Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

BBA - Molecular Basis of Disease 1868 (2022) 166514

Contents lists available at ScienceDirect

BBA - Molecular Basis of Disease

journal homepage: www.elsevier.com/locate/bbadis

Decoding molecular factors shaping human angiotensin converting enzyme 2 receptor usage by spike glycoprotein in lineage B beta-coronaviruses

Sandipan Chakraborty [a,b,*], Sanjana Ghosh [b], Trisha Mondal [b]

[a] Center for Innovation in Molecular and Pharmaceutical Sciences (CIMPS), Dr. Reddy's Institute of Life Sciences, University of Hyderabad Campus, Gachibowli, Hyderabad 500046, India
[b] Amity Institute of Biotechnology, Amity University, Kolkata 700135, India

ARTICLE INFO

Keywords:
ACE2 usage
Beta-coronaviruses
Recombination
Selection forces
Interfacial hydrogen bonds

ABSTRACT

Acquiring the human ACE2 receptor usage trait enables the coronaviruses to spill over to humans. However, the origin of the ACE2 usage trait in coronaviruses is poorly understood. Using a multi-disciplinary approach combining evolutionary bioinformatics and molecular dynamics simulation, we decode the principal driving force behind human ACE2 receptor recognition in coronaviruses. Genomic content, evolutionary divergence, and codon usage bias analysis reveal that SARS-CoV2 is evolutionarily divergent from other human ACE2-user CoVs, indicating that SARS-CoV2 originates from a different lineage. Sequence analysis shows that all the human ACE2-user CoVs contain two insertions in the receptor-binding motif (RBM) that directly interact with ACE2. However, the insertion sequences in SARS-CoV2 are divergent from other ACE2-user CoVs, implicating their different recombination origins. The potential of mean force calculations reveals that the high binding affinity of SARS-CoV2 RBD to human ACE2 is primarily attributed to its ability to form a higher number of hydrogen bonds than the other ACE2-user CoVs. The adaptive branch-site random effects likelihood method identifies positive selection bias across the ACE2 user CoVs lineages. Recombination and selection forces shape the spike evolution in human ACE2-using beta-CoVs to optimize the interfacial hydrogen bonds between RBD and ACE2. However, these evolutionary forces work within the constraints of nucleotide composition, ensuring optimum codon adaptation of the spike (*S*) gene within the host cell.

1. Introduction

Understanding the evolutionary dynamics of coronavirus is a topic of recent great importance. We are witnessing a Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV2) pandemic, initially reported from Wuhan, China, in December 2019. The virus evolves into various variants with enhanced receptor recognition and immune evasion abilities during the pandemic [1–3]. As a result, we have witnessed many waves of upsurge cases coming one after another, led by different variants [4,5]. Thus studying the evolutionary arms race dynamics of SARS-CoV2 for its receptor usage is very important to understand the pathway of diversifications of the virus. Identifying crucial residues for receptor recognition under strong evolutionary bias is essential in predicting future variants with improved host cell invasion abilities.

Coronaviruses belong to the family Coronaviridae in the order Nidovirales and further classified into four genera: Alphacoronavirus, Betacoronavirus, Gammacoronavirus, and Deltacoronavirus [6]. Alpha- and beta-coronaviruses primarily infect mammals [6], while gamma-coronaviruses infect avian species [7,8]. Notably, the first evidence of avian species infection by an emerging mammalian alphacoronavirus also has been demonstrated very recently [9]. Deltacoronaviruses can infect both mammalian and avian species [10]. All CoVs encode a surface glycoprotein spike, which binds to the host-cell receptor to mediate viral entry into the cell [11,12]. Structurally, the spike glycoprotein is comprised of two domains, S1 and S2. Three monomers of spike proteins entangle to form a homo-trimeric large clover-shaped protrusion [13]. Three S1 domains entwine to form the ectodomain, and the S2 domains intertwine to create the stalk, transmembrane, and small intracellular domains [14]. The S1 region contains the receptor-binding domain (RBD), which binds the host receptor to open up the cleavage sites [15]. Cleavage by the host proteases mediates the fusion of the viral membrane to the host membrane.

* Corresponding author at: Center for Innovation in Molecular and Pharmaceutical Sciences (CIMPS), Dr. Reddy's Institute of Life Sciences, University of Hyderabad Campus, Gachibowli, Hyderabad 500046, India.
E-mail address: sandipanc@drils.org (S. Chakraborty).

https://doi.org/10.1016/j.bbadis.2022.166514
Received 12 February 2022; Received in revised form 20 July 2022; Accepted 25 July 2022
Available online 4 August 2022