



## Appendix 1 to the Grant Proposal of BiCIKL. Letters of Support and Descriptions of the Collaborating Infrastructures



6 March 2020

To whom it may concern:

Letter of Support: BiCIKL

I am writing on behalf of the Biodiversity Heritage Library (BHL, <https://biodiversitylibrary.org/>), the world's largest open access digital library for biodiversity literature and archives. BHL operates as a worldwide consortium of natural history, botanical, research, and national libraries working together to digitize the natural history literature held in their collections and making it freely available for open access as part of a global biodiversity community.

BHL makes use of the taxonomic framework provided by the Catalogue of Life (CoL) to facilitate discovery and use of literature associated with any species. BHL is also working to increase metadata and mark-up of content within the library to allow the information contained in these publications to be integrated into international databases such as CoL and GBIF. BHL supports the efforts of Plazi to automate the processing of publications to extract structured content for wider use.

If the BiCIKL proposal is funded, BHL will participate in its efforts to develop standards, tools and pipelines and seek opportunities for BHL contents to become more accessible in forms compatible with those developed by Pensoft, Plazi and EJT.

Sincerely yours,

A handwritten signature in black ink, appearing to read 'Martin Kalfatovic', written over a white background.

Martin Kalfatovic  
BHL Program Director | Assistant Director, Smithsonian Libraries  
MRC 154 PO Box 37012  
Washington, D.C. 20013-7012  
202-633-1705  
[kalfatovicm@si.edu](mailto:kalfatovicm@si.edu)

[https://drive.google.com/open?id=1brwQRPDmXe4820DDHDc3ZDQ2CPFZG\\_qI](https://drive.google.com/open?id=1brwQRPDmXe4820DDHDc3ZDQ2CPFZG_qI)



Mar 10, 2020

Centre for Biodiversity Genomics  
University of Guelph  
50 Stone Road East  
Guelph, ON  
N1G 1Y2

Letter of Support: BiC IKL

To whom it may concern

I am writing on behalf of the Centre for Biodiversity Genomics in Guelph (CBG), which serves as the Secretariat for the International Barcode of Life Consortium (iBOL) and maintains the key informatics platforms supporting this global community. These platforms include the Barcode of Life Data Systems (BOLD, <http://boldsystems.org>), which maintain the reference system for DNA barcode sequences, associated specimen records, and taxonomic units they support, and the Multiplex Barcode Research And Visualization Environment (mBRAVE, <http://www.mbrave.net>), which supports large-scale projects based on high-throughput sequencing (HTS) instruments.

CBG has recently received funds from CANARIE to increase FAIR and open access to data held in BOLD and mBRAVE and to increase interoperability between this infrastructure and related biodiversity infrastructures globally.

If the BiC IKL proposal is funded, CBG will be well positioned to support European partners, including Species 2000 (for Catalogue of Life), GBIF, EMBL and UNITE, in their efforts to integrate data products and services from BOLD and mBRAVE as part of an interoperable framework of biodiversity infrastructures.

Yours sincerely,

A handwritten signature in black ink that reads 'Sujeevan Ratnasingham'.

Sujeevan Ratnasingham  
Associate Director, Informatics



To whom it may concern,

Letter of Support: BiC IKL

This letter serves to confirm our intent to support this proposal. The Global Genome Biodiversity Network (GGBN, <http://www.ggbn.org>) recognises that the development of interoperable FAIR data infrastructure requires research. The Botanic Garden and Botanical Museum Berlin is hosting the Technical Secretariat of GGBN approved by the GGBN members and their experience in many different biodiversity informatics projects predestines them to perform this project successfully.

The GGBN Data Portal links globally distributed databases and bridges the gap between biodiversity repositories, sequence databases and research results. It is maintained and hosted by the BGBM and currently provides access to more than 3 million tissue and DNA samples provided by 27 partners.

A critical bottleneck is the missing link between published sequences at INSDC and the underlying biological material. Standardized sequence data workflows and tools would improve data quality and traceability of molecular data enormously. Such traceability is crucial for the implementation of the Nagoya Protocol on a global scale and one key research topic of GGBN.

The GGBN therefore agrees on collaborating with the proposed project, by: (i) evaluating possibilities for establishing communication channels within and abroad GGBN for a sustainable infrastructure development, (ii) organizing telephone conferences and meetings with partners, such as EMBL, NCBI, DDBJ, GBIF, and (iii) promoting the project in presentations.

GGBN aims at bridging gaps between existing biodiversity infrastructures. The BiC IKL proposal serves as a great example to reach this goal.

A handwritten signature in blue ink, appearing to read 'Katharine Barker'.

Katharine Barker  
GGBN Coordinator  
Smithsonian Institution, National  
Museum of Natural History

A handwritten signature in blue ink, appearing to read 'Christina Löfdahl'.

GGBN Executive Committee Chair  
Natural History Museum Denmark

[https://drive.google.com/open?id=16th-o8g60QGO3z0nveBQg0\\_24LXP6dV6](https://drive.google.com/open?id=16th-o8g60QGO3z0nveBQg0_24LXP6dV6)

The [International Barcode of Life Consortium](http://ibol.org)<sup>1</sup> (iBOL) is a partnership between member nations, represented by key institutions active in biodiversity genomics. iBOL's goal is to implement a cost-effective DNA-based identification system for all species on earth and to advance the use of this system to benefit science and society. iBOL is currently working to deliver its second major program of activity, [BIOSCAN](http://bioscan.org), bringing together a wide range of stakeholders to accelerate development of the reference

---

<sup>1</sup> <http://ibol.org>



library for DNA barcode sequences, to include at least two million species and to optimise protocols to use this library as a tool to survey and monitor ecosystems and probe community interactions globally.

The Centre for Biodiversity Genomics in Guelph, ON, serves as the Secretariat for iBOL and hosts two major data platforms. The [Barcode of Life Data Systems](#)<sup>2</sup> (BOLD) platform, maintains the barcode reference library, comprising millions of records for curated specimens and their sequences, and delivering the Barcode Index Number (BIN) system that clusters these sequences, associates them with scientific names, and underpins DNA-based species identification. The [Multiplex Barcode Research And Visualization Environment](#)<sup>3</sup> (mBRAVE) is a workbench for multiplexed projects using high-throughput sequencing platforms for metabarcoding or library construction.

iBOL will participate in BiCIKL by facilitating FAIR and open access to data from BOLD and mBRAVE and contributing to design and integration of these data within the taxonomic framework services.

[Biodiversity Heritage Library](#)<sup>4</sup> (BHL) is the world's largest open access digital library for biodiversity literature and archives. BHL is revolutionizing global research by providing free, worldwide access to knowledge about life on Earth. BHL operates as a worldwide consortium of natural history, botanical, research, and national libraries working together to digitize the natural history literature held in their collections and making it freely available for open access as part of a global biodiversity community.

The BHL portal provides free access to hundreds of thousands of volumes, comprising over 57 million pages, from the 15th-21st centuries. In addition to public domain content, BHL works with rights holders to obtain permission to make in-copyright materials openly available under Creative Commons licenses. The BHL consortium works with the international taxonomic community, publishers, bioinformaticians, and information technology professionals to develop tools and services to facilitate greater access, interoperability and reuse of content and data. BHL provides a range of services, data exports, and APIs to allow users to download content, harvest source data files, and reuse materials for research purposes. Through the Global Names Recognition and Discovery (GNRD) service, BHL indexes the taxonomic names throughout the collection, allowing researchers to locate publications about specific taxa.

BHL will contribute to BiCIKL by participating in efforts to develop standardised pipelines for presenting structured content from digital literature and incorporating this content for use within the taxonomic framework services.

[The Global Genome Biodiversity Network](#)<sup>5</sup> (GGBN) is based on a Memorandum of Cooperation and is an unincorporated, international network of member organizations, which share the aim of making high-quality, well-documented and vouchered genomic samples of the Earth's biodiversity discoverable for research. The mission and objective of GGBN is to foster collaborations among biodiversity repositories in order to comply with quality standards, policies, best practices, interoperability and exchange of material in accordance with national and international legislation and conventions, thereby benefiting society through additional research contributing to development and biodiversity conservation.

Most successful research programs depend on easily accessible and standardized research infrastructures. GGBN fills the gap for other biodiversity collections, with a goal of standardized high-quality samples for genomic research, a major bottleneck for many sequencing initiatives. Presently GGBN's core activity is enabling access to searchable DNA and tissue collections across natural history

---

<sup>2</sup> <http://www.boldsystems.org/>

<sup>3</sup> <http://www.mbrave.net/>

<sup>4</sup> <http://biodiversitylibrary.org/>

<sup>5</sup> <http://www.ggbn.org>



museums and botanic gardens. Activities are gradually being expanded to encompass all kinds of biodiversity biobanks such as culture collections, zoological gardens, aquaria, arboreta, and environmental biobanks, etc. providing these collections with additional opportunities for making samples discoverable for research, increasing their value and relevance. GGBN facilitates sample search and discovery for its distributed member collections through a single entry point. It stores standardized information on mostly geo-referenced, vouchered samples, their physical location, availability, quality, and the necessary legal information on over 70,000 species of Earth's biodiversity (predominantly eukaryotes), from unicellular to multicellular organisms. The GGBN Data Portal and the GGBN Data Standard are complementary to existing infrastructures such as the Global Biodiversity Information Facility (GBIF) and International Nucleotide Sequence Database (INSDC). GGBN continues to grow its collections strategically, based on the needs of the research community, adding over one million online records in 2019 alone, and today over three million sample data are available through GGBN.