated by the host cell. On entry, the virus hijacks the cell's translation machinery, causing inhibition of cellular protein synthesis in favor of virus-specific protein production. Unlike the host cell's mRNAs, the 5’ end of poliovirus RNA is extremely long—over 700 nucleotides—and highly structured. This region of the viral genome is called internal ribosomal entry site (IRES), and it directs translation of the viral RNA. Genetic mutations in this region prevent viral protein production. The first IRES to be discovered was found in poliovirus RNA.

Poliovirus mRNA is translated as one long polypeptide. This polypeptide is then autocleaved by internal proteases into about 10 individual viral proteins. Not all cleavages occur with the same

efficiency. Therefore, the amounts of proteins produced by the polypeptide cleavage vary: for example, smaller amounts of $3D^{pol}$ are produced than those of capsid proteins, VP1–4. These individual viral proteins are:

- $3D^{pol}$, an RNA dependent RNA polymerase whose function is to make multiple copies of the viral RNA genome
- $2A^{pro}$ and $3C^{pro}/3CD^{pro}$, proteases which cleave the viral polypeptide
- VPg (3B), a small protein that binds viral RNA and is necessary for synthesis of viral positive and negative strand RNA
- 2BC, 2B, 2C (an ATPase), *3AB, 3A, 3B* proteins which comprise the protein complex needed for virus replication.
- VP0, which is further cleaved into VP2 and VP4, VP1 and VP3, proteins of the viral capsid

After translation, transcription and genome replication which involve a single process (synthesis of (+) RNA) is realized. For the infecting (+) RNA to be replicated, multiple copies of (−) RNA must be transcribed and then used as templates for (+) RNA synthesis. Replicative intermediates (RIs) which are an association of RNA molecules consisting of a template RNA and several growing RNAs of varying length, are seen in both the replication complexes for (−) RNAs and (+) RNAs. The primer for both (+) and (−) strand synthesis is the small protein VPg, which is uridylylated at the hydroxyl group of a tyrosine residue by the poliovirus RNA polymerase at a cis-acting replication element located in a stem-loop in the virus genome. Some of the (+) RNA molecules are used as templates for further (−) RNA synthesis, some function as mRNA, and some are destined to be the genomes of progeny virions.

In the assembly of new virus particles (i.e. the packaging of progeny genome into a procapsid which can survive outside the host cell), including, respectively:

- Five copies each of VP0, VP3, and VP1 which its N termini and VP4 form interior surface of capsid, assemble into a ‘pentamer’ and 12 pentamers form a procapsid. (The outer surface of capsid is consisting of VP1, VP2, VP3; C termini of VP1 and VP3 form the canyons which around each of the vertices; around this time, the 60 copies of VP0 are cleaved into VP4 and VP2.)
- Each procapsid acquires a copy of the virus genome, with VPg still attached at the 5' end.

Fully assembled poliovirus leaves the confines of its host cell by lysis 4 to 6 hours following initiation of infection in cultured mammalian cells. The mechanism of viral release from the cell is unclear, but each dying cell can release up to 10,000 polio virions.

Drake demonstrated that poliovirus is able to undergo multiplicity reactivation. That is, when polioviruses were irradiated with UV light and allowed to undergo multiple infections of host cells, viable progeny could be formed even at UV doses that inactivated the virus in single infections.

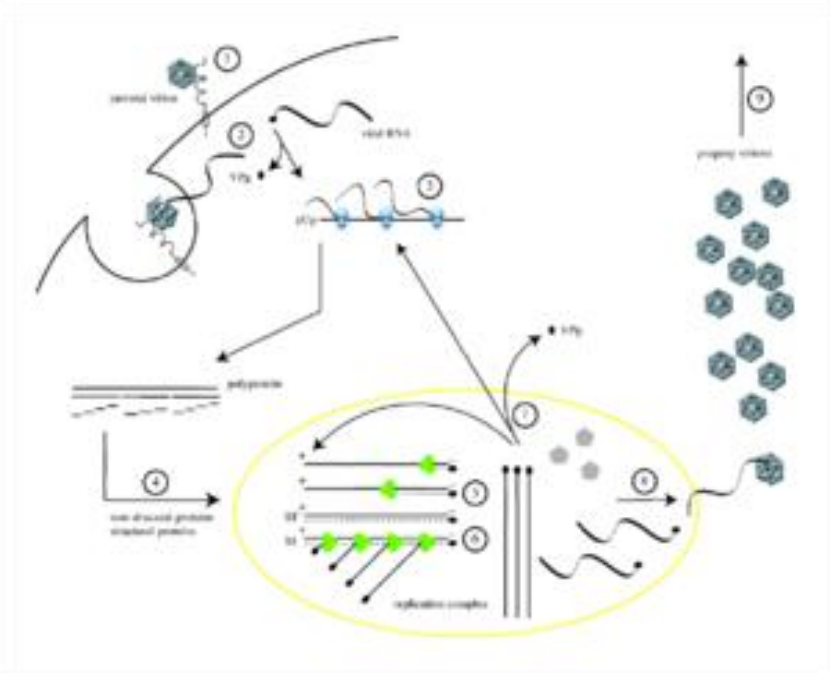


Fig 1.The replication cycle of poliovirus is initiated (1) by binding to the cell surface receptor CD155. The virion is taken up via endocytosis, and the viral RNA is released (2). Translation of the viral RNA occurs by an IRES-mediated mechanism (3). The polyprotein is cleaved, yielding mature viral proteins (4). The positive-sense RNA serves as template for complementary negative-strand synthesis, producing double-stranded replicative form (RF) RNA(5). Many positive strand RNA copies are produced from the single negative strand (6). The newly synthesized positive-sense RNA molecules can serve as templates for translation of more viral proteins (7) or can be enclosed in a capsid (8), which ultimately generates progeny virions. Lysis of the infected cell results in release of infectious progeny virions (9) **[27]**

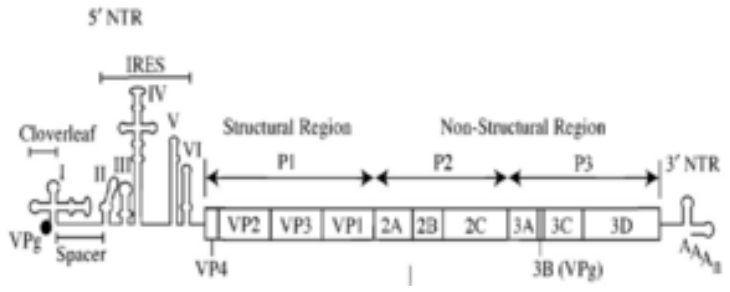


Fig 2.The genomic structure of poliovirus type 1 **[27]**

2.3 Origin and Serotypes

Poliovirus is structurally similar to other human enteroviruses (coxsackievirus, echovirus, and rhinovirus), which also use immunoglobulin-like molecules to recognize and enter host cells. Phylogenetic analysis of the RNA and protein sequences of poliovirus suggests that it may have evolved from a C-cluster Coxsackie A virus ancestor, that arose through a mutation within the capsid. The distinct speciation of poliovirus probably occurred as a result of change in cellular receptor specificity from intercellular adhesion molecule 1 (ICAM-1), used by C-cluster Coxsackie A viruses, to CD155; leading to a change in pathogenicity, and allowing the virus to infect nervous tissue.

The mutation rate in the virus is relatively high even for an RNA virus with a synonymous substitution rate of 1.0×10^{-2} substitutions/site/year and non synonymous substitution rate of 3.0×10^{-4} substitutions/site/year.Base distribution within the genome is not random with adenosine being less common than expected at the 5' end and higher at the 3' end. Codon use is not random with codons ending in adenosine being favoured and those ending in cytosine or guanine being avoided. Codon use differs between the three genotypes and appears to be driven by mutation rather than selection.

The three serotypes of poliovirus, PV1, PV2, and PV3, each have a slightly different capsid protein. Capsid proteins define cellular receptor specificity and virus antigenicity. *PV1* is the most common form encountered in nature, but all three forms are extremely infectious. As of November 2015, wild PV1 is highly localized to regions in Pakistan

and Afghanistan. Wild PV2 was declared eradicated in September 2015 after last being detected in October 1999 in Uttar Pradesh, India.As of November 2015, wild PV3 has not been seen since its 2012 detection in parts of Nigeria and Pakistan.

Specific strains of each serotype are used to prepare vaccine against polio. Inactive polio vaccine is prepared by formalin inactivation of three wild, virulent reference strains, Mahoney or Brunenders (PV1), MEF-1/Lansing (PV2), and Saukett/Leon (PV3). Oral polio vaccine contains live attenuated (weakened) strains of the three serotypes of poliovirus. Passaging the virus strains in monkey kidney epithelial cells introduces mutations in the viral IRES, and hinders (or attenuates) the ability of the virus to infect nervous tissue.

2.4.POLIOVIRUS VACCINES

Two different poliovaccines have been developed ,the IPV of Salk and Youngner**[1]** and the live, attenuated OPV of Sabin. Both vaccines are safe and effective, each with its particular advantages and disadvantages**[2]**, and both have played an important role in the control of poliomyelitis.

2.4.1 Inactivated Poliovirus Vaccine

IPV was the first poliovaccine to be licensed. Its development followed several key advances in virology**[1]**:*(a)* the cultivation of poliovirus in non-neural cells,(*b*) the identification of three poliovirus serotypes ,(*c*) the finding that viremia precedes paralysis,and (*d*) the demonstration that administration of immune globulin protects against paralytic polio**[3].** IPV is prepared by formalininactivation of three wild, virulent references trains, Mahoney (type 1), MEF-1(type 2), and Saukett (type 3). A less virulent type 1 strain, Brunenders (50), is used in IPV production in Sweden and Denmark. Although antigenic sites 1 (**Figure 1**) of types2 and 3 are modified by formalin inactivation, immunization with IPV can induce high titers of neutralizing antibodies protective against all poliovirus strains. After the Cutter incident (134), conditions for IPV manufacture were modified, resulting in a reduction in the immunogenicity of IPV preparations. However, improvements in cell culture technology in the 1970s led to the development of an enhanced-potency IPV, similar in immunogenicity to the original product, which has replaced the second generation IPV.

IPV was licensed for use in the United States, Canada, and Western Europe in 1955 and was the only poliovaccine available until licensure of OPV in 1961–1962. IPV use in the United States declined after the introduction of OPV, but it has been used continuously by some countries in Western Europe(Finland, Iceland, Sweden, and The Netherlands) and some provinces of Canada**[1]**.In 1997, in response to the eradication of wildpolioviruses in the Americas and the continuing occurrence of cases of VAPP ,the United States shifted from an all-OPV immunization schedule to a sequentialIPV/OPV schedule, which was replaced in2000 by an all-IPV schedule .

2.4.2 Oral Poliovirus Vaccine

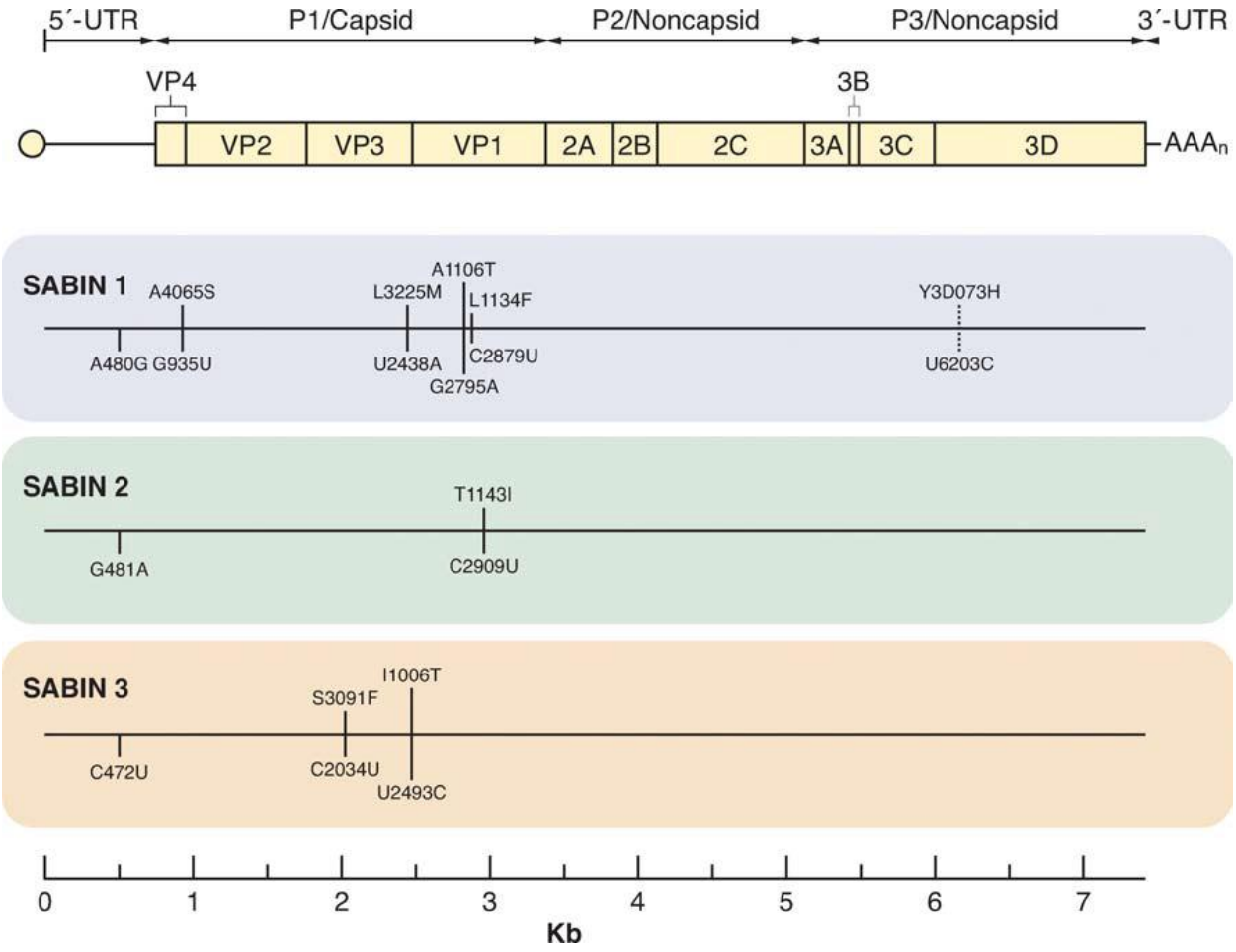
OPV was developed in the 1950s using the well-established empirical approach of rapid passage of virus at subphysiologic temperatures in cells and tissues of nonhuman origin. Molecular biology was then in its infancy, and although virologists of that period were aware of the rapid advances in bacteriophage genetics**[4]**, they were guided primarily by the principles of genetic selection for spontaneous mutants. OPV development occurred against the backdrop of the disastrous attempts to develop poliovaccines two decades earlier, when investigators failed to distinguish between chemical inactivation of infectivity and attenuation of neurovirulence by genetic selection and were unaware of the existence of multiple poliovirus serotypes . However, the rising incidence of paralytic polio in developed countries in the early 1950s greatly increased the urgency of developing and deploying effective poliovaccines. It would be another two decades after licensure of OPV that the molecular basis of OPV attenuation became amenable to systematic investigation.

2.4.3 Genetic Determinants of Attenuation of the Sabin OPV Strains

Identification of the genetic determinants of attenuation of the Sabin OPV strains has been comprehensively reviewed . The first reports of the sequences of complete poliovirus genomes in the early 1980s and the development of infectious poliovirus cDNAclones**[5]** opened the way for

systematic investigation of the critical mutations responsible for the attenuated and temperature-sensitive phenotypes of the Sabin OPV strains. A common feature of the Sabin strains is the presence of nucleotide substitutions in the IRES, which in serotypes 1 and 3 are critical attenuating mutations. Additional mutations encoding amino acid substitutions in the capsid region contribute to and stabilize the attenuated phenotype.

Sabin 1.The 57 nucleotide substitutions distinguishing the Sabin 1 strain from its neurovirulent parent, Mahoney, are scattered throughout the genome**[6].** Six map to the 5_-UTR, 49 map to the coding region (21 of which encode amino acid substitutions), and 2 map to the 3_-UTR. Infectious cDNA constructs containing different combinations of blocks of Sabin 1 and Mahoney sequences were tested for neurovirulence in monkeys or transgenic mice expressing the CD155 receptor, for temperature sensitivity, and for other phenotypic properties distinguishing the two strains .The single most important determinant of the attenuated phenotype of Sabin 1 was the A→G substitution atposition 480 (abbreviated A480G) in the IRES. Four other substitutions contributing to the attenuated phenotype mapped to the capsid region (one in VP4, one in VP3, and two in VP1), and one substitution contributing to the temperature-sensitive phenotype (but not to the attenuated phenotype) mapped to the 3Dpol region .



Location of principal attenuating nucleotide (*lower bars*) and amino acid (*upper bars*) substitutions in each of the three Sabin OPV strains. Abbreviations of nucleotide residues: A, adenine; C, cytosine; G, guanine; U, uracil. Abbreviations of amino acid residues: A, alanine; C, cysteine; F, phenylalanine; H, histidine; I, isoleucine; L, leucine; M, methionine; S, serine; T, threonine; Y, tyrosine. Substitutions are shown as nonattenuated parent–position–Sabin strain; nucleotide positions are numbered consecutively from residue 1 of the RNA genome; amino acid positions are indicated by the abbreviated name of the viral protein (4, VP4; 2, VP2; 3, VP3; 1, VP1; 3D, 3D-polymerase) and numbered consecutively from residue 1 of each protein. For example, a guanine (Mahoney)→uracil (Sabin 1) substitution at RNA position 935 (G935U) encodes an alanine (Mahoney)→serine (Sabin 1) replacement at residue 65 of VP4 (A4065S). The Y3D073H substitution in Sabin 1 and S3091F substitution in Sabin 3 are important determinants of temperature sensitivity**. [7,8]**

Sabin 2. Only two nucleotide substitutions (G481A in the IRES, and C2909U encoding a T → I substitution at position 143 of VP1) appear to be responsible for the attenuated phenotype of Sabin 2 (**Figure 2**).

The total number of sequence differences between the parental P712 strain and Sabin 2 is uncertain. However, because P712 has inherently low neurovirulence, identification of critical attenuating sites in Sabin 2 involved determination of the effects of introduction of sequences derived from a minimally divergent neurovirulentrevertant of Sabin 2 (obtained from a case of VAPP) into infectious cDNA constructs derived from Sabin 2. The precise contribution of the G481A substitution, which is not found in the IRES sequences of type 2 wild polioviruses, to the attenuated phenotype of Sabin 2 is unclear, as variants and even vaccine seed stocks containing the wild-type allele are attenuated in transgenic mice **[9].**

Sabin 3. Detailed analysis of the attenuated phenotype of Sabin 3 has been possible because the neuro virulent parental strain, Leon, differs from Sabin 3 by only 10 nucleotide substitutions (182). In addition, numerous neuro virulent revertants of the Sabin 3 strain have been isolated from patients with VAPP(203) and from healthy OPV recipients (109). Only three substitutions (C472U in the IRES, C2034U encoding an S → F substitution at position 91 of VP3, and U2493C encoding an I→ T substitution at position 6 of VP1) appear to be the main determinants of the attenuated phenotype (**Figure 2**). In all three Sabin strains, the attenuated phenotype is determined by multiple substitutions. The substitutions in the IRES, which alter stem-loop structures (63, 68, 126, 132) and reduce the efficiency of initiation of translation of the poliovirus RNA template (68, 188), contribute most to the attenuated phenotype of the Sabin 1 and Sabin 3 strains. Mutations that restore the original stem-loop structure in the IRES [Sabin 1: G480A (backmutation) or U525C (suppressor); Sabin 2: A481G; Sabin 3: U472C] are frequently found in vaccine-related isolates from healthy OPV recipients (127, 132) and patients with VAPP (61), as well as from the environment **[10]**. The role of the IRES to the neurovirulent phenotype of the type 1 Mahoney strain was demonstrated by the reduced neurovirulence of chimeras in which the IRES of Mahoney had been replaced with that of human rhinovirus type 2 . In Sabin 3, the critical C472U substitution reduces the efficiency of binding of the PTB, required for initiation of translation, to the IRES. The translational deficit for Sabin 3 is moderate in intestinal cells,

where PTB levels are high, but severe in neurons, where PTB levels are low. The precise mechanisms by which the capsid mutations contribute to the attenuated phenotype are less clear. Impairment of the efficiency of binding to the CD155 receptor and reductions in the stability of the capsid**[11]**may play a role.

Sabin and other developers of OPV strains struck a balance between low neuropathogenicity, good immunogenicity, and acceptable levels of genetic stability (150). The high genetic stability of the Sabin type 1 strain is probably attributable to the greater number of substitutions contributing to the attenuated phenotype. This property is especially important for the Sabin 1 vaccine strain, because type 1 wild polioviruses typically have high paralytic attack rates and can spread over wide geographic areas in explosive outbreaks**[12]**. Sabin 2 may revert more rapidly, but its immunogenicity is high and the paralytic attack rates of type 2 wild polioviruses are low . Sabin 3 is associated with the highest rates of VAPP (see below), which is probably a result of low genetic stability of the critical attenuating substitution (28), relatively low immunogenicity , and an intermediate paralytic attack rate for type 3 polioviruses. Nonetheless, all three Sabin strains normally have low pathogenic potentials, and incidence of VAPP in countries with high rates of OPV coverage are >3000-fold lower than the incidence of paralytic poliomyelitis in areas with circulating wild polioviruses.

2.4.4. Molecular Approaches Toward the Development of More Stable OPV Strains

Advances in the understanding of the genetic basis for attenuation of the Sabin OPV strains have generated renewed interest in developing OPV strains with improved genetic stabilities. Nomoto and colleagues constructed chimeras with the capsid regions of Sabin 2 and Sabin 3 inserted into the genetic background of the more highly substituted and stable Sabin 1 strain. The same group also developed less neurovirulent strains by introducing deletions into the IRES**[13].**Also constructed attenuated strains by modification of IRES sequences, while attenuated Mahoney by substitution of its IRES with the homologous rhinovirus IRES sequences. Finally, Macadam et al.constructed IRES mutants based upon Sabin 3 that had enhanced genetic stabilities in vitro. Despite the sophistication of these approaches, the development of attenuated virus vaccines

remains highly empirical, and the biological properties of these strains in the field cannot be predicted.

2.4.5. CONTROL OF POLIO BY IMMUNIZATION

The key biological requirements form poliovirus eradication are (*a*) absence of a persistent carrier state, (*b*) virus spread is by person-to-person transmission, (*c*) immunization interrupts virus transmission, (*d*) absence of any nonhuman reservoir hosts for the virus, and (*e*) finite virus survival time in the environment. An additional important nonbiological requirement for any disease eradication effort is political will, arising from the perceived benefits of eradication, and expressed internationally through resolutions passed by the World Health Assembly.

Essential to the success of the Global PEI has been the strong alliance among national governments, international agencies such as UNICEF, and private partners, including Rotary International (whose international PolioPlus campaign began in 1979)**[14],** the UN Foundation, and the Bill and Melinda Gates Foundation.

The basic strategy of the PEI is (*a*) high routine immunization coverage of infants with OPV, (*b*) supplementary OPV immunization through NIDs and SNIDs, (*c*) targeted doorto- door "mop-up" OPV immunization in areas of focal transmission, and (*d*) sensitive surveillance for poliovirus (80). High rates of routine OPV immunization are required to block poliovirus circulation in areas where the risk factors converge, conditions under which routine OPV coverage rates exceeding 90% may be insufficient to block poliovirus circulation. Such rates are currently unattainable through routine immunization in the least developed countries. Supplementary immunization is the mainstay of polio eradication in developing countries, and has been instrumental in raising population immunity rates above the thresholds required to block poliovirus transmission**[15].** Supplementary immunization strategies are driven by poliovirus surveillance, which is used to guide the intensified SNIDs and mop-up campaigns to the reservoir communities where the chains of poliovirus transmission continue to survive and propagate.

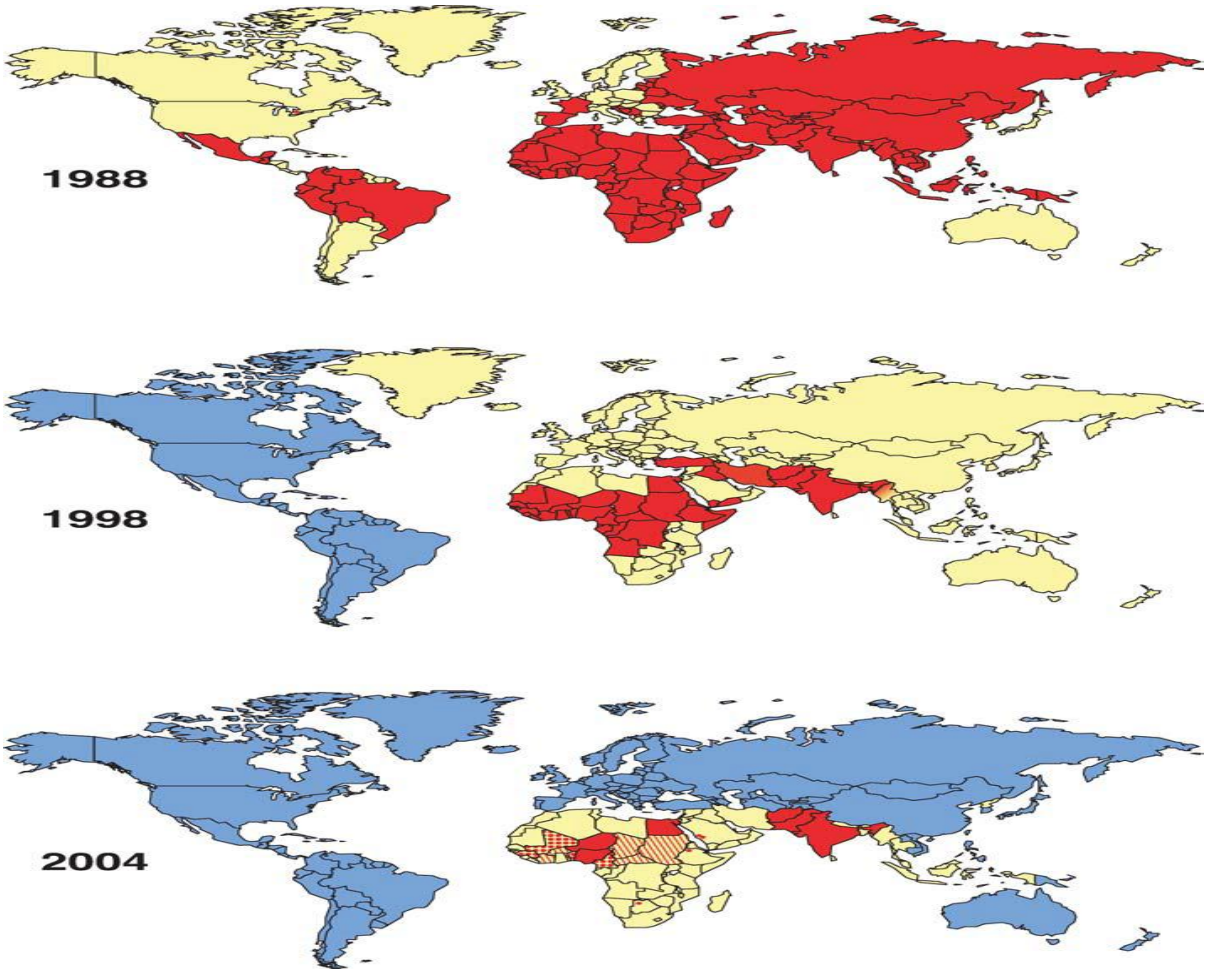


Fig 3.Geographic distribution of polio-endemic and polio-epidemic countries in 1988, 1998, and 2004. Polio-endemic countries are colored red, nonendemic countries are colored yellow, and countries within WHO regions certified as polio-free are colored blue. Countries experiencing repeated cross-border importation of wild poliovirus (Iran and Myanmar in 1998) are indicated by a red-to-yellow gradient, countries with re-established transmission of more than 1 year (2004 map) are indicated with red and yellow hatch pattern, countries with sporadic importation within a single year (2004 map) are indicated by red shading over yellow background, and countries with single cases associated with imported wild poliovirus (Canada in 1988; Botswana and Ethiopia in 2004) are shown with a red dot at the location of each case. The two red dots in Saudi Arabia (2004) represent two cases associated with separate importations of wild poliovirus from epidemic areas in Sudan.**[16]**

2.4.6. VACCINE-DERIVED POLIOVIRUSES

Three categories of poliovirus isolates are recognized: (*a*) OPV-like isolates, which differ from the respective parental Sabin strains by <1% of VP1 nucleotides; (*b*) VDPVs, which differ from the parental Sabin strains at 1% to 15% of VP1 nucleotides (i.e., having ≥10 nucleotide substitutions); and (*c*) wild polioviruses, which differ from the Sabin strains at >15% of VP1 nucleotides **[17].**

The demarcation between OPV-like isolates and VDPVs of >1% VP1 divergence implies that replication of vaccine virus, either within an individual and during personto- person transmission, had occurred for approximately 1 year or more, in contrast to the normal period of poliovirus excretion of 4 to 8 weeks. The definition is based upon the likelihood that VDPVs have had a history of prolonged replication since administration of the initiating OPV dose, but it makes no inference about the biological properties of VDPVs. It does not imply that isolates having <1% divergence would lack the capacity to cause paralytic disease in humans or be unable to initiate sustained person-to-person transmission in poorly immunized populations. Indeed, the critical attenuating mutations of the Sabin strains frequently revert well before nucleotide substitutions accumulate to the level of 1%**[10]**.

A small proportion of individuals with Bcellimmunodeficiencies exposed to OPV develop iVAPP. A much smaller proportion (∼16%) of iVAPP patients become long-term excretors of iVDPVs**[18]**. To date, 23 longtermiVDPVexcretors have been identified worldwideby WHO since 1962 . Although most of the iVDPVexcretors had paralysis, a substantial proportion (30%) has shown no signs of paralysis. More than half (57%) of the patients for which the outcome is known have survived, and most survivors have spontaneously stopped iVDPV excretion. Deaths are usually attributable to VAPP or other complications of immunodeficiency**[18].**

2.4.7. Properties Shared by iVDPV and cVDPV Isolates

In addition to the unusual degree of genetic divergence from their respective Sabin strains, VDPVs share other important properties. The most important of these is the capacity to cause paralytic polio in humans, documented for cVDPVs and for many, but not all, iVDPVs.

Sequence analysis of both categories of VDPV isolates has shown that the critical determinants of the attenuated and temperature-sensitive phenotypes either have reverted or have been exchanged out by recombination. These genetic properties correlate with high neurovirulence for transgenic mice expressing the CD155 receptor**[19]**and the capacity to replicate to high titers in cell culture at supraoptimal temperatures . Moreover, all VDPV isolates characterized so far have antigenic properties distinct from the original Sabin strains. The antigenic differences from the Sabin strains are less pronounced for type 2 cVDPVs than for type 1 cVDPVs, possibly because selection against the Sabin 2 antigenic sites is less intense during replication in immunocompetent persons. While these experimentally determined properties may correlate with clinically srelevant properties, they are not unique to VDPV isolates, as OPV-like polioviruses isolated from healthy individuals and VAPP patients may share some or all of these traits**[20].**

2.4.8. Independence of iVDPVs and cVDPVs

The shared properties of iVDPV and cVDPV isolates have prompted the suggestion that the recognized cVDPV outbreaks may have been triggered by iVDPV chronic excretors. Because the early events in cVDPV outbreaks have not been observed (except possibly in Guizhou and Romania, 2002), it is difficult to rule out this possibility rigorously. However, two key observations suggest that iVDPV infections and cVDPV outbreaks have so far been independent. First, cVDPV outbreaks have occurred in communities where the expected survival times for persons with B-cell immunodeficiencies are short **[21].** Second, the extent of amino acid substitution in the neutralizing antigenic sites is typically lower for cVDPV isolates than for iVDPV isolates having similar levels of total capsid sequence divergence from the vaccine . Past experience, however, may not be a predictor of future conditions, as iVDPVs from chronically infected individuals may present a risk for spread into the community in some settings in the post-OPV era, and additional surveillance measures may then be needed to detect long-term iVDPVexcretors .

2.4.9. VDPVs AND THE ENDGAME STRATEGY FOR POLIO ERADICATION

The restriction of indigenous wild poliovirus circulation to only a few parts of the world has greatly reduced the risk of poliovirus importation. However, this favorable circumstance places increased importance on polio immunization to maintain population immunity in countries that are now polio-free. In areas where the risks for cVDPV emergence are highest, periodic NIDs or SNIDs may be necessary to close any gaps in population immunity that may arise from deficiencies in routine OPV coverage. The required frequency and extent of the immunization campaigns is determined by the magnitude of the risk factors for poliovirus circulation in each population. In the current pre-eradication phase, the same strategies implemented to limit the spread of imported wild polioviruses would similarly limit the spread of VDPVs, and all poliovirus outbreaks can be controlled by mass immunization campaigns using OPV. The greatest challenges arise in the posteradication phase, when the risks of wild poliovirus infection are greatly reduced, public concern about polio as a disease wanes, population immunity can be maintained only by immunization, and the only continuous source of poliovirus infection and polio adverse events is from OPV. While the original plan for global polio eradication envisioned a straightforward endgame strategy**[22]**, that phase has become far more complex in view of the risks presented by VDPVs.

The occurrence of VDPVs appears to violate two key biological assumptions for poliovirus eradication: (*a*) the absence of a persistent carrier state and (*b*) the implicit assumption that the vaccine would not circulate and spread. When these principles were first explicitly detailed in 1997**[23]**, few examples of immunodeficient long-term chronic excretors had been described, and in every known case excretion of vaccinederived poliovirus had stopped. Moreover, the cVDPV outbreaks in Hispaniola, the Philippines, and Madagascar had not yet occurred, and the cVDPV outbreak in Egypt had not yet been recognized. The subsequent recognition of long-term iVDPVexcretors and the potential for cVDPV emergence has prompted in depth reassessment of the endgame strategies for the Global PEI .

2.5. Indefinite use of OPV.

The potential advantage of indefinite OPV use in countries currently using OPV would be the maintenance of population immunity, thereby affording protection from reintroduced wild poliovirus or the spread of iVDPVs. The disadvantages would be a sustained annual incidence of up to 500 cases of VAPP, and the necessity to maintain indefinite public investment in and community support for immunization against a disease that had been certified as eradicated. The practical challenge to sustain adequate levels of OPV coverage even in the current era was underscored by the recent cVDPV outbreaks. Indefinite use ofOPis not a viable policy because it is likely that OPV coverage rates would fall below safe levels in many countries because of competing health priorities (as in Madagascar), because of social instability (as in Haiti), or because susceptibles may accumulate in the highest risk communities (as in the Philippines).

The primary appeal of this scenario is that it adheres to the principle that each country should develop its own immunization policies. However, as illustrated by the Hispaniola outbreak, this scenario presents unacceptable risks for emergence of cVDPVs, as persons infected with OPV virus could reseed countries (or communities) where polio immunization had stopped.

2.5.1. Sequential removal of Sabin strains from OPV

A variation on coordinated cessation would be the sequential removal of Sabin strains from OPV as the corresponding wild poliovirus serotypes are eradicated (29). In this scenario, coordinated transition would first be to a bivalent (type 1 + type 3) OPV, followed by coordinated transition to a monovalent OPV, and then to the sole remaining serotype (most likely type 1). Coordinated transition to a bivalent OPV would permit a critical type-specific test of the coordinated OPV cessation strategy, especially because derivatives of the Sabin 2 strain have the highest potential for secondary spread and represent the VDPV serotype most frequently found. A variation on this approach would be to use monovalent vaccines to the surviving serotype(s). Despite its scientific appeal **[24],** this approach faces three key obstacles: (*a*) securing regulatory approval for a type 1+ 3 bivalentOPVmay prove difficult, as it would represent an untested OPV product with no record of performance in the field; (*b*) logistical implementation of simultaneous

withdrawal of all preceding OPV formulations will likely be arduous; and (*c*) the urgency to stop OPV use (see below) may not favor implementation of a more protracted endgame strategy.

2.5.2. Development of new OPV strains.

It would be highly desirable to apply the dramatic advances in the understanding of the molecular mechanisms for attenuation of the Sabin strains to the development of safer, more effective, and genetically more stable OPV strains. Use of the engineered Sabin strain variants already developed**[13]** might greatly reduce the incidence of VAPP. Less clear is to what extent these new strains could reduce the risk of emergence of VDPVs. VDPVs with increased replicative fitness might still arise (but possibly at much lower frequencies) through loss of the engineered IRES sequences by recombination, especially in countries with high rates of enterovirus carriage. Moreover, it is difficult to predict the evolution of any engineered constructs during chronic infection of an immunodeficient person. Nonetheless, new OPV strains would likely offer multiple advantages over the current Sabin strains. The critical challenge to the introduction of new attenuated OPV strains would be the design and implementation of field trials sufficiently large to detect a reduction from the current VAPP risk or from the even less readily quantifiable VDPV risk.

2.5.3. Current Endgame Perspectives and Challenges

After review of the threat that VDPVs present to polio eradication**[25]**, WHO has outlined a strategic plan for cessation of all routine use of OPV. The strategy has the following components: (*a*) Routine OPV use will stop in a globally synchronized manner approximately two to three years after the last detection of wild poliovirus, (*b*) IPV will not be used routinely in all countries, (*c*) AFP and poliovirus surveillance will be maintained for at least three years after detection of the last cVDPVs, (*d*) all laboratory stocks of poliovirus will be contained upon discontinuation of OPV, and (*e*) a global OPV stockpile and response capacity will be established to stop any poliovirus transmission that may occur after cessation of OPV use. Despite the urgent need to stop wild poliovirus circulation as quickly as possible, the WHO endgame strategy is developing rapidly.

2.5.4. Detection and control of iVDPV infections.

The number of long-term iVDPVexcretors is expected to steadily decline in countries that have shifted to IPV. The continued presence of iVDPVexcretors may not represent a serious health risk within countries that maintain high rates of IPV coverage. By contrast, iVDPVexcretors in middle-income countries or within high-income enclaves in low-income countries could present a serious health risk in the post-OPV era. It is important that these risks be properly assessed and that surveillance strategies are developed to detect iVDPV infections wherever they may occur (see below). Effective treatments for clearing iVDPV infections must also be found. Although attempts to clear one chronic iVDPVinfection by combined treatment with oral immunoglobulin and pleconaril (a firstgeneration antiviral effective against some poliovirus strains) were unsuccessful**[26],** alternative protocols and the use of newer antiviral agents may yield more promising results.

Because poliovirus could be reintroduced into the population by the emergence and spread of VDPVs, by a breach of containment in a laboratory or poliovaccine production facility, or even by intentional release, it is essential that poliovirus surveillance is maintained for the foreseeable future. Immediately after cessation of OPV use, it will be necessary to intensify poliovirus surveillance, especially in the traditional reservoirs for wild poliovirus circulation where the biological risks for poliovirus circulation (and potentially of cVDPV emergence) are highest. This would include populous, tropical, developing countries whose effective immunization programs had overcome the inherent biological risks. The sensitivity of environmental surveillance in developing countries would increase because the high background of OPV-like viruses would fade. The technical aspects of poliovirus surveillance would simplify six months into the post-OPV period, as any poliovirus isolate would raise concerns and be regarded as a potential public health emergency, especially if there is any geographic clustering of isolates of the same serotype. If, as is hoped, the early post-OPV years prove uneventful, it will still be necessary to maintain surveillance for AFP cases and poliovirus infections in the coming years. AFP and poliovirus surveillance will be integrated into other disease surveillance activities as efforts to control other vaccine-preventable diseases broaden and as the global

network of laboratories first built to support polio eradication extends its support to wider infectious disease control initiatives.

2.5.5. Vaccine Prediction

Predictionsfor the specific vaccine and the processes which is used to develop them been an integral part of the selection scheme. These predictions which are required to calculate the benefits of health which are expected from new vaccine and with its related costs. The features of vaccine (eg., live attenuated virus) which may improve its efficacy, and the twists and turns of development process which determine the concerned costs alongwith introducing health benefits and the time by that they could be achieved.

Predictions were made separately for every vaccine/ combination of disease. The criteria for addition was to detect whether a valid consensus would be easily identified on the components of vaccine potential nature. The description of vaccine candidates for the fast going development have been detected for each and every disease. The descriptions of vaccines which are based on recent researches in related areas.Predictions which are based on possible moves and not related to the judgments about the interest of courses of action. The predictions related to the efficacy of vaccine presents the cost of population based rather than cost of production of antibody in the individual.

2.5.6. Requirement of New Vaccine

The significant impact for the introduction of new vaccine which is based on both pros and cons on the whole programme of national immunization and related to health system of the country.
The decision whether to introduce a new vaccine into the programme of national immunization, decision makers take into account the impact of the starting on both the programme. The entry of vaccine into the programme of national immunization will get the chance to improve the immunization and health system.

However it is necessary to view again the immunization programme so that before the introduction of new vaccine weak areas can be identified. The vaccine introduction may get the chance to improve the whole immunization programme and the health system. For the vaccine

introduction there are many activities which is done to prepare, execute and function. The increased demand for the new vaccine - can introduce in children and adolescence who have not been immunized before in their schedule of immunization. In many countries, the entry of new vaccine which starts with an initial launch campaign.In this the new vaccine which is heavily circulated.

The vaccine which is going to develop for the polio which can stop the reversal. In the few years back the polio has been completely eradicated but due to some outbreaks which lead to the introduction of polio again. By using various tools for the four serotypes of polio virus it is being tried to develop a such kind of vaccine which can stop the virus which is replicating in the system. By calculating the various parameters like immunogenecity score, population coverage of a serotype , hope it is possible to give the decision whether a protein will attack the immune system or not.

3. METHODOLOGY

3.1 DATA COLLECTION

Sequences of four serotypes of polio are retrieved i.e., VP1, VP2, VP3, VP4 from the Uniprot and run it in a FASTA format.

3.2 Epitope prediction of T-cytotoxic cell

The antigen sequences are first divided into small peptide which binds to MHC1 then these MHC1-peptide complexes bind to the CD8+ receptor of T-cytotoxic cell. So we first predicted MHC1 binding peptide then CD8+ receptors binding.

3.2.1 MHC I binding peptide prediction

We use IEDB prediction tool Tepitool for the prediction of MHC I binding peptide. because this is a user friendly tool. And we can manually select the different attributes here.

3.2.1.1 Alleles selection

A set of alleles selected on the bases of their population frequencies. We use most abundant alleles set of HLA class l which covers almost 99% population of the world.

3.2.1.2 IC50 value selection

The ic50 value of all these alleles are more than 100 nm, so we select a threshold of less than or equal to 100 nm for best prediction results.

3.2.1.3 Prediction method selection

We used netMHC1pan method for the prediction of mhc1 binding peptides. This is the only method which predict peptide on the bases of IC50 value.

3.2.2 CD8+ immunogenicity prediction

We used IEDB CD8+ immunogenicity prediction tool with the default setting. In this tool automatically masking of 1,2 and C-terminal of amino acid take place. Prediction score is calculated for all MHC I binding peptide and select only those peptide whose prediction score is more than zero.

3.2.3 Population coverage of T-cytotoxic epitope

We used POPULATION ANALYSIS tool for the calculation of population coverage of selected peptide and select those peptide which cover most of the population.

3.3 T-helper cell epitope prediction

These are the internal peptides so presented by mhc1 but they phagocytose by dendritic cells which are antigen presenting cells. They present these antigen peptides to the T-helper cell.

3.3.1 MHC II binding peptide prediction

We use IEDB prediction tool Tepitool for the prediction of MHC II binding peptide. And we can select different attributes here.

3.3.1.1 Alleles selection

We use 26 most abundant alleles of MHC II. These alleles IC50 value threshold cutoff selected here is 1000.

3.3.1.2 Prediction method selection

We use here a set of methods for prediction of most promisiouse binding of HLA ll alleles. In this only those peptide are selected which bind to at least 50% of alleles from these 26 alleles. For this we use the tepitool as the prediction tool and IEDB recommended as the prediction method.

3.3.2 CD4+ immunogenicity prediction

We use IEDB prediction tool for the prediction of immunogenicity of the MHC II binding peptides. In this immunogenicity score of all the peptide are given and we have to select only those peptide which have immunogenicity score more than zero.

3.3.3 Population coverage

We used IEDB population coverage prediction tool for the analysis of population coverage in India. We selected those peptide which cover most population.

3.4 Combined population coverage of MHC1 and MHC11

We calculated combined population coverage of selected peptide of both MHC I and MHC ii by PopulationCoverage Tool. The problem of population coverage in respect to MHC polymorphism which is further tricky by the facts that the different types of MHC which are performed at different frequencies. If the careful consideration is not taken, the vaccine which is having the biased population coverage would come out as a result. To overcome this issue the binding capacity of

potential epitopes which is actual with many different types of MHC molecules which are possible. The vaccines which are epitope based are designed which will maximize the population coverage .

4. Results

Vaccines have not been changed greatly. When the killed or weakened virus is injected it revive up the body's immune system for the defence in future. The newly developed type of vaccine is basically a product of biology's recently crafted tools of gene manipulation. When the Salk polio vaccine first came into market in 1955, some of the children were died from shots in which the virus has not been properly activated. This experiment targets to detect the best possible epitope from the four serotypes of polio virus to prevent and stop the reversal which is caused by the virus which is replicating in polio. After protein sequences are retrieved, T cell epitopes with the population coverage and the conservancy are mentioned. Various regions of proteins are identified. These peptides are capsid proteins which will interact with the host cell receptor that will provide the attachment of virion which will target the host epithelial cell. After binding with its receptor, the capsid will go through conformational changes.

4.1. T-cytotoxic cell epitope prediction

The antigen sequences are first divided into small peptide which binds to MHC1 then these MHC1-peptide complexes bind to the $CD8_+$ receptor of T-cytotoxic cell. So we first predicted MHC1 binding peptide then $CD8_+$ receptors binding.

4.1.1 MHC I binding predictions

Now from these consensus sequences we have to predict the mhc I binding peptide. We predicted them on the bases of their IC_{50} value. We selected only those peptides whose predicted IC_{50} value <100. Because we have to select best binders. For this we used TEPITOOL as a prediction tool. It include 6 steps in it to predict mhc1 binding peptides.

4.1.1.1 Input antigen sequence

In first step we have to upload the file containing sequence of the viral proteins in fasta format.

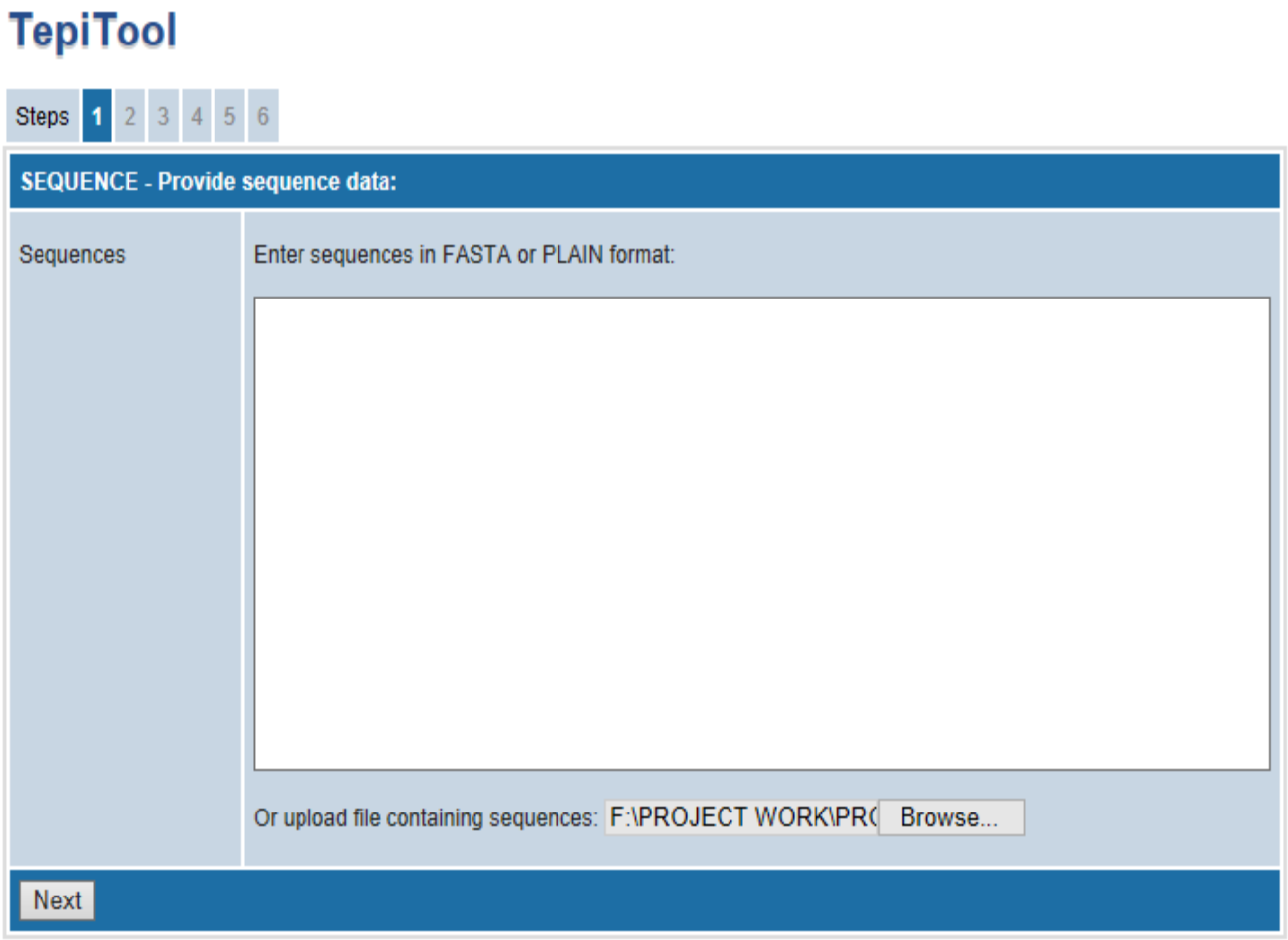


Figure 1 Tepitool prediction – input sequence

4.1.1.2 Selection of host species and MHC allele class

In second step we have to select the host species in our case that is human.And we have to select mhc allele class in this case that is mhc1 allele.

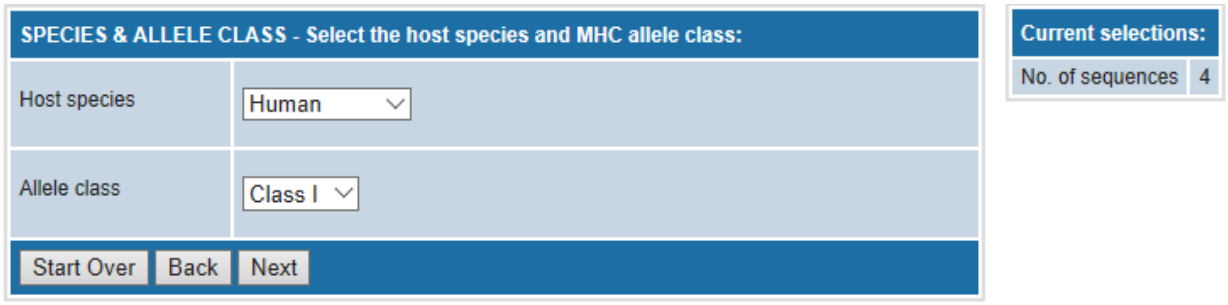


Figure 2 Tepitool prediction- species selection &alleles selection

4.1.1.3 Input list of selected alleles

In third step we have to put the list of mhc1 allele which have to bind to the peptide so we select the most frequently occurring alleles in human population.

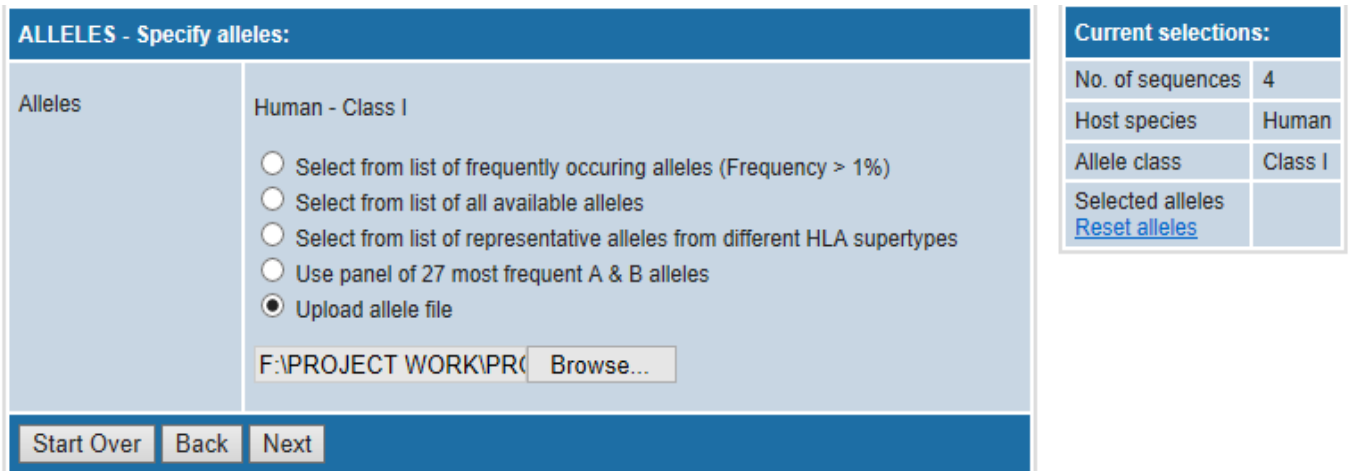


Figure 3 Tepitool prediction- input alleles list

4.1.1.4 Selection of peptides to be included in prediction

In forth step we have to select peptide which have to be included in prediction. We have to remove duplicate to reduce the no. of peptide. Have we have only predict 9 mer peptides only.

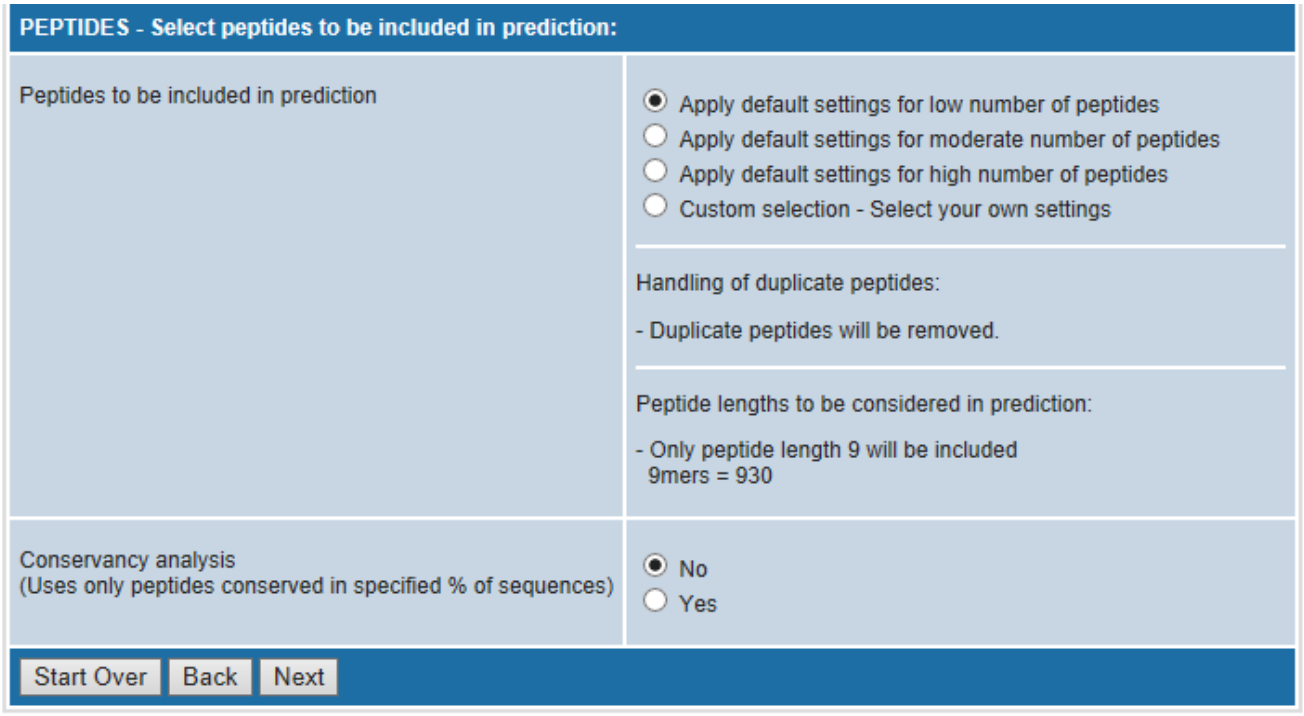


Figure 4 Tepitool prediction- selection of peptide

4.1.1.5 Select prediction & peptide selection and cutoff values

In this step we have to prediction and peptide selection methods and cutoff values. We selectpeptide on the base of predicted IC50 value and select the cutoff value 100.

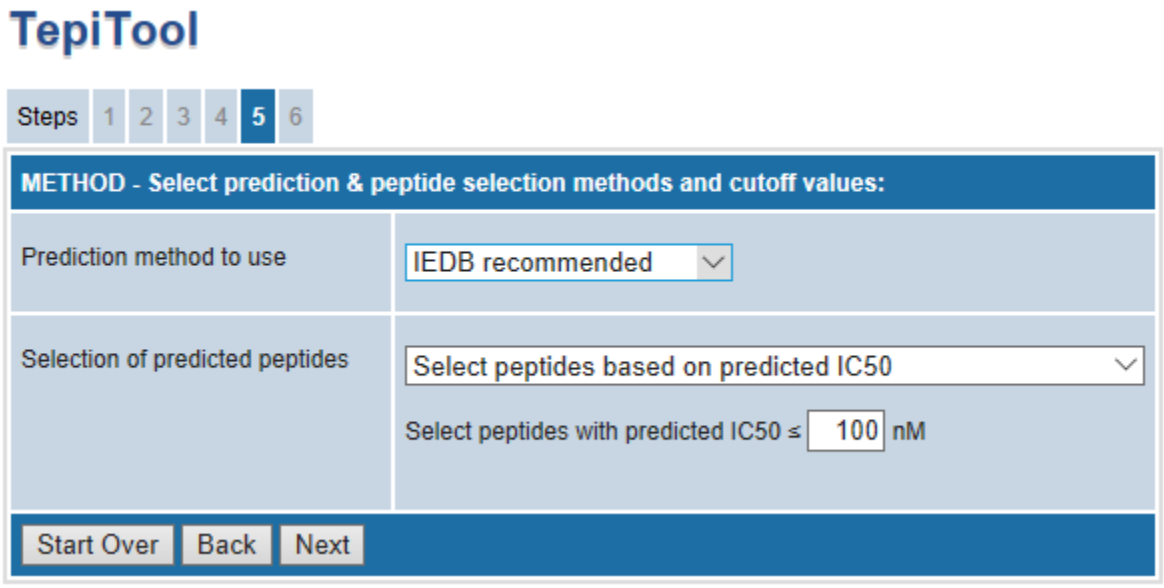


Figure 5 Select prediction method & selection of predicted peptide

4.1.1.6 Summary of all the selections & Antigen name

No. of sequences 1

Host species Human

Allele class Class I

Alleles A*01:01

A*02:01

A*02:06

A*03:01

A*11:01

A*23:01

A*24:02

A*25:01

A*26:01

A*29:02

A*30:01

A*30:02

A*32:01

A*33:01

A*33:03

A*68:01

A*68:02

Duplicate peptides **Removed**

Peptide lengths selected 9mers

Peptide overlap N/A

Conservancy analysis No

Prediction methodIEDB recommended

Peptide selection criterion Predicted IC50

Cutoff for peptide selection criterion100

Job nameVP1

Figure 6.Tepitool- summary of all selection

After that we have to download results. You can get them on your mail id too.

4.1.1.7 Results of MHC I Binding peptide

The results of**VP4** MHCI Binding results are as follows:

Table 1 Result of MHC I binding peptide and alleles combination of VP4

PEPTIDE	IC50	ALLELE
TINYTTINY	35.20	HLA-A*29:02
VLIKTSPML	58.20	HLA-A*02:01
VLIKTSPML	112.90	HLA-A*02:06
INYTTINYY	126.60	HLA-A*29:02
TINYTTINY	153.30	HLA-A*11:01
AYGGSTINY	208.90	HLA-A*29:02
MGAQVSSQK	318.50	HLA-A*11:01
DVLIKTSPM	380.90	HLA-A*26:01
MGAQVSSQK	453.60	HLA-A*30:01
YYRDSASNA	474.10	HLA-A*30:01
PSKFTEPIK	499.50	HLA-A*30:01

4.1.2. Class 1 immunogenicity prediction

These MHC I binding ligands have to bind to the CD8+ receptors of the cytotoxic T cell . All these peptide- mhc complex does not have the ability to bind with the CD8+ receptor so they can produce immune response. So for that we have to find immunogenicity. For immunogenicity prediction we select all the unique peptide sequence from the MHC I binding results. Then use them as the input sequence for immunogenicity prediction.

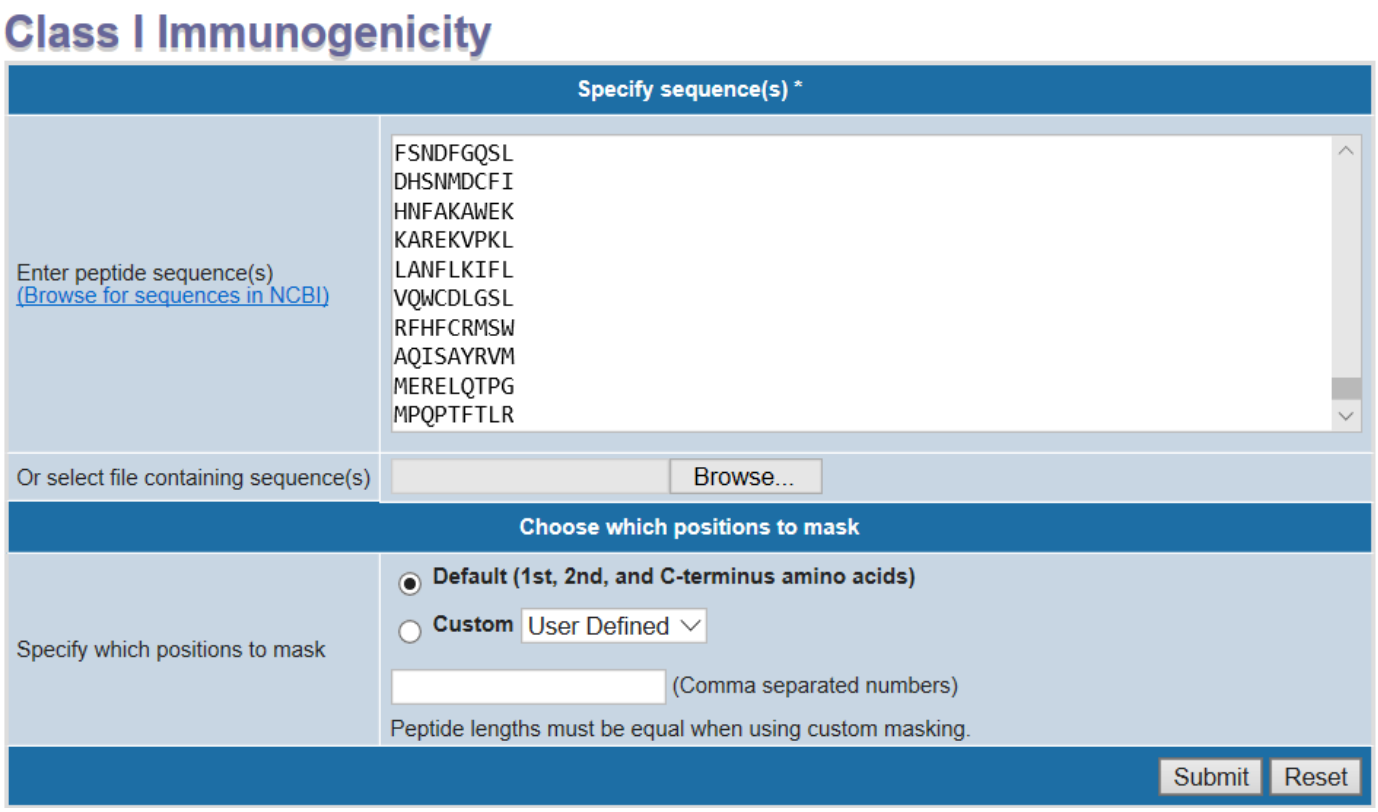


Figure 7. CD8+ Immunogenicity prediction

4.1.2.1. CD8+ Immunogenicity Prediction for tumor antigens binding to MHC 1

Immunogenicity of these peptide is predicted on the bases of immunogenicity score. We selected only those peptide which have immunogenicity score more than zero. Immunogenecity is generated by T cells recognition of peptide epitopes which will be displayed on the MHC. A function is created which will assign a score to each residue which will measure its tendency to be a part of a T cell epitope.

Table 2 Results of immunogenicity prediction

PEPTIDE	**COMBINED SCORE**	**IMMUNOGENECITY SCORE**
DFSQDPSKFTEPIKD	48.12384	97.7346
ESMIDNTVR	39.4846	87.4236
SSIESFFAR	40.9876	88.4236
SLNDFGILA	42.5678	86.5678
ISKFIDWLK	34.4567	84.5467
KHIRVWCPR	46.5678	82.4536
PSYFTEPIK	42.2341	75.4534
GTLTPLSTK	58.7856	88.2135
HLKMIAYD	56.3421	83.8745
YLTADNFQS	50.5437	81.6537
SCTMVVPWI	46.5873	79.5427

4.1.3 Population coverage results of class1 epitopes

T cells recognize a complex between a specific major histocompatibility complex (MHC) molecule and a particular pathogen-derived epitope. A given epitope will elicit a response only in individuals that express an MHC molecule capable of binding that particular epitope. MHC molecules are extremely polymorphic and over a thousand different human MHC (HLA) alleles are known. A disproportionate amount of MHC polymorphism occurs in positions constituting the peptide-binding region, and as a result, MHC molecules exhibit a widely varying binding specificity. In the design of peptide-based vaccines and diagnostics, the issue of population coverage in relation to MHC polymorphism is further complicated by the fact that different HLA types are expressed at dramatically different frequencies in different ethnicities. Thus, without careful consideration, a vaccine or diagnostic with ethnically biased population coverage could result[29].

Table 3. VP1 population coverage

EPITOPE LIST	**POPULATION COVERAGE**
Epitope #1: SSIESFFAR	42.29%
Epitope #2: ESMIDNTVR	35.68%

Table 4VP2 population coverage

EPITOPE LIST	POPULATION COVERAGE
Epitope #1: SLNDFGILA	42.47%

Table 5 VP3 population coverage

EPITOPE LIST	POPULATION COVERAGE
Epitope#1: ISKFIDWLK	18.98%
Epitope#2: KHIRVWCPR	36.98%

Table 6VP4 population coverage

EPITOPE LIST	POPULATION COVERAGE
Epitope#1: PSYFTEPIK	6.89%

Here it is desirable which will design a method to select a desirable sets of epitopes which will allow the detection of T and B cells. In this study, a method is developed which will automatically select the related epitope sets related to categorization system which is employed by IEDB. A web application will be applied in the IEDB that will be created which will allow the users to generate the epitope sets.

4.2. T-helper cell epitope prediction

These are the internal peptides so presented by mhc1 but these cell phagocytosis by dendritic cells which are antigen presenting cells they present these antigens peptides to the T-helper cell. We have to predict mhc II binding peptides followed by their CD4+ immunogenicity prediction

4.2.1 MHC II binding prediction

Now from these consensus sequences we have to predict the mhc11 binding peptide. We predicted them on the bases of their IC50 value. We selected only those peptides whose predicted IC50 value <200. Because we have to select best binders. For this we used tepitool as a prediction tool. It include 6 steps in it to predict mhc11 binding peptides.

4.2.1.1 Input antigen sequence

We have to select the sequence file of the viral antigen protein in the fasta format

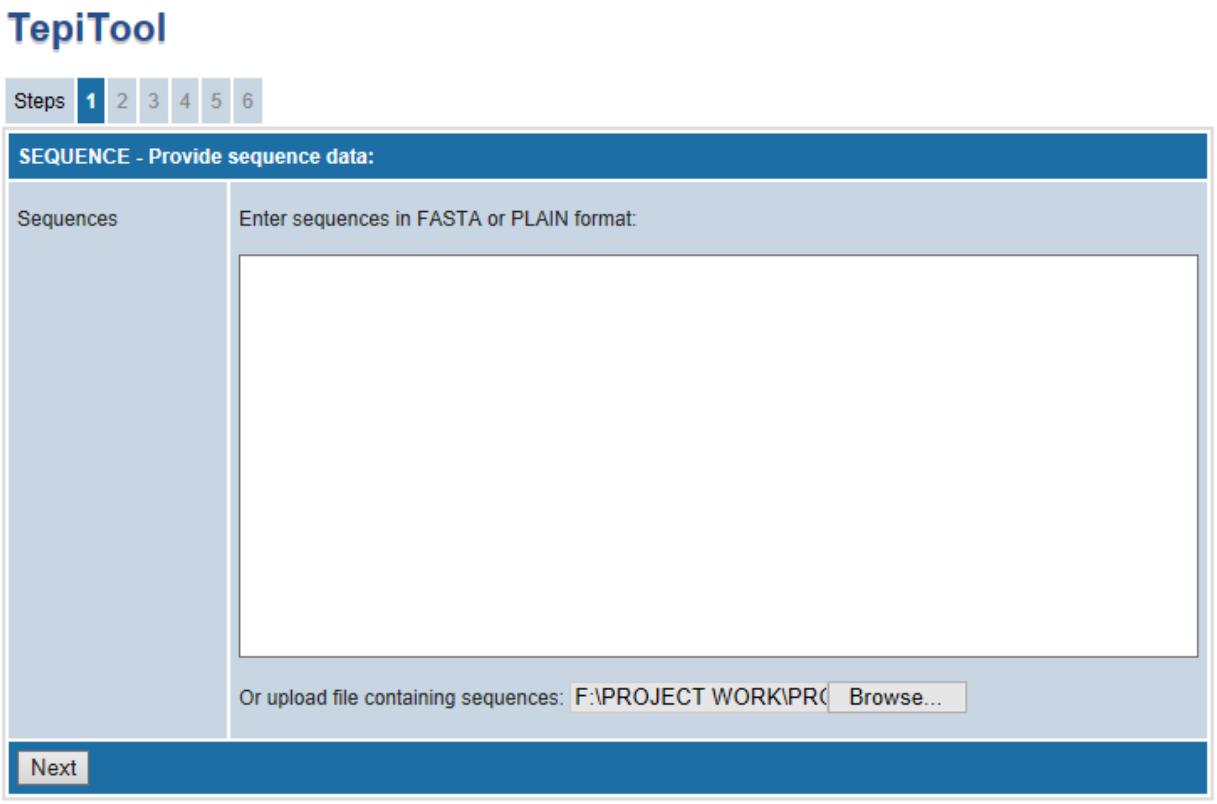


Figure 8.Tepitool MHC II binding prediction - input sequence

4.2.1.2 Selection of host species and MHC allele class

Here we have to select the host species in our case that is human and we have to select mhc allele class in this case that is MHC II class.

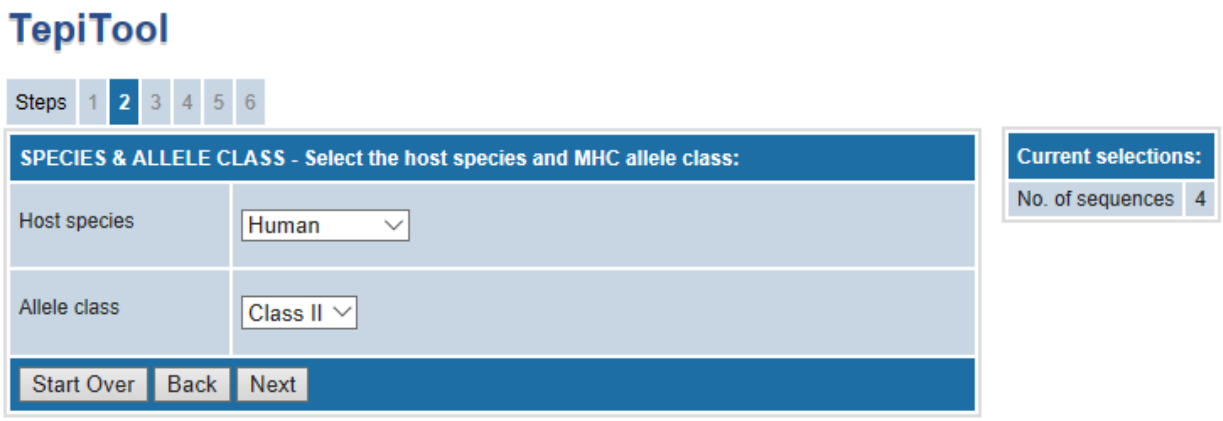


Figure 9.Tepitool mhc11 binding prediction-species &alleles selection

4.2.1.3 Selection of MHC II alleles

Here we have to select MHC II alleles which have to bind to the peptide in this case we select 26 most frequent occurring allele.

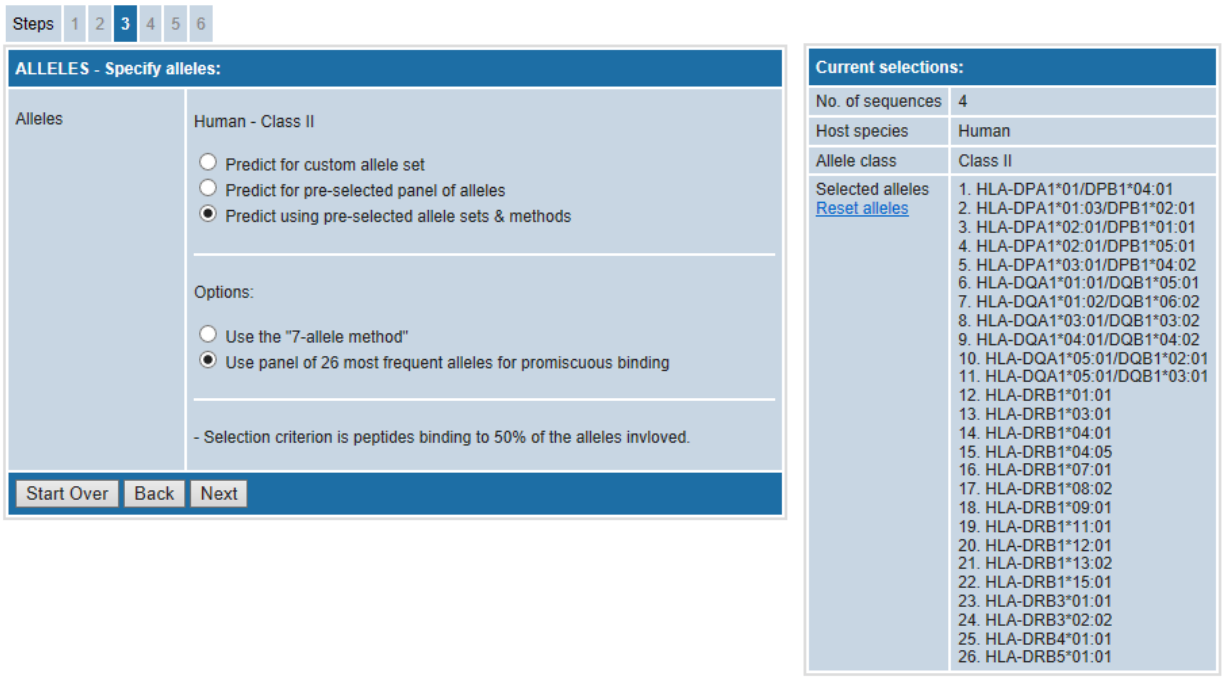


Figure 10 alleles selection

4.2.1.4 Selection of peptide

In this step we have to select peptide which have to be binding to the mhc alleles. We have to remove duplicate peptides. No. of overlapping peptides for 15 mer peptide to be generated (peptide length is fixed at 15 for class 2).

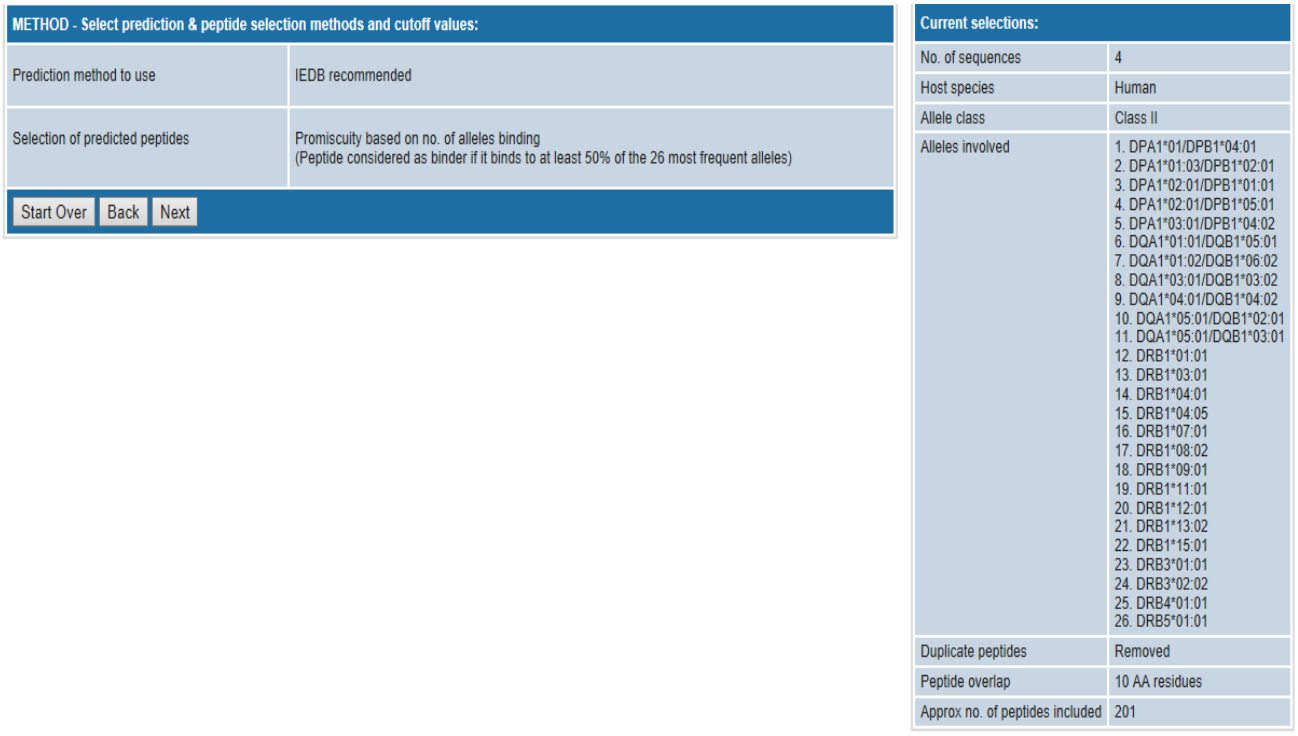
PEPTIDES - Select peptides to be included in prediction:	
Handling of duplicate peptides	Duplicate peptides will be removed
No. of overlapping residues for 15mer peptides to be generated (Peptide length is fixed at 15 for class II)	10
Approximate no. of peptides to be considered for prediction	145
Start Over Back	

Current selections:	
No. of sequences	1
Host species	Human
Allele class	Class II
Alleles involved	1. DPA1*01/DPB1*04:01 2. DPA1*01:03/DPB1*02:01 3. DPA1*02:01/DPB1*01:01 4. DPA1*02:01/DPB1*05:01 5. DPA1*03:01/DPB1*04:02 6. DQA1*01:01/DQB1*05:01 7. DQA1*01:02/DQB1*06:02 8. DQA1*03:01/DQB1*03:02 9. DQA1*04:01/DQB1*04:02 10. DQA1*05:01/DQB1*02:01 11. DQA1*05:01/DQB1*03:01 12. DRB1*01:01 13. DRB1*03:01 14. DRB1*04:01 15. DRB1*04:05 16. DRB1*07:01 17. DRB1*08:02 18. DRB1*09:01 19. DRB1*11:01 20. DRB1*12:01 21. DRB1*13:02 22. DRB1*15:01 23. DRB3*01:01 24. DRB3*02:02

	25. DRB4*01:01 26. DRB5*01:01

4.2.1.5 Selection of prediction method and cutoff values

Prediction method in this case is the IEDB recommended and selection of predicted peptide is based on promiscuity based on No. of alleles binding. Peptide considered as binders if it binds to 50% of 26 most frequently occurring alleles.



METHOD - Select prediction & peptide selection methods and cutoff values:	
Prediction method to use	IEDB recommended
Selection of predicted peptides	Promiscuity based on no. of alleles binding (Peptide considered as binder if it binds to at least 50% of the 26 most frequent alleles)

Start Over | Back | Next

Current selections:	
No. of sequences	4
Host species	Human
Allele class	Class II
Alleles involved	1. DPA1*01/DPB1*04:01 2. DPA1*01:03/DPB1*02:01 3. DPA1*02:01/DPB1*01:01 4. DPA1*02:01/DPB1*05:01 5. DPA1*03:01/DPB1*04:02 6. DQA1*01:01/DQB1*05:01 7. DQA1*01:02/DQB1*06:02 8. DQA1*03:01/DQB1*03:02 9. DQA1*04:01/DQB1*04:02 10. DQA1*05:01/DQB1*02:01 11. DQA1*05:01/DQB1*03:01 12. DRB1*01:01 13. DRB1*03:01 14. DRB1*04:01 15. DRB1*04:05 16. DRB1*07:01 17. DRB1*08:02 18. DRB1*09:01 19. DRB1*11:01 20. DRB1*12:01 21. DRB1*13:02 22. DRB1*15:01 23. DRB3*01:01 24. DRB3*02:02 25. DRB4*01:01 26. DRB5*01:01
Duplicate peptides	Removed
Peptide overlap	10 AA residues
Approx no. of peptides included	201

Figure 11.prediction method and cutoff value

4.2.1.6 Review of all the selection and Enter Antigen name

In this step we have to review selection, enter Antigen name in job detail& submit data.

No. of sequences	1
Host species	Human
Allele class	Class II
Alleles	1. DPA1*01/DPB1*04:01 2. DPA1*01:03/DPB1*02:01 3. DPA1*02:01/DPB1*01:01 4. DPA1*02:01/DPB1*05:01 5. DPA1*03:01/DPB1*04:02 6. DQA1*01:01/DQB1*05:01 7. DQA1*01:02/DQB1*06:02 8. DQA1*03:01/DQB1*03:02 9. DQA1*04:01/DQB1*04:02 10. DQA1*05:01/DQB1*02:01 11. DQA1*05:01/DQB1*03:01 12. DRB1*01:01 13. DRB1*03:01 14. DRB1*04:01 15. DRB1*04:05 16. DRB1*07:01 17. DRB1*08:02 18. DRB1*09:01 19. DRB1*11:01 20. DRB1*12:01 21. DRB1*13:02 22. DRB1*15:01 23. DRB3*01:01 24. DRB3*02:02 25. DRB4*01:01 26. DRB5*01:01
Duplicate peptides	Removed
Peptide lengths selected	15mers (Only one length for class II)
Approx no. of peptides included	145
Peptide overlap	10 AA residues
Conservancy analysis	No
Prediction method	IEDB recommended

Peptide selection criterion	Promiscuity method
Job details:	
Job name (optional)	
Email (optional - will notify when job is finished)	
Start Over Back	

Figure 12. Review all selection and job details

4.2.2 Class II immunogenicity prediction

Specify Sequence(s)	
Enter epitope sequence(s) in PLAIN or FASTA format	
Or upload epitope sequence(s) from a file	
Choose a prediction method	
Prediction method:	
Specify Output	
Sort Peptides by:	
Select maximum percentile rank threshold:	
Enter the Job Name (Optional)	
Email address (optional)	

Figure 13 CD4+ Immunogenicity prediction

4.2.3 Class II population coverage results

Table 7 VP1 Class II population coverage results

EPITOPE LIST	POPULATION COVERAGE
Epitope#1: GTLTPLSTK	60.98%

Table 8 VP2 class II population coverage results

EPITOPE LIST	POPULATION COVERAGE
Epitope#1: HLKMIAYD	55.23%

Table 9 VP3 class II population coverage results

EPITOPE	POPULATION COVERAGE
Epitope#1: YLTADNFQS	45.87%
Epitope#3: SCTMVVPWI	50.47%

Table 10 VP4 class 11 population coverage results

EPITOPE	POPULATION COVERAGE
Epitope#1:DFSQDPSKFTEPIKD	63.32%

4.3. Combined population coverage results of MHC class l and class ll epitopes

The cumulative Population protection coverage of class I epitopes and class II epitopes was calculating using the IEDB PC tool. The epitopes with the combination of HLA alleles is selected so that the population coverage at least >80% achieved. So different combination of epitopes are test for the achievement of highest population coverage.

4.3.1 VP1's selected epitopes

Table 11 VP1 MHC class 1 and class 11 epitope's population coverage

EPITOPE LIST	POPULATION COVERAGE
Epitope#1: SSIESFFAR	42.29%
Epitope#2: ESMIDNTVR	35.68%
Epitope#4: GTLTPLSTK	60.98%
Epitope set	**90.50%**

4.4.2 VP2's selected epitope

Table 12 VP2 MHC class I and class II epitope's population coverage

EPITOPE LIST	POPULATION COVERAGE
Epitope#1: SLNDFGILA	42.47%
Epitope#2: HLKMIAYDA	55.23%
Epitope set	**80.76%**

4.4.3. VP3's selected epitopes

Table 13 VP4 MHC class I and class II epitope's population coverage

EPITOPE LIST	POPULATION COVERAGE
Epitope#1: ISKFIDWLK	18.96%
Epitope#2: KHIRVWCPR	36.98%
Epitope#6: YLTADNFQS	45.87%
Epitope#7: SCTMVVPWI	50.47%
Epitope set	**85.32%**

4.4.4. VP4's selected epitopes

Table14 VP4 MHC class 1 and class 11 epitope's population prediction

EPITOPE LIST	POPULATION COVERAGE
Epitope#1: PSKFTEPIK	6.84%
Epitope#4: DFSQDPSKFTEPIKD	63.92%
Epitope set	**80.32%**

The immunogenic peptides - MHC class I complexes are important and are presented on nucleated cells which are detected by the cytotoxic CD8+ T cells. This class of proteins are used for presenting peptides which are present on cell surface for the recognition of T cells. The pattern of binding is composed of two domains, which will originate from the single heavy alpha chain and in the case of MHC II there are two chains alpha and beta chains. More scientifically B and T cells will identify the portions within their antigens which are known as epitopes. The identification of epitopes is time taking and money consuming because it requires screening of experiments of large assays of potential epitope candidates. Here the aspects of antigen presentation by T and B cells which are useful for epitope prediction.

5. Conclusion

Live attenuated oral polio vaccine and inactivated polio vaccine were the tools which are being used to achieve the eradication of wild polio virus. There are four serotypes of polio where the VP2 is announced to be eradicated which was detected in India in 1999. VP3 is about to eradicate which have no reported cases since November 2012.

In the whole work it has been tried to some extent that a new vaccine can be design so that the reversal in the polio can be stopped by using the various computational tools. The vaccine would be effective, safe and have proved to be frequent in eliminating the disease from most parts of the world. The tools which are used prove to be good and will enhance the process of achieving and getting eradication in developing countries.

TABLE 15 COMBINED RESULTS

PEPTIDES	**POPULATION COVERAGE**
SSIESFFAR	42.29%
ESMIDNTVR	35.68%
SLNDFGILA	42.47%
ISKFIDWLK	18.98%
KJIRVWCPR	36.98%
PSYFTEPIK	6.89%
GTLTPLSTK	60.98%
HLKMIAYD	55.23%
YLTADNFQS	45.87%
SCTMVVPWI	50.47%
DFSQDPSKFTEPIKD	63.32%

REFERENCES

1. Plotkin SA, Vidor E. 2004. Poliovirus vaccine: inactivated. See Ref. 151a, pp. 625–49

2. Melnick JL. 1978. Advantages and disadvantages of the killed and live poliomyelitis vaccines. Bull. WHO 56:21–38

3. Hammon WM, Coriell LL, Wehrle PF, Stokes J Jr. 1953.Evaluation of Red Cross gamma globulin as a prophylactic agent for poliomyelitis. IV. Final report of results based on clinical diagnoses. JAMA 151:1272–85

4. Delbr ¨ uck M. 1958. Bacteriophage genetics. In Poliomyelitis: Papers and Discussion Presented at the Fourth International Poliomyelitis Conference, ed. IP Conference, pp. 342–53. Philadelphia: Lippincott

5.Racaniello VR, Baltimore D. 1981. Cloned poliovirus complementaryDNAis infectious in mammalian cells. Science 214:916–19

6.Nomoto A, Omata T, Toyoda H, Kuge S, Horie H, et al. 1982. Complete nucleotide sequence of the attenuated poliovirus Sabin 1 strain genome. Proc. Natl. Acad. Sci. USA 79:5793–97

7.Bouchard MJ, Lam DH, Racaniello VR. 1995. Determinants of attenuation and temperature sensitivity in the type 1 poliovirus Sabin strain. J. Virol. 69:4972–78

8.Kawamura N, Kohara M, Abe S, Komatsu T, Tago K, et al. 1989. Determinants in the 5 noncoding region of poliovirus Sabin 1 RNA that influence the attenuation phenotype. J. Virol. 63:1302–9.

9.Rezapkin GV, Fan L, Asher DM, Fibi MR, Dragunsky EM, Chumakov KM. 1999. Mutations in Sabin 2 strain of poliovirus and stability of attenuated phenotype. Virology 258:152–60

10.Yoshida H, Horie H, Matsuura K, Kitamura T, Hashizume S, Miyamura T. 2002. Prevalence of vaccine-derived polioviruses in the environment. J. Gen. Virol. 83:1107–11

11.Minor PD, Dunn G, Evans DMA, Magrath DI, John A, et al. 1989. The temperature sensitivity of the Sabin type 3 vaccine strain of poliovirus: molecular and structural effects of a mutation in the capsid protein VP3. J. Gen. Virol. 70:1117–23

12.Patriarca PA, Sutter RW, Oostvogel PM. 1997. Outbreaks of paralytic poliomyelitis, 1976–1995. J. Infect. Dis. 175(Suppl. 1):S165–72
13.Gromeier M, Alexander L,Wimmer E. 1996. Internal ribosomal entry site substitution eliminatesneurovirulence in intergeneric poliovirus recombinants. Proc. Natl. Acad. Sci. USA 93:2370–75
14.Sever J. 2001. Rotary International: a partner in polio eradication. Dev. Biol. 105:105–8
15.WHO. 2002. Progress towards the global eradication of poliomyelitis, 2001. Wkly.Epidemiol. Rec. 77:98–107
16. Kew OM, Pallansch MA. 2002. The mechanism of polio eradication. See Ref. 175a, pp. 481–91
17.WHO. 2002. Expanding contributions of the Global Laboratory Network for Poliomyelitis Eradication. Wkly. Epidemiol. Rec. 77:133–37
18. Khetsuriani N, Prevots DR, Quick L, Elder ME, Pallansch M, et al. 2003. Persistence of vaccine-derived polioviruses among immunodeficient persons with vaccine-associated paralytic poliomyelitis. J. Infect. Dis. 188:1845–52
19.Yang C-F, Naguib T, Yang S-J, Nasr E, Jorba J, et al. 2003. Circulation of endemic type 2 vaccine-derived poliovirus in Egypt, 1983 to 1993. J. Virol. 77:8366–77
20.Minor PD, Almond JW. 2002. Poliovirus vaccines: molecular biology and immune response. See Ref. 175a, pp. 381–90
21.Halsey NA, Pinto J, Espinosa-Rosales F, Faure-Fontenia MA, da Silva E, et al. 2004. Search for poliovirus carriers in persons with primary immune deficiency diseases in the United States, Mexico, Brazil, and the United Kingdom. Bull. WHO 82:3–8
22. Hull HF,Ward NA, Milstien JB, de Quadros C. 1994. Paralytic poliomyelitis: seasoned strategies, disappearing disease. Lancet 343:1331–37
23.Dowdle WR, Birmingham ME. 1997. The biologic principles of poliovirus eradication. J. Infect. Dis. 175(Suppl. 1):S286–92
24.Fine PEM, Carneiro IAM. 1999. Transmissibility and persistence of oral polio vaccine viruses: implications for the global poliomyelitis eradication initiative. Am. J. Epidemiol. 150:1001–21
25.Fine PEM, Oblapenko G, Sutter RW. 2004. Polio control after certification: major issues outstanding. Bull. WHO 82:47–52

26.MacLennan C, Dunn G, Huissoon AP, Kumararatne DS, Mart´ın J, et al. 2004. Failure to clear persistent vaccine-derived neurovirulent poliovirus infection in an immunodeficient man. Lancet 363:1509–13

27. De Jesus NH (2007). "Epidemics to eradication: the modern history of poliomyelitis". *Virol. J.* **4** (1): 70. doi:10.1186/1743-422X-4-70. PMC 1947962. PMID 17623069.