BioSpectrum
the business of Bio & Health Sciences

“Only the anonymised summary of clinical variants in the Bharat Cancer Genome Atlas is accessible to clinicians”

31 March 2025 | Views | By Anusha Ashwin

IIT-Madras embarked on an initiative to build India’s Cancer Genome Database, which can completely transform cancer research in India. IIT-M kicked off the Indian breast cancer genome sequence generation a few years ago. In line with World Cancer Day, which falls on February 4 every year, IIT-M unveiled an open source database called the Bharat Cancer Genome Atlas (BCGA). So far, whole exome sequencing has been done on 958 tumour-normal samples from 479 breast cancer patients across India. BioSpectrum India spoke to Prof S Mahalingam, Faculty, Department of Biotechnology, IIT Madras to understand about BCGA and more.

With the launch of the BCGA as open-source data, tools like AI/ML and data analytics software will be deployed. How will these technologies support informed decision-making in diagnosis and treatment?

BCGA contains whole-exome sequencing data from both tumour tissue and matched normal tissue samples. Researchers can leverage this data, along with Artificial Intelligence (AI) or other machine learning (ML) techniques, to explore various combinations and permutations of the genetic variants and extract the specific information they need. Integrating this data with the existing healthcare system is a top priority. BCGA is readily available for all young researchers who want to conduct cancer genomic research. Also, we want to make this genomic information readily accessible to clinicians across the country. This would greatly assist clinicians in making efficient decisions about the best course of treatment for each individual patient.

What significant trends or patterns have you observed in the cancer genomes of the 500 samples collected until now? Are they indicating the probable cause of cancer incidence in these patients?

We observed some unique mutations specific to the Indian population and noticed different frequencies of mutation profiles compared with Western counterparts. BCGA will be very helpful in identifying mutations involved in hereditary cancers.

What other cancers will be considered for building a genome database, and why these cancers?

We are in the process of sequencing other prevalent cancers (such as colorectal, head and neck, pancreatic, and leukaemia) in India.

Scaling up this database will call for more financial aid, and maintaining an open-source database will also require continuous monetary support. What sustainable plans are ahead for this scaleup?

We certainly need more funding support to scale up the genomic sequencing of other prevalent cancer types and also maintain BCGA. We are planning to approach the Department of Biotechnology (DBT), Department of Science & Technology (DST), and the Indian Council of Medical Research (ICMR) for support.

Will IIT-M partner with Tata Memorial Hospital and other large cancer care hospitals in the country to build an extensive database?

We are currently collaborating with 10 different hospitals and are in the process of discussing this with other hospitals. We hope to get samples from across India very soon.

While the database does not reveal patient details, health data privacy is a significant concern. What measures have been taken to ensure data privacy?

We want to emphasise that all samples included in the database are collected and sequenced only after obtaining ethical clearance. So, there are absolutely no ethical concerns regarding the data in the BCGA. Only the anonymised summary of clinical variants in the BCGA is accessible to clinicians and other stakeholders. BCGA follows global best practices in data sharing and security. We have also implemented strict ethical review processes to ensure compliance with regulatory frameworks like the Personal Data Protection Bill of India.

Anusha Ashwin