GOVERNMENT OF INDIA
MINISTRY OF SCIENCE & TECHNOLOGY
DEPARTMENT OF BIOTECHNOLOGY

RAJYA SABHA
UNSTARRED QUESTION NO-2205
ANSWERED ON 18.12.2025

GenomeIndia project

2205 Dr. Parmar Jashvantsinh Salamsinh:

Will the Minister of Science and Technology be pleased to state:

(a) the current completion status of the GenomeIndia project, including total genomes sequenced and participating institutions;
(b) the data-sharing, storage and ethical-governance systems adopted for its genomic datasets;
(c) whether the Indian Biological Data Centre (IBDC) has been expanded;
(d) if so, the new databases, storage capacity and research services added during 2024– 25;
(e) the expected benefits of wider IBDC use in health, agriculture and biodiversity;
(f) whether the Indian TB Genomic Surveillance Consortium has been set up; and
(g) if so, the number of TB samples sequenced and its impact on diagnosis and drug-resistance tracking?

ANSWER

THE MINISTER OF STATE (INDEPENDENT CHARGE) FOR THE MINISTRY OF SCIENCE AND TECHNOLOGY & EARTH SCIENCES (DR. JITENDRA SINGH)

(a) The current completion status of GenomeIndia Project is as follows:

(i) A Biobank of 20195 samples collected during GenomeIndia project has been established at Centre for Brain Research (CBR), IISc campus, Bengaluru.
(ii) The whole genome sequencing (WGS) of 10174 samples has been completed.
(iii) Joint calling of sequenced samples has been carried out to create a catalogue of genetic variations, enhancing reference genome efforts for India.
(iv) A comprehensive list of markers has been compiled for developing a genome-wide array for research and diagnostics.

Twenty institutions participated in GenomeIndia project are:

1. Centre for Brain Research (CBR), IISc Campus, Bengaluru
2. CSIR - Centre for Cellular and Molecular Biology (CSIR-CCMB), Hyderabad
3. CSIR - Institute of Genomics & Integrative Biology (CSIR-IGIB), New Delhi
4. BRIC - National Institute of Biomedical Genomics (BRIC-NIBMG), Kolkata
5. All India Institute of Medical Sciences (AIIMSJ), Jodhpur
6. Gujarat Biotechnology Research Centre (GBRC), Gandhinagar
7. BRIC - Institute of Bioresources & Sustainable Development (BRIC-IBSD), Imphal
8. Indian Institute of Science Education and Research (IISER), Pune
9. BRIC - Institute of Life Sciences (BRIC-ILS), Bhubaneshwar
10. Mizoram University (MZU), Aizawl
11. National Institute of Mental Health & Neurosciences (NIMHANS), Bengaluru
12. BRIC - Rajiv Gandhi Centre for Biotechnology (BRIC-RGCB), Thiruvananthapuram
13. Sher-i-Kashmir Institute of Medical Sciences (SKIMS), Srinagar
14. BRIC - Centre for DNA Fingerprinting and Diagnostics (BRIC-CDFD), Hyderabad
15. Indian Institute of Information Technology (IIITA), Allahabad
16. Indian Institute of Science (IISc), Bengaluru
17. Indian Institute of Technology Delhi (IITD), New Delhi
18. Indian Institute of Technology Jodhpur (IITJ), Jodhpur
19. Indian Institute of Technology Madras (IITM), Chennai
20. National Centre for Biological Sciences (NCBS), Bengaluru

(b) The GenomeIndia data has been archived in Indian Biological Data Centre (IBDC) at NCR Biotech Cluster Faridabad. The sharing of GenomeIndia data will be governed as per the Biotech PRIDE (Promotion of Research and Innovation through Data Exchange) Guidelines 2021 and FeED (Framework for Exchange of Data) Protocols; and as per the extant rules and regulations issued by the Government of India from time to time.

(c) & (d) Yes, the amount and type of data archived at IBDC have increased during 2024-2025 in terms of databases and research services. The Metabolomics Portal (IMDA), Proteomics Data Portal (IPD), Biological Image Data Portal (IBIA), and GenomeIndia Data Access Portal were operationalised. During 2024-25, IBDC also (a) served as the central data hub for large-scale projects in health (such as INSACOG) and agriculture, (b) hosted training programs to skill

students and researchers in biological data analysis, (c) provided access to high-end computing capacity to researchers across India, and (d) developed a suite of analysis tools/software for the research community called 'Bionode'. The amount of storage space and computing capacity available in IBDC is at 5.5 petabytes and 961 teraflops, respectively.

(e) The IBDC supports health, agriculture, and biodiversity through its archived large-scale biological datasets. In health, these datasets enables early disease detection, outbreak monitoring, precision medicine, and antimicrobial resistance studies. In agriculture, genomic, soil, and environmental data accelerate the development of climate-resilient, high-yield crops and improved livestock while promoting precision farming and soil health management. Its biodiversity datasets document India's rich biological diversity and support conservation strategies.

(f) & (g) Yes, the Indian TB Genomic Surveillance Consortium has been established. It comprises of DBT, ICMR and CSIR Laboratories, as well as Clinical sites. So far, 17,517 tuberculosis isolates have been sequenced by the consortium. The analyzed mutation data for 15749 samples is available at the InTGS portal (https://intgs.nii.ac.in/InTGS/index.php) for use by researchers. Analysis of the sequence data generated so far offers insights into the drug resistance mutation patterns in the country. It reveals a high prevalence of resistance-inducing mutations in first-line drugs and in the fluoroquinolones. However, the number of Extreme Drug Resistance cases is very low. The distribution of various mutations that induce resistance to first-line drugs underscores the importance of adopting targeted next-generation sequencing- based assays and personalised drug regimens for tuberculosis treatment.

*****