

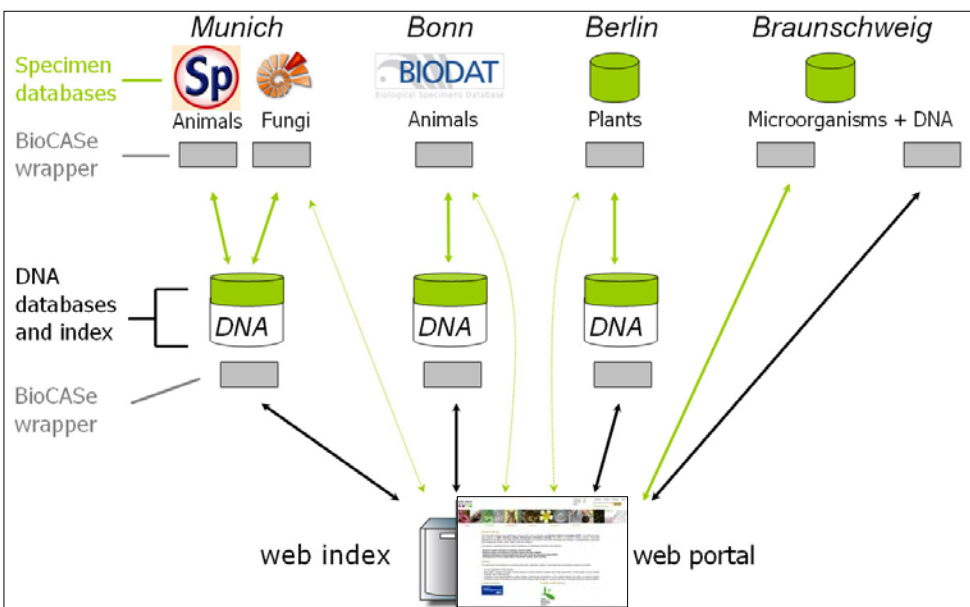
Gabriele Dröge, Holger Zetsche, Anton Güntsch & Birgit Gemeinholzer

Botanic Garden and Botanical Museum Berlin-Dahlem, Freie Universität Berlin, Königin-Luise-Str. 6-8, 14195 Berlin, Germany  
 email: g.droege@bgbm.org, h.zetsche@bgbm.org, a.guentsch@bgbm.org, b.gemeinholzer@bgbm.org

## Background

The DNA Bank Network connecting DNA banks of the Bavarian State Collection of Zoology, Munich (ZSM), the Botanic Garden and Botanical Museum Berlin-Dahlem (BGBM), the German Collection of Microorganisms and Cell Cultures, Braunschweig (DSMZ) and the Zoologisches Forschungsmuseum Alexander König, Bonn (ZFMK) was initiated under participation of GBIF Germany (Global Biodiversity Information Facility) and is funded by the DFG (German Research Foundation).

The network data architecture is based on the GBIF infrastructure and offers a worldwide unique concept. The separate DNA banks of the four network partners are linked and accessible via a central webportal, providing access to DNA samples with full online documentation. DNA sample documentation includes information about DNA quality, the related voucher and links to published molecular data.



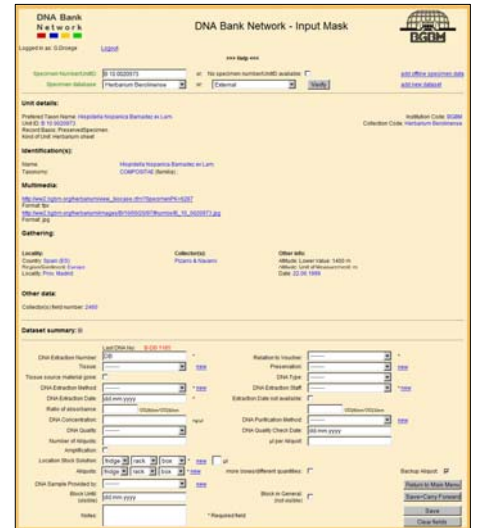
**Fig. 1. Data flow within the DNA Bank Network.**

Specimen data (green lines) are transferred to the DNA Module and the web portal by BioCAsE wrappers (grey boxes). Wrappers applied for specimen data of the network's databases are as well used by GBIF. DNA data (black lines) are transferred to the web index and web portal via a second BioCAsE wrapper installed on the DNA Modules and the database in Braunschweig.

## DNA Module

The DNA Module was developed as an open access product at the BGBM. It consists of a MySQL database as well as PHP forms for DNA data input and management. It is furthermore possible to add links to GenBank accessions and references of publications in which DNA bank samples have been used.

Additionally, customer requests like order or shipping data can be administered and the inventory can be managed.

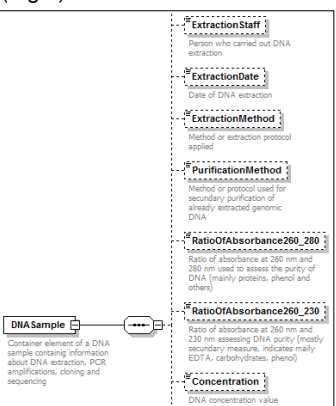


**Fig. 2. Input Mask of the DNA Module.**

Specimen data for *Hispidella hispanica* collected in Spain shown in the upper part of the screenshot are imported live from a specimen database (here Herbarium Berolinense) via the same wrapper used by GBIF. Related DNA data to be recorded using the input mask are shown in the lower part of the screenshot. Any GBIF compliant specimen database worldwide can be connected to the DNA Module following the BioCAsE or DIGIR protocol.

## DNA Extension for ABCD (ABCDDNA)

To present DNA data at the internet a BioCAsE wrapper had to be installed mandatory (Fig.1) on all connected DNA databases. The currently preferred ABCD 2.06 schema lacks important features in its existing DNA part 'Sequences'. A new DNA extension (ABCDDNA) was developed for that matter similar to the ABCDEFG extension for geosciences. So, the 'Sequences' part of the ABCD 2.06 should not be used any longer. Since the DNA mapping provides Triple Identifier for both the DNA sample and the specimen (Fig. 4 & 5) the central web portal is able to send two separate queries, one to the DNA database and a second one to the original specimen database (Fig.1).



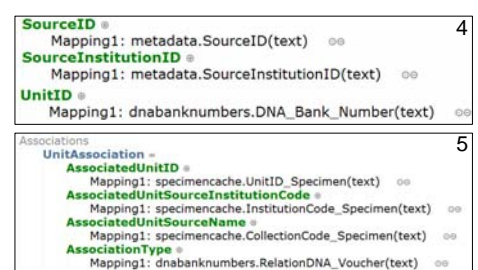
**www.dnabank-network.org**

Find out more about the DNA Module as well as the installation and mapping of the ABCDDNA schema.

A prototype with full search functionality will be published on our web portal, presumably by January 2009!

**Fig. 3. Detail of the DNA extension for ABCD.**

The base ABCD 2.06 version remains unmodified. The new DNA part covers more than 35 elements such as 'ExtractionDate', 'ExtractionMethod' for DNA extraction as well as an 'Amplification Container' for sequences, GenBank Numbers, clone strains etc.



**Fig. 4 & 5. Mapping the DNA database.**

Mapping the ABCDDNA schema requires two Triple Identifiers: one for the DNA sample (Fig.4) and another one for the related specimen (Fig.5). Since the ABCD Unit represents the DNA sample the UnitID has to be mapped as the (unique) 'DNA Bank Number'. The SourceID could, for example, be mapped as 'DNA Bank'. The Triple Identifier for the related specimen is defined by 'UnitAssociation' and must be identical to the Triple Identifier used in the original specimen database. 'AssociationType' is a mandatory field and should be mapped as shown in Fig.5.