

FOUNDERS ARRANGEMENT

BETWEEN

**THE NATIONAL LIBRARY OF MEDICINE, NATIONAL CENTER FOR
BIOTECHNOLOGY INFORMATION,**

AND

**THE RESEARCH ORGANIZATION OF INFORMATION AND SYSTEMS, NATIONAL
INSTITUTE OF GENETICS,**

AND

**THE EUROPEAN MOLECULAR BIOLOGY LABORATORY - EUROPEAN
BIOINFORMATICS INSTITUTE,**

CONCERNING

**OPERATIONAL INTENT OF THE INTERNATIONAL NUCLEOTIDE SEQUENCE
DATABASE COLLABORATION (INSDC)**

This Founders Arrangement is between the National Library of Medicine-National Center for Biotechnology Information, (hereafter, NLM-NCBI), a component of the National Institutes of Health, U.S. Department of Health & Human Services; the Research Organization of Information and Systems, National Institute of Genetics (hereafter, ROIS-NIG) in Japan; and the European Molecular Biology Laboratory - European Bioinformatics Institute (hereafter EMBL-EBI), hereinafter referred to individually as Founder or collectively as Founders.

Whereas the Founders build and maintain repositories and services for the collection and dissemination of nucleotide sequence data that is freely accessible and redistributable without any restriction to promote science and human health; and

Whereas, the Founders have collaborated successfully in the past and are convinced of the importance of furthering this collaboration; and

Whereas the Founders share an interest in further strengthening relations and broadening the reach of the collaboration; and

Whereas the Founders believe that this arrangement provides the general framework for collaborating on such joint activities.

NOW, THEREFORE, the Founders have reached the following understanding:

1. Purpose

The International Nucleotide Sequence Database Collaboration (INSDC) is a global collaboration of independent government or non-profit organizations that manage nucleotide sequence databases that capture and preserve nucleotide sequence information and annotations to create a comprehensive collection that preserves the scientific record and enables broad sharing of such data. Founders archive nucleotide sequence data, from raw to assembled and annotated sequences, from around the world and connect these data, where applicable, to associated scientific publications and the biological specimens from which the data were derived.

In support of this purpose, INSDC members coordinate their efforts and exchange data to create comprehensive publicly available sequence data resources in accordance with mutually agreed upon standards to ensure that the data are findable, accessible, interoperable, and reusable. In addition, they coordinate the issuance of accession numbers to provide a unique persistent identifier to refer to the sequences in participating databases, publications, and other resources in which data may be cited.

Originally established in the 1980s,¹ the INSDC is committed to open science and data sharing. The INSDC values free and unrestricted access to data and is supported by journal publishers as the recommended repositories for nucleotide sequence data. INSDC helps preserve nucleotide sequence data as part of the scientific record and serves as a core element of the global data infrastructure for bioinformatics. Sequence data publicly available through INSDC supports science necessary for the study of human health, food safety and security, and biodiversity, and other important domains.

2. Principles

The Founders are committed to the following principles:

1. Acquiring, processing, accessioning, distributing, and exchanging nucleotide sequence data in ways that ensure data are findable, accessible, interoperable, and reusable to the greatest extent possible;
2. Free, broad, and timely sharing and dissemination of nucleotide sequence data and associated annotations and metadata without limitations on use or redistribution;
3. Adopting quality control procedures, as well as formats and conventions, that enable consistent data submission and support regular data exchange among Founders so that users can access the content from any Founder organization;
4. Collecting rich descriptive metadata describing the biological specimen from which submitted sequences are derived and rich feature annotations describing sequences to enhance reusability;
5. Providing data infrastructure that is responsive to changes in technology and evolving needs of the user community;

¹ Stevens, H. Globalizing Genomics: The Origins of the International Nucleotide Sequence Database Collaboration *J Hist Biol.* 2020 Mar;53(1):209-210.

6. Operating with sustainable funding models to support long-term acquisition and hosting of nucleotide sequence data, with contingency planning for smooth and controlled closure of services and handover of data in cases of catastrophic loss of funding;
7. Ensuring the security of the data collected and shared by and amongst Founders; and
8. Maintaining regular communications amongst Founders to ensure broad awareness of INSDC activities.

3. Governance

INSDC is governed by two committees: an Executive Committee and an Implementation Committee, as detailed below.

a. INSDC Executive Committee (EC)

The Founders plan to each appoint not more than two individuals to the Executive Committee (EC) which has responsibility for the overall governance of the INSDC. The EC is responsible for guiding the strategic direction of the INSDC.

The functions of the EC comprise:

- Establishing the strategic objectives of INSDC;
- Directing INSDC administrative activities, including organizing meetings and ensuring that content for the INSDC website is comprehensive and reflects current policy;
- Maintaining awareness of new technologies to inform strategic direction of INSDC;
- Resolving any disputes that arise among INSDC members by reconciling differences of opinion or approach through consensus;
- Setting expectations and objectives for membership in the INSDC in consultation with the Implementation Committee (IC). The EC may evaluate potential INSDC members based on capacity and proficiency in the activities outlined in a separate Membership Arrangement and make recommendations to the IC; and
- Convening at least 4 meetings per year.

b. INSDC Implementation Committee (IC)

The INSDC Implementation Committee (IC) is expected to consist of not more than two representatives from each data resource or activity supported by each Founder. The IC intends to meet at least once per year. Meetings will be chaired by an EC member. The EC will decide which agenda items require a vote.

The functions of the IC comprise:

- Reviewing and providing formal acceptance of INSDC-wide policies, protocols, and procedures, including the INSDC Membership Arrangement;
- Preparing for and responding to changes in sequencing and assembly technologies;
- Approving new modification for infrastructure requirements, features, and schemas used to build INSDC-participating repositories and their services;

- Training, assessing, and approving new INSDC members based on the recommendations of the EC by majority vote, with a single vote per Founder; and
- Maintaining website content and internal documentation including meeting notes.

4. International Stakeholders Committee (ISC)

The INSDC is intended to create an ad hoc International Stakeholders Committee (ISC) comprised of external experts who are expected to promote the principles and activities of INSDC and provide input on different stakeholders' needs. Each INSDC Founder is intended to appoint two ISC members to serve in this role for a term of 3 years. A chair is expected to be chosen by ISC members. The ISC is expected to convene annually to report on the needs of the external stakeholder community.

5. INSDC Expansion beyond Founders

- Founders may invite to participate in INSDC organizations that represent regions or communities not represented by the Founders; manage government-recognized or multinational non-commercial sequence databases that capture, preserve, and present nucleotide sequence information and annotations; and meet the criteria for membership as specified in a separate Membership Arrangement. Standing invitations for representatives in the EC, IC and ISC may be issued to additional organizations, as defined in the Membership Arrangement.

6. General Provisions

- The Founders may modify this Arrangement by mutual consent in writing. Any Founder may discontinue cooperation under this Arrangement at any time but should endeavor to give the other Founders sixty (60) days advance written notice.
- Only the EC may authorize public communications to be issued on behalf of the INSDC. Any public communications issued on behalf of the INSDC are expected to be reviewed and approved by the EC prior to publication.
- This Arrangement imposes no funding obligations on the Founders to support INSDC functions. Nothing in this Arrangement authorizes or is intended to obligate the Founders to expend, exchange, or reimburse funds, services, or supplies for INSDC operations. Any financial arrangement between the Founders is expected to be subject to a separate agreement.
- No indemnification for any loss, claim, damage, or liability is intended or provided by the Founders under this Arrangement. Each Founder is responsible for its own activities under this Arrangement.
- The Founders intend to resolve any differences arising under this Arrangement amicably through mutual consultations and negotiations.
- This Arrangement reflects the intentions of the Founders and does not create any legally binding obligations among the Founders.

- This Arrangement may be executed in counterparts, each of which should be deemed to be an original and all of which together constitute a single document.