Department of Biotechnology

Ministry of Science & Technology,

Government of India

****

A. Indian SARS-CoV-2 Genome Sequencing Consortia (INSACOG)

1. Research Laboratories across the country had started sequencing of the SARS-CoV-2 genome in April 2020, however these were mostly samples collected randomly across the country. A total of nearly 4,000 samples were sequenced. At that time detailed reports of variants were not generated. Primarily the sequences reported presence of just one clade D614G which was not a major variant of any interest. The first report of a UK Variant of Concern was published in December and the Government immediately initiated the network of large-scale sequencing of viral genome collected in a systematic manner with complete metadata.
2. INSACOG is a national multi agency consortium set up by the Government in 30th December 2020. Ten laboratories of Department of Biotechnology, CSIR, ICMR and Ministry of Health & Family Welfare, were brought together. National Centre for Disease Control (NCDC), MoHFW was given the responsibility to correlate the clinical aspects and coordinate collection of samples from the states. Sequencing of all samples of UK Travellers and a defined percentage of positive samples of others in the community were initiated in January 2021.
3. All laboratories being Government of India Research Institutes, were asked to immediately start work. The capacities of existing centres, the required reagents and consumables were augmented and protocols were prepared in January itself so that the activity could get started straight away. Additionally, funds of Rs. 70.00 crores were made available in March 2021. The capacity is being continuously enhanced to include more sequencing centres, considering the increasing number of positive cases.
4. The sequencing is being currently conducted on samples from all states. So far about 20,000 samples have been sequenced and about 1,500 variants have been identified. These are mainly UK and South Africa Variant and a small number of Brazil Variants. In addition to these some other variants were noted including one which on clinical correlation appear to be a variant of concern which is being studied in detail.

Data on variants is being clinically correlated and regularly shared with respective States by NCDC.

B. The B.1.617 lineage of SARS-CoV-2

5. Mutations at sites E484 and L452 in the virus' Spike protein have been observed separately, but this is the first major viral lineage that combines the two into the same virus The B:1:167 variant was identified first time in India from the sample received from Maharashtra by NCCS Laboratory in Pune on 12th March 2021, but at the time there was no such lineage and also there were no associated outbreaks and the data on transmissibility plus immune escape properties of L452R was not available. The recognition of potential importance of this variant happened, when it was correlated with international data on L452R and outbreaks in Maharashtra. By March 2021, this information was shared with more precise clinical correlation and definitive molecular in vitro test data. Its genomic, epidemiological, and clinical characteristics are being studied. The sequences were deposited in the global Database GISAID in March 2021 and based on Indian sequencing information, the variant has been given a Global Lineage. GISAID lineage shows entry dated 7th December 2020 to be the earliest one in the lineage B.1.617.

6. B.1.617, referred to as the 'Indian' variant or the 'double' mutant, although first report from India but it is not exclusive to India.

7. As on 19 April 2021, there are now 770 reported sequences from India (39%), UK (18%), USA (11%), Switzerland (5%), Germany (5%) .The virus has been cultured, assays are being developed to test these vaccines against these variants for their efficacy .

8. Genomic Analysis of B.1.617 showed that there are 15 lineage defining mutations on the B1 background (D614G) in the Indian variant. Of these, 13 are most common in global sequences (from Outbreak.info). However, two – L452R and E484Q in the Receptor Binding Motif (RBM) of the Spike protein that binds to the ACE2 receptor on human cells – are relatively well described as immune escape mutations. The E484Q mutation is similar to E484K, a mutation found in the UK (lineage B.1.1.7)

and South Africa (lineage B.1.351) variants. The L452R mutation has been found in fast spreading variants in California (B.1.427 and B.1.429). It can increase the binding power of spike proteins with ACE2 receptors on human cells, making it more transmissible. L452R can also potentially enhance viral replication. It can decrease sensitivity to neutralization by a few therapeutic monoclonal antibodies. This is because many neutralizing antibodies preventing the virus from infecting human cells also bind to this area. Such mutations have been seen in many global variants and are not new in themselves. Only the combination is new. Other than these two mutations, the P681R mutation has also been shown to increase infectivity of the virus. So far, **genomic analysis of B.1.617 virus shows no reason for it to escape detection by RT-PCR tests**.

9. The increase in B.1.617 is correlated epidemiologically with some outbreaks (but not all) in India. However, unlike B.1.1.7, near complete displacement of other variants has not been observed. This suggests that the transmissibility of B.1.617 is higher than B.1 strains but is lower than B.1.1.7. **Clinical correlation** suggests that there is no evidence of increased severity or mortality. Overall B.1.617 meets most criteria for VoC but does not appear to be more dangerous than circulating global VoCs such as B.1.1.7, B.1.351, or P1, of which B.1.1.7 is now at community level transmission in India. Vaccination and public health measures are expected to stay effective in reducing transmission, severe disease and mortality.

C. <u>Triple Mutant</u>

10. The terms double or triple mutants are colloquial. Double or triple mutations signifies the number of mutations relevant as immune escape mutant. In fact, these variants harbours 15 lineage-defining mutations. Uniform variant calling nomenclature should be used as per WHO .B.1.617, initially termed as double mutant, has three new spike protein mutation, namely S: E484Q, L452R and P681R on the background of D614G lineage that was the dominant lineage since last year. Technically double or triple mutant refer to same variant. These mutations were found in some of the retrospective sample sequences.

D. <u>Data Deposit</u>

11. INSACOG consortium has data repository at IGIB ,New Delhi and and NIBMG, Kalyani In the IGIB portal, fastq sequences are submitted by sequencing labs. After variant calling and assembly, FASTA sequences along with metadata are submitted in INSACOG data hub in INBMG portal, where time analysis of sequences with locations are performed. Complete. Viral genome sequences are uploaded in "Global initiative on sharing all influenza data" (GISAID) with the tag INSACOG. This is the global registry of SARS CoV 2 sequence data. However, rules of GISAID are such that if the sequence data do not meet the minimal quality standards, then they cannot be deposited. About 50% of the data has been uploaded in GISAID.
12. The distribution of the lineages of B.1.617 and B.1.618 are available on:
    https://outbreak.info/situation-reports?pango=B.1.617
    https://outbreak.info/situation-reports?pango=B.1.618

******