JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

IDENTIFICATION AND COMPARISON OF POSSIBLE EPITOPE – DESIGNED TARGETS USING IN-SILICO TECHNIQUES FOR CORONA VIRUS

[1] M. Nithya, [2] Dr. Horne Iona Averal

[1] Research Scholar, [2]Associate Professor and IQAC Coordinator

Department of Zoology

Holy Cross College (Autonomous), Tiruchirappalli-620 002

Abstract: The SARS Coronavirus-2 (SARS-CoV-2) epidemic has become a global issue that has raised concerns for the scientific community to design and find a way to combat this deadly virus. To date, the epidemic has claimed hundreds of thousands of lives as a result of infection and spread. Growing evidence suggests that T cells may play a key role in the fight against acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Therefore, COVID-19 vaccines that can obtain a strong T cell response may be very important. The design, development and evaluation of vaccine trials help to understand the T cell epitopes of SARS-CoV-2, which is less well known. Because of the challenges of diagnosing epitopes by experimentation, many studies have suggested the use of in-silico methods. Here, we present of the in-silico methods used to predict SARS-CoV-2 T cell epitopes. These methods use a different set of technical methods, which often focus on machine learning. Functional comparisons are based on the diagnostic power of a specific set of immunogenic epitopes determined by experiments targeted T cells in recovering COVID-19 patients, highlighting the relative functional relevance of the various methods adopted by in - Silico studies. The investigation also prioritizes ideas for future research guidelines.

KEYWORDS: Coronavirus, COVID-19, SARS-CoV-2, Epitope based techniques, In- silico techniques, MHC prediction, Bioinformatics

1.INTRODUCTION

There are extensive studies being conducted around the world to develop appropriate therapies to control the effects of Severe acute respiratory syndrome coronavirus 2 (SARS-cov-2) that causes COVID-19 in humans. The first patient infected with SARS-cov-2 was diagnosed in December 2019 in Wuhan, China (Guo et al., 2019). Later, the virus spread to 187 countries and regions due to its high quality and infected 10,710,005 people as of July 3, 2020 with a total death of 517,877 (COVID-2019, 2020).

Coronaviruses belong to the Family, coronaviridae, and a bunch of Nidovirales is a big family presented with positive strains Coronaviruses are zoonotic pathogens that infect both animals and Humans, which can cause diseases of the intestines, liver, respiratory system and central nervous system.

The World Health Organization (WHO) has declared the outbreak of SARS-cov-2 as a global public health emergency of international concern (PHEIC) on January 30, 2020, and the epidemic on March 11, 2020 (Jeyaprakash et al., 2020). The unavailability of drugs or medically proven vaccines is a major concern of the

COVID-19 epidemic (Zhou et al., 2020). Therefore, effective measures such as early detection of people with the virus, isolation or isolation, isolation from society, the use of masks and hand sanitizers, etc. Are being taken to Fig.ht the epidemic.

It is established that is hiding under the known coronavirus cov-229e (alpha-coronavirus and NL63 Coronavirus (Alpha-coronavirus and OC43 coronavirus (Beta-version of coronavirus and coronavirus-HKU1 (Beta-version of coronavirus and coronavirus – severe acute respiratory syndrome (SARS-COV), Middle east respiratory syndrome Coronavirus (MERS-COV) and Acute Respiratory Infections Viral Infections-cov-2, which can affect people (wu et al., 2020).

In the Sars-cov-2 Genome, the 5'-end Polyprotein (Pp1ab) codes and then cleaves into 15-protein Structures (Nsp-1-a, NSP-10, NSP-12-NSP-16), And 3 - for the latter codes Four Structural Proteins, including, Spike (S), Coating (E), Membrane Surface (M) and Nucleocapsid (N) Protein, as well as six additional proteins (3a, 3b) 6, (7a, 7b), 8a, 9, and Orf 14) (Wu et al., 2020). Comparative Genomic Analysis Showed that SARS-COV-2 is larger than the number of Nucleotides that SARS-COV is larger than MERS-COV (Xu et al., 2020); (Zheng et al., 2020).

The virus is mainly composed of a membrane and nucleocapsid. The membrane contains three proteins: (1) spike (S) (2) membrane protein (M) and (3) envelope protein (E) (Weiss and Navas *et al.,* 2005). The virus controls the cell by initially binding to the angiotensin converting enzyme II (ACE2) in the alveoli of the lungs. As a result, the infected person suffers from respiratory problems (Babadaei *et al.,* 2020).

In Addition, it was found that SARS-COV-2 is associated with the human receptor ace2 higher affinity when SARS-COV and only a few of them in the framework of polybasic the collapse of the SARS-COV-2, to increase the infectious power of the virus (Boni et al., 2020);

(Ye, ZW et al.,2020). It has also been observed that in RDO, SARS-COV-2 binds to the human ace2 receptor with higher affinity than SARS-COV, as well as being polybasic in the sars-cov-2 cleavage region, increasing viral transmission (Boni et al., 2020); (Hans Christian Andersen et al., 2020); (Z. U. et al., 2020).

The SARS-CoV-2 life cycle in the host cells Every life cycle involves three steps. Inflammation (endocytosis, fusion) Dissociation Similarly, COVID-19 begins its life cycle when S-protein binds to the cellular receptor ACE2 and cell membranes using the endosomal method. SARS-CoV-2 releases RNA from the host cell Genome RNA is translated into viral replicase polyprotein ppl a and lab, which is then broken down into smaller products by viral proteins. The polymerase produces a series of subgenomic mRNAs continuously transcribed and eventually translated into the appropriate viral proteins and the genome RNA is sequentially synthesized into virions in the ER and Golgi and transported by vesicles and excreted outside the cell. All pathology pathogenicity I -SARS-CoV-2, from attachment to replication (Maya Datt Joshi *et al., 2021).*

In view of the critical need to develop safe, more effective ways to combat SARS-CoV-2 it has undergone some clinical trials worldwide .There is no doubt that any of the vaccines produced can be very important to the potential outbreaks and re-emergence of the season that relies heavily on the evolution of long-term defense. The advancement of MERS-CoV and SARS-CoV-1 vaccines over the years are important keys in view of its genetic similarity which provides important insights into the development of the SARS-CoV-2 vaccine (Moreno-Fierros *et al.,* 2020; Barnard et al., 2007; Yang *et al.,*2020; Chakraborty *et al.,*2020).

Therefore, multiple platforms have been developed since its emerging, including DNA-based and RNA-based forums and recombinant-subunit vaccines. However, the development of the SARS-CoV-2 vaccine poses some challenges and even new forums. For example, pre-clinical studies of SARS vaccine participants and MERS have raised concerns about the spread of lung disease as a result of the development of an autoimmune dependence or a direct effect. Therefore, testing with a suitable animal model and strict safety precautions in clinical trials will be critical (Amanat *et al.,* 2020; Zhang *et al.,* 2020)

Traditional methods of vaccination based on laboratory tests in the event of an outbreak could not meet emergency needs, and many medical agents are being investigated. Bioinformatics research is a robust tool identified for sorting, organizing, and processing large amounts of available data generated from other experiments to provide a large field of immunology over a limited period of time. As the virus genome and protein sequence information are available, presented epitopes and viral traits can be predicted by silico analysis, which greatly accelerates progress in vaccine development Chakraborty *et al.,*2020; Bhattacharya *et al.,*2020; Chen *et al.,*2020; Kiyotani *et al.,*2020).

The new SARS-COV-2 coronavirus that causes COVID-19 was recorded in Wuhan, China's Hubei province in December 2019, and since then COVID-19 has spread from China to 211 different countries, to more than 28 million people, including more than 900,000 deaths, according to the world health organization (WHO 2020).

Approaches are used to develop vaccines or recombinant vaccines. Traditional methods are based on inactivated or live attenuated computer threats that can be used for vaccine development, but it has been noted that these methods have some limitations, such as Labour intensity, as well as problems in producing many proteins and pathogens (Dangi et al.,2018).

This provision is aimed at preventing the development of new vaccines against the pathogen causing the outbreak, which leads to a pandemic. On the other hand, to solve this problem, when preparing a recombinant vaccine, all genes obtained from various pathogens were cloned, expressed and purified for use as a vaccine, which can be applied (Nascimento et al.,2012).

To create a new recombinant vaccine design, otherwise Vaccinology is an out-of-silicon approach that provides detailed predictions of the vaccine's genome and sequence, which may ultimately be the protein. The use of silicon preparation method is very important because it will allow predicting antigenicity, epitope regions on T cells, and other parameters such as peptide, subcellular localization, and salicylic acid in the target protein (Dangi et al., 2018; May et al., 2020). For this analysis, the unifying force between the predicted epitopes of the selected MHC-I and MHC-II genes is an important part of obtaining a silicon approach (Dangi et al.,2018). At present, insufficient knowledge of latency and SARS-CoV-2 infection increases the uncertainty of viral persistence. Some antiretroviral drugs are currently not available.

Vaccination is still a cost-effective and effective way to prevent infection. The selection and design of immunosuppressive immunogens is a major challenge in vaccine development, especially in newly developed viruses (Qiu *et al.,* 2018); (Chakraborthy *et al., 2020).* Traditional methods based on laboratory tests have not been able to meet the needs of a stressful situation in the event of violence (Rauch *et al., 2018).*

SARS-CoV-2 and SARS-CoV were found to be related which suggests that their biochemical interactions may be similar (Chen et al., 2020). Like SARS-CoV, SARS- CoV-2 invades cells carrying the ACE2 receptors present in the pneumocyte II type in the lungs. After entering the first line of the immune system such as mucus and molten cells, the PAMPS-associated cell patterns present in the virus show antibodies such as alveolar macrophages, neutrophils, monocytes and natural killer cells (NK cells) to Fig.ht off an infectious pathogen (Hoffmann et al., 2020).

Many studies suggest that, as a result of SARS-CoV 2 infection, there is an increase in the number of antibody-producing cells (ASCs). IgA and IgM appear approximately 5 days (ranging from between 2 to 6 days) while 14 days in the case of IgG (from 10 to 18 days) after symptoms have commenced. Antibodies against NP (nucleoprotein) and RBD (receptor binding domain) have a neutralizing function. In diseases such as SARS-CoV-2, various isotypes of the immune system are stimulated to show a high humorous response.

SARS-CoV-2 has been identified as a new species from the 2B Coronaviruses group, with almost 70% genetic similarity to SARS-CoV, since the 2003 outbreak (WHO). The virus has a 96% similarity to bat coronavirus, so it is widely suspected to be derived from bats (Cohen et al., 2020; Eschner et al., 2020). The epidemic has led to travel restrictions and national closures in several countries and led to economic damage (Hui et al., 2019). Genome sequencing is included in many online biorepositories selected to find solutions and to make the vaccine available worldwide is important.

ACE2, a homolog of angiotensin-converting enzyme (ACE), expressed in various organs and tissues, has many biological functions and can counteract the negative role of the renin-angiotensin (RAS) system in many diseases (Donoghue et al., 2000; Patal et al., 2017; santos et al., 2018). Given that SARS-CoV-2 spike proteins interact with ACE2, as does SARS-CoV, COVID-19 may have a pathogenic mechanism similar to that of SARS. In this review, we will discuss the role of ACE2 in COVID-19, as well as its potential therapeutic goals, which are intended to provide additional information on the management of the epidemic. Infiltration of participating cells is the first step in infection. Spike glycoprotein in the coronavirus virus envelope can bind to specific receptors in the host cell membrane.

In particular, the binding relationship between SARS-CoV-2 and ACE2 has been found to be 10 to 20 times higher than that between SARS-CoV and ACE2 (Wrapp et al., 2019). In addition, the ACE2 receptor expression in human tissues suggested that SARS-CoV-2 could invade or damage the digestive tract (Zhang et al., 2020) and reproductive organs (Fan et al., 2020). Therefore, cells and tissues expressing ACE2 may serve as targets for SARS-CoV-2 infection.

In this study, we examined the ACE2 expression pattern in the mouse and human heart using social data. In addition, we used bioinformatics analysis combined with the structure to determine the underlying mechanism of cardiovascular function and myocardial injury in patients with COVID-19. The findings revealed that ACE2 is not only a SARS-CoV-2 receptor but also an important component in the pathogenesis of COVID-19-related cardiac injury. We highlighted the need for doctors to pay attention to heart problems related to SARS-CoV-2

and provide new clinical clinical evidence.

Patients with severe COVID-19 have an additional IgG titer indicating significant antibody-dependent enhancement (ADE) developmental function in this disease (Cao et al., 2020). However, weakened immune systems weakened and internal responses to uncontrolled inflammation cause systemic and spinal tissue damage. Since antibodies present in COVID-19 patients are highly involved in viral production, convalescent plasma of COVID-19 survivors or monoclonal antibodies may be recommended in clinical management of the disease.

Sexual hormones such as estrogen and testosterone affect the immune response to both sexes as it leads to a decrease in male immune response and subsequent increase in respiratory infections and death. In women, the X chromosome contains a very high number of genes related to the immune system in their genome, which may also lead to a higher immune response in women (Schurz et al., 2019). One interesting fact is that the ACE2 gene is found in the X chromosome. ACE2 receptors in the pulmonary endothelium are the entry point for the corona virus. In vivo studies have shown that the male hormone testosterone increases the activity of ACE2 (Patel et al., 2014).

The mortality rate of severe COVID-19 patients could be as high as 49 percent, as demonstrated by a recent epidemiological report by China (Wang et al., 2020). The majority of patients with COVID-19 were elderly patients in a complex group with basic illnesses. Chronic obstructive pulmonary disease, high blood pressure, malignant tumors, heart disease and chronic kidney disease were more common in the critical group than in the small group. Patients over the age of 60 and chronic diseases, especially high blood pressure are more likely to be SARS-CoV-2 (Yin et al., 2020).

Patients with diabetes, especially those with uncontrolled glycemia are at greater risk of developing COVID-19 (Abu et al., 2018). High incidence of urinary tract infections, foot infections and surgical infections are often present in diabetic patients (Critchley et al.., 2018). In diabetes, the release of cytokines by macrophages and T-cells is altered and impairs neutrophil activation leading to a weakened immune response (Smith et al., 2009).

Traditional vaccines, based on biochemical chemical experiments, create strong neutral and protective responses in vaccinated animals, which can be expensive, allergenic, and time-consuming and require an in vitro culture of pathogenic bacteria leading to serious safety concerns (Y et al.,). 2014; w et al., 2013)]. Therefore, the need for safe and effective vaccines is highly recommended.

Peptide-based vaccines do not require an in vitro culture that makes them biologically safe, and their selection allows for the precise functioning of the immune response (A. W. Purcell et al., 2007; N. L. Dudek et al., 2010). The core pathway of peptide vaccines is built on a chemical pathway to combine known B-cell epitopes and T-cells that are immunodominant and can trigger specific immune responses. The B-cell epitope of the target molecule can be linked to the T-cell epitope to make it immunogenic. T-cell epitopes are shorter peptide fragments (8–20 amino acids), while B-cell epitopes may be proteins (S. Dermime et al., 2004; H. Meloen et al., 2001). Therefore, in this study, we aimed to design a peptide-based vaccine to predict epitopes from the corona envelope protein (E) using immunoinformatics analysis (V. Brusic et al., 2014; N. Tomar et al. , 2014; S. Khalili et al., 2017; NR Hegde et al., 2017). Further rapid research is recommended to confirm the efficacy of the reported epitopes as a peptide vaccine against these emerging infections.

Symptoms of COVID-19 can range from mild to severe including cough, fever and shortness of breath. Most people have asymptomatic. Symptoms may appear two to fourteen days after exposure. About 20% progress to serious illness, such as respiratory failure, pneumonia and death in some cases . Many patients appear to have less illness (Gu et al., 2020; Kooraki et al., 2020). Other symptoms may include runny nose, vomiting, fatigue, sore throat, diarrhea some people may lose their sense of smell or taste (Lai et al., 2020). Current information has shown that it spreads from person to person among those close to a distance of 6 meters, or 2 meters. The virus is spread through respiratory droplets when an infected person coughs or sneezes (Peeri et al., 2020).

The WHO and the CDC have recommended various safety measures to prevent COVID-19 infections in many healthy people. Some of these are:

- Avoid big events and large gatherings
- Avoid close association with anyone
- Wash your hands often with soap and water for at least 20 seconds or use a 60 percent alcohol-based disinfectant.
- If you cough or sneeze, cover your mouth and nose with your elbow or tissue.
- Used tissues should be disposed of properly
- Avoid touching your eyes, nose and mouth (Cascella et al., 2020).

In the case of gender, COVID 19 diseases are equally seen in both the apparent gender differences and

the risk of disease. In this regard, it is suggested that men are more likely to die than women due to differences in sexual immunology and gender differences such as the prevalence of smoking (Liu et al., 2016).

There are Various subjects in which there are many explanations regarding bioinformatics has been suggested. In accordance with lopresti, 2008, bioinformatics use of techniques form computer science to emerging problems biology (Lopresti *et al.*,2008). Bioinformatics refer to the field you are dealing with collection and storage of biological information. It all matters related to the biological site is called bioinformatics(rana *et al., 2012).*

Unlike the most dangerous human coronaviruses that continue to circulate in humans, animal-borne coronaviruses can be deadly viruses by crossing the boundaries of species. SARS-CoV in 2003, MERS-CoV since 2012, and SARS-CoV-2 currently all caused major epidemic problems. Effective and economic measures are urgently needed in the current context of the epidemic.

Compared with the development of traditional vaccine, stronger epitopers can be predicted by bioinformatics analysis, which makes policy design more straightforward and faster (Rappuoli R et al., 2018). Since most spike proteins are produced outside of the virion, it may be a suitable target for B-cell epitopes. Spike proteins in MERS-CoV and SARS-CoV have been shown to create a strong immune response (Zhou Y et al., 2018; Du L et al., 2008) Barah and Bose studies have also shown that epitopi in spike protein can be promising people. with the development of the SARS-CoV-2 vaccine (Bhattacharya M et al., 2018). In another study, experimental epitopes taken from SARS-CoV were tested and those with similar sequences in SARS-CoV-2 were identified (Ahmed SF et al., 2020). A similar strategy has been followed in Grifoni research (Grifoni A et al., 2020). Although the SARS-CoV-2 spike protein showed 77.38% sequence ownership in that of SARS-CoV, most antibodies against SARS-CoV spike proteins indicate a negative affinity for SARS-CoV-2 (Tian X et al ., 2020), indicating that spike proteins have a highly variable structure. Therefore, we used a different strategy and tested the localization of the predicted B cell epitopes to extract those buried within the protein. Although we have used the same tools and resources for bioinformatics, the peptides we have chosen for vaccine development are not the same. In our analysis of the SARS-CoV-2 spike protein, four peptides identified in multiple step tests showed excellent local accessibility.

Outbreaks of infectious diseases such as COVID-19 pose a serious challenge to the scientific community as they often occur in unknown zoonotic sources or due to a lack of data. Viruses can develop from one type of animal to another that can infect humans by acquiring their own receptors and biosynthetic machinery. Most newly developed viruses are difficult to treat due to a lack of specific treatment options (Coleman et al., 2014). To date, there is no vaccine for SARS-CoV, MERS-CoV, or SARS-CoV-2 currently on the market, although some clinical trials are still ongoing (Lee P-I et al., 2019).

Bioinformatics is a computer tool for analyzing biological data. It is a branch of Biology, Physics, Computer science and Mathematics. Bioinformatics is important in medical data management and biology. It is a solution to biological problems based on existing research results. Provides a way to store all biological data. By predicting the outcome of a laboratory experience, it makes certain laboratory tests easier. Computers have become an integral part of biology. They contribute to the management of large and expanded amounts of biological data and play a key role in testing new biological connections.

Bioinformatics is a multidisciplinary field that specializes in organizing, storing, and processing large amounts of data generated in biological experiments. The collection of high-level immunological data has resulted in a field known as immunoinformatics, which provides information about the body's processes. As genome information and the sequence of SARS-CoV-2 proteins become available, viral characteristics, and epitopes introduced into the pathogen, can be predicted by silico analysis, which will greatly accelerate vaccine development (U1Qumar *et al.,*2018 ; Ahmad *et al., 2019*; U1Qumar *et al.,2019).* predicting B-cell epitopes in spike protein and T-cell epitopes within the nucleocapsid protein of SARS-CoV-2 using bioinformatics methods and immunoinformatics tools.

Silicon analysis and its use for analysing the influence of physicochemical properties, antigenicity, subcellular localization of viral proteins, predicting the t-cell epitope, predicting the MHC-I and MHC-II Epitopes, predicting the effect of solvent on the location of epitopes, finally, the effects of changes often occur, and structural protein that have been predicted to have only one peptide per antigenicity. Epitopes have been studied for comparison in various organisms such as Ghost Shark, Zebra finch, Mouse, Polar Bear and Human.

TOOLS AND SOFTWARE

There are many tools involved in these studies

FASTA stands for FAST homology search All sequences. The protein sequencing alignment system was created by Pearsin and Lipman in 1988. The system is one of the many heuristic algorithms proposed to speed

up comparisons. The basic idea is to add a quick pre-screen step to get the most similar segments between two sequences, and then extend these same segments into local alignment using more robust algorithms like Smith-Waterman. FASTA can be very specific when identifying long regions of low similarity especially in very different sequences. You can also perform sequential search comparisons with a nucleotide site or fill in proteome / genome information using FASTA programs.

ANALYSIS OF PHYSIOCHEMICAL PROPERTIES

Vaccines do not have to be on the Physico-chemical properties of an important factor, as do vaccines interacting with the environment (Iwasaki *et al*., 2020). Protparum is a tool that allows you to calculate various physical and chemical parameters for a given protein, stored in swiss-prot or trembl, or by providing the user with a protein sequence. Calculated parameters include molecular weight, Theoretical PI, amino acid composition, nuclear composition, extinction coefficient, estimated period and a half life, instability index, aliphatic index, and average cost of hydropathicity (SOUS) (Gasteiger *et al*., 2005).

PREDICTION THE LOCALISATION AND NUMBER OF TRANSMEMBRANE HELICES.

The transmembrane helix is part of the 17-25 amino acid proteins associated with the structure of the α - helix across the membrane cell. most of the time, candidates- vaccine are expressed in biological systems that are different from the original source; In the middle if so, 3-dimensional (3d) protein structure can be modified or defined vomiting if it has a transmembrane helix, due to fractures in membrane formation.

The low number of transmembrane Helix is a major factor in vaccine candidate selection choice. According to the etiology of the disease being studied, the protein for cellular production, adhesin properties, antigenicity, lack of human proteinology to prevent the formation of a potential autoimmune response, and low- or non-transmissible structures of the transmembrane helix are the most common properties to be identified (Meunier *et al*., 2016).

The sequence and structure of the motif of the polypeptide chains that need to be classified by comparative analysis that make up the protein conserved domain. It is through molecular evolution that the building blocks, different shapes can undergo organization and production, make protein production of different functions.

CONSERVED DOMAINS DATABASE (CDD)

Significance of the evolution domain note factors brought to us by the 2019-ncov major protease epitope protection degree assessment. This was done using the conservative domain database (CDD). The conserved area of proteins includes 29 amino acid sequences along with that of the predicted epitope sequences (Marchler *et al*., 2017).

PREDICTION OF T-CELL EPITOPES WITH HLA

Peptide prediction of potential T-cell epitopes schedule-MHC I selected the best proximity binds with confidence 1/0. 89, and lest ic50 for each allele, and in the same MHCpred used an additive method to predict forced major histocompatibility complex (MHC) Class I (HLA- (A*) And II (DRB*) Molecules and Also Transporter Associated with Processing (TAP). The allele-specific quantitative structure-activity relationship (QSAR) model was are found, with partial least squares (pls) method (Kobayashi H *et al*., 2000).

PREDICTION OF B CELL EPITOPES

B Epitopes plays an important role in the development of epitope-based vaccines and the surrounding research area. the dominant B- cell line epitope, which can be used in the treatment of autoimmune diseases, in which the goal is a neutralizing antibody response it can also induce antibodies that cross-react with the parent's protein (Saha *et al*.,2006).

PREDICTION OF MHC-I AND MHC-II EPITOPES

B-cell Epitopes are part of an antigen for induced immunoglobulin or antibody, and in order to activate single-cells to give the immune system a response (Sanchez-trincado, *et al*., 2017). T cells recognize epitopes when these are presented to them bound to mhc molecules. therefore, epitopes can be predicted by computing Their MHC-binding Profile. because of the differences in the molecular interactions between epitopes and MHC I And II Complexes, the prediction of epitopes binding to MHC I is more accurate than to MHC II . for both types, we used IEDB tools as detailed in materials and methods (Sanchez-trincado, *et al*., 2017).

Genetic sequences make it possible to detect coded protein. Bioinformatics tools can be used to find the key regions for these proteins. Epitopes are regions identified by cells T and B. Immunoinformatics developed as a branch of bioinformatics responsible for predicting T and B cell epitopes that can be used for vaccine formulation (Slathia, P.S et al., 2018). There is a lot of research on vaccine formulation using this biological method; virus (West Nile virus and Japanese Encephalitis virus) (Slathia, PS et al., 2019), parasites (Trypanosoma cruzi) (Slathia, PS et al., 2018), viruses (Listeria monocytogenes) (Jahangiri, A ., et al.., 2011). B epitopes of B cells

can also be used in diagnostic construction where epitopic regions can be recorded in their own immune system. These antibodies if they have a high titanium in the patient's serum can be used in the manufacture of serological kits. In the present study, we predicted T and B cell epitopes on the SARS-CoV-2 proteome using different bioinformatics tools that could find use in the development of vaccines and diagnoses.

Antigenic protein regions identified by binding sites or paratope of immunoglobulin molecules called B-cell epitopes. When such a direct binding (between the antigen epitope and the antibody paratope) is detected by experimentation, a specific immunoglobulin establishes the epitope nature of the protein. Epitopes are therefore related entities that can only be defined in a practical sense (i.e. in immunoassay) by binding paratopes parallel (Van Regenmortel MH et al., 1989).

These epitopes play an important role in the development of peptide-based vaccines and in the diagnosis of diseases (Wiesmuller KH et al., 2001; Zauner W et al., 2001; Van Regenmortel et al., 2001). MH 2–4 B-cell epitopes are also important in the study of allergies and in determining the reactivation of IgE epitopes for allergies (Negroni L et al., 1998; Cells I et al., 1999; Clement G et al. , 2002) .These epitopes may be linear (continuous) or parallel (stop). When synthetic peptides are found to interact with anti-protein antibodies or when they are able to induce antibodies that respond to the parent's protein, then these peptides are labeled (continuous) epitopes (Van Regenmortel MH et al., 1993).

B-cell protective epitopes may lead to the development of an effective peptide vaccine against viral infections (Langeveld JP et al., 2001). B-cell regulatory epitope is used as a target to reduce antibody responses in autoimmune diseases (Castelletti D JP et al., 2004) .A continuous or concomitant epitope is formed in a few different, localized sequences. This sequence forms an accessible cohesive region where the protein is folded.

Defining these episodes is a difficult task, but it can provide insight into the basic structure of antigen-antibody recognition (Estienne et al., 2001). Recently, the Conformational epitope prediction server (CEP) was designed to predict conformational epitopes using 3D structural data for antigen proteins (Kulkarni-Kale et al., 2002). Predicting immunogenic epitopes remains an important and challenging task using bioinformatic tools. The natural complexity of antigen recognition makes it difficult to predict epitope (Flower et al., 2001). In the past, a number of algorithms have been developed to predict continuous B-cell epitopes based on physico-chemical amino acids, (Westhof et al., 1991) but their level of effective prediction is not very high.

B-cell epitopes are higher levels of antigen, in which certain antibodies detect and bind, triggering an immune response. This interaction is central to the adaptive immune system, where among other things we are responsible for immunological memory and direct antigen responses in spinal cord animals. The ability to identify these binding sites in antigen sequence or formation is important in the development of synthetic vaccines (EE Hughes et al., 1993; JP Tam et al., 1989; RC Russi et al., 2018), diagnostic tests (GA Schellekens et al., 2000) and immunotherapeutics (AJ Chirino et al., 2000; H. ShiraiI et al., 2002). The focus of these applications, through the epitope acquisition lens, has received attention over the years, especially with regard to the safety benefits of artificial goal development (E.H. Nardin et al., 2001).

Typically, B-cell epitopes are divided into two categories: straight (continuous) epitopes, consisting of linear fossils and fusion epitopes (discontinuous), consisting of fossils that are not consistent with the sequence of the main protein, but delivered. together in the structure of coated proteins (DJ Barlow et al., 1986). In addition, approximately 90 percent of B-cell epitopes are estimated to be consistent and only about 10 percent should be linear (J.L. Pellequer et al., 1991). However, it has been shown that most persistent epitopes contain several groups of progressive residues that reunite in a high protein structure (M.H. Van Regenmortel et al., 2006), making the distinction between them unclear. All of the above-mentioned immunology applications share the need for the detection of all potential epitopes of any antigen, a process called “Epitope mapping”.

Although epitope mapping can be done using a few experimental techniques (U. Reineke et al., 2009), it is time-consuming and expensive, especially on the genomic scale. To solve this problem and access the ever-growing data of epitopes embedded in the biological site on a daily basis, a few calculations to predict parallel or linear B-cell epitopes published decades ago (Y. El-Manzalawy et al., 2010; JL Sanchez- Trincado et al., 2017; N. Tomar et al., 2014). Despite the relatively small percentage of B-cell epitopes, many of the methods developed in the last few years focus on their predictions. This is mainly due to the requirement of the 3D structure of the antigen when it predicts its corresponding epitopes (D.R. Flower et al., 2007). Therefore, in this review we will only discuss the effectiveness of the linear B-cell epitope (BCE) predictions.

The importance of non-peptide epitopes, such as lipids and carbohydrates, has been increasingly recognized, the precise prediction of B-cells and T-cell epitopes (on which modern epitope-based protocols are built) remains central. challenge to informatics with immunology. While the prediction of the B-cell epitope remains absurd (Alix et al., 1999), or relies on commonly defined information of a three-dimensional protein structure (Thornton, JM et al., 1986), many advanced methods T-cell prediction epitopes have emerged (Flower, DR et al., 2002). It is acknowledged that only peptides binding to histocompatibility complexes (MHCs) with a

parallel boundary [usually 500 nM) (Sette, A., et al., 1994)] act as T-cell epitopes and that peptide-MHC affinity is approximately related T-cell response. Many current methods of predicting T-cell epitopes rely on predicting peptides binding to MHCs.

T-cell epitope prediction aims to identify the shortest peptides within the antigen that can stimulate CD4 or CD8 T-cells (R. K. Ahmed et al., 2009). This ability to regenerate T-cells is called immunogenicity, and it is confirmed in experiments that require peptides to be obtained from antigens (T. A. Ahmad et al., 2016; L. Malherbe et al., 2009). There are many different peptides within antigens and T-cell predictive methods are aimed at identifying those that are immune. T-cell epitope immunogenicity depends on three basic steps: (i) antigen processing, (ii) peptide binding to MHC molecules, and (iii) cognate TCR recognition. Of the three cases, MHC-peptide binding is the most preferred in determining T-cell epitopes (EM Lafuente et al., 2009; PE Jensen et al., 2007).

The MHC I and MHC II molecules have similar 3D structures with bounded peptides located in a duct-shaped two α-helice above the ground that comprises eight β-antiparallel fibers. However, there is also a major difference between the binding holes of MHC I and II which we should highlight because it sets the predictor conditions for peptide binding. The peptide-binding cleft of MHC I molecules is blocked as it forms a single α chain. As a result, MHC I molecules can bind only short peptides of 9 to 11 amino acids, their N- and C-terminal terminals remain pinned to the retained residues of the MHC molecule I through a network of hydrogen bonds (LJ Stern et al., 1994; DR Madden et al., 1995). The MHC I peptide binding peptide also contains deep binding packets with strong physicochemical preferences that help bind binding.

Peptides of different sizes and binding to the same MHC I molecule usually use different binding packets (D. R. Madden et al., 1993). Therefore, methods for predicting the binding of peptide-MHC I require a fixed peptide length. However, since most MHC I peptide ligands have 9 residues, it is usually best to predict peptides of that size. In contrast, the groove binding peptide of MHC II molecules is open, allowing N- and C-terminal end of peptide to extend beyond the binding canal (LJ Stern et al., 1994; DR Madden et al., 1995). As a result, the peptides binding to MHC II vary widely in length (residues 9–22), although only the nucleus of the nine residues (the peptide-binding spinal cord) resides in the binding duct of MHC II. Therefore, binding mechanisms of peptide-MHC II targets often target the identification of these peptide binding signals. The binding packets of the MHC II molecule are also shallow and less demanding than those of the MHC I molecules binding peptide to MHC II molecules are less accurate than those of the MHC I molecules.

As with other viruses, T cells act by detecting SARS-CoV-2 peptides (short amino acid sequences) introduced on the surface of infected cells by leukocyte antigen (HLA) molecules. The goal of the COVID-19 vaccine is to mimic this process by rejuvenating SARS-CoV-2 antibodies, thus preparing the immune system to Fig.ht off the virus in the natural environment. A common challenge facing modern vaccination methods is to identify certain SARS-CoV-2 peptides that can deliver a strong and protective T cell response.

This is important for the skin of the vaccine development pipeline above all else as it provides basic recommendations for all the downstream efforts to be followed in vaccine compilation, laboratory tests, and clinical trials. Experimental identification of SARS-CoV-2 peptides that provide T cell responses is difficult, due in part to the large number of potential experimental options, and the high genetic variation of large histocompatibility complex (MHC) codecs of HLA molecules. Although each person has 12 different types of HLA alleles, currently more than 27,000 known HLA allele are listed in the immune polymorphism database (Robinson et al., 2020) and these vary in peptide binding details.

With the availability of large amounts of data related to peptide-HLA binding, several attempts to solve the problem of T cell epitope detection (i.e., predicting peptides to detect T cell response) have suggested the use of this data in silico methods. (V. Jurtz et al., 2017; J.G. Abelin et al., 2017; c, B. Alvarez et al., 2020; J.G. Abelin et al., 2019)for SARS-CoV-2, soon after the first genetic sequence was discovered in January 2020, silico methods began to be used to predict and recommend T cell epitopes as a possible target for the SARS-CoV-2 vaccine.

In addition to guiding the development of vaccines, many of these predictions have been helpful in informing experimental studies aimed at understanding the immune responses that occur naturally in recovering COVID-19 patients. This review discusses the reasons and features of the silico methods and tools used to date in predicting SARS-CoV-2 T cell epitope. As we explain, a different set of composite techniques has been used, usually using mechanical learning methods, and in some cases the expected cross-reactivity of epitopes between genes such as genetics is used.

These silico methods and tools are often developed independently and in many cases have been trained using viral or other virus-related data sets, making it difficult to understand the epitope-related functionality of SARS-CoV-2. To help clarify these questions, this review provides a pre-diction comparison of 61 SARS-CoV-2 in silico studies, showing similarities and differences between certain SARS-CoV-2 epitopes predicted in different ways. We also evaluated and compared predictions using the emerging data from nine experimental studies that identified SARS-CoV-2 T cell epitopes targeted to COVID-19 recovering patients.

By predicting the B-cell epitope, hydrophilicity is an important factor that is common in the beta region. These tests strengthen the chances of people going to be vaccinated.By examining SARS-CoV-2-derived epitopes derived from SARS-CoV-2 immunogenic proteins and by using bioinformamatic tools, we produced a list of potential epitopes for vaccine formulation. Our findings can help reduce the strong targeted demand for an effective vaccine based on the SARS-CoV-2 peptide and help guide development-focused studies to find a vaccine for COVID-19 infection as soon as possible.

AIM AND OBJECTIVES

The main objectives is to

search for ligand using in-silico tools that would elicit response to form antibodies and hence protect individual from the infection.

check various parameters of ligand - targets which will be most possibly used to produce a vaccine.

identify and compare the parameters that will give the possible target. thus, there is a scope by epitope-designing and recombinant DNA technology to study and to produce vaccine models.

II. METHODOLOGY

1.KEGG

Kyoto Encyclopaedia of Genes and Genomes is a database resource that integrates genomic, chemical and systemic functional information. KEGG pathway for Corona infection to observe various stage Entry Keywords: SARS-COV-

Map05171 click

PATHWAY: map05171 and map03230 Search Ligand for Corona virus – type the site https://www.guidetopharmacology.org/GRAC/CoronavirusForward

choose all ligands and target.

KEGG pathway for corona infection to observe various infectious stages and protein molecules/ligands for targeting using https://www.genome.jp

1.COVID-19 disease & 2. SARS-CoV-2 (SARS-CoV-2 sequences Human, MOUSE, POLAR BEAR, GHOST SHARK AND ZEBRA FINCH.SARS coronavirus (SARS-CoV-RaTG13).

COLLECTION OF TARGET PROTEIN SEQUENCE

Collecting Protein Sequences of These Possible Ligands From Uniprot KB Database To Provide The Scientific Community With A Comprehensive, High-quality And Freely Accessible Resource Of Protein Sequence And Functional Information. **Https://Www.Uniprot.Org** .

A MULTIPLE ALIGNMENT SEQUENCING (MSA) USING CLUSTAL OMEGA

Collect 5-7 protein sequences of each protein ligand of both corona virus – human , mouse, Polar Bear, Ghost Shark and zebra fish conduct a multiple alignment sequencing (msa) using clustal omega. **https://www.ebi.ac.uk/Tools/msa/clustalo/**

And View The Result In Jalview 2.11.1.0 To Observed The Conserved Regions In Each Protein For Jalview You Need To Launch Manually, Install Use File > Open URL And Paste.

ANALYSIS OF PHYSIOCHEMICAL PROPERTIES

The reference genome proteins were investigated using expay protparm online server https://web.expasy.org/protparam/. **Https://Doi.Org/10.1385/1-59259-890-0:571** (Humana Press *et al.*, 2005).

PREDICTION THE LOCALISATION AND NUMBER OF TRANSMEMBRANE HELICES.

(Transmembrane Helices [Tmhs] Present & Possible Location Of Peptide In Membrane)

To predict the Number of Transactions, use TMHMM Server V. 2.0 **Http://Www.Cbs.Dtu.Dk/Services/TMHMM/** where they are used. (Krogh *et al.*,2001).

CDD

Conserved Domains Database of Peptides & Within Them, Name of Superfamily, Length Covered & E-value Of Specific Hits Used **https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi** (Marchler-bauer *et al.*,2012

PREDICTION OF T-CELL EPITOPES WITH HLA

Mhcpred (Peptide Prediction to find Possible T-cell Epitopes With HLA MHCI Best Affinity Binding Alleles With Confidence 1/0.89 And One With Least IC50 Value For Each Allele Were Selected And Listed. **http://www.ddg-pharmfac.net/mhcpred/MHCPred/**

PREDICTION OF B CELL EPITOPES.

Linear B Cell Epitopes Of The Reference Genome Structural Proteins, Variant Proteins, And The Proteins That Have A Signal Peptide Were Predicted By Abcpred

http://crdd.osdd.net/raghava/abcpred/ (Saha *et al.*, 2004) (Vita *et al.*,2019).

PREDICTION OF MHC I AND MHC II EPITOPES

E Prediction Of MHC-I And MHC-II Epitopes of the reference genome structural proteins, Variant proteins, and the proteins that have a Signal Peptide Were Analysed By IEDB http://tools.iedb.org/bcell/ (Vita *et al.*,2019).

III RESULT& DISCUSSION

Fig..1. KEGG PATHWAY INFECTIOUS STAGES

Entry **Keywords : SARS-CoV-**

PATHWAY: map05171 and map03230

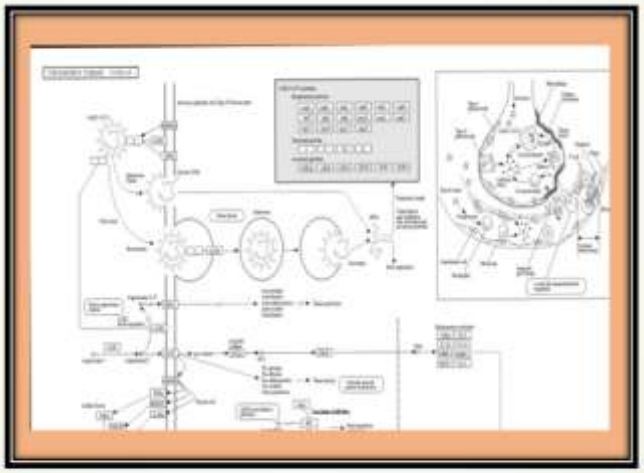


Fig. 2a. ACE2 IN CALLORHINCHUS MILLI (GHOST SHARK)

```
>tr|A0A4W3HYM0|A0A4W3HYM0_CALMI Angiotensin-converting enzyme OS=Callorhinchus milii OX=7868 GN=ace2 PE=3 SV=1
MFLQWLLLLSLAAAALSLSPVEQEATAFLKEFDTKSQDLVYKSSLASWEYNTNITDENIDKMNEESAKWSAFYQQASDDSSKFNINEISDNIIKLQLNSLQDKGSGVLSKEEQSQHIMEVQ
NEMSKIYSTGTVCKPNNPSDCLGLEPGLTILLAESKDYNERIMAMEGWRHNVGKALRPLYEDYADLKNKAAKLNGYQDYGDYWRGNYETKDIGEYAYSRDDLVKDVESLFEEVKPLYREL
HAYVRAKLMEYFGSEHISRTGGLPAHLLGDMWGRFWANLYPWSIPYPSEEDIDVTQAMVEQGWTAKRMFESADKFFQSVGLQPMNDNFWKNSMIELPTDGRKVVCHPTAWDMGNRVDFRI
KMCTKINMEDFLTVHHEMGHIQYDMEYAHLPTLLRDGANEGFHEGVGEIMSLSAATPKHLKSLGLLPIDINFLLKQALSIVGTLPFTYMMEQWRWKMFRGEIPKDQWMKKFWEMKREIVG
VVEPVPHDETYCDPAALFHIANDYSFIRYYTRTIFQFQFQEALCQAAGHTGPLHKCDITNSTKAGTKLSNMLKLGKSKSWTRALEEVTGQTRMNARPLLNYFKPLYEWLKKDNQDKGRHV
GWDPTWTPSAEAFKVRISLKTALGEKAYEMNANEEYFFQATVAYSMRKYWAEVKSETLNFEITHVHMSNSTQRISFYFIVKNPKDNTTIPKADVEQAIRMNKHRFNSAFLLDDKTLEFVG
IPPTLAPQSKSSVTVWLILFGVVMGMVCIALALLIITGQRCNPQKFKNDDTYLMRFQYVMFNWKVQNSD

tr|A0A4W3HYL6|A0A4W3HYL6_CALMI Angiotensin -converting enzyme OS= Callorhinchus milii OX=7868 GN=ace2 PE=3 SV=1
MFLQWLLLLSLAAAALSLSPVEQEATAFLKEFDTKSQDLVYKSSLASWEYNTNITDENIDKMNEESAKWSAFYQQASDDSSKFNINEISDNIIKLQLNSLQDKGSGVLSKEEQSQHV
IQNEMSKIYSTGTVCKPNNPSDCLGLEPGLTILLAESKDYNERIMAMEGWRHNVGKALRPLYEDYADLKNKAAKLNGYQDYGDYWRGNYETKDIGEYAYSRDDLVKDVESLFEEVK
PLYRELHAYVRAKLMEYFGSEHISRTGGLPAHLLGDMWGRFWANLYPWSIPYPSEEDIDVTQAMVEQGWTAKRMFESADKFFQSVGLQPMNDNFWKNSMIELPTDGRKVVCHPTAW
DMGNRVDFRIKMCTKINMEDFLTVHHEMGHIQYDMEYAHLPYLLRDGANEGFHEGVGEIMSLSAATPKHLKSLGLLPASFIETSKIDINFLLKQALSIVGTLPFTFMMEQWRWKMF
RGKIPKDQWMKKFWEMKREIVGVVEPVPHDETYCDPAALFHVANDYSFIRYYTRTIFQFQFQEALCQAAGHTGPLHKCDITNSTKAGTKLSNMLKLGKSKSWTRALEEVTGQTRMNARPLLNYFKPLYEWLKKDNQDKGRHVGWDPTWTPSAEAFKVRISLKTALGEKAY
EMNANEEYFFQATVAYSMRKYWAEVKSETLNFEITHVHMSNSTQRISFYFIVKNPKDNTTIPKADVEQAIRMNKHRFNSAFLLDDKTLEFVGIPPTLAPQSKSSVTVWLILFGVVM
GMVCIALALLIITGQRAKKHKVEEEGFENVP

tr|A0A4W3HYL1|A0A4W3HYL1_CALMI Angiotensin -converting enzyme OS= Callorhinchus milii OX=7868 GN=ace2 PE=3 SV=1
MFLQWLLLLSLAAAALSLSPVEQEATAFLKEFDTKSQDLVYKSSLASWEYNTNITDENIDKMNEESAKWSAFYQQASDDSSKFNINEISDNIIKLQLNSLQDKGSGVLKFLHFYSV
LQQSVSEYFTGNKLFMNLFDYNERLMAMEGWRHNVGKALRPLYEDYADLKNKAAKLNGYQUYGDYWRGNYETKDIGEYAYSRDDLVKDVESLFEEVKPLYRELHAYVRAKLMETFG
SEHISRTGGLPAHLLGDMWGRFWANLYPWSIPYPSEEDIDVTQAMVEQGWTAKRMFESADKFFQSVGLQPMNDNFWKNSMIELPTDGRKVVCHPTAWDMGNRVDFRIKMCTKINME
DFLTVHHEMGHIQYDMEYAHLPYLLRDGANEGFHEIVGEIMSLSAATPKHLKSLGLLPASFIETKIDINFLLKQALSIVGTLPFTPMMEQMRWKMFRGEIPKDQWMKKFWEMKREF
VGVVEPVPHDETYCDPAALFHIANDYSFIRYYTRTIFQFQFQEALCQAAGHTGPLHKCDITNSTKAGTKLSNMLKLGKSKSWTRALEEVTGQTRMNARPLLNYFKPLYEWLKKDNQ
DKGRHVGWDPTWTPSAEAFKVRISLKTALGEKAYEMNANEEYFFQATVAYSMRKYWAEVKSETLNFEITHVHMSNSTQRISFYFIVKNPKDNTTIPKADVEQAIRMNKHRFNSAFL
LDDKTLEFVGIPPTLAPQSKSSVTVWLILFGVVMGMVCIALALLIITGQRAKKHKVEEEGFENVP
```

Fig. 2a. ACE2 IN CALLORHINCHUS MILLI (GHOST SHARK)

>tr|A0A4W3HYJ6|A0A4W3HYJ6_CALMI Angiotensin-converting enzyme OS=Callorhinchus milli OX=7868 GN=ace2 PE=3 SV=1
MFLQWLLLLSLAAAALSLSPVEQEATAFLKEFDTKSQDLVYKSSLASWEYNTNITDENIDKMNEESAKWSAFYQQASDDSSKFNINEISDNIIKLQLNSLQDKGSGVL
SKEEQDHLNEVQNEMSKIYSTGTVCKPNNPSDCLGLEPGLTILLAESKDYNERLWAWEGWRHNVGKALRPLYEDYADLKNKAAKLNGYQDYGDYWRGNYETKDIGEYA
YSRDDLVKDVESLFEEGWTAKRMFESADKFFQSVGLQPMNDNFWKNSMIELPTDGRKVVCHPTAWDMGNRVDFRIKMCTKINMEDFLTVHHEMGHIQYDMEYAHLPYL
LRDGANEGFHEGVGEIMSLSAATPKHLKSLGLLPASFIETSKIDINFLLKQALSIVGTLPFTFMMEQWRWKMFRGEIPKDQWMKKFWEMKREFVGVVEPVPHDETYCD
PAALFHIANDYSFIRYYTRTIFQFQFQEALCQAAGHTGPLHKCDITNSTKAGTKLSNMLKLGKSKSWTRALEEVTGQTRMNARPLLNYFKPLYEWLKKDNQDKGRHVG
WDPTWTPYADRHHVDLISKSDEEQFDQKSVSEAAEAFKVRISLKTALGEKAYEWNANEEYFFQATVAYSMRKYWAEVKSETLNFEITHVHMSNSTQRISFYFIVKNPK
DNTTIPKADVEQAIRMNKHRFNSAFLLDDKTLEFVGIPPTLAPQSKSSVTVWLILFGVVMGMVCIALALLIITGQRAKKQKAKETDVYENPSSIEEPDFDKGVKNSAF
TIEESLNNTAM

>tr|A0A4W3I1M1|A0A4W3I1M1_CALMI Angiotensin-converting enzyme OS=Callorhinchus milli OX=7868 GN=ace2 PE=3 SV=1
MFLQWLLLLSLAAAALSLSPVEQEATAFLKEFDTKSQDLVYKSSLASWEYNTNITDENIDKMNEESAKWSAFYQQASDDSSKFNINEISDNIIKLQLNSLQDKGSGVL
SKEEQDHLNEVQNEMSKIYSTGTVCKPNNPSDCLGLEPGLTILLAESKDYNERLWAWEGWRHNVGKALRPLYEDYADLKNKAAKLNGYQDYGDYWRGNYETKDIGEYA
YSRDDLVKDVESLFEEVKPLYRELHAYVRAKLMETFGSEHISRTGGLPAHLLGDMWGRFWANLYPWSIPYPSEEDIDVTQAMVEQGWTAKRMFESADKFFQSVGLQPM
NDNFWKNSMIELPTDGRKVVCHPTAWDMGNRVDFRIKMCTKINMEDFLTVHHEMGHIQYDMEYAHLPYLLRDGANEGFHEGVGEIMSLSAATPKHLKSLGLLPASFIE
TSKIDINFLLKQALSIVGTLPFTFMMEQWRWKMFRGEIPKDQWMKKFW
EMKREFVGVVEPVPHDETYCDPAALFHIANDYSFIRYYTRTIFQFQFQEALCQAAGHTGPLHKCDITNSTKAGTKLSNMLKLGKSKSWTRALEEVTGQTRMNARPLLN
YFKPLYEWLKKDNQDKGRHVGWDPTWTPSAEAFKVRISLKTALGEKAYEWNANEEYFFQATVAYSMRKYWAEVKSETLNFEITLSPLTILNSKYFNTILLIFYPNRMN
KHRFNSAFLLDKTLEFVGIPPTLAPQSKSSVTVWLILFGVVMGMVCIALALLIITGQRAKKQLKRRVLRMCLEGGKGGSEAKGMWDGVPKVEGRKAEGLSASGGAD

Fig. 2b. ACE2 IN TAENIPYGIA GUTTATA (ZEBRA FINCH)

- >tr|H0ZCK6|H0ZCK6_TAEGU Angiotensin-converting enzyme OS=Taeniopygia guttata OX=59729 GN=ACE2 PE=3 SV=2
- LSGVSSCLGVKVLGLVAVVTPQDVTQQAQIFLEEFNRRAENISYENSIASWNYNTNITEENANKMSEADARWSAFYEEASRNASTFQVDSI
  ADDPTKLQIQILQERGSSVLSPEKYNRLGTVLNTMSTIYSTGTVCKINNPSECLVLEPGLDAIMSGSTDYYERLWAWEGWRADVGRMMRPL
  YEEYVELENEVARLNGYSDYGDYWRANYEAKSPENYKYSRDQLIKDVEKTFEQIKPLYEQLHAYVRHKLGQVYGPKLISSTGGLPAHLLGD
  MWGRFWTNLYALTVPYPAKPNIDVTSAMVEKKWDEIKIFKSAEAFFVSIGLNNMTDGFWENSMLTEPTDNRKVVCHPTAWDLGKNDYRIKM
  CTKVTMDDFLAAHHEMGHIEYDMAYAGQPYLLRSGANEGFHEAVGEIMSLSVATPQHLKSLNLLEPTFQDDEETEINFLLKQALTIVGTMP
  FTYMLEKWRWMVFKGEITQQEWTKRWWEMKRAIVGVVEPVPHDETYCDAATLFHVASDYSFIRYYTRTIYQFQFQEALCKAANHTGPLHKC
  DISNSTEAGQKLRQMLELGRSKPWTQALESVTGEKYMNAAPLLHYFEPLYEWLKRNNSGRFVGWKTDWTPYSSDAIKVRISLKSALGDQAY
  EWDESELFLFKSSVAYAMRKYFAEVKKQKAAFDITDIHVGEETQRISFYITVSMPGNISDMVPKADVENAIRMSRGRMNEAFRLDDSTLEF
  VGILPTLAAPYEPPVTIWLIVFGVVISLVVIGIIALIVTGLRDRAAGSNCEEVNPYGEEGRSNLGFEPAEDTQTSF
- >tr|H0ZYW8|H0ZYW8_TAEGU Angiotensin-converting enzyme OS=Taeniopygia guttata OX=59729 GN=ACE2 PE=3 SV=2
- MLVCFWLLCGLSAVVTPQDVTQQAQIFLEEFNRRAENISYENSIASWNYNTNITEENANKMSEADARWSAFYEEASRNASTFQVDSIADDP
  TKLQIQILQERGSSVLSPEKYNRLGTVLNTMSTIYSTGTVCKINNPSECLVLEPGLDAIMSGSTDYYERLWAWEGWRADVGRMMRPLYEEY
  VELENEVARLNGYSDYGDYWRANYEAKSPENYKYSRDQLIKDVEKTFEQIKPLYEQLHAYVRHKLGQVYGPKLISSTGGLPAHLLGDMWGR
  FWTNLYALTVPYPAKPNIDVTSAMKWDEIKIFKSAEAFFVSIGLNNMTDGFWENSMLTEPTDNRKVVCHPTAWDLGKNDYRIKMCTKVTMD
  DFLAAHHEMGHIEYDMAYAGQPYLLRSGANEGFHEAVGEIMSLSVATPQHLKSLNLLEPTFQDDEETEINFLLKQALTIVGTMPFTYMLEK
  WRWMVFKGEITQQEWTKRWWEMKRAIVGVVEPVPHDETYCDAATLFHVASDYSFIRYYTRTIYQFQFQEALCKAANHTGPLHKCDISNSTE
  AGQKLRQMLELGRSKPWTQALESVTGEKYMNAAPLLHYFEPLYEWLKRNNSGRFVGWKTDWTPYSSDAIKVRISLKSALGDQAYEWDESEL
  FLFKSSVAYAMRKYFLFVKVCALFEFAFDGRMSRGRMNEAFRLDDSTLEFVGILPTLAAPYEPPVTIWLIVFGVVISLVVIGIIALIVTGL
  RDRAAGSNCEEVNPYGEEGRSNLGFEPAEDTQTSF

Fig. 2b. ACE2 IN TAENIPYGIA GUTTATA (ZEBRA FINCH)

>tr|A0A674GHV0|A0A674GHV0_TAEGU Angiotensin-converting enzyme OS=Taeniopygia guttata OX=59729 GN=ACE2 PE=3 SV=1

MLVCFWLLCGLSAVVTPQDVTQQAQIFLEEFNRRAENISYENSIASWNYNTNITEENANKMSEADARWSAFYEEASRNASTFQVDSIADDPTKLQIQILQERGSSVLSPEKYNRLGTVLWTMSTIYST
GTVCKINNPSECLVLEPGLDAIMSGSTDYYERLWAWEGWRADVGRMMRPLYEEYVELENEVARLNGYSDYGDYWRANYEAKSPENYKYSRDQLIKDVEKTFEQIKPLYEQLHAYVRHKLGQVYGPKLI
SSTGGLPAHLLGDMWGRFWTNLYALTVPYPAKPNIDVTSAMVEKKWDEIKIFKSAEAFFVSIGLNNMTDGFWENSMLTEPTDNRKVVCHPTAWDLGKNDYRIKMCTKVTMDDFLAAHHEMGHIEYDMA
YAGQPYLLRSGANEGFHEAVGEIMSLSVATPQHLKSLNLLEPTFQDDETEINFLLKQALTIVGTMPFTYMLEKWRWMVFKGEITQDEWTKRWWEMKREIVGVVEPVPHDETYCDAATLFHVASDYSF
IRYYTRTIYQFHKFLTYNKSECNKFWTQALESVTGEKYNMAAPLLHYFEPLYEWLKRNNSGRFVGWKTDWIPCTDAIKVRISLKSALGDQAYEWDESELFLFKSSVAYAMRKYFAEVKKQKAAPDITD
IHVGEETQRISFYITVSMPGNISDMVPKADVENAIRMSRGRMNEAFRLDDSTLEFVGILPTLAAPYEPPVTIWLIVFGVVISLVVIGIIALIVTGLRDRSNCEEVNPYGEEGRSNLGFEPAEDTQTSF

>tr|A0A674GKE4|A0A674GKE4_TAEGU Angiotensin-converting enzyme OS=Taeniopygia guttata OX=59729 GN=ACE2 PE=3 SV=1

MLVCFWLLCGLSAVVTPQDVTQQAQIFLEEFNRRAENISYENSIASWNYNTNITEENANKMSEADARWSAFYEEASRNASTFQVDSIADDPTKLQIQILQERGSSVLSPEKYNRLGTVLWTMSTIYST
GTVCKINNPSECLVLEPGLDAIMSGSTDYYERLWAWEGWRADVGRMMRPLYEEYVELENEVARLNGYSDYGDYWRANYEAKSPENYKYSRDQLIKDVEKTFEQIKPLYEQLHAYVRHKLGQVYGPKLI
SSTGGLPAHLLGQSPSRDIEPTDMRKVVCHPTAWDLGKNDYRIKMCTKVTMDDFLAAHHEMGHIEYDMAYAGQPYLLRSGANEGFHEAVGEIMSLSVATPQHLKSLNLLEPTFQDDETEINFLLKQA
LTIVGTMPFTYMLEKWRWMVFKGEITQDEWTKRWWEMKRAIVGVVEPVPHDETYCDAATLFHVASDYSFIRYYTRTIYQFQFQEALCKAANHTGPLHKCDISNSTEAGQKLRQMLELGRSKPWTQALE
SVTGEKYMNAAPLLHYFEPLYEWLKRNNSGRFVGWKTDWTPYSSDAIKVRISLKSALGDQAYEWDESELFLFKSSVAYAMRKYFAEVKKQKAAFDITDIHVGEETQRISFYITVSMPGNISDMVPKAD
VEMAIRMSRGRMNEAFRLDDSTLEFVGILPTLAAPYEPPVTIWLIVFGVVISLVVIGIIALIVTGLRDRSNCEEVNPYGEEGRSNLGFEPAEDTQTSF

>tr|A0A674GJP6|A0A674GJP6_TAEGU Angiotensin-converting enzyme OS=Taeniopygia guttata OX=59729 GN=ACE2 PE=3 SV=1

MSTIYSTGTVCKINNPSECLVLEPGLDAIMSGSTDYYERLWAWEGWRADVGRMMRPLYEEYVELENEVARLNGYSDYGDYWRANYEAKSPENYKYSRDQLIKDVEKTFEQIKPLYEQLHAYVRHKLGQ
VTGPKLISSTGGLPAHLLGDMWGRFWTNLYALTVPYPAKPNIDVTSAMVEKKWDEIKIFKSAEAFFVSIGLNNMTDGFWENSMLTEFTDNRKVVCHPTAWDLGKNDYRIKMCTKVTMDDFLAAHHEMG
HIEYDMAYAGQPYLLRSGANEGFHEAVGEIMSLSVATPQHLKSLNLLEPTFQDDEETKINFLLKQALTIVGTMPFTYMLEKWRWMVFKGEITQQEWTKRWWEMKREIVGVVEPVPHDETYCDAATLFH
VASDYSFIRYYTRTIYQFQFQEALCKAANHTGPLHKCDISNSTEAGQKLRQMLELGRSEPWTQALESVTGEKYMNAAPLLHYFEPLYEWLKRNNSGRFVGWKTDWTPYSSDAIKVRISLKSALGDQAY
EWDESELFLFKSSVAYAMRKYFAEVKQKAAFDITDIHVGEETQRISFYITVSMPGNISDMVPKADVENAIRMSRGRMNEAFRLDDSTLEFVGILPTLAAPYEPPVTIWLIVFGVVISLVVIGIIALI
VTGLRDRKKRARESRGAASSNCEEVNPYGEEGRSNLGFEPAEDTQTSF

Fig.2c.ACE2 IN MUS MUSCULUS(MOUSE)

>tr|Q3URC9|Q3URC9_MOUSE Angiotensin-converting enzyme OS=Mus musculus OX=10090 GN=Ace2 PE=2 SV=1

MSSSSWLLLSLVAVTTAQSLTEENAKTFLNNFNQEAEDLSYQSSLASWNYNTNITEENAQKMSEAAAKWSAFYEEQSKTAQSFSLQEIQTPIIKRQLQALQQSGSSAL
SADKNKQLNTILNTMSTIYSTGKVCNPKNPQECLLLEPGLDEIMATSTDYNSRLWAWEGWRAEVGKQLRPLYEEYVVLKNEMARANNYNDYGDYWRGDYEAEGADGYN
YNRNQLIEDVERTFAEIKPLYEHLHAYVRKKLMDTYPSYISPTGCLPAHLLGDMWGRFWTNLYFLTVPFAQKPNIDVTDAMMNQGWDAERIFQEAEKFFVSVGLPHMT
QGFWANSMLTEPADGRKVVCHPTAWDLGHGDFRIKMCTKVTMDNFLTAHHEMGHIQYDMAYARQPFLLRNGANEGFHEAVGEIMSLSAATPKHLKSIGLLPSDFQEDS
ETEINFLLKQALTIVGTLPFTYMLEKWRWMVFRGEIPKEQWMKKWWEMKREIVGVVEPLPHDETYCDPASLFHVSNDYSFIRYYTRTIYQFQFQEALCQAAKYNGSLH
KCDISNSTEAGQKLLKMLSLGNSEPWTKALENVVGARNMDVKPLLNYFQPLFDWLKEQNRNSFVGWNTEWSPYADQSIKVRISLKSALGANAYEWTNNEMFLFRSSVA
YAMRKYFSIIKNQTVPFLEEDVRVSDLKPRVSFYFFVTSPQNVSDVIPRSEVEDAIRMSRGRINDVFGLNDNSLEFLGIHPTLEPPYQPPVTIWLIIFGVVMALVVVG
IIILIVTGIKGRKKKNETKREENPYDSMDIGKGESNAGFQNSDDAQTSF

>tr|Q9D836|Q9D836_MOUSE Angiotensin-converting enzyme 2 (Fragment) OS=Mus musculus OX=10090 GN=Ace2 PE=2 SV=2

XCHISNSTEAGQKLLKMLSLGNSEPWTKALENVVGARNMDVKPLLNYFQPLFDWLKEQNRNSFVGWNTEWSPYADQSIKVRISLKSALGANAYEWTNNEMFLFRSSVA
YAMRKYFSIIKNQTVPFLEEDVRVSDLKPRVSFYFFVTSPQNVSDVIPRSEVEDAIRMSRGRINDVFGLNDNSLEFLGIHPTLEPPYQPPVTIWLIIFGVVMALVVVG
IIILIVTGIKGRKKKNETKREENSYDSMDIGKGESNAGFQNSDDAQTSF

>tr|F6X479|F6X479_MOUSE Angiotensin-converting enzyme (Fragment) OS=Mus musculus OX=10090 GN=Ace2 PE=1 SV=2

MDTYPSYISPTGCLPAHLLGDMWGRFWTNLYPLTVPFAQKPNIDVTDAMMNQGWDAERIFQEAEKFFVSVGLPHMTQGFWANSMLTEPADGRKVVCHPTA
WDLGHGDFRIKMCTKVTMDNFLTAHHEMGHIQYDMAYARQPFLLRNGANEGFHEAVGEIMSLSAATPKHLKSIGLLPSDFQEDSETEINFLLKQALTIVGTLPFTYML
EKWRWMVFRGEIPKEQWMKKWWEMKREIVGVVEPLPHDETYCDPASLFHVSNDYSFIRYYTRTIYQFQFQEALCQAAKYNGSLHKCDISNSTEAGQKLLKMLSLGNSE
PWTKALENVVGARNMDVKPLLNYFQPLFDWLKEQNRNSFVGWNTEWSPYADQSIKVRISLKSALGANAYEWTNNEMFLFRSSVAYAMRKYFSIIKNQTVPFLEEDVRV
SDLKPRVSFYFFVTSPQNVSDVIPRSEVEDAIRMSRGRINDVFGLNDNSLEFLGIHPTLEPPYQPPVTIWLIIFGVVMALVVVGIIILIVTGIKGRK

Fig.2c.ACE2 IN MUS MUSCULUS(MOUSE)

>tr|Q3UXR1|Q3UXR1_MOUSE Angiotensin-converting enzyme (Fragment) OS=Mus musculus OX=10090 GN=Ace2 PE=2 SV=1
FLISRVAYWKLKYSWSKTFQIKPLYEHLHAYVRRKLMDTYPSYISPTGCLPAHLLGDMWGRFWTNLYPLTVPFAQKPNIDVTDA
MMNQGWDAERIFQEAEKFFVSVGLPHMTQGFWANSMLTEPADGRKVVCHPTAWDLGHGDFRIKMCTKVTMDNFLTAHHEMGHIQYDMA
YARQPFLLRNGANEGFHEAVGEIMSLSAATPKHLKSIGLLPSDFQEDSETEINFLLKQALTIVGTLPFTYMLEKWRWMVFRGEIPKEQ
WMKKWWEMKREIVGVVEPLPHDETYCDPASLFHVSNDYSFIRYYTRTIYQFQFQEALCQAAKYNGSLHKCDISNSTEAGQKLLKMLSL
GNSEPWTKALENVVGARNMDVKPLLNYFQPLFDWLKEQNRNSFVGWNTEWSPYADQSIKVRISLKSALGANAYEWTNNEMFLFRSSVA
YAMRKYFSIIKNQTVPFLEEDVRVSDLKPRVSFYFFVTSPQNVSDVIPRSEVEDAIRMSRGRINDVFGLNDNSLEFLGIHPTLEPPYQ
PPVTIWLIIFGVVMALVVVGIIILIVTGIKGRK
>sp|Q8R0I0|ACE2_MOUSE Angiotensin-converting enzyme 2 OS=Mus musculus OX=10090 GN=Ace2 PE=1 SV=1
MSSSSWLLLSLVAVTTAQSLTEENAKTFLNNFNQEAEDLSYQSSLASWNYNTNITEENAQKMSEAAAKWSAFYEEQSKTAQSFSLQEI
QTPIIKRQLQALQQSGSSALSADKNKQLNTILNTMSTIYSTGKVCNPKNPQECLLLEPGLDEIMATSTDYNSRLWAWEGWRAEVGKQL
RPLYEEYVVLKNEMARANNYNDYGDYWRGDYEAEGADGYNYNRNQLIEDVERTFAEIKPLYEHLHAYVRRKLMDTYPSYISPTGCLPA
HLLGDMWGRFWTNLYPLTVPFAQKPNIDVTDAMMNQGWDAERIFQEAEKFFVSVGLPHMTQGFWANSMLTEPADGRKVVCHPTAWDLG
HGDFRIKMCTKVTMDNFLTAHHEMGHIQYDMAYARQPFLLRNGANEGFHEAVGEIMSLSAATPKHLKSIGLLPSDFQEDSETEINFLL
KQALTIVGTLPFTYMLEKWRWMVFRGEIPKEQWMKKWWEMKREIVGVVEPLPHDETYCDPASLFHVSNDYSFIRYYTRTIYQFQFQEA
LCQAAKYNGSLHKCDISNSTEAGQKLLKMLSLGNSEPWTKALENVVGARNMDVKPLLNYFQPLFDWLKEQNRNSFVGWNTEWSPYADQ
SIKVRISLKSALGANAYEWTNNEMFLFRSSVAYAMRKYFSIIKNQTVPFLEEDVRVSDLKPRVSFYFFVTSPQNVSDVIPRSEVEDAI
RMSRGRINDVFGLNDNSLEFLGIHPTLEPPYQPPVTIWLIIFGVVMALVVVGIIILIVTGIKGRKKKNETKREENPYDSMDIGKGESN
AGFQNSDDAQTSF

Fig.2d.ACE2 IN URSUS MARITIMUS (POLAR BEAR)

>tr|A0A452TT37|A0A452TT37_URSMA Angiotensin-converting enzyme OS=Ursus maritimus OX=29073 GN=ACE2 PE=3 SV=1

MDLAETFLEKFNYEAEDLYYQSSLASWNYNTNITNENIQKMNDAGAKWSAFYEEQSKHAKTYPLEEIHNSTVKRQLQALQHSGSSVLSADKSQRLNTILNAMSTIYSTGKAC
NFNNPQECLLLEPGLDDIMENSKDYNERLWAWEGWRSEVGKQLRPLYEEYVALKNEMARANNYEDYGDYWRGDYEEEWTDGYNYSRNQLIEDVEHTFTQIKALYEHLHAYVR
AKLMDTYPSRISPTGCLPAHLLGDMWGRFWTNLYPLTIPFGQKPNIDVTDAMVNQNWDARRIFEEAEKFFVSVGLPNMTQEFWENSMLTEPGDGQKVVCHPTAWDLGKGDFR
IKMCTKVTMDDFLTAHHEMGHIQYDMAYAEQPFLLRNGANEGFHEAVGEIMSLSAATPNHLKNIGLLPPGFSEDNETEINFLLKQALTIVGTLPFTYMLEKWRWMVFQGKIP
KEQWMKKWWEMKRDIVGVVEPLPHDETYCDPASLFHVANDYSFIRYYTRTIYQFQFQEALCQIAKHEGPLHKCDISNSSEAGKTLLQMLRLGRSKPWTLALEHVVGAKNMDV
RPLLNYFEPLFTWLKEQNRNSFVGWNTDWSPYADQSIKVRISLKSALGEKAYEWNDNEMYLFRSSIAYAMRKYFSEAKNQMIPFVEDNVWVNDLKPRISFNFFVTSPGNVSD
VIPRADVEGAIKMSRDRINDAFQLDDNSLEFLGIQPTLGPPYQPPVTIWLIVFGVVMGLVVIGIILLIFSGIRNRRKNDQARSEENPYASVDLSKGENNPGFQNADDVQTSF

>tr|A0A452TTF7|A0A452TTF7_URSMA Angiotensin-converting enzyme OS=Ursus maritimus OX=29073 GN=ACE2 PE=3 SV=1

ALSITEDLYYQSSLASWNYNTNITNENIQKMNDAGAKWSAFYEEQSKHAKTYPLEEIHNSTVKRQLQALQHSGSSVLSADKSQRLNTILNAMSTIYSTGKACNPNNPQECLL
LEPGLDDIMENSKDYNERLWAWEGWRSEVGKQLRPLYEEYVALKNEMARANNYEDYGDYWRGDYEEEWTDGYNYSRNQLIEDVEHTFTQIKALYEHLHAYVRAKLMDTYPSR
ISPTGCLPAHLLGDMWGRFWTNLYPLTIPFGQKPNIDVTDAMVNQNWDARRIFEEAEKFFVSVGLPNMTQEFWENSMLTEPGDGQKVVCHPTAWDLGKGDFRIKMCTKVTMD
DFLTAHHEMGHIQYDMAYAEQPFLLRNGANEGFHEAVGEIMSLSAATPNHLKNIGLLPPGFSEDNETEINFLLKQALTIVGTLPFTYMLEKWRWMVFQGKIPKEQWMKKWWE
MKRDIVGVVEPLPHDETYCDPASLFHVANDYSFIRYYTRTIYQFQFQEALCQIAKHEGPLHKCDISNSSEAGKTLLQMLRLGRSKPWTLALEHVVGAKNMDVRPLLNYFEPL
FTWLKEQNRNSFVGWNTDWSPYADQSIKVRISLKSALGEKAYEWNDNEMYLFRSSIAYAMRKYFSEAKNQMIPFVEDNVWVNDLKPRISFNFFVTSPGNVSDVIPRADVEGA
IKMSRDRINDAFQLDDNSLEFLGIQPTLGPPYQPPVTIWLIVFGVVMGLVVIGIILLIFSGIRNRRKNDQARSEENPYASVDLSKGENNPGFQNADDVQTSF


Fig.2d. ACE2 IN URSUS MARITIMUS (POLAR BEAR)
>tr|A0A452TTE2|A0A452TTE2_URSMA Angiotensin-converting enzyme OS=Ursus maritimus OX=29073 GN=ACE2 PE=3 SV=1
FSGSLFFLKDTEDLYYQSSLASWNYNTNITNENIQKMNDAGAKWSAFYEEQSKHAKTYPLEEIHNSTVKRQLQALQHSGSSVLSADKSQRLNTILNAMSTIYSTGKACNPNN
PQECLLLEPGLDDIMENSKDYNERLWAWEGWRSEVGKQLRPLYEEYVALKNEMARANNYEDYGDYWRGDYEEEWTDGYNYSRNQLIEDVEHTFTQIKALYEHLHAYVRAKLM
DTYPSRISPTGCLPAHLLGDMWGRFWTNLYPLTIPFGQKPNIDVTDAMVNQNWDARRIFEEAEKFFVSVGLPNMTQEFWENSMLTEPGDGQKVVCHPTAWDLGKGDFRIKMC
TKVTMDDFLTAHHEMGHIQYDMAYAEQPFLLRNGANEGFHEAVGEIMSLSAATPNHLKNIGLLPPGFSEDNETEINFLLKQALTIVGTLPFTYMLEKWRWMVFQGKIPKEQW
MKKWWEMKREIVGVVEPLPHDETYCDPASLFHVANDYSFIRYYTRTIYQFQFQEALCQIAKHEGPLHKCDISNSSEAGKTLLQMLRLGRSKPWTLALEHVVGAKNMDVRPLL
NYFEPLFTWLKEQNRNSFVGWNTDWSPYADQSIKVRISLKSALGEKAYEWNDNEMYLFRSSIAYAMRKYFSEAKNQMIPFVEDNVWVNDLKPRISFNFFVTSPGNVSDVIPR
ADVEGAIKMSRDRINDAFQLDDNSLEFLGIQPTLGPPYQPPVTIWLIVFGVVMGLVVIGIILLIFSGIRNRRKNDQARSEENPYASVDLSKGENNPGFQNADDVQTSF
>tr|A0A452TTE1|A0A452TTE1_URSMA Angiotensin-converting enzyme OS=Ursus maritimus OX=29073 GN=ACE2 PE=3 SV=1
CSSCLSEDLYYQSSLASWNYNTNITNENIQKMNDAGAKWSAFYEEQSKHAKTYPLEEIHNSTVKRQLQALQHSGSSVLSADKSQRLNTILNAMSTIYSTGKACNPNNPQECL
LLEPGLDDIMENSKDYNERLWAWEGWRSEVGKQLRPLYEEYVALKNEMARANNYEDYGDYWRGDYEEEWTDGYNYSRNQLIEDVEHTFTQIKALYEHLHAYVRAKLMDTYPS
RISPTGCLPAHLLGDMWGRFWTNLYPLTIPFGQKPNIDVTDAMVNQNWDARRIFEEAEKFFVSVGLPNMTQEFWENSMLTEPGDGQKVVCHPTAWDLGKGDFRIKMCTKVTM
DDFLTAHHEMGHIQYDMAYAEQPFLLRNGANEGFHEAVGEIMSLSAATPNHLKNIGLLPPGFSEDNETEINFLLKQALTIVGTLPFTYMLEKWRWMVFQGKIPKEQWMKKWW
EMKRDIVGVVEPLPHDETYCDPASLFHVANDYSFIRYYTRTIYQFQFQEALCQIAKHEGPLHKCDISNSSEAGKTLLQMLRLGRSKPWTLALEHVVGAKNMDVRPLLNYFEP
LFTWLKEQNRNSFVGWNTDWSPYADQSIKVRISLKSALGEKAYEWNDNEMYLFRSSIAYAMRKYFSEAKNQMIPFVEDNVWVNDLKPRISFNFFVTSPGNVSDVIPRADVEG
AIKMSRDRINDAFQLDDNSLEFLGIQPTLGPPYQPPVTIWLIVFGVVMGLVVIGIILLIFSGIRNRRKNDQARSEENPYASVDLSKGENNPGFQNADDVQTSF
>tr|A0A452TT60|A0A452TT60_URSMA Angiotensin-converting enzyme OS=Ursus maritimus OX=29073 GN=ACE2 PE=3 SV=1
TQESSQQDLAETFLEKFNYEAEDLYYQSSLASWNYNTNITNENIQKMNDAGAKWSAFYEEQSKHAKTYPLEEIHNSTVKRQLQALQHSGSSVLSADKSQRLNTILNAMSTIY
STGKACNPNNPQECLLLEPGLDDIMENSKDYNERLWAWEGWRSEVGKQLRPLYEEYVALKNEMARANNYEDYGDYWRGDYEEEWTDGYNYSRNQLIEDVEHTFTQIKALYEH
LHAYVRAKLMDTYPSRISPTGCLPAHLLGDMWGRFWTNLYPLTIPFGQKPNIDVTDAMVNQNWDARRIFEEAEKFFVSVGLPNMTQEFWENSMLTEPGDGQKVVCHPTAWDL
GKGDFRIKMCTKVTMDDFLTAHHEMGHIQYDMAYAEQPFLLRNGANEGFHEAVGEIMSLSAATPNHLKNIGLLPPGFSEDNETEINFLLKQALTIVGTLPFTYMLEKWRWMV
FQGKIPKEQWMKKWWEMKREIVGVVEPLPHDETYCDPASLFHVANDYSFIRYYTRTIYQFQFQEALCQIAKHEGPLHKCDISNSS
EAGKTLLQMLRLGRSKPWTLALEHVVGAKNMDVRPLLNYFEPLFTWLKEQNRNSFVGWNTDWSPYADQSIKVRISLKSALGEKAYEWNDNEMYLFRSSIAYAMRKYFSEAKN
QMIPFVEDNVWVNDLKPRISFNFFVTSPGNVSDVIPRADVEGAIKMSRDRINDAFQLDDNSLEFLGIQPTLGPPYQPPVTIWLIVFGVVMGLVVIGIILLIFSGIRNRRKND
QARSEENPYASVDLSKGENNPGFQNADDVQTSF


Fig. 2e. ACE2 IN HOMO SAPIENS(Human)
>BAB40370.1 ACE2 [Homo sapiens]
MSSSSWLLLSLVAVTAAQSTIEEQAKTFLDKFNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDKWSAFLKEQSTLAQMYPLQEIQNLTV
KLQLQALQQNGSSVLSEDKSKRLNTILNTMSTIYSTGKVCNPDNPQECLLLEPGLNEIMANSLDYNERLWAWESWRSEVGKQLRPLYEEYVVL
KNEMARANHYEDYGDYWRGDYEVNGVDGYDYSRGQLIEDVEHTFEEIKPLYEHLHAYVRAKLMNAYPSYISPIGCLPAHLLGDMWGRFWTNLY
SLTVPFGQKPNIDVTDAMVDQAWDAQRIFKEAEKFFVSVGLPNMTQGFWENSMLTDPGNVQKAVCHPTAWDLGKGDFRILMCTKVTMDDFLTA
HHEMGHIQYDMAYAAQPFLLRNGANEGFHEAVGEIMSLSAATPKHLKSIGLLSPDFQEDNETEINFLLKQALTIVGTLPFTYMLEKWRWMVFK
GEIPKDQWMKKWWEMKREIVGVVEPVPHDETYCDPASLFHVSNDYSFIRYYTRTLYQFQFQEALCQAAKHEGPLHKCDISNSTEAGQKLFNML
RLGKSEPWTLALENVVGAKNMNVRPLLNYFEPLFTWLKDQNKNSFVGWSTDWSPYADQSIKVRISLKSALGDRAYEWNDNEMYLFRSSVAYAM
RQYFLKVKNQMILFGEEDVRVANLKPRISFNFFVTAPKNVSDIIPRTEVEKAIRMSRSRINDAFRLNDNSLEFLGIQPTLGPPNQPPVSIWLI
VFGVVMGVIVVGIVILIFTGIRDRKKKNKARSGENPYASIDISKGENNPGFQNTDDVQTSF
>AAQ89076.1 ACE2 [Homo sapiens]
MSSSSWLLLSLVAVTAAQSTIEEQAKTFLDKFNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDKWSAFLKEQSTLAQMYPLQEIQNLTV
KLQLQALQQNGSSVLSEDKSKRLNTILNTMSTIYSTGKVCNPDNPQECLLLEPGLNEIMANSLDYNERLWAWESWRSEVGKQLRPLYEEYVVL
KNEMARANHYEDYGDYWRGDYEVNGVDGYDYSRGQLIEDVEHTFEEIKPLYEHLHAYVRAKLMNAYPSYISPIGCLPAHLLGDMWGRFWTNLY
SLTVPFGQKPNIDVTDAMVDQAWDAQRIFKEAEKFFVSVGLPNMTQGFWENSMLTDPGNVQKAVCHPTAWDLGKGDFRILMCTKVTMDDFLTA
HHEMGHIQYDMAYAAQPFLLRNGANEGFHEAVGEIMSLSAATPKHLKSIGLLSPDFQEDNETEINFLLKQALTIVGTLPFTYMLEKWRWMVFK
GEIPKDQWMKKWWEMKREIVGVVEPVPHDETYCDPASLFHVSDDYSFIRYYTRTLYQFQFQEALCQAAKHEGPLHKCDISNSTEAGQKLL

Fig. 2e. ACE2 IN HOMO SAPIENS(Human)

- >NP_001375381.1 angiotensin -converting enzyme 2 isoform 4 [Homo sapiens]

- MREAGWDKGRIIMCTKVTMDDFLTAHHEMGHIQYDMAYAAQPFLLRNGANEGFHEAVGEIMSLSAATPKHLKSIGLLSPDFQEDNETEINFLLKQALTIVGTLPFTYMLEKWRWMVF
KGEIPKDQWMKKWWEMKREIVGVVEPVPHDETYCDPASLFHVSNDYSFIRYYTRTLYQFQFQEALCQAAKHEGPLHKCDISNSTEAGQKLFNMLRLGKSEPWTLALENVVGAKNMNVR
PLLNYFEPLFTWLKDQNKNSFVGWSTDWSPYADQSIKVRISLKSALGDKAYEWNDNEMYLFRSSVAYAMRQYFLKVKNQMILFGEEDVRVANLKPRISFNFFVTAPKNVSDIIPRTEV
EKAIRMSRSRINDAFRLNDNSLEFLGIQPTLGPPNQPPVSIWLIVFGVVMGVIVVGIVILIFTGIRDRKKKNKARSGENPYASIDISKGENNPGFQNTDDVQTSF

- >BAD99267.1 angiotensin -converting enzyme 2 [Homo sapiens]

- MSSSSWLLLSLVAVTAAQSTIEEQAKTFLDKFNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDKWSAFLKEQSTLAQMYPLQEIQNLTVKLQLQALQQNGSSVLSEDKSKRLNT
ILNTMSTIYSTGKVCNPDNPQECLLLEPGLNEIMANSLDYNERLWAWESWRSEVGKQLRPLYEEYVVLKNEMARANHYEDYGDYWRGDYEVNGVDGYDYSRGQLIEDVEHTFEEIKPL
YEHLHAYVRAKLMNAYPSYISPIGCLPAHLLGDMWGRFWTNLYSLTVPFGQKPNIDVTDAMVDQAWDAQRIFKEAEKFFVSVGLPNMTQGFWENSMLTDPGNVQKAVCHPTAWDLGKG
DFRILMCTKVTMDDFLTAHHEMGHIQYDMAYAAQPFLLRNGANEGFHEAVGEIMSLSAATPKHLKSIGLLSPDFQEDNETEINFLLKQALTIVGTLPFTYMLEKWRWMVFKGEIPKDQ
WMKKWWEMKREIVGVVEPVPHDETYCDPASLFHVSNDYSFIRYYTRTLYQFQFQEALCQAAKHEGPLHKCDISNSTEAGQKLFNMLRLGKSEPWTLALENVVGAKNMNVRPLLNYFEP
LFTWLKDQNKNSFVGWSTDWSPYADQSIKVRISLKSALGDKAYEWNDNEMYLFRSSVAYAMRQYFLKVKNQMILFGEEDVRVANLKPRISFNFFVTAPKNVSDIIPRTEVEKAIRMSR
SRINDAFRLNDNSLEFLGIQPTLGPPNQPPVSIWLIVFGVVMGVIVVGIVILIFTGIRDRKKKNKARSGENPYASIDISKGENNPGFQNTDDVQTSF

- >BAD99266.1 angiotensin -converting enzyme 2 [Homo sapiens]

- MSSSSWLLLSLVAVTAAQSTIEEQAKTFLDKFNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDKWSAFLKEQSTLAQMYPLQEIQNLTVKLQLQALQQNGSSVLSEDKSKRLNT
ILNTMSTIYSTGKVCNPDNPQECLLLEPGLNEIMANSLDYNERLWAWESWRSEVGKQLRPLYEEYVVLKNEMARANHYEDYGDYWRGDYEVNGVDGYDYSRGQLIEDVEHTFEEIKPL
YEHLHAYVRAKLMNAYPSYISPIGCLPAHLLGDMWGRFWTNLYSLTVPFGQKPNIDVTDAMVDQAWDAQRIFKEAEKFFVSVGLPNMTQGFWENSMLTDPGNVQKAVCHPTAWDLGKG
DFRILMCTKVTMDDFLTAHHEMGHIQYDMAYAAQPFLLRNGANEGFHEAVGEIMSLSAATPKHLKSIGLLSPDFQEDNETEINFLLKQALTIVGTLPFTYMLEKWRWMVFKGEIPKDQ
WMKKWWEMKREIVGVVEPVPHDETYCDPASLFHVSNDYSFIRYYTRTLYQFQFQEALCQAAKHEGPLHKCDISNSTEAGQKLFNMLRLGKSEPWTLALENVVGAKNMNVRPLLNYFEP
LFTWLKDQNKNSFVGWSTDWSPYADQSIKVRISLKSALGDKAYEWNDNEMYLFRSSVAYAMRQYFLKVKNQMILFGEEDVRVANLKPRISFNFFVTAPKNVSDIIPRTEVEKAIRMSR
SRINDAFRLNDNSLEFLGIQPTLGPPNQPPVSIWLIVFGVVMGVIVVGIVILIFTGIRDRKKKNKARSGENPYASIDISKGENNPGFQNTDDVQTSF

Fig. 3a, CALLORHINCHUS MILII (GHOST SHARK)

>tr|A0A4W3HYJ6|A0A4W3HYJ6_CALMI Angiotensin-converting enzyme OS=Callorhinchus milii OX=7868 GN=ace2 PE=3 SV=1

Number of amino acids: 767
Molecular weight: 88250.04
Theoretical pI: 5.58
Total number of negatively charged residues (Asp + Glu): 107
Total number of positively charged residues (Arg + Lys): 88

Atomic composition:

Carbon C	3968
Hydrogen H	6066
Nitrogen N	1038
Oxygen O	1180
Sulfur S	33

Formula: C3968H6066N1038O1180S33
Total number of atoms: 12285

Amino acid composition:

Ala (A)	55	7.2%
Arg (R)	26	3.4%
Asn (N)	42	5.5%
Asp (D)	47	6.1%
Cys (C)	8	1.0%
Gln (Q)	30	3.9%
Glu (E)	60	7.8%
Gly (G)	37	4.8%
His (H)	19	2.5%
Ile (I)	37	4.8%
Leu (L)	68	8.9%
Lys (K)	62	8.1%
Met (M)	25	3.3%
Phe (F)	39	5.1%
Pro (P)	28	3.7%
Ser (S)	54	7.0%
Thr (T)	43	5.6%
Trp (W)	21	2.7%
Tyr (Y)	29	3.8%
Val (V)	37	4.8%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%
(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

Extinction coefficients:
Extinction coefficients are in units of M cm-1, at 280 nm measured in water.
Ext. coefficient 159210
Abs 0.1% (=1 g/l) 1.804, assuming all pairs of Cys residues form cystines
Ext. coefficient 158710
Abs 0.1% (=1 g/l) 1.798, assuming all Cys residues are reduced
Estimated half-life:
The N-terminal of the sequence considered is M (Met).
The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
>20 hours (yeast, in vivo).
>10 hours (Escherichia coli, in vivo).
Instability index:
The instability index (II) is computed to be 36.46
This classifies the protein as stable.

Aliphatic index: 74.55

Grand average of hydropathicity (GRAVY) 0.495

Fig.4a TAENIOPYGIA GU TTATA (ZEBRA FINCH)

>tr|H0ZCK6|H0ZCK6_TAEGU Angiotensin -converting enzyme OS=Taeniopygia guttata OX=59729 GN=ACE2 PE=3 SV=2

- Number of amino acids: 804
- Molecular weight: 91738.87
- Theoretical pI: 5.10Total number of negatively charged residues (Asp + Glu): 108
- Total number of positively charged residues (Arg + Lys): 79
- Atomic composition:
- Carbon C 4124
- Hydrogen H 6297
- Nitrogen N 1075
- Oxygen O 1232
- Sulfur S 34
- Formula: C4124H6297N1075O1232S34
- Total number of atoms: 12762

Amino acid composition :

Ala (A)	61	7.6%
Arg (R)	37	4.6%
Asn (N)	38	4.7%
Asp (D)	40	5.0%
Cys (C)	9	1.1%
Gln (Q)	28	3.5%
Glu (E)	68	8.5%
Gly (G)	47	5.8%
His (H)	15	1.9%
Ile (I)	43	5.3%
Leu (L)	66	8.2%
Lys (K)	42	5.2%
Met (M)	25	3.1%
Phe (F)	32	4.0%
Pro (P)	34	4.2%
Ser (S)	55	6.8%
Thr (T)	50	6.2%
Trp (W)	22	2.7%
Tyr (Y)	40	5.0%
Val (V)	52	6.5%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%
(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

Extinction coefficients:
Extinction coefficients are in units of M-1 cm-1, at 280 nm measured in water.
Ext. coefficient 181100
Abs 0.1% (=1 g/l) 1.974, assuming all pairs of Cys residues form cystines
Ext. coefficient 180600
Abs 0.1% (=1 g/l) 1.969, assuming all Cys residues are reduced
Estimated halflife:
The N-terminal of the sequence considered is L (Leu).
The estimated halflife is: 5.5 hours (mammalian reticulocytes, in vitro).
3 min (yeast, in vivo).
2 min (Escherichia coli, in vivo).
Instability index:
The instability index (II) is computed to be 41.99
This classifies the protein as unstable.
Aliphatic index: 79.22
Grand average of hydropathicity (GRAVY): -0.347

>tr|Q9D836|Q9D836_MOUSE Angiotensin-converting enzyme 2 (Fragment) OS=Mus musculus OX=10090 GN=Ace2 PE=2 SV=2

Fig. 5b MUS MUSCULUS (MOUSE)

- Number of amino acids: 265
- Molecular weight: 30079.48
- Theoretical pI: 5.66
- Total number of negatively charged residues (Asp + Glu): 31
- Total number of positively charged residues (Arg + Lys): 29
- Atom composition:
- As there is at least one ambiguous position (B,Z or X) in the sequence
- considered, the atomic composition cannot be computed.

Amino acid composition:

Ala (A)	13	4.9%
Arg (R)	13	4.9%
Asn (N)	20	7.5%
Asp (D)	14	5.3%
Cys (C)	1	0.4%
Gln (Q)	9	3.4%
Glu (E)	17	6.4%
Gly (G)	15	5.7%
His (H)	1	0.4%
Ile (I)	18	6.8%
Leu (L)	21	7.9%
Lys (K)	16	6.0%
Met (M)	7	2.6%
Phe (F)	15	5.7%
Pro (P)	13	4.9%
Ser (S)	24	9.1%
Thr (T)	11	4.2%
Trp (W)	6	2.3%
Tyr (Y)	8	3.0%
Val (V)	22	8.3%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%
(B)	0	0.0%
(Z)	0	0.0%
(X)	1	0.4%

Extinction coefficients:
Extinction coefficients are in units of M-1 cm-1, at 280 nm measured in water.
Ext. coefficient 44920
Abs 0.1% (=1 g/l) 1.493, assuming all pairs of Cys residues form cystines
Ext. coefficient 44920
Abs 0.1% (=1 g/l) 1.493, assuming all Cys residues are reduced
Estimated half-life:
The N-terminal of the sequence considered is X ().
Due to the presence of an N-terminal ambiguity, the estimated half-life can not be computed.
Instability index:
The instability index (II) is computed to be 34.77
This classifies the protein as stable.
Aliphatic index: 86.38
Grand average of hydropathicity (GRAVY): -0.261

>tr|A0A452TTF7|A0A452TTF7_URSMA Angiotensin-converting enzyme OS=Ursus maritimus OX=29073 GN=ACE2 PE=3 SV=1

Fig.6b URSUS MARITIMUS (POLAR BEAR)

Number of amino acids: 774
Molecular weight: 89360.95
Theoretical pI: 5.23

Total number of negatively charged residues (Asp + Glu): 101
Total number of positively charged residues (Arg + Lys): 75
Atomic composition:

Carbon C	4022
Hydrogen H	6084
Nitrogen N	1064
Oxygen O	1185
Sulfur S	33

Formula: $C_{4022}H_{6084}N_{1064}O_{1185}S_{33}$
Total number of atoms: 12388

Amino acid composition

Ala (A)	69	6.3%
Arg (R)	34	4.4%
Asn (N)	53	6.8%
Asp (D)	44	5.7%
Cys (C)	8	1.0%
Gln (Q)	34	4.4%
Glu (E)	57	7.4%
Gly (G)	43	5.6%
His (H)	18	2.3%
Ile (I)	41	5.3%
Leu (L)	68	8.8%
Lys (K)	41	5.3%
Met (M)	25	3.2%
Phe (F)	33	4.3%
Pro (P)	39	5.0%
Ser (S)	48	6.2%
Thr (T)	38	4.9%
Trp (W)	24	3.1%
Tyr (Y)	35	4.5%
Val (V)	40	5.2%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%
(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

Extinction coefficients:
Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water
.Ext. coefficient 184650
Abs 0.1% (=1 g/l) 2.066, assuming all pairs of Cys residues form cystines
Ext. coefficient 184150
Abs 0.1% (=1 g/l) 2.061, assuming all Cys residues are reduced

Estimated half -life:
The N-terminal of the sequence considered is A (Ala).
The estimated half-life is: 4.4 hours (mammalian reticulocytes, in vitro).
>20 hours (yeast, in vivo).
>10 hours (Escherichia coli, in vivo).

Instability index:
The instability index (II) is computed to be 42.44
This classifies the protein as unstable.

Aliphatic index: 76.24
Grand average of hydropathicity (GRAVY): -0.485

>NP_001375381.1 angiotensin-converting enzyme 2 isoform 4 [Homo sapiens]

Number of amino acids: 808

Molecular weight 9895437

Theoretical pI: 4.68

Total number of negatively charged residues

(Asp + Glu): 107

Total number of positively charged residues

(Arg + Lys): 46

Atom composition

As there is at least one ambiguous position (B,Z or X) in the sequence considered,the atomic compositioncannot be computed.

Amino acid composition

Ala (A)	59	7.3%
Arg (R)	45	5.6%
Asn (N)	50	6.2%
Asp (D)	23	2.8%
Cys (C)	41	5.1%
Gln (Q)	2	0.2%
Glu (E)	84	10.4%
Gly (G)	21	2.6%
His (H)	27	3.3%
Ile (I)	81	10.0%
Leu (L)	41	5.1%
Lys (K)	1	0.1%
Met (M)	30	3.7%
Phe (F)	25	3.1%
Pro (P)	16	2.0%
Ser (S)	62	7.7%
Thr (T)	62	7.7%
Trp (W)	3	0.4%
Tyr (Y)	21	2.6%
Val (V)	11	1.4%
Pyl (O)	53	6.6%
Sec (U)	28	3.5%
(B)	12	1.5%
(Z)	1	0.1%
(X)	9	1.1%

Fig.7b HOMO SAPIENS (HUMAN)

Extinction coefficients:

Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water.

Ext. coefficient 50290

Abs 0.1% (=1 g/l) 0.508, assuming all pairs of Cys residues form cystines

Ext. coefficient 47790

Abs 0.1% (=1 g/l) 0.483, assuming all Cys residues are reduced

Estimated half-life:

The N-terminal of the sequence considered is N (Asn).

The estimated half-life is: 1.4 hours (mammalian reticulocytes, in vitro).

3 min (yeast, in vivo).

>10 hours (Escherichia coli, in vivo).

Instability index:

The instability index (II) is computed to be 51.74

This classifies the protein as unstable.

Aliphatic index: 70.14

Grand average of hydropathicity (GRAVY): -0.185

Fig.8aCALLORHINCHUS MILII(GHOST SHARK)

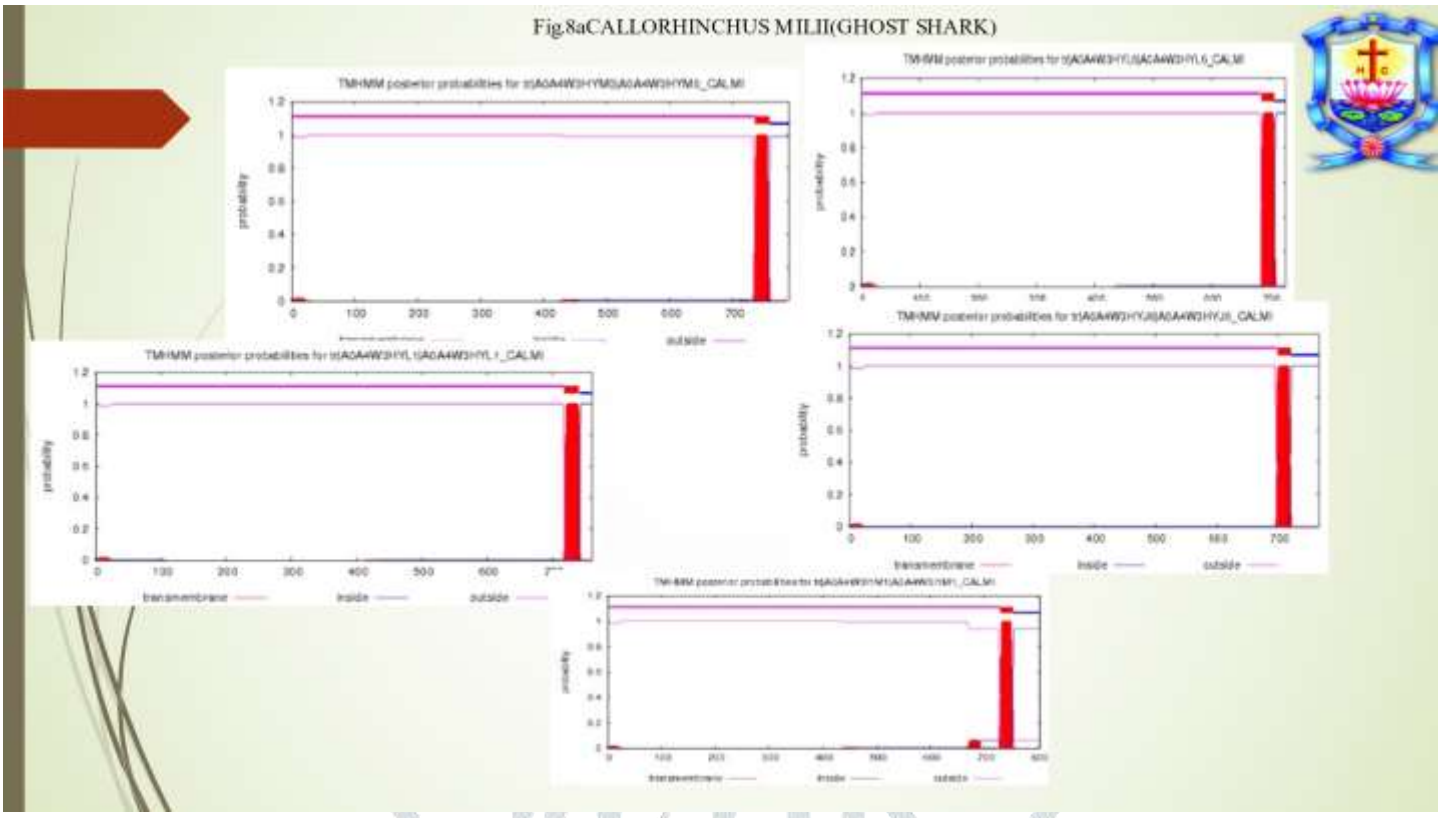
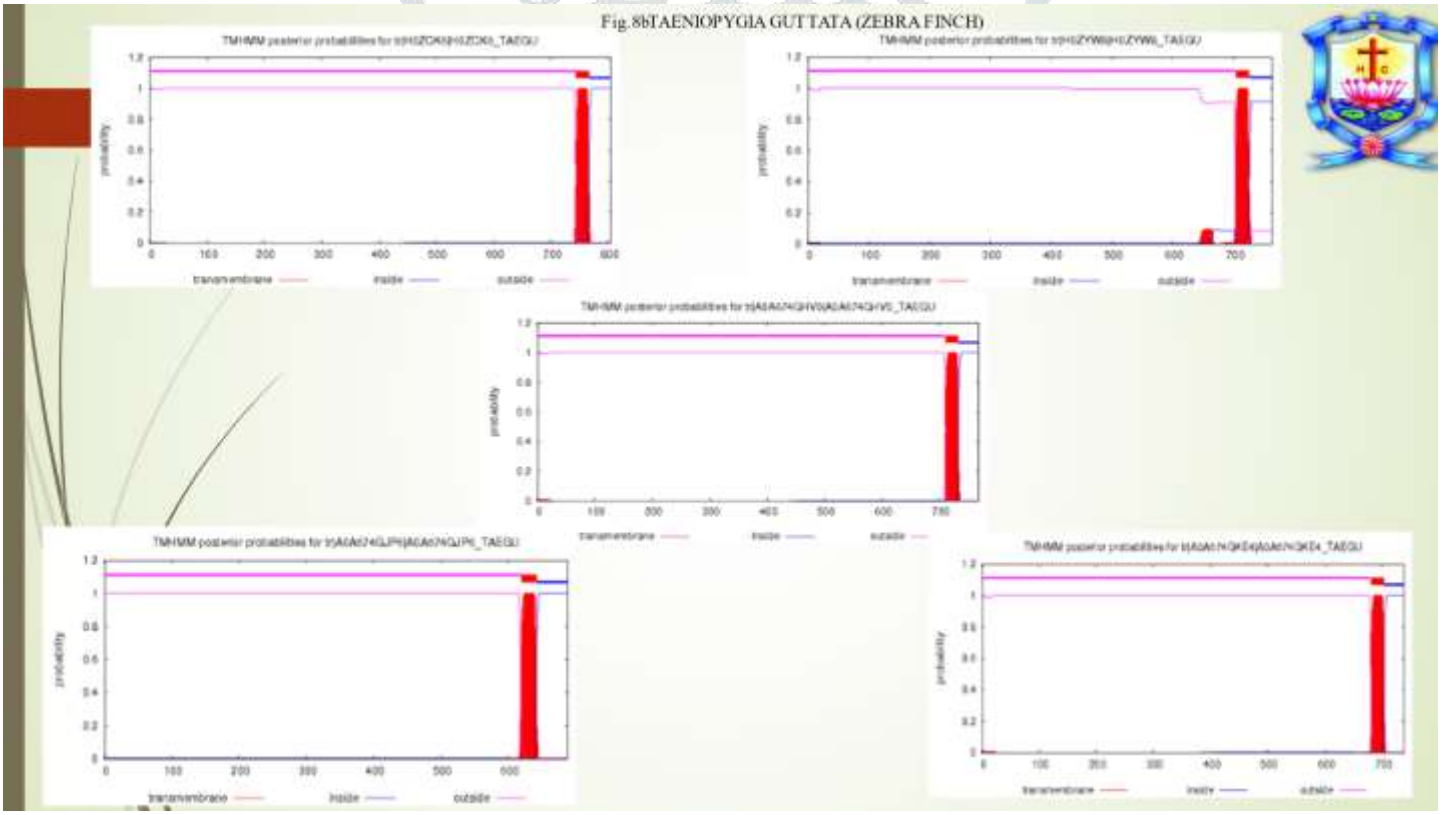


Fig.8bTAENIOPYGIA GUTTATA (ZEBRA FINCH)



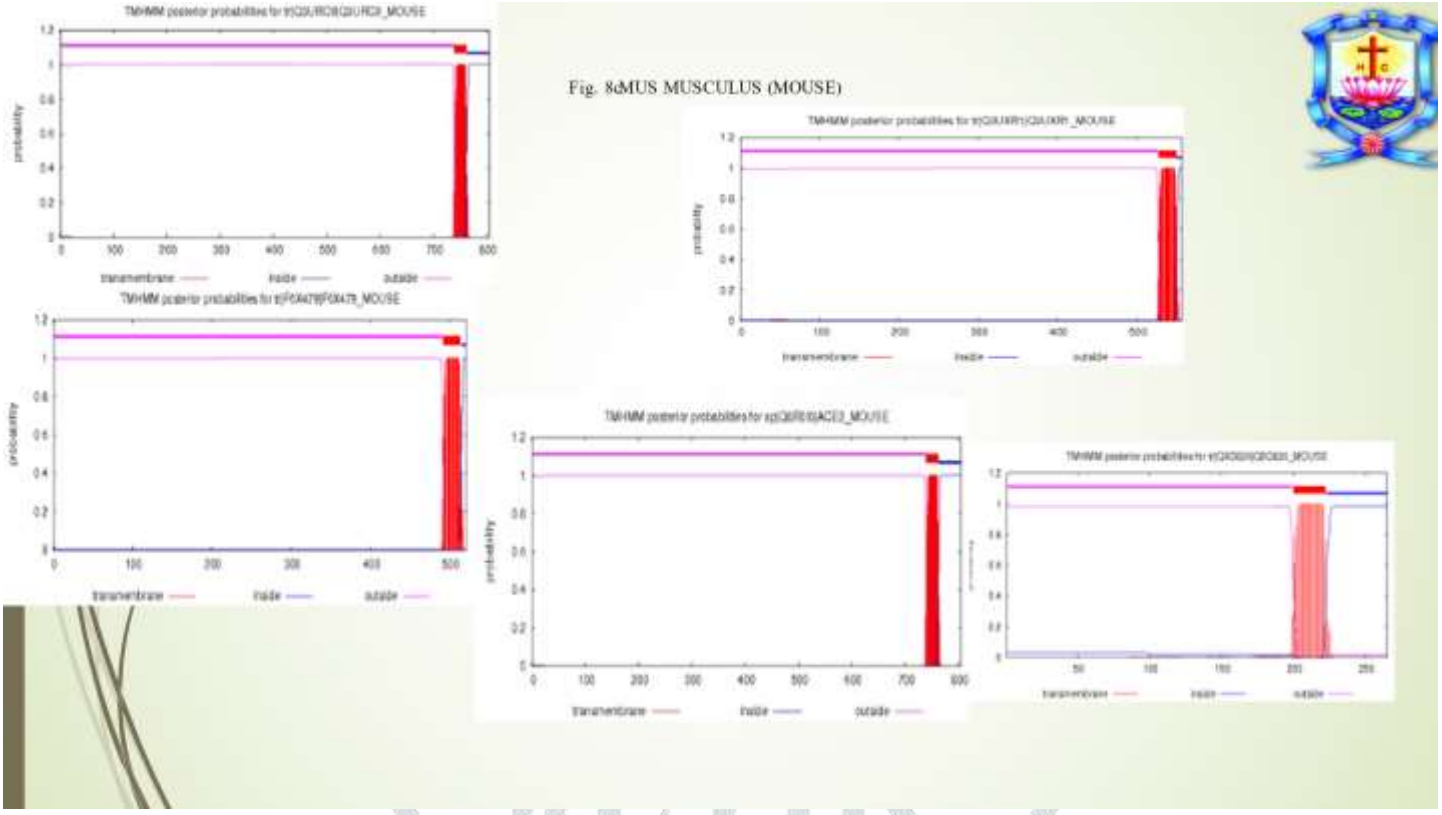


Fig. 8cMUS MUSCULUS (MOUSE)

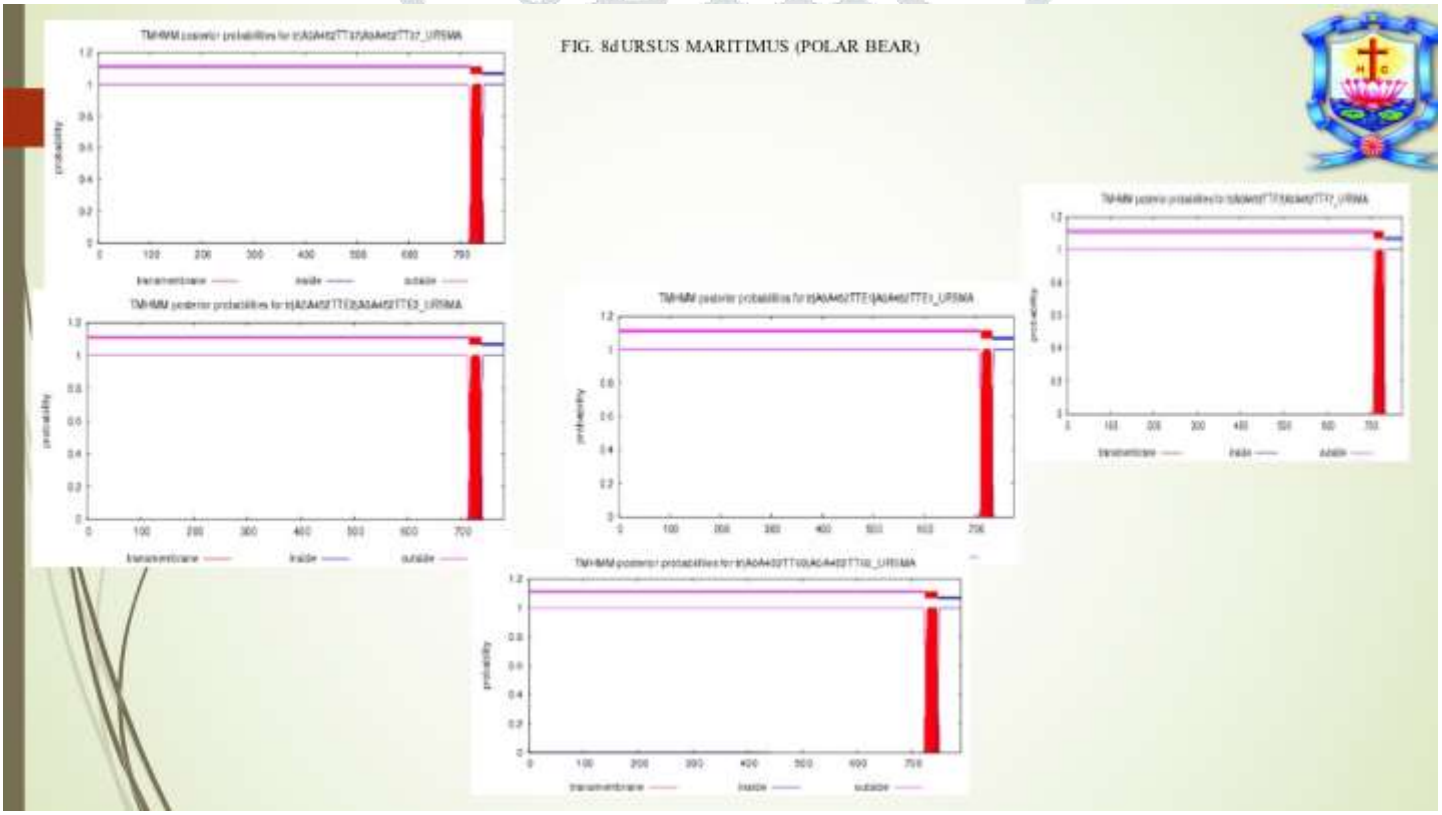


FIG. 8d URSUS MARITIMUS (POLAR BEAR)

Fig.8dHOMO SAPIENS (HUMAN)

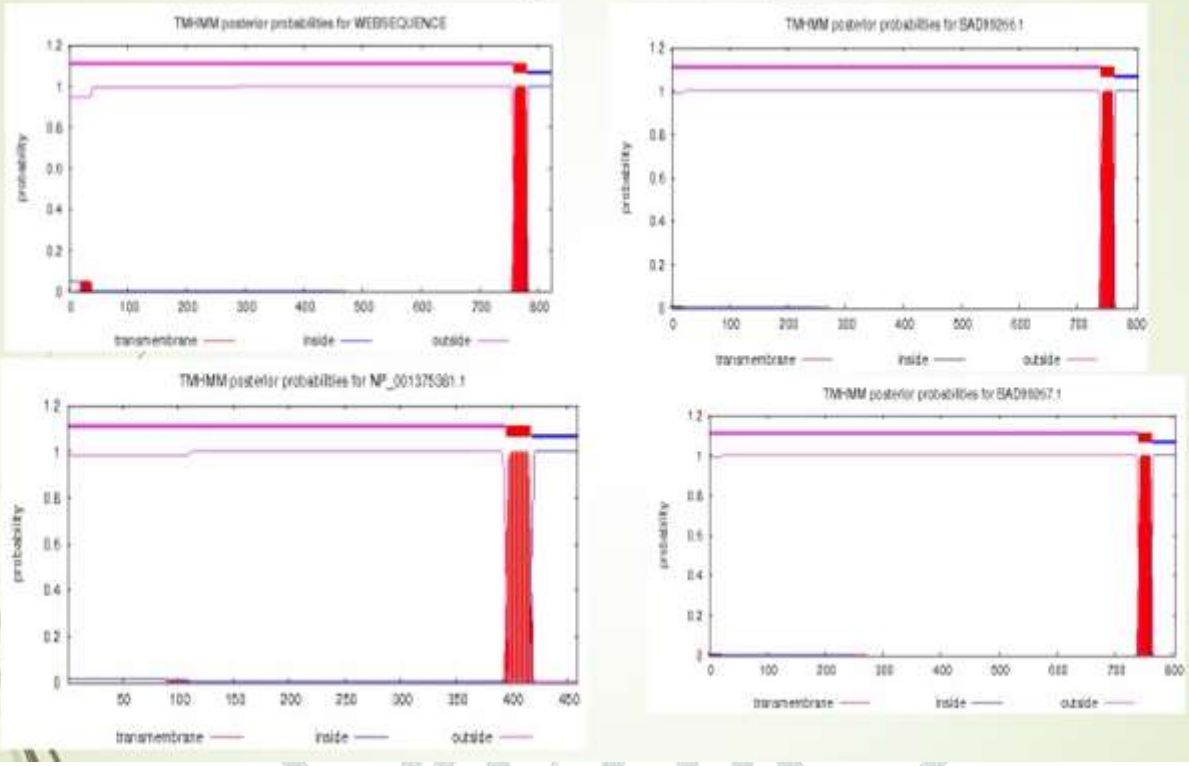


Fig.9aCALLORHINCHUS MILII (GHOST SHARK)

>tr | A0A4W3I1M1 | A0A4W3I1M1_CALMI Angiotensinconverting enzyme OS=CallorhinchusmiliiOX=7868 GN=ace2 PE=3 SV=1

LIST OF DOMAIN HITS

NAME	ACCESSION	DESCRIPTION	INTERVAL	E-VALUE
peptidaseM2	Pfam01401	Angiotensin-converting enzyme; members of this family are dipeptidyl carboxydipeptidases...	22-610	0e+00
Collectrin	pfam16959	Renal amino acid transporter; Collectrin is a single-pass transmembrane protein that is ...	620-758	1.01e-60

>tr | A0A4W3HYL6 | A0A4W3HYL6_CALMI Angiotensinconverting enzyme OS=CallorhinchusmiliiOX=7868 GN=ace2 PE=3 SV=1

LIST OF DOMAIN HITS

NAME	ACCESSION	DESCRIPTION	INTERVAL	E- VALUE
peptidase M2	Pfam01401	Angiotensin-converting enzyme; members of this family are dipeptidyl carboxydipeptidases...	22-554	0e+00
Collectrin	pfam16959	Renal amino acid transporter; Collectrin is a single-pass transmembrane protein that is ...	564-715	2.31e-85

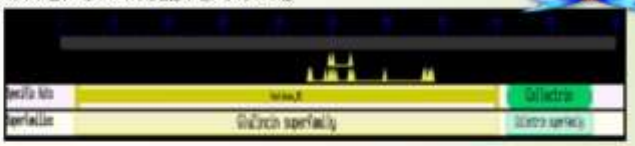
Fig 9b TAENIOPYGIA GUTTATA (ZEBRA FINCH)

>tr | H0ZCK6 | H0ZCK6_TAEGU Angiotensin-converting enzyme OS=Taeniopygia guttata OX=59729 GN=ACE2 PE=3 SV=2

LIST OF DOMAIN HITS

NAME	ACCESSION	DESCRIPTION	INTERVAL	E- VALUE
peptidase-m2	Pfam01401	Angiotensin-converting enzyme; Members of this family are dipeptidyl carboxydipepti dases ...	25-610	0e+00
Collectrin	pfam16959	Renal amino acid transporter; Collectrin is a single-pass transmembran e protein that is ...	621-772	6.78e-96

>tr | H0ZYW8 | H0ZYW8_TAEGU Angiotensin converting enzyme OS=taeniopygiaguttata OX=59729 GN=ACE2 PE=3 SV=2



LIST OF DOMAIN HITS

NAME	ACCESSION	DESCRIPTION	INTERVAL	E- VALUE
peptidase-m2	Pfam01401	Angiotensin-converting enzyme; Members of this family are dipeptidyl carboxydipepti dases ...	21-603	0e+00
Collectrin	pfam16959	Renal amino acid transporter; Collectrin is a single-pass transmembrane protein that is ...	614-731	6.28e-64

Fig.9c MUS MUSCULUS (MOUSE)

>tr|F6X479|F6X479_MOUSE Angiotensin converting enzyme (Fragment) OS=Mus musculus OX=10090 GN=Ace2 PE=1 SV=2

LIST OF DOMAIN HITS

NAME	ACCESSION	DESCRIPTION	INTERVAL	E- VALUE
GluZincin super family	cl14813	Gluzincin Peptidase family (thermolysinlike proteinases, TLPs) which includes peptidases M1, ...	1-358	0e+00
Collectrin	pfam16959	Renal amino acid transporter; Collectrin is a single-pass transmembrane protein that is ...	369-521	3.19e-88

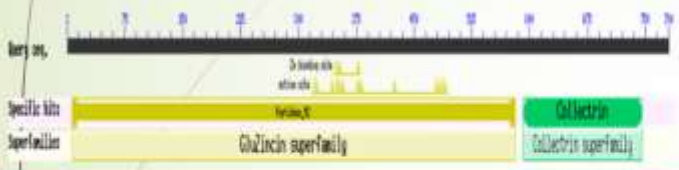
>tr | Q3UXR1 | Q3UXR1_MOUSE Angiotensin converting enzyme (Fragment) OS=Mus musculus OX=10090 GN=Ace2 PE=2 SV=1

LIST OF DOMAIN HITS

NAME	ACCESSION	DESCRIPTION	INTERVAL	E- VALUE
GluZincin super family	cl14813	Gluzincin Peptidase family (thermolysinlike proteinases, TLPs) which includes peptidases M1, ...	7-394	0e+00
Collectrin	pfam16959	Renal amino acid transporter; Collectrin is a single-pass transmembrane protein that is ...	405-557	1.24e-87

Fig.9dURSUS MARITIMUS (POLAR BEAR)

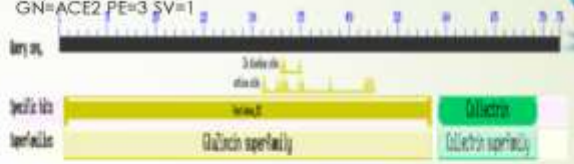
>tr | A0A452TTF7 | A0A452TTF7_URSMA Angiotensinconverting enzyme OS=Ursus maritimus OX=29073 GN=ACE2 PE=3 SV=1



LIST OF DOMAIN HITS

NAME	ACCESSION	DESCRIPTION	INTERVAL	E- VALUE
peptidase-m2	Pfam01401	Angiotensin-converting enzyme: Members of this family are dipeptidyl carboxydipeptidases ...	7-581	0e+00
Collectrin	pfam16959	Renal amino acid transporter; Collectrin is a single-pass transmembrane protein	592-745	5.52e-102

>tr | A0A452TTE1 | A0A452TTE1_URSMA Angiotensin-converting enzyme OS=Ursus maritimus OX=29073 GN=ACE2 PE=3 SV=1

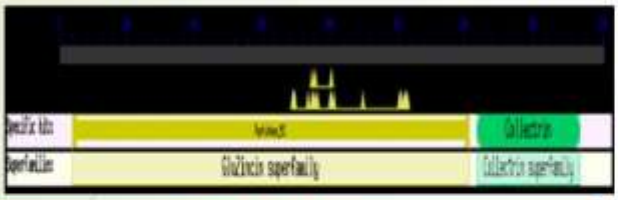


LIST OF DOMAIN HITS

NAME	ACCESSION	DESCRIPTION	INTERVAL	E- VALUE
peptidase-m2	Pfam01401	Angiotensin-converting enzyme: Members of this family are dipeptidyl carboxydipeptidases...	8-577	0e+00
Collectrin	pfam16959	Renal amino acid transporter; Collectrin is a single-pass transmembrane protein that is ...	588-741	3.12e-102

Fig. 9e HOMO SAPIENS (HUMAN)

>BAB40370.1 ACE2 [Homo sapiens]



LIST OF DOMAIN HITS

NAME	ACCESSION	DESCRIPTION	INTERVAL	E- VALUE
peptidase-m	Pfam01401	Angiotensin-converting enzyme; Members of this family are dipeptidyl carboxydipeptidases..	19-606	0e+00
Collectrin	pfam16959	Renal amno acid transporter;Collectrin is a single-pass transmembrane protein that is ...	617-770	3.28e-83

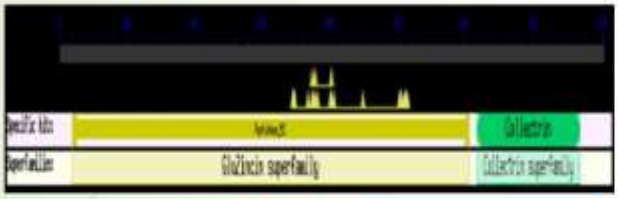
>AAQ89076.1 ACE2 [Homo sapiens]

LIST OF DOMAIN HITS

NAME	ACCESSION	DESCRIPTION	INTERVAL	E- VALUE
peptidaseM2	Pfam01402	Angiotensin-converting enzyme; members of this family are dipeptidyl carboxydipeptidases..	19-554	0e+00

Fig. 9e HOMO SAPIENS (HUMAN)

>BAB40370.1 ACE2 [Homo sapiens]



LIST OF DOMAIN HITS

NAME	ACCESSION	DESCRIPTION	INTERVAL	E- VALUE
peptidase-m	Pfam01401	Angiotensin-converting enzyme; Members of this family are dipeptidyl carboxydipeptidases..	19-606	0e+00
Collectrin	ptam16959	Renal amno acid transporter;Collectrin is a single-pass transmembrane protein that is ...	617-770	3.28e-83

>AAQ89076.1 ACE2 [Homo sapiens]

LIST OF DOMAIN HITS

NAME	ACCESSION	DESCRIPTION	INTERVAL	E- VALUE
peptidaseM2	Pfam01402	Angiotensin-converting enzyme: members of this family are dipeptidyl carboxydipeptidases..	19-554	0e+00

Fig. 10a CALLORHINCHUS MILII(GHOST SHARK)

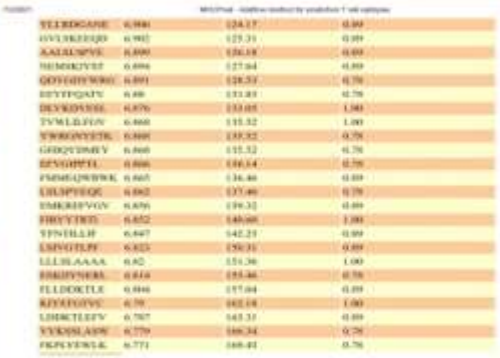


Fig.10b TAENIOPYGIA GUTTATA (ZEBRA FINCH)

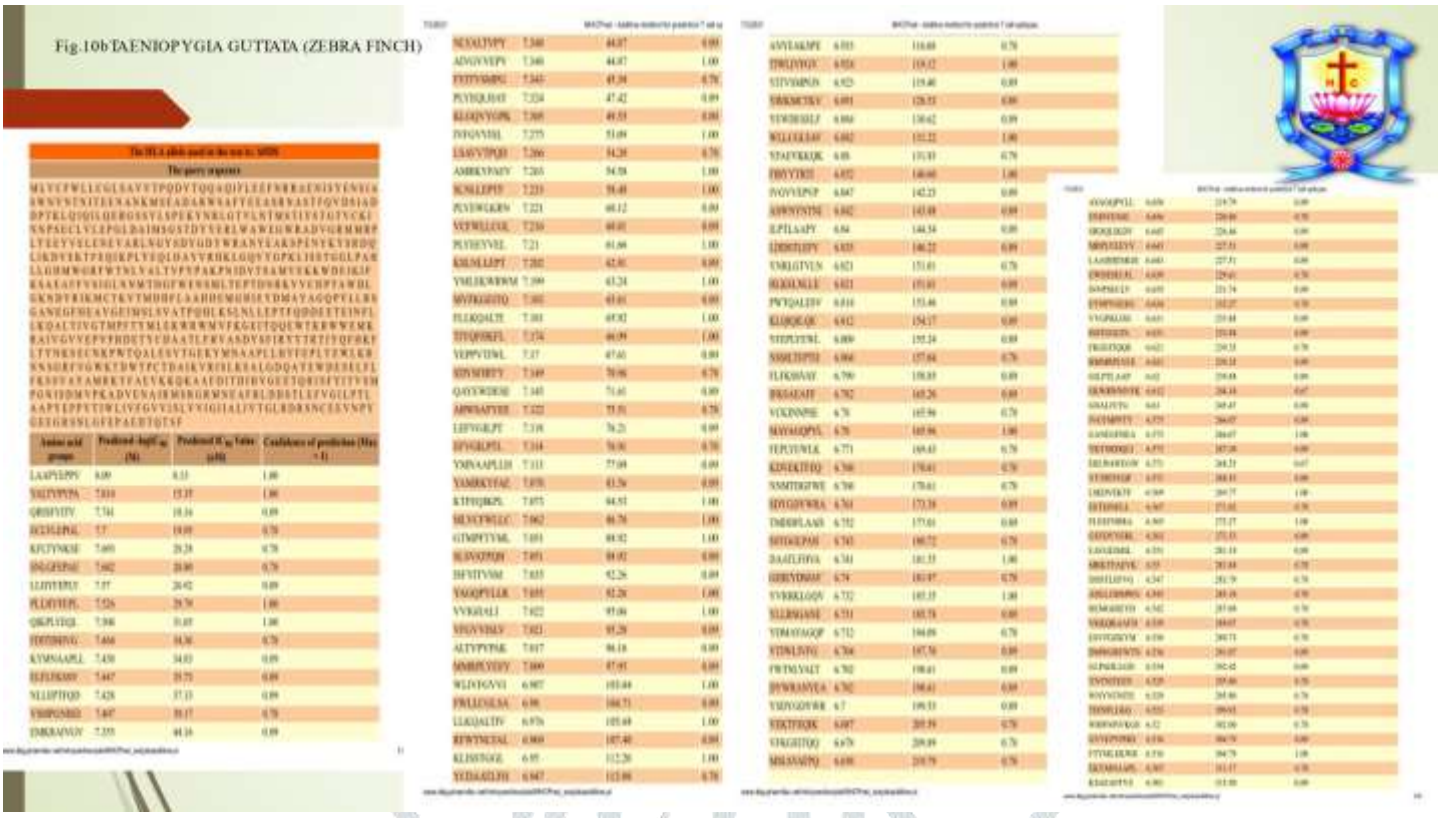


Fig.10c MUS MUSCULUS (MOUSE)

Fig.10c MUS MUSCULUS (MOUSE)

GLNDINSLEF	5.985	1035.14	1.00
SFVGWNTEW	5.963	1088.93	0.78
NAYEWTNSE	5.956	1106.62	0.89
RINDVFGLN	5.954	1111.73	0.89
SDLKPRVSF	5.910	1207.81	0.89
ALGANAYEW	5.901	1256.03	0.89
TNNEMFLFR	5.885	1303.17	0.89
ADQSIKVRI	5.879	1321.30	0.89
SNAGFQNSD	5.861	1377.21	0.78
PPYQPPVTI	5.85	1412.54	1.00
VTSPQNVSD	5.854	1465.55	0.89
IWLEIFGVV	5.822	1586.61	0.89
MALVVVGII	5.81	1548.82	1.00
EWSPYADQS	5.809	1552.39	0.78
RKVFSIKN	5.8	1584.89	0.78
QNRNSFVGW	5.798	1592.21	0.78
LIEFGVVMA	5.785	1640.59	1.00
PRSEVEDAI	5.772	1690.44	0.89
WLKEQNRNS	5.763	1725.84	1.00
DNSLEFLGI	5.761	1733.80	0.89
DVKPLENYF	5.76	1737.80	1.00
SIKVRISLK	5.755	1757.92	0.89
VIPRSEVED	5.728	1870.68	0.89
VGWNTEWSP	5.711	1945.36	0.89
CDISNSTIA	5.7	1995.26	0.89
EPWTKALEN	5.694	2023.02	0.78
KVRISLKSA	5.688	2051.16	1.00
IILIVTGIK	5.684	2070.14	0.89
PVADQSIKV	5.683	2074.91	0.89
LFRSSVAYA	5.671	2133.04	0.89
SEVEDAIRM	5.662	2177.71	0.89
EDVRVSDLK	5.656	2208.00	0.67
NSFVGWNTE	5.652	2228.44	0.78
YDSMDHGKG	5.649	2243.88	0.78
KKNETKREE	5.641	2295.60	0.78
FGLNDINSLE	5.629	2349.63	0.89
FQNSDDAQT	5.622	2387.81	1.00
NSEPWTKAL	5.617	2415.46	0.89
NEMFLFRSS	5.617	2415.46	0.89
DISNSTEAG	5.614	2432.20	0.89
TEAGQKLLK	5.606	2477.42	0.78

FQPLFDWLK	5.602	2506.35	0.89
RIVAYAMRK	5.601	2506.11	0.78
NSYDSMDHG	5.598	2523.48	0.78
SEKNQTVF	5.564	2728.94	0.89
FLEEDVRVS	5.561	2747.89	1.00
VFGLNDNSL	5.556	2770.71	0.89
TLEPFVQPF	5.55	2814.38	0.89
RKKKNETKR	5.549	2824.88	0.89
NSEDAQTTF	5.546	2844.46	0.89
WSPYADQSI	5.484	3286.95	0.89
NSTLAGQKL	5.459	3475.36	0.89
VEDAIRMSR	5.453	3523.71	0.89
PYQPPVTIW	5.442	3614.10	0.78
WTNNEMFLF	5.426	3748.73	1.00
KNETKREEN	5.345	4128.98	0.78
SRGRINDVF	5.339	4175.22	0.89
VAYAMRKYF	5.340	4487.45	1.00
NNEMFLFRS	5.344	4526.98	0.89
GFQNSDDAQ	5.342	4549.88	0.78
VSDLKPRVS	5.333	4645.15	0.89
ANAYEWTNN	5.33	4677.35	0.78
LKPRVSFVF	5.31	4677.35	0.89
AGQKLLKML	5.326	4726.63	1.00
YSPQNVSDV	5.324	4742.42	0.89
DVIPRSEVE	5.324	4742.42	0.89
RVAYAMRKY	5.315	4864.07	0.89
LKSALGANA	5.312	4875.28	0.89
ENVVGAROM	5.31	4897.79	0.78
SEPWTKALE	5.31	4897.79	0.78
VVGARNMDV	5.300	4977.37	1.00
GNSEPWTKA	-	-	-
WTKALENVV	-	-	-
LKMLSLGNS	-	-	-
TKALENVVG	-	-	-
DVRVSDLKP	-	-	-
LGIIPTLIF	-	-	-
HMSRGRIN	-	-	-
GARNMDVKF	-	-	-
RPLKSALG	-	-	-
SNETEAGQK	-	-	-

Fig.10d URSUS MARITIMUS (POLAR BEAR)

The HLA allele used in the test is: A0201

The query sequence

LCSSCLSEDLYYQSSLASWNYNTNITNENIQKMNDAGAKWSAFY
EEQSKHAKTYPLEEIHNSTVKRQLQALQHSGSSVLSADKSQRLN
TILNAMSTIYSTGKACNPNNPQECLLLEPGLDDIMENSKDYNERL
WAWEGWRSEVGKQLRPLYEEYVALKNEMARANNYEDYGDYWR
GDYEEEWTDGYNYSRNQLIEDVEHTFTQIKALYEHLHAYVRAKL
MDTVPSHRISPTGCLPAHLLGDMWGRFWTNLYPLTIPFGQKPNID
VTDAMVNQNWDARRIFEEAEKFFVSVGLPNMTQEFWENSMLTE
PGDGQKVVCHPTAWDLGKGDFRIKMCTKVTMDDFLTAHHEMG
HIQYDMAYAEQPFLLRNGANEGFHEAVGEIMSLSAATPNHLKNI
GLLPPGFSEDNETEINFLLKQALTIVGTLPFTYMLEKWRWMVFQ
GKIPKDQWMKKWWEMKREIVGVVEPLPHDETYCDPASLFHVAN
DYSFIRYYTRTIYQFQFQEALCQIAKHEGPLHKCDISNSSEAGKT
LLQMLRLGRSKPWTLALEHVVGAKNMDVRPLLNYFEPLFTWLK
EQNRNSFVGWNTDWSPYADQSIKVRISLKSALGEKAYEWNDNE
MYLFRSSIAYAMRKYFSEAKNQMIPFVEDNVWVNDLKPRISFNF
FVTSPGNVSDVIPRADVEGAIKMSHDRINDAFQLDDNSLEFLGIQ
PTLGPPYQPPVTIWLIVFGVVMGLVVIGIILLIFSGIRNRRKNDQA
RSEENPYASVDLSKGENNPGFQNADDVQTSF

Amino acid groups	Predicted -logIC50 (M)	Predicted IC50 Value (nM)	Confidence of prediction (Max = 1)
ALYSHLHAV	8.091	8.11	0.89
KNKILLPPG	8.059	8.73	0.78
LLNVFEPLF	7.94	11.48	1.00
AKNQMHPFV	7.902	12.53	0.89
KLMDTYPSH	7.823	15.03	1.00
NLYFLTIPF	7.814	15.35	1.00
ECLLLEPGL	7.595	25.41	0.78
PRISFNFFV	7.575	26.61	0.89
PLLNYFEPL	7.495	31.99	1.00
YAMRKYFSE	7.44	36.31	0.89
RFWTNLYPL	7.402	39.63	0.89
KNMDVRPLL	7.396	40.18	0.89
GLLPPGFSE	7.377	41.98	0.89
TIYQFQFQE	7.356	44.06	0.89
ACNPNNPQE	7.344	45.29	0.78

IVFGVVMGL	7.215	60.42	1.00
FWTNLYPLT	7.209	61.83	0.89
KHAKTYPLE	7.259	55.08	0.78
PLFTWLKEQ	7.25	56.23	0.89
YDMAYAEQP	7.202	62.81	0.78
YMLEKWRWM	7.199	63.24	1.00
EMKREIVGV	7.187	65.01	0.89
PLLKQALTI	7.181	65.92	1.00
QMHPFVEPN	7.18	66.07	0.89
AEQPFLLRN	7.16	69.18	0.78
PLYEEYVAL	7.158	69.59	1.00
GILLPPSG	7.154	70.15	0.89
VFQGKIPKE	7.146	71.45	0.78
IVGVVEPLP	7.135	73.28	0.89
EMYLFRSSI	7.134	73.45	0.89
AKNWSAFVEI	7.122	75.51	0.78
NDYSFRYYT	7.095	80.35	0.78
FFVSVGLPN	7.089	81.47	0.78
YVQSSLASW	7.081	82.99	0.78
DISNSLEFLG	7.078	83.56	0.78
IKALYEHLH	7.067	85.70	0.78
VVKGILLI	7.063	86.50	1.00
YEWNDNEMY	7.032	92.90	0.78
MVFQGKIPK	7.029	93.54	0.89
WLIVFGVVM	7.025	94.41	1.00
LMDTYPSHI	7.022	95.06	1.00
YCDPASLFH	7.021	95.28	0.78
NVFEPLFTW	7.019	95.72	0.78
LIKQALTIV	6.976	105.68	1.00
DIVGVVEPL	6.97	107.15	1.00
VMGLVVIGI	6.966	108.14	1.00
EFLGIQPTL	6.963	108.89	0.78
GTLPFTYML	6.948	112.72	1.00
ISFNFFVTS	6.946	113.24	0.89
GDYEEEWTD	6.938	115.35	0.78
TIWLIVFGV	6.924	119.12	1.00
QLRPLYEEY	6.92	120.25	0.89
LALEHVVGA	6.92	120.23	1.00
KEYPLEEBI	6.914	121.90	0.89
NSMLTEPGD	6.91	123.05	0.78
FEPLFTWLK	6.907	123.88	0.78

Fig.10e HOMO SAPIENS (HUMAN)

Amino acid groups	Predicted IC_{50} (nM)	Predicted IC_{50} Value (nM)	Confidence of prediction (Max = 1)
KLMNAVPSV	8.326	4.72	0.89
FLDKFNHEA	7.869	10.74	1.00
YSCTVPFQQ	7.94	11.49	0.78
PLYEELHAY	7.704	19.77	0.89
NLVSLTVPF	7.704	19.77	1.00
TLYQFQFQE	7.638	23.01	0.89
ECLLLEPGL	7.595	25.41	0.78
VCNPDNPQE	7.400	39.68	0.78
MVTKGEIPK	7.403	39.54	0.89
RFWTSEYSL	7.331	46.67	0.89
GLLSPDFQE	7.265	54.33	0.89
DDYSFHEYV	7.262	54.70	0.78
ESIGLLSPD	7.256	55.46	0.78
TLAQMYPLQ	7.255	55.65	0.89
EIVGVVEPV	7.215	60.95	0.89
YDMAYAAQP	7.206	62.52	0.78
YMLEKWAWM	7.199	63.24	1.00
FLLKQALTE	7.181	65.92	1.00

PLYEEYVVL	7.172	67.30	1.00
SSWLLLSLV	7.156	69.82	0.89
VFKGEIPKD	7.153	70.31	0.78
ALQQNGSSV	7.117	76.38	1.00
QNLTVKLQL	7.115	76.74	0.89
EMKREIVGV	7.091	81.10	0.89
FFVSVGLPN	7.089	81.47	0.78
EIKPLYEHL	7.072	84.72	0.89
MAYAAQPFL	7.049	89.33	1.00

Fig.11b TAENIOPYGIA GUTTATA (ZEBRA FINCH)

INPUT INFORMATION

OVERLAP DISPLAY

Fig.11c MUS MUSCULUS (MOUSE)

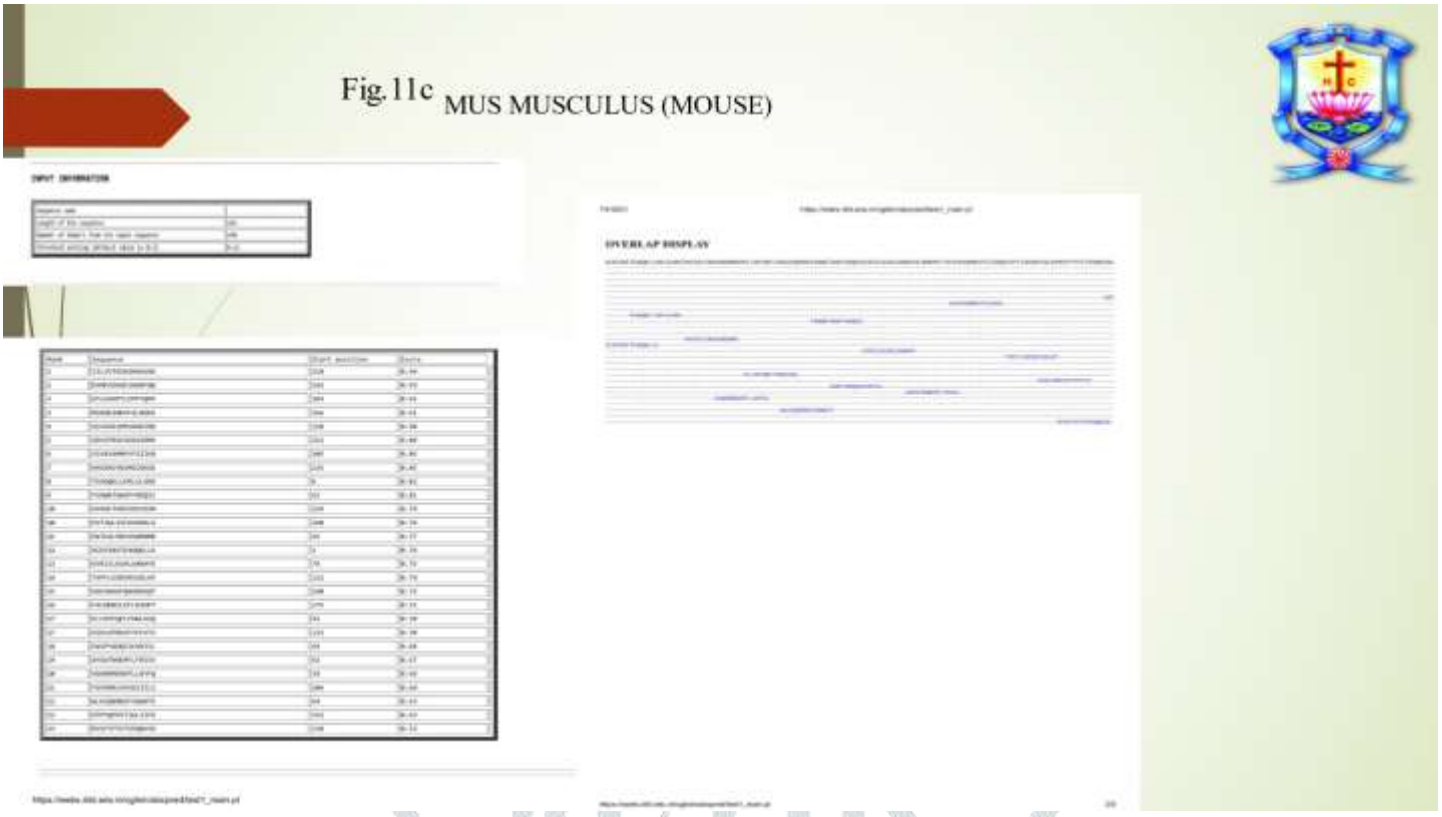


Fig.11d

URSUS MARITIMUS (POLAR BEAR)

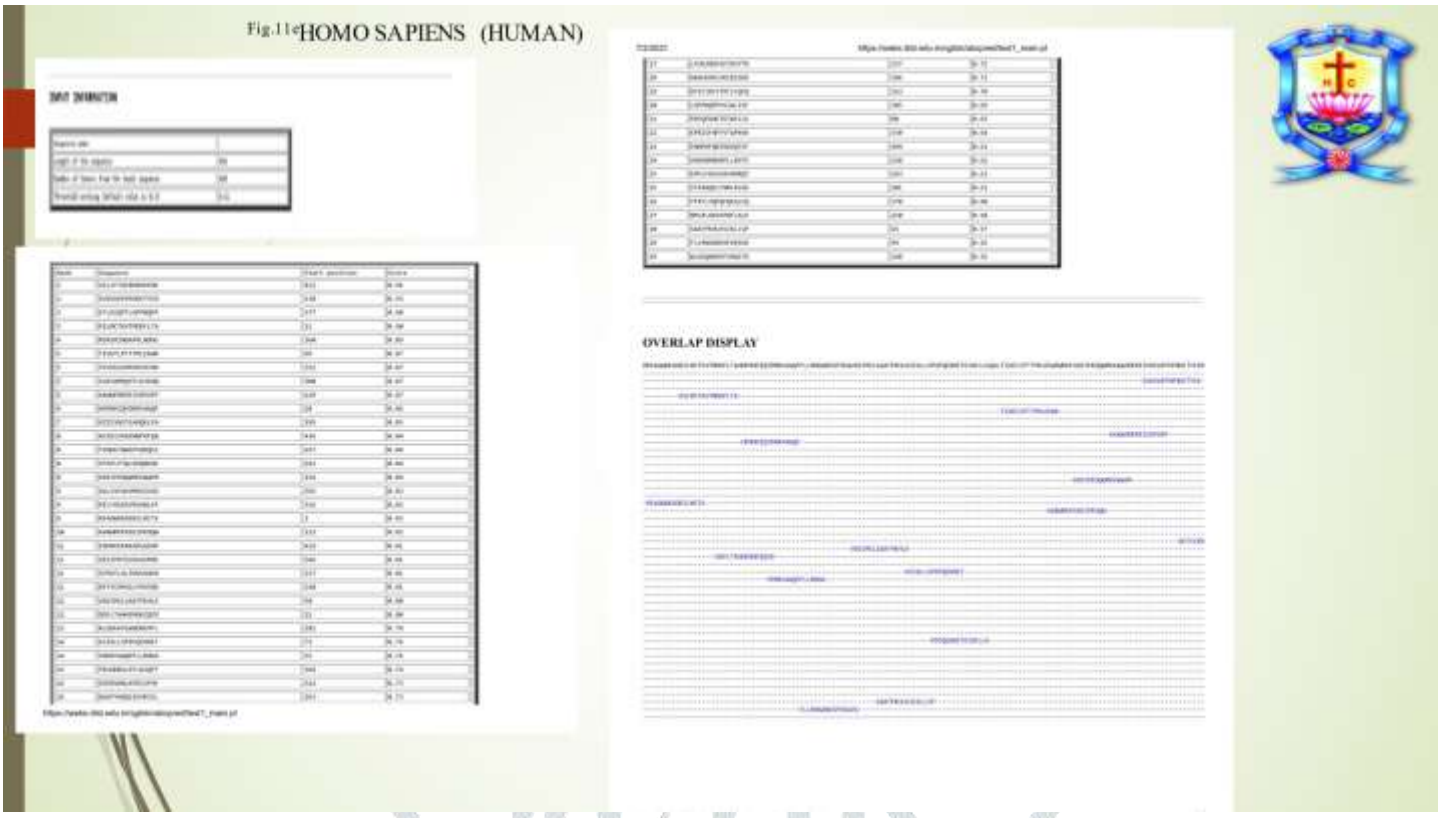


Fig.11eHOMO SAPIENS (HUMAN)
OVERLAP DISPLAY

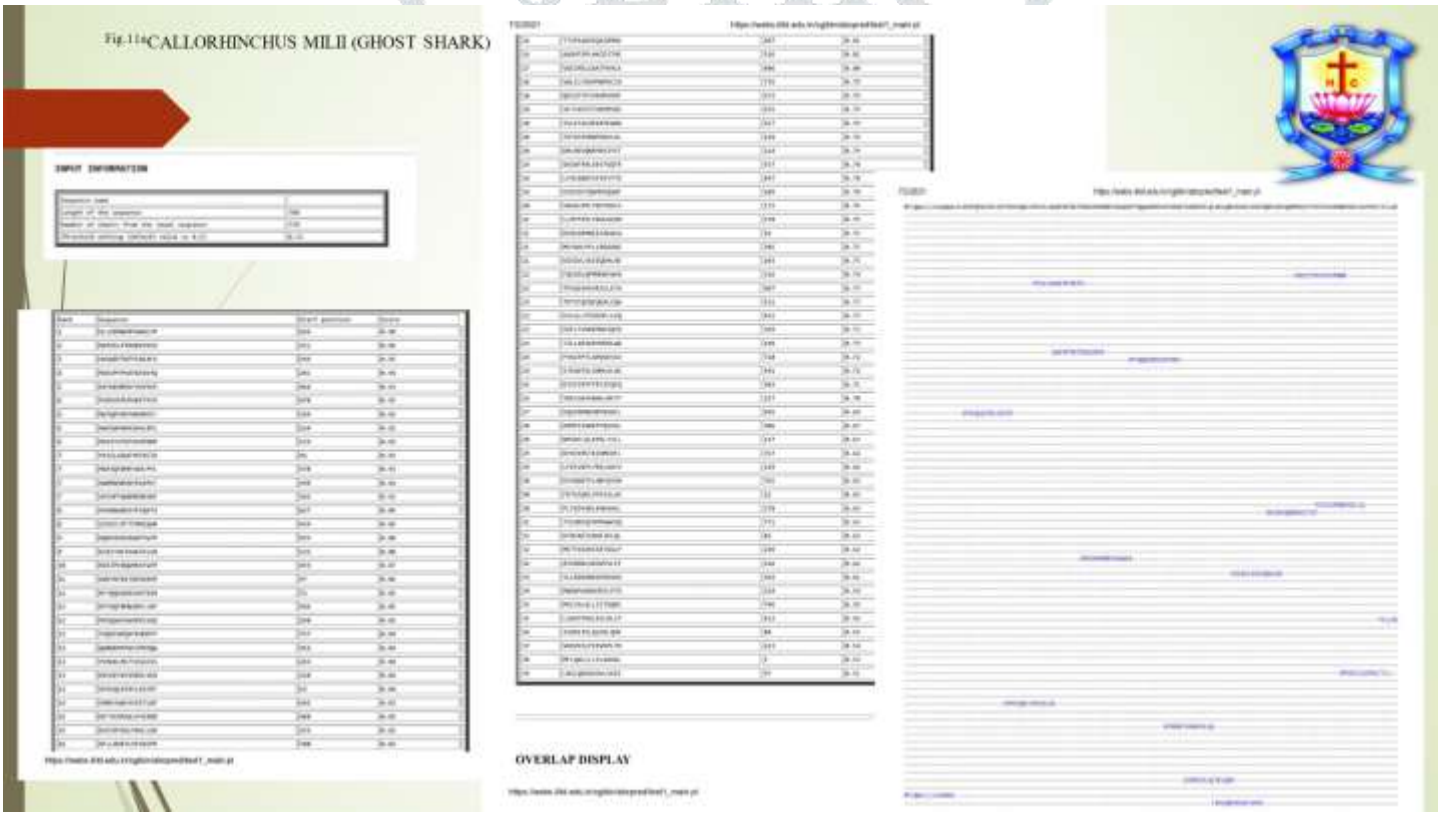


Fig.11aCALLORHINCHUS MILII (GHOST SHARK)
OVERLAP DISPLAY

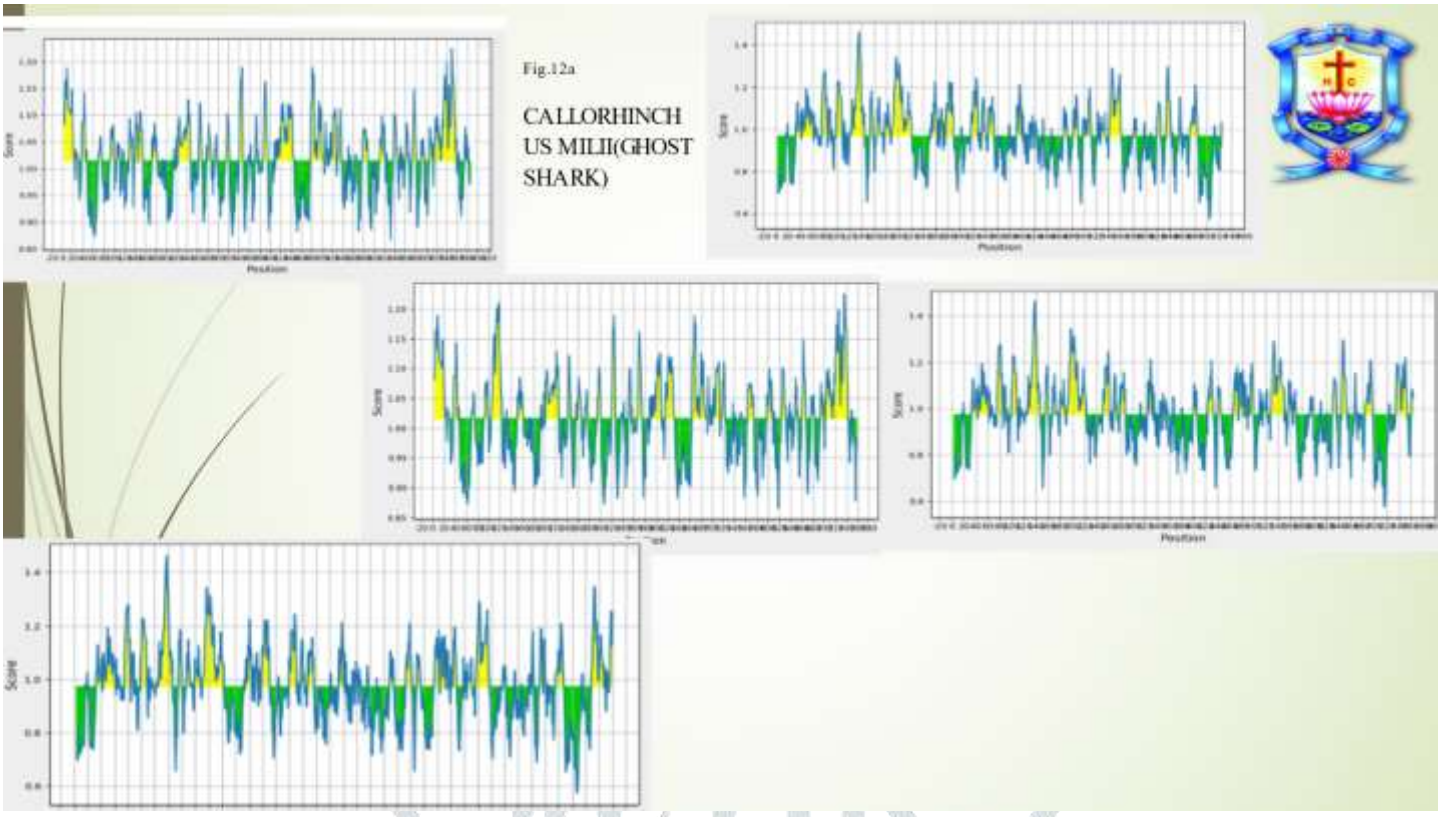


Fig.12a

CALLORHINCH US MILII(GHOST SHARK)

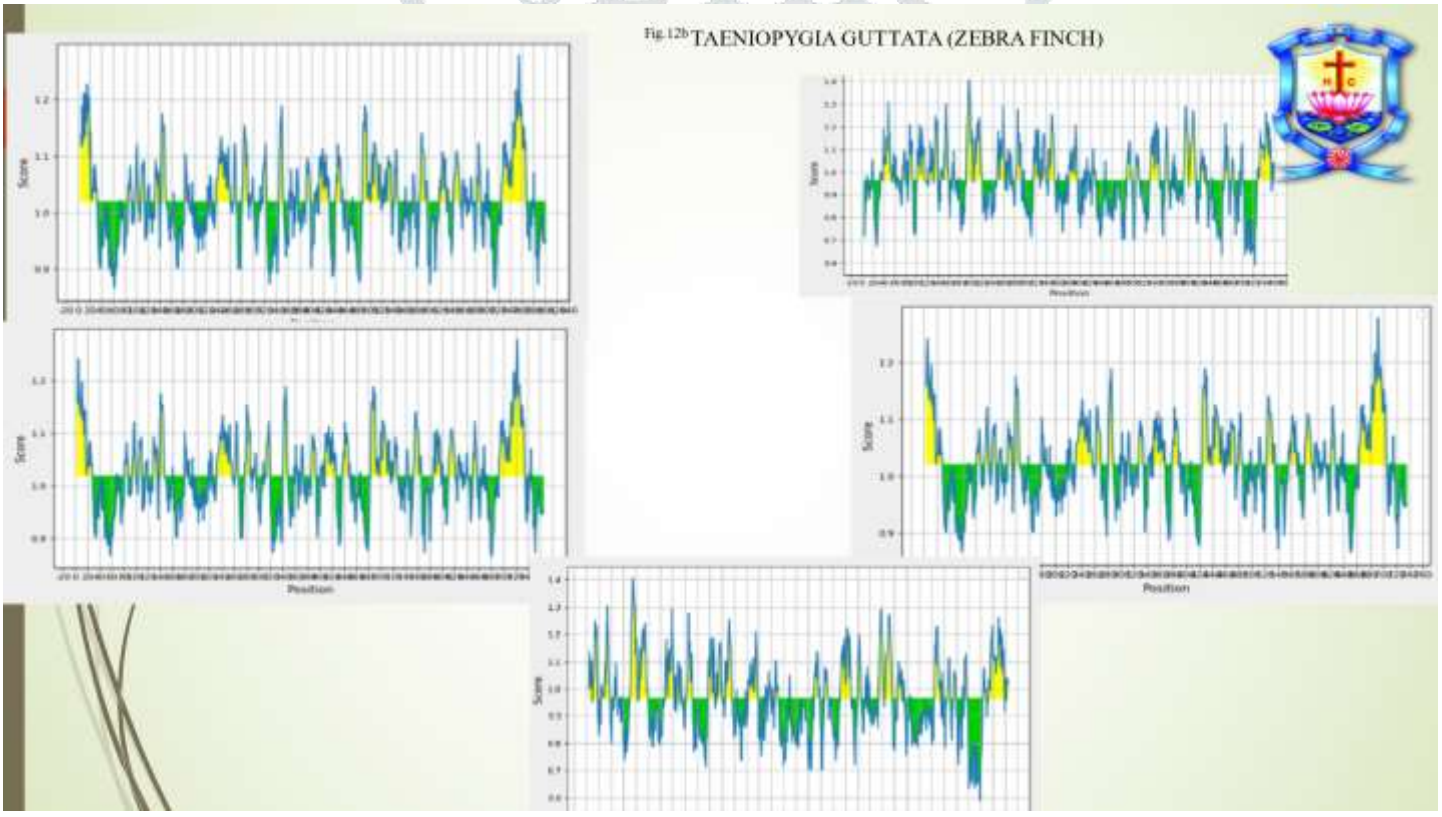


Fig.12b TAENIOPYGIA GUTTATA (ZEBRA FINCH)

Fig.12c MUS MUSCULUS (MOUSE)

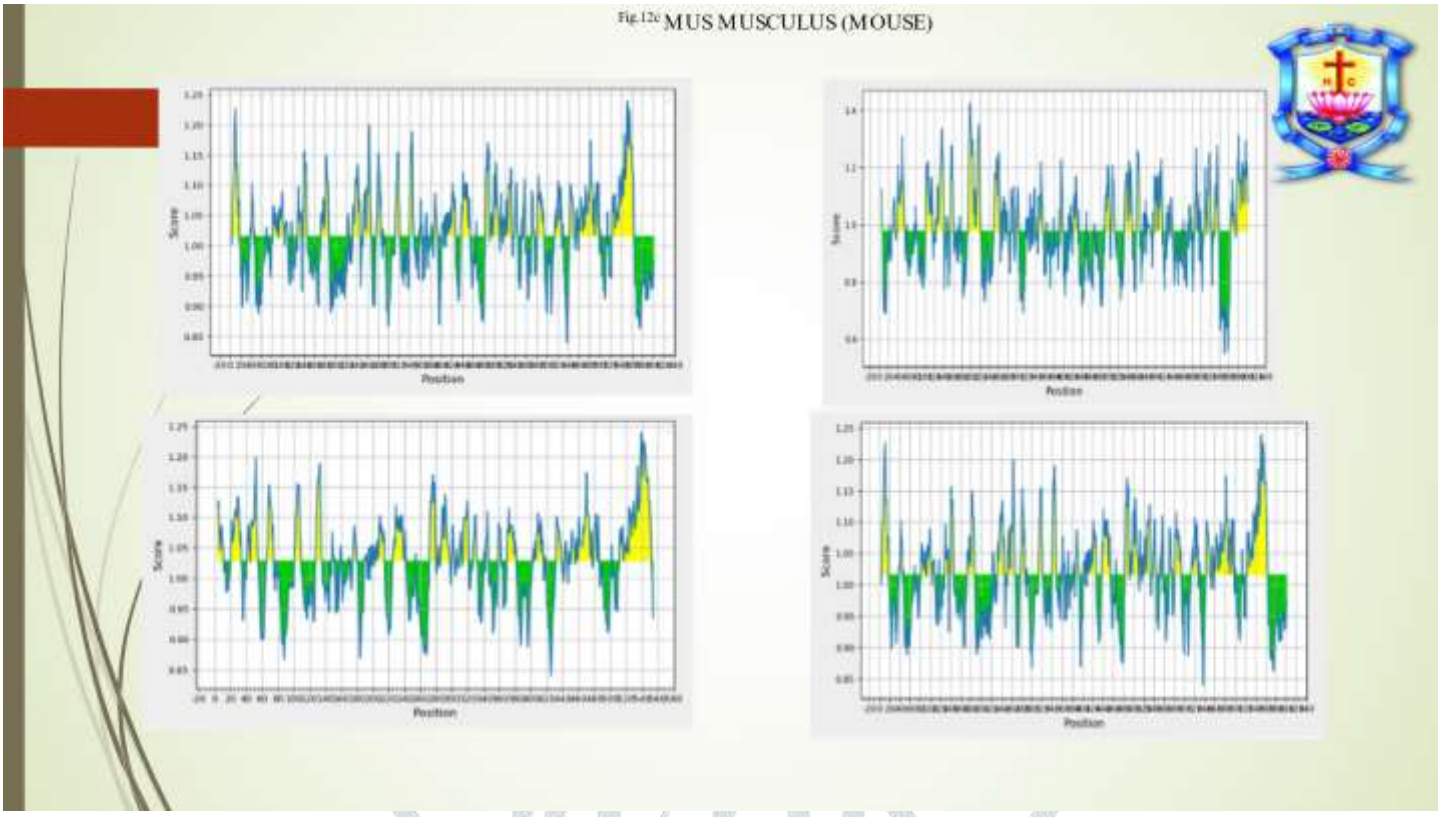


Fig.12d URSUS MARITIMUS (POLAR BEAR)

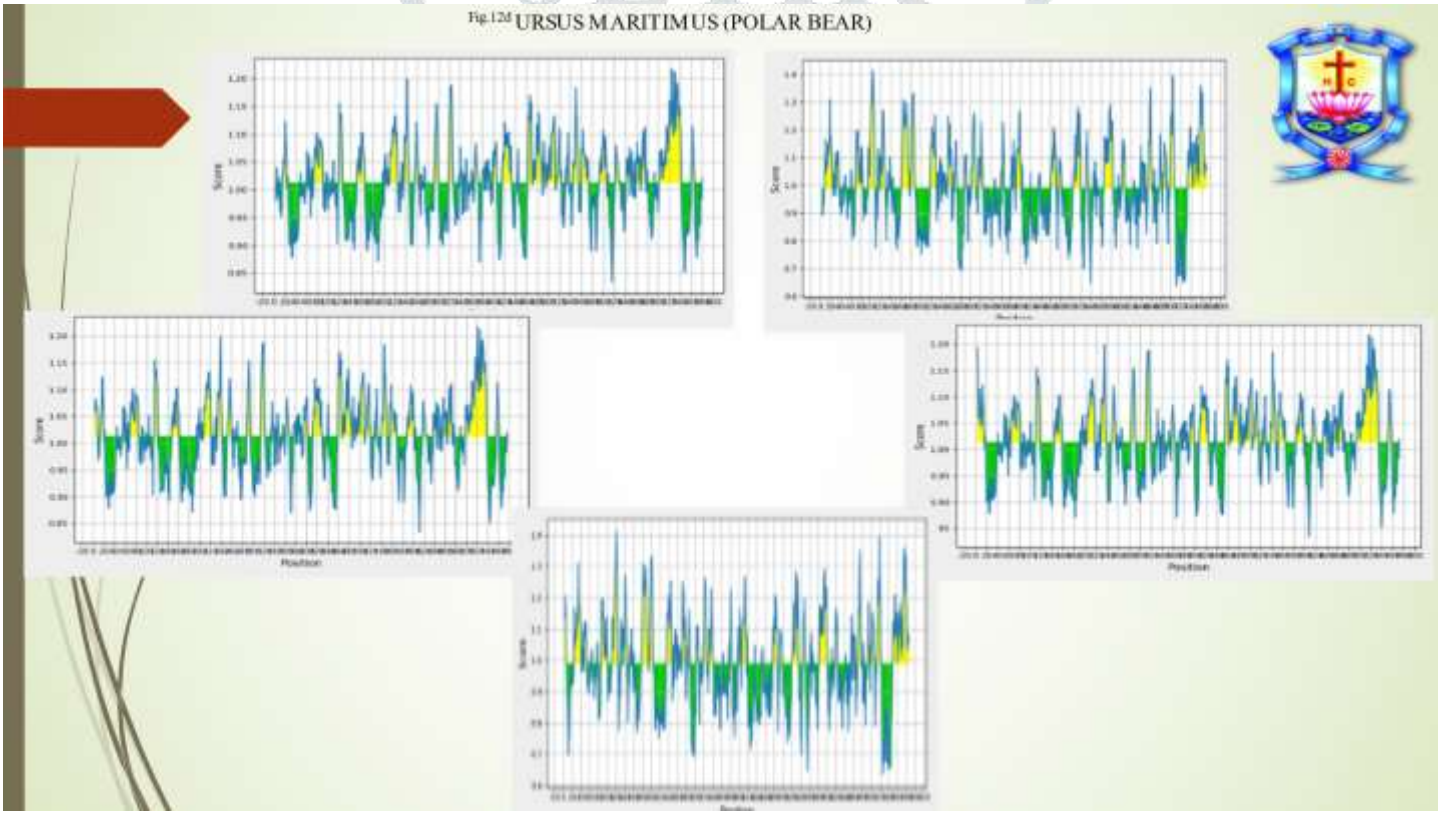


Fig.12e HOMO SAPIENS(HUMAN)

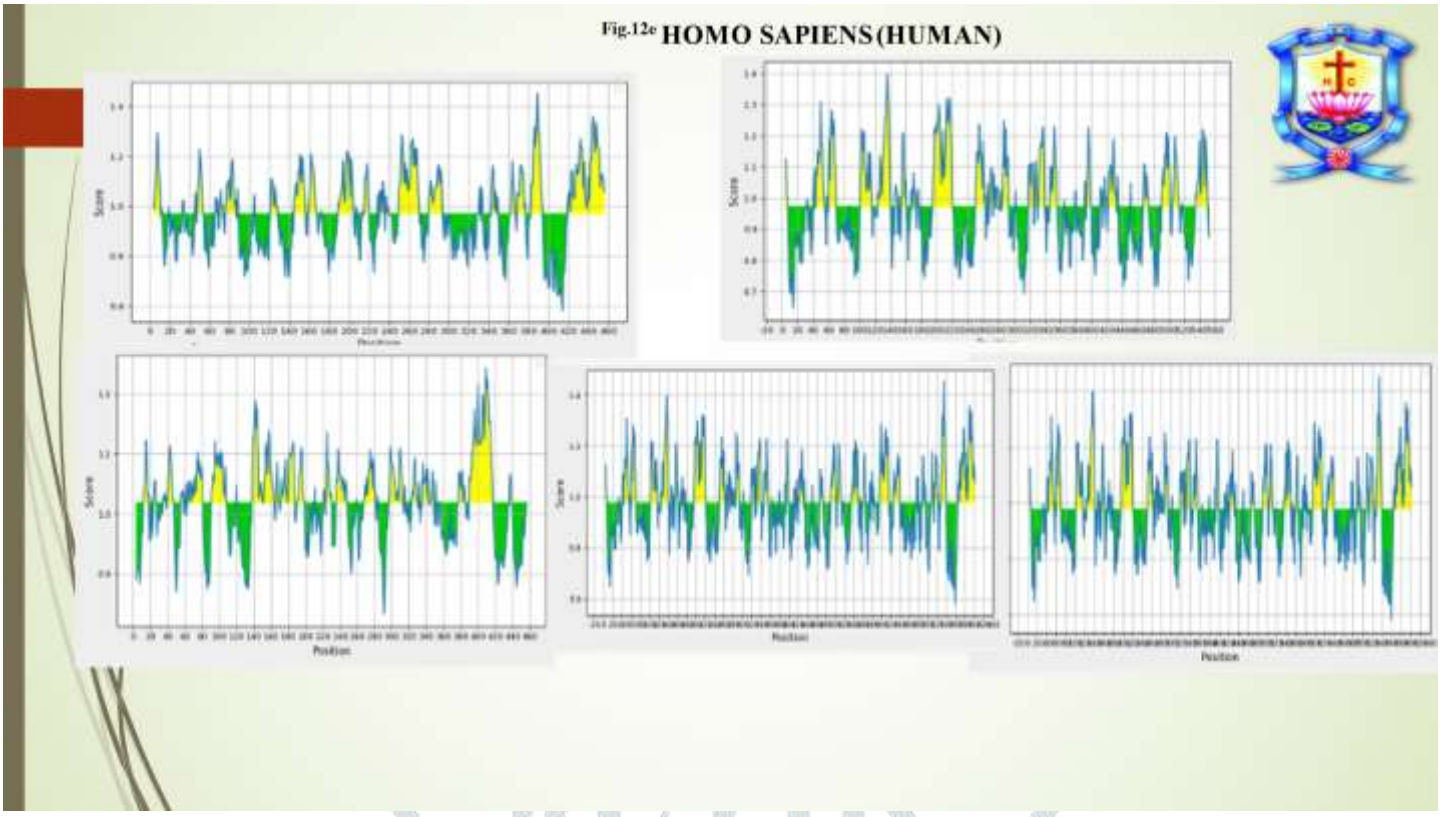


Fig.13a CALLORHINCHUS MILII (GHOST SHARK)

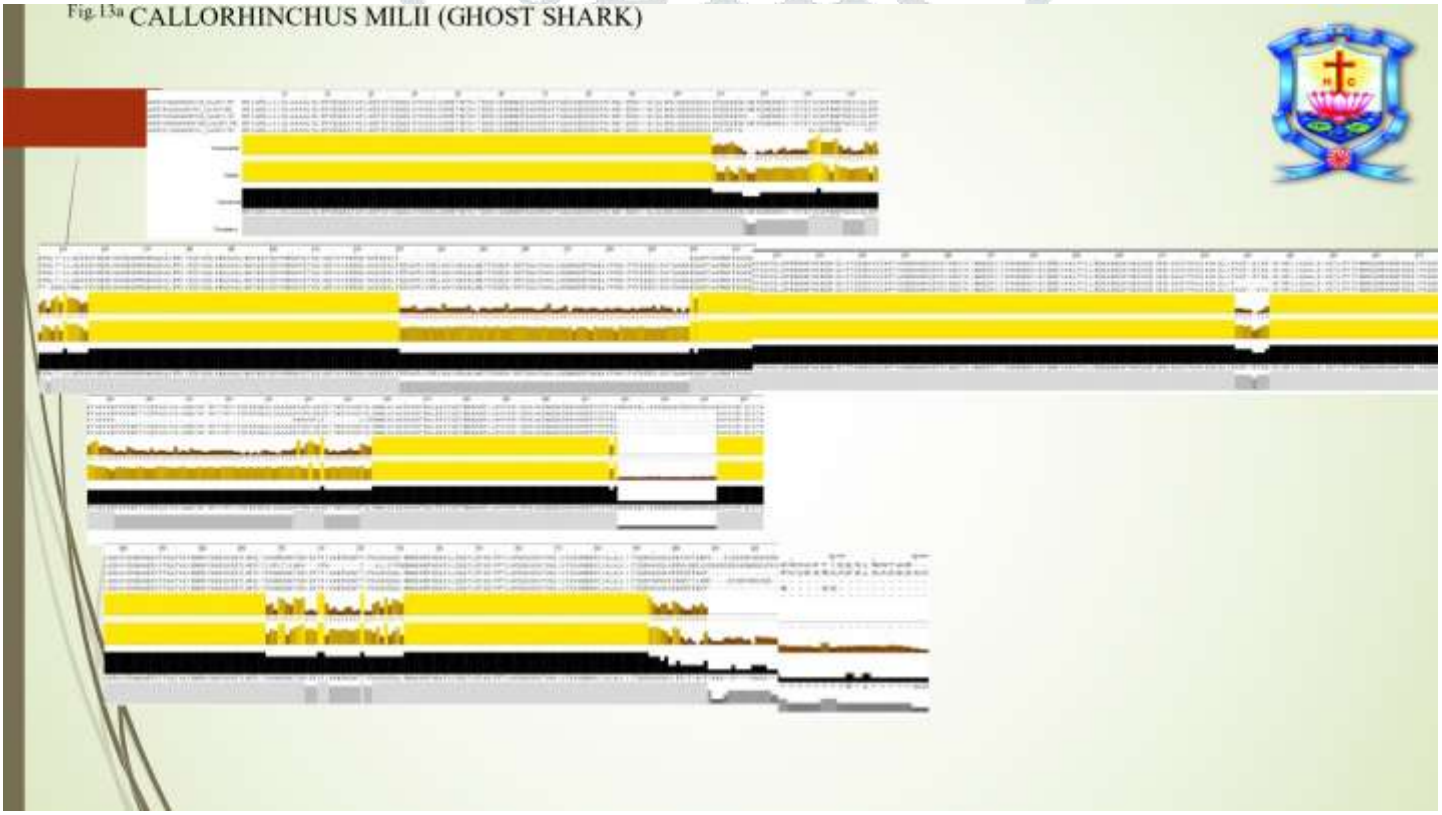


Fig. 13b TAENIOPYGIA GUTTATA (ZEBRA FINCH)

Fig. 13c MUS MUSCULUS (MOUSE)

Fig.13d URSUS MARITIMUS (POLAR BEAR)

Fig.13e HOMO SAPIENS (HUMAN)

We obtained the FASTA sequence of ACE2 in Zebra Finch, Ghost Shark, Mouse, Polar Bear, Human. Protein of SARS- COV2 from the NCBI database. The result was shown in the Fig.. (2a,2b,2c,2d,2e).After the complete physiochemical analysis of antigenic proteins were predicted for all ACE2. Protparum tool was used to predicit a physiochemical analysis.

In Ghost Shark the amino acid was present in 727-803. The theoretical pI was calculated to be 5.10-6. which shows that the final protein is Acidic. The estimated half - life is 30hours for mammalian reticulocytes in vitro condition. Also, the instability index (II) is computed to be 33.24-36.92. This classifies the protein as stable. The aliphatic index is displayed in the range 74.55-78.62. Grand average of hydropathicity (GRAVY) in range of -0.398- -0.495. the result was shownn in Fig. (3a,3b).

In Zebra Finch the amino acid was present in 688-804. The theoretical pI was calculated to be 5.10-6. indicating that the final protein is Acidic. The average half-life is 5.5 hours for mammalian reticulocytes in vitro. 3 minutes half-life of yeast in vivo, and 2 minutes half - life of Escherichia coli in vivo. Also, the instability index (II) is calculated to be 37.81-46.22. This classifies protein as stable and unstable. In II it shows more than 40.01 protein classifying as unstable. The aliphatic index is shown in range 77.27-79.22. GRAVY range -0.321- -0.386. the result was a shown in Fig. (4a, 4b).

In Mouse the amino acid was present in 265-805. Theoretical PI was calculated to be 5.10-6. Which indicates that the final protein is Acidic. The average half-life is 30 hours in mammalian reticulocytes in vitro condition. In the main yeast there is 20 and a half hours - life in vivo state, and Escherichia coli also more than 10 half hours life in vivo. Also, the instability index (II) is calculated to be 34.77-44.55. This distinction indicates protein as stable and unstable. In II it shows more than 40.01 protein classifying as unstable. the aliphatic index is shown in range 76.58-86.38. Maximum hydropathicity (GRAVY) range -0.193- -0.416. the result was a show in Fig. (5a, 5b).

In the Polar Bear the amino acid was present in 774-790. Theeretical pI was calculated to be 5.10-6. Which indicates that the final protein is Acidic. The average half-life is 30 hours in mammalian reticulocytes in vitro condition. In the main yeast there is 20 and a half hours - life in vivo state, and Escherichia coli also more than 10 half hours - life in vivo. Also, the instability index (II) is calculated to be 41.84-43.27. This distinction indicates protein as stable and unstable. In II it shows more than 40.01 protein classifying as unstable. the aliphatic index is shown in range 74.82-76.24. GRAVY range -0.482- -0.508. the result was a demonstration in Fig. (6a, 6b).

In humans the amino acid was present in 555-808. Theoretical pI was calculated to be 5.10-6.so the final protein is Acidic. The average half-life is 30 hours in mammalian reticulocytes in vitro condition. In the main yeast there is 20 and a half hours - life in vivo state, and Escherichia coli also more than 10 half hours - life in vivo. Also, the instability index (II) is calculated to be 40.09-51.74. This indicates protein as stable and unstable. In II it shows more than 40.01 protein classifying as unstable. The aliphatic index is displayed in the range 70.14-80.55. Maximum hydropathicity (GRAVY) range -0.185 - -0.423. the result was a demonstration in Fig. (7a, 7b).

Predicting the localization and number of Transmembrane helices (TMHMM) is a tool used to predict the presence of transmembrane helices in proteins. The results will show the protein components that lie inside, outside or inside the membrane.

In Ghost Shark the segment of the protein that lie outside the membrane is (1-734) (1-684) (1-718) (1-698) (1-729) and the protein was present with the helix is (735-757) (685-707) (719-741) (699-721) (730-752) The protein membrane inside the helix is (708-727) (722-767) (742-761) (753-803) (758-789). The result was shown in Fig. (8a).

In Zebra Finch the segment of the protein that lie outside the membrane (1-763) (1-702) (1-709) (1-679) (1-619) and the protein was present with the helix is (764-766) (703-725) (710-732) (680-702) (620-642) and inside the helix is (767-804) (726-763) (733-768) (703-738) (643 -688). The result was shown in Fig.(8b).

In Mouse the segment of the protein that lie outside the membrane is (1-739) (1-199) (1-419) (1-527) (1-739) and the protein was present with the helix is (740-762) (200-222) (492-514) (528-550) (740-762) The protein membrane inside the helix is (763-805) (223-265) (515- 521) (551-557) (763-805). The result was shown in Fig. (8c).

In Polar Bear the segment of the protein that lie outside the membrane is (1-718) (1-708) (1-714) (1-710) (1-724) and the protein was present with the helix is (719-741) (709-731) (714-737) (711-733) (725-747) The protein membrane inside the helix is (742-784) (732-774) (738- 780) (734-776) (748-790). The result was shown in Fig.(8d).

In Human the segment of the protein that lie outside the membrane is (1-757) (1-740) (1-394) (1-740) and the protein was present with the helix is (758-780) (740-763) (395-417) (741-763) The protein membrane inside the helix is (781-805) (764-805) (418- 459) (764-805). The result was shown in Fig. (8e).

The sequence and structure of the polypeptide chains that need to be classified by comparative analysis that make up the protein conserved domain. It is through molecular evolution that the building blocks, different shapes can undergo organization and production, make protein production of different functions. Significance of the evolution domain note factors brought to us by the 2019-ncov major protease epitope protection degree assessment. This was done using the conservative domain database (CDD).

In Ghost Shark conservation domain databases shows the peptidaseM2 accession no Pfam01401 shows the interval value 22-610 and E-value 0e+00. The expect value was a true positives. But the collectrin was contain a accession no pfam16959 and internal value was 620-750 and e-value was 1.01e-60. The result was shown in Fig.(9a).

In Zebra Finch the databases shows the peptidaseM2 accession no Pfam01401 shows the interval value621-772 and E-value 0e+00. The expect value was a true positive. But the collectrin was contain a accession no pfam16959 and internal value was 621-772 and e-value was 6.78e-96. The result was shown in Fig. (9b).

In mouse CDD show the Gluzin super family accession no cl14813 shows the interval value 1-358 and E-value 0e+00. The expect value was a true positives. But the collectrin was contain a accession no pfam16959 and internal value was 369-521 and e-value was 3.19e-88. The result was shown in Fig. (9c).

In Polar Bear conservation domain databases shows the peptidaseM2 accession no Pfam01401 shows the interval value 8-557 and E-value 0e+00. The expect value was a true positive. But the collectrin was contain a accession no pfam16959 and internal value was 558-741 and e-value was 3.12e-102. The result was shown in Fig. (9d).

In Human the databases shows the peptidase M2 accession no Pfam01401 shows the interval value 19- 554and E-value 0e+00. The result was shown in Fig. (9e).

In Ghost Shark GMWDGVPKV sequence contain highest IC50 VALUE (8.083). FJPLYEWLK sequence contain lowest IC50 value (6.771). The IC50 value lower than 10 shows the strong binding among the epitope. The result was shown in Fig.(10a).

In Zebra Finch LAAPYEPPV sequence contain a IC50 VALUE (8.09) is higher. But WLKRNNSGR sequence contain IC50 value (6.091)is lower.The result was shown in Fig.(10b).

In Mouse KLMDTYPSY sequence contain highest IC50 VALUE (8.047). RKLMDTYPS sequence contain lowest IC50 value (6.027). The result was shown in Fig.(10c).

In Polar Bear ALYEHLHAY sequence contain higher IC50 VALUE (8.091). YWRGDYEEE sequence contain lower IC50 value (6.027). The result was shown in Fig.(10d).

In Human KLMNAYPSY sequence contain high IC50 VALUE (8.326). MAYAAQPFL sequence contain low IC50 value (7.049). The result was shown in Fig.(10e).

In Ghost Shark 0.98 higher score. They show the peptide higher probability to as the epitope. Threshold values are 0.51. Some sequences are shown in overlay display." MSKIYSTGTVCKPNNP, YKSSLASWEYNTNITD, SWEYNTNITDENIDKM, AFYQQASDDSSKFNIN, SPVEQEATAFLKEFDT, TGTVCKPNNPSDCLGL, DHLNEVQNEMSKIYST, DENIDKMNEESAKWSA, KGSGVLSKEEQDHLNE, TILLAES, NPSDCLGLEPGLTILL, FDTKSQDLVYKSSLAS, KFNINEISDNIIKLQL, ISDNIIKLQLNSLQDK, MFLQWLLLLSLAAAAL LNSLQDKGSGVLSKEE". The result was shown in Fig. (11a).

In zebra Finch show 0.95 is a higher score. They show the peptide higher probability to as the epitope. Threshold values are 0.51. Some sequences are shown in overlay display "MSTIYSTGTVCKINNP, DARWSAFYEEASRNAS, LDAIMSGST, GSSVLSPEKYNRLGTV, NISYENSIASWNYNTN, TVCKINNPSECLVLEP, SIASWNYNTNITEENA, QQAQIFLEEFNRRAEN, ASRNASTFQVDSIADD, VDSIADDPTKLQIQIL, LVCFWLLCGLSAVVTP, NPSECLVLEPGLDAIM, LSAVVTPQDVTQQAQI, NITEENANKMSEADAR, KLQIQILQERGSSVLS". The result was shown in Fig. (11b).

In Mouse show 0.94s a higher score. They show the peptide higher probability to as the epitope. Threshold values are 0.51. Some sequences are shown in overlay display "SDVI, SSVAYAMRKYFSIIKN, TEAGQKLLKMLSLGNS, FVGWNTEWSPYADQSI, PWTKALENVVGARNMD, XCDISNSTEAGQKLLK, KVRISLKSALGANAYE, PLLNYFQPLFDWLKEQ, VSDLKPRVSFYFFVTS, EWSPYADQSIKVRISL, AYEWTNNEMFLFRSSV, VGARNMDVKPLLNYFQ, WLKEQNRNSFVGWNTE, RVSFYFFVTSPQNVSD". The result was shown in Fig.(11c).

In Polar Bear show 0.95 is a higher score. They show the peptide higher probability to as the epitope. Threshold values are 0.51. Some sequences are shown in overlay display "MSTIYSTGKACNPNNP, ECLLLEPGLDDIMENS, KQ, SKHAKTYPLEEIHNST, DLYYQSSLASWNYNTN, YPLEEIHNSTVKRQLQ,

NENIQKMNDAGAKWSA, RLWAWEGWRSEVGKQ, KWSAFYEEQSKHAKTY, DDIMENSKDYNERLWA, GSSVLSADKSQRLNTI, SLASWNYNTNITNENI, TGKACNPNNPQECLLL, LQALQHSGSSVLSADK, QRLNTILNAMSTIYST, TFLEKFNYEAEDLYYQ". The result was shown in Fig. (11d).

In Human show 0.96 is a higher score. They show the peptide higher probability to as the epitope. Threshold values are 0.51. Some sequences are shown in overlay display "MSTIYSTGKACNPNNP, ECLLLEPGLDDIMENS, KQ, SKHAKTYPLEEIHNST, DLYYQSSLASWNYNTN, YPLEEIHNSTVKRQLQ, NENIQKMNDAGAKWSA, RLWAWEGWRSEVGKQ, KWSAFYEEQSKHAKTY, DDIMENSKDYNERLWA, GSSVLSADKSQRLNTI, SLASWNYNTNITNENI, TGKACNPNNPQECLL, LQALQHSGSSVLSADK, QRLNTILNAMSTIYST, FLEKFNYEAEDLYYQ". The result was shown in Fig. (11e).

The IDEB program evaluates segments of the protein sequence that are likely to be antigenic by obtaining an antibody response. The result was shown in graph. The graph has x and y axes. The X and Y axes indicate the sequence position and the antigenic score, the threshold value being 1.01. The area above the entry value has a yellow antigenic appearance. And the score selected to estimate the Redline. An epitope is drawn at the residual position plot at the score threshold value. The locations of the expected epitope residues are yellow. The estimated peptide table below the plot lists all the estimated lines of epitopes and their positions in the protein The result was shown in Fig. (13 a, b,c,d,e).

SARS-COV-2, The causative agent of respiratory distress syndrome, has infected more than 10,000 people worldwide, leading to numerous deaths. First detected in Wuhan, Hubei Province, China, COVID-19 spread uncontrollably, eventually becoming a global threat. Scientists around the world are working to find a solution to this evil outbreak (manojit bhattacharya et al., 2019). Based on the World Health Organization (WHO) website (https://who.sprinklr.com/) 23 July, 2021, there have been 192,284,207 confirmed cases of covid-19, including 4,136,518 deaths, reported to WHO. As of 24 july 2021, a total of 3,605,386,928 vaccine doses have been administered (WHO).

Epitope-based vaccines provide a new strategy for the prophylactic and therapeutic use of germline-specific immunity. A multi-epitope vaccine containing peptides or overlapping peptides appears to be an appropriate solution for the prevention and treatment of viral infections (Zhang et al.,2018). Appropriate multi-epitope vaccine engineering should be done to include cytotoxic T lymphocytes (CTLs), T-cells, and B-cells, which contain active epitopes that can activate and induce successful responses to specific viruses.

Today, researchers are exploring ways to develop subunit vaccines from a complete gene / protein of pathogens. Epitope assessment for antibodies has become more important with the development of computational tools for vaccine design (Dubey et al., 2018). There is a subdivision in the field of bioinformatics, which includes many tools and databases. Immunological dataset prediction and in silico analysis is done with the help of tools. Advances in tools and the availability of a variety of data, such as genetic, proteomic, and various algorithms, have made it more effective for scientists to accurately estimate the epitopes that are most effective in the development of subunits of vaccines (De Gregorio De et al., 2012).

The SARS-CoV-2 gene sequence distinguishes with an accession number used as a source. Sequence of proteins from this genome was used in this study. Protein sequence was found on the NCBI Protein website. Each protein sequence was re-searched on the Identical Protein Sequence (IPG) Database and all sequence of each protein was matched to Clustal Omega. Interestingly, all sequences were the same which means that the sequence of all the proteins was maintained. Since this is a novel virus that is incorporated in sequence is low in value, which could be the reason for the same protein sequence. An IPG website is more useful than other protein sequence information sites as a single sequence can be used as a representative of the same sequence group. Thus, unnecessary duplication of sequence data can be avoided. List of protein sequences used to predict epitope. Genetic variation is one of the main reasons for the inhibition of the immune system by pathogens. Protein sequence means that epitopes remain unchanged and as a result the immune system continues to detect them.

Two tables are produced by ProtParam, the first showing computer values based on the assumption that all cysteine residues appear as cystines, and the second assumes that no cysteine appears as cystine.IN VIVO HALF-LIFE Part of life is the prediction of the time it takes for part of the amount of protein in a cell to disappear after its incorporation into a cell. Predictability is given to three organisms (Human, Yeast, and E. coli), but it is possible that the effect is transmitted to the same organisms. ProtParam measures half-life by looking at N-terminal amino acid sequences under investigation. INDICATOR OF SUSTAINABILITY (II) Protein in its index of instability less than 40 is predicted to be stable; a value of more than 40 predicts that protein may be unstable. ALIPHATIC INDICATOR The aliphatic protein index is defined as the average volume taken by separate aliphatic chains (alanine, valine,

isoleucine, and leucine). It may be considered a positive factor in the increase in thermostability of globular proteins(Elisabeth Gasteiger et al., 2003).

The predictable results of transmembrane helices are TMHMM and TMPred. (A) Predicting effect of transmembrane helices with TMHMM Server ver. 2.0. Red bars indicate transmembrane domains, blue lines indicate intracellular loops and magenta lines indicating extracellular loops. (B) PgLHT Hydrophobicity plot obtained using TMPred. Ordinate represents TMPred points divided by 1,000. Positive scores elevate the hydrophobic region. Expandable membrane regions are marked with arrows.

Conserved Domain BLAST: - The CDD 27036 PSSMs website used to search for saved regions using the E-value parameter of 0.01 and keep the key OPEN with a 'sophisticated filter' that removes all of those edits that did not show evolutionary relationships. Result with e-value in the range of 1 above it should be considered putative false positives. Expect value is a parameter that describes the number of hits one can expect to see by chance when searching a database of a particular size. The E value setting can be modified to adjust the statistical significance threshold used for reporting matched against in the database. False positive results should be very rate with the default setting of 0.01(Marchler et al., 2004).

Prediction Of T-Cell Epitopes with HLA using the MHC prediction server. Program results are shown in a three-column table. The first column shows the peptide sequence, the second and third columns show the IC50 and IC50 values inverted by the IC, respectively. If the IC50 value is above 5000, the peptide will not bind to MHC atoms. Arrays of peptides are sorted with IC50 values. Peptides with lower IC50 values (or higher IC50 values than IC50 values) are listed first and non-binders are given at the bottom of the Table. (http://www.ddgpharmfac.net/mhcpred/MHCPred/)

The future development of MHC Pred will improve both the scope and use of the server and the sub-system. First, we expect to significantly increase the number of allele models, with increased focus on both human Class II and HLA-B and HLA-C loci, as well as non-human allele, particularly murine, i -bovine and primate. Although the peptide binding data binding to class I alleles of length outside 9 is limited, we will also seek to produce peptide binding models of lengths 8, 10 and 11. We are also looking at technological advances aimed at the automatic excavation of epitope genomes. Similarly, by combining a user-defined set of allele models, we will be able to address the problem of identifying contaminated peptides that can bind several different MHC alleles. Second, the additive method used in MHC Pred, itself, relies on the presence of certain amino acids in specific areas within the set of training peptides to reliably predict the effect of that residue on that condition in any experimental peptides.

MHC Class I predictor episodes: The MHC Pred server predicts peptides binding to different alleles of Class I MHC. Selection of binding peptides is made on the basis of binding points. Peptides selected as epitopes have a threshold score of 0.8 or more. Such peptides showed a predictable binding correlation of less than 10nM. The server uses a default threshold score of 0.5 however, a higher score value has been used to detect epitopes with high binding relationships of MHC alleles. Predicted episodes and their schools. Higher numbers of epitopes were predicted by protein sequence(Van Regenmortel et al., 1993).

MHCPred is one of the most effective among the available binding methods for class I MHC.Therefore, this method was chosen to make predictions. There have been highly contaminated epitopes predicted in the study. The immoral nature of epitopes is a desirable asset as one epitope can bind to different alleles. Only spike proteins have shown the presence of four contaminated epitopes. All peptides showing high binding may not be able to activate Tc cells. In order for the peptide to act as a Tc cell epitope it must be approved for proteasome processing, indicating the correlation to be mediated by TAP in addition to the Class I MHC affinity. The server predicts CTL epitopes among a group of binding peptides in Class I MHC.

The combined score greater than 1 was used as the selective condition for CTL epitopes (although the default limit value is 0.75). As MHC Pred is trained in human data it should therefore provide better functioning of the human proteasome and TAP (Larsen et al., 2004). Therefore, peptides complementing both conditions (MHCPred score of 0.8 or more and CTL scores over 1) were selected as epitopes. These epitopes have a high binding affinity for MHC molecules and are capable of processing the cytosolic pathway, transported by TAP with high potential to act as epitopes. We analyzed five different SARS-COV-2 proteins in the current study (due to their availability on the NCBI-GenBank website and their role in the structural role in SARS-COV-2 and finally revealed T-Cell epitopes that could be used for wet laboratory observations In the most recent study, different episodes of SARS-COV-2 were discovered, based on In-silico methods and focus, but in our study there are many differences as we analyze a group of proteins from SARS-COV-2 to classify Ts -Cell epitopes with short lengths straight to MHC I and MHC II.

Prediction of B-Cell Epitopes using ABC prediction server. Server for estimating linear B cell epitope regions in antigen sequence using artificial neural network. This server assists in selecting synthetic vaccine candidates, identifying epitope areas that can be used in diagnosis and allergy research. High score of the peptide means higher probability to as epitope. All the peptide shown here threshold value chosen. (http://crdd.osdd.net/raghava/abcpred/).The threshold applied to both servers was 0.5 and the peptides had high score they were considered epitopes. The common epitopes predicted by both servers were used. ACE2 prediction could be used as ABC Pred server can process predictions for less than 6000 protein residues of amino acids. B epitopes of B cells fall into two categories - linear and continuous. For the lack of all SARS-CoV-2 protein components, in our study we have predicted only for episodes. The B cell epitopes through two servers that strengthen the chances of detection by immune system.

Prognosis of B-cell epitopes in antigen sequence is an important and complex problem. Although, most antigenic protein selections do not persist, it is possible to mimic epitopes with synthetic peptides (Van Regenmortel et al., 1993;1994). One of the major problems faced in developing B-cell epitope predictors is the variable epitope length. Performance is much better than random, despite the fact that B-cell epitope prediction is a complex problem. It is therefore advisable to use an ABC pred server to detect B-cell epitopes in the antigen.

The immune epitope site (IEDB) (Peters et al., 2005) is probably the most complete site for B-cell and T-cell epitopes tested. The IEDB provides users with access to epitope-related analyzes and predictive tools including: (i) a few ways to predict accurate and consistent B-cell epitopes; (ii) a visual tool for predicted conformational epitopes in a 3D antigen structure; (iii) several epitope data analysis tools (e.g., computerized epitope preservation and epitope population inclusion). The IEDB allows users to obtain both internal biochemical information and external epitope-based information (Peters et al., 2005). This makes it possible to easily integrate customized data sets (e.g., a set of security data. In addition, several researchers have used the IEDB to perform meta-pathogens analysis of interest, thereby improving the use of the IEDB in the analysis and prediction of B-cell epitopes.

Chou and Fasman's method is based on arithmetic the potential for the expanded remains to form part of the second curve structure of B turn. First of the two experimental defensive epitopes were predicted with seven to six lists therethe window size was seven and nine, respectively. Kolaskar and Tongaonkar calculated the frequency of occurrences of each type of amino acid (fAg) in 156 experimental epithets from 34 different proteins. Then, using the Parker's scale, moderate levels of hydrophilicity, accessibility, and variability were calculated across heptapeptide dispersed at 156 epitopes, and the frequency of amino acids increased. (fs) was calculated. The antigenic propensity (Ap) value of each amino acid is calculated. About 75% of epitopes have been correctly predicted using an antigenic propensity scale for epitopes for which the scale has been developed. In HA1, the route predicted epitope 91–108, while the second protective episode was not predicted.

Reverse vaccinology plays an important role in the development of recombinant vaccines by allowing the in silico analysis of the viral genome. In-silico analysis it helps to identify the most antigenic and hidden proteins that are important for vaccine development before the start of a wet laboratory study (Dangi et al., 2018). Using this approach, the current study aims to identify potentially antigenic proteins and epitope regions targeted by both the B and T cell flexible immune systems to improve vaccine or sero diagnostic testing as described.

The recent global pandemic of SARS-CoV-2 has claimed hundreds of precious lives in various parts of the world and weakened the economies of many countries. A fully effective vaccine against the approved drug or SARS-CoV-2 has not yet been reported. In this study, there was a successful attempt to make ACE2 against SARS-COV-2. The current study describes ACE2 as a potential candidate for vaccine production. However, current research is the result of an integrated vaccine omics approach. Therefore, more experimental research by future is needed to demonstrate the efficacy of the developed vaccine.

Antibodies is currently the most promising class of biopharmaceuticals. The primary purpose of epitope detection is to replace antigen in vaccines, immune production, and serodiagnosis. Accurate identification of B-cell epitopes and large-scale data collection remains major challenges for immunogenicity. The development of epitope mapping of B cells and computer prediction has revealed molecular details in the biological process and the formation of the Ag-Ab complex, which can help to make more accurate algorithms to predict localization in antigen. However, based on statistics it is not possible to accurately determine the characteristics of the epitope, which allows biological recognition. One should keep in mind that epitopes are not an internal protein component and antibodies that are ignored by antibodies predict epitope placement where an unspecified Ab may bind. True epitopics cannot be predicted regardless of the impact of the structure on the complex structure of Ag-Ab.

REFERENCE:

1) A.J. Chirino, M.L. Ary, S.A. Marshall. Minimizing the immunogenicity of protein therapeutics. Drug Discov Today 2004; 9:82-90.

2) Ahmad B, Ashfaq UA, Rahman MU, Masoud MS, Yousaf MZ. Conserved B and T cell epitopes prediction of ebola virus glycoprotein for vaccine development: an immuno-informatics approach. Microb Pathog. 2019; 132:243–53.

3) Ahmed SF, Quadeer AA, McKay MR. Preliminary identification of potential vaccine targets for the COVID-19 coronavirus (SARS-CoV-2) based on SARS-CoV immunological studies. Viruses. 2020;12(3):254.

4) Alix,A.J. (1999) Predictive estimation of protein linear epitopes by using the program PEOPLE. Vaccine, 18, 311–314.

5) Amanat, F., & Krammer, F. (2020). SARS-CoV-2 vaccines: Status Report. *Immunity, 52,* 583–589.

6) Andersen, K. G., Rambaut, A., Lipkin, W. I., Holmes, E. C., & Garry, R. F. (2020). The proximal origin of SARS-Cov-2. Nature Medicine, 26(4), 450-452. doi:10.1038/s41591-020-0820-9

7) Bhattacharya M, Sharma AR, Patra P, Ghosh P, Sharma G, Patra BC, et al. Development of epitope-based peptide vaccine against novel coronavirus 2019 (SARS-COV-2): Immunoinformatics approach. J Med Virol. 2020;92(6):618–631.

8) Blythe M, Flower D: Benchmarking B cell epitope prediction: underperformance of existing methods. Protein Science: A Publication of the Protein Society. 2005, 14: 246-254

9) Boni, M. F., Et Al. The Evolutionary Origin Of Sars-cov-2, The Sarbecovirus Lineage Behind The COVID-19 Pandemic. Microbiol. Https://Doi.Org/10.1038/S41564-020-0771-4 (2020)).

10) Carver, C. and N. Jones, Comparative accuracy of oropharyngeal and nasopharyngeal swabs for diagnosis of COVID-19. 2020, Oxford COVID-19 evidence service team centre for evidence based medicine

11) Castelletti D, Fracasso G, Righetti S, Tridente G, Schnell R, Engert A, Colombatti M. A dominant linear B-cell epitope of ricin A-chain is the target of a neutralizing antibody response in Hodgkin's lymphoma patients treated with an anti-CD25 immunotoxin. Clin Exp Immunol 2004;136:365–372.

12) Chakraborty C, Sharma AR, Sharma G, Bhattacharya M, Lee SS. SARS-CoV-2 causing pneumonia-associated respiratory disorder (COVID-19): diagnostic and proposed therapeutic options. Eur Rev Med Pharmacol Sci. 2020;24(7):4016–26.

13) Chakraborty, C., Sharma, A. R., Bhattacharya, M., Sharma, G., Lee, S.-S., & Agoramoorthy, G. (2020). Consider TLR5 for new therapeutic development against COVID-19. *Journal of Medical Virology, 92,* 2314–2315.

14) Chakraborty, C., Sharma, A. R., Sharma, G., Bhattacharya, M., Saha, R. P., & Lee, S.-S. (2020). Extensive partnership, collaboration, and teamwork is required to stop the COVID-19 outbreak. *Archives of Medical Research, 51,* 728–730.

15) Chen, H., Tang, L., Yu, X., Zhou, J., Chang, Y., & Wu, X. (2020). Bioinformatics analysis of epitope-based vaccine design against the novel SARS-CoV-2. *Infectious Diseases of Poverty, 9*(1), 1–10.

16) Clement G, Boquet D, Frobert Y, Bernard H, Negroni L, Chatel JM, Adel-Patient K, Creminon C, Wal JM, Grassi J. Epitopic characterization of native bovine b-lactoglobulin. J Immunol Methods 2002;266:67–78.

17) Coleman CM, Frieman MB. Coronaviruses: important emerging human pathogens. J Virol. (2014) 88:5209–12. doi: 10.1128/JVI.03488-13

18) D. C. Benjamin, "B-cell epitopes: fact and fiction," Advances in Experimental Medicine and Biology, vol. 386, pp. 95–108, 1995.

19) D. Lopresti, "Introduction to Bioinformatics", Biol. S 95, pp. 1-43, 2008.

20) D. R. Davies and G. H. Cohen, "Interactions of protein antigens with antibodies," Proceedings of the National Academy of Sciences of the United States of America, vol. 93, no. 1, pp. 7– 12, 1996.

21) D. R. Madden, "The three-dimensional structure of peptide-MHC complexes," Annual Review of Immunology, vol. 13, no. 1, pp. 587–622, 1995.

22) D. R. Madden, D. N. Garboczi, and D. C. Wiley, "The antigenic identity of peptide-MHC complexes: a comparison of the conformations of five viral peptides presented by HLA-A2," Cell, vol. 75, no. 4, pp. 693–708, 1993.
23) D.J. Barlow, M.S. Edwards, J.M. Thornton. Continuous and discontinuous protein antigenic determinants. Nature 1986; 322:747-748.
24) D.R. Flower. Immunoinformatics. Predicting immunogenicity in silico. Preface. Methods Mol Biol 2007; 409:v-vi.
25) Dangi, M., Kumari, R., Singh B., & Chhillar, A. K., Advanced In Silicon Tools For The Design Of The Antigenic Epitope, As A Potential Candidate Vaccines Against The Coronavirus. In Bioinformatics: Sequence, Structure, Phylogeny (Ed. Shanker A.) 329-357 (Springer, Singapore, 2018)).

26) del Rio, C. and P.N. Malani, COVID-19—new insights on a rapidly changing epidemic. Jama, 2020.

27) DeLano WL. Pymol: An open-source molecular graphics tool. CCP4 Newsl Protein Crystallogr. 2002;40:82–92.

28) Doytchinova,I.A., Blythe,M.J. and Flower,D.R. (2002) An additive method for the prediction of protein–peptide binding affinity. Application to the MHC class I molecule HLA-A*0201. J. Proteome Res., 1, 263–272

29) Dubey KK, Luke GA, Knox C, Kumar P, Pletschke BI, Singh PK, et al. Vaccine and antibody production in plants: developments and computational tools. Brief Funct Genomics. 2018;17(5):295–307. pmid:29982427

30) E. M. Lafuente and P. A. Reche, "Prediction of MHC-peptide binding: a systematic and comprehensive overview," Current Pharmaceutical Design, vol. 15, no. 28, pp. 3209–3220, 2009.

31) E.E. Hughes, H.E. Gilleland, Jr. Ability of synthetic peptides representing epitopes of outer membrane protein F of Pseudomonas aeruginosa to afford protection against P. aeruginosa infection in a murine acute pneumonia model. Vaccine 1995; 13:1750-1753.

32) E.H. Nardin, J.M. Calvo-Calle, G.A. Oliveira, R.S. Nussenzweig, M. Schneider, J.M. Tiercy, L. Loutan, D. Hochstrasser, K. Rose. A totally synthetic polyoxime malaria vaccine containing Plasmodium falciparum B cell and universal T cell epitopes elicits immune responses in volunteers of diverse HLA types. J Immunol 2001; 166:481-489.

33) EL-Manzalawy Y, Dobbs D, Honavar V: Predicting flexible length linear B-cell epitopes. 7th International Conference on Computational Systems Bioinformatics. 2008, 121-131.
34) Eschner K. We're still not sure where the COVID-19 really came from. Popular Science; 2020. pmid:31919200
35) Estienne V Kulkarni-Kale U, Bhosle S, Kolaskar AS. CEP: a conformational epitope prediction server. Nucleic Acids Res 2005;33:W168– W171. Web server issue.
36) Estienne V, Duthoit C, Blanchin S, Montserret R, Durand-Gorde JM, Chartier M, Baty D, Carayon P, Ruf J. Analysis of a conformational B cell epitope of human thyroid peroxidase: identifica1tion of a tyrosine residue at a strategic location for immunodo1minance. Int Immunol 2002;14:359–366.
37) Fan C, Li K, Ding Y, et al. ACE2 Expression in Kidney and Testis May Cause Kidney and Testis Damage After 2019-nCoV Infection. medRxiv 2020.
38) Feng Y, Jacobs F, Van Craeyveld E, Lievens J, Snoeys J, Van Linthout S, De Geest B: The impact of antigen expression in antigen-presenting cells on humoral immune responses against the transgene product. Gene therapy. 2009, 17: 288-293. 10.1038/gt.2009.125.
39) Flower D: . Bioinformatics for vaccinology. 2009, Wiley Black-well
40) Flower D: . Immunoinformatics: predicting immunogenicity in silico. 2007, Quantum distributor, 1
41) Flower DR. Towards in silico prediction of immunogenic epitopes. Trends Immunol 2003;24:667
42) Flower,D.R., Doytchinova,I.A., Paine,K., Taylor,P., Blythe,M.J., Lamponi,D., Zygouri,C., Guan,P., McSparron,H. and Kirkbride,H. (2002) Computational vaccine design. In Flower,D.R. (ed.), Drug Design: Cutting Edge Approaches. RSC Publications, London, pp. 136–180.
43) G.A. Schellekens, H. Visser, B.A. de Jong, F.H. van den Hoogen, J.M. Hazes, F.C. Breedveld, W.J. van Venrooij. The diagnostic properties of rheumatoid arthritis antibodies recognizing a cyclic citrullinated peptide. Arthritis Rheum 2000; 43:155-163.
44) Gasteiger E., Et Al. John M. Walker, Protein Identification And Analysis Tools, Expasy. Proteomics. Protoc. Hanb. Https://Doi.Org/10.1385/1592598900 (2005)).

45) Gorbalenya, A.E., Severe acute respiratory syndrome-related coronavirus–The species and its viruses, a statement of the Coronavirus Study Group. BioRxiv, 2020.

46) Greenbaum J, Sidney J, Chung J, Brander C, Peters B, Sette A. Functional classification of class II human leukocyte antigen (HLA) molecules reveals seven different supertypes and a surprising degree of repertoire sharing across supertypes. Immunogenetics. 2011;63(6):325–35.

47) Grifoni A, Sidney J, Zhang Y, Scheuermann RH, Peters B, Sette A. A sequence homology and bioinformatic approach can predict candidate targets for immune responses to SARS-CoV-2. Cell Host Microbe. 2020;27(4):671–680.e2.

48) Guan, W., Ni, Z., Hu, Y., Liang, W., Ou, C., He, J., … Zhong, N. (2020). Clinical characteristics of coronavirus disease 2019 in China. New England Journal of Medicine, 382(18), 1708-1720. doi:10.1056/nejmoa2002032

49) Guo Y.-R., Q.-D. Cao, et al., The origin, transmission and clinical therapies on coronavirus disease 2019 (COVID-19) outbreak – an update on the status, Mil. Med. Res. 7 (2020) 11.

50) H. Shirai, C. Prades, R. Vita, P. Marcatili, B. Popovic, J. Xu, J.P. Overington, K. Hirayama, S. Soga, K. Tsunoyama, D. Clark, M.P. Lefranc, K. Ikeda. Antibody informatics for drug discovery. Biochim Biophys Acta 2014; 1844:2002-2015.

51) Haste Andersen P, Nielsen M, Lund O: Prediction of residues in discontinuous B-cell epitopes using protein 3D structures. Protein Sci. 2006, 15: 2558-10.1110/ps.062405906.

52) Huang, C., et al., Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. The Lancet, 2020. 395(10223): p. 497-506.

53) Iwasaki A. And Yang Yu. The Potential Risk Of Their Study Responding To Antibodies During COVID-19. Rheumatol. Https://Doi.Org/10.1038/S41577-020-0321-6 (2020)).

54) J. A. Greenbaum ,D. J. Barlow, M. S. Edwards, and J. M. Thornton, "Continuous and discontinuous protein antigenic determinants," Nature, vol. 322, no. 6081, pp. 747–748, 1986.

55) J. Xiang, S. Zhang, A. Cheng et al., "Expression and characterization of recombinant VP19c protein and N-terminal from duck enteritis virus," Virology Journal, vol. 8, article 82, 2011.

56) J.L. Pellequer, E. Westhof, M.H. Van Regenmortel. Predicting location of continuous epitopes in proteins from their primary structures. Methods Enzymol 1991; 203:176-201.

57) J.L. Sanchez-Trincado, M. Gomez-Perosanz, P.A. Reche. Fundamentals and Methods for T- and B-Cell Epitope Prediction. J Immunol Res 2017; 2017:2680160.

58) J.M.S. Rana & K. S. Vaisla- Introduction to Bioinformatics-

59) J.P. Tam, Y.A. Lu. Vaccine engineering: enhancement of immunogenicity of synthetic peptide vaccines related to hepatitis in chemically defined models consisting of T- and B-cell epitopes. Proc Natl Acad Sci U S A 1989; 86:9084-9088.

60) Kobayashi H, Wood M, Song Y, Appella, E, Celis E, (2000), Promiscuous Helper T-cell Epitopes Of The Class II MHC Tumor Antigen HER2/Neu. Cancer Res. 60:5228-5236). .

61) Kulkarni-Kale U, Bhosle S, Kolaskar A: CEP: a conformational epitope prediction server. Nucleic Acids Res. 2005, 33: W168-10.1093/nar/gki460.

62) Kuo,, Wang C., Badakhshan T., Chilukuri S. & BenMohamed L. The challenges and opportunities for the development of a T-cell epitope-based herpes simplex vaccine. Vaccine 32, 6733–6745, doi:10.1016/j.vaccine.2014.10.002 (2014).

63) L. J. Stern and D. C. Wiley, "Antigenic peptide binding by class I and class II histocompatibility proteins," Structure, vol. 2, no. 4, pp. 245–251, 1994.

64) L. Malherbe, "T-cell epitope mapping," Annals of Allergy, Asthma & Immunology, vol. 103, no. 1, pp. 76–79, 2009.

65) Langeveld J, martinez Torrecuadrada J, boshuizen R, Meloen R, Ignacio C: Characterisation of a protective linear B cell epitope against feline parvoviruses. Vaccine. 2001, 19: 2352-2360. 10.1016/S0264-410X(00)00526-0.

66) Langeveld JP, Martinez-Torrecuadrada J, Boshuizen RS, Meloen RH, Ignacio CJ. Characterisation of a protective linear B cell epitope against feline parvoviruses. Vaccine 2001;19:2352– 2360.

67) Lanza, L. Perillo, C. Landi, F. Femiano, F. Gombos, and N. Cirillo, "Controversial role of antibodies against linear epitopes of desmoglein 3 in pemphigus vulgaris, as revealed by semiquantitative living cell immunofluorescence microscopy and in-cell ELISA," International Journal of Immunopathology and Pharmacology, vol. 23, no. 4, pp. 1047–1055, 2010.

68) Larsen J, Lund O, Nielsen M: Improved method for predicting linear B-cell epitopes. Immunome Res. 2006, 2: 2-10.1186/1745-7580-2-2.

69) Lu I. N., Farinelle S., Sausy A. & Muller C. P. Identification of a CD4 T-cell epitope in the hemagglutinin stalk domain of pandemic H1N1 influenza virus and its antigen-driven TCR usage signature in BALB/c mice. Cellular & Molecular Immunology 14, 511–520, doi:10.1038/cmi.2016.20 (2017).

70) Lu, H. (2020). Drug treatment options for the 2019-new coronavirus (2019- nCoV). BioScience Trends, 14(1), 69-71. https://doi.org/10.5582/bst.2020.01020

71) M. H. V. Van Regenmortel, "Immunoinformatics may lead to a reappraisal of the nature of B cell epitopes and of the feasibility of synthetic peptide vaccines," Journal of Molecular Recognition, vol. 19, no. 3, pp. 183–187, 2006.

72) M. Yadav, E. Liebau, C. Haldar, and S. Rathaur, "Identification of major antigenic peptide of filarial glutathione-Stransferase," Vaccine, vol. 29, pp. 1297–1303, 2011.

73) Madeira, F. Et Al. Te Embl-ebi Search And Sequence Analysis Tools Apis In 2019. Nucl. Acids Res. 47(1), 636–641. Https://Doi. Org/10.1093/Nar/Gkz268 (2019)).

74) Marchler, bauer A Et Al (2017) CDD/SPARCLE: Functional Classification Of Proteins Via Subfamily Domain Architectures. Nucleic Acids Res. 45(d):200–203)

75) Maxwell, C., McGeer, A., Tai, K. F., & Sermer, M. (2017). No. 225-Management guidelines for obstetric patients and neonates born to mothers with suspected or probable severe acute respiratory syndrome (SARS). Journal of Obstetrics and Gynaecology Canada, 39(8), e130-e137. doi:10.1016/j.jogc.2017.04.024

76) Meunier M, Guyard - Nicodème M, Hirchaud E, Parra A, Chemaly M, Dory Dseparation Of Members Of The Novel Vaccine Against Campylobacter Using Reverse Vaccinology. Immunology Research Journal. 2016; 2016: 5715790. Part: 10.1155 / 2016/5715790

77) Moreno-Fierros, L., García-Silva, I., & Rosales-Mendoza, S. (2020). Development of SARS-CoV-2 vaccines: should we focus on mucosal immunity? *Expert Opinion on Biological Therapy, 20*(8), 831–836.

78) N. L. Dudek, P. Perlmutter, M. I. Aguilar, N. P. Croft, and A. W. Purcell, "Epitope discovery and their use in peptide based vaccines," *Current Pharmaceutical Design*, vol. 16, no. 28, pp. 3149–3157, 2010.

79) N. R. Hegde, S. Gauthami, H. M. Sampath Kumar, and J. Bayry, "The use of databases, data mining and immunoinformatics in vaccinology: where are we?" *Expert Opinion on Drug Discovery*, vol. 13, pp. 117–130, 2017.

80) N. Tomar and R. K. De, "Immunoinformatics: a brief review," *Methods in Molecular Biology*, vol. 1184, pp. 23–55, 2014.

81) Nascimento, I. P. & Leite, L. C. Recombinant vaccines and the development of new vaccine strategies. *Braz. J. Med. Biol. Res.* **45**(12), 1102–1111. https://doi.org/10.1590/s0100-879x2012007500142 (2012).

82) Negroni L 1998 Selo I , Clement G , Bernard H, Chatel J, Creminon C, Peltre G, Wal J. Allergy to bovine b-lactoglobulin: specificity of human IgE to tryptic peptides. Clin Exp Allergy 1999;29:1055–1063. 664.2005

83) O. S. Andersen, P. Boisguerin, S. Glerup et al., "Identification of a linear epitope in sortilin that partakes in proneurotrophin binding," Journal of Biological Chemistry, vol. 285, no. 16, pp. 12210–12222, 2010.

84) P. E. Jensen, "Recent advances in antigen processing and presentation," Nature Immunology, vol. 8, no. 10, pp. 1041–1048, 2007.

85) Patel S, Rauf A, Khan H, Abu-Izneid T. Renin-angiotensin-aldosterone (RAAS): the ubiquitous system for homeostasis and pathologies. Biomed Pharmacother. 2017;94:317–25.

86) Peters B, Sidney J, Bourne P, Huynh-Hoa B, Buus S, et al: The immune epitope database and analysis resource: From vision to blueprint. PLoS Biol. 2005, 3 (3): e91-10.1371/journal.pbio.0030091.

87) Ponomarenko J, Bourne P: Antibody-protein interactions: benchmark datasets and prediction tools evaluation. BMC Structural Biology. 2007, 7: 64-10.1186/1472-6807-7-64.
88) R. K. Ahmed and M. J. Maeurer, "T-cell epitope mapping," Methods in Molecular Biology, vol. 524, pp. 427–438, 2009.
89) Raj, V. S., Smits, S. L.,Provacia, L. B., Van den Brand, J. M., Wiersma, L., Ouwendijk, W. J. Haagmans, B. L. (2013). Adenosine Deaminase acts as a natural antagonist for Dipeptidyl peptidase 4-Mediated entry of the Middle East respiratory syndrome coronavirus. *Journal of Virology*, *88*(3), 1834-1838. doi:10.1128/jvi.02935- 13

90) Rakib, A., Sami, S. A., Mimi, N. J., Chowdhury, M. M., Eva, T. A., Nainu, F., et al. (2020). Immunoinformatics-guided design of an epitope-based vaccine against severe acute respiratory syndrome coronavirus 2 spike glycoprotein. *Computers in Biology and Medicine, 124,* 103967.

91) Rothan, H.A. and S.N. Byrareddy, The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. Journal of autoimmunity, 2020: p. 102433.

92) S. Dermime, D. E. Gilham, D. M. Shaw et al., "Vaccine and antibody-directed T cell tumour immunotherapy," *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*, vol. 1704, no. 1, pp. 11–35, 2004.View at: Publisher Site | Google Scholar

93) S. Khalili, A. Jahangiri, H. Borna, K. Ahmadi Zanoos, and J. Amani, "Computational vaccinology and epitope vaccine design by immunoinformatics," *Acta Microbiologica et Immunologica Hungarica*, vol. 61, no. 3, pp. 285–307, 2014.View at: Publisher Site | Google Scholar

94) Saha, A., Sharma, A. R., Bhattacharya, M., Sharma, G., Lee, S.-S., & Chakraborty, C. (2020). Tocilizumab: A therapeutic option for the treatment of cytokine storm syndrome in COVID-19. *Archives of Medical Research, 51,* 595–597.

95) Saha, S.; Raghava, G. P. S., Predicting Continuous-cell Epitopes In An Antigen Using A Recurrent Neural Network. Structure Of Proteins. Function. Bioinform. 2006, 65, 40-48. [Google Scholar] [Crossref]).
96) Santos R, Sampaio W, Alzamora A, Motta-Santos D, Alenina N, Bader M, et al. The ACE2/angiotensin-(1-7)/MAS axis of the renin-angiotensin system: focus on angiotensin-(1-7). Physiol Rev. 2018;98:505–53.
97) Schwartz, D. A., & Graham, A. L. (2020). Potential maternal and infant outcomes from coronavirus 2019-nCoV (SARS-Cov-2) infecting pregnant women: Lessons from SARS, MERS, and other human coronavirus infections. Viruses, 12(2), 194. doi:10.3390/v12020194

98) Seyed sajjad hasheminasab et al., immunoinformatic prediction about potential novel vaccine in surface antigen fragment protein of toxoplasma gondii. international journal of health. 4(1)(2016)1-5
99) Söllner J, Mayer B: Machine learning approaches for prediction of linear B-cell epitopes on proteins. J. Mol. Recognit. 2006, 19: 200-208. 10.1002/jmr.771.
100) T. A. Ahmad, A. E. Eweida, and L. H. El-Sayed, "T-cell epitope mapping for the design of powerful vaccines," Vaccine Reports, vol. 6, pp. 13–22, 2016.
101) Tan, W., Zhao, X., Ma, X., Wang, W., & Niu, P. (2020). A novel coronavirus genome identified in a cluster of pneumonia cases — Wuhan, China 2019−2020. *China CDC Weekly*, *2*(4), 61-62. doi:10.46234/ccdcw2020.017
102) Thornton,J.M., Edwards,M.S., Taylor,W.R. and Barlow,D.J. (1986) Location of 'continuous' antigenic determinants in the protruding regions of proteins. EMBO J., 5, 409–413.

103) Tyrrell, D. A., & Bynoe, M. L. (1965). Cultivation of a novel type of common-cold virus in organ cultures. *BMJ*, *1*(5448), 1467-1470. doi:10.1136/bmj.1.5448.1467
104) U, Schutkowski M: . Epitope mapping protocols. Preface. 2009, Humana Press, 2
105) U. Reineke, M. Schutkowski. Epitope mapping protocols. Preface. Methods Mol Biol 2009; 524:v-vi.
106) Ul Qamar MT, Saleem S, Ashfaq UA, Bari A, Anwar F, Alqahtani S. Epitope-based peptide vaccine design and target site depiction against Middle East respiratory syndrome coronavirus: an immune-informatics study. J Transl Med. 2019;17(1):362.

107) V. Brusic and N. Petrovsky, "Immunoinformatics and its relevance to understanding human immune disease," *Expert Review of Clinical Immunology*, vol. 1, no. 1, pp. 145–157, 2014.

108) Vaduganathan, M., Vardeny, O., Michel, T., McMurray, J. J., Pfeffer, M. A., & Solomon, S. D. (2020). Renin–angiotensin–Aldosterone system inhibitors in patients with COVID-19. New England Journal of Medicine, 382(17), 1653-1659. doi:10.1056/nejmsr2005760

109) Van Gorkom, G., Klein Wolterink, R., Van Elssen, C., Wieten, L., Germeraad, W., & Bos, G. (2018). Influence of vitamin C on lymphocytes: An overview. *Antioxidants, 7*(3), 41. doi:10.3390/antiox7030041

110) Van Regenmortel MH. Pitfalls of reductionism in the design of peptide-cased vaccines. Vaccine 2001;19:2369–2374.

111) Van Regenmortel MH. Synthetic peptides versus natural antigens in immunoassays. Ann Biol Clin (Paris) 1993;51:39– 41.

112) Van Regenmortel MH. The concept and operational definition of protein epitopes. Philos Trans R Soc Lond B Biol Sci 1989;323: 451–466

113) W. Li, M. Joshi, S. Singhania, K. Ramsey, and A. Murthy, "Peptide vaccine: progress and challenges," *Vaccines*, vol. 2, no. 3, pp. 515–536, 2014.View at: Publisher Site | Google Scholar

114) W. Purcell, J. McCluskey, and J. Rossjohn, "More than one reason to rethink the use of peptides in vaccine design," *Nature Reviews Drug Discovery*, vol. 6, no. 5, pp. 404–414, 2007.View at: Publisher Site | Google Scholar

115) Walter G: Production and use of antibodies against synthetic peptides. J. Immunol. Methods. 1986, 88: 149-61. 10.1016/0022-1759(86)90001-3.

116) Wang, Y., Jiang, W., He, Q., Wang, C., Wang, B., Zhou, P., … Tong, Q. (2020). Early, low- dose and short-term application of corticosteroid treatment in patients with severe COVID-19 pneumonia: Single-center experience from Wuhan, China. doi:10.1101/2020.03.06.20032342

117) Weiskopf D, Angelo MA, de Azeredo EL, Sidney J, Greenbaum JA, Fernando AN, et al. Comprehensive analysis of dengue virus-specific responses supports an HLA-linked protective role for CD8+ T cells. Proc Natl Acad Sci U S A. 2013;110(22):E2046–53.

118) Weiss, S. R., & Navas-Martin, S. (2005). Coronavirus pathogenesis and the emerging pathogen severe acute respiratory syndrome coronavirus. *Microbiology and Molecular Biology Reviews*, *69*(4), 635-664. https://doi.org/10.1128/mmbr.69.4.635-

119) Westhof E, Regenmortel MHV. Predicting location of continuous epitopes in proteins from their primary structures. Methods Enzymol 1991;203:176–201

120) World Health Organization (WHO). Statement Regarding Cluster of Pneumonia Cases in Wuhan, China. <www.who.int>. (2020)

121) Y. El-Manzalawy, V. Honavar. Recent advances in B-cell epitope prediction methods. Immunome Res 2010; 6 Suppl 2:S2.

122) Yang X, Yu X. An introduction to epitope prediction methods and software. Rev Med Virol. 2009;19(2):77–96. pmid:19101924

123) Zauner W, Lingnau K, Mattner F, Von Gabain A, Buschle M. Defined synthetic vaccines. Biol Chem 2001;382:581–595.

124) Zhang, J., Zeng, H., Gu, J., Li, H., Zheng, L., & Zou, Q. (2020). Progress and prospects on vaccine development against SARS-CoV-2. *Vaccines, 8,* 153.

125) Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A novel coronavirus from patients with pneumonia in China, 2019. N Engl J Med. 2020;382(8):727–33.