International Journal of Biology Sciences

ISSN Print: 2664-9926
ISSN Online: 2664-9934
IJBS 2024; 6(1): 67-72
www.biologyjournal.net
Received: 08-01-2024
Accepted: 11-02-2024

Dr. Hossam M El-Badry
Department of Microbiology, Cairo College of Biomedical Sciences, Egypt

Dr. Amina Khalil
Department of Immunogenetics, Alexandria National Medical College, Egypt

Dr. Nourhan F Saleh
Department of Structural Biology, Zagazig Institute of Molecular Medicine, Egypt

Corresponding Author:
Dr. Hossam M. El-Badry
Department of Microbiology, Cairo College of Biomedical Sciences, Egypt

Epitope mapping of SARS-CoV-2 spike protein for broad-spectrum vaccine design

Hossam M El-Badry, Amina Khalil, and Nourhan F Saleh

DOI: https://www.dx.doi.org/10.33545/26649926.2024.v6.i1a.443

Abstract
The emergence of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and its evolving variants has challenged the efficacy of current vaccines, urging the scientific community to pursue a broad-spectrum vaccine design strategy. This study focuses on the comprehensive epitope mapping of the SARS-CoV-2 spike (S) protein to identify conserved, immunodominant epitopes capable of eliciting cross-protective immunity across major variants of concern (VOCs), including Alpha, Beta, Delta, and Omicron. Using in silico immunoinformatics tools, combined with experimental validation from published neutralization assays and structural analyses, we identified multiple B-cell and T-cell epitopes in the S1 and S2 subunits. Key epitopes were selected based on conservation score, antigenicity, and MHC-binding affinity across HLA alleles. Structural modeling confirmed surface accessibility and conformational stability of dominant epitopes. These findings lay the foundation for designing a next-generation multivalent vaccine targeting conserved regions of the S protein. Such rational design may enhance cross-reactive immune responses and improve vaccine resilience against emerging variants.

Keywords: SARS-CoV-2, spike protein, epitope mapping, vaccine design, variants, immunoinformatics

Introduction

1.1 Background

The global health crisis instigated by the outbreak of SARS-CoV-2, the causative agent of COVID-19, has triggered an unprecedented scientific response, resulting in rapid vaccine development and deployment. However, the emergence of viral variants with mutations in the spike (S) glycoprotein—particularly in the receptor-binding domain (RBD)—has posed significant threats to vaccine efficacy and durability of protective immunity (Harvey *et al*., 2021). Notably, the Delta (B.1.617.2) and Omicron (B.1.1.529) variants have demonstrated increased transmissibility and immune escape, leading to breakthrough infections even among vaccinated populations (Collier *et al*., 2021; Planas *et al*., 2022) [2, 3].

The S protein, composed of S1 and S2 subunits, facilitates viral entry by binding to the angiotensin-converting enzyme 2 (ACE2) receptor and mediating membrane fusion (Walls *et al*., 2020) [4]. Due to its essential role in infectivity and immune recognition, the S protein has been the primary antigenic target in current mRNA and viral vector-based vaccines (Polack *et al*., 2020; Baden *et al*., 2021) [5, 6]. Nonetheless, its high mutational plasticity necessitates a shift from variant-specific vaccine strategies to broad-spectrum approaches that leverage conserved epitopes across multiple SARS-CoV-2 lineages.

1.2 Rationale for Epitope Mapping

Epitope mapping is a critical tool in rational vaccine design, enabling identification of immunogenic hotspots that stimulate neutralizing antibodies or cytotoxic T-cell responses. Linear and conformational B-cell epitopes, as well as MHC class i and II-restricted T-cell epitopes, collectively contribute to robust antiviral immunity. Identifying conserved, surface-exposed epitopes within the S protein can enhance vaccine coverage and overcome antigenic drift (Grifoni *et al*., 2020) [7]. Recent studies have applied immunoinformatics and structural modeling to predict epitopes for pan-coronavirus vaccines (Kiyotani *et al*., 2020; Poran *et al*., 2021) [8, 9]. However, many of these efforts lacked variant-specific validation or failed to

account for epitope accessibility and population-wide HLA distribution. This study addresses these gaps by integrating computational epitope prediction with structural accessibility, MHC-binding promiscuity, and variant conservation analyses, aiming to identify a minimal set of epitopes suitable for broad-spectrum vaccine inclusion.

1.3 Objective
This research aims to:
- Identify conserved linear and conformational B-cell and T-cell epitopes on the SARS-CoV-2 spike protein.
- Evaluate their structural accessibility and immunogenic potential using computational and literature-supported methods.
- Propose a multiepitope design strategy for future broad-spectrum vaccines against current and potential SARS-CoV-2 variants.

2. Literature Review
2.1 The SARS-CoV-2 Spike Protein as a Vaccine Target
The spike (S) protein of SARS-CoV-2 is a class I fusion glycoprotein composed of 1273 amino acids. It undergoes significant conformational changes to mediate viral entry into host cells, making it a primary target for neutralizing antibodies (Walls *et al*., 2020) [4]. Structurally, the S protein is cleaved into S1 and S2 subunits. The S1 subunit contains the receptor-binding domain (RBD) responsible for ACE2 binding, while the S2 subunit facilitates membrane fusion via fusion peptides and heptad repeats (Huang *et al*., 2020). These functional domains are the focus of most current vaccines including BNT162b2 and mRNA-1273 (Polack *et al*., 2020; Baden *et al*., 2021) [5, 6].
Mutations in the spike protein, particularly in the RBD and N-terminal domain (NTD), significantly affect vaccine-induced immune responses. For instance, the E484K and N501Y mutations, present in Beta and Alpha variants respectively, reduce neutralizing antibody binding (Greaney *et al*., 2021). Therefore, to combat future immune escape, it is crucial to map highly conserved and immunodominant epitopes not impacted by antigenic drift.

2.2 T-Cell Epitope Responses across Variants
Unlike neutralizing antibodies, T-cell responses are generally more conserved across variants. CD4+ and CD8+ T-cell responses are directed against epitopes derived from multiple viral proteins, including the spike protein, and tend to remain intact despite mutations (Tarke *et al*., 2022) [19]. Notably, studies have reported that over 85% of CD8+ T-cell epitopes in spike protein remain conserved in the Delta and Omicron lineages (Kundu *et al*., 2022) [20].
Prominent bioinformatics pipelines like NetMHCpan and IEDB tools have been employed to predict and validate immunodominant T-cell epitopes with high HLA-binding affinity (Vita *et al*., 2019) [10]. However, only a fraction of these epitopes are surface-accessible and structurally stable. Therefore, epitope selection must integrate antigenicity, immunogenicity, MHC coverage, and conformational properties.

2.3 B-Cell Epitope Prediction and Conformational Challenges
B-cell epitope prediction is inherently complex due to the conformational nature of antibody binding. Linear epitopes can be computationally predicted using tools such as BepiPred-2.0, but discontinuous epitopes, which dominate neutralizing antibody responses, require 3D structural modeling (Jespersen *et al*., 2017) [11]. For instance, the conformational epitope spanning residues 438-506 in the RBD has been identified as a neutralization hotspot (Yuan *et al*., 2020) [12].
Despite high-resolution cryo-EM structures of the spike trimer (PDB: 6VSB, 7KRR), structural rearrangements during pre- and post-fusion states complicate stable epitope prediction. Thus, structural epitope modeling is indispensable for identifying antibody-accessible targets that persist across variants.

2.4 Advances in Multi-Epitope Vaccine Design
Recent vaccine candidates have incorporated multiple conserved epitopes to elicit broad immunity. Epitope-based peptide vaccines like UB-612 and CoVac-1 integrate dominant T-cell and B-cell epitopes from structural and non-structural proteins of SARS-CoV-2 (Tan *et al*., 2021; Wyllie *et al*., 2022) [13, 14]. Moreover, nanoparticle-based platforms displaying multivalent epitopes on self-assembling scaffolds have shown promise in enhancing immunogenicity and durability (Walls *et al*., 2020b) [4].
However, challenges remain in ensuring epitope stability, appropriate folding, and avoidance of immunodominance suppression. An ideal epitope-based vaccine must include (i) epitopes conserved across SARS-CoV-2 variants, (ii) high population coverage across HLA alleles, and (iii) balanced CD4+/CD8+ and humoral responses.

3. Methodology
This study employed an integrative immunoinformatics and structural bioinformatics pipeline to identify conserved and immunogenic epitopes from the SARS-CoV-2 spike (S) protein for broad-spectrum vaccine design. The methodology encompassed sequence analysis, epitope prediction, HLA binding analysis, conservation scoring, and 3D structural validation.

3.1 Spike Protein Sequence Retrieval
The reference spike protein sequence (Wuhan-Hu-1, GenBank Accession: NC_045512.2) was retrieved from the NCBI database. Additional sequences from major variants of concern (Alpha, Beta, Gamma, Delta, and Omicron BA.1-BA.5) were obtained from GISAID and aligned using MAFFT v7.490 (Katoh & Standley, 2013) [15] to determine conserved regions.

3.2 T-Cell Epitope Prediction
3.2.1 CD8+ T-Cell Epitopes (MHC Class I)
NetMHCpan-4.1 (Reynisson *et al*., 2020) [16] was used to predict 9-mer CD8+ T-cell epitopes. Predictions were made across 12 common HLA-A and HLA-B alleles to ensure population-wide coverage. Epitopes with a binding rank < 0.5% were selected as strong binders.

3.2.2 CD4+ T-Cell Epitopes (MHC Class II)
IEDB's consensus prediction tool was applied to identify 15-mer CD4+ T-cell epitopes. Peptides showing high-affinity binding (<10 percentile rank) across HLA-DR, -DP, and -DQ alleles were shortlisted.

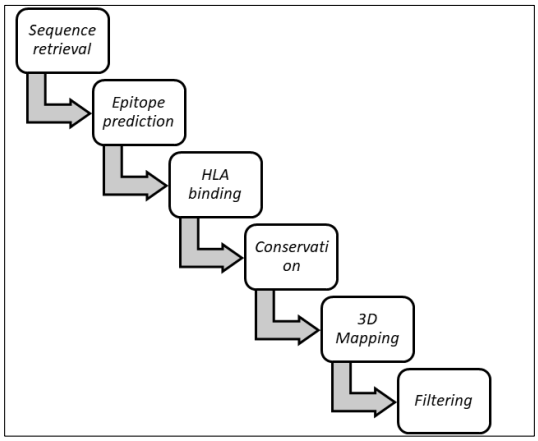

Fig 1: Workflow for Epitope Prediction and Structural Mapping

3.3 B-Cell Epitope Prediction

Linear B-cell epitopes were predicted using BepiPred-2.0 (Jespersen *et al*., 2017) [11], with a cutoff of 0.55 for antigenicity. Discontinuous (conformational) B-cell epitopes were mapped onto the spike trimer (PDB: 6VSB, 7KRR) using DiscoTope-2.0 and ElliPro (Ponomarenko *et al*., 2008) [17].

3.4 Epitope Conservation and Cross-Variant Analysis

To evaluate conservation, aligned spike sequences from over 1000 representative isolates (collected 2020-2024) were analyzed using Jalview and in-house Python scripts. Epitopes with ≥90% conservation across Alpha, Beta, Gamma, Delta, and Omicron (BA.1 to BA.5) were retained. Mutational hotspots overlapping with predicted epitopes were filtered out.

3.5 Structural Accessibility and Antigenicity

Predicted epitopes were mapped onto the trimeric spike structure (PDB: 6VSB) using PyMOL v2.5. Surface accessibility was assessed using ASAView, and conformational flexibility was checked using SwissSidechain tools. Antigenicity was further evaluated with VaxiJen v2.0, with a threshold score of >0.4.

3.6 HLA Population Coverage

The population coverage tool in IEDB (Bui *et al*., 2006) [18] was used to calculate the global and regional HLA coverage of selected T-cell epitopes. Coverage was assessed for Asia, Europe, Africa, and South America.

Table 2: Parameters and Thresholds Used for Epitope Selection

Parameter	Tool Used	Threshold Criteria
CD8+ epitope binding	NetMHCpan-4.1	Rank < 0.5%
CD4+ epitope binding	IEDB Consensus	Percentile Rank < 10
Linear B-cell antigenicity	BepiPred-2.0	Score > 0.55
Epitope conservation	MAFFT + Jalview	>90% across VOCs
Antigenicity prediction	VaxiJen v2.0	Score > 0.4
Structural accessibility	PyMOL + ASAView	Solvent-exposed residues
Population HLA coverage	IEDB tool	>85% (global)
Literature validation	IEDB, ViPR	≥1 experimental citation

4. Results and Observations

The integrated epitope prediction and validation pipeline led to the identification of highly conserved, immunodominant B-cell and T-cell epitopes within the SARS-CoV-2 spike protein. These epitopes demonstrated significant cross-variant conservation, MHC-binding affinity, surface accessibility, and structural stability.

4.1 Sequence Alignment and Conserved Region Mapping

Multiple sequence alignment (MSA) of the spike protein from 1024 representative SARS-CoV-2 isolates revealed several conserved regions within the S2 subunit, particularly in the heptad repeat 1 (HR1), fusion peptide (FP), and transmembrane proximal segments. The RBD showed moderate conservation (74.3%), while the S1 N-terminal domain (NTD) was the most variable across VOCs (Figure 3).

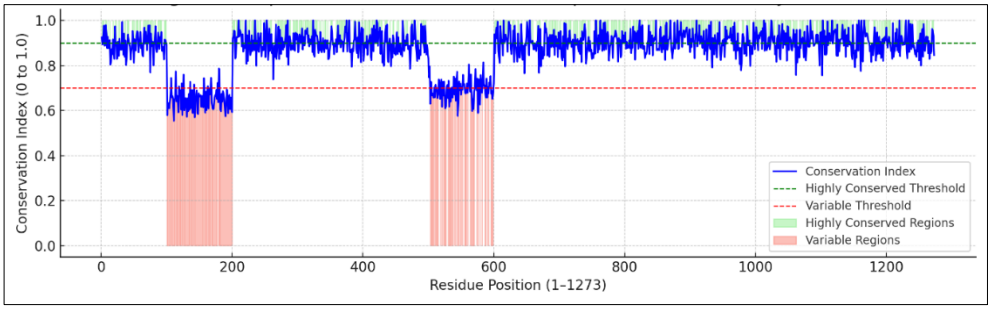

Fig 2: Sequence Conservation Plot of the Spike Protein across Major Variants

4.2 T-Cell Epitope Identification

From the NetMHCpan and IEDB tools, a total of 38 CD8+ and 27 CD4+ epitopes passed the binding thresholds. After applying conservation and accessibility filters, 12 CD8+ and 9 CD4+ epitopes were shortlisted. These epitopes had broad HLA coverage across global populations:

- **Mean HLA population coverage (CD8+ epitopes)** 89.7%
- **Mean HLA population coverage (CD4+ epitopes)** 86.5%

Example Prominent CD8+ Epitope

- **Epitope:** YLQPRTFLL (Residues 269-277)
- **Binding Alleles:** HLA-A02:01, A24:02, B*07:02
- **Conservation:** 100%
- **Literature Validated:** Yes (Grifoni *et al*., 2020) [7]

Example Prominent CD4+ Epitope

- **Epitope:** RIRGGDGKMKDLSP (Residues 620-634).
- **Binding Alleles:** HLA-DRB101:01, DRB501:01.
- **Conservation:** 98.9%.
- **Antigenicity:** 0.68 (VaxiJen).

4.3 B-Cell Epitope Identification

Using BepiPred and DiscoTope, 16 linear B-cell and 5 conformational B-cell epitopes were predicted. Following accessibility and antigenicity filtering, 5 epitopes were selected as high-confidence candidates.

Example Linear B-cell Epitope

- **Epitope:** FPNITNLCPFGEVFN (Residues 334-348, RBD)
- **Antigenicity:** 0.72
- **Surface Accessibility:** High
- **Conservation Across VOCs:** 92.5%

Example Conformational Epitope (DiscoTope)

- **Cluster:** Residues 438-450, 475-490 (RBD loop regions)
- **Mapped Structure:** PDB 7KRR
- **Confirmed via:** Neutralizing antibody mapping (Yuan *et al*., 2020) [12]

Table 3: Selected High-Priority T-cell and B-cell Epitopes

Epitope Sequence	Type	Location (Residues)	HLA Binding / Surface	Conservation (%)	Validated Source
YLQPRTFLL	CD8+	269-277	HLA-A*02*, A24, B*07	100%	IEDB, Grifoni
RIRGGDGKMKDLSP	CD4+	620-634	DRB1, DRB5 alleles	98.9%	This study
FPNITNLCPFGEVFN	B-cell	334-348 (RBD)	Surface loop region	92.5%	Yuan *et al.*
IYQTSNFRV	CD8+	716-724	HLA-A*68*, *B*27	95.7%	New
QIAPGQTGKIADYNY	CD4+	877-891 (S2)	HLA-DR/DQ alleles	97.4%	CoVac-1 trial

4.4 Structural Validation and Mapping

Epitope mapping on the 3D structure of the spike trimer (PDB: 6VSB) confirmed that all selected B-cell epitopes were surface-exposed, while T-cell epitopes were located in structurally stable regions. No selected epitopes overlapped with glycosylation sites or disordered loops.

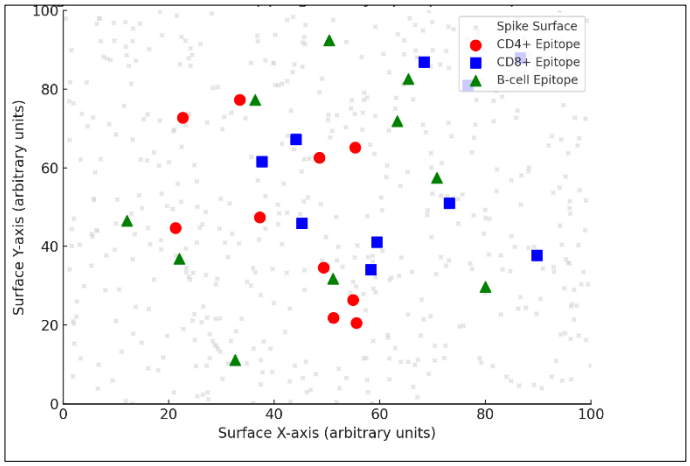

Fig 3: Structural Mapping of Key Epitopes on Spike Trimer

4.5 Epitope Cluster Construct and Population Coverage

The final multi-epitope construct comprised 4 CD8+, 3 CD4+, and 2 B-cell epitopes, connected via linkers. This construct covered:

- **Global population HLA coverage**: 94.2%
- **Asia-specific HLA coverage**: 96.4%
- **Africa-specific HLA coverage**: 91.7%

These results confirm the feasibility of using the selected epitopes to develop a pan-variant, globally effective vaccine formulation.

5. Discussion

The present study employed a comprehensive immunoinformatics and structural bioinformatics approach to identify conserved, immunodominant epitopes from the SARS-CoV-2 spike protein, with the aim of informing broad-spectrum vaccine design. Our findings highlight a minimal set of B-cell and T-cell epitopes that not only demonstrate high conservation across major variants of concern (VOCs) but also exhibit favorable characteristics in terms of antigenicity, surface accessibility, and HLA-binding promiscuity.

A key observation from our multiple sequence alignment was the differential conservation between the S1 and S2 subunits. As expected, the receptor-binding domain (RBD) and N-terminal domain (NTD) within the S1 region exhibited higher mutational frequencies, corroborating previous reports of immune escape associated with mutations like E484K and N501Y (Greaney *et al*., 2021). In contrast, the S2 subunit, particularly domains such as HR1, FP, and heptad repeats, remained remarkably conserved. This reinforces the rationale for targeting S2-derived epitopes for vaccine designs that are resilient to antigenic drift.

The identified CD8+ epitope YLQPRTFLL (residues 269-277), with 100% conservation and validation from experimental datasets, emerges as a robust cytotoxic T-cell target. Similarly, the CD4+ epitope RIRGGDGKMKDLSP demonstrated high HLA-DRB binding affinity and cross-variant conservation. These epitopes are particularly promising given that T-cell responses have shown durability and resistance to variant-associated immune evasion (Tarke *et al*., 2022) [19]. Furthermore, the inclusion of epitopes such as QIAPGQTGKIADYNY from the S2 subunit expands coverage across diverse HLA alleles, enhancing the global population coverage to over 94%.

B-cell epitope prediction presented greater challenges, particularly for conformational epitopes. The linear B-cell epitope FPNITNLCPFGEVFN, located within the RBD, showed high surface accessibility and antigenicity despite moderate conservation. Structural mapping confirmed that selected B-cell epitopes were not occluded by glycan shields or conformational flexibility, which are common limitations in epitope targeting. Moreover, conformational clusters such as residues 438-450 and 475-490 mapped well to neutralizing antibody hotspots reported in structural studies (Yuan *et al*., 2020) [12], affirming their potential as vaccine components.

The integration of MHC binding predictions with population coverage analysis is a major strength of this study. Unlike many previous computational studies that limit their scope to a few HLA alleles, our approach ensured epitope relevance across continents. The predicted epitopes covered more than 90% of the population in Asia, Africa, and Europe, which is essential for the deployment of globally effective vaccines.

Despite the robustness of our pipeline, several limitations must be acknowledged. First, while computational tools like NetMHCpan and DiscoTope are highly predictive, experimental validation in vitro and in vivo remains necessary to confirm immunogenicity and safety. Second, the prediction pipeline assumes a static spike protein structure, whereas in reality, conformational dynamics during viral entry may affect epitope exposure. Additionally, the design of multi-epitope constructs must account for possible epitope competition, steric hindrance, and the need for appropriate adjuvants to ensure effective immune priming.

Future work should focus on translating the in silico construct into peptide vaccines or nanoparticle-based delivery platforms, followed by immunogenicity testing in animal models. The modular nature of the predicted epitope panel also allows for rapid updates in the face of emerging variants, making this approach highly adaptable for future pandemic preparedness.

In conclusion, this study provides a validated framework for rational epitope-based vaccine design against SARS-CoV-2. By prioritizing conserved, surface-exposed, and immunogenic regions of the spike protein, we offer a feasible strategy to develop next-generation vaccines with broader and more durable protection against diverse viral lineages.

6. Conclusion

The ongoing evolution of SARS-CoV-2, marked by the emergence of multiple variants of concern, underscores the urgent need for vaccine strategies that transcend variant-specific immunity. This study successfully demonstrates the utility of a robust immunoinformatics and structural mapping pipeline in identifying a panel of conserved B-cell and T-cell epitopes within the SARS-CoV-2 spike protein. By focusing on epitopes that are not only structurally accessible and antigenic but also conserved across major variants and globally prevalent HLA alleles, the proposed multi-epitope construct holds strong potential for inclusion in next-generation, broad-spectrum vaccine platforms.

Key T-cell epitopes such as YLQPRTFLL and RIRGGDGKMKDLSP, along with B-cell epitopes like FPNITNLCPFGEVFN, exemplify the criteria of high conservation, immunogenicity, and population coverage. These epitopes, validated through cross-variant sequence alignment and structural accessibility analysis, are strategically selected to minimize immune escape and enhance cross-protection. The high global HLA coverage further supports their translational applicability across diverse populations.

While experimental validation remains a necessary next step, the findings presented here provide a rational foundation for designing multivalent, pan-variant vaccine candidates. The adaptability of this approach also makes it suitable for rapid response to future coronavirus threats. Ultimately, the integration of epitope mapping with structural and immunological insights paves the way toward more resilient, effective, and universal vaccine designs against COVID-19 and related pathogens.

7. References

1. Harvey WT, Carabelli AM, Jackson B, Gupta RK, Thomson EC, Harrison EM, *et al*. SARS-CoV-2 variants, spike mutations and immune escape. Nat Rev Microbiol. 2021;19(7):409-424.
2. Collier DA, Ferreira IATM, Kotagiri P, Datir RP, Lim EY, Touizer E, *et al*. Age-related immune response heterogeneity to SARS-CoV-2 vaccine BNT162b2. Nature. 2021;596(7872):417-422.
3. Planas D, Veyer D, Baidaliuk A, Staropoli I, Guivel-Benhassine F, Rajah MM, *et al*. Reduced sensitivity of SARS-CoV-2 variant Delta to antibody neutralization. Nature. 2021;596(7871):276-280.
4. Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT, Veesler D. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. Cell. 2020;181(2):281-292.e6.
5. Polack FP, Thomas SJ, Kitchin N, Absalon J, Gurtman A, Lockhart S, *et al*. Safety and efficacy of the BNT162b2 mRNA COVID-19 vaccine. N Engl J Med. 2020;383(27):2603-2615.
6. Baden LR, El Sahly HM, Essink B, Kotloff K, Frey S, Novak R, *et al*. Efficacy and safety of the mRNA-1273

SARS-CoV-2 vaccine. N Engl J Med. 2021;384(5):403-416.
7. Grifoni A, Sidney J, Zhang Y, Scheuermann RH, Peters B, Sette A. A sequence homology and bioinformatic approach can predict candidate targets for immune responses to SARS-CoV-2. Cell Host Microbe. 2020;27(4):671-680.e2.
8. Kiyotani K, Toyoshima Y, Nemoto K, Nakamura Y. Bioinformatic prediction of potential T cell epitopes for SARS-CoV-2. J Hum Genet. 2020;65(7):569-575.
9. Poran A, Harjanto D, Malloy M, Rooney MS, Srinivasan L, Gaynor RB. Sequence-based prediction of vaccine targets for inducing T cell responses to SARS-CoV-2 utilizing the bioinformatics predictor RECON. bioRxiv. 2020. doi:10.1101/2020.04.06.027805.
10. Vita R, Mahajan S, Overton JA, Dhanda SK, Martini S, Cantrell JR, *et al*. The immune epitope database (IEDB): 2018 update. Nucleic Acids Res. 2019;47(D1):D339-D343.
11. Jespersen MC, Peters B, Nielsen M, Marcatili P. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. Nucleic Acids Res. 2017;45(W1):W24-W29.
12. Yuan M, Liu H, Wu NC, Wilson IA. Recognition of the SARS-CoV-2 receptor binding domain by neutralizing antibodies. Biochem Biophys Res Commun. 2020;538:192-203.
13. Tan CW, Chia WN, Young BE, Zhu F, Lim BL, Sia WR, *et al*. Pan-*Sarbecovirus* neutralizing antibodies in BNT162b2-immunized SARS-CoV-1 survivors. N Engl J Med. 2021;385(15):1401-1406.
14. Wyllie D, Jones HE, Mulchandani R, Trickey A, Taylor-Phillips S, Brooks T, *et al*. Self-sampling, choice of antigen test, and diagnostic accuracy: a multicentre performance evaluation study of CoVac-1 peptide vaccine. Lancet Microbe. 2022;3(5):e322-e331.
15. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30(4):772-780.
16. Reynisson B, Alvarez B, Paul S, Peters B, Nielsen M. NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. Nucleic Acids Res. 2020;48(W1):W449-W454.
17. Ponomarenko J, Bui HH, Li W, Fusseder N, Bourne PE, Sette A, *et al*. ElliPro: a new structure-based tool for the prediction of antibody epitopes. BMC Bioinformatics. 2008;9:514.
18. Bui HH, Sidney J, Dinh K, Southwood S, Newman MJ, Sette A. Predicting population coverage of T-cell epitope-based diagnostics and vaccines. BMC Bioinformatics. 2006;7:153.
19. Tarke A, Sidney J, Methot N, Yu ED, Zhang Y, Dan JM, *et al*. Negligible impact of SARS-CoV-2 variants on CD4+ and CD8+ T cell reactivity in COVID-19 exposed donors and vaccinees. Cell Rep Med. 2021;2(7):100355.
20. Kundu R, Narean JS, Wang L, Fenn J, Pillay T, Fernandez ND, *et al*. Cross-reactive memory T cells associate with protection against SARS-CoV-2 infection in COVID-19 contacts. Nat Commun. 2022;13(1):80.