The Digital Botanical Gardens Initiative

Authors

- **The Digital Botanical Gardens Initiative Consortium** ✉
    · ⓞ [digital-botanical-gardens-initiative](#)

✉ — Correspondence possible via [GitHub Issues](#) or email to The Digital Botanical Gardens Initiative Consortium <dbgi@protonmail.com>.

Abstract

The Digital Botanical Gardens Initiative (DBGI) ambitions to explore innovative solutions for the collection, management and sharing of digital information acquired on living botanical collections. A particular focus will be placed on the large scale characterization of the chemodiversity of living plants collections through mass spectrometric approaches. The acquired data will be structured, organized and connected with relevant metadata through semantic web technology. The gathered knowledge will then inform ecosystem functioning research and orient biodiversity conservation projects. The DBGI initially aims to take advantage of the readily available living collections of Swiss botanical gardens to establish robust and scalable chemo- and biodiversity digitisation workflows. The ultimate goal is to apply these approaches in the field and at the global scale in wild ecosystems.
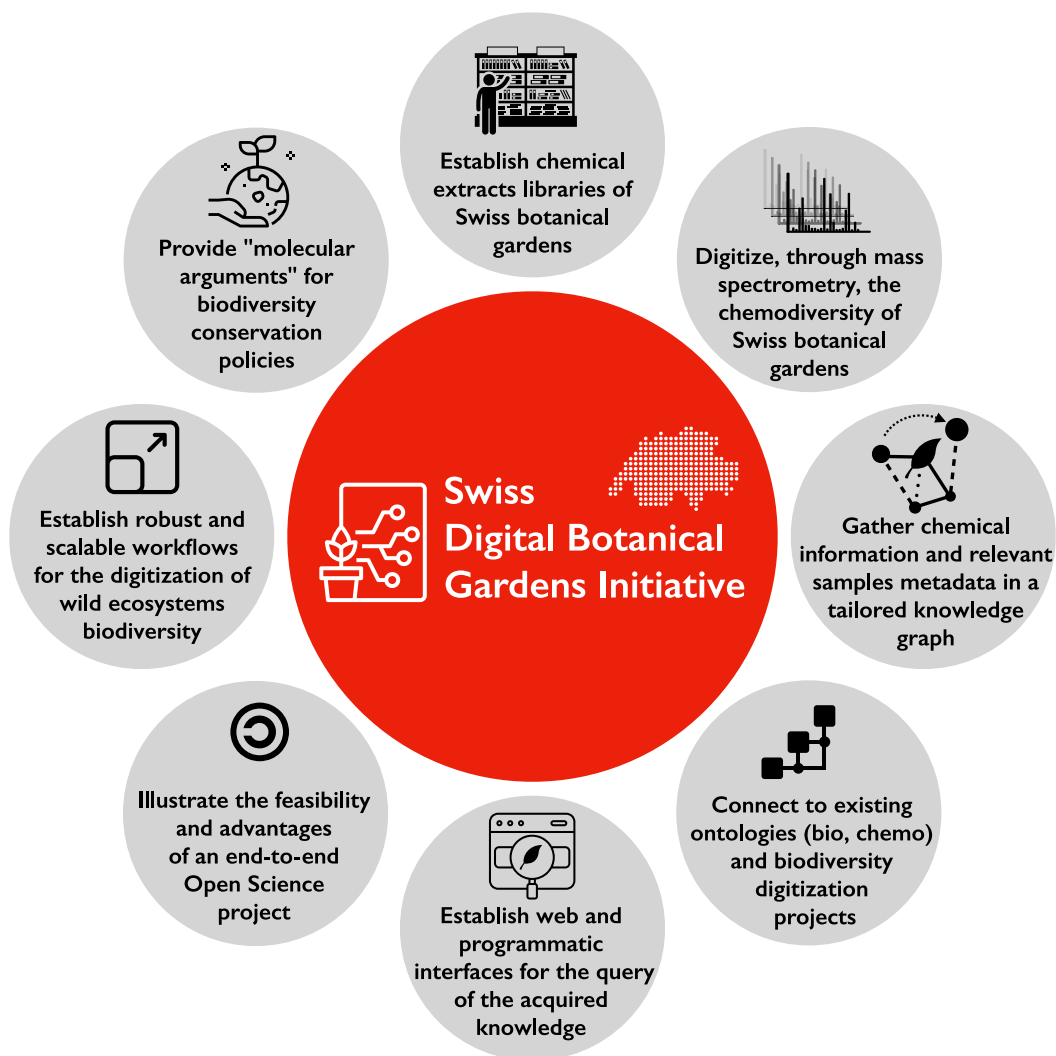
Goals

Figure 1: Main goals of the Digital Botanical Gardens Initiative.

The main goals of the DBGI are resumed in Figure 1 Some details are given hereafter:

1. Establish chemical extracts libraries of Swiss botanical gardens. These chemical libraries can be considered as complementary to herbarium samples. They are easily conserved over time and in a reasonable space. They represent the chemical diversity of a sample. They can be easily aliquoted. They can be screened for bioassays.

2. Digitize, through mass spectrometry, the chemodiversity of Swiss botanical gardens. Here high-resolution mass spectrometry is considered as a digital scanner allowing to capture the chemical fingerprint of the profiled sample. State-of-the-art computational metabolomics solutions are used to organize and annotate the acquired spectra with molecular data.

3. Gather chemical information and relevant samples metadata in a tailored knowledge graph. Chemical information acquired at the previous step (spectra and chemical structures) are connected to relevant samples metadata (taxonomy, phenology, geolocalisation, time of collection etc.). For this semantic web technologies (namely the RDF data model) are employed and a tailored knowledge graph is established.

4. Connect to existing ontologies (bio, chemo) and biodiversity digitization projects. Chemical and biological objects of the graph are connected to relevant pre-established ontologies (e.g. CHEBI,

[Plants Ontology](#)) and data graphs (e.g. [Wikidata](#)). Connection with complementary biodiversity digitization efforts will be done (e.g [BiCIKL](#), [Dissco](#))

5. Establish web and programmatic interfaces for the query of the acquired knowledge. A web interface will allow a convenient access to the data acquired within the framework of the project. A dashboard will allow simple visualizations (e.g pie charts, barplots, treemap) to interpret the data. In addition a SPARQL endpoint and an application programming interface (API) will allow retrieval of data programmatically.

6. Illustrate the feasibility and advantages of an end-to-end Open Science project. The DBGI will strictly follow the [Open Science](#) guidelines by using open-source software and making available the acquired data and scripts under an open license agreement. In addition the DBGI results will be published at the moment they are acquired (previous to formal publication or even pre-prints) thus following the [Open Notebook Science](#) concepts.

7. Establish robust and scalable workflows for the digitization of wild ecosystems biodiversity. The DBGI, albeit ambitious, is a pilot project. The future objective is to propose digitization workflows for the characterization of the wild ecosystems chemodiversity, on a global scale. These workflow will be central to the future Earth Metabolome Initiative.

8. Provide "molecular arguments" for biodiversity conservation policies. The ultimate goal of DBGI is to use all the gathered metabolic information to support, extend or implement conservation efforts worldwide. We believe that by providing chemical maps of the landscape it will be possible to contribute to the prioritization of conservation and restoration targets . In other words, by establishing large scale chemical maps we expect to provide "molecular arguments" to biodiversity conservation endeavors (e.g. *"This piece of land has a high content of antibacterial scaffolds."* or *"This place might be poor in species diversity but rich in a rare chemodiversity".*)

Background and context

Biodiversity is a major determinant of ecosystem stability [1]. Hundreds of studies spanning terrestrial and aquatic ecosystems support that higher levels of biodiversity, in all its forms, promote better ecosystem functions, such as carbon sequestration, underpinning human well-being [2,3]. Sadly, earth is experiencing a major biodiversity crisis and of the estimated nine million species of fungi, plants, and animals have been described [4], about a million species are currently at risk of extinction and may go extinct before the end of the century [5]. One major issue is that more than 80% of the estimated biodiversity still awaits to be described. We are in fact facing what is now called the Anthropocene extinction (sixth mass extinction) [6]. In order to try to deviate from these alarming trends, all possible efforts must be made by the responsible (i.e. our species) for the conservation of biodiversity. For this, the characterisation and documentation of biodiversity is a fundamental prerequisite. Over 3.5 billion years of evolution, natural selection, the craftsman of biodiversity, has created an overwhelming array of molecular entities. Myriad compounds are produced by all living organisms from bacteria to whales, forming the backbone of the ever-growing tree of life. Through the lens of chemistry, every species, biotic interaction, and community, reveals a unique ensemble of molecular structures: the metabolome. These chemical assemblages are a valuable, yet largely unexplored reflection of biodiversity and ecosystem functioning. To go beyond the simple quantitative representation of biodiversity provided by species inventories, and to reinforce our understanding of links between biodiversity and ecosystem functioning, we see chemical diversity as an alternative and highly complementary view of the diversity of our planet's ecosystems.

Rationale

With these urgent biodiversity characterization objectives in mind we are setting up the Digital Botanical Gardens Initiative (DBGI), which aims to develop robust and scalable workflows for the digitization of botanical gardens using several approaches. The central one being the use of analytical chemistry strategies to build information-rich chemical maps that will guide researchers focusing on biodiversity characterization and conservation. Sampling plants for chemical characterization can be done in three ways, either from natural living collections (botanical gardens), in the wild, or by growing plants in highly controlled settings (See Figure 2 for an overview of the advantages and inconveniences of each biodiversity source). For this first project, the sampling in botanical gardens is the chosen option, as within a very accessible location, thousands of species, which are already identified, labeled and organized, can be readily sampled.
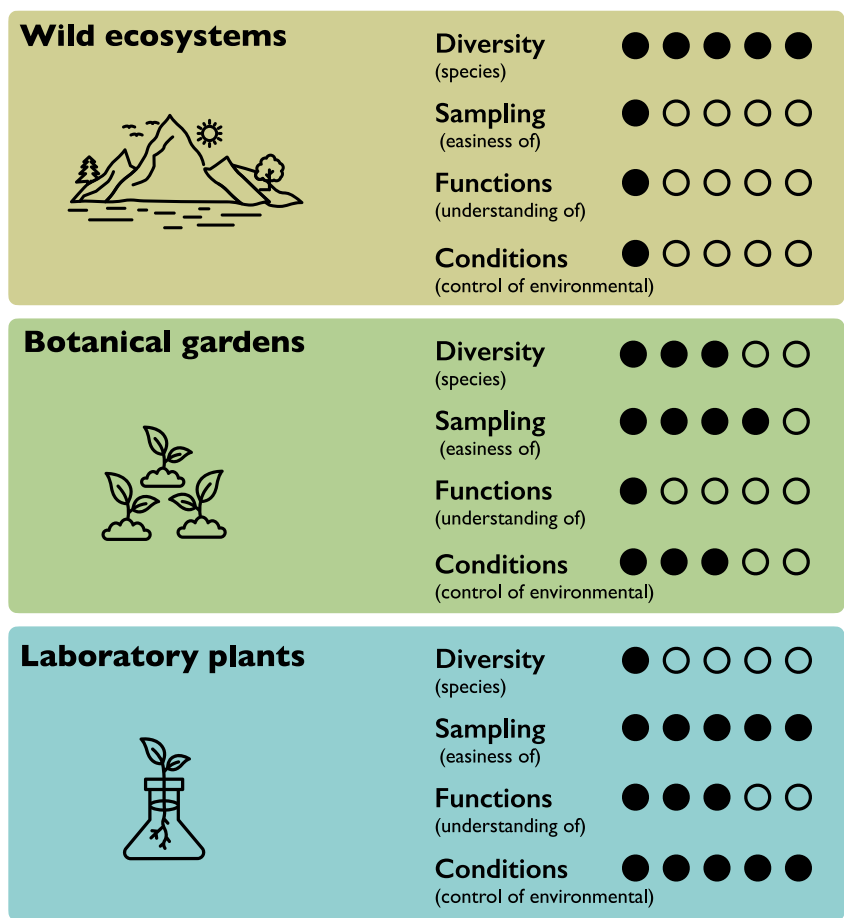

Figure 2: Characteristics of three main plant biodiversity sources for the scientist.

Outline of the research

The aim of the DBGI is to characterize the chemodiversity of the all botanical gardens of Switzerland. This ambitious objective requires us to evaluate and test the entire workflow and methodology. For this, the initiators of the DBGI started gathering preliminary data from two botanical collections: the Jardin Botanique de l'Université de Fribourg (JBUF) and the Jardin Botanique de Neuchâtel (JBN). This choice was made for practical reasons (these are the respective working places of the DBGI initiators) but also because these two gardens each offer their unique characteristics. On one hand the JBUF presents over 5000 species systematically organized to reflect the Angiosperm Phylogeny Group (APG) system. The JBUF researchers are specialized in conservation biology. On the other hand, the JBN presents over 2000 species organized in 8 sub-gardens. The JBN has a focus on ethnomedicinal plants. Below, we will briefly outline the research workflow envisioned for the DBGI. The main steps have been schematized in Figure 3. The overall workflow can be divided into two main parts: one dealing with physical objects (upper part of Figure 3) and a second one dealing with data and metadata acquired from these objects (lower part of Figure 3).

Physical objects. Starting from a botanical garden (a living specimens collection), the aim is to sample each specimen and to build two libraries: a library of dried plant material and a library of chemical extracts. The dried plant material library will serve as a backup for extractions to be repeated, further characterization of compounds, or for orthogonal analysis (e.g. genetic sequencing). The chemical extracts library will be the source material for the mass-spectrometry digitization stage. This library will also be available for backups and orthogonal analysis (e.g. NMR profiling or bioassays campaigns). Complementary to herbarium, these two libraries offer an efficient way (reduced space, long term storage) to capture and conserve the chemistry of the botanical gardens.
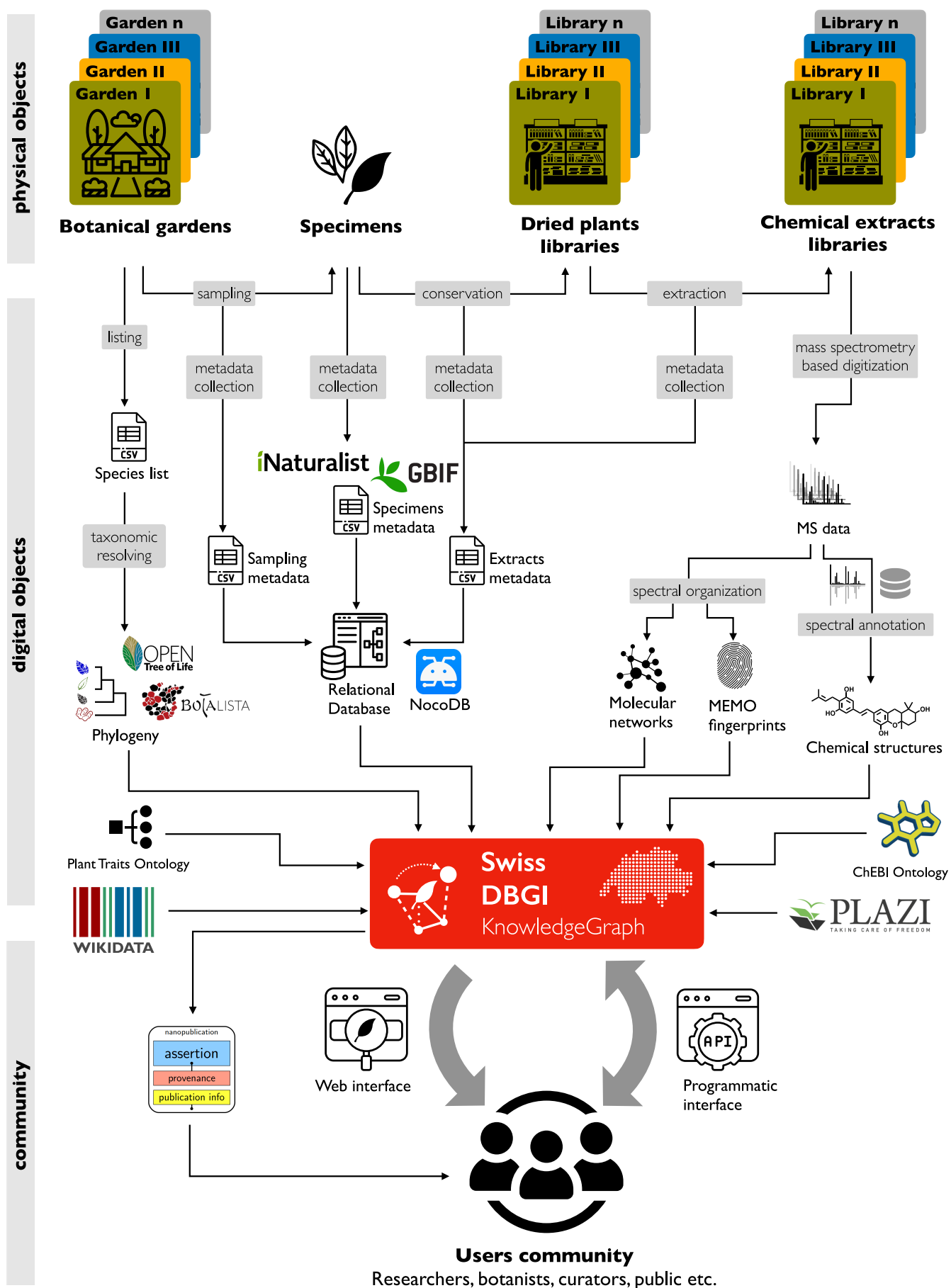
Figure 3: Data types and fluxes in the DBGI.

Digital objects. For all operations occurring on physical objects (sampling, conservation, extraction), metadata are collected to document the experiment. For each botanical garden, data collection is made at the species level. A species is collected for each garden, curated and taxonomically resolved (using the Open Tree of Life framework) in order to be compared across gardens. The Botalista platform will also be exploited at this step. For each collected sample, data are acquired at a finer granularity (namely at the specimen level). Here we take advantage of the iNaturalist platform and app. Using a smartphone, pictures of the sampled specimen (including eventual label in the botanical gardens, sampled organ and collection label), collector identity, date and geolocation are conveniently captured at the time of the collection. This data is then automatically collected by the iNaturalist DBGI project. The data of the project can then be programmatically accessed via the iNaturalist API. All species, specimens and experimental metadata will be collected and managed through an SQL database and accessed through a NocoDB instance for a convenient tracking of the samples by the DBGI participants.

The mass-spectrometry digitization then constitutes the core of the chemical information acquisition process. We use Ultra High Performance Liquid Chromatography coupled to High Resolution Mass Spectrometry (UHPLC-HRMS) to acquire fragmentation data in an untargeted fashion. Building on our computational metabolomics expertise we then organize and annotate the acquired pool of MS data. Here we will take advantage of five central tools (four of which were conceived by us). Molecular networking [7] will serve as a basis for spectral organization. The metabolite annotation will be performed by spectral matching against a theoretical natural products spectral database [8] and via a taxonomically informed scoring process [9] fed by the widest open resource of natural products biological occurrences (LOTUS [10]).

Dissemination of the results

The DBGI conducts this project following the Open Science principles, and by following the [Open Notebook Science](#) concepts. This approach allows all research artifacts (e.g. research proposals, drafts, ideas, source code, raw and processed data etc.) produced in the frame of the DBGI to be publicly available immediately, from the moment of their production, and not only after peer-reviewed publication. To implement the DBGI Open Notebook we employ [Dendron](#), an open-source and lightweight note-taking and knowledge management software. Dendrons are built using an ingenious system of markdown files hierarchically organized based on their filename. This allows for extremely efficient note searching and refactoring of the hierarchies. Dendrons can be conveniently shared across members of the DBGI, versioned via git and automatically published as websites. The DBGI Dendron can be browsed at http://www.dbgi.org/dendron-dbgi/. Regarding raw data sharing, specimen-related information are hosted on the iNaturalist DBGI project home page and pictures are shared under a permissive CC0 license allowing further reuse in Wikidata for example. Mass spectrometry profiles and metabolite annotation files will be hosted on the [MassIVE data platform](#) where they will benefit from a permanent DOI.
All the code written in the frame of the DBGI will be publicly shared and versioned through the [DBGI Github organization](#). Ideas, comments and issues will be collected using [Github discussions](#).

The DBGI Consortium

Note: the DBGI is welcoming any researchers interested to join the initiative. We expect people to participate according to their ressource/time/interest at the level they suit.

Initiators

- Pierre-Marie Allard ([COMMONS Lab](#), University of Fribourg)
- Emmanuel Defossez (Institute of Biology, University of Neuchâtel & JBN)

Core team

> **⇥ Join the core team !**
> You have ca. 2 hours / week to dedicate to the DBGI plus time for 6 meetings a year ? You are willing to lead a group of Collaborators (see below) or gather Expert knowledge (see below). Join us in one of the following areas. We need you !

⚒ Sampling & Collections Management

🔬 Metabolomics & Computational Mass Spectrometry

⌁ Linked data & Knowledge management

🔓 Open Science & Dissemination

👀 Fundamental & Applied Research (from functional ecology to drug discovery)

🌍 Biodiversity Conservation

Collaborators

> **⇥ Collaborate !**
> You are willing to put your hands in the dirt and concretely contribute to the DBGI under any of the following areas. We need you ! If you have expert knowledge to share but not that much time, please see Expert section.

⚒ Sampling & Collections Management

- Edouard Bruelhart ([COMMONS Lab](#), University of Fribourg)

🔬 Metabolomics & Computational Mass Spectrometry

⌁ Linked data & Knowledge management

- Maëlle Wannier ([COMMONS Lab](#), University of Fribourg)

🔓 Open Science & Dissemination

👀 Fundamental & Applied Research (from functional ecology to drug discovery)

🌍 Biodiversity Conservation

Experts

🧪 Sampling & Collections Management

🔬 Metabolomics & Computational Mass Spectrometry

🕸 Linked data & Knowledge management

🔓 Open Science & Dissemination

👀 Fundamental & Applied Research (from functional ecology to drug discovery)

🌍 Biodiversity Conservation

Contacts

- ☑ Donat Agosti (Plazi)
- ☑ Nadir Alvarez (Université de Genève)
- ☑ Caterina Barillari (ETHZ)
- ☑ Cyril Boillat (Botalista)
- ☑ Edouard Bruelhart ([COMMONS Lab](), University of Fribourg)
- ☐ Philippe Cudré-Mauroux (University of Fribourg)
- ☐ Patrice Descombes (University of Lausanne)
- ☑ Christophe Dessimoz (University of Lausanne)
- ☐ Jakub Galgonek (Elixir Prague)
- ☐ Gaétan Glauser (University of Neuchâtel)
- ☑ Gregor Kozlowski (JBUF)
- ☑ Henry Lütcke (ETHZ)
- ☐ Laurent Kneubuhl (Botalista)
- ☑ Tarcisio Mendes (SIB)
- ☑ Blaise Mulhauser (JBN)
- ☐ Marco Pagni (SIB)

- ☐ Brian Sedio (University of Texas at Austin)
- ☑ Sergio Rassmann (University of Neuchâtel)
- ☑ Bernd Rinn (ETHZ)
- ☑ Ana-Claudia Sima (SIB)
- ☑ Thomas Walker (University of Neuchâtel)
- ☑ Ming Wang (University of Riverside)
- ☑ Maëlle Wannier ([COMMONS Lab](), University of Fribourg)

Botanical Gardens

🌱 **Botanical Gardens**
The following botanical gardens of Switzerland are part of the DBGI.
⧗ Collections and profiling are ongoing.
⧗ Botanical garden team has been contacted, next on the list !
☐ Botanical garden to be contacted.

- ⧗ Jardin Botanique de l'Université de Fribourg (JBUF)
- ⧗ Jardin Botanique de Neuchâtel (JBN)
- ⧗ Jardin Botanique de Lausanne (JBL)
- ☐ Conservatoire et Jardins Botanique de Genève (CJBG)

References

1. **Biodiversity and Ecosystem Functioning**
   David Tilman, Forest Isbell, Jane M Cowles
   Annual Review of Ecology, Evolution, and Systematics (2014-11-23) https://doi.org/gf5tpq
   DOI: 10.1146/annurev-ecolsys-120213-091917

2. **Biodiversity and Ecosystem Functioning: Current Knowledge and Future Challenges**
   M Loreau, S Naeem, P Inchausti, J Bengtsson, JP Grime, A Hector, DU Hooper, MA Huston, D Raffaelli, B Schmid, … DA Wardle
   Science (2001-10-26) https://doi.org/cnf65t
   DOI: 10.1126/science.1064088 · PMID: 11679658

3. **Linking biodiversity to ecosystem function: implications for conservation ecology**
   MW Schwartz, CA Brigham, JD Hoeksema, KG Lyons, MH Mills, PJ van Mantgem
   Oecologia (2000-02-23) https://doi.org/cjwdg4
   DOI: 10.1007/s004420050035 · PMID: 28308280

4. **Number of species on Earth tagged at 8.7 million**
   Lee Sweetlove
   Nature (2011-08-23) https://doi.org/djwm2d
   DOI: 10.1038/news.2011.498

5. **Global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services**
   IPBES
   Zenodo (2019-05-04) https://doi.org/grbxsr
   DOI: 10.5281/zenodo.6417333

6. **The misunderstood sixth mass extinction**
   Gerardo Ceballos, Paul R Ehrlich
   Science (2018-06-08) https://doi.org/ggf4kk
   DOI: 10.1126/science.aau0191 · PMID: 29880679

7. **Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking**
   Mingxun Wang, Jeremy J Carver, Vanessa V Phelan, Laura M Sanchez, Neha Garg, Yao Peng, Don Duy Nguyen, Jeramie Watrous, Clifford A Kapono, Tal Luzzatto-Knaan, … Nuno Bandeira
   Nature Biotechnology (2016-08) https://doi.org/f8zxbs
   DOI: 10.1038/nbt.3597 · PMID: 27504778 · PMCID: PMC5321674

8. **Integration of Molecular Networking and *In-Silico* MS/MS Fragmentation for Natural Products Dereplication**
   Pierre-Marie Allard, Tiphaine Péresse, Jonathan Bisson, Katia Gindro, Laurence Marcourt, Van Cuong Pham, Fanny Roussi, Marc Litaudon, Jean-Luc Wolfender
   Analytical Chemistry (2016-03-01) https://doi.org/f8v86w
   DOI: 10.1021/acs.analchem.5b04804 · PMID: 26882108

9. **Taxonomically Informed Scoring Enhances Confidence in Natural Products Annotation**
   Adriano Rutz, Miwa Dounoue-Kubo, Simon Ollivier, Jonathan Bisson, Mohsen Bagheri, Tongchai Saesong, Samad Nejad Ebrahimi, Kornkanok Ingkaninan, Jean-Luc Wolfender, Pierre-Marie Allard
   Frontiers in Plant Science (2019-10-25) https://doi.org/fm73
   DOI: 10.3389/fpls.2019.01329 · PMID: 31708947 · PMCID: PMC6824209

10. **The LOTUS initiative for open knowledge management in natural products research**
Adriano Rutz, Maria Sorokina, Jakub Galgonek, Daniel Mietchen, Egon Willighagen, Arnaud Gaudry, James G Graham, Ralf Stephan, Roderic Page, Jiří Vondrášek, … Pierre-Marie Allard
eLife (2022-05-26) https://doi.org/gpzrr8
DOI: 10.7554/elife.70780 · PMID: 35616633 · PMCID: PMC9135406