Confederation of Indian Industry

INDIA @ 75 MISSION 2022 a CII initiative

Azadi Ka Amrit Mahotsav

CII National Task Force on Science and Data (Genomics)

REPORT

December 2021

TABLE OF CONTENTS

1. Executive Summary 3

2. Introduction and Context Setting 5

3. Major Barriers and Policy Challenges 8

4. Specific Recommendations 9

4.1. Recommendations w.r.t. Science, Research and Data 10

4.2. Recommendations w.r.t. key policy and regulatory changes 13

4.3. Recommendations w.r.t. Strengthening science establishments and critical scientific collaborations 15

4.4. Recommendations w.r.t. Capability Building / skilling 18

1. EXECUTIVE SUMMARY

As the threat of existing SARS-CoV2 variants continue and new ones emerge, there is a very clear need for robust genomic sequencing to boost pandemic management. As the country ramps up preparedness to investigate the role of variants and their symptomatology, genomics will be an integral and crucial component. Besides, genomic sequencing is not only unique to COVID-19 but applies to any pathogen in the context of a pandemic, as well as to more localized epidemics caused by viruses and bacterial infections of public health importance.

The COVID-19 outbreak expedited genome sequencing and over 1.2 million coronavirus genome sequences, were collected from over 172 countries, and shared on a popular online data platform, GISAID, as of April 2021. However, India ranked 102 in terms of sharing sequencing data through GISAID, and sequencing from India amounted to 0.13 % of total infections. Bangladesh, Sri Lanka and UK had sequenced 0.19%, 0.2% and 9.25% respectively, at the same time point.

The global genomics market is projected to reach USD 54.4 billion by 2025 from USD 22.7 billion in 2020, at a CAGR of 19.0% during the forecast period. In 2019, North America accounted for the largest share of the overall market, followed by Europe. Increasing government initiatives, adoption of several strategies by key players to provide efficient genomics solutions in this region, and the presence of prominent players are the major drivers of the North American market.

Although India holds a huge potential in genomics, there remains a challenge in terms of funding, data access, data sharing, resources required for scale up, awareness about the potential, inappropriate policies and lack of appropriate skill sets.

There is a recognition that India should aggressively invest in genomics technology not only for COVID-19 management but also for better understanding of rare genetic diseases, cancers, non-communicable diseases and other communicable diseases and to inform development of effective vaccines.

With this in view, CII has formed a **Task Force on Genomics** comprising of industries, academia, research institutions and INSACOG labs. The Task Force would like to submit the following recommendations which is grouped into four categories:

A. Recommendations w.r.t. Science, Research and Data
B. Recommendations w.r.t. key policy and regulatory changes
C. Recommendations w.r.t. strengthening science establishments and critical scientific collaborations
D. Recommendations w.r.t. Capability building / skilling

While the specific recommendations with actionable items are provided in greater detail in the report, some of the key recommendations are listed below.

1. Boost funding and investment to **enhance Genome sequencing capacity**- allocate a minimum fixed percentage of the health budget for surveillance

including genome sequencing
2. Create **infrastructure and policies** for gene synthesis, which is very important
3. Industry to increase investments in newer techniques, **reagents and equipment manufacture for international quality equivalent sequencing** and govt to support the same through incentives, subsidies etc.
4. Enhance research capacity by enriching the capabilities of existing institutes, and create new centrally funded and connected institutes **for networking of surveillance and data sharing and management**
5. Focus on development of new genomic related technologies, particularly **computational methods**. There is a major need for investments in **supercomputing** infrastructure
6. Creation of **National level consortium** comprising of sequencing facilities; Planning and developing **Joint Centres of Excellence** on infectious diseases and health monitoring, vaccine design and development, and new emerging technologies etc.
7. Clear and consistent policies and processes for **robust data collection** and open data sharing mechanisms
8. Creation of a publicly accessible and a **centralized database**, where ease of access to data from external stakeholders is a key performance indicator
9. Ensure uniform data standards, and standardized nomenclature for genomic sequencing data
10. Develop / strengthen **infrastructure and enabling mechanisms** for clinical research and trials
11. Encouragement for **domestic industry to invest in genomics**, with enabling data related policies, clear healthcare data privacy rules and support for generation of intellectual property
12. Set up an integrated **National Level Task Force** for human and animal genomics to build and maintain leadership in technology and data utilization
13. Need for **capacity building at all levels** including Technicians, Bioinformaticians, Nurses, Computational Biologists, Scientists etc.
14. **Skilling** and capabilities in Artificial intelligence (AI), Machine Learning (ML), big data analytics, Bioinformatics capabilities, Computational Biology and Bioinformatics degrees particularly for pathogen genomics need to be developed and strengthened
15. There is a need to strengthen the **intellectual property / patent laws** in India, particularly related to genomics and data

Therefore, the power of genomic surveillance and its ability to meaningfully guide and inform public health response will require collaboration between multiple stakeholders. Going forward, the industries, government, academic and R&D labs would need to step up and collaborate to strengthen the country's genomic surveillance and sequencing capacity.

2. INTRODUCTION AND CONTEXT SETTING

Genomics is the study of genes and their functions, and related techniques used to scrutinize the functioning and composition of the single gene. The role of genomics and related biotechnologies have the potential to achieve a number of public health goals, including providing countries with an efficient, cost effective means of predicting, preventing, diagnosing and treating major diseases burdening their populations. [1] Genomics allows us to understand disease in the human body in a new level of detail. While this field and its tools and techniques are not new, COVID-19 and the global response to it has brought these to the fore. However, it has only scratched the surface as the future seems to be leading towards predictive and personalised healthcare.[2] We are faced with a significant opportunity to develop targeted interventions and to prepare for new pandemics and threats to public health.

Genomic solutions will be an integral and crucial component of the future of healthcare.

With COVID-19 there has been a significant transformation in the use of viral genomics in disease outbreaks. The most salient development is a shift from retrospective analysis of the past, to an ability to investigate genomic epidemiology in near real-time. Public health emergencies have served a catalytic role in fuelling developments in viral genomic sequencing and molecular epidemiology.

Beyond COVID 19: The field has been extremely helpful for infectious disease and communicable disease specialists to identify mutations associated with pathogens or human tissue samples for cancer or other diseases in order to diagnose particular conditions in a more precise way.

Genomic surveillance is not unique to COVID-19 but applies to any pathogen in the context of a pandemic or even localized epidemics, or infections. When a new pathogen is encountered for the first time; genomic sequencing enables its characterization and aids the development of diagnostics and therapeutics. This is done in combination with other biological disciplines such as structural biology and virology.

During a pandemic, it is vitally important to understand how pathogens mutate and change and how these changes affect host-pathogen interactions in terms of clinical manifestations of diseases and immune responses both natural and vaccine induced to be able to strengthen public health response. Additionally, genomic sequencing through a pandemic helps to identify regions of the pathogen that remain more stable than others thereby facilitating the development of better therapeutics and diagnostics.[3,4]

Further, with various viral infections (Dengue, Chikungunya, Zika, Japanese Enchepalitis

[1] https://www.who.int/news-room/q-a-detail/genomics
[2] https://kennedyslaw.com/thought-leadership/reports/life-science-in-the-era-of-pandemics-part-4-fighting-covid-19-with-genomics/
[3] Genomic sequencing in pandemics- Lancet, Vol 397 February 6, 2021
[4] Genomic sequencing of SARS-CoV-2 A guide to implementation for maximum impact on public health 8 January 2021

etc.,) affecting various regions of India, it is important to characterize the novel variant that have potential to mutate and become aggressive. Therefore, genomic sequencing of local strains in very critical for managing and treating such infections.

In order to contextualise current trends in scientific collaboration, data sharing and genomics, and to better understand their interplay, it is worthwhile comparing differences in speed and scale between the genomic responses seen in two distinct and notable infectious disease outbreaks, is outlined.

a) In 2003, with the SARS outbreak, only three virus genomes were publicly shared in the first month following the detection of the causative pathogen; and only 31 were available in 3 months. Genomics was not advanced enough to allow virus epidemiology to be studied in real-time and at scale.

b) Fast forward to 2019, metagenomic sequencing was used to identify the causative pathogen within a week of the first reported case of SARS-CoV-2. Within two weeks, six genomes were shared publicly enabling extensive virus genomic sequencing. Since then, sequencing efforts have acquired a global, collaborative shape. A constantly growing data set of more than 60,000 near-complete viral genomes was available within 6 months following the identification of SARS-CoV-2.

Real time global genomic surveillance requires databases, data standards and data sharing standards and policies. The Indian SARS-CoV-2 Genomic Consortium (INSACOG) was only set up in December 2020. Prior to this sequencing in India was random and India's contribution to global databases such as Nextstrain was very low.

As of April 28, 2021, **India ranked 102 in terms of sharing sequencing data through GISAID**[5]. **This continues to be the case with the INSACOG having been able to sample only a fraction of its initial target of 5% of positive cases.** As of 6th July 2021, 40777,[6], genomes have been sequenced. India's case count is 3.06 crores[7]. That amounts to 0.13 % of total infections. Bangladesh, Sri Lanka and the UK have sequenced 0.19%, 0.2% and 9.25% respectively[8].

India should aggressively invest in genomics technology, gene synthesis, and creation of genomic reagents (primers, genes, vectors, plasmids) necessary to create solutions for such infections by creating multiple institutes geared for research in these fields.

Therefore, it is important to strengthen the country's genomic surveillance and sequencing capacity.

With this in view, **CII has formed a Task Force on Genomics**, the composition of which is appended:

[5] https://www.newindianexpress.com/nation/2021/may/02/the-meta-blocwhy-india-faces-a-data-shortage-on-genome-sequencing-of-coronavirus-2297425.html
[6] http://clingen.igib.res.in/covid19genomes/
[7] https://coronavirus.jhu.edu/region/india
[8] https://scroll.in/article/997327/why-is-india-failing-to-use-its-genome-sequencing-capacity-for-covid-19

Chairman:

Mr Kris Gopalakrishnan
Past President, CII and Chairman, CII National Task Force on Genomics

List of Members:

1. Prof. Ajit Kumar Chaturvedi, Director, IIT- Roorkee
2. Dr Ashutosh Raghuvanshi, MD & CEO, Fortis Healthcare
3. Dr Ashok Kumar FRSC, President, Centre for R&D, IPCA
4. Prof B.S. Murty, Director, IIT Hyderabad
5. Dr D. Sundar, Institute Chair Professor, Head, Department of Biochemical Engineering and Biotechnology Coordinator, IIT Delhi
6. Prof Gagandeep Kang, FRS Professor, Christian Medical College
7. Dr Harish Iyer, Lead, R&D and innovation, Gates Foundation
8. Dr Krishna Ella, MD, Bharat Biotech
9. Dr Lakshmikanth Gandikota, Senior Vice President, Biopharma, Intas
10. Dr Mahesh Bhalgat, Chief Operating Officer, Syngene
11. Ms Mahima Datla, MD& CEO, Biological E
12. Dr Manoj Kumar Bhat, Director, National Centre for Cell Science, Pune
13. Mr. Mayur Abhaya, Managing Director, Life Cell International
14. Dr Nanda Kumar Jairam, Chairman & CEO, Columbia Asia Hospital
15. Dr Prabuddha Kundu, Co-founder and MD, Premas Biotech
16. Prof Ramgopal Rao, Director, IIT Delhi
17. Mr RamPrasad V, CEO, Medgenome
18. Prof Rintu Banerjee, Head and Professor, IIT Kharagpur
19. Dr Sanjeev Kumar, President, Zydus Cadila
20. Mr Sanjay Murdeshwar, Country President & Managing Director, Novartis
21. Dr K Thangaraj, Director, Centre for DNA Fingerprinting and Diagnostics
22. Ms Vani Manja, Country MD India, Boehringer Ingelheim
23. Dr Vinay Nandicoori, Director, CSIR-CCMB
24. Mr Vivek Kamath, MD, Abbott
25. Prof V. K Tewari, Director, IIT- Kharagpur

The aforesaid Task Force would like to submit the following recommendations broadly grouped into four categories:

A. Recommendations w.r.t. Science, Research and Data
B. Recommendations w.r.t. key policy and regulatory change
C. Recommendations w.r.t. strengthening science establishments and critical scientific collaborations
D. Recommendations w.r.t. Capability building / skilling

Each of the above are explained in detail in subsequent sections of the report.

3. MAJOR BARRIERS AND POLICY CHALLENGES

Limited Resources and Institutional Barriers

Financial, infrastructural and technical constraints pose some of the main barriers to being able to perform and share results of genome sequencing and allied services. **Genome sequencing is resource intensive – in addition to high cost of establishing and maintaining a sequencing facility, it requires high levels of computational power, massive data storage, expensive reagents, and skilled personnel.**[9]

There is a need to explore broader partnerships to optimize utilization of available resources. While the INSACOG has scaled up its activities and the laboratories have been expanded to 28, it would further benefit from a deeper collaboration between public and private entities.

Governments with diverse healthcare systems are looking to increase public funding for genomics related projects in the wake of the pandemic.

- The Biden Administration, recognising the importance of genomic sequencing, made a nearly $200 million investment to help increase sequencing capacity in April this year.[10]

- UK Government announced new funding of GBP 37 million for genomics and data driven initiatives. These will be delivered through the Genome UK implementation Plan and the Functional Genomics Initiative. [11] In terms of public private partnerships, the UK Biobank is a prime example of how the public and private entities can come together to fund large scale genomics initiatives. In addition to support from public health bodies, it has attracted GBP 200 million in industry funding. Over 2,000 researchers worldwide use UK Biobank to study genetics and other determinants of disease in a diverse set of projects.

Legal and Ethical Framework for Data Sharing

Biomedical research handles large quantities of potentially sensitive data. With the growth of cloud services, data sharing across jurisdictions has become easier and cheaper. Open data enables quick and easy access to information by research teams, reducing the need to seek multiple permissions from owners of discrete datasets. In this context, the need to establish a framework to safely and securely share data without giving rise to ethical or legal concerns should be a policy priority. Public participation and willingness are of course crucial in creating and maintaining genomic databases. However, apprehensions relating to privacy and data security will create hindrances. To

[9] https://www.sciencedirect.com/science/article/pii/S2212066116300205

[10] https://globalbiodefense.com/2021/07/14/us-announces-plans-for-public-health-pathogen-genomics-centers-of-excellence/

[11] https://www.drugtargetreview.com/news/90361/uk-government-to-fund-37-million-into-genomics-projects/

implement a quick ethical and data privacy framework which also encourages the industry participation, India needs to formally align with existing international data privacy regulations which have been well tested like The general data protection regulation (GDPR) of European Union. At the same time fine tune these existing regulations to the specific needs of our country and come up with long term balanced regulation.

A collaborative framework for responsible data sharing between sequencing facilities, data scientists, clinics, and healthcare providers, and industry must be developed placing ideals of privacy and security at the centre of policy. For instance, a massive international cancer project to identify the genetic traits associated with various types of cancer made their data available to investigators all over the world after placing safeguards to protect the privacy of patients and volunteers. Since such projects source a vast amount of data from across countries, it is not always clear how to protect privacy, and comply with international and national data protection regulations/laws, especially considering that such legislation is subject to ongoing changes.

4. SPECIFIC RECOMMENDATIONS

It is very clear that an increased genomic sequencing, gene synthesis capacity and sharing of data is likely to benefit the entire at-risk population who might contract an infection during an outbreak of an infectious disease.

As mentioned earlier, other applications of sequencing and synthesis include understanding of rare genetic diseases, cancers, non-communicable diseases and development of cell and gene therapies against the targets etc. Applications of genomic sequencing and gene synthesis can serve as tools in clinical medicine, to drug research and development, devise options for cell and gene therapy, and help guide healthcare policy generally.[12] It is because of genomic sequencing and synthesis and open data sharing that a vaccine for COVID could be developed in record time. The Pfizer and Moderna vaccines also represent the first in a new generation that harness genomics to train the human immune system against the SARS-CoV-2 virus using 'mRNA' molecules which can be developed in a matter of weeks from DNA templates without needing to grow the virus.[13]

With this in view, we submit the following recommendations which is grouped into four categories:

4.1. Recommendations w.r.t. Science, Research and Data
4.2. Recommendations w.r.t. key policy and regulatory changes
4.3. Recommendations w.r.t. strengthening science establishments and critical scientific collaborations
4.4. Recommendations w.r.t. Capability building / skilling

[12] https://www.sciencedirect.com/topics/medicine-and-dentistry/genome-sequencing

[13] https://kennedyslaw.com/thought-leadership/reports/life-science-in-the-era-of-pandemics-part-4-fighting-covid-19-with-genomics/

4.1. Recommendations w.r.t. Science, Research and Data

4.1.1. Boost funding and investment to enhance genome sequencing capacity

By realising the importance of genomic research and its applications for health, drugs and food security, the Government and industry should prioritize funding and investments in genomic research. Genomic surveillance and sequencing are expensive. Currently, it costs between INR 5000-8000 to sequence SARS-CoV2 (the number will depend on the throughput as well). Hence while genomic surveillance is indeed beneficial, it requires massive investments in infrastructure, new evolving technologies, development of new assays, enhancing capacity of human resources. Additionally, the reagents needed for sequencing are expensive and hence the running costs of sequencing facilities are high.

Towards this, following are the recommendations:

4.1.1.1. A **concrete incentive structure and subsidies** to encourage industry participation for genome sequencing needs to be built. Insurance coverage for genome sequencing tests need to be provisioned and encouraged by IRDAI (keeping the non-discriminatory provisions) so that more precision medicine programs are implemented and also doesn't negatively impact the pockets of patients as many of these sequencing tests are expensive.

4.1.1.2. Government needs to decide the **percentage of the health budget that will be allocated to surveillance, including genomic sequencing both as a part of routine genomic surveillance as well as during a pandemic**. The amount will depend on disease prevalence, context, etc.

4.1.1.3. There is a need to build more **regional and state capacity** for genomic sequencing and gene synthesis that can be geographically well distributed.

4.1.1.4. The industry should **invest in newer techniques that allow for high throughput genomic sequencing** such as amplicon sequencing protocols that can be used to identify functionally important regions in pathogens[14]. This combined with barcoding for high throughput sequencing based detection can be used for pooled sequencing. This can help reduce costs and timelines and is efficient for population-based sequencing which is important in the context of a pandemic.

4.1.1.5. The industry should also develop capacity to **manufacture quality reagents** used for sequencing, gene synthesis, oligonucleotide, and plasmid synthesis to cut down dependencies on global supply chains that may be strained during a pandemic

4.1.1.6. Indian industry needs to take up the challenging task of **manufacturing of instruments/ consumables/kits** so that same is available at affordable price to the Indian scientists / researchers. Government incentives and subsidies

[14] Next generation sequencing for pandemic preparedness- Indian Chemical Engineer, 62:4, 351-358

would encourage the industries.

4.1.1.7. It would be important to channelize **CSR funding** from industries to fund important projects. **Industry should be involved in research programs of common interests**. Existing grant schemes are High-risk high-reward, IMPRINT, UAY etc. can be expanded and be made more competitive.

4.1.1.8. A strong focus on **Genome synthesis** is essential. Genome synthesis is not at all available in our country. We need to ensure that our data does not go out of the country. We must have mechanisms to create, maintain and monitor Centralized Gene Synthesis capability.

4.1.2. Enhance research capacity in genomics

It is recommended to enhance research capacity in disciplines as molecular virology, bacteriology, immunology, structural and computational biology through the following:

4.1.2.1. Increase the **number of laboratories and researchers** within institutes specialized in genomic research.

4.1.2.2. The laboratories should be capable of handling research with pathogens in terms of their **biosafety requirements**.

4.1.2.3. The **National Institutes specialized in research on pandemic diseases** (like NIV), cell and gene therapy, immunology (like NII) may be further assisted to develop expertise in Omics (genomics, proteomics, metabolomics, metagenomics and transcriptomics) and data mining.

4.1.2.4. The **networking amongst the institutes** should be such that the data generated can be utilized to draw purposeful conclusions.

4.1.2.5. Create a **centrally funded Institute** focused on research on genomics, gene delivery platforms such as recombinant viral vectors, cellular therapy, synthetic biology (mRNA synthesis and production), gene delivery formulations that are collocated to address emerging needs for India during such situations.

4.1.2.6. Since mutations can continuously generate new and more virulent strains, **extensive and vigorous focus on vaccine and drug development** for existing and emerging strains by establishing high throughput gene synthesis platforms in each of the research institutes for rapid reagent and vector development.

4.1.2.7. Focus on **development of new genomic related technologies/ analytical/computational methods** including supercomputing, data science, electrophoresis, northern blot, Bioinformatics, gene synthesis, chip sequencing etc. It is important to develop capabilities in Supercomputers and computational efficiency to interpret the available data.

4.1.3. Data collection and Open data sharing

Robust data collection and open data sharing is key to scientific collaborations and successful genome sequencing as it enables instantaneous and easy access to critical information. In this direction, following are the recommendations:

4.1.3.1. Data Collection and annotation of samples for sequencing: For the sequencing data to generate meaningful and actionable information during a pandemic there is a need for a **robust sampling strategy** (that considers geographic spread and skewed distribution of testing sites), **sample collection** (what sample should be sent for sequencing, how it should be collected) and **sample annotation**. The sample annotation (Metadata) includes information on the patient; age, sex, vaccination history, symptoms, clinical outcomes (mortality and morbidity).

There is a need to have a **designated pathway for sample collection**. This should span the who, what, when where and how of data collection. There is a requirement to have personnel other than those administering healthcare to oversee collection of meta-data so that it does not interfere with routine care. There is a scope to **leverage new technologies** such as app-based technologies and have an integrated database for capturing and correlating metadata. There is a need for a **data quality assessment of the collected samples** ideally before they reach the site. Uniform protocols for data collection, storage and transportation to sequencing labs with inbuilt quality checks at each point should be undertaken.

4.1.3.2. **Publicly accessible databases -** While genomic sequencing technologies are becoming accessible to the general population, information on variants associated with several major diseases are not available in publicly accessible databases. Creating a **centralized database** of variants of pathogens that facilitates early detection and mitigation of health risks in individuals must be prioritised. Open access to data is key to enabling greater collaboration and innovation.

4.1.3.3. **Data from labs/ institutions**: A lot of Indian labs and institutions (CSIR-IGIB, NIBMG, CCMB, etc) generate biological data from projects funded by public money. However, most of this data is not accessible to India's own researchers. Data sharing should be brought under the ambit of RTI and clear policies should be established that allow for sharing of data between govt institutions.

4.1.3.4. Constitute **uniform data standards and standardized nomenclature for genomic sequencing data**[15]. The Genomic Standards Consortium is developing metadata standards for harmonized data for genomics, software tools, data formats etc. India could build on these and contextualize it to our local situation.

4.1.3.5. **Finalize policy for biological data storage, access and sharing** - The

[15] https://www.ebi.ac.uk/ena/browser/about/data-standards

Government has given emphasis on an open data policy through the National Data Sharing and Accessibility Policy (NDSAP). Pursuant to the NDSAP, the Ministry of Science and Technology had released a zero draft on biological data storage, access and sharing in 2019, however it is yet to be finalized. Further, a national repository for life science data has just been launched in India - The Indian Biological Data Centre (IBDC). An appropriate policy should mandate that the data generated from publicly funded research be archived in IBDC.

4.1.3.6. **Harmonization with legal framework for data sharing -** With the Personal Data Protection Bill, pending before the Parliament, any proposed policy for genomic data sharing must be harmonized with the anticipated legislative mandate.

4.1.3.7. Create a **reagent repository** such as Addgene (https://www.addgene.org/) of all genomics reagents such as plasmid, primers accessible to researcher and industries for rapid circulation and supply.

4.2. Recommendations w.r.t. key policy and regulatory changes

Open Science Policies play a crucial role in removing obstacles to the free flow of research data and ideas. For genomics research to advance, data must be accurate, verifiable and robust in order to address the trust deficit which presently stems from various issues including lack of specific standards, low levels of co-ordination and interoperability, as well as poor data quality and interpretation.

Towards this, there is a need to have a **stronger policy and regulatory regime** addressing various aspects of genomic sequencing. To address all such issues, following are the recommendations:

4.2.1. In order to enable and strengthen research and open science, policy makers need to ensure development of **adequate data governance models, interoperable standards, sustainable data sharing agreements** involving public sector, private sector and civil society. We need to strengthen **healthcare data privacy laws**. Personal health record must be accessible to healthcare providers through interoperable networks. There is a need for legislative and technical safeguards to protect such sensitive data on the one hand. On the other, there is a need for fostering interoperability and avoiding information blocking to facilitate access in emergency situations.

4.2.2. Develop **policies for genomic surveillance** which incorporates routine surveillance and pandemic surveillance and lays down specific organizational structures and responsibilities.

4.2.3. Draft policies to tackle the pandemic by way of **increased repurposing of drugs**, or new interventions should be clear and transparent enough to give an idea about the pre-clinical and clinical studies required for regulatory approval. The shortest possible, yet robust, path for approval is the

requirement for emergency use authorization.

4.2.4. A national policy needs to be created utilizing Government, Universities, Medical institutions both public and private to create a **policy on pandemic management**. The current policy needs to be revisited as it has been formulated many years ago.

4.2.5. Adoption of policies to enable **faster clinical trials** for novel research by promoting / insisting industry and academia collaboration in India. Policies enabling the IRBs to approve such clinical trials and DCGI giving fast NOD to such trials needs to be envisaged. Programs developed in India should be given preferences to accelerate and promote research in India for addressing local needs.

4.2.6. There is a need for development of a concrete **policy for encouragement of domestic industry for production of kits/consumables** for genomics research.

4.2.7. Finalization and implementation of various other policies related to **Orphan drug designation** and regulatory review process and guideline.

4.2.8. Finalization and implementation of **Gene therapy regulatory review** process and guideline.

4.2.9. Setting up a **Subject Expert Committee** with experts on genomics, or inclusion of such experts in existing committees.

4.2.10. Development of policy/guideline to include **real-world data, use of machine-learning, AI based algorithms, and medical devices.**

4.2.11. Need for **certification and quality assurance mechanisms for devices and reagents produced indigenously**. In genetic services domain proficiency testing that checks the accuracy of the lab processes should be made mandatory, currently NABL doesn't have the proficiency testing component available except for Covid testing.

4.2.12. Review of **Intellectual Property Law Framework**: In order to foster the environment of research and innovation, there is a need to strengthen the intellectual property laws in India, particularly related to genomics and data. Scope of data that must be made publicly available and shared must be clearly defined, with sufficient protection/incentives provided for promoting scientific developments. On 23rd July, the Standing Committee on Commerce tabled a report on 'Review of the Intellectual Property Rights Regime in India' which recommends a holistic review of the IPR policy. The report recommends structural changes such as expanding innovation ecosystem in India, organising awareness drives on IPR, comprehensive advisories, increasing R&D activities, encouraging IP financing and increasing involvement of state governments in evolving a robust IPR regime. In line with this, any framework for genomic research in India must also address the

legitimate concerns of the scientific community owing to weak protections for intellectual property.

4.2.13. Another important area to focus upon is the **security at entry points** that includes airports, sea- ports or land routes or border areas, through which there is a possibility of transmission of pathogenic organisms. Our country should reframe rules for import or export of biological samples such as serum, seeds, food, plant samples etc. and strict rules should be implemented at all entry and checking points. Moreover, the suspicious samples collected at entry points if any, could be checked by the national laboratories before allowing its entry in the country.

4.2.14. Create an independent **monitoring and evaluation mechanism** with an appropriate appellate mechanism.

4.2.15. Follow **international biohazard and surveillance policies** for control and transport of synthetic genetic materials. (e.g, IGSC International Gene Synthesis Consortium: https://genesynthesisconsortium.org).

4.3. Recommendations w.r.t. Strengthening science establishments and critical scientific collaborations

Preparedness and response systems to manage diseases need pre-defined responsibilities and quality collaboration among all stakeholders *i.e.*, researchers, healthcare service providers, caregivers, patients, industry, policy makers etc.

This research necessitates a grand alliance of interdisciplinary / transdisciplinary researchers. For example, Biodiversity researchers will need to lead specimen acquisition; population biologists will need to track ecological interactions in the field; pathobiologists will need to optimize sampling protocols; genomicists will need to generate complete genomic libraries for all species and screen samples for pathogens. To be fully effective, epidemiologists, computer scientists, and mathematicians will then need to assemble these data to evaluate risks and propose mitigation measures for implementation by public health managers and policymakers.

Following recommendations are submitted towards Strengthening Science establishments and critical scientific collaborations:

4.3.1. A **National Level Consortium** comprising of Next-Gen Sequencing facilities, Cryo-EM facilities, and computational facilities can be formed so that everybody can work in harmony to address the issue quickly.

4.3.2. Set up an integrated **National Level Task Force** for pandemic preparedness (to assess the healthcare resource utilization, contingency plan, etc). The aforesaid task force should aim to promote, support and implement genomics and data analytics in all aspects of drug discovery, development, manufacturing, diagnostics and medicine. The objective is to bring genomics into mainstream medicine by developing robust and standardised health genomic pathways into

all aspects of health and medicine. The task force with appropriate representation from all sectors would aim to develop and foster innovation and research opportunities.

4.3.3. Planning and developing **Joint Centers of Excellence** on **infectious diseases and health monitoring** may be very useful. Further constant monitoring of health for a variety of other existing infectious diseases including seasonal flus through such Centres will help in combating or prevention of recurrence of pandemics.

4.3.4. Establishing **Joint Centers of excellence on vaccine design and development** including scientists (computational and structural biologist, protein design, immunologist, virologist, cell biologist etc) from both industry and academia/research institutes.

4.3.5. **Open access platforms for scientific publications** - Having open access to publications helps to form a fuller picture of the current understanding of a disease within the scientific community. In the COVID context, while there were certain initiatives to open access to publications in the early months of the pandemic, it remains to be seen how these might endure in the long run.

4.3.6. Funding Infrastructure for digital collaboration - **Adequately funded infrastructure** to enable easy data exchange, availability of data storage repositories, creation of digital collaboration platforms will allow preservation and utilisation of disease data and can help prepare for future emergencies.

4.3.7. Newer emerging technologies shall enable quick handling of the pandemic. Further, development and administration of vaccine to masses require efficient data handling and analysis. **Artificial intelligence (AI), Machine Learning (ML) and data science expertise** may help finding early solutions for controlling the pandemic relatively faster. Therefore, Institutes or establishments with capabilities of using or AI, ML or robotics may be bolstered. A continuous research on platform development to handle and foster such situation is much needed.

4.3.8. Strengthening of **partnerships between research labs and hospitals**. Possibly all major biology labs should adopt 2-3 tertiary care hospitals to promote genomic understanding and research in the medical field.

4.3.9. Implement **One Health approach** for identifying emerging pathogens and threats requiring strong collaborations between biodiversity researchers, population biologists, genomics experts, epidemiologists, structural and computational biologists, mathematicians and modeling experts, public health experts (both advisors and implementers), policy makers and state and central government. This may be tied in with the One health institution envisaged under the PMASBY for better impact.

4.3.10. India can look at **newer programs such as the Pandemic Genomics**

program[16] and pandemic interception systems[17] and set up mechanisms to be a part of and contribute to these networks.

4.3.11. Country should at least **establish five National Bioagent Research Laboratories** that would be working extensively on highly pathogenic organisms having potential to cause future outbreaks. Locations of setting up these 5 national laboratories is most significant, as these should be set up at places that they can extend their research over entire country and could also co-ordinate well with each other. Keeping these points in mind, these 5 laboratories should be planned and setup in North (Uttarakhand), East (West Bengal), West (Rajasthan), Sea shore (Maharashtra/Pune) South (Chennai). The laboratories should be well-equipped with all essential genomics, proteomics and metabolomics facilities and equipments for pathogen surveillance, detection, development of treatment (vaccine/antidote/drugs etc) research. Trained technical staff/scientists having immense expertise in field of disease diagnosis and surveillance strategies to plan and conduct research accordingly.

4.3.12. Government should specifically work on **identification of potential pathogens that can cause pandemic or epidemic**. Dedicated labs should regularly have surveillance, diagnosis of potential pathogens, Containment of pathogens when there is a outbreak to control the spread, Development of treatment of infections caused by the pathogenic agents. There is a need to have appropriate biocontainment mechanisms for facilities that engage in this kind of research so as to prevent accidental leal.

4.3.13. Both **research and production capacity building** are required to handle such pandemic spread. The recent learning suggests a big gap in terms of vaccines production or manufacturing capabilities. Upscale existing facilities and create new facilities to address this gap. It is also important to emphasize that we need to foster local production of quality raw materials and equipment's needed for vaccine/drug manufacturing. For diagnostics too having a biobank with well annotated samples would be very valuable.

4.3.14. Focused and dedicated basic research towards the development of novel platform technologies, which with the little modifications can produce vaccines for the novel emerging variants of a pandemic pathogen is the most sought capability if we really want to reduce dependence on western countries or China.

4.3.15. Separate **faculty to train on pandemics and their management** should be a part of every course for all stakeholders managing pandemics. Global interactions and bodies like WHO should provide opportunities and create a workforce of trainers using the train the trainer concept.

4.3.16. Develop **patient advocacy groups** for genetic disorders to generate more awareness and disseminate government schemes/policies

[16] https://www.oxfordmartin.ox.ac.uk/pandemic-genomics/

[17] https://www.pnas.org/content/117/25/13852

4.3.17. **Uniform Electronic Health Records** across private and public institutions (common platform for exchange of health information) as envisaged under ABDM. EHR can be linked with Unique Health ID for ease of data capture.

4.3.18. Certain facilities are required at the **Institute level / Hospital levels**: (IITs/IISc/IISER/AIIMS/Medical Schools/Hospitals) such as:

- Availability of the NexGen Sequence facility
- Availability of Cryo-EM experimental facility
- Availability of Gene and Plasmid Synthesis facility
- Availability of computational resources for data storage and high-end parallel computation.
- Availability of cell culture facility of laboratory level trials
- India is a treasure trove of flora, a library can be made about the resources and their genome to proteome database can be created for this purpose.
- Genome to proteome database can be established for the medicinal plants and their therapeutic metabolites.
- Proper animal facility is required for producing therapeutic raw materials or animal trials.

4.3.19. Certain facilities are required at the **industry level**:

- Support for designing new drug molecules
- Software for genetic sequence analysis and visualisation
- Support for clinical trials especially Phase I units
- Support for mass production of the drugs
- Support for building new industries for production of raw material and cultivation of therapeutic raw materials for the same purpose.

4.4. Recommendations w.r.t. Capability Building / skilling

Genomic surveillance will need a skilled and trained workforce that includes laboratory personnel, data analysts, epidemiologists, computational biologists, scientists, healthcare staff etc. In view of this, we have the following recommendations:

4.4.1. There is urgent need to **equip medical colleges with Departments of medical genetics** so that infrastructure and trained manpower is available to take up genomic sequencing as and when needed for infectious diseases or other medical conditions.

4.4.2. **Capacity building** needs to be taken up at all levels including Technicians, Bioinformaticians, Scientists etc. National Genomics core (funded by DBT) at CDFD is a good example and same can be emulated at least one centre in a large state.

4.4.3. As a part of routine training, young researchers and MSc students could have a **module on genomic sequencing** to understand its importance. This should also be integrated into public health specializations.

4.4.4. **Specific skill training** and continuing education for laboratory personnel and data analysts on newer techniques as they emerge.

4.4.5. Capabilities in **Artificial intelligence (AI), Machine Learning (ML) and big data analytics** need to be developed and strengthened. Artificial intelligence based newer methods such as Google's alpha fold should be leveraged for predicting pathogen structure based on available data.[18]

4.4.6. **Bioinformatics capabilities** need to be strengthened to generate heat maps and perform mathematical modelling for improvised predictions on spread of pandemic, likely emergence of resistant variants and next outbreak/wave. Accordingly, the next generation bioinformatics tools, data mining and storage facilities and adequate resources may be built.

4.4.7. In a pandemic situation, many Government officers are assigned to handle tasks quite different from their regular duties. Such officers should have **formal training in managing pandemics** and their role should be specified.

4.4.8. Training programmes for scientists for **oligo, gene and plasmid synthesis**, sequencing, quantification techniques designing advances in bio-nanotechnology like protein nanocages as vaccine and drug delivery vehicles and also to develop such nanocages, and development of appropriate mammalian cell lines for scaleup production of vaccine protein nanocages.

4.4.9. Training **computational scientists** to analyze the proteomic, genomic and immune-informatics data, and rapid screening of vaccine candidates against serum collected from the population.

4.4.10. IITs and similar academic institutions could possibly consider offering **programs in Computational Biology and Bioinformatics degrees**. There is a dire need for trained manpower with high-end quantitative skills in this area.

4.4.11. Another important aspect is the **need for skilled nurses and clinicians**. This group of trained staff should have specific certified training about the management of critical illnesses and potential pandemics.

Conclusion:

The power of genomic surveillance and its ability to meaningfully guide and inform appropriate public health response in a pandemic will require collaboration between multiple stakeholders. Going forward, the industries, ministries, academia and R&D labs would need to step up and collaborate to strengthen the country's genomic surveillance and sequencing capacity.

****

[18] https://deepmind.com/blog/article/AlphaFold-Using-AI-for-scientific-discovery

Confederation of Indian Industry

The Confederation of Indian Industry (CII) works to create and sustain an environment conducive to the development of India, partnering Industry, Government and civil society, through advisory and consultative processes.

CII is a non-government, not-for-profit, industry-led and industry-managed organization, with over 9000 members from the private as well as public sectors, including SMEs and MNCs, and an indirect membership of over 300,000 enterprises from 294 national and regional sectoral industry bodies.
For more than 125 years, CII has been engaged in shaping India's development journey and works proactively on transforming Indian Industry's engagement in national development. CII charts change by working closely with Government on policy issues, interfacing with thought leaders, and enhancing efficiency, competitiveness and business opportunities for industry through a range of specialized services and strategic global linkages. It also provides a platform for consensus-building and networking on key issues.

Extending its agenda beyond business, CII assists industry to identify and execute corporate citizenship programmes. Partnerships with civil society organizations carry forward corporate initiatives for integrated and inclusive development across diverse domains including affirmative action, livelihoods, diversity management, skill development, empowerment of women, and sustainable development, to name a few.

As India marches towards its 75th year of Independence in 2022, CII, with the Theme for 2021-22 as *Building India for a New World: Competitiveness, Growth, Sustainability, Technology*, rededicates itself to meeting the aspirations of citizens for a morally, economically and technologically advanced country in partnership with the Government, Industry and all stakeholders.

With 62 offices, including 10 Centres of Excellence, in India, and 8 overseas offices in Australia, Egypt, Germany, Indonesia, Singapore, UAE, UK, and USA, as well as institutional partnerships with 394 counterpart organizations in 133 countries, CII serves as a reference point for Indian industry and the international business community.

CONFEDERATION OF INDIAN INDUSTRY

The Mantosh Sondhi Centre

23, Institutional Area, Lodi Road, New Delhi – 110 003 (India)

T: 91 11 45771000 / 24629994-7

E: info@cii.in • W: www.cii.in