

# Data Sharing of Genetic Resources

perspectives from an INSDC member

Masanori Arita  
Bioinformation & DDBJ Center  
National Institute of Genetics, Japan

[arita@nig.ac.jp](mailto:arita@nig.ac.jp)

10.5281/zenodo.14411962

# INSDC for open data sharing

International Nucleotide Sequence Database Collaboration (INSDC) started in 1982.



NCBI



## Databases

The same platform is used for personal genomes, not openly shared among countries.

Data type	DDBJ	EMBL-EBI	NCBI
Next Generation reads	<a href="#">Sequence Read Archive</a>	European Nucleotide Archive	<a href="#">Sequence Read Archive</a>
Assembled Sequences	DDBJ		<a href="#">GenBank</a>
Samples	<a href="#">BioSample</a>		<a href="#">BioSample</a>
Studies	<a href="#">BioProject</a>		<a href="#">BioProject</a>

Metadata includes BioProject and BioSample with sequencing method details.

# INSDC details

New website: insdc.org

Click “Announcements”  
to see our message  
mandating geolocation.

**INSDC spatiotemporal metadata – minimum standards update  
(03-03-2023)**

The screenshot shows the INSDC website homepage. At the top left is the INSDC logo. To the right of the logo are navigation links: "About INSDC", "Global Participation", "Technical Specifications", and "Announcements". The "Announcements" link is circled in red. Below the navigation is a large blue banner with the text "International Nucleotide Sequence Database Collaboration". Underneath the banner is a paragraph of text: "The International Nucleotide Sequence Database Collaboration (INSDC) archives nucleotide sequence data, from raw to assembled and annotated sequences, from around the world." Below this text is a search bar with the placeholder "Search term" and a magnifying glass icon. At the bottom of the page, there are two featured sections: "Global Participation" with an image of a person at a microscope, and "Technical Specifications" with an image of server racks. At the very bottom, there is a section titled "Access INSDC Member Participating Databases" with a subtext: "For full details of how to submit data or access data, please select an INSDC Member." A red arrow points from the "Announcements" link to the "INSDC spatiotemporal metadata" text box on the left.

We expand our membership to other continents / regions to cover more global information (membership criteria are available).

# Convention on Biological Diversity (<https://cbd.int>)



Convention on  
Biological Diversity

International treaty of 193 parties, no US (since 1993)

*Three objectives:*

1. The conservation of biological diversity
2. The sustainable use of the components of biological diversity
3. The fair and equitable sharing of the benefits arising out of the utilization of genetic resources

Nagoya Protocol (2014) is the procedure on **Access** to *Genetic Resources* and the *Fair and Equitable Sharing of Benefits* Arising from their Utilization. (**ABS**)

# Digital Sequence Information

COP15 (Kunming-Montreal; Dec 2022):

- Biodiversity finance of \$20 billion / year (by 2025)
- Benefit sharing from DSI

COP16 is now ongoing in Columbia (10/21 – 11/1).

BS from DSI is in dispute:

**B**iodiversity on areas **B**eyond  
**N**ational **J**urisdiction (BBNJ; 2023

WHO Pandemic Treaty, and

FAO **P**lant **G**enetic **R**esources for  
**F**ood & **A**griculture (ITPGR-FA).



# DSI is nationally owned and managed



CBD and its Nagoya Protocol is legally binding. Now the COP says we need to share benefits from DSI.

- Does DSI include *images*?
- How shall we license data?
- Compatibility with ongoing principles, e.g., FAIR, CARE and TRUST.

UNESCO recommends open science but it is not legally binding.

<https://www.unesco.org/en/open-science/about>



# Challenges

International treaties are *NOT* harmonized in terms of data sharing policy and management.

Treaty	CBD	BBNJ	WHO Pandemic	ITPGRFA
Target information	DSI (not yet clear)	DSI (not yet clear)	Virus genomes	Crop genomes
Repository or Indexing site	Clearing House in each nation	Global clearing house	GISAID	Global Information System (GLIS)
Open Access / FAIR data	Nation-wise	YES	NO	?
Benefit Sharing mechanism	Under discussion	multilateral	No benefit; COVAX?	multilateral
Related Documents	<a href="https://www.cbd.int/abs/">https://www.cbd.int/abs/</a>	IISD Bulletin: <a href="http://bit.ly/bbnj5res">bit.ly/bbnj5res</a>	UN Doc SSA2/CONF./ 1Rev.1	CBD.INT ITPGRFA-DSI. pdf

# Why not open?

**COVID-19 pandemic told us important issues.**

1. Travel ban to South Africa who quickly reported the Omicron variant
2. Only 0.3% vaccine distribution to LMIC (when 30% is disposed)
3. Disinformation on vaccines and other remedies



**Lessons learned:** Incentivize and credit data depositors. Enable high-quality metadata. Invest for the long term.

# Re-balancing to avoid data fragmentation

From equality to equity in data sharing

✓ INSDC guarantees data backup and exchange.

China and India are important partners.

We keep all submitted records permanently.

✓ INSDC guarantees FAIR and TRUST.

We also commit to inclusion, but not authority.

✓ Capacity building is the key.

# Conclusions

- All scientists should know about CBD discussions.
  - Data and 'knowledge' are not easily separated.
  - DSI needs labels for countries and communities (IPLC).
- The current system is not sustainable. We need more discussion for a change.
- Future repositories support data provenance. FAIR and equitable access are important.
  - Capacity building in data literacy and analysis
  - Cooperation with multiple fora