Dear Editor,

The outbreak of the COVID-19 that occurred in late 2019 has posed a remarkable threat to public health around the world. It is known that SARS-CoV-2 is a genetically diverse group that mutates continuously, leading to the emergence of multiple variants. Potential variants of concern (VOCs), variants of interest (VOIs), or variants under monitoring (VUMs) are regularly assessed based on the risk posed to global public health.

Here, we propose that the prototype Omicron variant B.1.1.529 might be derived from the recombination of two early PANGO lineages of SARS-CoV-2, where the continued decline in the prevalence of the previous Delta and other variants is observed. Despite its prompt predominance, knowledge gaps remain in their origin and evolution, fueling worldwide interests and speculations. As the minor parents, hybridized the SARS-CoV-2/human/IRN/Ir-3/2019-like strains. Another substitution of G/D339D may come from a back mutation after recombination. All these substitutions locate in the NTD (N-terminal domain, residues 18–300) and RBD (receptor-binding domain, residues 331–528) of the 51 subunit of spike protein, especially the latter where up to 16 substitutions occurred. The consistency of amino acid residues encoded by Omicron and its minor parent in the corresponding fraction and the difference between it and the major parent proved at the level of amino acids that the recombination event may have happened.

It is known that the SARS-CoV-2 Omicron variant encodes 37 amino acid substitutions (including insertions and deletions) in the spike protein, and then the recombination event alone leads to 22 of them. Spike protein is the most crucial structural protein of SARS-CoVs. It plays a pivotal role in viral attachment, fusion, and entry and is also a target for developing antibodies, entry inhibitors, and vaccines. The recently emerged substitutions within the SARS-CoV-2 Omicron variant thereby raised concerns about the effectiveness of available vaccines and antibody therapeutics. It has been reported that these substitutions have led to some subtle variations in the spatial structure and the affinity to the hACE2 receptor of the spike protein. More importantly, it has caused the immune escape of the Omicron variant to the available vaccines and antibody therapeutics.

According to the isolation frequency, BA.1 is a lineage with medium or high circulating intensity. There are 402,436 and 209,127 isolates in the GISAID’s EpiCov™ database and SARS-CoV-2 data (NCBI). These genome sequences belong to 1,263 PANGO lineages, including 29 lineages of VOCs, VOIs, VUMs, and formerly monitored variants (FMVs) according to WHO’s Tracking SARS-CoV-2 variants (https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/), accessed December 18, 2021, and are with the earliest collection times within each PANGO lineage (Supplementary Table S1, and S2). By assessing the extent of sequencing completion, 2609 whole-length genomes of SARS-CoV-2 were used for the first round rapid screen (Extended Data 1, Genome sequence matrix 1). Subsequently, the genomic sequences involved in all putative recombination events identified in the first round of screening were singled out for further validation (Extended Data 2, Genome sequence matrix 2). Taking SARS-CoV-2/human/USA/UT-UPHL-211211887190/2021 (Accession, OL920485) as the representative of early prototype Omicron variant B.1.1.529 for querying, recombination events were detected and verified by Recombination Detection Program (RDP) v4.101 and the SimPlot Program package.

We confirmed that at least one recombination event occurred in the origin and evolutionary history of the prototype Omicron variant of SARS-CoV-2. In this event, strains belonging to PANGO lineage BA.1, like SARS-CoV-2/human/USA/COR-21-434196/2021 (Accession, OL849989), provided the fundamental genome for VOC Omicron and served as its major parents. While strains like SARS-CoV-2/human/IRN/Ir-3/2019 (Accession, MW737421) belonging to PANGO lineage B.35, as the minor parents, hybridized the genomic fractions into the major genome at the position of 21593–23118 nt (Fig. 1a and d and Supplementary Fig. S1). This fraction encodes 144–505 amino acid residues of SARS-CoV-2's spike protein (S). As a result of the recombination, VOC Omicron did derive the substitutions of N211I, L212V, V213R, R214E, deletion215P, deletion216E, R346K, S371L, S373P, S375F, K417N, N440K, G446S, S477N, T478K, E484A, Q493R, G496S, Q498R, N501Y, Y505H, from the minor parent of SARS-CoV-2/human/IRN/Ir-3/2019-like strains. Another substitution of G/D339D may come from a back mutation after recombination. All these substitutions located in the NTD (N-terminal domain, residues 18–330) and RBD (receptor-binding domain, residues 331–528) of the 51 subunit of spike protein, especially the latter where up to 16 substitutions occurred. The consistency of amino acid residues encoded by Omicron and its minor parent in the corresponding fraction and the difference between it and the major parent proved at the level of amino acids that the recombination event may have happened (Supplementary Table S3).

Received: 8 February 2022 Revised: 27 February 2022 Accepted: 3 March 2022 Published online: 17 March 2022
variants (Accessions, OL920485, OL901845, and OL902308). It also indicated that BA.1 lineage should be the recombination parent of VOC Omicron rather than the descendant of it (Supplementary Table S3). However, the recombination event for VOC Omicron did not occur until recently. The reason may be that the circulating frequency of the lineage B.35 is too low, and its striking area is too limited, which reduces the chance of recombination between it and other lineages of SARS-CoV-2. After all, the prerequisite for recombination is that no less than two lineages of viruses co-infect an individual simultaneously, but the minor parent lineage B.35 is so rare that it hardly has the opportunity to infect an individual with a prior SARS-CoV-2 infection, and vice versa.
Our study suggested that chimerism is an important characteristic of the Omicron PANGO lineage SARS-CoV-2. It was generated by genomic recombination of two early SARS-CoV-2 lineages in the spike protein Coding Sequence (CDS). Recombination is proposed to be critical for coronavirus diversity and the emergence of SARS-CoV-2, MERS-CoV, SARS-CoV (2002), and other zoonotic CoVs. It allows viruses to overcome selective pressure and adapt to new hosts and environments. Viral recombination between different CoVs within animal populations may lead to the emergence of novel zoonotic CoVs that are lethal to humans. Studies comparing coronavirus strains closely related to SARS-CoV-2 had proposed that SARS-CoV-2 acquired the ability to infect human cells through recombination within the spike protein sequence.

Recombination may occur during infections in humans. The global spread and explosive growth of the SARS-CoV-2 in the human population have contributed additional mutational variability into this genome, increasing opportunities for future recombination. It has been reported that recombination among SARS-CoV-2 is associated with the increased spread and severe disease and has resulted in vaccine failure. Thus, targeting the ability of the virus to recombine is a critical consideration for vaccine development in the ongoing SARS-CoV-2 pandemic as well as future animal and zoonotic CoVs.

LIMITATION
This study was carried out only based on bioinformatics analysis, therefore further laboratory analyses needed to confirm the in silico findings.

DATA AVAILABILITY
The data are available from the corresponding author on reasonable request, but GISAID data access, if needed, requires registration.

REFERENCES
1. Bano, I., Sharif, M. & Alam, S. Genetic drift in the genome of SARS-CoV-2 and its global health concern. J. Med. Virol. 94, 88–98 (2022).
2. CDC. SARS-CoV-2 B.1.1.529 (Omicron) Variant—United States, December 1-8, 2021. MMWR Morb. Mortal. Wkly. Rep. 70, 1731–1734 (2021).
3. WHO. Weekly epidemiological update on COVID-19—18 January 2022 (Edition 75). https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19—18-january-2022 (2022).
4. Cameroni, E. et al. Broadly neutralizing antibodies overcome SARS-CoV-2 Omicron antigen shift. Nature 602, 664–670 (2022).
5. Ma, W. et al. Genomic perspectives on the emerging SARS-CoV-2 omicron variant. Genomics Proteomics Bioinformatics. 2022. https://doi.org/10.1007/s41392-022-00949-5.
6. Wang, Y. et al. The significant immune escape of pseudotyped SARS-CoV-2 variant Omicron. Emerg. Microbes Infect. 11, 1–5 (2022).
7. Patino-Galindo, J. A. et al. Recombination and lineage-specific mutations linked to the emergence of SARS-CoV-2. Genome Med. 13, 124 (2021).
