JOURNAL OF
BIOLOGICAL
INSIGHTS

Vol. 1| Issue 1|pp:94-109| 2025 |
DOI: 10.4197/jbi.1-1.5

Sharing Knowledge, Shaping the Future

Published by Scientific Publishing Center, King Abdulaziz University

Review Article

Received: 2025/11/10| **Accepted**: 2025/12/21| **Published**: 2025/12/29|

Handling Editor: Irfan A. Rather | Department of Biological Sciences| King Abdulaziz University |Jeddah| Saudi Arabia

MERS-CoV Spike Protein Variants: A Computational Perspective

Raneem S. alamri[1,*], Najla Ali Alburae[1,*]

[1]Department of Biological Sciences, Faculty of Science, King Abdulaziz University, Jeddah 21589, Saudi Arabia

*Corresponding authors: Ralamri0314@stu.kau.edu.sa (R.S.A)| nalbaurai@kau.edu.sa (N.A.A)

Abstract

Since its identification in 2012, Middle East Respiratory Syndrome Coronavirus (MERS-CoV) has remained a persistent zoonotic and public health threat, with a case fatality rate exceeding 35%. The viral spike (S) protein, responsible for facilitates host cell entry and immune recognition, represents the primary target for vaccine and therapeutic development. Comprehensive characterization of spike protein genetic variability is therefore crucial for anticipating viral evolution, assessing immune escape potential, and guiding the development of pan-coronavirus vaccine strategies. This review compiles recent *in silico* studies examining variations in the MERS-CoV spike protein. The methods summarized encompass sequence-based phylogenetic studies, structural modeling, epitope prediction algorithms, and machine learning tools employed to assess the effects of mutations. These approaches together offer swift and economical frameworks for worldwide monitoring and functional forecasting. Computational analyses have delineated conserved residues with potential for vaccine development, identified emerging variant lineages, and simulated the effects of spike mutations. Nonetheless, significant challenges remain, such as limited global sampling, a lack of experimental validation, and insufficient representation of recombination dynamics. Future initiatives should aim to better integrate computational and experimental methods, create AI-enhanced surveillance systems, and design more comprehensive pan-coronavirus immunogens.

Keywords: MERS-CoV | spike protein | computational biology | phylogenetics | structural modeling

1. Introduction

As a novel betacoronavirus, Middle East Respiratory Syndrome Coronavirus (MERS-CoV) has caused severe respiratory illnesses with high mortality rates in 2012 [1,2]. The MERS-CoV virus has been responsible for over 2,500 confirmed cases across 27 countries [3,4] The disease has been fatal in 30-40% of cases [5,6] , making it one of the deadliest human coronaviruses we have ever encountered [7]. Unlike SARS-CoV-2, MERS-CoV exhibits limited human-to-human transmission [8,9], but maintains persistent circulation in dromedary camels [10,11], providing a continuous reservoir zoonotic reservoir [12,13].

Spike (S) protein a 1,353 amino acid trimeric glycoprotein [14,15], plays a central role in MERS-CoV pathogenesis by binding dipeptidyl peptidase-4 (DPP4) receptors to facilitate virus entry into host cells [16–18]. CD26, which is also referred to as dipeptidyl peptidase-4 (DPP4), serves as the cellular receptor targeted by the MERS-CoV spike RBD. For the sake of clarity, this review consistently uses the term DPP4.

The spike protein is composed of two functional subunits: S1, which contains the receptor-binding domain (RBD) [19,20], and S2, which mediates membrane fusion [21,22]. Antibodies neutralize the spike protein because it is the most exposed viral antigen [20,23], which is why most vaccines are developed around it [24–26].

A viral spike protein's genetic variability affects its fitness, host range expansion, immune evasion, and therapeutic efficacy [27–29]. Mutations in the RBD may affect receptor binding affinity and specificity [18,30], potentially facilitating cross-species transmission and adaptation [30,31]. It is also possible for changes in neutralizing epitopes to compromise vaccine effectiveness or the binding of therapeutic antibodies [29]. Genetic diversity is essential for predicting viral evolution trajectories and designing robust countermeasures [32,33].

As a result of computing biology, large-scale sequence analyses, predictive modeling, and rapid hypothesis generation have revolutionized viral genetics research [28,34,35]. Due to advances in bioinformatic algorithms and machine learning approaches, genomic sequencing capacity has grown exponentially [36,37]. Data-driven methods now play a central role in tracking viral evolution and guiding countermeasures [35]. MERS-CoV offers unique advantages using data-driven approaches, including global sequence analysis, prediction of mutational effects [30,35], identification of conserved epitopes [35,38], and modelling of viral evolution [28,33]. Reviewing the pattern of genetic variability and structural-functional relationships of the MERS-CoV spike proteins, we aim to provide insight into the evolution of the virus and establish how it adapts. In addition, it outlines future directions for developing effective therapeutic strategies, vaccines, and data-driven approaches to monitor Coronavirus evolution.

2. Materials and Methods

This review is a narrative compiled using PubMed, Google Scholar with search terms including "MERS-CoV", "Middle East Respiratory Syndrome", "spike protein", "computational", "phylogenetic", "structural modeling", "epitope prediction", and "machine learning". Some of the work cited is included for its methodological contribution only. AI was used in language polishing.

3. Spike Protein Genetic Variants

There have been over 300 complete spike protein sequences deposited in public databases since the first MERS-CoV genome was sequenced in 2012 [2,8,28]. Five major lineages have been identified based on phylogenetic analyses (A, B, C, D, and E) with distinct geographic distributions and evolutionary histories [8,13,25].

3.1 Variable domains

Genetic variation in the MERS-CoV spike protein is heterogeneous, with elevated mutation rates in several regions [28,33].

3.1.1 Receptor-Binding Domain (RBD, residues 484-567)

Approximately 2-3 times more mutations are accumulating in the RBD than in the overall spike protein [27,28], making it one of the most variable regions of the spike protein [18,19]. The DPP4 receptor is directly contacted at positions 509, 511, and 526 of the variable sites [16,18]. Some mutations such as D510G, R511H, and A522V alter receptor binding affinity and may affect host adaptation[30,31].

3.1.2 N-terminal Domain (NTD, residues 18-353)

In contrast to the RBD, the NTD contains several hypervariable loops that may serve as immune escape points[39]. Multiple lineages have shown clade-specific deletions and insertions in loops 1 and 3, which modify antigenic surface and contribute to immune evasion [8,25,39], potentially affecting antibody recognition.

3.1.3 Fusion peptide and heptad repeats: There is moderate variability in the fusion machinery [22,40], with mutations affecting the heptad repeat regions HR1 (residues 984-1104) and HR2 (residues 1246-1295)[21,41]. Despite strong functional constraints, the fusion peptide (residues 955-885) remains highly conserved[40,41].

3.2 Conserved Regions

In spite of overall diversity, several spike protein regions remain highly conserved as shown in Table 1.

- **Furin cleavage site (residues 748-751):** All sequenced isolates possess the RXXR motif [42,43]
- **Fusion peptide core (residues 955-975):** Maintaining >95% sequence identity[40,41]
- **Transmembrane domain (residues 1296-1318):** There are only synonym substitutions, indicating that it is highly conserved [14]
- **Cytoplasmic tail (residues 1319-1353):** Signals associated with trafficking and assembly are conserved [44].

Table 1. Summary of conserved and variable regions in MERS-CoV spike protein

Residue range	Conserved regions	Varibal regions
residues 18-353		√
residues 484-567		√
residues 984-1104		√
residues 955-885		√
residues 1246-1295		√
residues 748-751	√	
residues 955-975	√	
residues 1296-1318	√	

3.3 Temporal evolution patterns

While numerous prior studies have characterized the MERS-CoV spike gene as being mainly limited by purifying selection, this view does not entirely reflect the intricate nature of coronavirus evolutionary dynamics. Global dN/dS ratios frequently fail to accurately reflect adaptive changes because a small number of positively selected sites can be masked by numerous structurally conserved residues. Additionally, coronaviruses experience episodic adaptation bursts rather than steady directional evolution. The recent surge in available camel and human spike sequences has uncovered more RBD variations, especially at residues that interact with DPP4, indicating continuous refinement of receptor interactions. Furthermore, research on structural and molecular dynamics has repeatedly demonstrated that certain substitutions can impact binding affinity and conformational stability, even when the traditional selection criteria remain low. The latest evidence proposes a model in which the spike protein is predominantly conserved at a global level; however, it faces localized, intermittent, and functionally important adaptive pressures, especially within the receptor-binding domain.[13,28,30,45].

3.4 Recombination in MERS-CoV

Recombination is a significant evolutionary mechanism influencing the diversity of the MERS-CoV spike protein and must be considered when analyzing its genetic composition. Multiple independent studies have revealed that the spike gene, particularly the S1 region, harbors several recombination breakpoints originating from co-circulating camel coronaviruses.[25]. offered the initial definitive proof of recombination among lineages present in Saudi Arabia, uncovering mosaic genomes created through exchanges between different parental viral strains. Since this discovery, both phylogenetic analyses and breakpoint-scanning research have verified that recombination plays a role in the development of new spike configurations, speeds up antigenic diversification, and aids in the introduction of mutations that would not occur solely through gradual point substitutions. Including recombination in evolutionary evaluations is crucial because it affects lineage development, host adaptation, and the potential for transmission across species.[13,25,46].

3.5 Camel–human adaptation signatures

Adaptation to specific hosts also plays a role in the diversity of spike proteins figure 1, especially in the receptor-binding domain, where sequence variations indicate optimization for species-specific DPP4 variants. Structural comparisons revealed differences in DPP4 contact residues between camels and humans, and certain mutations in the RBD, such as D510G, R511H, and A522V, influenced the binding affinity depending on the host. Phylogenetic studies have further indicated that many changes in the spike protein originate in camels, the main reservoir, with only some being retained during occasional transmission to humans[13,30,31]. The results of the study highlight the importance of incorporating indicators of camel–human adaptation when analyzing the development of the MERS-CoV spike.

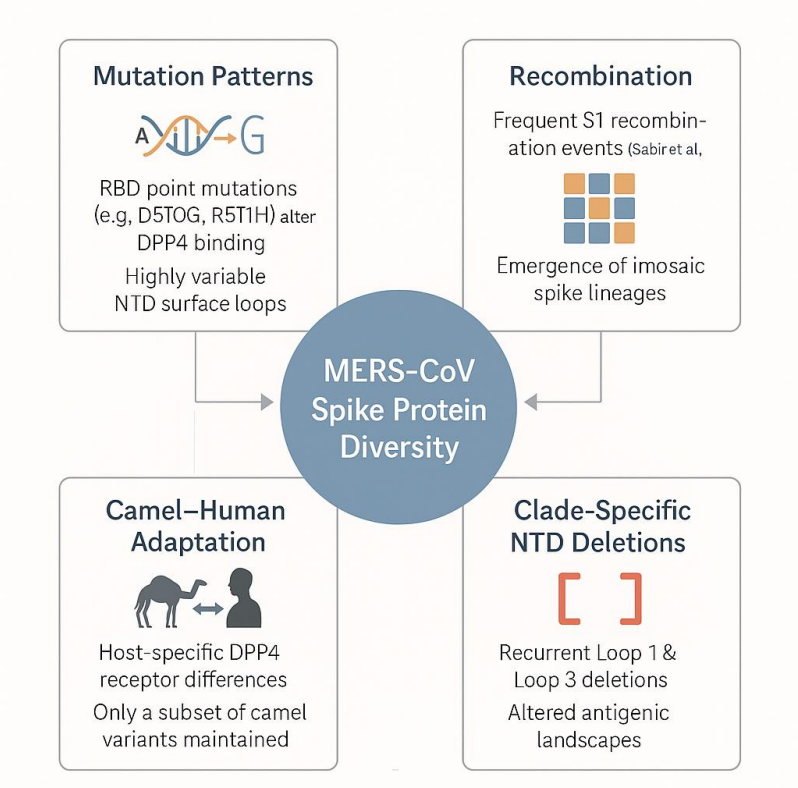

Figure 1. Evolutionary and genetic drivers of MERS-CoV spike protein diversity.

4. Computational approaches

4.1 Sequence analysis methods

4.1.1 Multiple sequence alignment (MSA)

To align MERS-CoV spike sequences, standard MSA algorithms such as MUSCLE, ClustalW, and MAFFT have been extensively used[34,35]. The alignment of hypervariable regions has been improved using advanced methods such as PRANK and GUIDANCE. In order to identify conserved motifs and functional domains, profile-based methods using Hidden Markov Models (HMMs) have been particularly effective[36,37].

4.1.2 Mutation detection and tracking: The development of real-time mutation surveillance pipelines incorporated tools such as GISAID's mutation tracking system and custom databases for MERS-CoV variants. A set of automated algorithms identifies emerging mutations, calculates their frequencies, and flags any potential significant changes based on structural or functional predictions[35,47].

4.2 Phylogenetic analysis

4.2.1 Maximum likelihood methods

MERS-CoV phylogenies have been constructed primarily using RAxML, IQ-TREE, and PhyML [8,13,48,49]. BEAST-based time-calibrated phylogenies have provided insights into evolutionary timescales and transmission dynamics[13]. For phylogenetic uncertainty and estimating confidence intervals, Bayesian approaches have proved especially useful[8,25].

4.2.2 Phylogeographic analysis

Geographic metadata and phylogenetic data were integrated to reveal patterns of viral spread and transmission [8,13]. A discrete phylogeography package such as BEAST has been used to map viral migration routes between nations and continents [13,25].

4.3 Structural modeling

4.3.1 Homology modeling

High-resolution crystal structures of MERS-CoV spike domains have enabled accurate homology modeling of variant sequences [15,18,19]. Mutations have been predicted using SWISS-MODEL, Phyre2, and AlphaFold2 [34,35,38].

4.3.2 Molecular dynamics simulations

The kinetics of receptor binding and protein stability have been examined using MD simulations using GROMACS, AMBER, and NAMD [18,30].

4.3.3 Protein-protein interaction modeling

Modeling spike-DPP4 interactions with docking algorithms (HADDOCK, ClusPro, Rosetta) has provided insight into how mutations affect binding affinity and specificity [16,18,30,31].

4.4 Epitope prediction

4.4.1 B-cell epitope prediction

The spike protein has been identified as a potential antibody binding site by linear and conformational B-cell epitope prediction tools such as BepiPred, ABCpred, and DiscoTope [34,35,47]. Improved prediction accuracy has been achieved through the use of support vector machines and neural networks [36,37,50].

4.4.2 T-cell epitope prediction

A number of MHC class I and II binding prediction algorithms (NetMHCpan, IEDB tools) have been developed to identify potential T-cell epitopes for different HLA allotypes [37,50–52]. The applicability of predicted epitopes has been assessed based on population coverage analysis [34,35,38].

4.4.3 Epitope conservation analysis

In order to identify potential targets for pan-coronavirus vaccines, sequence-based and structure-based approaches were used to assess cross-coronavirus epitope conservation [53,54].

4.5 Machine learning and artificial intelligence

4.5.1 Deep learning approaches

The prediction of mutation effects, epitope immunogenicity, and viral fitness has been achieved with the use of convolutional neural networks and transformer models [50,52]. In order to predict the functional impact of novel mutations, large language models trained on protein sequences show promise [55].

4.5.2 Ensemble methods.

A combination of random forests and gradient boosting algorithms can predict mutation pathogenicity and immune escape potential by integrating multiple features (sequence, structure, evolution) [36–38].

5. Insights from computational studies

5.1 Viral evolution and adaptation

Based on bioinformatic studies, MERS-CoV spike protein evolution is primarily shaped by purifying selection [33,48], with dN/dS ratios consistently below 1.0 for most spike regions [45]. There is, however, evidence of positive selection at specific locations in the RBD, particularly those that are in contact with the DPP4 receptor [18,30]. Therefore, it is likely that overall structural integrity will be maintained while receptor binding is optimized [30,31].see Table 2

Based on phylogenetic analyses, MERS-CoV evolution follows a "ladder-like" pattern rather than the "star-like" phylogenies observed in rapidly spreading viruses like SARS-CoV-2. [13,55] It is evident that this is due to the limited human-to-human transmission capacity of MERS-CoV. This suggests that the majority of evolutionary changes in the camel population are due to zoonotic maintenance during this period [8–11,13].

5.2 Structural-functional relationships

In molecular dynamics simulations, key insights have been obtained about how mutations affect spike protein function. The following are some of the key findings:

- **Receptor binding affinity:** A D510G or R511H mutation in the RBD changes the electrostatic landscape of the receptor-binding interface [16,18], resulting in differences in DPP4 binding affinity ranging from a 2-fold decrease to a 3-fold increase [30]
- **Conformational stability:** Mutations can shift the equilibrium between "closed" and "open" conformations of the spike protein [15,56]. Variants with an increased propensity to open may be more infectious but less stable [14,56]
- **Allosteric effects:** Considering the entire spike protein structure is important because mutations outside the RBD can affect receptor binding through long-range allosteric mechanisms [30,56].

5.3 Immune escape predictions

Researchers have identified several areas of potential immune vulnerability based on epitope mapping studies:

5.3.1 Conserved neutralizing epitopes

Several studies have identified highly conserved epitopes in the RBD and stem region that remain accessible across variants [20,23,29]. The DPP4-binding interface and cryptic epitopes exposed during conformational changes are examples of these [15,16,18,56].

5.3.2 Variable epitopes

The regions with high mutation rates often overlap with predicted B-cell epitopes[35,48], indicating immune-driven evolution. In contrast to other viruses, MERS-CoV has relatively limited genetic diversity [33,48].

5.3.3 Cross-reactive epitopes

Epitopes that share epitopes with other coronaviruses have been identified using bioinformatic approaches [35,53,54], which have the potential to serve as targets for the development of pan-coronavirus vaccines [26,57].

5.4 Host Range and zoonotic potential

Based on a comparative analysis of spike-receptor interactions across species, it has been found that :

- **DPP4 Conservation:** DPP4 is highly conserved across mammalian species, explaining MERS-CoV's broad host range.
- **Species-Specific Adaptations:** The transmission patterns between camels and humans may be explained by certain mutations that optimize binding to specific DPP4 variants in the host[10,30,31].
- **Barrier Identification:** Molecular models have identified amino acid positions that may serve as barriers to cross-species transmission[30,58], allowing surveillance to be targeted.

Table 1. Summary of computational approaches, functions, and key findings in MERS-CoV spike protein studies.

Computational Approach	Tools / Methods	Function / Application	Key Findings / Insights	References
Sequence Analysis	MUSCLE, ClustalW, MAFFT, PRANK, GUIDANCE, HMMER	Align spike protein sequences; identify conserved motifs and variable regions	Revealed high variability in RBD and NTD; conserved regions in fusion peptide and transmembrane domains	[34–36]
Mutation Tracking	GISAID pipeline, custom MERS databases	Detect and monitor emerging spike mutations	Tracked D510G, R511H, and A522V mutations affecting receptor binding and host adaptation	[35,47]
Phylogenetic Analysis	RAxML, IQ-TREE, PhyML, BEAST	Construct phylogenetic trees and infer evolutionary relationships	Identified five major lineages (A–E); evolution dominated by purifying selection with adaptive hotspots in RBD	[8,13,48]
Phylogeographic Mapping	BEAST discrete model	Integrate phylogeny with geographic metadata	Mapped MERS-CoV spread between countries; confirmed camel-to-human maintenance cycle	[13,25]
Structural Modeling	SWISS-MODEL, Phyre2, AlphaFold2	Build 3D models of spike variants	Explained structural effects of RBD mutations on receptor	[18,19,34]

			binding and spike conformation	
Molecular Dynamics (MD)	GROMACS, AMBER, NAMD	Simulate spike–receptor interactions and structural stability	Showed mutations alter conformational stability and receptor affinity	[15,30,56]
Protein–Protein Docking	HADDOCK, ClusPro, Rosetta	Model spike–DPP4 binding	Quantified mutation-driven changes in binding affinity across species	[16,18,31]
B-cell Epitope Prediction	BepiPred, ABCpred, DiscoTope	Predict linear and conformational antibody-binding sites	Identified variable loops overlapping with neutralizing epitopes	[34,35]
T-cell Epitope Prediction	NetMHCpan, IEDB, NetMHCIIpan	Predict T-cell epitopes and HLA binding	Discovered conserved epitopes suitable for multi-epitope vaccine design	[35,37,52]
Epitope Conservation Analysis	Sequence/structure-based tools	Assess conservation across coronaviruses	Revealed cross-reactive epitopes for pan-coronavirus vaccine design	[53,54]
Machine Learning / AI	CNNs, Random Forest, Gradient Boosting	Predict mutation pathogenicity and immune escape	Improved accuracy in predicting immunogenicity and variant fitness	[36,38]
Large Language Models (LLMs)	ProtBERT, ESM, transformer models	Predict functional effects of novel mutations	Emerging tools showing strong potential for mutation impact prediction	[52,55]
Cloud / Automated Pipelines	Cloud genomic platforms	Enable real-time variant detection and risk assessment	Accelerated global genomic surveillance and analysis	[35]

5.5 Enhanced Computational Infrastructure

5.5.1 Cloud-based platforms

A scalable cloud computing platform could enable broader participation in surveillance activities by democratizing access to computationally intensive analyses [35].

5.5.2 Automated pipelines

By automating analysis pipelines, characterization of variants could be completed in a matter of hours rather than weeks [35].

5.5.3 Data integration Hubs

The generation and testing of hypotheses could be accelerated using centralized platforms that integrate genomic, structural, and functional data [34,35].

5.6 Pan-Coronavirus approaches

5.6.1 Universal vaccine design

A computational approach could be used to identify conserved epitopes across coronavirus genera [35,53,54], which could lead to the development of broadly protective vaccines [26,57].

5.6.2 Cross-species modeling: Incorporating host species diversity into advanced models could improve the prediction of zoonotic potential and spillover risks [30,58,59].

5.6.3 Phylogenetic forecasting: Through machine learning, proactive countermeasures could be developed by forecasting likely evolutionary trajectories [33,55].

5.7 Experimental integration

5.7.1 High-throughput validation

It would be possible to systematically validate bioinformatic predictions by deep mutational scanning and other high-throughput experimental techniques [35,47].

5.7.2 Computational-experimental feedback loops

A continuous improvement in computational accuracy could be achieved through iterative cycles of prediction, testing, and model refinement [34,35].

5.7.3 Standardized benchmarks: By developing standardized datasets and benchmarks, data-driven methods can be evaluated and improved systematically [36,37].

6. Conclusions

Computational methodologies have profoundly altered our comprehension of MERS-CoV spike protein variants, elucidating evolutionary trajectories, pinpointing conserved intervention targets, and forecasting the functional impacts of mutations. Through the utilization of sequence analysis, structural modeling, and machine learning techniques, scholars have delineated the landscape of spike protein variability and identified both susceptible and stable domain that may function as therapeutic targets. The discovery of highly conserved epitopes within the receptor-binding domain and fusion machinery presents promising avenues for the development of pan-coronavirus vaccines. Similarly, computational forecasts regarding the effects of mutations on receptor affinity and immune recognition have shaped risk assessment frameworks for newly emerging variants. Nevertheless, considerable constraints persist within the realm of contemporary data-driven methodologies. The sparse and geographically skewed sequence data, the limited experimental corroboration of predictions, and the complexities involved in modeling intricate evolutionary phenomena such as recombination all inhibit our understanding. Moreover, the integration of molecular-level predictions with population-level dynamics is yet to be fully realized. Looking ahead, the confluence of sophisticated artificial intelligence techniques with enhanced surveillance systems and high-throughput experimental validation presents substantial promise as a complomentry to experiments done with either bioinformatics or wet lab. The establishment of real-time variant assessment frameworks, platforms for pan-coronavirus vaccine design, and predictive evolutionary models could substantially transform our readiness for future coronavirus threats. Ultimately, computational biology should not be perceived as a substitute for experimental methodologies, but rather as an enhancer and navigator for hypothesis generation and examination. The most significant advancements will emerge from the close integration of computational predictions with empirical validation, fostering virtuous cycles of discovery that deepen our understanding of

viral evolution and inform the development of robust countermeasures. The COVID-19 pandemic has underscored the paramount significance of swift and precise variant assessment. Drawing upon insights gained from SARS-CoV-2, the coronavirus research community is strategically positioned to cultivate next-generation computational instruments capable of rapidly characterizing emerging MERS-CoV variants and directing public health interventions. The groundwork established by existing computational investigations of MERS-CoV spike variants serves as a pivotal stepping stone toward achieving this objective.

Author Contributions

Raneem S. alamri perfomerd the Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization. Najla Ali Alburae carried out supervision and review. AI tools were used for improving language and grammar.

Funding

No external funding was received for this study.

Institutional Review Board Statement

This study did not involve human participants or animal subjects.

Informed Consent Statement

Not applicable.

Data Availability Statement

The data supporting this study are included within the manuscript.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] De Groot RJ, Baker SC, Baric RS, Brown CS, Drosten C, Enjuanes L, et al. Commentary: Middle East Respiratory Syndrome Coronavirus (MERS-CoV): Announcement of the Coronavirus Study Group. J Virol 2013;87:7790–2. https://doi.org/10.1128/JVI.01244-13.

[2] Zaki Ali M., van Boheemen Sander, Bestebroer Theo M., Osterhaus Albert D.M.E., Fouchier Ron A.M. Isolation of a Novel Coronavirus from a Man with Pneumonia in Saudi Arabia. New England Journal of Medicine n.d.;367:1814–20. https://doi.org/10.1056/NEJMoa1211721.

[3] Memish ZA, Perlman S, Van Kerkhove MD, Zumla A. Middle East respiratory syndrome. The Lancet 2020;395:1063–77. https://doi.org/10.1016/S0140-6736(19)33221-0.

[4] Sharif-Yakan A, Kanj SS. Emergence of MERS-CoV in the Middle East: Origins, Transmission, Treatment, and Perspectives. PLOS Pathogens 2014;10:e1004457. https://doi.org/10.1371/journal.ppat.1004457.

[5] Hui DS, Azhar EI, Kim Y-J, Memish ZA, Oh M, Zumla A. Middle East respiratory syndrome coronavirus: risk factors and determinants of primary, household, and nosocomial transmission. The Lancet Infectious Diseases 2018;18:e217–27. https://doi.org/10.1016/S1473-3099(18)30127-0.

[6] M.S N, M.A B, S.A. M, M.S. A, W.A. Z. MERS-CoV infection: epidemiology, pathogenesis and clinical characteristics 2018.

[7] Chafekar A, Fielding BC. MERS-CoV: Understanding the Latest Human Coronavirus Threat. Viruses 2018;10:93. https://doi.org/10.3390/v10020093.

[8] Cotten M, Watson SJ, Zumla AI, Makhdoom HQ, Palser AL, Ong SH, et al. Spread, Circulation, and Evolution of the Middle East Respiratory Syndrome Coronavirus. mBio 2014;5:e01062-13. https://doi.org/10.1128/mBio.01062-13.

[9] Drosten C, Meyer B, Müller MA, Corman VM, Al-Masri M, Hossain R, et al. Transmission of MERS-Coronavirus in Household Contacts. N Engl J Med 2014;371:828–35. https://doi.org/10.1056/NEJMoa1405858.

[10] Hemida MG, Perera RA, Wang P, Alhammadi MA, Siu LY, Li M, et al. Middle East Respiratory Syndrome (MERS) coronavirus seroprevalence in domestic livestock in Saudi Arabia, 2010 to 2013. Euro Surveill 2013;18:20659. https://doi.org/10.2807/1560-7917.es2013.18.50.20659.

[11] Reusken CB, Haagmans BL, Müller MA, Gutierrez C, Godeke G-J, Meyer B, et al. Middle East respiratory syndrome coronavirus neutralising serum antibodies in dromedary camels: a comparative serological study. The Lancet Infectious Diseases 2013;13:859–66. https://doi.org/10.1016/S1473-3099(13)70164-6.

[12] Azhar EI, Hui DSC, Memish ZA, Drosten C, Zumla A. The Middle East Respiratory Syndrome (MERS). Infectious Disease Clinics of North America 2019;33:891–905. https://doi.org/10.1016/j.idc.2019.08.001.

[13] Dudas G, Carvalho LM, Rambaut A, Bedford T. MERS-CoV spillover at the camel-human interface. eLife 2018;7:e31257. https://doi.org/10.7554/eLife.31257.

[14] Pallesen J, Wang N, Corbett KS, Wrapp D, Kirchdoerfer RN, Turner HL, et al. Immunogenicity and structures of a rationally designed prefusion MERS-CoV spike antigen. Proceedings of the National Academy of Sciences 2017;114:E7348–57. https://doi.org/10.1073/pnas.1707304114.

[15] Yuan Y, Cao D, Zhang Y, Ma J, Qi J, Wang Q, et al. Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor binding domains. Nature Communications 2017;8:15092. https://doi.org/10.1038/ncomms15092.

[16] Lu G, Hu Y, Wang Q, Qi J, Gao F, Li Y, et al. Molecular basis of binding between novel human coronavirus MERS-CoV and its receptor CD26. Nature 2013;500:227–31. https://doi.org/10.1038/nature12328.

[17] Raj VS, Mou H, Smits SL, Dekkers DHW, Müller MA, Dijkman R, et al. Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. Nature 2013;495:251–4. https://doi.org/10.1038/nature12005.

[18] Wang N, Shi X, Jiang L, Zhang S, Wang D, Tong P, et al. Structure of MERS-CoV spike receptor-binding domain complexed with human receptor DPP4. Cell Res 2013;23:986–93. https://doi.org/10.1038/cr.2013.92.

[19] Chen Y, Rajashankar KR, Yang Y, Agnihothram SS, Liu C, Lin Y-L, et al. Crystal Structure of the Receptor-Binding Domain from Newly Emerged Middle East Respiratory Syndrome Coronavirus. J Virol 2013;87:10777–83. https://doi.org/10.1128/JVI.01756-13.

[20] Mou H, Raj VS, Van Kuppeveld FJM, Rottier PJM, Haagmans BL, Bosch BJ. The Receptor Binding Domain of the New Middle East Respiratory Syndrome Coronavirus Maps to a 231-Residue Region in the Spike Protein That Efficiently Elicits Neutralizing Antibodies. J Virol 2013;87:9379–83. https://doi.org/10.1128/JVI.01277-13.

[21] Bosch BJ, Van Der Zee R, De Haan CAM, Rottier PJM. The Coronavirus Spike Protein Is a Class I Virus Fusion Protein: Structural and Functional Characterization of the Fusion Core Complex. J Virol 2003;77:8801–11. https://doi.org/10.1128/JVI.77.16.8801-8811.2003.

[22] Millet JK, Whittaker GR. Host cell proteases: Critical determinants of coronavirus tropism and pathogenesis. Virus Research 2015;202:120–34. https://doi.org/10.1016/j.virusres.2014.11.021.

[23] Du L, Zhao G, Kou Z, Ma C, Sun S, Poon VKM, et al. Identification of a Receptor-Binding Domain in the S Protein of the Novel Human Coronavirus Middle East Respiratory Syndrome Coronavirus as an Essential Target for Vaccine Development. J Virol 2013;87:9939–42. https://doi.org/10.1128/JVI.01048-13.

[24] Muthumani K, Falzarano D, Reuschel EL, Tingey C, Flingai S, Villarreal DO, et al. A synthetic consensus anti–spike protein DNA vaccine induces protective immunity against Middle East respiratory syndrome coronavirus in nonhuman primates. Sci Transl Med 2015;7. https://doi.org/10.1126/scitranslmed.aac7462.

[25] Sabir JSM, Lam TT-Y, Ahmed MMM, Li L, Shen Y, E. M. Abo-Aba S, et al. Co-circulation of three camel coronavirus species and recombination of MERS-CoVs in Saudi Arabia. Science 2016;351:81–4. https://doi.org/10.1126/science.aac8608.

[26] Zhang N, Jiang S, Du L. Current advancements and potential strategies in the development of MERS-CoV vaccines. Expert Review of Vaccines 2014;13:761–74. https://doi.org/10.1586/14760584.2014.912134.

[27] Farrag MA, Amer HM, Bhat R, Almajhdi FN. Sequence and phylogentic analysis of MERS-CoV in Saudi Arabia, 2012–2019. Virol J 2021;18:90. https://doi.org/10.1186/s12985-021-01563-7.

[28] Sohrab SS, Azhar EI. Genetic diversity of MERS-CoV spike protein gene in Saudi Arabia. Journal of Infection and Public Health 2020;13:709–17. https://doi.org/10.1016/j.jiph.2019.11.007.

[29] Tai W, Wang Y, Fett CA, Zhao G, Li F, Perlman S, et al. Recombinant Receptor-Binding Domains of Multiple Middle East Respiratory Syndrome Coronaviruses (MERS-CoVs) Induce Cross-Neutralizing Antibodies against Divergent Human and Camel MERS-CoVs and Antibody Escape Mutants. J Virol 2017;91:e01651-16. https://doi.org/10.1128/JVI.01651-16.

[30] Wang Q, Qi J, Yuan Y, Xuan Y, Han P, Wan Y, et al. Bat Origins of MERS-CoV Supported by Bat Coronavirus HKU4 Usage of Human Receptor CD26. Cell Host & Microbe 2014;16:328–37. https://doi.org/10.1016/j.chom.2014.08.009.

[31] Yang Y, Du L, Liu C, Wang L, Ma C, Tang J, et al. Receptor usage and cell entry of bat coronavirus HKU4 provide insight into bat-to-human transmission of MERS coronavirus. Proc Natl Acad Sci USA 2014;111:12516–21. https://doi.org/10.1073/pnas.1405889111.

[32] Cui J, Li F, Shi Z-L. Origin and evolution of pathogenic coronaviruses. Nature Reviews Microbiology 2019;17:181–92. https://doi.org/10.1038/s41579-018-0118-9.

[33] Su S, Wong G, Shi W, Liu J, Lai AC, Zhou J, et al. Epidemiology, genetic recombination, and pathogenesis of coronaviruses. Trends in Microbiology 2016;24:490–502.

[34] Ikram A, Zaheer T, Awan FM, Obaid A, Naz A, Hanif R, et al. Exploring NS3/4A, NS5A and NS5B proteins to design conserved subunit multi-epitope vaccine against HCV utilizing immunoinformatics approaches. Sci Rep 2018;8:16107. https://doi.org/10.1038/s41598-018-34254-5.

[35] Mahmud S, Rafi MO, Paul GK, Promi MM, Shimu MSS, Biswas S, et al. Designing a multi-epitope vaccine candidate to combat MERS-CoV by employing an immunoinformatics approach. Sci Rep 2021;11:15431. https://doi.org/10.1038/s41598-021-92176-1.

[36] Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinformatics 2007;8:4. https://doi.org/10.1186/1471-2105-8-4.

[37] 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 2007;8:424. https://doi.org/10.1186/1471-2105-8-424.

[38] Kaliamurthi S, Selvaraj G, Junaid M, Khan A, Gu K, Wei D-Q. Cancer Immunoinformatics: A Promising Era in the Development of Peptide Vaccines for Human Papillomavirus-induced Cervical Cancer. CPD 2019;24:3791–817. https://doi.org/10.2174/1381612824666181106094133.

[39] Jiaming L, Yanfeng Y, Yao D, Yawei H, Linlin B, Baoying H, et al. The recombinant N-terminal domain of spike proteins is a potential vaccine against Middle East respiratory syndrome coronavirus (MERS-CoV) infection. Vaccine 2017;35:10–8. https://doi.org/10.1016/j.vaccine.2016.11.064.

[40] Madu IG, Roth SL, Belouzard S, Whittaker GR. Characterization of a Highly Conserved Domain within the Severe Acute Respiratory Syndrome Coronavirus Spike Protein S2 Domain with Characteristics of a Viral Fusion Peptide. J Virol 2009;83:7411–21. https://doi.org/10.1128/JVI.00079-09.

[41] Xu Y, Lou Z, Liu Y, Pang H, Tien P, Gao GF, et al. Crystal Structure of Severe Acute Respiratory Syndrome Coronavirus Spike Protein Fusion Core*. Journal of Biological Chemistry 2004;279:49414–9. https://doi.org/10.1074/jbc.M408782200.

[42] Kleine-Weber H, Elzayat MT, Hoffmann M, Pöhlmann S. Functional analysis of potential cleavage sites in the MERS-coronavirus spike protein. Sci Rep 2018;8:16597. https://doi.org/10.1038/s41598-018-34859-w.

[43] Park J-E, Li K, Barlan A, Fehr AR, Perlman S, McCray PB, et al. Proteolytic processing of Middle East respiratory syndrome coronavirus spikes expands virus tropism. Proc Natl Acad Sci USA 2016;113:12262–7. https://doi.org/10.1073/pnas.1608147113.

[44] J Alsaadi EA, Jones IM. Membrane Binding Proteins of Coronaviruses. Future Virol 2019;14:275–86. https://doi.org/10.2217/fvl-2018-0144.

[45] Forni D, Cagliani R, Clerici M, Sironi M. Molecular Evolution of Human Coronavirus Genomes. Trends in Microbiology 2017;25:35–48. https://doi.org/10.1016/j.tim.2016.09.001.

[46] Su S, Wong G, Shi W, Liu J, Lai ACK, Zhou J, et al. Epidemiology, Genetic Recombination, and Pathogenesis of Coronaviruses. Trends in Microbiology 2016;24:490–502. https://doi.org/10.1016/j.tim.2016.03.003.

[47] Shey RA, Ghogomu SM, Esoh KK, Nebangwa ND, Shintouo CM, Nongley NF, et al. In-silico design of a multi-epitope vaccine candidate against onchocerciasis and related filarial diseases. Sci Rep 2019;9:4409. https://doi.org/10.1038/s41598-019-40833-x.

[48] Sohrab SS, Azhar EI. Genetic diversity of MERS-CoV spike protein gene in Saudi Arabia. Journal of Infection and Public Health 2020;13:709–17. https://doi.org/10.1016/j.jiph.2019.11.007.

[49] Jo Y-S, Tamuri AU, Towers GJ, Goldstein RA. SARS-CoV-2 convergent evolution cannot be reliably inferred from phylogenetic analyses 2021:2021.05.15.444301. https://doi.org/10.1101/2021.05.15.444301.

[50] Andreatta M, Nielsen M. Gapped sequence alignment using artificial neural networks: application to the MHC class I system. Bioinformatics 2016;32:511–7. https://doi.org/10.1093/bioinformatics/btv639.

[51] Nielsen M, Lundegaard C, Worning P, Lauemøller SL, Lamberth K, Buus S, et al. Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. Protein Science 2003;12:1007–17. https://doi.org/10.1110/ps.0239403.

[52] 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:W449–54. https://doi.org/10.1093/nar/gkaa379.

[53] Fast E, Altman RB, Chen B. Potential T-cell and B-cell Epitopes of 2019-nCoV 2020:2020.02.19.955484. https://doi.org/10.1101/2020.02.19.955484.

[54] 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;27:671-680.e2. https://doi.org/10.1016/j.chom.2020.03.002.

[55] Tang X, Wu C, Li X, Song Y, Yao X, Wu X, et al. On the origin and continuing evolution of SARS-CoV-2. Natl Sci Rev 2020;7:1012–23. https://doi.org/10.1093/nsr/nwaa036.

[56] Gui M, Song W, Zhou H, Xu J, Chen S, Xiang Y, et al. Cryo-electron microscopy structures of the SARS-CoV spike glycoprotein reveal a prerequisite conformational state for receptor binding. Cell Res 2017;27:119–29. https://doi.org/10.1038/cr.2016.152.

[57] Ahmed SF, Quadeer AA, McKay MR. Preliminary Identification of Potential Vaccine Targets for the COVID-19 Coronavirus (SARS-CoV-2) Based on SARS-CoV Immunological Studies. Viruses 2020;12:254. https://doi.org/10.3390/v12030254.

[58] Luo C-M, Wang N, Yang X-L, Liu H-Z, Zhang W, Li B, et al. Discovery of Novel Bat Coronaviruses in South China That Use the Same Receptor as Middle East Respiratory Syndrome Coronavirus. J Virol 2018;92:e00116-18. https://doi.org/10.1128/JVI.00116-18.

[59] Yang L, Gong L, Wang P, Zhao X, Zhao F, Zhang Z, et al. Recent Advances in Lipid Nanoparticles for Delivery of mRNA. Pharmaceutics 2022;14:2682. https://doi.org/10.3390/pharmaceutics14122682.

How to cite this paper:

Raneem S. alamri, and Najla Ali Alburae. MERS-CoV Spike Protein Variants: A Computational Perspective. *J Biol Insights*. 2025;1(1):94–109.

Arabic translation of title page with Arabic abstract

متغيرات البروتين الشوكي لفيروس كورونا المسبب لمتلازمة الشرق الأوسط التنفسية: منظور حاسوبي

رنيم س. العمري[1,*] ، نجلاء البرعي[1,*]

[1] قسم العلوم البيلوجية، كلية العلوم، جامعة الملك عبد العزيز، جدة ٢١٥٨٩، المملكة العربية السعودية

*المؤلفون المراسلون : Ralamri0314@stu.kau.edu.sa (RSA)| nalbaurai@kau.edu.sa (NAA)

الملخص

يُعتبر فيروس متلازمة الشرق الأوسط التنفسية(MERS-CoV) ، الذي تم اكتشافه لأول مرة في عام ٢٠١٢، من القضايا الصحية العامة البارزة بسبب ارتفاع معدل الوفيات المرتبط به، والذي يتجاوز ٣٥%. ويُعد بروتين الشوكة (Spike protein)، المسؤول عن دخول الفيروس إلى الخلايا المضيفة، الهدف الرئيسي لتطوير اللقاحات والعلاجات المضادة. إن فهم التغيرات الجينية في بروتين الشوكة يُعتبر أساسيًا للتنبؤ بتطور الفيروس وآليات هروبه المناعي وتصميم لقاحات شاملة لعائلة فيروسات كورونا. أصبحت المقاربات الحاسوبية أدوات حيوية لتحليل متغيرات بروتين الشوكة في فيروس MERS-CoV، لما توفره من طرق سريعة وفعّالة من حيث التكلفة للمراقبة العالمية والنمذجة التنبؤية. يستعرض هذا البحث أبرز الدراسات الحاسوبية المتعلقة بمتغيرات بروتين الشوكة، بما في ذلك التحليلات التسلسلية المعتمدة على علم الوراثة التطوري، ودراسات النمذجة البنائية، وخوارزميات التنبؤ بالحواتم(epitopes) ، والمقاربات المعتمدة على التعلم الآلي. كما نُبرز كيف ساعدت هذه الأساليب في تحديد المناطق المحفوظة المناسبة لاستهداف اللقاحات، وتتبع أنماط تطور الفيروس، والتنبؤ بالأثر الوظيفي للطفرات. وعلى الرغم من أن الدراسات الحاسوبية قدمت رؤى قيمة حول تنوع وتطور بروتين الشوكة في فيروس MERS-CoV، إلا أن هناك فجوات كبيرة لا تزال قائمة في مجالات جمع العينات على المستوى العالمي، والتحقق التجريبي من التنبؤات، ودمج أحداث إعادة التركيب الجيني. وتشمل التوجهات المستقبلية تطوير أنظمة مراقبة مدعومة بالذكاء الاصطناعي، ومنصات تصميم لقاحات شاملة لفيروسات كورونا، وتحسين التكامل بين التحليل الحاسوبي والتجريبي لتسريع تطوير العلاجات المضادة للفيروسات.

الكلمات المفتاحية : النمذجة الهيلكية | علم الوراثة العرقي | علم الاحياء الحاسوبي | بروتين الشوكة | فيروس كرونا