

SUPPLEMENTAL TABLE

Data Availability

GISAID Identifier: EPI_SET_230110db

doi: [10.55876/gis8.230110db](https://doi.org/10.55876/gis8.230110db)

All genome sequences and associated metadata in this dataset are published in GISAID's EpiCoV database. To view the contributors of each individual sequence with details such as accession number, Virus name, Collection date, Originating Lab and Submitting Lab and the list of Authors, visit [10.55876/gis8.230110db](https://gisaid.org/WIV04)

Data Snapshot

- EPI_SET_230110db is composed of 6,701,432 individual genome sequences.
- The collection dates range from 2010-12-06 to 2021-12-31;
- Data were collected in 207 countries and territories;
- All sequences in this dataset are compared relative to hCoV-19/Wuhan/WIV04/2019 (WIV04), the official reference sequence employed by GISAID (EPI_ISL_402124). Learn more at <https://gisaid.org/WIV04>.