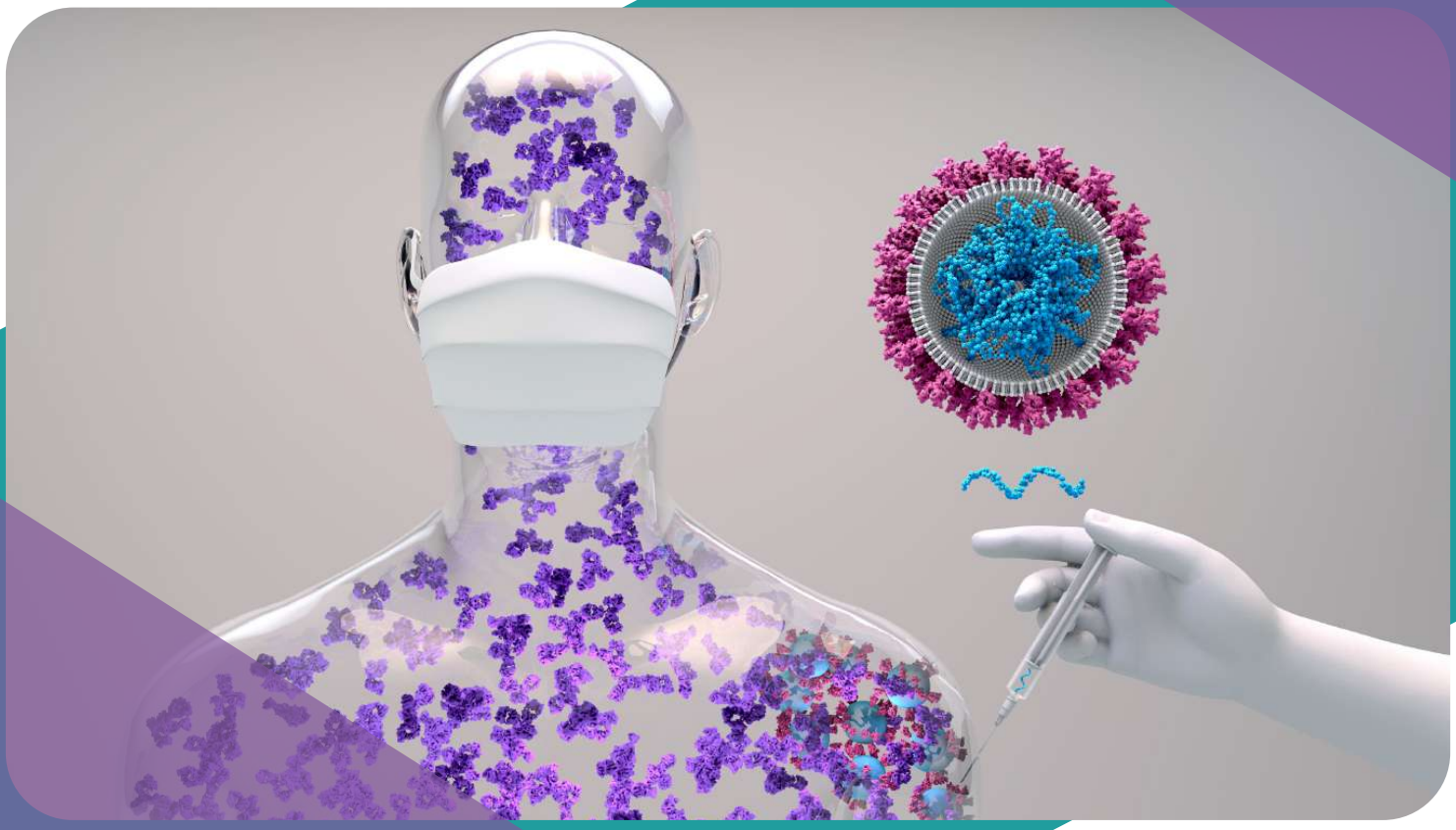
BioTechniques
The International Journal of Life Science Methods



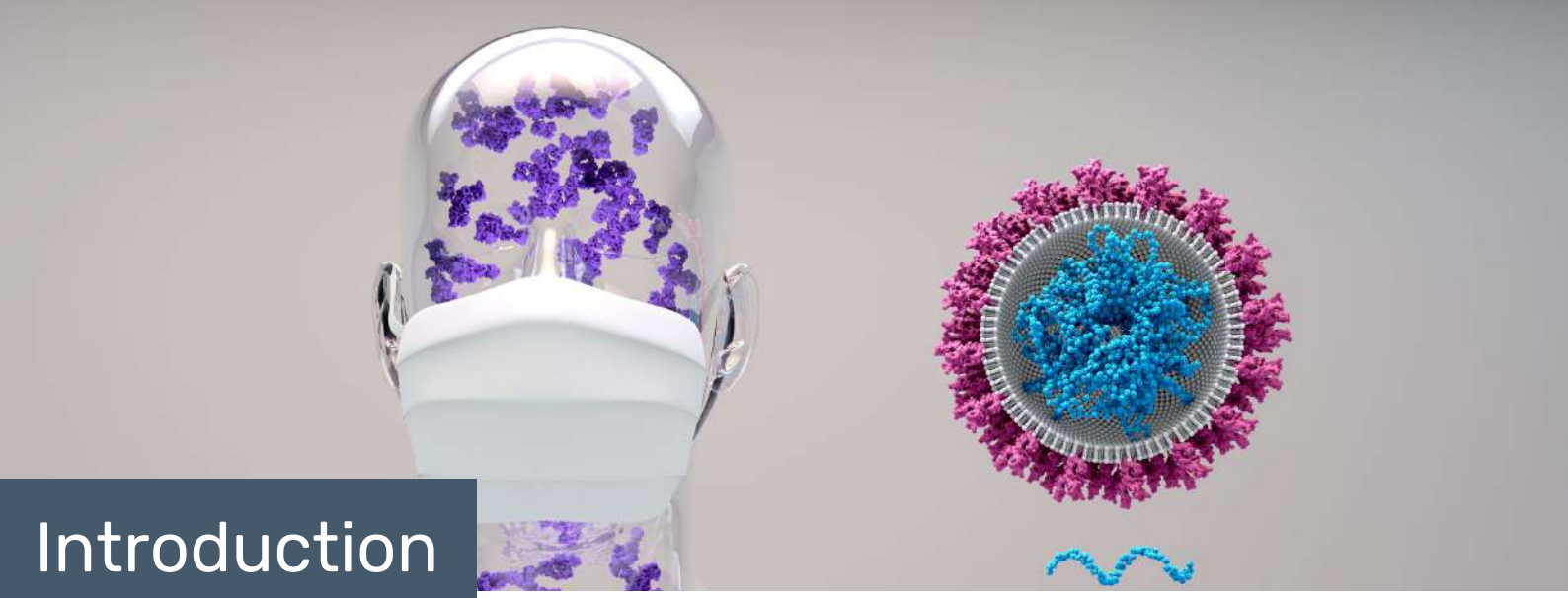
eBook: COVID-19 vaccines

TECAN

eBook: COVID-19 Vaccines

CONTENTS

TECAN.



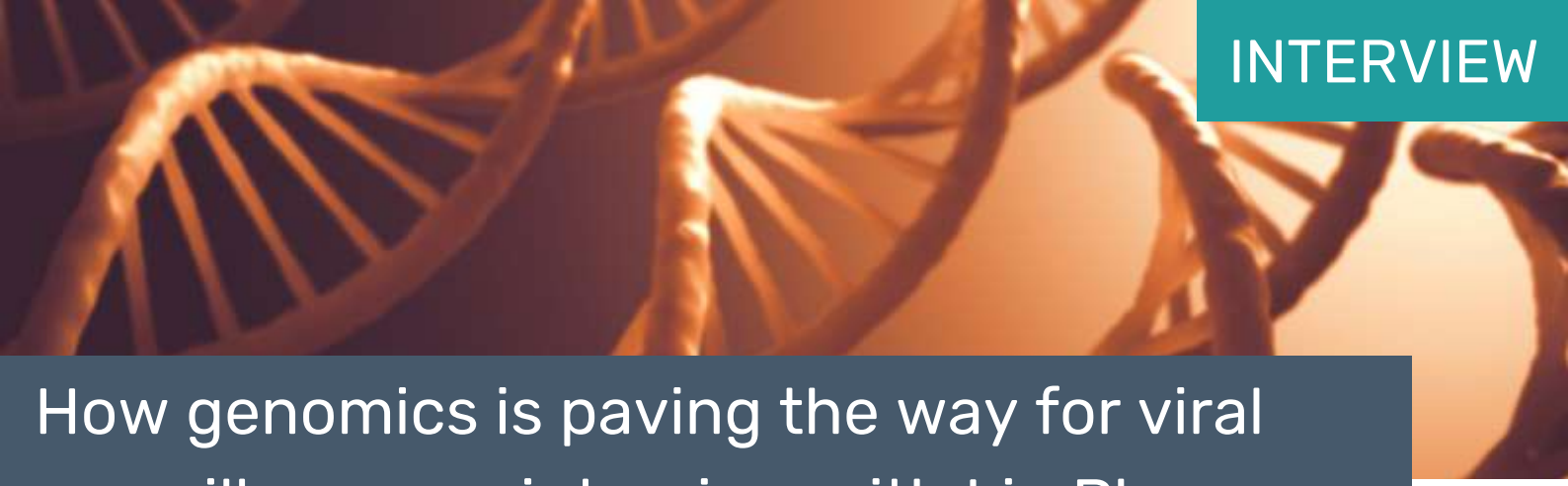
Introduction

Academics, industry and health services alike combined forces and sprang into action last year to develop, produce, trial, gain approval and rollout vaccines against COVID-19 in record-breaking time. While usually timelines for approval of novel vaccines can span time periods of up to 20 years, this feat was achieved by pharmaceutical companies in less than 1 year while maintaining safety protocols and high efficacies.

In the midst of the vaccine rollouts, there are still many unknowns surrounding the impact of COVID-19 vaccines on SARS-CoV-2 infection spread and protection of the population as a whole. Additionally, more research is required on the effect of the vaccine on many different population subgroups, including pregnant women, those with compromised immune systems and in some cases, the elderly.

In this eBook on COVID-19 vaccines, we will explore some of the foundational research into different COVID-19 vaccine compositions and efficacies. We will also explore the current and developing knowledge of the impact of these vaccines on the population, both as a whole and within subgroups, exploring methods and techniques for observation.

Abigail Sawyer
Senior Editor,
BioTechniques
asawyer@biotechniques.com



INTERVIEW

How genomics is paving the way for viral surveillance: an interview with Lin Pham

Lin Pham is the vice president of R&D at Tecan Genomics (CA, USA). Few people share Lin's incredibly versatile background in the life sciences; she has worked in R&D, marketing, project management, market development, as the general manager of the first genome company in China and was one of the pioneers in the genomics industry.

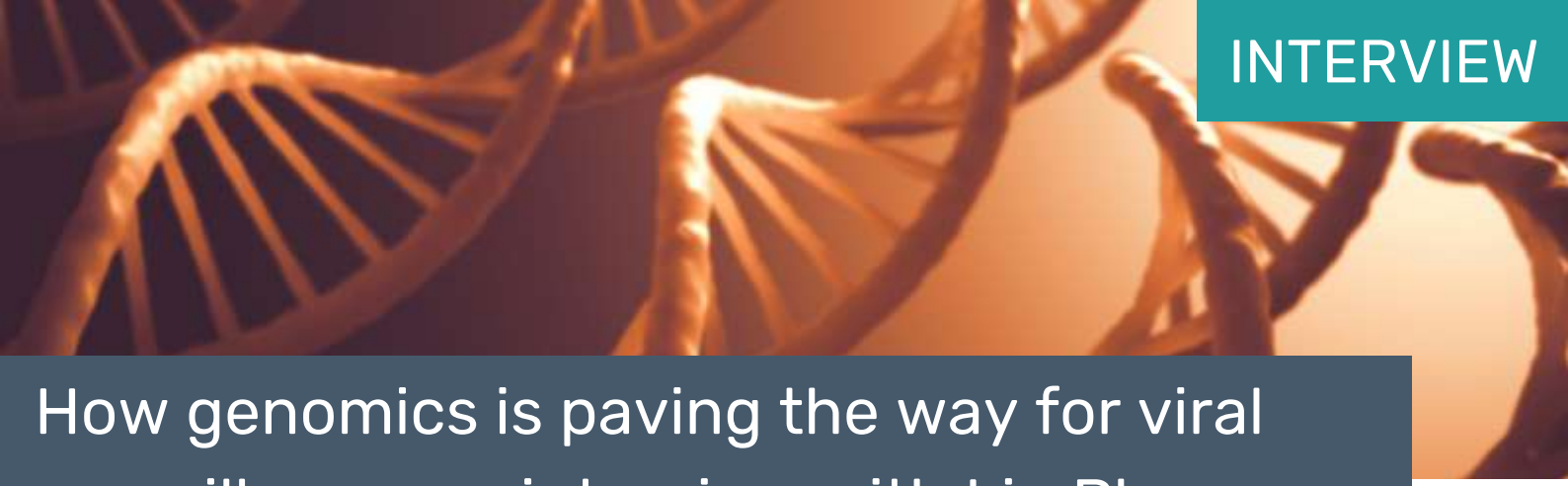
Q) Can you introduce yourself and give us a brief overview of your background?

I initially trained as an MD but, after I graduated from medical school, I soon realized that I didn't like to focus on treating symptoms without finding the underlying cause. I wanted to learn more about the fundamentals and the etiology of disease, which naturally led me into research and then quickly into genomics. I was there for the early days of DNA sequencing, before its evolution from cutting-edge research to translational research, to what it is now: the driving force of molecular diagnostics, pharmacogenetics, metagenomics and genetic testing, to name a select few. Nowadays, I consider myself a genomics research, development and applications insider and, yes, truly one of the 'dinosaurs' of the industry.

Q) You've been in the field during an incredibly exciting time, with advances that have reshaped research beyond recognition. How have these technologies been used during the COVID-19 pandemic?

These technologies were there at ground zero in Wuhan (China) when the virus first emerged. Because of modern genomics, researchers could quickly identify and isolate the causative pathogen and install public health measures to try to contain it. Unfortunately, SARS-CoV-2 (the causative agent of COVID-19) is proving harder to contain than the SARS outbreak in 2003, and thus we now find ourselves amid a global pandemic. However, the vast amount of genetic information we've rapidly uncovered, combined with recent technological breakthroughs, have allowed researchers to develop effective vaccines in record time. This has shown us that we are capable of responding quickly, and that when the next new virus emerges – because this will happen again – we will be much better prepared and hopefully be able to stop it before it becomes a pandemic.

So, to better answer your question, genomics played a pivotal role in the initial discovery of SARS-CoV-2 and is at the forefront of ongoing viral monitoring. Researchers in Wuhan were faced with two patients presenting with severe pneumonia-like symptoms, but who returned negative results to all the routine respiratory disease tests. They took a whole genome sequencing approach to discover what was causing the disease. RNA was extracted from the bronchoalveolar lavage fluid samples previously collected from the two patients and, using Tecan's Trio RNA-Seq™ Library Preparation Kit – designed to detect highly degraded and limited RNA – they constructed the library and sequenced the whole



INTERVIEW

How genomics is paving the way for viral surveillance: an interview with Lin Pham

transcriptome. They found a novel pathogen with high homology to SARS-coronavirus but with clear differences, which we now know as SARS-CoV-2.

To identify a new virus and nail it as the responsible agent – all within a week – is an utterly amazing result. What's even more incredible is the ability to do this with the abundant background noise of human ribosomal RNA and a whole bunch of microbes that reside in our body. Using next generation sequencing (NGS) technologies, they were able to deplete the unwanted noise and accurately sequence the entire viral genome.

Q) The pandemic has taken so many twists and turns. Could you explain the importance of ongoing viral monitoring and identifying new variants?

As we know, the virus has evolved and will continue to mutate as it infects more people. This makes ongoing monitoring – or surveillance – of the viral genome imperative to containing the pandemic and implementing appropriate public health measures. New variants may be more contagious, some more deadly, and, if one emerges that is both, this public health crisis will escalate even further.

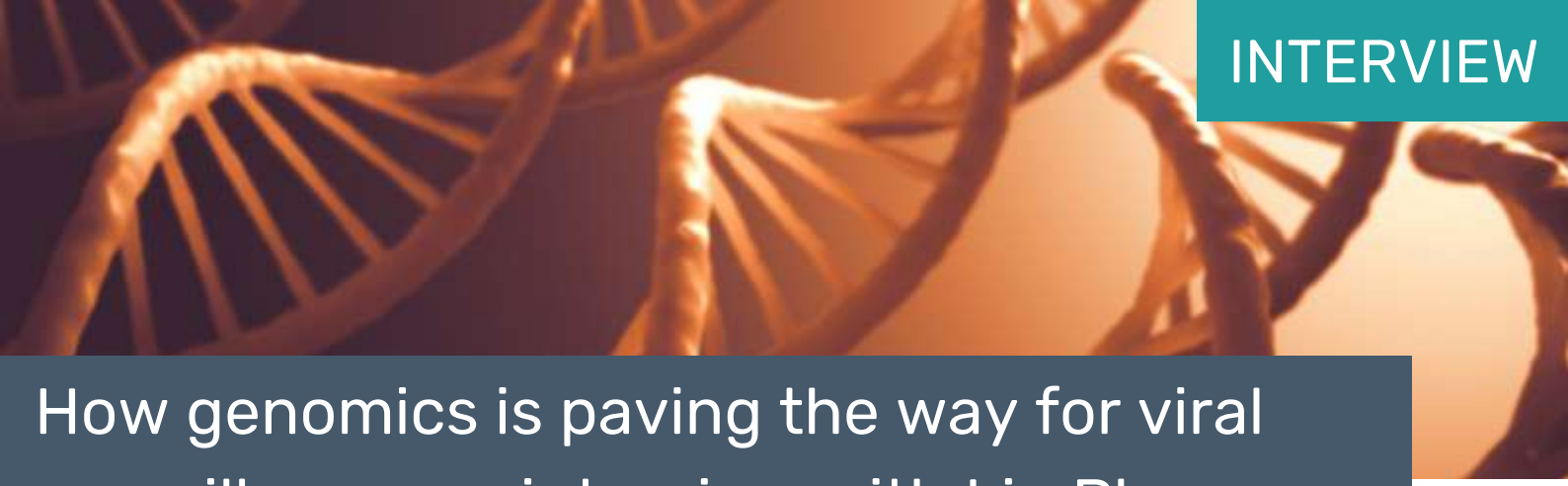
Ongoing surveillance is not only vital to tracking these variants, but also to answering the many unanswered questions that remain. Can we transmit the disease to others if we get infected while vaccinated? How long does immunity through vaccination last? How do we know if asymptomatic people are infectious? And, what do new variants mean for infected patients' recovery and long-term health? Ongoing viral surveillance with the use of NGS and other technologies will hopefully answer these questions, leading to a greater understanding of this virus and any future emerging pathogens.

Q) Do you think the pandemic has brought about a change in genomic research methods, or has it led to any new innovations?

Absolutely, yes. But NGS library construction methods are evolving all the time anyway, especially for greater sensitivity and speed, and reduced cost. When I visited multiple labs in China 2 years ago – before COVID-19 – one of the main desires for infectious agents testing was for a 24-hour turnaround time. So, if they collected a patient sample one day, the report was ready in the next day. Then, obviously cheaper and more robust kits with minimal hands-on time were also important goals.

Before that trip, I would've told you that the Tecan Genomics' Trio RNA-Seq kit was ideal for pathogen RNA-Seq. With its sensitivity and robustness, it is designed for super low input and degraded samples, but we realized the need for a faster workflow. This led to the development of Revelo™ RNA-Seq, which provides a robust solution meaning lower failure rates – with even better sensitivity – and the ability to finish in a 6-hour workflow. Although the work on this kit started before COVID-19, the pandemic certainly accelerated its development in response to the global need.

The sensitivity of these kits has also proven to be vitally important during the pandemic and is a crucial factor in helping to answer the questions above. Testing asymptomatic populations or potential



INTERVIEW

How genomics is paving the way for viral surveillance: an interview with Lin Pham

carriers may be inconclusive because of a lack of sufficient viral material to identify the virus using qPCR or to classify the variant through sequencing. This is where Tecan's Crescendo cDNA Synthesis for qPCR kit comes into play. Crescendo allows researchers to amplify precious samples in order to perform not only accurate qPCR, but also a range of additional downstream investigations like sequencing, gene expression and genotyping.

Q) What do you think is the next step? And what have we learnt most from the last 18 months?

In the research community and beyond, we see once again how powerful NGS technologies are in helping us during a global health crisis. We have learnt that we need to test more widely and much earlier, and the importance of having technologies that we can rapidly deploy globally. For Tecan Genomics, it has shown us the need to stay at the cutting edge, to have the right tools in place and to respond quickly. We simply don't know when the next SARS-CoV-2 variant will emerge, but we are better prepared to respond and to swiftly control the next infectious disease when it strikes.

Article

Bolstering the assay arsenal to investigate COVID-19

The COVID-19 pandemic has highlighted the importance of biosurveillance to understand, control and limit the spread of human diseases. In response to the pandemic, the laboratories at PTP Science Park (Lodi, Italy) quickly enhanced their setup, introducing various assays to gain a deeper understanding of the virus. Researchers now not only diagnose COVID-19 but measure anti-SARS-CoV-2 antibody levels after infection and vaccination to learn about the pathogenesis of the disease and how our immune systems respond to infection.

PTP Science Park is a business consortium that operates in a wide range of sectors, including agrifood, life sciences and bioeconomics. The company's position at the forefront of scientific research allowed it to ramp up its diagnostic and research capabilities in response to the COVID-19 pandemic. The science hub is still engaged in various projects, but its extensive research facilities, staff and resources have been focused on the battle against COVID-19 on multiple fronts.

Stefano di Giovine, COO of PTP Science Park, explained: "At the beginning of the COVID-19 pandemic in March 2020, PTP Science Park already had the facilities and the ability to analyze the virus. We started with a small number of samples but, thanks to Tecan's state-of-the-art automation solutions – such as the Freedom EVO® workstation – we already had the capacity to evaluate up to 2,000 samples per day. More recently, in March this year, we started a drive-in facility to enable us to take in even more samples. To date, we have processed about 160,000 samples, and are the primary laboratory in the Lodi area – supporting care homes and hospitals."

Located within PTP Science Park is the SMeL laboratory, which was added to the regional crisis unit in response to the pandemic. It is authorized to operate as a specialized medical laboratory and carries out various tests using a biosurveillance toolkit that includes qPCR for diagnostic testing of oropharyngeal swabs, and quantitative serological assays for SARS-CoV-2 IgG, IgM and IgA antibodies, to evaluate the immune response post-infection or post-vaccination.

Stefano continued: "Our activity with SARS-CoV-2 is not limited to diagnostics. We are one of the few laboratories in the region authorized to sequence and report new variants, partly thanks to Tecan's assays and the DreamPrep™ NAP workstation. We also have a program funded by the Lombardy Region and the Cariplo Foundation called 'COVID in pets', to investigate whether animals can catch and transmit the virus."

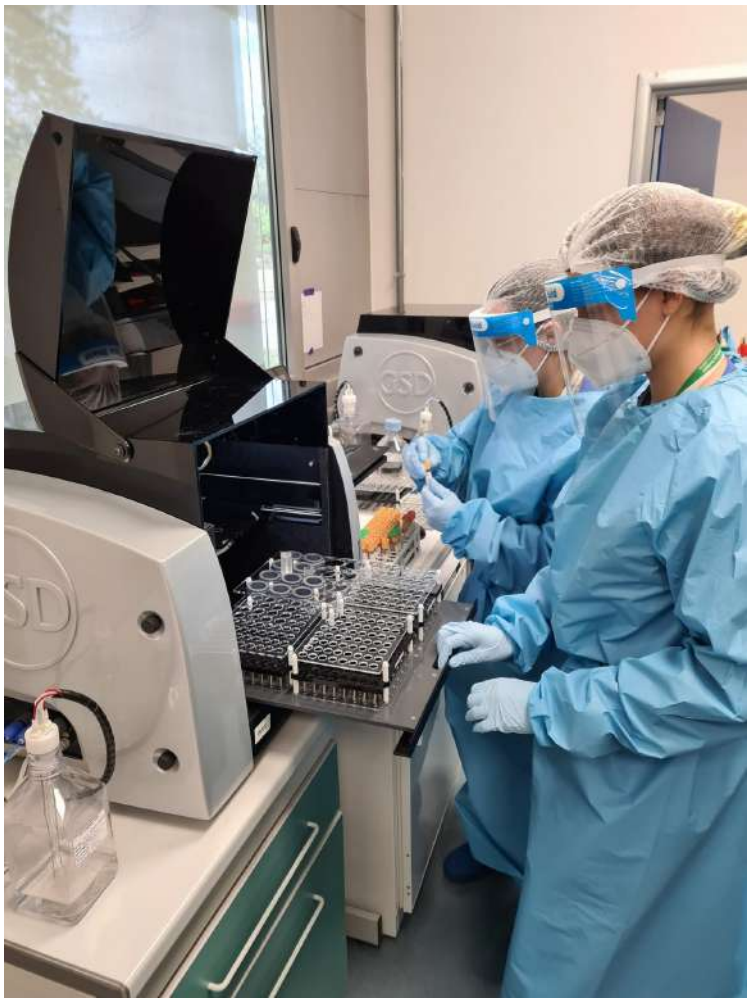
"The demand for serological testing to check for antibodies in response to the vaccine has increased dramatically. Using Tecan's SARS-CoV-2 S1/RBD IgG Ab ELISA, we can detect antibodies against the virus spike protein's receptor binding domain, which allows us to examine variations in the immune response to the virus. For example, we have found that some people gain immunity after contracting COVID-19, whilst others have been infected up to three times as they fail to generate immunizing antibodies. We expect a real boom in testing requests in the coming months with all of the clinical studies underway."

"The scope of our research continues to expand as

Bolstering the assay arsenal to investigate COVID-19

the pandemic progresses. Recently we started a small, local research project to evaluate the levels of IgG antibodies against the virus before and after vaccination. Again, using Tecan's SARS-CoV-2 S1/RBD IgG Ab ELISA we can accurately gather this data and find variations in peak antibody levels following the vaccine."

"Tecan's instruments and assay kits have allowed us to act swiftly in response to the pandemic and really make an impact, especially in the region. Ongoing viral surveillance with PCR testing and ELISA kits will help us discover new variants and continue building our knowledge on how our immune system responds to infection," concluded Stefano.



Left: Researchers at the PTP Science Park

To find out more about Tecan's SARS-CoV-2 S1/RBD IgG Ab ELISA kits, visit: www.tecan.com/interrogate-covid-19-with-tecan-sars-cov-2-igg-elisa To learn more about PTP Science Park, go to www.ptp.it

For reprint orders, please contact: reprints@future-science.com

Bioanalysis

Key points of technical review for the registration of SARS-CoV-2 antigen/antibody tests

Hongran Li[‡,1], Jingyun He[‡,1], Wen Bao[‡,1], Peirong Wang[‡,1], Yunfeng Lv[1], Chao Xu[1], Peng Hu[1], Yu Gao[1], Shengwei Zheng[1], Juanjuan An[1], Gang Deng[1] & Jinchun Dong*[,1]

[1]Center for Medical Device Evaluation, NMPA, Building 1, No. 50 Qixiang Road, Haidian District, Beijing 100081, PR China

*Author for correspondence: dongjch@cmde.org.cn

‡Authors contributed equally

Coronavirus disease-2019 (COVID-19), caused by the novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has spread globally since its first report and become a worldwide pandemic. In response to the outbreak of COVID-19, Center for Medical Device Evaluation, NMPA (CMDE) initiated emergency review and approval procedures to accelerate the process of reviewing emergent medical products and issued the Key Points of Technical Review for the Registration of SARS-CoV-2 Antigen/Antibody Tests (Key Points) to provide the requirements on the technical review of the tests. With uncontrolled spread and evolution of COVID-19 in the world, continuous prevention and measurements are necessary for fighting this pandemic and SARS-CoV-2 antigen/antibody tests are still urgently needed. This article is an attempt to expand clarification of the Key Points to wider audiences based on current understanding of SARS-CoV-2 to facilitate the development and application of SARS-CoV-2 antigen/antibody tests.

First draft submitted: 16 August 2020; Accepted for publication: 2 December 2020; Published online: 11 January 2021

Keywords: antibody • antigen • registration • SARS-CoV-2 • tests

A novel coronavirus pneumonia (coronavirus disease-2019, COVID-19), caused by a novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2, originally named 2019-nCoV), has spread globally and rapidly since its first report [1–4]. The WHO declared this outbreak as a Public Health Emergency of International Concern on 30 January 2020 [5] and characterized COVID-19 as a pandemic on 11 March 2020 [6]. As of 1 November 2020, there have been 45,942,902 confirmed cases of COVID-19, including 1,192,644 deaths globally according to the report from WHO [7]. This pandemic has still not been controlled effectively. Most of the patients with COVID-19 have mild or moderate disease, however up to 5–10% present with severe and even life-threatening disease course. Although different treatments are being evaluated, the most important way to control this pandemic is the development of an effective and safe vaccine widely available [8].

SARS-CoV-2 taxonomically belongs to the Sarbecovirus subgenus of the coronavirus family and is the seventh member of human coronavirus [9]. Four of these human coronaviruses, HCoV-229E, HCoV-NL63, HCoV-OC43 and HCoV-HKU1, usually cause mild-to-moderate symptoms and the other two human coronaviruses, severe acute respiratory syndrome coronavirus (SARS-CoV) and the Middle East respiratory syndrome coronavirus (MERS-CoV), generally cause more severe cases [10]. The SARS-CoV-2 genome is around 30 kb and contains *ORF1ab* (encoding polyproteins PP1ab and PP1a), structural proteins genes, which encode surface glycoprotein (S), an envelope protein (E), membrane protein (M), nucleocapsid (N) proteins and several accessory proteins [11].

Rapid, accurate and sensitive detection for SARS-CoV-2 is essential for the prevention of the COVID-19 pandemic and the treatment of COVID-19 patients. Real-time reverse transcription PCR (RT-PCR) has been used as gold standard for SARS-CoV-2 RNA detection. But in past clinical practice and research, the false negative results of nucleic acid tests may occur due to different factors, such as patient sampling quality, viral load and distribution, and the standardization of the detection process among others [12]. Reliable serological tests can provide more information about SARS-CoV-2 infection, and SARS-CoV-2-specific IgG and IgM have been used

newlands press

as evidence for COVID-19 diagnostics [13]. But what should be noted that the positive rate of SARS-CoV-2-specific IgM and IgG is low in the first week since symptoms onset, and usually the serological tests cannot directly diagnose the presence of the virus [13]. It just can be used as a supplementary method to increase the positive rate of the diagnosis for suspected cases with negative SARS-CoV-2 nucleic acid results and can also help to confirm the disease status [12].

In order to respond to this pandemic effectively, the National Medical Products Administration of China (NMPA) initiated the emergency review and approval procedures soon after its outbreak to accelerate the approval of SARS-CoV-2 tests. The Center for Medical Device Evaluation (CMDE) drafted the 'Key Points of Technical Review for the Registration of SARS-CoV-2 Antigen/Antibody Tests' (hereinafter referred to as the Key Points) on 11 February 2020 and issued it on 25 February [14] to provide a guidance for reviewer and commercial manufacturers. This Key Points is consistent with the Provisions for *In-Vitro* Diagnostic Reagent Registration of China (2014) [15] and provides requirements and methods for evaluation of SARS-CoV-2 antigen/antibody detection reagents, which mainly includes analytic performance (such as limit of detection, inclusive test, cross reactivity validation, etc.) and clinical trial requirements. Based upon this Key Points, 25 SARS-CoV-2 antibody and two SARS-CoV-2 antigen tests have been approved by NMPA (as of 10 November 2020), all of which are applied by Chinese manufacturers and play important roles in fighting against COVID-19 pandemic in China.

Since the regulatory requirements for SARS-CoV-2 products are different among different regulatory agents, correct and comprehensive understanding of each guidance will improve the design, development and application for marketing or emergency use during this pandemic. With the uncontrolled spread and evolution of SARS-CoV-2 in the world, continuous prevention and countermeasures are necessary, and comprehensive explanation of the requirement for SARS-CoV-2 tests are needed for wider audiences. In addition, continuous updated understanding of SARS-CoV-2 and COVID-19 put forward the necessity for analysis of the suitability of this Key Points issued at the beginning of this pandemic for current situation. This article is an attempt to provide clarification of the Key Points to wider audiences based on the current understanding of SARS-CoV-2 to facilitate the development and application of SARS-CoV-2 antigen/antibody tests.

Scope of application

The Key Points is applicable for SARS-CoV-2 antigen/antibody assays, which are used to conduct *in vitro* qualitative detection of SARS-CoV-2 antigen/antibody in the samples of serum, plasma, whole blood, throat swab, bronchoalveolar lavage fluid, sputum or other respiratory secretions. Semi-quantitative and quantitative assays are not included because there is no international standard reference material to demonstrate them. It is also difficult to be verified by clinical trials according to limited understanding of clinical significance of quantitative and semi-quantitative tests [16]. Currently, only qualitative assays have been approved for SARS-CoV-2 antibody detection. Since the Key Points is to provide the requirement for the application of SARS-CoV-2 antigen/antibody assay, the clinical intended purpose must be clear. With deep understanding of SARS-CoV-2 and its clinical significance, the semi-quantitative and quantitative assays may also be included in the guidance.

Analytical performance evaluation

According to Provisions for *In-Vitro* Diagnostic Reagent Registration of China (2014) [15], the analytical performance of candidate tests should be evaluated and verified, which includes limit of detection, inclusivity, analytical specificity, precision, hook effect and other suitable performance. Since the reagents used for research in the lab are quite different from those produced in a real production environment in many aspects, such as personnel, environment, operation and production batch, and the performance may not be represented by each other, the reagents used for performance evaluation should be manufactured under the control of quality management system. The analysis report submitted to CMDE should include detailed information such as research aim, materials and methods, results, statistical analysis and conclusion. It is suggested to focus on the following items.

Sample collection & processing

Correct sample collection and processing is very important for correct SARS-CoV-2 assay. Sample collection by individuals and by medical providers may lead to very different results. Although each collection method has its own advantages, here only sample collection by healthcare provider is discussed. The factors of disease course, clinical symptoms and medication treatment should be considered when determining the collection time point of clinical samples. For swab specimens, the requirement of sampling swab and sample preservation solution (sampling

solution) should be clarified, including sampling swab material (including swab head and swab rod), sample container and sample solution (such as composition, concentration and dosage of sample solution). For blood samples, anticoagulant should be investigated. The sample collection is suggested to be conducted in accordance with the Technical Guidelines for Laboratory Detection of SARS-CoV-2-Infected Pneumonia [17].

Limit of detection

Limit of detection (LoD) studies determine the lowest detectable concentration of SARS-CoV-2 antigen or antibody. According to the principle of method for LoD evaluation in EP17-A2 [18], the determination of LoD for the SARS-CoV-2 IgM or IgG antibody assays can be performed by detecting a serial of dilutions from samples with known antibody concentration and these dilutions are tested in replicates by the measurement procedure in which results are judged to be negative or positive. Each dilution should be divided into three to five copies and no less than 20 replicates are required for each copy. For each dilution, a ratio is computed as the number of replicates with a 'positive' outcome per the total number of replicates tested. The antibody level with 90%–95% positive detection rate should be used as the LoD. Antibody titer should be confirmed by suitable method. For antigen detection, the difference is that the sample containing the virus instead of antibody is diluted for the LoD detection. Three representative clinical samples from different source or virus stock from culture should be used for LoD analysis.

The established LoD should be verified using at least three virus strains or clinical samples with temporal and regional characteristics, which are different from those used for LoD determination. The concentration should be equal to the claimed LoD and the positive detection rate should reach 90–95%. The consistence of diluents with the matrix of applicable sample type will decrease the matrix effects. Negative samples can be used for dilution. For antigen detection, detailed confirmation method for virus titer, the method for virus identification and the results of samples should be provided. For antibody detection reagent, it is required to provide detailed methods and verification results of antibody type and titer.

Inclusivity

Inclusivity refers to the ability to detect different types of pathogens to be tested, including known types, pathogen subgroups (groups), serotype and genotype.

Five major groups (G_{614}, S_{84}, V_{251}, I_{378} and D_{392}) of SARS-CoV-2 viral populations have been defined [19] and the relationship between Gly614 replacement in the spike protein and increased infectivity has been reported [20–23]. This feature of genetic diversity and rapid evolution of SARS-CoV-2 virus might induce conformation change of proteins and probably lead to the antigenicity change [24]. Whether the diversity affects the performance of tests should be evaluated carefully.

The seroconversion of different types of antibody displayed distinct features according to several studies [12,25] and the time of seroconversion may vary between individuals [26]. The first seroconversion day of IgA has been reported at the second day since the onset of the initial symptoms, while for IgM and IgG it was reported at 5 days after onset. After 32 days since symptom onset, the seroconversion rate for IgA, IgM or IgG reached 100% [25]. IgM and IgG specific to N and S proteins have been reported to increase gradually after symptom onset and can be used for detection of SARS-CoV-2 infection [27]. The accuracy of antibody tests for people infected with COVID-19 varied at different time points since the onset of first symptoms, which was 30% at the first week, increased to 70% after 2 weeks and reached the highest in week 3 (more than 90% detected) [28]. These features of different types of antibodies request the detailed analysis for the samples collected at different time and different regions.

In order to evaluate the capacity of proposed assay to detect different isolates of SARS-CoV-2, the real clinical samples of COVID-19 patients from different sources with temporal and regional characteristics should be verified. It is difficult to obtain more clinical samples for so many manufacturers to verify the performance in such an emergent situation. Ten different clinical samples derived from COVID-19-positive patients with positive SARS-CoV-2 antibody collected at different times and from different places are regarded as capable for verifying the inclusivity. Those COVID-19-positive cases with negative real-time RT-PCR results should also be included if they are confirmed by criteria outlined by 'Guidelines for the Diagnosis and Treatment of Coronavirus Disease 2019 (8th Trial Version)' in China [13]. More samples may add burden on the manufacturer and delay the development process and accessibility of tests during an emergent period. So at least ten different samples are recommended for detection of antigens, IgM antibody and IgG antibody and the LoD and repeatability should be studied. The concentrations of SARS-CoV-2 antigen, IgG or IgM in clinical samples since the symptom onset or PCR positive without symptoms, as well as the determination methods, should also be provided

Although total antibodies (combination of IgA, IgG and IgM) have been used as targets of SARS-CoV-2 antibody assay, the function and clinical meanings of IgA itself still remain unclear. IgA has not been regarded as one of the evidences for COVID-19 diagnostics according to 'Guidelines for the Diagnosis and Treatment of Coronavirus Disease 2019 (8th Trial Version)' in China [13]. But with more and more investigation of IgA, the indication of IgA test can clarify and related information can be updated.

Analytical specificity

Analytical specificity reflects the ability of an assay to detect solely the measurant [29], including cross reactivity and interference. The specificity of the assay depends on the specific recognition of the SARS-CoV-2 antibody and antigen. For an immunoassay, special attention should be paid to the possibility of the cross reactivity or the recognition of nontarget components, which is different from inclusivity. For inclusivity, the measurant is a target antigen or antibody for SARS-CoV-2 virus in samples from different regions or different times. If the assay is unable to detect all of the types of SARS-CoV-2, false negative results may occur. But for specificity, non-SARS-CoV-2 pathogens with the potential to give false positive results should be investigated.

Cross reactivity

The cross reactivity for related pathogens should be verified. The sequence homology of SARS-CoV-2 and other human coronaviruses in their respective spike proteins make them most likely cross-reactive with each other and strong cross-recognition between SARS-CoV-2 and SARS/MERS-CoV antibody has been reported [30]. Individuals infected with influenza A and B may have similar symptoms as those infected with SARS-CoV-2 and coinfection of SARS-CoV-2 and influenza [31,32] or other common respiratory pathogens has been reported [33]. Immunopathological similarities between COVID-19 and influenza has been indicated [34].

In order to evaluate the specificity of tests comprehensively, the cross reactivity of SARS-CoV-2 antigen, IgM antibody or IgG antibody tests should be verified respectively. According to the multiple respiratory pathogens related guidance [35] and the opinion of experts, the following potential cross reactants are suggested: antigens of or antibodies against endemic human coronavirus (HKU1, OC43, NL63 and 229E), SARS coronavirus and MERS coronavirus; H1N1 (novel influenza A H1N1 virus 2009, seasonal H1N1 influenza virus), H3N2, H5N1, H7N9, influenza B Yamagata, Victoria, respiratory syncytial virus, rhinovirus group A, B and C, adenovirus 1, 2, 3, 4, 5, 7 and 55, enterovirus group A, B, C and D, Epstein–Barr virus, measles virus, human cytomegalovirus, rotavirus, norovirus, mumps virus, varicella zoster virus and *Mycoplasma pneumoniae*.

The source, negative/positive, species/type and concentration/titer confirmation of antigen, antibody and other samples used for verification of cross reactivity shall be provided. For antigen tests, medical decision concentrations of viruses, such as 10^5 plaque forming unitor/ml (PFU/ml) or higher, are recommended for cross-reactive verification. For antibody tests, the methods and results for antibody confirmation should be clarified.

The cross reactivity between IgG antibody and IgM antibody against SARS-CoV-2 with high concentration should be verified. No less than 20 normal human samples should also be evaluated.

What should be noted that this list is just a requirement for the premarket application of SARS-CoV-2 antigen/antibody detection tests and does not provide all of the potential cross-reactive pathogens. More and more evidence indicated that SARS-CoV-2 infection can lead to multiple organ failure, especially lungs, heart, kidney, gastrointestinal and hepatic system, brain and skin [36,37] and coinfection of SARS-CoV-2 and other pathogens have been reported [38,39], such as HIV [40] and hepatitis virus [41] etc. So potential cross reaction in such situations should be considered during clinical utility.

Interference of endogenous/exogenous substances

The potential interference of related endogenous/exogenous substances should be evaluated for the assay. Substances recommended can be referred to in the issued Key Points [14]. The potential highest concentration of interfering substances (regarded as 'worst case' conditions) and the weak positive concentration of virus antigen and antibody are recommended for the interference evaluation. If streptavidin–biotin system is used for the assay, the interference of free biotin in samples [42,43] should be analyzed carefully.

IgM antibody class specificity

At least five samples containing SARS-CoV-2-specific IgM antibody should be evaluated for class specificity. The detection for the samples treated with certain chemicals (e.g., 2-mercaptoethanol or dithiothreitol) should give a negative result.

Precision

According to EP05-A3, precision means the closeness of agreement between indications or measured quantity values obtained by replicate measurements on the same or similar objects under specified conditions [44]. It includes repeatability, intermediate precision, and reproducibility, which indicates the measurement precision under a set of repeatability conditions, intermediate precision conditions, or reproducibility conditions of measurements seperately [44].

Applicants should set up acceptable criteria, including the standard deviation or coefficient of variation, for the evaluation of precision. Multiple factors, such as detection time, equipment, operator, location, run and other elements, should be considered for the precision analysis. Reasonable evaluation period should be set that at least 20 or 5 days according to different experiment designs. The detailed information can be found in the EP05-A3 document [44]. Consequently, between-batch/within-batch, between-day/within-day and between-operator precision can be evaluated comprehensively.

The clinical samples of at least three levels of concentration should be analyzed for precision evaluation, which are negative samples, positive samples near the cutoff and (moderate/strong) positive samples. The appropriate precision requirements should be set in accordance with product characteristics.

For negative samples, the concentration of analytes is lower than LoD or equals to zero and the percentage of negative results should be 100% ($n \geq 20$). For positive samples near the cutoff, the concentration of analytes is slightly higher than LoD and ≥95% samples should yield positive results ($n \geq 20$). For moderate/strong positive samples, 100% samples should give positive results with CV ≤15% ($n \geq 20$).

What should be noted that precision study described in EP15-A3 is not suitable for the precision evaluation of the SARS-CoV-2 tests. The protocol described in EP15-A3 is used to verify a manufacturer's claims for precision and the trueness of the measurement procedure relative to the assigned values of materials with known concentrations [45], but not to establish or validate the precision performance of a measurement procedure, which has been described in EP05-A3 [44].

Hook effect study

The hook effect is observed in sandwich immunoassays, where at very high concentrations of the analyte, the assay signal is saturated and leveled off [46].

It is necessary to detect gradient dilutions diluted from multiple high-concentration samples to evaluate the hook effect. Three to five duplicates for each dilution should be tested. The results of the hook effect should be clearly indicated on the instructions for use.

Verification of reference materials

Three batches of products should be used to test the reference materials produced by the manufacturer described in the 'Research data of main raw materials' section. Considering the differences of materials, time, environment, operators and other occasional factors, the performance maybe different among batches. It is the common way to evaluate the performance of *in vitro* diagnostic (IVD) products using three or more batches to reduce the deviation caused by different batches, which is also described in the 'The Announcement of Registration Document Requirements and the Approved Certificate Format for *in vitro* Diagnostic Reagents (2014-09-05)' [47]. Comprehensive experimental data of the verification should be provided.

Instrument used for SARS-CoV-2 tests

In China, an IVD reagent and instrument used together as a system are registered separately as a different product and the matched instrument is required to be listed in the instructions for the use of the IVD reagent. It is the same for SARS-CoV-2 antigen/antibody tests that all of the instruments used for result reading and interpretation should be listed in the product instructions, and the analytical performance evaluation using each instrument should be provided.

Specimen types

If serum and plasma are the applicable specimen types for the product, the comparability between them can be verified by homologous comparison method. For whole blood, two methods can be chosen for evaluation, one is to study all of the analytical performance and another is to analyze at least the LoD, inclusivity of virus samples from different regions, precision and the homologous comparison test between different specimen types. The sample collection is suggested to be conducted in accordance with the 'Technical Guidelines for Laboratory Detection of SARS-CoV-2-Infected Pneumonia' [17].

Determination of cut-off value

It is required to provide the research data that determines the cut-off value for positive and negative results or a zone where the response is 'indeterminate', which can be referred to EP28-A3c [48]. Different geographic regions, different infection stages and physiological status should be considered when selecting samples for the establishment of the cut-off. Moreover, differences among sample types should be clarified and confirmed respectively. The receiver operating characteristic curve is recommended to determine the cut-off, and the equivocal zone should be investigated if it exists.

Research data of main raw materials, production process & reaction system

The main raw materials of this product include antigen, antibody, quality control materials and reference materials. It is required to provide the research data on the selection and source, preparation process and quality standard of main raw materials. For the main production process, the principle of reaction should be introduced and the main production process should be clarified. The flow chart can be used to make it clear.

The establishment, optimization and confirmation of reaction system are required. The contents include reaction time, reaction temperature, washing liquid volume and washing times (if involved), as well as loading method and loading amount, should be optimized. As to IgM antibody assay with indirect method, the interference of high-concentration-specific SARS-CoV-2 IgG on IgM results should be considered and the reasonable sample treatment should be designed to reduce this interference.

Evaluation of stability

Research data on stability mainly involves two parts, the stability of the reagent and the stability of applicable samples. The stability of applied reagent mainly include shelf-life stability, in-use stability and accelerated stability under damage of high temperature and transport simulation [49]. For shelf-life stability research, at least three lots of products are required to be evaluated. The sample stability mainly focuses on the validity of specimens under refrigerating and freezing conditions. Freeze–thaw effects should be evaluated if applicable. The treated samples, such as samples in preservation solution, should be studied. Moreover, if heated samples (e.g., thermal inactivation) are used for detection, the interference of heating factor should be verified for positive samples with SARS-CoV-2-specific IgM antibody before and after being heated. No less than five positive heated and unheated samples near the cut-off value should be compared.

Clinical studies

The clinical trial for the SARS-CoV-2 antigen/antibody tests should be conducted in at least three clinical sites (including Centers for Disease Control and Prevention) as required, which helps to ensure the diversity of the cases enrolled in the clinical trial [50]. Ideally, the clinical performance of SARS-CoV-2 antibody tests should be evaluated by comparing the tests with the results of the available clinical reference standard for COVID-19. Additionally, the clinical performance of SARS-CoV-2 antigen tests should be evaluated by comparing the tests with the results of the synchronous SARS-CoV-2 nucleic acid tests.

Determination of clinical reference standard

The SARS-CoV-2 antigen/antibody test is generally expected to be used for the auxiliary diagnosis of COVID-19. Therefore, in clinical trials, the clinical diagnosis criteria for COVID-19 and the judgement criteria of disease progression shall be used as reference in the comparative study of SARS-CoV-2 antibody tests. 'Diagnosis and Treatment Protocol for Novel Coronavirus Pneumonia' and other documents issued by the National Health Commission of the People's Republic of China have specified the diagnosis criteria for COVID-19 [13]. It is recommended to compare the positive detection rate of antibody test to that of the nucleic acid test of COVID-

19 in order to fully evaluate the clinical significance of antibody tests. For the clinical trial of the SARS-CoV-2 antigen tests, the SARS-CoV-2 nucleic acid tests will be the preferred comparator and the diagnosis criteria for COVID-19 will help to evaluate the clinical performance more scientifically and explain the inconsistent result.

Study population & inclusion criteria

Subjects or samples enrolled in the clinical trial should represent the target population of the test. The studied population of SARS-CoV-2 antigen or antibody tests should be the COVID-19 suspected cases. The definition of 'suspected case' can refer to the 'Diagnosis and Treatment Protocol for Novel Coronavirus Pneumonia' [13] issued by the National Health Committee of the People's Republic of China.

The subjects enrolled in the clinical trial should represent various types of target population, including the confirmed cases of COVID-19 (including some cases in recovery period) and excluded cases. For the antibody tests, the confirmed cases should include the patients in different course of disease, and for the antigen tests, the confirmed cases are suggested to be mainly of the patients in the early stage of infection.

Moreover, for the evaluation of SARS-CoV-2 antibody tests, it is required to include some subjects with multi-time-point surveillance data of different stage of disease, where SARS-CoV-2 nucleic acid and antibody transformation can be detected.

Sample types for clinical trial

Generally, the sample types of SARS-CoV-2 antibody detection reagent are serum, plasma and whole blood. The sample types of SARS-CoV-2 antigen detection reagent are throat swab and other respiratory samples. The sample collection is suggested to be conducted in accordance with the 'Technical Guidelines for Laboratory Detection of SARS-CoV-2-Infected Pneumonia' [17].

For different sample types, if it is proved that there is no difference in analytical performance (e.g., there is no difference between serum and plasma samples in analytical performance), the detection data for different types of sample can be analyzed and summarized together or homologous samples comparison can be conducted in clinical trials. In other cases, statistical analysis for each type of sample is recommended and the requirements for the number of cases should be met respectively.

Sample size for clinical trial

At present, the developed SARS-CoV-2 antigen/antibody test is qualitative detection product generally. In the process of clinical trial, it is suggested to estimate the minimum sample size using a reasonable statistical model according to the preset clinical evaluation measures, such as clinical sensitivity, clinical specificity and relevant statistical parameters. Among them, the clinical evaluation measures should be confirmed in accordance with the previous studies of this product. For instance, if the expected clinical sensitivity is 85%, the confirmed cases are estimated to be no less than 200 cases according to the sample size calculation formula for sampling precision. If the expected specificity can reach 95% and according to the clinical requirements, the specificity of such tests should be no less than 90%, based on the sample size calculation formula for single arm study with performance goal, the excluded cases enrolled should be no less than 300. Applicant may estimate the sample size according to the specific characteristics of product.

In order to conduct more scientific and comprehensive evaluation of clinical performance of product, a certain number of patients at different stages of disease should be included for the SARS-CoV-2 antibody tests. When necessary, the sample size of each important subgroup should meet the statistical requirements.

Statistical analysis of clinical trial results

The aim of the clinical trial for the SARS-CoV-2 antibody tests is to verify the consistency between the test product and the clinical reference standard. The test results are generally summarized in the form of table 2 × 2, by which the clinical sensitivity, clinical specificity and confidence interval are calculated.

For the SARS-CoV-2 antibody tests, stratification analysis should be conducted for the patients in different stages of disease course. For the multi-time-point surveillance samples of COVID-19 patients for antibody tests evaluation, it is required to compare the test results of nucleic acid simultaneously and evaluate the detection capability and window period of the antibody reagent for trial for SARS-CoV-2 infection.

In order to evaluate the clinical significance of IgG and IgM joint detection, in addition to separate statistics for IgG and IgM test results, a statistical analysis of comprehensive evaluation of the two indicators should also be performed.

All inconsistent results in the clinical trial should be analyzed fully by combining them with the epidemiological background, clinical symptoms, prognosis of disease and other information of patients.

Formal requirements for clinical evidences

In accordance with the requirements of 'Provision of *In-Vitro* Diagnostic Reagent Registration' [15], 'The Announcement of Registration Document Requirements and the Approved Certificate Format for *in vitro* Diagnostic Reagents' [47] and other regulatory documents, applicant should submit ethical review opinions of all institutions, clinical trial protocol, clinical trial report and summary report on clinical trial.

The summary sheet of clinical trial data should be submitted as the attachment of clinical trial report. The data sheet should include the record number, age, gender, sample type, sampling time, background information on clinical diagnosis, detection results of this product and the confirm or exclusion results of SARS-CoV-2 infection. Meanwhile, it is required to specify the nucleic acid test results (including the name of nucleic acid test reagent) used for the diagnosis of the disease. All the data in clinical trials should be traceable. Considering that the generation of antigen and antibody is closely related to the course of pathogen infection, it is suggested to specify the time of disease onset, the changes of symptoms and the disease prognosis of patients in clinical background information.

Conclusion

The Key Points for the SARS-CoV-2 test was issued to facilitate the availability of SARS-CoV-2 antigen and antibody tests to respond to the COVID-19 public health emergency effectively. It describes the requirements for the SARS-CoV-2 antigen and antibody tests for emergency use during the COVID-19 pandemic. It not only facilitates the reviewer to review such products under consistent standard and ensure safety, effectiveness and controllable quality of products in a short period, but also provides a guidance for the enterprise to develop the SARS-CoV-2 diagnostic products and prepare the application for approval. Although this Key Points was issued at the beginning of this pandemic, its principle is still applied for the situation. With further understanding of COVID-19 and SARS-CoV-2, new requirements may be needed for new situations and the Key Points may be revised accordingly.

Future perspective

Since the global outbreak of COVID-19, more than one wave of COVID-19 infections and different mutations of SARS-CoV-2 have been reported in different regions [24] and the spread of SARS-CoV-2 has not been controlled yet. The upcoming 2020–2021 flu season will bring more challenge to the prevention and control of SARS-CoV-2 infection. The accurate, specific and rapid SARS-CoV-2 tests and the combined detection of the potential coinfection are critically important for making effective measurements to respond and control the SARS-CoV-2 and other infections. SARS-CoV-2 antibody tests could provide useful information on the infection of the population, especially when SARS-CoV-2 RNA cannot be detected, but positive results of antibody tests can not be used to confirm the presence of the virus. SARS-CoV-2 antigen tests undergone stringent regulatory review is much less than nucleic acid and antibody tests. The rapid detection of SARS-CoV-2 antigen displays the potential merit of ease-usage, rapid turnaround time and wide range of application due to decentralized testing of patients, which may facilitate the diagnosis of patients with early symptoms and the guidance for antigen detection in the diagnosis of SARS-CoV-2 infection using rapid immunoassays has been provided by WHO [51]. Comprehensive understanding of the dynamic evolution of SARS-CoV-2 antibody and antigen will facilitate to the design of accurate, sensitive and specific SARS-CoV-2 tests, and help to choose a reasonable detection time and frequency. Normalized prevention and countermeasures, hierarchical management and immediate response will facilitate the early control of regional outbreak. Systematic strategy for prevention, control and treatment and effort from the whole world are needed to fight against this global COVID-19 pandemic.

Executive summary

Background

- Coronavirus disease-2019 (COVID-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has spread globally and more and more waves of COVID-19 infections have occurred in different regions. The coming winter and flu infection will make the control and treatment of COVID-19 pandemic more complicated. SARS-CoV-2 antigen/antibody tests, combined with SARS-CoV-2 nucleic acid tests with high efficiency and rapid accessibility will play an important role in prevention and treatment for this continuous pandemic. The purpose of this article is to provide more explanation on the requirement for the SARS-CoV-2 antigen/antibody tests to facilitate the development and accessibility of such tests.

Scope of application

- The Key Points are applicable for SARS-CoV-2 antigen/antibody assays, which are used to conduct *in vitro* qualitative detection of SARS-CoV-2 antigen/antibody in the samples of serum, plasma, whole blood, throat swab, bronchoalveolar lavage fluid, sputum or other respiratory secretions.

Analytical performance evaluation

- All analytical performance of candidate tests should be evaluated and verified using reagent manufactured under the control of quality management system.

Sample collection & processing

- Disease course, clinical symptoms and medication treatment should be considered.

Limit of detection

- Evaluation of limit of detection and verification of limit of detection claim should be carried out with enough replicates and different samples using suitable analysis method.

Inclusivity

- At least ten different real clinical samples of COVID-19 patients from different sources with temporal and regional characteristics should be verified.

Analytical specificity

- Cross reactivity should be analyzed for other human coronaviruses, influenza virus and other common respiratory viruses or coinfections of other viruses with SARS-CoV-2. The interference of suitable endogenous/exogenous substances should be evaluated.

IgM antibody class specificity

- IgM damage study should be carried out to verify the specification of IgM test.

Precision

- Multiple factors and reasonable evaluation period should be considered for precision evaluation. Negative samples, positive samples near the cutoff and moderate/strong positive clinical samples should be analyzed.

Hook effect study

- The concentration of targets without hook effect should be clarified.

Verification of reference materials

- Three batches of products should be used to test the reference materials.

Instrument used for SARS-CoV-2 tests

- Instruments used by proposed tests should be evaluated and clarified in production instruction for use.

Specimen types

- Applicable specimen types should be verified using suitable method.

Determination of cut-off value

- Different geographic regions, different infection stages, physiological status and sample types should be considered for cut-off value determination. Receiver operating characteristic curves are recommended.

Research data on main raw materials, production process & reaction system

- It is required to provide the research data on the selection, source, preparation and quality standard of main raw materials, such as antigen, antibody, quality control materials and reference materials.

Evaluation of stability

- Reagent stability and sample stability should be evaluated. Three lots of reagents are needed for shelf-life stability. The stability of treated samples by different method (preservation solution or heating) should also be studied.

Clinical studies

- Antibody tests should be evaluated by comparing the results of the available clinical reference standard for COVID-19. And the clinical performance of SARS-CoV-2 antigen tests should be compared with the results of SARS-CoV-2 nucleic acid tests.

Determination of clinical reference standard

- The positive detection rate of tested kits to that of the COVID-19 nucleic acid test can be used to fully evaluate the clinical performance of antibody tests.

Study population & inclusion criteria
- Subjects or samples enrolled in the clinical trial should represent the proposed target population of the test.

Sample types for clinical trial
- Statistical analysis for each type of sample is recommended and the requirements for the number of cases should be met respectively. For serum and plasma, homologous sample comparison can be conducted.

Sample size for clinical trial
- It is suggested to estimate the minimum sample size using a reasonable statistical model according to the preset clinical evaluation measures, such as clinical sensitivity, clinical specificity and relevant statistical parameters.

Statistical analysis of clinical trial results
- For the SARS-CoV-2 antibody tests, stratification analysis should be conducted for the patients in different stages of disease course. The clinical significance of separate statistics for IgG and IgM and joint detection of IgG and IgM should be evaluated.

Formal requirements for clinical evidences
- Ethical review opinions of all institutions, clinical trial protocol, clinical trial report and summary report on clinical trial.

Conclusion
- The Key Points for the SARS-CoV-2 test was issued to facilitate the availability of SARS-CoV-2 antigen and antibodies tests to response the COVID-19 public health emergency effectively.

Disclaimer

The official Key Points of Technical Review for the Registration of SARS-CoV-2 Antigen/Antibody Tests is the Chinese version issued by CMDE, and the expressions in this article are personal opinions of the authors.

Financial & competing interests disclosure

This work was supported by a grant from Ministry of Science and Technology, COVID-19 Pandemic Emergency Project 'Development of SARS-CoV-2 Tests' (No. 2020YFC0841200). The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

References

1. Zhou P, Yang XL, Wang XG *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 579(7798), 270–273 (2020).
2. Li H, Liu SM, Yu XH, Tang SL, Tang CK. Coronavirus disease 2019 (COVID-19): current status and future perspectives. *Int. J. Antimicrob. Agents* 55(5), 105951 (2020).
3. Amawi H, Abu Deiab GI, Aa AA, Dua K, Tambuwala MM. COVID-19 pandemic: an overview of epidemiology, pathogenesis, diagnostics and potential vaccines and therapeutics. *Ther. Deliv.* 11(4), 245–268 (2020).
4. Garrafa E, Brugnoni D, Barbaro M *et al.* Laboratory considerations amidst the coronavirus disease 2019 outbreak: the Spedali Civili in Brescia experience. *Bioanalysis* 12(17),1223- 1230 (2020).
5. WHO. WHO Director-General's statement on IHR Emergency Committee on Novel Coronavirus (2020). www.who.int/dg/speeches/detail/who-director-general-s-statement-on-ihr-emergency-committee-on-novel-coronavirus-(2019-ncov)
6. WHO. WHO Director-General's opening remarks at the media briefing on COVID-19 (2020). www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020
7. WHO. WHO coronavirus disease (COVID-19) dashboard (2020). https://covid19.who.int/
8. Gavriatopoulou M, Ntanasis-Stathopoulos I, Korompoki E *et al.* Emerging treatment strategies for COVID-19 infection. *Clin. Exp. Med.* doi:10.1007/s10238-020-00671-y (2020) (Epub ahead of print).
9. Zhu N, Zhang D, Wang W *et al.* A novel coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.* 382(8), 727–733 (2020).
10. Hoffmann M, Kleine-Weber H, Schroeder S *et al.* SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* 181(2), 271–280.e278 (2020).
11. Khailany RA, Safdar M, Ozaslan M. Genomic characterization of a novel SARS-CoV-2. *Gene Rep.* 19, 100682 (2020).
12. He J, Hu P, Gao Y *et al.* Comparison and application of different immunoassay methods for the detection of SARS-CoV-2. *J. Med. Virol.* 92(11), 2777–2784 (2020).
13. National Health Commission, National Administration of Traditional Chinese Medicine. Guidelines for the Diagnosis and Treatment of Coronavirus Disease 2019 (8th Trial Version). *Chinese J. Viral Dis.* 10(5), 321-328 (2020).

14. NMPA. Key Points of Technical Review for the Registration of SARS-CoV-2 Antigen/Antibody Tests (2020). www.cmde.org.cn/CL0189/20512.html
15. ChinaDaily. Provisions for *In-Vitro* Diagnostic Reagent Registration (2014). http://subsites.chinadaily.com.cn/nmpa/2019-07/25/c_390598.htm
16. Yang HS, Racine-Brzostek SE, Lee WT *et al.* SARS-CoV-2 antibody characterization in emergency department, hospitalized and convalescent patients by two semi-quantitative immunoassays. *Int. J. Clin. Chem.* 509, 117–125 (2020).
17. China CDC. Technical Guidelines for Laboratory Detection of SARS-CoV-2-Infected Pneumonia (2020). www.chinacdc.cn/jkzt/crb/zl/szkb_11803/jszl_11815/202003/W020200309540843062947.pdf
18. CLSI. Evaluation of detection capability for clinical laboratory measurement procedures; approved guideline – second edition. *CLSI Document EP17-A2* 32(8), 1–68 (2012).
19. Guan Q, Sadykov M, Mfarrej S *et al.* A genetic barcode of SARS-CoV-2 for monitoring global distribution of different clades during the COVID-19 pandemic. *Int. J. Infect. Dis.* 100, 216–223 (2020).
20. Becerra-Flores M, Cardozo T. SARS-CoV-2 viral spike G614 mutation exhibits higher case fatality rate. *Int. J. Clin. Practice.* 74(8), e13525 (2020).
21. Long SW, Olsen RJ, Christensen PA *et al.* Molecular architecture of early dissemination and massive second wave of the SARS-CoV-2 virus in a major metropolitan area. *mBio* 11(6), e02707–e02720 (2020).
22. Daniloski Z, Guo X, Sanjana NE. The D614G mutation in SARS-CoV-2 Spike increases transduction of multiple human cell types. *bioRxiv* doi:10.1101/2020.06.14.151357 (2020).
23. Korber B, Fischer WM, Gnanakaran S *et al.* Tracking changes in SARS-CoV-2 Spike: Evidence that D614G increases infectivity of the COVID-19 Virus. *Cell* 182(4), 812–827.e819 (2020).
24. Phan T. Genetic diversity and evolution of SARS-CoV-2. *Infect. Genet. Evol.* 81, 104260 (2020).
25. Yu HQ, Sun BQ, Fang ZF *et al.* Distinct features of SARS-CoV-2-specific IgA response in COVID-19 patients. *Eur. Respir. J.* 56(2), 2001526 (2020).
26. Ji T, Liu Z, Wang G *et al.* Detection of COVID-19: a review of the current literature and future perspectives. *Biosens. Bioelectron.* 166, 112455 (2020).
27. Sun B, Feng Y, Mo X *et al.* Kinetics of SARS-CoV-2 specific IgM and IgG responses in COVID-19 patients. *Emerg. Microbes Infect.* 9(1), 940–948 (2020).
28. Deeks JJ, Dinnes J, Takwoingi Y *et al.* Antibody tests for identification of current and past infection with SARS-CoV-2. *Cochrane Database Syst. Rev.* 6(6), Cd013652 (2020).
29. CLSI. Interference testing in clinical chemistry; approved guideline-second edition. CLSI document EP07-A2. *CLSI Document EP07-A2.* 25(27), 1–107 (2005).
30. Hicks J, Klumpp-Thomas C, Kalish H *et al.* Serologic cross-reactivity of SARS-CoV-2 with endemic and seasonal Betacoronaviruses. *medRxiv* doi:10.1101/2020.06.22.20137695 (2020).
31. Huang BR, Lin YL, Wan CK *et al.* Co-infection of influenza B virus and SARS-CoV-2: a case report from Taiwan. *J. Microbiol Immunol. Infect.* doi:10.1016/j.jmii.2020.06.011 (2020) (Epub ahead of print).
32. Wu X, Cai Y, Huang X *et al.* Co-infection with SARS-CoV-2 and influenza A virus in patient with pneumonia, China. *Emerg. Infect. Dis.* 26(6), 1324–1326 (2020).
33. Jiang S, Liu P, Xiong G *et al.* Coinfection of SARS-CoV-2 and multiple respiratory pathogens in children. *Clin. Chem. Lab. Med.* 58(7), 1160–1161 (2020).
34. Khorramdelazad H, Kazemi MH, Najafi A, Keykhaee M, Emameh RZ, Falak R. Immunopathological similarities between COVID-19 and influenza: investigating the consequences of co-infection. *Microb. Pathog.* doi:10.1016/j.micpath.2020.104554 104554 (2020) (Epub ahead of print).
35. Dong J. Key points of analysis performance of nucleic acid detection reagents for multiple respiratory pathogens. *Chinese Pharm. J.* 51(18), 1622–1625 (2016).
36. Devaux CA, Rolain JM, Raoult D. ACE2 receptor polymorphism: susceptibility to SARS-CoV-2, hypertension, multi-organ failure and COVID-19 disease outcome. *J. Microbiol. Immunol. Infect.* 53(3), 425–435 (2020).
37. Bal A, Agrawal R, Vaideeswar P, Arava S, Jain A. COVID-19: an up-to-date review – from morphology to pathogenesis. *Indian J. Pathol. Microbiol.* 63(3), 358–366 (2020).
38. Burrel S, Hausfater P, Dres M *et al.* Co-infection of SARS-CoV-2 with other respiratory viruses and performance of lower respiratory tract samples for the diagnosis of COVID-19. *Int. J. Infect. Dis.* 102, 10-13 (2020).
39. Sharov KS. SARS-CoV-2-related pneumonia cases in pneumonia picture in Russia in March-May 2020: secondary bacterial pneumonia and viral co-infections. *J. Global Health* 10(2), 020504 (2020).
40. Johnston R. The first 6months of HIV-SARS-CoV-2 coinfection: outcomes for 6947 individuals. *Current Opin. HIV AIDS* 16(1), 54–62 (2021).

41. Butt AA, Yan P. Rates and characteristics of SARS-CoV-2 infection in persons with hepatitis C virus infection. *Liver Int.* 41(1), 76–80 (2020).
42. Bowen R, Benavides R, Colón-Franco JM *et al.* Best practices in mitigating the risk of biotin interference with laboratory testing. *Clin. Biochem.* 74, 1–11 (2019).
43. Luong JHT, Vashist SK. Chemistry of biotin-streptavidin and the growing concern of an emerging biotin interference in clinical immunoassays. *ACS Omega* 5(1), 10–18 (2020).
44. CLSI. Evaluation of precision of quantitative measurement procedures; approved guideline. 3nd ed. *CLSI Document EP05-A3.* 34(13), 1–107 (2014).
45. CLSI. User verification of precision and estimation of bias; approved guideline-third edition. *CLSI Document EP15-A3.* 34(12), 1–93 (2014).
46. Vashist SK, Luong JHT.Chapter 4 – bioanalytical requirements and regulatory guidelines for immunoassays. In: *Handbook of Immunoassay Technologies*. Vashist SK, Luong JHT (Eds). Academic Press, 81–95 (2018).
47. NMPA. The Announcement of Registration Document Requirements and the Approved Certificate Format for *in vitro* Diagnostic Reagents (2020). https://www.cmde.org.cn/CL0097/3050.html
48. CLSI. Defining, establishing and verifying reference intervals in the clinical laboratory; approved guideline – third edition. *CLSI Document EP28-A3c* 28(30), 1–61 (2008).
49. CLSI. Evaluation of stability of *in vitro* diagnostic reagents; approved guideline. *CLSI Document EP25-A* 29(20), 1–41 (2009).
50. US FDA. Acceptance of Clinical Data to Support Medical Device Applications and Submissions: Frequently Asked Questions (2018). https://www.fda.gov/media/111346/download
51. WHO. Antigen-Detection in the Diagnosis of SARS-CoV-2 Infection using Rapid Immunoassays (2020). www.who.int/publications/i/item/antigen-detection-in-the-diagnosis-of-sars-cov-2infection-using-rapid-immunoassays

Review

For reprint orders, please contact: reprints@futuremedicine.com

Epigenomics

An immune epigenetic insight to COVID-19 infection

Bimal P Jit[1], Sahar Qazi[1], Rakesh Arya[1], Ankit Srivastava[2], Nimesh Gupta[3] & Ashok Sharma*,[1]

[1]Department of Biochemistry, All India Institute of Medical Sciences, New Delhi 110029, India
[2]Regional Institute of Ophthalmology, Institute of Medical Sciences, Banaras Hindu University, Varanasi 220115, India
[3]National Institute of Immunology, Aruna Asaf Ali Marg, New Delhi 110067, India
*Author for correspondence: Tel.: +91 112 659 3545; ashoksharma1202@gmail.com

Severe acute respiratory syndrome coronavirus-2 is a positive-sense RNA virus, a causal agent of ongoing COVID-19 pandemic. *ACE2R* methylation across three CpG sites (cg04013915, cg08559914, cg03536816) determines the host cell's entry. It regulates ACE2 expression by controlling the SIRT1 and KDM5B activity. Further, it regulates Type I and III IFN response by modulating H3K27me3 and H3K4me3 histone mark. SARS-CoV-2 protein with bromodomain and protein E mimics bromodomain histones and evades from host immune response. The 2′-O MTases mimics the host's cap1 structure and plays a vital role in immune evasion through Hsp90-mediated epigenetic process to hijack the infected cells. Although the current review highlighted the critical epigenetic events associated with SARS-CoV-2 immune evasion, the detailed mechanism is yet to be elucidated.

First draft submitted: 7 September 2020; Accepted for publication: 14 February 2021; Published online: 9 March 2021

Keywords: ACE2 • COVID-19 • DNA methylation • histone PTMs • host-virus interaction • immune evasion • immunoepigenetics • SARS-CoV-2

The advent of budding infectious agents like COVID-19 has resulted in a remarkable impact on the discrepancy of human resources, lifestyle, economy, and longevity in the modern era. COVID-19 infections have been associated with respiratory dysfunction that contributed to substantial alterations in the clinical manifestation and has become a significant public health concern [1]. The severe acute respiratory syndrome coronavirus (SARS-CoV-2), β-enveloped, segmented, positive-sense RNA belong to Orthocoronavirinae subfamily. Mostly attacks mammals and could cause potentially fatal respiratory dysfunctions. The previous form, regarded as 2019-nCoV, and COVID-19 have transmitted rapidly throughout China and received tremendous attention worldwide [2]. However, the SARS-CoV incidence was in 2002 followed by Middle East respiratory syndrome coronavirus (MERS-CoV) in 2012, followed by the subsequent entry of highly pathogenic large-scale epidemic coronavirus, COVID-19 as observed in the 21st century in the human population. Recent data as of 1 January 2020, indicates, a total of 82,579,768 cases globally, with mortality of 1,811,849 cases affecting 213 countries.

Out of the whole affected population, Bharat (India) accounts for 10,305,778 confirmed cases with a mortality of 149,218 cases [1]. Although the actual evidence regarding the virus' origin remains unclear, several studies indicate that the bat population in Wuhan city's seafood market in China is the potential reservoir [3]. Population genetic analysis revealed that the virus' L-type strain is more aggressive and contagious than the S strain [4]. The genome sequencing approach results indicate that COVID-19 shares 96.2% overall genomic similarity with Bat-CoV-RaTG13 suggests a typical ancestral relationship [5]. According to recent Lancet findings, Acute Respiratory Distress Syndrome is the most common pathological event of COVID-19 characterized by cytokine storm, resulting from the massive secretion of pro-inflammatory cytokines and chemokines during the infection by immune effectors cells [6]. Immunocompromised populations associated with several genetic and nongenetic diseases are highly vulnerable to infection and acts as the potential risk factor modulating the clinical severity. Evidence suggests persistent viral attacks associated with altering host epigenetic machinery, thus paving a way to evade and subvert the immune system for a successful infection strategy [7–9]. Therefore, it is highly crucial to understand the host-cell

Future Medicine

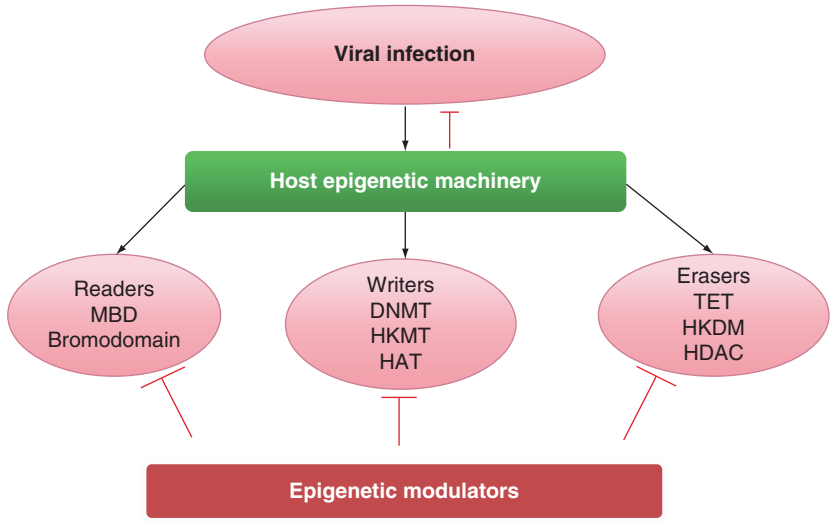


Figure 1. Epigenetic approach to viral infection.

chromatin dynamics to regulate the expression of specific genes adapted by COVID-19-host confrontation during its infection. Here, we provide a brief insight on possible epigenetic changes associated with coronavirus and other viruses to evade host immunity.

Host epigenetic architecture

Despite the Mendelian inheritance, which emphasizes the transmission of genes associated with a particular trait, discovery in several high throughput technologies contributed significant advancement to epigenetics. Epigenetics elucidates gene expression regulation and silencing events without any change in the DNA sequence. Although the changes instigated by this are stable, reversible modifications in the DNA sequences could induce a dramatic change in parent and progeny phenotypes. The consequences may persist for a lifelong period affecting cellular behavior [10]. Deciphering the fine-tuning mechanisms associated with an epigenetic signature could lead to identifying a potential target that will be suitable for identifying severity markers of a particular disease. Understanding the host-viral functional network is crucial to control the virus' pathogenesis and infectivity (Figure 1). More recently, epigenetic has become an emerging field in controlling the host innate and adaptive immune system induced by viral pathogenesis. Being originated by environmental stimuli, the field encompasses two major events methylation and acetylation, which play a crucial role in altering the chromatin packaging and position of regulatory elements like promoter or enhancers. These methylation and acetylation processes are being carried out by DNMTs and HDACs, which have varying expression levels in different cells under different conditions. In addition to this, other histone biochemical modifications includes phosphorylation, sumoylation, ubiquitination. Histone modification enzymes portray their role in coordination with other chromatin regulators like ATP-dependent chromatin remodeling complexes and the contrasting effects of the polycomb group and trithorax group genes [10]. Normal functional cells with transcriptionally active genes characterized by unmethylated promoter CpG islands, histone hyperacetylation, H3 lysine 4 (H3K4) di- and tri-methylation and H3K79 methylation. In addition to this, the transcriptionally repressed genes are discerned with promoter CpG island methylation, histone hypoacetylation and H3K9, H3K27 methylation [11].

Epigenetic basis of viral immune evasion

Many DNA and RNA viruses harbor epigenetic marks to maintain their persistency and latency. A previous study by Shamay *et al.* [12], revealed LANA expressed by Kaposi's sarcoma-associated virus (KSHV or HHV8) could interact with cellular DNMTs [13]. Recruitment of DNMT3A by LANA and initiation of DNA methylation in B cell lymphoma have been also reported [14]. Mechanically, LANA interacts with the gene promoter of TGF-β Type II receptor (TGF-βRII) inducing DNA methylation and causes transcriptional silencing [14]. It was also evident that promoters of several IFNγ-regulated genes with STAT1 binding sites are the potential target for KSHV associated with repression of IFNγ-induced gene activation [7]. Interaction of KSHV polyadenylated nuclear RNA with

host nuclear regulators, like PRC2 and IRF4, also shown to modulate the expression of several cytokines [14]. In particular, an increase in DNMT is crucial in regulating the viral latency by modifying the epigenetic reprogramming of infected cells. Furthermore, it was elucidated that trimethylation of H3K27 probably through PRC2 is associated with CpG methylation of the tumor suppressor gene, as shown in Epstein–Barr virus (EBV) latency in B cells [15]. As evident, Epstein–Barr-virus encodes a LMP1 and LMP2A, leading to activation of downstream signaling molecules like JNK, AP-1 and STAT3 causing upregulation of the host DNMT1 [16]. Besides, the role of oncoprotein E7 of human papillomavirus in modulating the DNMT1 and H3K27 methyltransferase EZH2 activity in cervical cancer has been reported earlier [8]. Furthermore, the modulation of host epigenome by simian vacuolating virus 40, adenoviruses, HIV and human T cell virus-1, paramecium bursaria chlorella virus, provides significant insight into the virus controls host epigenetic integrity and machinery for its efficient propagation by modulating gene repression [9].

Thus far, the most comprehensive evidence of the modification of chromatin at specific gene locus as observed in some viruses that play a crucial role in regulating host epigenetic and immune evasion. Interference in the Th-1 cell immune response stimulated by the respiratory syncytial virus infection, upregulates H3K4 demethylase KDM5B, results in Type 1 interferon and cytokine responses. As evidence elucidates the association of hepatitis B and C virus with hepatocellular carcinoma and their positive correlation with the host genome's aberrant DNA methylation [17–19]. DNA methylation-dependent repression of IL-4 receptor was observed during hepatitis B virus infection, mediated by the recruitment of DNMT3A by HBX [20]. Recent results indicate the HIV encoded tat protein regulates RNA polymerase II activity and fine-tune the expression of early response genes and makes the host cellular environment hospitable for the virus [21]. Furthermore, the association of HIV infection with upregulated DNA methylation event at CpG site of FOXOP3 locus leading to decreased production of TGF-β and increased production of interleukin IL-4, thus alters the Treg cell function [22].

Notably, epigenetic processes are incredibly crucial in controlling the viral DNA-based processes like genome replication, DNA damage response, a temporal cascade of transcription provides a footprint for understanding and targeting the viral pathogenesis and infection [23]. Certain viruses develop a common strategy where the early proteins of viruses target the cellular processes by binding the cellular regulatory factors. As observed in DNA tumor viruses; polyomaviruses, adenoviruses and papillomaviruses propagate their productive infection and cell transformation by binding to the p53 gene and regulating the cell cycle [24]. For instance, adenovirus e1a protein binds the host acetyltransferase, p300/CBP promoting viral replication by disrupting histone acetylation. Late-expressed protein VII binds host chromatin and sequesters the danger signaling HMGB proteins in chromatin resulting in downregulation of HMGB-induced immune response [25]. Further evidence showed that e1a blocks hBRe1 ubiquitin ligase complex formation required for ISGs in response to adenovirus infection [26].

Besides this, miRNA's role derived from RNA and DNA viruses plays a prominent role in immune evasion. miRNA derived from RNA viruses have been shown to play a significant role in controlling host immunity. miRNA from several human herpesviruses, polyomavirus JC and murine cytomegalovirus provide subtle evidence of miRNA's potential role in host immune evasion [27,28]. Ebola virus (EBOV/EBV) miRNA overexpression (EBOV-miR-1-5p) causes inhibition of importin-alpha5, leading to immune evasion from the host interferon response [29]. SV-40 encoded miRNA also regulates the viral gene expression and decreases the susceptibility to regulatory T cells [30]. Viral miRNA derived from the KSHV, herpes simplex virus 1, human cytomegalovirus shown to modulate the host innate and adaptive immune response [31]. Evasion from the adaptive response was observed from the miRNA derived from SV40 and murine polyomavirus by negatively regulating the early gene expression [32]. Similar results were observed in the KSHV and EBV virus [33]. Furthermore, miR-146a dampen the host IFN-β production-mediated by TRAF6, thus increasing the replication and infectivity ability of dengue virus [34].

Whether in cell culture or mouse model, epigenetic drugs are currently the potential candidates that act as prominent antiviral drugs. More recently, the HDAC inhibitors like panobinostat (NCT01680094), vorinostat (NCT01319383) and romidepsin (NCT02092116, NCT01933594), VPA (NCT00289952) are used in combination with antiretroviral therapy. However, this attribute is limited by the incomplete latency reversal or insufficient clearance of latency-reactivated cells, which further seeks immune enhancement treatments [35,36]. Interestingly, some HDAC inhibitors like arginine butyrate and ganciclovir could induce lytic phase gene expression that acts as a potent sensitizer observed in EBV-associated lymphoma [36,37]. DNMT inhibitor azacytidine has been observed to reverse the dense CpG methylation and induces gene re-expression in patients with EBV tumors [38]. The previous study has explained the potential role of DNMT2 in retrotransposon silencing, and the results revealed overexpression of DNMT2 altered several genes associated with viral infection [39].

The cross-talk between host and viral infection could be exemplified by the potential miRNA specific to the viral infection that could be used for the post-transcriptional regulation of target gene expression. Upadhya *et al.*, 2014 highlighted the association of CpG methylation as a genomic signature for large DNA viruses infecting invertebrates [40]. Further study also revealed the promoter hypermethylation as possible indicators of human papillomavirus-infected patients with head and neck squamous carcinomas [41]. Significant correlation of EBV oncoprotein EBNA3C and its positive association with transcriptional activation of autophagy genes by recruiting active epigenetic marks like H3K4me1, H3K4me3, H3K9ac and H3K27ac has been elucidated previously [42]. Furthermore, the prospective regulatory role of hsa-miR-374b-5p miRNA on Type 1 interferon expression in Japanese encephalitis virus-infected microglial cells was also reported earlier [43–45]. Immune metabolic findings show that viral protein ICP0 is an attenuator of toll-like receptor (TLR) signaling inhibiting innate response to herpes simplex virus [46].

Epigenetic approach of coronavirus immune evasion

Epigenetic alterations of *ACE2R* determines SARS-CoV entry into the host cell

Maintaining a latency stage inside the host, mimicking the host immune system requires manipulation of chromatin and heterochromatin assembly by viruses. It has become clear that the coronavirus' spike protein facilitates its entry into the target cells mostly by the surface unit s1 of S protein upon S protein priming. The molecular affinity between the ACE2 receptor, mostly expressed on the Type–II lung epithelial cells, is crucial in the viral entry [47–50]. Spike protein and receptor affinity is a crucial determinant of tissue tropism, which plays a vital role in the disease's etiopathogenesis. Therefore, it is crucial to understand the epigenetic signature of the *ACE2* gene for controlling the initial step of entry and fusion of the virus.

The genome-wide DNA methylation array and chip methylation pipelines study indicates the varied degree of DNA methylation of the *ACE2* gene in different tissue subtypes. The lowest *ACE2* gene methylation across three CpG sites (cg04013915, cg08559914, cg03536816) was predominant in lung epithelial cells compared with other tissues [51]. A subsequent study shows that *ACE2* gene hypo-methylations are mostly confined to the females compared with the male, suggesting angiotensin II metabolism and its association with hormonal differences or genetic differences in chromosome dosage [52]. Moreover, transcriptomic analysis shows a lack of possible association of ACE2 with race, age and gender. However, Asian smoker population exhibits higher ACE2 than the nonsmokers suggesting an epigenetic impact on ACE2 activity in the respiratory system [53–55]. However, the validation of the approach necessitates proteomic data. Previous evidence also highlighted hypo-methylation mediated overexpression of ACE2 and its association with the onset of severity in the patients of systemic lupus erythematosus, an autoimmune disease upon infection of SARS-CoV-2 with peripheral blood T cells [56]. Inconsistent with this, further study shows the prospective role of TNF-α in the regulation of ACE gene transcription and pathological complexity in endothelial cells. The results indicate that TNF-α enhances DNA methylation in the ACE promoter by decreasing the activity of DNMTs, DNMT3a and DNMT3b and TET1 [57]. Transcriptomic and system biology approach revealed the significant association of higher expression of ACE2 with RAB1A, HAT1, HDAC2 and KDM5B in patients with other comorbidities like hypertension, diabetes and chronic obstructive lung disease [58]. Also, the role of NAD^+ dependent histone deacetylase, SIRT1 in the induction of ACE2 activity by stimulating the ACE2 promoter was reported during energy stress indicates SIRT1 could be a target for epigenetic drugs in context to COVID-19 infection [54]. Pathway enrichment analysis also revealed the potential role of KDM5B in regulating the expression of several genes associated with ACE2 possibly by acetylation and methylation epigenetic marks like H3K4me1 and H3K4me3, as well as H3K27ac [56]. Although the literature is scanty about the COVID-19 ACE2 epigenetic pattern, the molecular mechanism regulating the ACE2 activity cannot be undermined so far as the current pandemics and pathogenesis are concerned. A recent study has observed that SARS-CoV-2 cross-reactive $CD4^+$ T cells are still present in 70% convalescent patients even after 5 months of the infection [57].

Coronavirus induces epigenetic modulation of immune cells, alters antigen presentation & interferon response

It is apparent that the generation of useful anti-inflammatory cytokines and chemokines inhibit viral replication and enhances antigen presentation [57–59]. The ISG response plays a prominent role in controlling the viral infection for efficient immune function. Type 1 IFN induces a cascade of signaling events which causes transcription of several ISG [60,61]. Although evidence is scanty from SARS-CoV-2 in this perspective, studies from influenza and

other RNA respiratory viruses provide significant insight. SARS-CoV-2 and MERS-CoV viruses are found to delay ISG expression significantly. Transcriptomics and proteomics findings in Calu3 cells revealed diverse virus-specific ISG expression signatures. SARS-CoV-2 infection of Calu3 cells revealed a strong induction of ISG effectors, but the response was significantly delayed with peak expression at 48 h post-infection. In 2012, the newly emerged MERS-CoV showed a dramatic delayed ISG released with effects visible at 18 h post-infection, for the decreased expression of potential ISG subsets [62].

Further evidence supports the notion that downregulation of ISGs is not due to any impairment in the signaling cascade, but by histone modification like methylation and acetylation induced by a pathogen. Upon viral infection like SARS-CoV-2, the host produces Type I and III IFN, which induces histone modulation complex, which renders removal of repressive histone mark (H3K27me3) inducing activating mark like H3K4me3. This conversion of inactive chromatin to active chromatin allows the binding of several transcription factors like STAT1 and IRF7 thus inducing the ISG expression [63,64]. However, incorporating repressive histone modifications like H3K27me3 and removing active mark H3K4me3 could impose a more condensed state of chromatin which prevents binding of transcription factor and thus reduces ISG expression. Additionally, inhibition or downregulation of an H3K79 methylase, Dot1L enzyme associated with decreased antiviral response and facilitates viral replication, suggests its crucial role in antiviral response [65].

The previous study's prospective interleukin role in regulating the epigenetic signature has been provided [66]. The author revealed that treatment with IL-1 induces STAT-6 activity; a major transcription factor for IL-4 mediated signaling that binds to the H3K27 demethylase Jmjd3 promoter. An elevated level of Jmjd3 decreases H3K27 dimethylation and trimethylation (H3K27me2/3), leading to transcriptional activation of the M2 marker gene. The significant association of H3K9me2 as a suppressor of IFN inducible antiviral response has been elucidated by Fang *et al.* (2012) [67,68]. However, inactivation of lysine methyltransferase G9a an inducer of H3K9me2 resulted in high IFN production, suggesting that methyltransferase could be an ideal therapeutic epigenetic target challenge the viral invasion. Interestingly, results of Menachery *et al.* (2018) revealed a possible association of H3K27me3 global methylation with downregulation of ISG and DNA methylation of antigen presentation gene upon MERS-CoV and H5N1VN1203 [66]. The possible association of TNF-α and H3K4me3 in the induction of trained innate immunity in monocytes and DC1 antigen presentation and TH1/TH17 immunity upon infection has been described previously [67,69]. Mechanistically, it was observed that an increase in either TNF-α or IFN-γ is sufficient to induce the MLL1 activity which stimulates H3K4 methylation and is required for DC stabilization [67]. Further evidence by Liu *et al.* supported the potential role of influenza virus NS1 in modulating the JAK-STAT signaling by facilitating the export of DNMT3b from the nucleus to the cytoplasm and its subsequent degradation by K48-linked polyubiquitination. Promoter demethylation leads to the expression of specific JAK-STAT signaling suppressors such as SOCS1, SOCS3, PIAS1 and induces inhibition of interferon signaling in an autocrine or paracrine manner [68].

Acetylation and deacetylation of histones play a significant role in macrophage activation and survival. HDAC mediates the regulation of non-histone proteins involved in mechanisms crucial for cellular functions like DNA repair, replication, P53 signaling, HIF-1α, STAT3 or p65. HDAC macrophages display a pro-inflammatory function by producing pro-inflammatory cytokines like TNF-α, MCP-1, IL-1α, IL-1β and IFN-γ and are a potential target for HDAC inhibitors [69]. The role of HDAC2 in modulating the NF-κB activity plays a significant role in the immune evasion strategy of SARS-CoV-2. Nuclear localization of HDAC2 makes it convenient to inhibit the NF-κB activity, thus altering the monocyte and macrophage function [70]. Overexpression and knockdown study with HDAC5, a Type II HDAC indicates the subsequent activation of TNF-α and MCP-1 in macrophages contributing to the inflammatory response. Respiratory dysfunctions associated with COVID-19 infection are exacerbated by inflammatory response stimulated by monocyte and macrophages. Monocytes play a crucial role in innate immune response, which migrated to the affected areas and differentiated into macrophage and plays a defensive role by producing pro-inflammatory cytokines like IL-1β, TNF-α, IL-6 and chemokines that facilitate migration [71]. HDAC5 localized in the nucleus and participated in the inflammatory response. However, drugs that could activate the HDAC2 and facilitate nuclear export of HDAC5 could be a novel strategy to control the inflammatory response associated with COVID-19. Current evidence indicates methylxanthine theophylline, macrolide antibiotics, the tricyclic antidepressant nortriptyline, the volatile anaesthetics isoflurane, the phenolic compounds gallic acid and curcumin as well as the plant bioactive molecule andrographolide have the potential to induce the HDAC2 activity by inhibiting PI3K-δ signaling (Figure 2) [72–75].

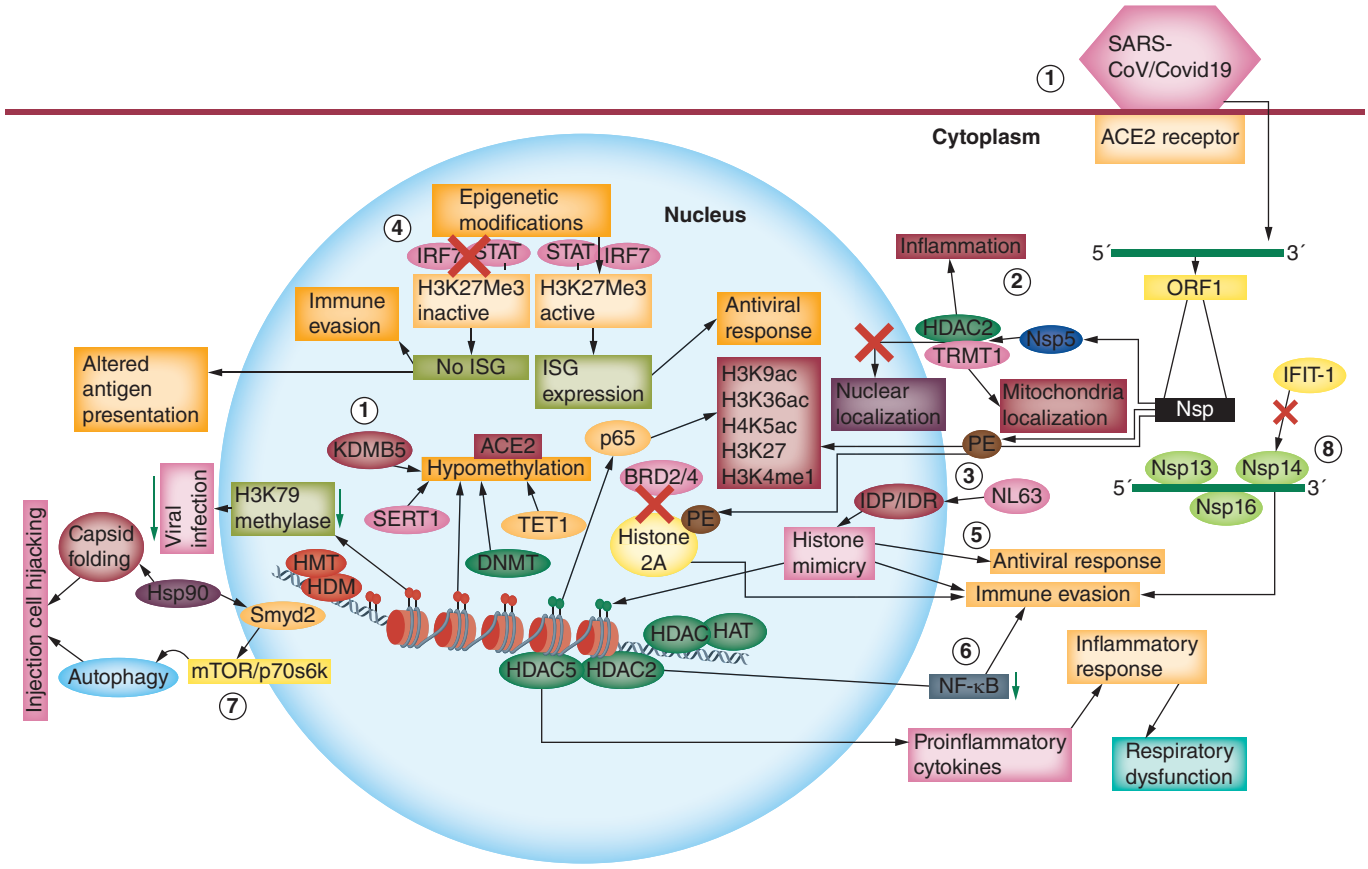


Figure 2. Interaction of coronavirus and modulation of host epigenetic machinery. (1) *ACE2R* hypomethylation across CpG sites in lung epithelial cells/other cell types is the first step for entering the coronavirus. DNA methylation in the ACE promoter regulated is under the strict regulation by DNMTs, DNMT3a and DNMT3b, TET1, MAX, KDM5B, HDAC2. **(2)** Virus fuses taking advantage of host cell membrane endocytic pathway. Upon the SARS-CoV entry, the positive strand RNA releases to the cytoplasm and serves as a template for negative strand, which subsequently produces more positive stand and mRNA leading to viral protein synthesis. The Nsp5 protein interacts with both TRMT1 and HDAC2 and prevents their nuclear entry. **(3)** The protein E of SARS-CoV-2 shares similarity with histone 2A and prevents its interaction with BRD2 leading to evasion from host defense. **(4)** Infection with SARS-CoV-2, host, produces Type I and III IFN by removing the repressive histone mark (H3K27me3) and activating H3K4me3, induces chromatin in an open state and permits the binding of transcription factors like STAT1, IRF7 leading ISG expression. However, the incorporation of repressive histone marks like H3K27me3 could impose a more condensed state of chromatin which prevents binding of transcription factor and thus reduces ISG expression inducing immune evasion. **(5)** IDP or IDR derived from SARS-CoV-2 or NL63 coronavirus interacts with host SLiMs proteins that share certain residues of histone proteins play an important role in modulating gene expression histone mimicry. SLiMs are also a potent target for binding of epigenetic readers. This contributes to viral pathogenicity and suppresses antiviral response. **(6)** Nuclear localization of HDAC5 a Type II HDAC induces pro-inflammatory cytokines and controls inflammatory response. In response to SARS-CoV infection, HDAC2 activity increases, leading to NF-κB activity inhibition, altering monocyte and macrophage function and modulating host cell response. **(7)** SARS-CoV-2 uses Hsp90, which hijacks infected cells and causes autophagy by regulating SMYD2 and mTOR/p70S6K signaling. **(8)** Ifit1, an interferon-induced RNA binding protein, binds to viral RNA lacking 2′-O-methylation at their 5′ end and preventing RNA translation. ORF1 of viral RNA translated into polyproteins consists of Nsp13, Nsp14, Nsp16. Nsp16 possess methyltransferase (2′-O-MT) activity, nsp13 function as 5′ triphosphatase and helicase activity, nsp14 exhibits N7MT activity. Viral Nsp16 with 2′-O MTases and mimics c.
IDR: Intrinsically disordered region; SARS-CoV: Severe acute respiratory syndrome coronavirus.

SARS-CoV-2 protein 3b as an epigenetic modulator

Characterization of viral protein components and their host interacting partners could provide novel insight to understand the molecular basis of immunomodulation strategy. Recent evidence indicates the potential role of SARS-CoV-2 protein 3b an accessory protein component, has been observed to interacts with the host protein machinery like RUNX1b. Given that RUNX1b stimulates transcription of genes involved in definitive hematopoiesis and T cell differentiation cytokines and chemokines including IL-2, IL-3, GM-CSF, MIP-1α, CSFR, etc. [74,75]. In response to SARS-CoV-2 infection protein, 3b stimulates phosphorylation of RUNX1b in an ERK-dependent

manner and activates IL-2 promoter [74]. This high degree of molecular linkage of RUNX1 is thought to be mediated by HDAC recruitment and T cell cytotoxic response [76]. The evidence supports the notion that efficient T cell function or exaggerated immune response might be epigenetically controlled.

Histone mimicry as a basis for modulation of gene expression & immune evasion

Coronavirus is enveloped single-stranded positive RNA viruses with a genome size ranging from 26.2 to 31.7 kb. The large capped polyadenylated genome constitutes a set of conserved genes arranged in a particular order: 5′ORF1a-ORF1b-S-ORF3-E-M-N-3′. Among these ORF; ORF1a/b exhibits two-thirds of the genome and produces an mRNA (mRNA1) which encodes different structural and functional protein components. The structural proteins include: S, E, M and N. The N proteins of CoV-2 consist of three conserved domains: N terminal domain (NTD), C terminal domain (CTD) separated by intrinsically disordered regions (IDR) or RNA binding domains. It was observed that the NTD preferentially binds the 3′ end of viral RNA by electrostatic interaction which stabilizes RNA structure and acts as a chaperone and helps in replication. The CTD plays a vital role in the dimer-dimer association, protein interaction and stress response [77]. The NTD and CTD are separated by IDR or IDP which lack 3D shape in their native conformation. These domains play a crucial role in DNA, RNA and protein binding and enhance the RNA binding activity of NTD and CTD. It was also pointed out that these IDPs play a predominant role in viral adaptation, evasion from the host immune system. Regulation of viral protein synthesis managing the economical use of genetic material via alternative splicing, overlapping genes and antisense transcription [78]. It is surprising to note that the human coronavirus NL63 has 7.3% of these disordered residues [79]. The IDP or IDR are also observed in the mammalian proteome characterized by a predominant hydrophilic amino acid with a low abundance of bulky hydrophobic amino acids. Post-translational modification and sharing of the particular motif of eukaryotic proteins like SLiMs make them ideal for regulating the host defensive strategy. Short linear motifs commonly known as SLiMs, also the components of eukaryotic linear motifs. SLiMs are observed to be a part of histone proteins and acts as a binding target for readers. As the binding affinity and specificity of SLiMs reside in 2–5 residues, it makes it easy to mimic the host SLiMs. It was observed that the presence of histone H3 like sequence within the C-terminal portion of NS1 of the H3N2 subtype of influenza virus [80].

The role of NS1 protein in suppressing Type I IFN response during infection has been described previously. NS1 protein consists of a sequence of 226-ARSK-229 which resembles the first four amino acids (1-ARTK-4) of histone H3. The tails of influenza may contain PDZ ligand (PL) motifs whereas, SUMOylation sequence in the NS1 tail of H1N1 strain [81]. The PL motif observed in ESEV and EPEV within the NS1 tails of avian-derived influenza strain are significantly associated with pathogenicity and could suppress the antiviral response. The PL motif can attenuate the apoptosis of infected cells and increases the viral load [82].

Furthermore, the histone mimicking ability of NS1 of influenza virus and its modulation of host epigenome has been well explained [80]. The study revealed carboxy-terminus of the H3N2 protein NS1 and tail of histone H3 shares homologs sequence. The NS1 protein interacts with the human PAF1 transcription elongation complex (hPAF1C) and decreases the PAF1-mediated antiviral response in a host. Furthermore, the binding affinity of histone H3 tail and H3N2 NS1 tail to PAF1 has been explained previously, contributing to RNA elongation co-transcriptional process [80]. Interaction between CHD1 and PAF1 in the regulation of transcriptional elongation has been elucidated previously. WDR5 a core subunit of the human MLL and SET1 histone H3K4 methyltransferase complex and is highly important for global H3K4 methylation and *HOX* gene activation in human cells. Human CHD1 preferentially binds to H3K4me3 a hallmark of actively transcribed chromatin. Recently, it was reported the CHD1 and WDR5 are a potential target for NS1 protein of influenza A H3N2 subtype possess a histone H3K4 like sequence at its CTD and adapts antiviral response [83].

The bromodomain (BRD) is a conserved structural module of chromatin-associated proteins, and histone acetyltransferases play a dynamic role in regulating chromatin-based gene transcription. BRD specifically binds to the acetylated histones and regulates gene expression [84]. Recent affinity purification-based mass spectrometry results indicate the interaction of E protein of SARS-CoV-2 with BRD-containing proteins BRD2, BRD4 disrupting the activity of BRD histones binding by mimicking histone structure. The N terminus of histone 2A shares local sequence similarity over an alpha helix about 15 residues some of which are in the transmembrane segment of protein E, which suggests mimicking protein E's action on histone which disrupts its interaction with BRD2, thus evading from host immune defense. Interaction analysis revealed the affinity of Nsp5 (C145A) with TRMT1 and wild-type Nsp5 with TRMT1 and HDAC2. Taking both wild-type and catalytic dead constructs (C145A) of Nsp5 of SARS-CoV-2 indicates, wild-type Nsp5 exhibits high confidence interaction with epigenetic regulator HDAC2

predicted a cleavage site between the nuclear localization sequence and HDAC domain and suggested an inhibitory effect of Nsp5 on HDAC2 transport into the nucleus. Furthermore, Nsp5 removes zinc finger and the nuclear localization signal of TRMT1 mediates mitochondrial localization (Figure 2) [83]. Fascinating results obtained by the same study from chemoinformatics data indicate valproic acid and the preclinical candidate apicidin possess HDAC2 inhibitory activity with an affinity of 5 and 120 nM, clinical compounds like ABBV-744 and CPI-0610 on BRD2 and BRD4 with affinities of 2 and 39 nM, respectively.

Coronavirus modulates innate epigenetic signaling

The potential role of the pathogen-related receptor (PRR) and pathogen-associated molecular pattern, TLR, JAK-STAT, NF-κB signaling in orchestration against the viral pathogenesis was previously known. TLR2 mRNA was observed to increase in patients with SARS infection. Previous evidence indicates an increase in NF-κB signaling in response to the TLR2 signaling in monocytes in *in vitro* conditions [84]. Spike proteins of SARS-CoV are cleaved by cathepsin L, factor Xa and trypsin which cleaves the spike protein into S1 and S2 and permits viral entry into the cytoplasm [85]. In epithelial and fibroblast cells displaying ACE2 receptor induces IL-8 production in response to S protein via the AP-1 pathway [86]. Subsequent evidence indicates crucial roles of specific accessory proteins encoded by MERS-CoV antagonize NF-κB signaling to evade the host defense. Transport of NF-κB into the nucleus is stimulated by the destruction of IkBs, where NF-κB induces activation of cytokine genes. Upon viral infection activation of TLR and retinoic acids inducible gene like receptors (RLRs) and nucleic acid sensors recognize the pathogen-associated molecular pattern. The role of *RIG-1* and MDA-5 via MAVS in priming IkB ubiquitination by recruiting TRAF and TAK1 which induces NF-κB activation [87]. However, previous evidence indicates the viral origin of different proteins and proteases causes inactivation of these adaptor molecules leading to silencing of the NF-κB paving the way for immune evasion [88]. A recent infection experiment of coronavirus with 229E cells indicates that p65 chromatin recruitment is highly crucial in NF-κB and its target gene induction. P65 occupying regions are enhancers elements, promoter-transcription start site regions characterized by increased acetylation of H3 and H4 histones which are stimulated by the activity of transcription factor induced after CoV-2 infection and leads to expression of genes associated with antiviral response [89]. Activation of NF-κB allows the synthesis of the A20 protein required for efficient viral replication [89]. Historically, recruitment of inducible transcription factors to the enhancer region has H3K4me1 and H3K27ac and increases the acetylation of H3K36 and H4K5 in the chromatin structure near the promoter region determines virus-induced host cell response in a nuclear transcription-dependent mechanism. The possible interaction of viral protein components with different DNMTs have been well established [61,90]. The potent role of IL-32 in exerting its antiviral response is confined to its ability by inducing the pro-inflammatory cytokines and differentiation of monocytes into macrophages.

Demethylation in the CREB binding site increases the binding of CREB to the promoter followed by *IL-32* transcriptional activation in influenza A virus-infected cells. Influenza virus activates *IL-32* expression by activating NF-κB and CREB with site-specific demethylation of CRE in the *IL-32* promoter region. Inactivation of DNMT1 and DNMAT3b causes hypomethylation of *IL-32* promoter in response to influenza virus infection, indicating a host's protective mechanism in preventing viral replication [91]. Similar findings were reported by Fang *et al.* (2012), where the downregulation of DNMT3a and DNMT3b, but not that of DNMT1, involves a COX2 dependent IFN-λ1 production by increase NF-κB signaling. The result of this study indicates increased activity of miRNA (mir29) in A549 cells and PBMC derived from the influenza patients induces PKA-mediated phosphorylation of CREB1 and inhibition of DNMTs activity and contributes to COX2 and PGE2 expression [62]. Mechanistic evidence indicates infection with influenza virus induces an increase in the expression of the host methyltransferase Setdb2, which mediates trimethylation of histone H3 Lys9 (H3K9) at the *Cxcl1* promoter and make the host susceptible to superinfection with *Streptococcus pneumonia* [92]. The epigenetic alterations associated with CoV-2 infection are summarized in Table 1.

The epigenetic mechanism controls viral RNA replication

Extracellular vesicles trigger epigenetic reprogramming in the host cell. Virally induced vesicle formation can trigger the multiplicity of infection. CoV-2 Nsp-3, -4 and -6 play a fundamental role in the rearrangement of the host cell membrane and required for the establishment of replication-transcription complexes, called replication organelles, which are nothing but the double-membrane vesicles are characteristic features of all RNA viruses including CoV-2 for making stable infection in the host [101,102]. The marginalization of host cell chromatin, a proliferation of nuclear membrane are the prime events during herpes virus infection. However, such molecular events in CoV-2

Table 1. Epigenetic alterations associated with COVID-19 infection.

Viral components	Host machinery	Epigenetic change	Response	Ref.
Protein 3b	RUNX1b	Recruitment	T cell function cytokine response	[93,94]
Covid19	NF-κB TNF-α MCP-1	HDAC2, HDAC5	Inflammation	[95,96]
Nsp16 Nsp13 Nsp14	2′-O-MT activity 5′-triphosphatase and helicase activity N7MT activity	Methylation and mimic of Cap1 structure	Immune evasion from interferon response by protecting of viral RNA from 5′-3′ exonuclease activity	[97,98]
CoV infection	Hsp90 induced mTOR pathway	SMYD2 (Lysine methyltransferase)	Autophagy	[99,100]
	p65 induced NF-κB activity	H3H4 acetylation	Antiviral response	
	BRD2	Mimics H2A histone	Host immune evasion	[83]
Nsp5	TRMT1 HDAC2	TRMT1 HDAC2	Prevents HDAC2 transport to nucleus induces mitochondrial localization of HDAC2	[83]
NL63/ IDP or IDR	Host SLiM protein	Mimics Histone H3	Immune evasion by surpassing Interferon response	[79,80]
MERS-CoV	TNF-α Interferon	H3K27me3	Antagonizes antigen presentation immune evasion	[70,75]
Spike protein	ACE2R	Methylation at CpG site	Viral entry and pathogenesis	[61]

IDR: Intrinsically disordered region; MERS-CoV: Middle East respiratory syndrome coronavirus.

vesicle are not reported. Vesicle fusion with other cellular components and chromatin disassembly in a GTPase Ran mediated manner was also described in the previous report [103].

5′ OMT as a molecular target plays a crucial role in evasion from the immune system

The exploitation of host synthetic machinery or encoding own proteins to counteract innate immune response is crucial for establishing a successful infection strategy. Methylation of transcriptome involving different RNAs like tRNA, mRNA, rRNA and other noncoding RNAs is crucial for regulating gene expression. Mechanistically eukaryotic mRNA is capped at 2′-O positions (or Nm where N can be any nucleotide) of the 5′-guanosine cap by methyltransferases (MTases) to distinguish endogenous self-capped RNA from exogenous nonself RNA encoded by pathogen lacking Nm. This mechanism has been well understood at the molecular level where IFIT1 interferon-induced RNA binding protein mediates this effect by preferentially binding to viral RNAs lacking 2′-O-methylation at their 5′ end and preventing RNA translation. Seminal findings suggest a wide variety of pathogens like flaviviruses, coronavirus, Japanese encephalitis virus, mouse virus, dengue virus, SARS-CoV virus and vaccinia virus adopts such strategy for the propagation of successful replication and evasion from interferon response [104]. IFIT1 has been observed to have a higher affinity for RNA lacking 2′-O methylation, IFIT1 can out-compete eIF4E or eIF4F for binding, remove cap 0 RNA from the actively translating pool [99]. Surprisingly, specific pathogens like coronaviruses encode their own viral 2′-O MTases and mimic the host's cap1 structure by different distinct mechanisms rendered them a potential target for drug development [105]. *ORF1* translated into polyproteins (ppla and ppl1ab) which undergoes co- and post-translational modifications and forms 16 nonstructural proteins nsp 1–16. Bioinformatics study revealed SAM-dependent RNA 2′-O-MT activity of nsp16, where nsp13 function as 5′triphosphatase and helicase activity whereas nsp14 exhibits N7MT activity. This 5′ cap protects the viral RNA from degradation by altering 5′-3′ exonuclease activity and induces its translation by stimulating the preinitiation complex formation. Although the exact role of SARS-CoV-2 2′-O-MT is unknown, it plays an integral part in a viral replicase-transcriptase complex on its interactions with other viral proteins implicated in the formation of a 3′ terminal protein complex. Recent findings from computational analysis study revealed that dolutegravir and bictegravir are potential drug candidates to target SARS-CoV-2 2′-O-MT activity [100].

Coronavirus uses Hsp90-mediated epigenetic process to hijack the infected cells

Historically, histone modification plays a significant role in the epigenetic silencing of endogenous retroviruses. It was observed that ZFP809, a member of the KRAB-ZFP family, induces the silencing of endogenous retroviruses in a sequence-specific manner via recruitment of heterochromatin-inducing complexes. Epigenetic mark involving Histone 3 Lys9 trimethylation (H3K9me3) in host-associated with tightly inactive repressed chromatin [103].

ZFP809 binding to proline tRNA primer-binding site used by some retroviruses to prime reverse transcription. ZFP809 recruits the KRAB domain binding corepressor KAP1 (TRIM28, TIF1b), which induces silencing via recruitment of histone deacetylases, HP1 and the histone methyltransferase SETDB1 (ESET, KMT1E) [106]. Hsp90 chaperon activity plays a vital role in TRIM28/KAP1-mediated epigenetic silencing of endogenous retroviral elements [97]. Hsp90 in influenza virus by binding to the PB2 subunit enhances the RNA polymerase activity. In the poliovirus, Hsp90 is required for proper folding of the capsid protein. Hsp90 enables viruses to hijack the infected cells through the process of autophagy in targeting the mTOR pathway by inducing the mTOR/p70S6K pathway [98]. It has also been proposed complex interaction between estrogen, Hsp90 and lysine methyltransferases (SMYD2) plays an essential role in autophagy [107]. A similar mechanism might have to go on in SARS-CoV-2-infected patients. Remarkably, a drug repositioning study suggests geldanamycin and its derivatives as the potential candidate to target Hsp90 during COVID-19 infection [95].

Conclusion

A viral threat to humanity represents a foremost public health concern and accounts for a significant cause of morbidity and mortality for decades. Currently, COVID-19 has raised many scientific and clinical questions. A growing body of evidence indicates an evolutionary arms race and molecular cross-talk between virus and host epigenetic landscape plays a pivotal role in encouraging the altered immune response. Another question is whether SARS-CoV-2 is evolutionarily adapted to subvert the host replication, transcription and proteome program by modifying the epigenetic machinery, thus evading host immune response. The evidence summarized and discussed here clearly demonstrates that RNA viruses like SARS-CoV-2 are equipped with a molecular entity to evade from host innate immunity by altering epigenetic architecture. Although SARS-CoV-2 modifies the host phenotype by potent epigenetic types of machinery like ACE2R methylation, interfering with host replication machinery, altering antigen presentation and interferon response, innate epigenetic signaling and histone mimicry, however, the mechanistic basis of altered chromatin dynamics and viral antagonism of SARS-CoV-2 is still not thoroughly investigated. Therefore, future studies should explore and generate a more comprehensive map of important epigenetic events in the histone induced by SARS-CoV-2. Integration of the epigenetically technological approach will allow understanding of the epigenetic landscape of immune response induces by SARS-CoV-2 and prediction of a pharmacological target for therapeutic efficacy.

Future perspective

Evidence from both MERS-CoV and SARS-CoV-2 infection indicate a significant decrease in the peripheral lymphocytes resulting in lymphopenia, thrombocytopenia, pneumonia and high CRP contributes respiratory dysfunction. Existing evidence also demonstrated the possible association of pulmonary injury with various immune cells like neutrophil and macrophage infiltration and concomitant increase in the concentration of interleukins, G-CSF, TNF-α, interferon-gamma IP10, MCP1 and MIP1A, D-dimers, alanine transaminases, lactate dehydrogenase, creatine kinase, amylase and ferritin [96,108].

Although the onset of upper acute respiratory symptoms like fever, fatigue, cough and digestive symptoms like nausea, vomiting, abdominal pain, diarrhoea is typical, however patients with severe phenotypes are characterized by pneumonia, sepsis, encephalopathy, acute respiratory distress syndrome, acute kidney injury, cardiomyopathy, pulmonary dysfunction, endothelial injury and coagulation abnormality. Considerable challenges with antimalarial drugs, antiviral agents, antiparasitic agents, immune modulators, anti-inflammatory agents, anti-angiotensin receptors and traditional herbal remedies have shown to be beneficial [109]. Moreover, till recently, no suitable US FDA-approved agent has demonstrated therapeutic efficacy in randomized clinical trials. Although cytokine storm seems crucial in modulating clinical severity, however, due to insufficient data and lack of strong evidence until recently, it is unknown whether cytokine storm is the major confounding factor modulating pathological complexity. In contrast, further emerging evidence supports the notion that the phenotype may be manifested due to endothelial dysfunction and systemic inflammatory response [93,94].

In the battle of nature and host immunity, the role of epigenetic is of tremendous importance elucidating a scientific heft to the current scenario. Although it was previously experienced that mechanistic underpinnings governing the epigenetic process are highly deterministic in affecting psychological traits like intelligence, personality, sexuality, now it is proved that epigenetics plays a significant role in modulating the infectivity and host-pathogen interaction. A host defense mechanism by manipulating epigenetic process is significant for survival and successful infection strategy by viruses. The evolution of multiple pathways played by the viruses' structural and accessory

proteins targeting the epigenetic readers, writers and erasers result in the variable of expression of proteins crucial for host defense. A viral infection is significantly associated with the exploitation of host cellular immune behavior and potent cellular signaling process. Although the immune strategy system has developed to tackle the viral attack from the perspective of the evolutionary point of view, its virus also owns the battle by developing a predominate strategy to counteract the immune system exaggerated by its adaptation and conducive environment. Epidemiological evidence indicates that viral infection and release are associated with several novel strategies induced by a virus sufficient to evade the host's innate and adaptive immunity. Intracellular nature, low fidelity of RNA polymerase, limited genome size, resistance to physical injury make the virus ideal to manipulate the host immune system renders it to evade the immune control mechanisms [110]. Although the epigenetic mechanism controlling gene behavior is well established, understanding the primary mechanism responsible for fine-tune the epigenetic target by coronaviruses and its association with the clinical severity and disease outcome is need to be understood in detail. COVID-19 has become the primary cause of morbidity and mortality worldwide. Owing to its large extent of infectivity, lack of sufficient knowledge in understanding the pathogenesis, devoid of ideal potent therapeutic agent, vaccine and population with asymptomatic phenotype rendered it as an emerging issue in the current scenario.

It is highly crucial to understand the potential cross-talk between viruses and host; for example mechanism of viral latency, reactivation of viral latency, virus localization to different chromatin sites, recognition of potent epigenetic

Executive summary

Overview

- Currently, COVID-19 is the significant public health concern and exhibits a prominent cause of morbidity and mortality worldwide.
- Viruses have improved several potent mechanisms to efficiently propagate inside the host by fine-tuning the host epigenetic program leading to evasion from innate and adaptive immunity.
- Both DNA and RNA viruses regulate the epigenetic players like HAT, DNMTs, HDAC culminating into activation and repression of the specific genetic mark in the promoter of innate signaling cascade and thus subvert from host defense.

The epigenetic perspective of SARS-CoV-2

***ACE2R* methylation determines SARS-CoV-2 entry into the host cell**

- *ACE2* gene methylation across three CpG sites (cg04013915, cg08559914, cg03536816) in lung epithelial cells is of paramount importance SARS-CoV-2 infection.
- SARS-CoV-2 may regulate ACE2 expression by controlling the SIRT1 and KDM5B activity.

SARS-CoV-2 epigenetically alters antigen presentation & Interferon response

- SARS-CoV-2 significantly delayed interferon-stimulated gene expression.
- SARS-CoV-2 may regulate Type I and III IFN response by modulating H3K27me3 and H3K4me3 histone mark.
- Monocyte and macrophage-mediated inflammatory response associated with COVID-19 are induced by acetylation and deacetylation of histones.

Histone mimicry as a basis for modulation of gene expression & immune evasion

- SARS-CoV-2 protein with bromodomain and protein E mimics bromodomain histones and evades from host immune response.
- However, there is a paucity of literature pertinent to this field.

Coronavirus modulates innate epigenetic signaling

- SARS-CoV-2 may regulate NF-κB signaling by p65 chromatin recruitment.
- However, very few works have been conducted in this field.
- SARS-CoV-2, 2′-O MTases mimics the host's cap1 structure and plays a vital role in immune evasion.
- Coronavirus uses Hsp90-mediated epigenetic process to hijack the infected cells.

Epigenetic tools to study the evasion strategy of SARS-CoV-2

- Several technological approaches have contributed to the understanding of nuclear architecture and chromatin network substantially.
- Advancement in the single cell-based chromatin analysis approach; microscopy-based single-molecule real-time imaging for chromatin dynamics, epigenome microarray, chromatin immune precipitation with next-generation sequencing, will be highly beneficial to understand the dynamism of host chromatin and virus interaction.

Conclusion & future perspective

- Although the current review highlighted some of the critical epigenetic events associated with SARS-CoV-2 immune evasion, the detailed mechanism is yet to be elucidated.
- Future studies should focus on the novel paradigm involved between host and SARS-CoV-2 and the epigenetic basis of immune evasion to predict pharmacological targets and therapeutic interventions.

signature as a result of association with different stages of infection and chromatin basis of immune regulatory genes. In the past few years, significant advancement in several technological approaches has substantially contributed to understanding nuclear architecture and chromatin network. Advancement in the single cell-based chromatin analysis approach; microscopy-based single-molecule real-time imaging for chromatin dynamics, epigenome microarray, chromatin immune precipitation with next-generation sequencing, and FISH approach will be of an immense requirement to understand the dynamism of host chromatin and virus interaction. It keeps the present alarming situation in mind; there are no epigenetic drugs with clinical efficacy to control the COVID-19 infection. Therefore, the future study should focus on the detailed understanding and elucidation of the novel paradigm involved between COVID-19 and host epigenetic to develop epigenetic agents with potential efficacy.

Financial & competing interests disclosure

This work was supported by grant Number: IPA/2020/000077 from the Science and Engineering Research Board, DST, India. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

References

1. WHO. Coronavirus disease (COVID-19) Weekly Epidemiological Update and Weekly Operational Update. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports
2. Lu R, Zhao X, Li J, Niu P *et al.* Genomic characterization and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* 395(10224), 565–574 (2020).
3. Zhou P, Yang XL, Wang XG *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 579(7798), 270–273 (2020).
4. Tang X, Wu C, Li X, Song Y *et al.* On the origin and continuing evolution of SARS-CoV-2. *Nat. Sci. Rev.* 7(6), 1012–1023 (2020).
5. Huang C, Wang Y, Li X *et al.* Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 395(10223), 497–506 (2020).
6. Wu JT, Leung K, Leung GM. Now casting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *Lancet* 395(10225), 689–697 (2020).
7. Lu F, Tsai K, Chen HS *et al.* Identification of host-chromosome binding sites and candidate gene targets for Kaposi's sarcoma-associated herpesvirus LANA. *J. Vir.* 86(10), 5752–5762 (2012).
8. Obata Y, Furusawa Y, Hase K. Epigenetic modifications of the immune system in health and disease. *Imm. Cell Biol.* 93(3), 226–232 (2015).
9. Sato A, Saito Y, Sugiyama K *et al.* Suppressive effect of the histone deacetylase inhibitor suberoylanilide hydroxamic acid (SAHA) on hepatitis C virus replication. *J. Cell. Biochem.* 114(9), 1987–1996 (2013).
10. Tzika E, Dreker T, Imhof A. Epigenetics and metabolism in health and disease. *Front. Gen.* 9, 361 (2018).
11. Marcos-Villar L, Díaz-Colunga J, Sandoval J *et al.* Epigenetic control of influenza virus: role of H3K79 methylation in interferon-induced antiviral response. *Sci. Rep.* 8(1), 1–13 (2018).
12. Shamay M, Krithivas A, Zhang J *et al.* Recruitment of the *de novo* DNA methyltransferase Dnmt3a by Kaposi's sarcoma-associated herpesvirus LANA. *Proc. Natl Acad. Sci.* 103(39), 14554–14559 (2006).
13. Di Bartolo DL, Cannon M, Liu YF *et al.* KSHV LANA inhibits TGF-β signalling through epigenetic silencing of the TGF-β Type II receptor. *Blood* 111(9), 4731–4740 (2008).
14. Rossetto CC, Pari GS. Kaposi's sarcoma-associated herpesvirus non-coding polyadenylated nuclear RNA interacts with virus-and host cell-encoded proteins and suppresses expression of genes involved in immune modulation. *J. Virol.* 85(24), 13290–13297 (2011).
15. Paschos K, Smith P, Anderton E *et al.* Epstein-barr virus latency in B cells leads to epigenetic repression and CpG methylation of the tumour suppressor gene Bim. *PLoS Pathog.* 5(6), e1000492 (2009).
16. Hino R, Uozaki H, Murakami N *et al.* Activation of DNA methyltransferase 1 by EBV latent membrane protein 2A leads to promoter hypermethylation of PTEN gene in gastric carcinoma. *Can. Res.* 69(7), 2766–2774 (2009).
17. Okamoto Y, Shinjo K, Shimizu Y *et al.* Hepatitis virus infection affects DNA methylation in mice with humanized livers. *Gastroenterol.* 146(2), 562–572 (2014).
18. Jain S, Chang TT, Chen S *et al.* Comprehensive DNA methylation analysis of hepatitis B virus genome in infected liver tissues. *Sci.Rep.* 5(1), 1–11 (2015).
19. Schinzari V, Barnaba V, Piconese S. Chronic hepatitis B virus and hepatitis C virus infections and cancer: synergy between viral and host factors. *Clin. Microbiol. Infect.* 21(11), 969–974 (2015).

20. Zheng DL, Zhang L, Cheng N *et al.* Epigenetic modification induced by hepatitis B virus X protein via interaction with *de novo* DNA methyltransferase DNMT3A. *J. Hepat.* 50(2), 377–387 (2009).
21. Reeder JE, Kwak YT, McNamara RP *et al.* HIV Tat controls RNA Polymerase II and the epigenetic landscape to transcriptionally reprogram target immune cells. *Elife* 4, e08955 (2015).
22. Pion M, Jaramillo-Ruiz D, Martínez A *et al.* HIV infection of human regulatory T cells downregulates Foxp3 expression by increasing DNMT3b levels and DNA methylation in the FOXP3 gene. *AIDS* 27(13), 2019–2029 (2013).
23. Nitzsche A, Paulus C, Nevels M. Temporal dynamics of cytomegalovirus chromatin assembly in productively infected human cells. *J. Virol.* 82(22), 11167–11180 (2008).
24. Levine AJ. The common mechanisms of transformation by the small DNA tumor viruses: the inactivation of tumor suppressor gene products: p53. *Virology* 384(2), 285–293 (2009).
25. Lynch KL, Gooding LR, Garnett-Benson C *et al.* Epigenetics and the dynamics of chromatin during adenovirus infections. *FEBS Lett.* 593(24), 3551–3570 (2019).
26. Fonseca GJ, Thillainadesan G, Yousef AF *et al.* Adenovirus evasion of interferon-mediated innate immunity by direct antagonism of a cellular histone post-translational modification. *Cell Host Microbe* 11(6), 597–606 (2012).
27. Bauman Y, Nachmani D, Vitenshtein A *et al.* An identical miRNA of the human JC and BK polyomaviruses targets the stress-induced ligand ULBP3 to escape immune elimination. *Cell Host Microbe* 9(2), 93–102 (2011).
28. Liu Y, Sun J, Zhang H *et al.* Ebola virus encodes a miR-155 analog to regulate importin-α5 expression. *Cell. Mol. Life Sci.* 73(19), 3733–3744 (2016).
29. Sullivan CS, Grundhoff AT, Tevethia S *et al.* SV40-encoded microRNAs regulate viral gene expression and reduce susceptibility to cytotoxic T cells. *Nature* 435(7042), 682–686 (2005).
30. Naqvi AR, Shango J, Seal A *et al.* Viral miRNAs alter host cell miRNA profiles and modulate innate immune responses. *Front. Immunol.* 6(9), 433 (2018).
31. Seo GJ, Chen CJ, Sullivan CS. Merkel cell polyomavirus encodes a microRNA with the ability to autoregulate viral gene expression. *Virology* 383(2), 183–187 (2009).
32. Nachmani D, Stern-Ginossar N, Sarid R *et al.* Diverse herpesvirus microRNAs target the stress-induced immune ligand MICB to escape recognition by natural killer cells. *Cell Host Microbe* 5(4), 376–385 (2009).
33. Wu S, He L, Li Y *et al.* miR-146a facilitates replication of dengue virus by dampening interferon induction by targeting TRAF6. *J. Infect.* 67(4), 329–341 (2013).
34. Shirakawa K, Chavez L, Hakre S *et al.* Reactivation of latent HIV by histone deacetylase inhibitors. *Trends Microbiol.* 21(6), 277–285 (2013).
35. Perrine SP, Hermine O, Small T *et al.* A Phase I/II trial of arginine butyrate and ganciclovir in patients with Epstein-Barr virus-associated lymphoid malignancies. *Blood* 109(6), 2571–2578 (2007).
36. Ghosh SK, Perrine SP, Williams RM *et al.* Histone deacetylase inhibitors are potent inducers of gene expression in latent EBV and sensitize lymphoma cells to nucleoside antiviral agents. *Blood* 119(4), 1008–1017 (2012).
37. Chan AT, Tao Q, Robertson KD *et al.* Azacitidine induces demethylation of the Epstein-Barr virus genome in tumors. *J. Clin. Oncol.* 22(8), 1373–1381 (2004).
38. Thiagarajan D, Dev RR, Khosla S. The DNA methyltransferase Dnmt2 participates in RNA processing during cellular stress. *Epigenetics* 6(1), 103–113 (2011).
39. Samantarrai D, Dash S, Chhetri B *et al.* Mallick. Genomic and epigenomic cross-talks in the regulatory landscape of miRNAs in breast cancer. *Mol. Can. Res.* 11(4), 315–328 (2013).
40. Upadhyay M, Sharma N, Vivekanandan P. Systematic CpT (ApG) depletion and CpG excess are unique genomic signatures of large DNA viruses infecting invertebrates. *PLoS ONE* 9(11), e111793 (2014).
41. Choudhury JH, Ghosh SK. Promoter hypermethylation profiling identifies subtypes of head and neck cancer with distinct viral, environmental, genetic and survival characteristics. *PLoS ONE* 10(6), e0129808 (2015).
42. Bazot Q, Paschos K, Skalska L et al.. Epstein-Barr virus proteins EBNA3A and EBNA3C together induce expression of the oncogenic microRNA cluster miR-221/miR-222 and ablate expression of its target p57KIP2.. *PLoS Pathog.* 11(7), e1005031 (2015).
43. Manocha GD, Mishra R,Sharma N et al.. Regulatory role of TRIM21 in the type-I interferon pathway in Japanese encephalitis virus-infected human microglial cells.. *J. Neuroinflammation* 11(1),1-12 (2014).
44. Bhattacharjee S, Bose P, Patel K *et al.* Transcriptional and epigenetic modulation of autophagy promotes EBV oncoprotein EBNA3C induced B-cell survival. *Cell Death Dis.* 9(6), 1–18 (2018).
45. Rastogi M, Singh SK. Modulation of type-I interferon response by hsa-miR-374b-5p during Japanese encephalitis virus infection in human microglial cells. *Front. Cell. Infect. Microbiol.* 9, 291 (2019).
46. Chattopadhyay D, Mukhopadhyay A, Ojha D *et al.* Immuno-metabolic changes in herpes virus infection. *Cytokine* 112, 52–62 (2018).

47. Hoffmann M, Kleine-Weber H, Schroeder S *et al.* SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* 181(2), 271–280 (2020).
48. Yan T, Xiao R, Lin G. Angiotensin-converting enzyme 2 in severe acute respiratory syndrome coronavirus and SARS-CoV-2: a double-edged sword? *FASEB J.* 34(5), 6017–6026 (2020).
49. Mudersbach T, Siuda D, Kohlstedt K *et al.* Epigenetic control of the angiotensin-converting enzyme in endothelial cells during inflammation. *PLoS ONE* 14(5), e0216218 (2019).
50. Zhang XH, Zeng ZP, Li HZ *et al.* Expression of renin-angiotensin-aldosterone system in human adipose tissues. Zhongguo Yi Xue ke Xue yuan xue bao. *Acta Academiae Medicinae Sinicae* 28(6), 766 (2006).
51. Corley MJ, Ndhlovu LC. DNA methylation analysis of the COVID-19 host cell receptor, angiotensin I converting enzyme 2 gene (ACE2) in the respiratory system reveal age and gender differences. 10.20944/preprints202003.0295.v1 (2020).
52. Joubert BR, Felix JF, Yousefi P *et al.* DNA methylation in newborns and maternal smoking in pregnancy: genome-wide consortium meta-analysis. *Am. J. Hum. Genet.* 98(4), 680–696 (2016).
53. Stueve TR, Li WQ, Shi J *et al.* Epigenome-wide analysis of DNA methylation in lung tissue shows concordance with blood studies and identifies tobacco smoke-inducible enhancers. *Hum. Mol. Genet.* 26(15), 3014–3027 (2017).
54. Zhao Y, Zhao Z, Wang Y *et al.* Single-cell RNA expression profiling of ACE2, the receptor of SARS-CoV-2. *BioRxiv* 202(5), 756–759 (2020).
55. Sawalha AH, Zhao M, Coit P *et al.* Epigenetic dysregulation of ACE2 and interferon-regulated genes might suggest increased COVID-19 susceptibility and severity in lupus patients. *Clin. Immunol.* 6(215), 108410 (2020).
56. Clarke NE, Belyaev ND, Lambert DW *et al.* Epigenetic regulation of angiotensin-converting enzyme 2 (ACE2) by SIRT1 under conditions of cell energy stress. *Clin. Sci.* 126(7), 507–516 (2014).
57. Ansari A, Arya R, Sachan S *et al.* Immune memory in mild COVID-19 patients and unexposed donors from India reveals persistent T cell responses after SARS-CoV-2 infection. *medRxiv* 10.1101/2020.11.16.20232967 (2020).
58. Pinto BG, Oliveira AE, Singh Y *et al.* ACE2 expression is increased in the lungs of patients with comorbidities associated with severe COVID-19. *MedRxiv* 222(4), 556–563 (2020).
59. Channappanavar R, Zhao J, Perlman S. T cell-mediated immune response to respiratory coronaviruses. *Immunol. Res.* 59(1–3), 118–128 (2014).
60. Yip MS, Leung NH, Cheung CY *et al.* Antibody-dependent infection of human macrophages by severe acute respiratory syndrome coronavirus. *Virol. J.* 11(1), 1–11 (2014).
61. Menachery VD, Eisfeld AJ, Schäfer A *et al.* Pathogenic influenza viruses and coronaviruses utilize similar and contrasting approaches to control interferon-stimulated gene responses. *MBio* 5(3), e01174–14 (2014).
62. Fang J, Hao Q, Liu L *et al.* Epigenetic changes mediated by microRNA miR29 activate cyclooxygenase 2 and lambda-1 interferon production during viral infection. *J. Virol.* 86(2), 1010–1020 (2012).
63. Marcos-Villar L, Nieto A. The DOT1L inhibitor pinometostat decreases the host-response against infections: considerations about its use in human therapy. *Sci. Rep.* 9, 9 (2019).
64. Ishii M, Wen H, Corsa CA *et al.* Epigenetic regulation of the alternatively activated macrophage phenotype. *Blood* 114(15), 3244–3254 (2009).
65. Kramer JM, Kochinke K, Oortveld MA *et al.* Epigenetic regulation of learning and memory by *Drosophila* EHMT/G9a. *PLoS Biol.* 9(1), e1000569 (2011).
66. Menachery VD, Schäfer A, Burnum-Johnson KE *et al.* MERS-CoV and H5N1 influenza virus antagonize antigen presentation by altering the epigenetic landscape. *Proc. Natl Acad. Sci.* 115(5), E1012–E1021 (2018).
67. Quintin J, Saeed S, Martens JH *et al.* Candida albicans infection affords protection against reinfection via functional reprogramming of monocytes. *Cell Host Microbe* 12(2), 223–232 (2012).
68. Liu S, Liu L, Xu G *et al.* Epigenetic modification is regulated by the interaction of influenza A virus non-structural protein 1 with the *de novo* DNA methyltransferase DNMT3B and subsequent transport to the cytoplasm for K48-linked polyubiquitination. *J. Virol.* 93(7), e01587–18 (2019).
69. Eastman AJ, Xu J, Bermik J, Potchen N *et al.* Epigenetic stabilization of DC and DC precursor classical activation by TNFα contributes to protective T cell polarization. *Sci. Advan.* 5(12), eaaw9051 (2019).
70. Weiss U, Möller M, Husseini SA *et al.* Inhibition of HDAC Enzymes Contributes to Differential Expression of Pro-Inflammatory Proteins in the TLR-4 Signaling Cascade. *Int. J. Mol. Sci.* 21(23), 8943 (2020).
71. Leoni F, Zaliani A, Bertolini G *et al.* The antitumor histone deacetylase inhibitor suberoylanilide hydroxamic acid exhibits anti-inflammatory properties via suppression of cytokines. *Proc. Natl Acad. Sci.* 99(5), 2995–3000 (2002).
72. Ito K, Hanazawa T, Tomita K *et al.* Oxidative stress reduces histone deacetylase 2 activity and enhances IL-8 gene expression: role of tyrosine nitration. *Biochem. Biophy. Res. Comm.* 315(1), 240–245 (2004).

73. Drexhage RC, Knijff EM, Padmos RC *et al.* The mononuclear phagocyte system and its cytokine inflammatory networks in schizophrenia and bipolar disorder. *Expert Rev. Neurotherapeut.* 10(1), 59–76 (2010).
74. Bristow CA, Shore P. Transcriptional regulation of the human MIP-1α promoter by RUNX1 and MOZ. *Nucleic Acids Res.* 31(11), 2735–2744 (2003).
75. Law HK, Cheung CY, Ng HY *et al.* Chemokine up-regulation in sars-coronavirus–infected, monocyte-derived human dendritic cells. *Blood* 106(7), 2366–2374 (2005).
76. Varshney B, Agnihotram S, Tan YJ *et al.* SARS coronavirus 3b accessory protein modulates transcriptional activity of RUNX1b. *PLoS ONE* 7(1), e29542 (2012).
77. Taniuchi I, Sunshine MJ, Festenstein R *et al.* Evidence for distinct CD4 silencer functions at different stages of thymocyte differentiation. *Mol. Cell* 10(5), 1083–1096 (2002).
78. Xue B, Dunker AK, Uversky VN. Orderly order in protein intrinsic disorder distribution: disorder in 3500 proteomes from viruses and the three domains of life. *J. Biomol. Str. Dynam.* 30(2), 137–149 (2012).
79. Pyrc K, Berkhout B, Van Der Hoek L. The novel human coronaviruses NL63 and HKU1. *J. Virol.* 81(7), 3051–3057 (2007).
80. Marazzi I, Ho JS, Kim J *et al.* Suppression of the antiviral response by an influenza histone mimic. *Nature* 483(7390), 428–433 (2012).
81. Bao Y, Bolotov P, Dernovoy D *et al.* The influenza virus resource at the National Center for Biotechnology Information. *J. Virol.* 82(2), 596–601
82. Golebiewski L, Liu H, Javier RT *et al.* The avian influenza virus NS1 ESEV PDZ binding motif associates with Dlg1 and Scribble to disrupt cellular tight junctions. *J. Virol.* 85(20), 10639–10648 (2011).
83. Gordon D, Jang GM, Bouhaddou M *et al.* A SARS-CoV-2-Human Protein-Protein Interaction Map Reveals Drug Targets and Potential Drug Repurposing. *Nature* 583(7816), 459–468 (2020).
84. Josling GA, Selvarajah SA, Petter M *et al.* The role of bromodomain proteins in regulating gene expression. *Genes* 3(2), 320–343 (2012).
85. Dosch SF, Mahajan SD, Collins AR. SARS coronavirus spike protein-induced innate immune response occurs via activation of the NF-κB pathway in human monocyte macrophages *in vitro*. *Virus Res.* 142(1–2), 19–27 (2009).
86. Chang YJ, Liu CY, Chiang BL *et al.* Induction of IL-8 release in lung cells via activator protein-1 by recombinant baculovirus displaying severe acute respiratory syndrome-coronavirus spike proteins: identification of two functional regions. *J. Immunol.* 173(12), 7602–7614 (2004).
87. Lee JY, Bae S, Myoung J. Middle East respiratory syndrome coronavirus-encoded accessory proteins impair MDA5-and TBK1-mediated activation of NF-κB. *J. Microbiol. Biotechnol.* 29(8), 1316–1323 (2019).
88. Deng L, Zeng Q, Wang M *et al.* Suppression of NF-κB activity: a viral immune evasion mechanism. *Viruses* 10(8), 409 (2018).
89. Poppe M, Wittig S, Jurida L *et al.* NF-κB-dependent and-independent transcriptome and chromatin landscapes of human coronavirus 229E-infected cells. *PLoS Pathog.* 13(3), e1006286 (2017).
90. Zheng DL, Zhang L, Cheng N *et al.* Epigenetic modification induced by hepatitis B virus X protein via interaction with *de novo* DNA methyltransferase DNMT3A. *J. Hepat.* 50(2), 377–387 (2009).
91. Li W, Sun W, Liu L *et al.* IL-32: a host pro-inflammatory factor against influenza viral replication is up-regulated by aberrant epigenetic modifications during influenza A virus infection. *J. Immunol.* 185(9), 5056–5065 (2010).
92. Schliehe C, Flynn EK, Vilagos B *et al.* The methyltransferase Setdb2 mediates virus-induced susceptibility to bacterial superinfection. *Nature Immunol.* 16(1), 67–74 (2015).
93. Kox M, Waalders NJ, Kooistra EJ *et al.* Cytokine levels in critically ill patients with COVID-19 and other conditions. *JAMA* 324(15), 1565–1567 (2020).
94. Sinha P, Matthay MA, Calfee CS. Is a "cytokine storm" relevant to COVID-19? *JAMA Int. Med.* 180(9), 1152–1154 (2020).
95. Sultan I, Howard S, Tbakhi A. Drug repositioning suggests a role for the heat shock protein 90 inhibitor geldanamycin in treating COVID-19 infection. 10.21203/rs.3.rs-18714/v1 (2020).
96. Yuki K, Fujiogi M, Koutsogiannaki S. COVID-19 pathophysiology: a review. *Clin. Immunol.* 10.1016/j.clim.2020.108427 (2020).
97. Hummel B, Hansen EC, Yoveva A *et al.* The evolutionary capacitor HSP90 buffers the regulatory effects of mammalian endogenous retroviruses. *Nature Struct. Mol. Biol.* 24(3), 234 (2017).
98. Liu G, Zhong M, Guo C *et al.* Autophagy is involved in regulating influenza A virus RNA and protein synthesis associated with both modulation of Hsp90 induction and mTOR/p70S6K signalling pathway. *Int. J. Biochem. Cell Biol.* 72, 100–108 (2016).
99. Habjan M, Hubel P, Lacerda L *et al.* Sequestration by IFIT1 impairs translation of 2′ O-unmethylated capped RNA. *PLoS Pathog.* 9(10), e1003663 (2013).
100. Khan RJ, Jha RK, Amera GM *et al.* Targeting SARS-CoV-2: a systematic drug repurposing approach to identify promising inhibitors against 3C-like proteinase and 2′-O-ribose methyltransferase. *J. Biomol. Struct. Dynam.* 10.1080/07391102.2020.1753577 1–14 (2020).
101. Angelini MM, Akhlaghpour M, Neuman BW *et al.* Severe acute respiratory syndrome coronavirus non-structural proteins 3, 4, and 6 induce double-membrane vesicles. *MBio* 10.1128/mBio.00524-13. 4(4), (2013).

102. Doyle N, Hawes PC, Simpson J *et al.* The porcine Deltacoronavirus replication organelle comprises double-membrane vesicles and zippered endoplasmic reticulum with double-membrane spherules. *Viruses* 11(11), 1030 (2019).
103. Prunuske AJ, Ullman KS. The nuclear envelope: form and reformation. *Curr. Opin. Cell Biol.* 18(1), 108–116 (2006).
104. Miorin L, Maestre AM, Fernandez-Sesma A *et al.* Antagonism of Type I interferon by flaviviruses. *Biochemi. Biophy. Res. Comm.* 492(4), 587–596 (2017).
105. Li SH, Dong H, Li XF *et al.* Rational design of a flavivirus vaccine by abolishing viral RNA 2′-O methylation. *J. Virol.* 87(10), 5812–5819 (2013).
106. Wolf G, Yang P, Füchtbauer AC *et al.* The KRAB zinc finger protein ZFP809 is required to initiate epigenetic silencing of endogenous retroviruses. *Genes & development* 29(5), 538–554 (2015).
107. Obermann WM. A motif in HSP90 and P23 that links molecular chaperones to efficient estrogen receptor α methylation by the lysine methyltransferase SMYD2. *J. Biol. Chem.* 293(42), 16479–16487 (2018).
108. Zhou Y, Fu B, Zheng X *et al.* Pathogenic T-cells and inflammatory monocytes incite inflammatory storms in severe COVID-19 patients. *Nat. Sci. Rev.* 7(6), 998-1002(2020).
109. Hassan SA, Sheikh FN, Jamal S *et al.* Coronavirus (COVID-19): a review of clinical features, diagnosis, and treatment. *Cureus* 12(3), e7355 (2020).
110. Hilleman MR. Strategies and mechanisms for host and pathogen survival in acute and persistent viral infections. *Proc. Natl Acad. Sci.* 5(101), 14560–14566 (2004).

Research Article

Future VIROLOGY

Scouting the receptor-binding domain of SARS coronavirus 2: a comprehensive immunoinformatics inquisition

Zaira Rehman[1], Ammad Fahim*,[2] & Muhammad Faraz Bhatti**,[3]

[1]Department of Virology, National Institute of Health (NIH), Islamabad, Pakistan

[2]Department of Multidisciplinary Sciences, National University of Medical Sciences (NUMS), Rawalpindi, Pakistan

[3]Atta-ur-Rahman School of Applied Biosciences (ASAB),National University of Sciences & Technology (NUST), Sector H-12, Islamabad, Pakistan

*Author for correspondence: Tel.: +92 51 927 2896; ammad.fahim@numspak.edu.pk

**Author for correspondence: Tel.: +92 51 9085 6136; mfbhatti@asab.nust.edu.pk

Aim: December 2019 witnessed the emergence of a worldwide outbreak of a novel strain of coronavirus (CoV) termed SARS-CoV-2. Several preventive strategies are being developed, such as vaccines, to stop the spread of infection. **Materials & methods:** A comprehensive immunoinformatics approach was used to map conserved peptide sequences on the receptor binding domain of SARS-CoV-2 for their B-cell, T-helper & T-cytotoxic cell epitope profiles. **Results & conclusion:** The antigenic B-cell epitopes were LFRKSN and SYGFQPT. Among T-cell epitopes, CVADYSVLY and FTNVYADSF exhibited affinity for MHC class I, while YRLFRKSNL and VYAWNRKRI exhibited affinity for of MHC class II alleles. The overlapping epitope between B- and T-cells was YRLFRKSNL. The deployment of these epitopes in potential vaccine development against COVID-19 may help in slowing down the SARS-CoV-2 spread.

First draft submitted: 3 August 2020; Accepted for publication: 9 February 2021; Published online: 22 February 2021

Keywords: coronaviruses • COVID-19 • epitope mapping • human ACE2 receptor

The COVID-19 pandemic [1] with its widespread pathogenesis on human body reminds us of the SARS and Middle East respiratory syndrome (MERS) epidemics [2,3]. Before the emergence of SARS family-mediated epidemics, coronaviruses (CoVs) were not a notable pathogenic entity, although they have been cited in the literature since the 1960s [4]. Besides SARS-CoV-2, the SARS-CoV and MERS-CoV were other famous coronaviral strains known for crossing the species barrier and infecting humans, while 229E, NL63(alpha), OC43 and HKU1 (beta) strains of CoVs caused common cold in humans [5]. The SARS-CoV that surfaced during 2002–2003 in China, infected almost 8000 individuals and caused 774 causalities in 37 countries [6,7] while MERS-CoV emerged in 2012 and infected 2494 humans causing 858 mortalities [2,8,9]. The ongoing epidemic of SARS-CoV-2 has affected more than 13 million and caused 0.5 million deaths globally up to 15 July 2020 [10]. The detailed viral sequence of SARS-CoV-2 indicates a genetic similarity to its earlier family members SARS and MERS, harboring six open reading frames [11]. The genome of CoVs is generally 27–35 kb, stuffed inside a nucleocapsid protein envelope. There are three structural proteins associated with viral envelope: membrane, envelope and spike (S) proteins [12]. The S protein is essential for viral entry into host but also impels tissue tropism, host diversity as well as host immune responses [13]. The S protein constitutes S1 and S2 subunits where S1 harbors receptor-binding domain (RBD) while S2 aids in virus–host cell fusion. Interestingly, the S gene of CoVs has been noticeable for the most recurring recombination breakpoints in SARS-CoV [14]. In case of SARS CoV, the main entry receptor on host cell is ACE2 [15] while in case of MERS-CoV, the primary entry receptor is dipeptidyl peptidase 4 (also known as CD26) [16]. Comparatively, the genomes of SARS-CoV and MERS-CoV although harboring significant similarity, they potentially harbor variation, particularly from the antigenic response they might generate with reference to RBD [17].

Future Medicine

Receptor-binding protein has been previously targeted for vaccinable solutions [18,19]. However, from the evolutionary standpoint, coronaviral RBD is reported as a hypervariable region [20]. There are no approved treatments or vaccines available to tackle COVID-19 thus far; however, development has been fast tracked. Contextually, vaccinomics approach exploits immunogenomics of SARS-CoV-2 RBD that can lead to potential vaccine candidate identification. Immunoinformatics driven in-depth analysis can aid in identifying repertoire of viral antigenic epitopes that may be either linear or discontinuous and it also helps in identifying whether these epitopes are immunogenic or virulent [21]. A comprehensive immunoinformatics data mining of SARS-CoV-2 RBD can increase our understanding for its antigenic profile. Encouragingly enough, epitope prediction analyses were reported earlier for SARS and MERS strains [19,22] and more recently for SARS-CoV-2 [23–25].

Numerous studies have reported bats as the primary reservoirs of SARS and MERS viruses [14,26–29]; however, rodent origin has also been reported [13,30]. A recently reported strain derived from bats, the bat CoV (BatCoV RaTG13) shares more than 96% homology with SARS-CoV-2 and 93% with its S protein, rendering it to be a close relative of SARS-CoV-2 [11]. Anatomically, conservative peptide sequence of coronaviral RBD compared with the closest known zoonotic coronaviral strain can provide better potential vaccine candidates for human testing. After the emergence of COVID-19 pandemic, SARS-CoV-2 surface protein has been repeatedly utilized for the identification of potential vaccine epitopes for SARS-CoV-2 [23–25]. However, there has also been simultaneous speculation regarding potential existence of cross resistance epitopes between SARS-CoV-2 and SARS-CoV [17,31,32]. Earlier, the computational probing of protein structures for respiratory infections via employment of docking methods has added useful information to stereochemical properties, virus binding mediated host receptor conformational transformation and binding preferences [33–37]. Contextually, viral receptor interactions were considered valuable in the instances of picornaviruses, influenza, HIV and CoVs [20,38–40]. For the SARS-CoV-2-induced infection, the basic reproduction number for viral transmissibility (R_o) as per various estimates is around 1.1–5.5 [41,42]. Since this threshold suggests very high infectivity rate, it is pertinent to target the viral binding region with vaccines to prevent infection.

Vaccinable peptide sequence for epitope based vaccine in case of alphaviruses, hepatitis B and C, HIV, HPV, and influenza viruses for recognition of potent immunogens has given propitious results [43–46]. Several studies reported on SARS and MERS CoV strains provided useful information regarding the potential epitopes retained by these strains [47–54], while, the data in context pertinent to SARS-CoV-2 are insufficient. The global COVID-19 pandemic has sparked rigorous R&D activity for vaccine development, and in a matter of just 4 months, various potential vaccine candidates are in the preclinical and clinical development phases [55]. The clinical behavior of SARS-CoV-2, infecting people around the world, with varied clinical symptomatology, ranging from completely asymptomatic to rapidly progressing lethal respiratory insufficiency demands the utilization of thorough and rapid novel technology platforms with more vaccinable options against SARS-CoV-2 [55,56].

This study reports key findings vis-a-vis SARS-CoV-2 RBD for its variable and conservative residues in comparison with BatCoV RaTG13 strain, which can be considered immunogenic epitopes for potential multi-epitope vaccine candidate for SARS-CoV-2 in the backdrop of its binding orientation.

Materials & methods

To identify the presence of antigenic epitopes within the RBD of S glycoprotein of SARS-CoV-2, *in silico* analysis was performed (Figure 1). The antigenicity of RBD was determined through VaxiJen v2.0. It is an alignment-free method for prediction of antigenicity. It predicts the antigenicity of proteins on the basis of physicochemical properties of amino acids [57].

Sequence retrieval & multiple sequence alignment

Protein sequences of S glycoprotein of SARS-CoV-2 (reported till 31 March 2020) were retrieved from National Center for Biotechnology Information. The sequences reported from China, Australia, USA, Taiwan, India, Pakistan, Nepal, Italy, Sweden, Brazil, Vietnam, Spain, Colombia, Peru and Japan were selected for analysis. The sequence of S glycoprotein from BatCoV RaTG13 (GISAID accession no. EPI_ISL_402131) and SARS CoV ZJ02 (Accession No. ABB29898) were used as reference for comparison. Sequence analysis was performed to ascertain the changes in the RBD of S glycoprotein. Multiple sequence alignments were performed using Clustal X. The consensus sequence of SARS-CoV-2 was used as input for epitope prediction.

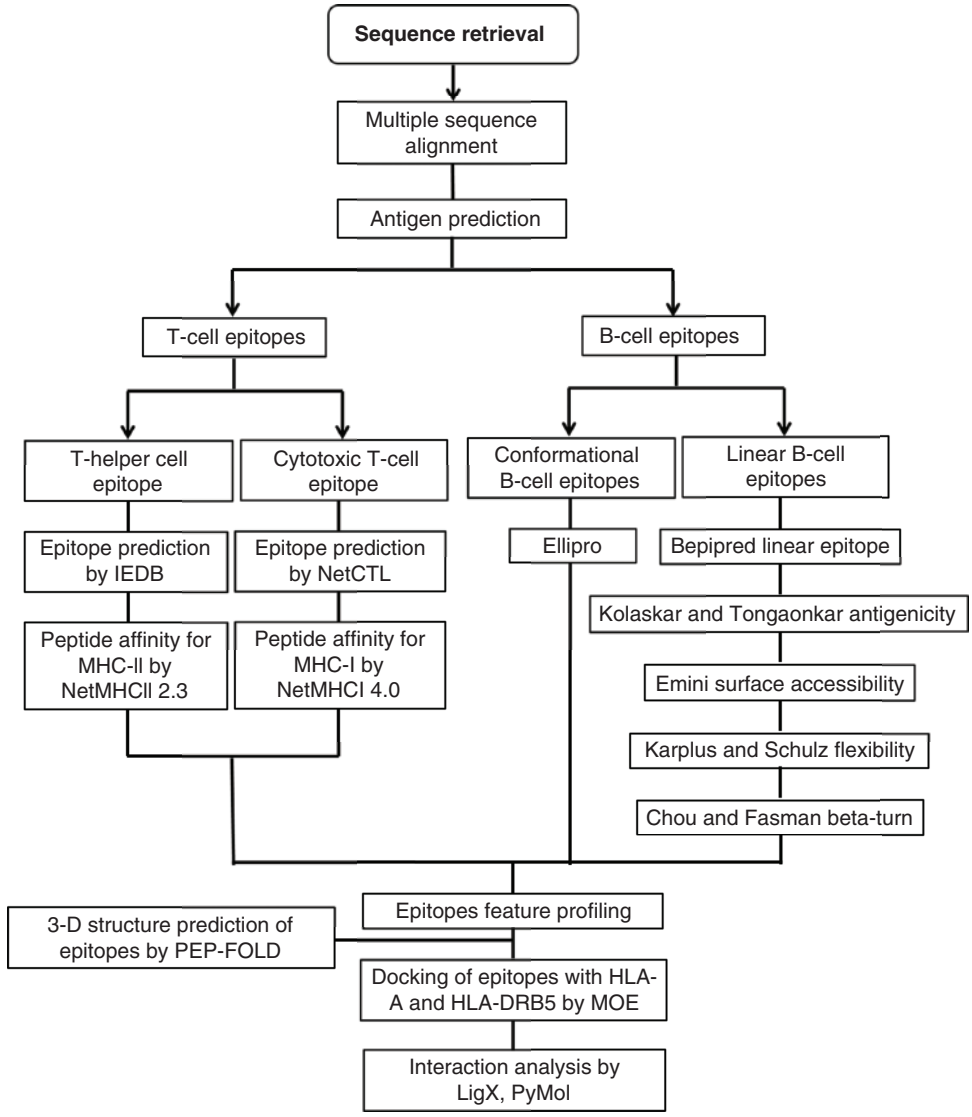


Figure 1. Detailed workflow for analysis of receptor binding domain of SARS coronavirus 2. The sequences of SARS-CoV-2, BatCoV RaTG13 and SARS-CoV were downloaded from the NCBI followed by multiple sequence alignment through ClustalX and JalView. Within the receptor-binding domain, the antigenic epitopes were also determined using immune-informatics approach. Linear and conformational B-cell epitopes were determined. Linear B-cell epitopes were determined through different methods that included BepiPred, Karplus & Schulz, Kolaskar & Tongaonkar, Emini, and Chou & Fasman methods. Conformational B-cell epitopes were determined through Ellipro. Cytotoxic and helper T-cell epitopes were also determined. CTL epitopes were determined through NetCTL followed by their affinity for specific MHC-I alleles through NetMHCI 4.0. Helper T-cell epitopes were identified through IEDB followed by identification of allele specificity through NetMHCII 2.3. All the B- and T-cell epitopes were then subjected to feature profiling by determining their toxicity, mutation, antigenicity, conservation, charge, molecular weight and nondigesting enzymes. Selected T-cell epitopes were then modeled through PEPFOLD. The interaction analysis of peptides with HLA-A and HLA-DR was done through MOE. Interacting residues of HLA-A and HLA-DR with peptides were determined through LigX in MOE and PyMOL.
BatCoV: Bat CoV; CoV: Coronavirus; CTL: Cytotoxic T lymphocyte; IEDB: Immune epitope database; MHC-I: MHC class I; MOE: Molecular operating environment; NCBI: National Center for Biotechnology Information.

B-cell epitope prediction

For identification of B-cell epitopes, Immune Epitope Database (IEDB) and BepiPred-2.0 were used [58,59]. The consensus sequence of SARS-CoV-2 was used to predict B-cell epitopes through Karplus & Schulz flexibility, Kolaskar & Tongaonkar antigenicity, Chou & Fasman beta-turn, Emini surface accessibility and BepiPred linear epitope prediction methods. Kolaskar & Tongaonkar method predicted antigenicity on the basis of amino acid abundances in naturally occurring epitopes as well as their physicochemical properties. The default threshold was set to 1.00 for antigenicity determination [60]. Emini surface accessibility method predicts the surface accessibility of epitopes as the surface accessible peptides recognized by the immune system [61]. Chou & Fasman beta-turn method was used to predict the antigenic regions, exhibiting beta turn as the beta turns are usually hydrophilic in nature and highly accessible [62]. Karplus & Schulz flexibility method predicts those antigens that are exhibiting flexible amino acids in nature as flexibility is correlated with antigenicity [63]. BepiPred prediction method based on hidden Markov model predicts linear epitopes in protein [64]. The B-cell epitopes were also predicted using ElliPro, which identifies linear and discontinuous epitopes in protein structure. It calculates the protrusion index (PI) of the residues and then cluster the residues on the basis of protrusion index values [65].

T-cell epitope prediction

For vaccine development, cytotoxic T-lymphocyte (CTL) epitopes play an important role. Hence the T-cell epitopes were identified that have the ability to bind with MHC class I (MHC-I) and class II (MHC-II). CTL epitopes were identified through NetCTL 1.2 server. It is based on artificial neural network and trained on different human MHC alleles for prediction of epitopes [66]. The IEDB and NetMHCI 4.0 server were used to predict the binding of epitopes with MHC-I. NetMHCI 4.0 predicts the binding affinity through artificial neural network by schooling 81 distinct HLA-A, -B, -C and -E human MHC alleles [67]. T-helper cell epitopes were predicted through IEDB and NetMHCII 2.3 server [68]. The epitopes were predicted having high affinity toward HLA-DR, -DQ and -DP. The length of the predicted epitopes was set to be 9-mer epitopes as it is reported that most HLA molecules have a strong preference for binding 9-mer epitopes [33]. For all the T-cell epitopes the threshold for predicting the binding affinity was set to $\leq$500 nM.

B- & T-cell epitopes feature profiling

B- and T-cell epitopes were further scrutinized for their enzymatic digestion, toxicity, hydrophobicity and physiochemical properties. The digestion of peptides with enzymes is an important parameter in vaccine development as the peptides that are digested by many enzymes are usually rendered unstable. Hence, the digestion of peptides by different enzymes was predicted through protein-digest server. AntiangioPred was used to predict the mutational variability and other physicochemical properties of peptides. It is based on machine learning model that is generated on the basis of already reported anti-angiogenic peptide [69]. ClanTOX predicted toxicity of peptides [70]. It is trained on set of ion-channel inhibitors and a set of nontoxin proteins. Based on the model, the program predicts the toxicity of input peptide. Antigenicity of peptides was predicted through Immunomedicine group server. For a peptide to be antigenic the threshold is 1.0. Antigenic prediction is based on physicochemical properties of epitopes like hydrophilicity, flexibility and accessibility [60].

Human proteome analysis for nonhuman homologues

To avert autoimmunity, vaccine contenders were screened for human and nonhuman homologues. The nonhuman homologues were identified by scrutinizing selected epitopes sharing <30% identity with human proteome, via BLASTp analysis.

Docking of T-cell epitopes with MHC-I & MHC-II alleles

The peptides that showed affinity for maximum number of MHC-I and MHC-II alleles were selected for interaction analysis. The structure of peptides was modeled through PEPFOLD server [71] followed by energy minimization. In case of MHC-I, the common allele between the peptides was selected for docking. Hence, the crystal structure of human *HLA-A*0101* was downloaded from Protein Databank (PDB; PDB ID: 6AT9: resolution = 2.9 Å). Same criterion was followed for MHC-II alleles and for that the crystal structure of *HLA-DRB5*01:01* (PDB ID: 1FV1: resolution = 1.9 Å) was also downloaded from PDB. Both the structures after ligand removal underwent protonation followed by energy minimization by AMBER99 force field. In order to analyze the inhibitory potential of peptides the docking of MHC-I specific peptides was executed with the *HLA-A*0101* alleles, while the MHC-II

specific peptides were docked with HLA-DRB5_0101. The docking studies were accomplished with induced fit docking protocol implemented in molecular operating environment (MOE) version 2016.08. By using, triangle match as placement method with London-dG scoring and GBVI/WSA dG rescoring function, 50 poses of each peptide were generated. Interaction analysis was done using MOE and PyMOL v2.3.

Interaction analysis of BatCoV RaTG13 with bat ACE2

After identification of vaccine epitopes, we further explored whether these vaccine epitopes harbor any important residues that could be involved in binding of SARS-CoV-2 with human ACE2 (hACE2) and BatCoV RaTG13 with bat ACE2 (bACE2). The interactions of SARS-CoV-2 with hACE2 were recently reported by Lan *et al* [72]. The interaction analysis of BatCoV RaTG13 with bACE2 was performed in the current study. To the best of our knowledge, the structure of RBD of BatCoV RaTG13 as well as bACE2 has not been determined yet. Hence the structure of both the proteins was determined through homology modeling using Modeller V9.23. BatCoV RaTG13 was modeled using SARS-CoV as template (PDB ID: 2GHV). While bACE2 was modeled using hACE (PDB ID: 1R42) as template. The generated model was subjected to model evaluation and structural validation via Ramachandran plot, PROSA, ERRAT, QMEAN and MolProbity. Ramachandran plot calculates the presence of amino acid residues in allowed, favored and outlier regions on the basis of torsional angles (Ψ and Φ) of amino acids [73]. PROSA reveal the quality of model by estimating any error in the models. It also calculates the score of model on the basis of experimentally reported (x-ray and nuclear magnetic resonance) structures of proteins [74]. Qualitative model energy analysis (QMEAN) apprises the geometry of protein structure by measuring the torsion angles on three consecutive amino acid residues [75]. MolProbity evaluates the protein structure by assessing its geometry [76]. ERRAT gauged the quality of model by analyzing the statistics between nonbonded interactions and different type of atoms and compared these values with the extremely refined structures [77]. The best model was then subjected to energy minimization using AMBER99 force field implemented in MOE. Docking of BatCoV RaTG13:bACE2 was performed using HADDOCK web server [78]. Analysis of protein–protein interactions was performed through pdbSum [79] and PyMOL v2.3.

Results

Multiple sequence alignment

RBD of SARS-CoV-2 is 192 amino acids long (within position 330–522 amino acids) lying in S1 region of S glycoprotein. When comparing the receptor binding motif with the BatCoV RaTG13 there was variation between the two virus strains (Supplementary Figure 1). The important changes were observed at position 439 (Lys→Asn), 440 (His→Asn), 441 (Ile→Leu), 443 (Thr→Lys), 445 (Glu→Val), 449 (Phe→Tyr), 459 (Ala→Ser), 478 (Lys→Thr), 483 (Gln→Val), 484 (Thr→Glu), 486 (Leu→Phe), 490 (Tyr→Phe), 492 (Ile→Leu), 493 (Tyr→Gln), 494 (Arg→Ser), 498 (Tyr→Gln), 501 (Asp→Asn) and 505 (His→Tyr). The changes at positions 441, 486, 492, 493, 498 and 505 may not have any obvious effect on binding due to similar properties of amino acids.

B-cell epitopes within RBD

Continuous B-cell epitopes were predicted using B-cell epitope prediction methods on IEDB server. The Kolaskar & Tongaonkar method predicted 11 antigenic epitopes in the RBD (Table 1), which can prompt B-cell responses. Surface accessibility analysis revealed four epitopes with surface accessibility (Table 1). Flexibility of epitopes is a measure of antigenicity [53]. The flexible epitopes in RBD were at positions 352–362, 380–392, 400–410, 421–433, 434–451, 454–473, 472–487 and 495–506. Beta-turns are the more flexible regions of the protein. According to Chou & Fasman predictions the beta-turn epitopes were at positions 437–443, 455–468, 422–428 and 495–500. Linear epitopes determined by BepiPred method are shown in Table 1. On the basis of consensus of all the methods, the peptides that can induce B-cell response were placed at positions 423–428, 455–461 and 494–500. The 423–428, 455–461 and 494–500, were the peptides that may prompt B-cell responses as predicted by ElliPro method. The mapping of epitopes on 3D structure of protein is shown in Supplementary Figure 2.

Cytotoxic T-cell epitope prediction

The default setting in the NetCTL server was used to predict T-cell epitopes. On the basis of highest combinatorial scores, five epitopes (with NetCTL score = 1.1–2.5 nM) were opted for subsequent analysis (Table 2). On the basis of NetCTL scores the peptide with the highest score (2.5 nM) had sequence CVADYSVLY. Further analysis of all the

Table 1. Prediction of B-cell epitopes by Kolaskar & Tongaonkar, Emini surface accessibility and BepiPred methods.

S. No.	Start position	End position	Peptide	Peptide length
Kolaskar & Tongaonkar method				
1.	334	341	NLCPFGEV	8
2.	347	353	FASVYAW	7
3.	358	372	ISNCVADYSVLYNSA	15
4.	374	385	FSTFKCYGVSPT	12
5.	387	404	LNDLCFTNVYADSFVIRG	18
6.	407	412	VRQIAP	6
7.	429	436	FTGCVIAW	8
8.	455	461	LFRKSNL	6
9.	423	428	YKLPDD	5
10.	470	478	TEIYQAGST	9
11.	494	500	SYGFQPT	6
Emini method				
1.	419	428	ADYNYKLPDD	10
2.	437	442	NSNNLD	6
3.	455	468	LFRKSNLKPFERDI	14
4.	495	500	YGFQPT	6
BepiPred method				
1.	382	385	VSPT	4
2.	407	420	VRQIAPGQTGKIAD	14
3.	423	428	YKLPDD	6
4.	439	447	NNLDSKVGG	9
5.	461	463	LKP	3
6.	466	467	RD	2
7.	469	469	S	1
8.	473	483	YQAGSTPCNGV	11
9.	495	506	YGFQPTNGVGYQ	12

five peptides for their binding with MHC-I showed that peptide CVADYSVLY illustrated binding with maximum MHC-I alleles (*HLA-A*26:01, HLA-A*01:01, HLA-A*30:02, HLA-B*35:01, HLA-A*11:01, HLA-B*15:01, HLA-A*68:01, HLA-A*03:01, HLA-B*53:01 and HLA-C*07:01*). The next peptide showing the binding with maximum number of alleles was FTNVYADSF.

Helper T-cell epitope prediction

A total of nine peptides were predicted which exhibited strong affinity for MHC-II alleles (Table 2). Among these the peptide YRLFRKSNL and VYAWNRKRI reflected affinity for maximum number of alleles. YRLFRKSNL held strong affinity with large number of MHC-II allele including: *DRB1_0103, DRB1_0701, DRB1_0801, DRB1_0802, DRB1_1602, DRB4_0103, DRB1_1001, DRB1_1101, DRB1_1501, DRB4_0103 and DRB5_0101.*

B- & T-cell epitopes feature profiling

To identify the best epitope for vaccine construction, different features of T-cell epitopes were determined (Table 3). The identified epitopes did not exhibit any homology with human proteins, were conserved and predicted to be nontoxic. The peptides which were digested by fewer enzymes have been considered good potential vaccine candidates (Table 3). Antigenicity of the peptides depicted that CTL specific peptides can be antigenic except the peptide ERDISTEIY. In case of helper T-cell epitopes, FELLHAPAT, TGCVIAWNS and VLYNSASFS were highly antigenic. In case of B-cell epitopes, all the three peptides were antigenic.

Interaction analysis of CTL epitopes with MHC-I specific alleles

Two peptides (CVADYSVLY and FTNVYADSF) with the maximum number of bindings with HLA alleles were selected for interaction analysis. PEPFOLD created five models for each peptide and the model with the best score

Table 2. T-cell epitopes predicted to be recognized by MHC class I and class II alleles.

T-cell epitopes bind to specific MHC-I alleles		
S. No.	**Peptide**	**MHC-I binding**
1.	NATRFASVY	*HLA-B*35:01* *HLA-A*01:01*
2.	RISNCVADY	*HLA-A*30:02, HLA-B*15:01, HLA-A*03:01 and HLA-A*01:01*
3.	CVADYSVLY	*HLA-A*26:01, HLA-A*01:01, HLA-A*30:02, HLA-B*35:01, HLA-A*11:01, HLA-B*15:01, HLA-A*68:01, HLA-A*03:01, HLA-B*53:01 and HLA-C*07:01*
4.	FTNVYADSF	*HLA-A*01:01, HLA-B*15:01, HLA-B*15:03, HLA-B*15:17, HLA-A*25:01, HLA-A*26:01, HLA-B*08:03, HLA-B*58:01, HLA-B*53:01 and HLA-C*03:03*
5.	ERDISTEIY	*HLA-A*01:01*
T-cell epitopes bind to specific MHC-II alleles		
S. No.	**Peptide**	**MHC-II binding**
1.	FELLHAPAT	*DRB1_0101 and DRB1_1001*
2.	FNATRFAS	*DRB1_0402*
3.	TGCVIAWNS	*DRB1_0403 and DRB3_0202*
4.	FRKSNLKPF	*DRB1_0701*
5.	YRLFRKSNL	*DRB1_0103, DRB1_0701, DRB1_0801, DRB1_0802, DRB1_1602, DRB4_0103, DRB1_1001, DRB1_1101, DRB1_1501, DRB4_0103 and DRB5_0101*
6.	VYAWNRKRI	*DRB1_1101, DRB1_1301, DRB4_0103, DRB5_0101, DRB1_0402 and DRB3_0202*
7.	FERDISTEI	*DRB3_0101*
8.	IRGDEVRQI	*DRB3_0101*
9.	VLYNSASFS	*DRB3_0202*

MHC-I: MHC class I; MHC-II: MHC class II.

was selected for further analysis. *HLA-A*0101* was selected as representative allele for interaction analysis. The docking of HLA-A with the top ranking peptide resulted in generation of 50 complexes for each peptide. The peptide 3 (CVADYSVLY) binds with HLA with binding energy of -11.3 KJ/mol, while a detailed interaction analysis revealed that peptide had H-bonding interactions with Asp-115, Arg-156, Tyr-99 and Asn-77 (Figure 2A). Peptide 4 (FTNVYADSF) bound with the HLA-A with an energy value of -12.6 KJ/mol. The peptide exhibited H-bonding interactions with Glu-63, Arg-156, Gln-155, Asn-77, Thr-143 and Lys-146 (Figure 2B).

Interaction analysis of helper T lymphocyte epitopes with MHC-II specific alleles

Two of the peptides (YRLFRKSNL and VYAWNRKRI) that showed affinity with maximum number of MHC-II alleles were selected for interaction analysis. Docking of peptides with HLA-DRB5 depicted both the peptides to be binding with strong affinity. The peptide 6 and peptide 5 with maximum number of alleles were used to study affinity of these peptides with *DRB5_0101*. Peptide 5 (YRLFRKSNL) bound with the energy of -10.2 KJ/mol while peptide 6 (VYAWNRKRI) bound with the energy of -11.3 KJ/mol. The residues Asp-70, Asn-62, Arg-71 and Glu-55 of HLA-DRB5 were having H-bonding interactions with peptide 5 (Figure 2C). Interaction analysis revealed that Ser-53, Asp-11, Asp-30 and Asp-70 of HLA-DRB5 exhibited H-bonding interactions with peptide 6 (Figure 2D).

Interaction analysis of BatCoV RaTG13 with bACE2

The interactions of BatCoV RaTG13 with bACE2 were done using HADDOCK. There were 155 different complexes of BatCoV RaTG13:bACE2 that were generated, which clustered into 12 groups. Supplementary Table 1 shows the Z scores of all the seven clusters, size of each cluster, root-mean-square deviation (RMSD) from the overall lowest energy structure, and energy values of electrostatic, Van der Waals, and desolvation. The cluster with the best HADDOCK score (-178.9 ± 3.6) was further used for analysis. Detailed interaction analysis showed that 26 residues of bACE2 and nine residues of BatCoV RaTG13 were present at the interface. These residues were involved in 12 H-bonded, two salt bridges and 157 nonbonded contacts. The detailed interactions are shown in Table 4 and Figure 3.

The important residues of BatCoV RaTG13 that were involved in interactions with bACE2 are Lys-417, Leu-455, Phe-456, Asn-487, Tyr-489, Asp-501 and His-505. The important conserved residues of BatCoV RaTG13 that involved in interactions and demonstrated overlapping with epitopes were Leu-455 and Phe-456. These residues

Table 3. Profiling of MHC class I, class II and B-cell specific peptides.

S. No.	Peptide	Homology with human proteins (Y/N)	Mutation (Y/N)	Toxicity (Y/N)	Hydrophobicity	Charge	Molecular weight	Antigenicity	Nondigesting enzymes
MHC-I peptide									
1.	NATRFASVY	N	N	N	-0.13	1.0	1028.24	1.03	Cyanogen bromide, idosobenzoate, proline endopept, staph protease, trypsin K and AspN
2.	RISNCVADY	N	N	N	-0.20	0.00	1040.27	1.07	Cymotrypsin, cyanogen bromide, trypsin K, endopept, staph protease andidosobenzoate
3.	CVADYSVLY	N	N	N	0.11	-1.00	1032.29	1.18	Trypsin, clostripain, cyanogen bromide, idosobenzoate, proline endopept, staph protease, trypsin K and trypsin R
4.	FTNVYADSF	N	N	N	0.03	-1.0	10632.24	1.03	Trypsin, clostripain, cyanogen bromide, idosobenzoate, proline endopept, staph protease, trypsin K and trypsin R
5.	ERDISTEIY	N	N	N	-0.30	-2.0	1125.30	0.95	Cymotrypsin, cyanogen bromide, idosobenzoate, proline endopept and trypsin K
MHC-II peptide									
1.	FELLHAPAT	N	N	N	0.10	-0.5	998.27	1.09	Trypsin, clostripain, cyanogen bromide, idosobenzoate, trypsin K, trypsin R and AspN
2.	FNATRFAS	N	N	N	-0.14	1.0	913.09	0.98	Cyanogen bromide, idosobenzoate, proline endopept, staph protease, N trypsin K and AspN
3.	TGCVIAWNS	N	N	N	0.11	0.0	950.20	1.05	Trypsin, clostripain, cyanogen bromide, proline endopept, staph protease, trypsin K, trypsin R and AspN
4.	FRKSNLKPF	N	N	N	-0.35	3.0	1136.48	0.99	Cyanogen bromide, idosobenzoate, staph protease and AspN
5.	YRLFRKSNL	N	N	N	-0.43	3.0	1196.54	0.99	Cyanogen bromide, idosobenzoate, proline endopept, staph protease and AspN
6.	VYAWNRKRI	N	N	N	-0.03	3.0	1205.55	0.99	Cyanogen bromide, proline endopept, staph protease and AspN
7.	FERDISTEI	N	N	N	-0.23	-2.0	1109.32	0.95	Cyanogen bromide, idosobenzoate, proline endopept and trypsin K
8.	IRGDEVRQI	N	N	N	-0.38	0.00	1085.36	0.98	Cymotrypsin, cyanogen bromide, idosobenzoate, proline endopept and trypsin K
9.	VLYNSASFS	N	N	N	0.06	0.00	987.19	1.09	Trypsin, clostripain, cyanogen bromide, idosobenzoate, proline endopept, staph protease, trypsin K, trypsin R and AspN
B-cell epitopes									
1.	YKLPDD	N	N	N	-0.34	-1.00	749.89	1.05	Clostripain, cyanogen bromide, idosobenzoate, staph protease, trypsin R and elastase
2.	LFRKSN	N	N	N	-0.44	2.00	763.97	1.03	Cyanogen bromide, idosobenzoate, proline endopept, staph protease and AspN
3.	SYGFQPT	N	N	N	-0.06	0.00	798.94	1.03	Trypsin, clostripain, cyanogen bromide, idosobenzoate, staph protease, trypsin K, trypsin R, elastase and AspN

Prediction of their toxicity, mutation, homologs present in human proteome, antigenicity, digestion with enzymes and other physiochemical properties as charge, weight and hydrophobicity.
MHC-I: MHC class I; MHC-II: MHC class II; N: No; Y: Yes.

lay within the B-cell epitope (455-LFRKSN-461) and T-cell epitope (453-YRLFRKSNL-461). The important interacting residues of BatCoV RaTG13 with bACE2 and a comparison with SARS CoV2:hACE2 are shown in Figure 4.

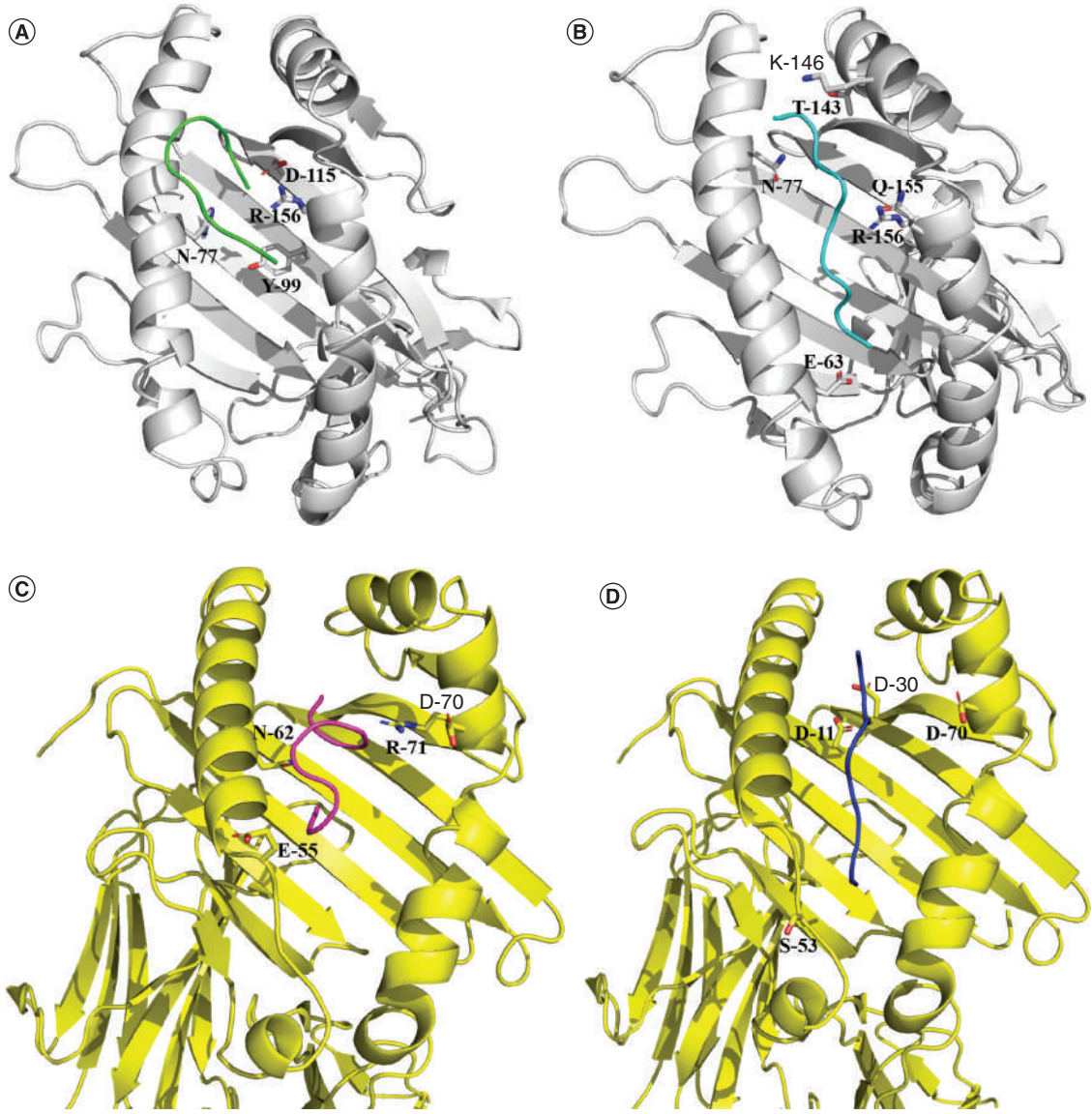


Figure 2. Interaction analysis of peptides with HLA-A & HLA-DR. (A) CVADYSVLY:*HLA-A*01:01*. **(B)** FTNVYADSF:*HLA-A*01:01*. **(C)** YRLFRKSNL:*DRB5_0101*. **(D)** VYAWNRKRI:*DRB5_0101*. HLA-A protein is shown in grey while HLA-DR protein is shown in yellow. The residues involved in interactions are highlighted as stick representation.

Discussion

Research on various CoV species earlier has been continually reported since the last one and a half decade after the emergence of SARS, for annotating signatures and virulence factors [80]. Viral entry receptors are crucial in viral life cycle, sustenance and egress [81]. Realizing their particular tissue tropism further augments their importance for therapeutic targeting and restricting viral entry into cell which can abolish infectivity altogether.

In case of SARS-CoV and MERS-CoV, S protein and specially S1 regions have been the prime focus in developing immune strategies against these strains [52]. Similar strategy can be employed by investigating the S protein for identifying immune epitopes against SARS-COV-2. Vaccines against SARS-CoV-2 can serve as one of the most promising modes of containing COVID-19 pandemic. To date no reliable treatment options are available for COVID-19, so logically vaccine against it is a much needed solution. As the global burden of infectivity by COVID-19 pandemic is increasing every day, computational biology aided vaccine design for SARS-CoV-2 with removal of unnecessary antigenic load and screened allergic response can provide the characteristic immune

Table 4. Interactions of bat coronavirus RaTG13 with bat angiotensin-converting enzyme2.

bACE2			BatCoV RaTG13	
S. No.	Residue name	Type of interactions	Residue name	Distance (Å)
1.	Glu-24	H-bond	Tyr-473	2.73
2.	Asp-30	H-bond	Lys-417	2.63
3.	Thr-30 Lys-353	H-bond	His-505	2.78
4.	Lys-35	H-bond	Arg-494	3.08
5.	Asp-38	H-bond	Tyr-498	2.55
6.	Gln-42	H-bond	Glu-445	2.73
7.	Lys-61	H-bond	Glu-445	2.86
8.	Leu-79	H-bond	Tyr-489	2.84
9.	Asn-82	H-bond	Asn-487	3.03
10.	Lys-31	H-bond	Leu-455 Phe-456	2.01 3.8
11.	Gln-42	H-bond	Tyr-489	3.4
12.	His-41	H-bond	Asp-501	3.9
13.	Asp-30	Salt bridge	Lys-417	2.63
14.	Lys-61	Salt bridge	Glu-445	2.77

bACE2: Bat ACE2; BatCoV: Bat coronavirus.

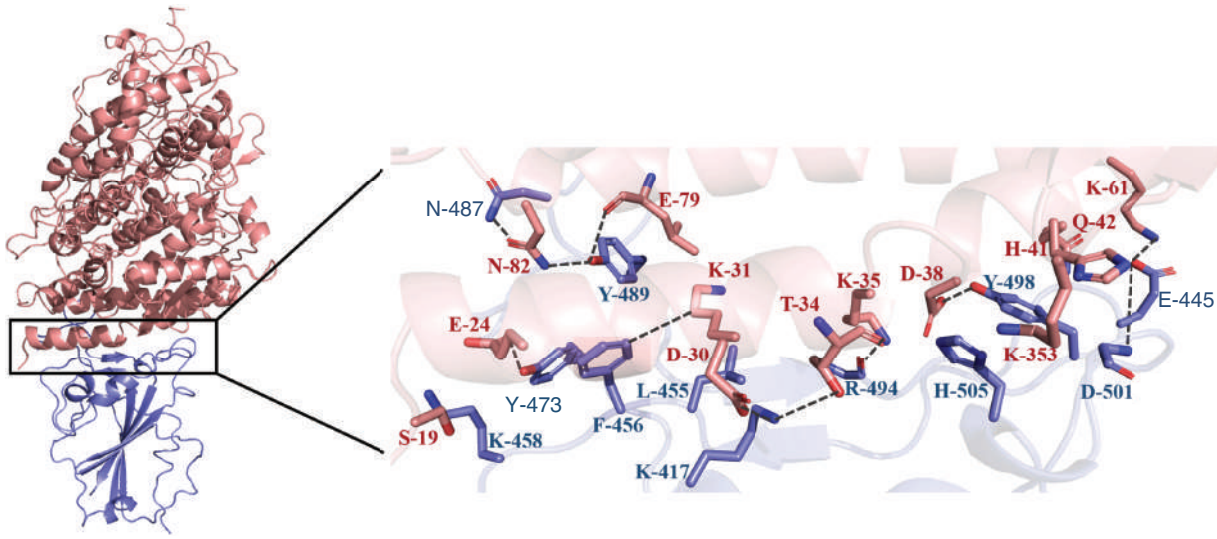


Figure 3. Top scoring complex of bat coronavirus RaTG13:bat ACE2. BatCoV RaTG13 is shown in blue while bACE2 is shown in pink. bACE2: Bat ACE2; BatCoV: Bat coronavirus.

response required for preventing SARS-CoV-2 infection. Similar to SARS-CoV and MERS-CoV, S1 region of CVOID-19 harbors the RBD which is involved in the entry of virus into a host cell. Hence, the identification of antigenic peptides within RBD can be a good strategy to forestall infection. The present study focused on deriving immunogenic epitopes capable of triggering both humoral and cell mediated immune response, on the basis of high degree of comparative sequence similarity of RBD from BatCoV RaTG13 with SARS-CoV-2. Previously, immunogenic epitopes for SARS-CoV-2 by immunoinformatics method have been reported in comparison with SARS-CoV [23–25]. Using this approach may yield rather specific epitopes against SARS-CoV-2. Contextually, cross-reactivity of SARS-CoV antibodies against SARS-CoV-2 epitopes is also under debate and is providing useful information against potential SARS-CoV-2 host immune response [17,31,32].

Figure 4. Sequence alignment of receptor binding domain of SARS coronavirus 2 with bat coronavirus RaTG13. B-cell epitopes are highlighted with orange boxes. CTL epitopes are highlighted with blue boxes. T-helper cell epitopes are highlighted as green boxes. Green dots represent SARS-CoV-2 interacting residues with hACE2 while blue dots represent interacting residues of BatCoV RaTG13 with bACE2. bACE2: Bat ACE2; BatCoV: Bat CoV; CoV: Coronavirus; CTL: Cytotoxic T lymphocyte; hACE2: Human ACE2.

Current results from analysis of RBD of SARS-CoV-2 revealed that 423-YKLPDD-428, 455-LFRKSN-461 and 494-SYGFQPT-500 are the B-cell epitopes that are highly antigenic. The two epitopes lie in receptor binding motif and also muddled in interactions with hACE2 [72]. The Leu-455, Phe-456, Ser-494 and Thr-500 were the conserved residues among all the SARS-CoV-2 sequences hooked in interactions with hACE2. In case of T-cell epitopes, several epitopes were identified from IEDB and NetCTL and the peptides manifesting affinity with maximum number of MHC-I/MHC-II alleles were further selected for analysis. Previously, it has been reported that peptides showing affinity with maximum number of HLA alleles could be very antigenic [53,54]. In case of CTL epitopes, two epitopes (peptide 3: CVADYSVLY and peptide 4: FTNVYADSF) were found to be highly antigenic and also showed strong binding with HLA-A allele. HLA are polymorphic proteins with variable expression in different population. Therefore, a vaccine which is suited for all population without inciting any autoimmunity met by for T-cell epitopes by HLA selectivity is crucial for an effective vaccine candidate [82]. In case of HLA-A*0101, the already docked peptide in crystal structure (PDB ID: 6AT9) has shown interactions with Tyr-7, Phe-9, Phe-33, Val-34, Tyr-59, Glu-63, Asn-77, Tyr-84, Ile-97, Tyr-99, Thr-143, Lys-146, Trp-147, Gln-155, Arg-156 and Tyr-159 of HLA-A [83]. Peptide 3 (CVADYSVLY) has paraded interactions with Arg-156, Tyr-99 and Asn-77 while peptide 4 (FTNVYADSF) showed interactions with maximum number of amino acids (Glu-63, Arg-156, Gln-155, Asn-77, Thr-143 and Lys-146). In case of helper T-cell epitopes, YRLFRKSNL and VYAWNRKRI illustrated affinity with maximum number of MHC-II alleles. One of the MHC-II allele HLA-*DRB5*0101* was used to study possible interactions of peptide with HLA-DR. In case of HLA-*DRB5*0101* (PDB ID: 1FV1) the crystallized structure displayed interactions with Asp-9, Phe-12, Glu-55, Met-23, Ser-53, Asn-62, Asp-66, Asn-69, Asp-11, Tyr-13, Asp-70, Arg-71, Tyr-78 and Asn-82 [84]. When we studied peptide 5 (YRLFRKSNL), Asn-62, Arg-71 and Glu-55 were residues important for binding of peptide. In case of peptide 6 (VYAWNRKRI), the Ser-53, Asp-11 and Asp-70 were involved in interactions with peptide. The interaction as well as binding energy data depicted these two peptides to be binding with HLA-DRB5 and can be used in vaccine construction. The peptide 453-YRLFRKSNL-461, present within the receptor binding motif and Leu-455 and Phe-456 illustrated interactions with hACE2 [72] as well as bACE2.

On the basis of different properties of these selected B- and T-cell epitopes, it can be observed that these epitopes were 100% conserved from the reported data, as predicted by conservation analysis. These peptides also did not exhibited homology with any human protein hence may not incite any autoimmunity. These peptides did not display any toxicity. The digesting enzyme data showed these peptides to be indigestible by a range of enzymes and hence are safer to use for vaccine development (Table 3). The B- and T-cell epitopes (LFRKSN and YRLFRKSNL) are overlapping, hence the presence of YRLFRKSNL in vaccine construct along with other peptides may enhance the efficacy of a vaccine. The analysis on the sequences submitted until October 2020 has shown that all the predicted epitopes have been fairly conserved except one change observed in T-cell epitope (453-YRLFRKSNL-461) region at position 453 as Y453F. This mutation was labeled as mink mutation and found in SARS-CoV-2 sequences obtained from Denmark, South Africa and The Netherlands [85,86]. The results proposed herein are preliminary and further *in vitro* and *in vivo* testing is required for the proposed vaccinable targets of SARS-CoV-2. Based on the study results, the predicted epitopes harbor attractive capability to be considered for ascertaining therapeutic potency. The identified epitopes from this study can further be investigated along with conserved region epitopes of

S protein for molecular simulation studies, in construction of experimental vaccine constructs and for considering its potential as a peptide based vaccine.

Cumulatively, the detailed interaction analysis of BatCoV RaTG13 has shown that hydrophobic and charged residues have been involved in binding with bACE2. Analogous pattern has been observed in binding of SARS-CoV-2 with hACE2 receptor [72]. The important residues of BatCoV RaTG13 that were involved in interactions with bACE2 are Lys-417, Leu-455, Phe-456, Asn-487, Tyr-489, Asp-501 and His-505. The residues at these positions in SARS-CoV-2 were also involved in interactions with hACE2 [72]. Among them Lys-417, Leu-455, Phe-456, Asn-487 and Tyr-489 were the conserved residues (Figure 4).

Designing an effective vaccine against viral infection such as COVID-19 is tricky. On one hand, it has to be ensured that the vaccinable epitopes hold enough antigenic potential to mount a befitting yet specific immune response so as to rapidly clear the infection if the need arises; on the other hand, the host immune response should not be strong enough to trigger chronic inflammation which in case of COVID-19 might significantly deteriorate lung infection. Lung as an organ is highly sensitive to inflammatory changes initiated by a surge of cytokine response [87]. Mutations in the viral genome are capacitating CoVs to breach species barrier repeatedly. As the CoVs harbor an error-prone RNA dependent-RNA polymerase, it may engender recombination events with mutational diversity, concocting therapeutic challenges and survival advantage to the virus [80]. This has been the case observed in SARS-CoV epidemic of 2004 [88,89]. It is of grave concern that SARS-CoV-2 has the potential to reach pandemic proportions while considering the low persistent R_o estimates. As bats are considered primary hosts for CoV species, it will be interesting to scrutinize how bats evade viral entry as previous studies have identified bats evolving a mechanism for defying interferon pathway activation by the STING interferon pathway [90].

Conclusion

The current study proposed the identification of potential multiple epitopes for vaccine development against SARS-CoV-2. The potential vaccine epitopes have been rigorously screened for multiple HLA, B-Cell, CTL and helper T lymphocyte epitopes thus augmenting its capability in inducing both humoral and cellular immune responses. Furthermore, it can be co-opted with adjuvant treatment in further enhancing viral disease control. The epitopes were further screened and validated for 100% conservancy with nonoverlapping human proteome thus additionally reducing the chances of autoimmunogenic side effects. The molecular docking of epitopes with HLA alleles were further validated for their mode of binding patterns and analysis of binding energy affinities.

Summary points

- The December of 2019 witnessed emergence of worldwide outbreak by a novel strain of coronavirus (CoV) termed SARS-CoV-2.
- There is no therapeutic or preventive strategy like vaccine developed so far to overcome infection.
- The receptor binding domain of SARS-Cov-2 for any potential vaccine epitopes were explored using immunoinformatics approach.
- The B-cell epitopes LFRKSN and SYGFQPT, were found to be highly antigenic. Among T-cell epitopes, the epitope CVADYSVLY and FTNVYADSF were antigenic and exhibited affinity for maximum number of MHC class I alleles. The T-cell epitopes YRLFRKSNL and VYAWNRKRI displayed affinity for maximum number of MHC class II alleles. The overlapping epitope among B- and T-cells was YRLFRKSNL.
- The epitopes were further screened and validated for 100% conservancy with nonoverlapping human proteome thus additionally reducing the chances of autoimmune side effects.
- The important conserved residues of BatCoV RaTG13, the Leu-455 and Phe-456 that have been involved in interactions, also demonstrated overlap with epitopes. These residues lay within the B-cell epitope (455-LFRKSN-461) and T-cell epitope (453-YRLFRKSNL-461).
- The deployment of these epitopes in potential vaccine against COVID-19 may help in sweeping the COVID-19 infectious spread.

Supplementary data

To view the supplementary data that accompany this paper please visit the journal website at: www.futuremedicine.com/doi/suppl/10.2217/fvl-2020-0269

Acknowledgments

We would like to acknowledge the Shanghai Public Health Clinical Center and School of Public Health, Central Hospital of Wuhan, Huazhong University of Science and Technology, Wuhan Center for Disease Control and Prevention, National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control, and University of Sydney Australia for releasing the sequence of COVID-19 genome, and Department of Chemistry, COMSATS, Abbottabad, for computational lab support.

Declaration of interest statement

The authors declare no conflict of interest. All the data used in the study are publicly available and have been cited properly. No human participant was directly involved in the study.

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

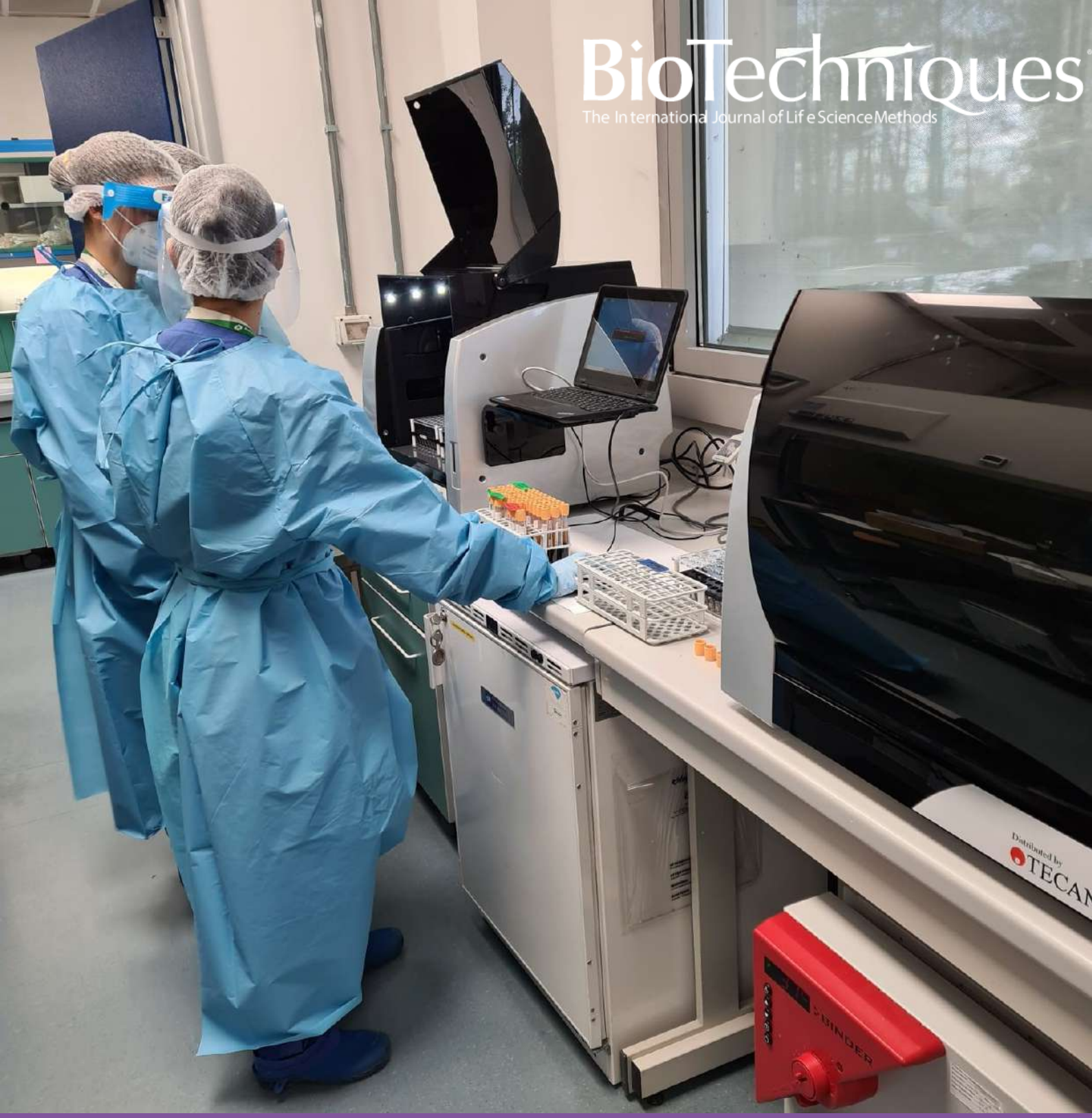
References

1. Coronavirus latest: WHO officially names disease COVID-19. www.nature.com/articles/d41586-020-00154-w
2. Zaki AM, Van Boheemen S, Bestebroer TM, Osterhaus AD, Fouchier RA. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N. Engl. J. Med.* 367(19), 1814–1820 (2012).
3. Drosten C, Gunther S, Preiser W *et al.* Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N. Engl. J. Med.* 348(20), 1967–1976 (2003).
4. Fouchier RA, Kuiken T, Schutten M *et al.* Aetiology: Koch's postulates fulfilled for SARS virus. *Nature* 423(6937), 240 (2003).
5. Wang Y, Sun J, Zhu A, Zhao J. Current understanding of middle east respiratory syndrome coronavirus infection in human and animal models. *J. Thorac. Dis.* 10(Suppl. 19), S2260–S2271 (2018).
6. Chan-Yeung M, Xu RH. SARS: epidemiology. *Respirology* 8(Suppl. 1), S9–S14 (2003).
7. Peiris JS, Guan Y, Yuen KY. Severe acute respiratory syndrome. *Nat. Med.* 10(Suppl. 12), S88–S97 (2004).
8. Lee J, Chowell G, Jung E. A dynamic compartmental model for the Middle East respiratory syndrome outbreak in the Republic of Korea: a retrospective analysis on control interventions and superspreading events. *J. Theor. Biol.* 408, 118–126 (2016).
9. Lee JY, Kim YJ, Chung EH *et al.* The clinical and virological features of the first imported case causing MERS-CoV outbreak in South Korea, 2015. *BMC Infect. Dis.* 17(1), 498 (2017).
10. Worldometer. Coronavirus (COVID-19) cases statistics. www.worldometers.info/coronavirus/coronavirus-cases/
11. Zhou P, Yang XL, Wang XG *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 579(7798), 270–273 (2020).
12. Wu A, Peng Y, Huang B *et al.* Genome composition and divergence of the novel coronavirus (2019-nCoV) originating in China. *Cell Host Microbe* 27(3), 325–328 (2020).
13. Su S, Wong G, Shi W *et al.* Epidemiology, genetic recombination, and pathogenesis of coronaviruses. *Trends Microbiol.* 24(6), 490–502 (2016).
14. Hu B, Zeng LP, Yang XL *et al.* Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus. *PLoS Pathog.* 13(11), e1006698 (2017).
15. Jeffers SA, Tusell SM, Gillim-Ross L *et al.* CD209L (L-SIGN) is a receptor for severe acute respiratory syndrome coronavirus. *Proc. Natl Acad. Sci. USA* 101(44), 15748–15753 (2004).
16. Meyerholz DK, Lambertz AM, Mccray PB Jr. Dipeptidyl peptidase 4 distribution in the human respiratory tract: implications for the Middle East respiratory syndrome. *Am. J. Pathol.* 186(1), 78–86 (2016).
17. Walls AC, Park YJ, Tortorici MA, Wall A, Mcguire AT, Veesler D. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell* 181(2), 281–292 e286 (2020).
18. Jiang S, Lu L, Liu Q, Xu W, Du L. Receptor-binding domains of spike proteins of emerging or re-emerging viruses as targets for development of antiviral vaccines. *Emerg. Microbes Infect.* 1(8), e13 (2012).
19. Zhu X, Liu Q, Du L, Lu L, Jiang S. Receptor-binding domain as a target for developing SARS vaccines. *J. Thorac. Dis.* 5(Suppl. 2), S142–S148 (2013).
20. Du L, Yang Y, Zhou Y, Lu L, Li F, Jiang S. MERS-CoV spike protein: a key target for antivirals. *Expert Opin. Ther. Targets* 21(2), 131–143 (2017).
21. Davies MN, Flower DR. Harnessing bioinformatics to discover new vaccines. *Drug Discov. Today* 12(9–10), 389–395 (2007).

22. Srivastava S, Kamthania M, Singh S, Saxena AK, Sharma N. Structural basis of development of multi-epitope vaccine against Middle East respiratory syndrome using *in silico* approach. *Infect. Drug Resist.* 11, 2377–2391 (2018).
23. 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* 12(3), 254 (2020).
24. Baruah V, Bose S. Immunoinformatics-aided identification of T cell and B cell epitopes in the surface glycoprotein of 2019-nCoV. *J. Med. Virol.* 92(5), 495–500 (2020).
25. Bhattacharya M, Sharma AR, Patra P *et al.* Development of epitope-based peptide vaccine against novel coronavirus 2019 (SARS-COV-2): immunoinformatics approach. *J. Med. Virol.* 92(6), 618–631 (2020).
26. Kuehn BM. More evidence emerges that bats may have spread SARS. *JAMA* 310(20), 2138 (2013).
27. Lau SK, Woo PC, Li KS *et al.* Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. *Proc. Natl Acad. Sci. USA* 102(39), 14040–14045 (2005).
28. Kan B, Wang M, Jing H *et al.* Molecular evolution analysis and geographic investigation of severe acute respiratory syndrome coronavirus-like virus in palm civets at an animal market and on farms. *J. Virol.* 79(18), 11892–11900 (2005).
29. Ge XY, Li JL, Yang XL *et al.* Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature* 503(7477), 535–538 (2013).
30. Forni D, Cagliani R, Clerici M, Sironi M. Molecular evolution of human coronavirus genomes. *Trends Microbiol.* 25(1), 35–48 (2017).
31. Tian X, Li C, Huang A *et al.* Potent binding of 2019 novel coronavirus spike protein by a SARS coronavirus-specific human monoclonal antibody. *Emerg. Microbes Infect.* 9(1), 382–385 (2020).
32. Yuan M, Wu NC, Zhu X *et al.* A highly conserved cryptic epitope in the receptor binding domains of SARS-CoV-2 and SARS-CoV. *Science* 368(6491), 630–633 (2020).
33. Bakkers MJ, Zeng Q, Feitsma LJ *et al.* Coronavirus receptor switch explained from the stereochemistry of protein–carbohydrate interactions and a single mutation. *Proc. Natl Acad. Sci. USA* 113(22), E3111–3119 (2016).
34. Kannan S, Kolandaivel P. Computational studies of pandemic 1918 and 2009 H1N1 hemagglutinins bound to avian and human receptor analogs. *J. Biomol. Struct. Dyn.* 34(2), 272–289 (2016).
35. Pan D, Xue W, Wang X, Guo J, Liu H, Yao X. Molecular mechanism of the enhanced virulence of 2009 pandemic influenza A (H1N1) virus from D222G mutation in the hemagglutinin: a molecular modeling study. *J. Mol. Model.* 18(9), 4355–4366 (2012).
36. Tran-to Su C, Ouyang X, Zheng J, Kwoh CK. Structural analysis of the novel influenza A (H7N9) viral neuraminidase interactions with current approved neuraminidase inhibitors oseltamivir, zanamivir, and peramivir in the presence of mutation R289K. *BMC Bioinformatics* 14(Suppl. 16), S7 (2013).
37. Xu X, Chen P, Wang J *et al.* Evolution of the novel coronavirus from the ongoing Wuhan outbreak and modeling of its spike protein for risk of human transmission. *Sci. China Life Sci.* 63(3), 457–460 (2020).
38. Skehel JJ, Wiley DC. Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. *Annu. Rev. Biochem.* 69, 531–569 (2000).
39. Kwong PD, Wyatt R, Robinson J, Sweet RW, Sodroski J, Hendrickson WA. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* 393(6686), 648–659 (1998).
40. Rossmann MG, He Y, Kuhn RJ. Picornavirus-receptor interactions. *Trends Microbiol.* 10(7), 324–331 (2002).
41. Chen J. Pathogenicity and transmissibility of 2019-nCoV-A quick overview and comparison with other emerging viruses. *Microbes Infect.* 22(2), 69–71 (2020).
42. Zhao S, Lin Q, Ran J *et al.* Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: a data-driven analysis in the early phase of the outbreak. *Int. J. Infect. Dis.* 92, 214–217 (2020).
43. Sominskaya I, Skrastina D, Dislers A *et al.* Construction and immunological evaluation of multivalent hepatitis B virus (HBV) core virus-like particles carrying HBV and HCV epitopes. *Clin. Vaccine Immunol.* 17(6), 1027–1033 (2010).
44. Rodrigues RL, Menezes GL, Saivish MV *et al.* Prediction of MAYV peptide antigens for immunodiagnostic tests by immunoinformatics and molecular dynamics simulations. *Sci. Rep.* 9(1), 13339 (2019).
45. Stanekova Z, Vareckova E. Conserved epitopes of influenza A virus inducing protective immunity and their prospects for universal vaccine development. *Virol. J.* 7, 351 (2010).
46. Jabbar B, Rafique S, Salo-Ahen OMH *et al.* Antigenic peptide prediction from E6 and E7 oncoproteins of HPV types 16 and 18 for therapeutic vaccine design using immunoinformatics and MD simulation analysis. *Front. Immunol.* 9, 3000 (2018).
47. Dutta NK, Mazumdar K, Lee BH *et al.* Search for potential target site of nucleocapsid gene for the design of an epitope-based SARS DNA vaccine. *Immunol. Lett.* 118(1), 65–71 (2008).
48. Liu SJ, Leng CH, Lien SP *et al.* Immunological characterizations of the nucleocapsid protein based SARS vaccine candidates. *Vaccine* 24(16), 3100–3108 (2006).

49. Zhao J, Huang Q, Wang W, Zhang Y, Lv P, Gao XM. Identification and characterization of dominant helper T-cell epitopes in the nucleocapsid protein of severe acute respiratory syndrome coronavirus. *J. Virol.* 81(11), 6079–6088 (2007).
50. Zhou M, Xu D, Li X *et al.* Screening and identification of severe acute respiratory syndrome-associated coronavirus-specific CTL epitopes. *J. Immunol.* 177(4), 2138–2145 (2006).
51. Gupta V, Tabiin TM, Sun K *et al.* SARS coronavirus nucleocapsid immunodominant T-cell epitope cluster is common to both exogenous recombinant and endogenous DNA-encoded immunogens. *Virology* 347(1), 127–139 (2006).
52. Wang L, Shi W, Joyce MG *et al.* Evaluation of candidate vaccine approaches for MERS-CoV. *Nat. Commun.* 6, 7712 (2015).
53. Shi J, Zhang J, Li S *et al.* Epitope-based vaccine target screening against highly pathogenic MERS-CoV: an *in silico* approach applied to emerging infectious diseases. *PLoS ONE* 10(12), e0144475 (2015).
54. Tahir Ul Qamar M, 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.* 17(1), 362 (2019).
55. Thanh Le T, Andreadakis Z, Kumar A *et al.* The COVID-19 vaccine development landscape. *Nat. Rev. Drug Discov.* 19(5), 305–306 (2020).
56. Kang S, Peng W, Zhu Y *et al.* Recent progress in understanding 2019 novel coronavirus (SARS-CoV-2) associated with human respiratory disease: detection, mechanisms and treatment. *Int. J. Antimicrob. Agents* 55(5), 105950 (2020).
57. Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics* 8, 4 (2007).
58. Vita R, Overton JA, Greenbaum JA *et al.* The Immune Epitope Database (IEDB) 3.0. *Nucleic Acids Res.* 43(Database issue), D405–412 (2015).
59. Jespersen MC, Peters B, Nielsen M, Marcatili P. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Res.* 45(W1), W24–W29 (2017).
60. Kolaskar AS, Tongaonkar PC. A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS Lett.* 276(1–2), 172–174 (1990).
61. Emini EA, Hughes JV, Perlow DS, Boger J. Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. *J. Virol.* 55(3), 836–839 (1985).
62. Chou PY, Fasman GD. Prediction of the secondary structure of proteins from their amino acid sequence. *Adv. Enzymol. Relat. Areas Mol. Biol.* 47, 45–148 (1978).
63. Vihinen M, Torkkila E, Riikonen P. Accuracy of protein flexibility predictions. *Proteins* 19(2), 141–149 (1994).
64. Larsen JE, Lund O, Nielsen M. Improved method for predicting linear B-cell epitopes. *Immunome Res.* 2, 2 (2006).
65. Ponomarenko J, Bui HH, Li W *et al.* ElliPro: a new structure-based tool for the prediction of antibody epitopes. *BMC Bioinformatics* 9, 514 (2008).
66. Larsen MV, Lundegaard C, Lamberth K, Buus S, Lund O, Nielsen M. Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. *BMC Bioinformatics* 8, 424 (2007).
67. Andreatta M, Nielsen M. Gapped sequence alignment using artificial neural networks: application to the MHC class I system. *Bioinformatics* 32(4), 511–517 (2016).
68. Jensen KK, Andreatta M, Marcatili P *et al.* Improved methods for predicting peptide binding affinity to MHC class II molecules. *Immunology* 154(3), 394–406 (2018).
69. Ettayapuram Ramaprasad AS, Singh S, Gajendra PSR, Venkatesan S. AntiAngioPred: a server for prediction of anti-angiogenic peptides. *PLoS ONE* 10(9), e0136990 (2015).
70. Naamati G, Askenazi M, Linial M. ClanTox: a classifier of short animal toxins. *Nucleic Acids Res.* 37(Web Server issue), W363–W368 (2009).
71. Maupetit J, Derreumaux P, Tuffery P. PEP-FOLD: an online resource for *de novo* peptide structure prediction. *Nucleic Acids Res.* 37(Web Server issue), W498–W503 (2009).
72. Lan J, Ge J, Yu J *et al.* Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* 581(7807), 215–220 (2020).
73. Sheik SS, Sundararajan P, Hussain AS, Sekar K. Ramachandran plot on the web. *Bioinformatics* 18(11), 1548–1549 (2002).
74. Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* 35(Web Server issue), W407–W410 (2007).
75. Benkert P, Tosatto SC, Schomburg D. QMEAN: a comprehensive scoring function for model quality assessment. *Proteins* 71(1), 261–277 (2008).
76. Williams CJ, Headd JJ, Moriarty NW *et al.* MolProbity: more and better reference data for improved all-atom structure validation. *Protein Sci.* 27(1), 293–315 (2018).

77. Colovos C, Yeates TO. Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci.* 2(9), 1511–1519 (1993).
78. Van Zundert GCP, Rodrigues J, Trellet M *et al.* The HADDOCK2.2 web server: user-friendly integrative modeling of biomolecular complexes. *J. Mol. Biol.* 428(4), 720–725 (2016).
79. Laskowski RA, Jablonska J, Pravda L, Varekova RS, Thornton JM. PDBsum: structural summaries of PDB entries. *Protein Sci.* 27(1), 129–134 (2018).
80. Cui J, Li F, Shi ZL. Origin and evolution of pathogenic coronaviruses. *Nat. Rev. Microbiol.* 17(3), 181–192 (2019).
81. Marsh M, Helenius A. Virus entry: open sesame. *Cell* 124(4), 729–740 (2006).
82. Maenaka K, Jones EY. MHC superfamily structure and the immune system. *Curr. Opin. Struct. Biol.* 9(6), 745–753 (1999).
83. Toor JS, Rao AA, Mcshan AC *et al.* A recurrent mutation in anaplastic lymphoma kinase with distinct neoepitope conformations. *Front. Immunol.* 9, 99 (2018).
84. Li Y, Li H, Martin R, Mariuzza RA. Structural basis for the binding of an immunodominant peptide from myelin basic protein in different registers by two HLA-DR2 proteins. *J. Mol. Biol.* 304(2), 177–188 (2000).
85. Dyer O. Covid-19: denmark to kill 17 million minks over mutation that could undermine vaccine effort. *BMJ* 371, m4338 (2020).
86. Tegally H, Wilkinson E, Lessells RJ *et al.* Sixteen novel lineages of SARS-CoV-2 in South Africa. *Nat. Med.* doi:10.1038/s41591-021-01255-3 (2021) (Epub ahead of print).
87. Liu Q, Zhou YH, Yang ZQ. The cytokine storm of severe influenza and development of immunomodulatory therapy. *Cell. Mol. Immunol.* 13(1), 3–10 (2016).
88. Graham RL, Baric RS. Recombination, reservoirs, and the modular spike: mechanisms of coronavirus cross-species transmission. *J. Virol.* 84(7), 3134–3146 (2010).
89. Dawson P, Malik MR, Parvez F, Morse SS. What have we learned about Middle East respiratory syndrome coronavirus emergence in humans? A systematic literature review. *Vector Borne Zoonotic Dis.* 19(3), 174–192 (2019).
90. Xie J, Li Y, Shen X *et al.* Dampened STING-dependent interferon activation in bats. *Cell Host Microbe* 23(3), 297–301 e294 (2018).



Contact us

Editorial Department

Senior Editor
Abigail Sawyer
asawyer@biotechniques.com

Business Development and Support

Commercial Director
Sarah Mayes
s.mayes@future-science-group.com

TECAN.