



Barcoding projects

Kessy Abarenkov
University of Tartu, Estonia



What is DNA Barcoding?

Hebert et al. 2003. Biological identifications through DNA barcodes.

“... the sole prospect for a sustainable identification capability lies in the construction of systems that employ DNA sequences as taxon ‘barcodes’.”

Hebert et al. proposed a new system of species **identification** and **discovery** using a short section of DNA from a **standardized** region of the genome.

What is DNA Barcoding?

Hebert et al. 2003. Biological identifications through DNA barcodes.

“... the sole prospect for a sustainable identification capability lies in the construction of systems that employ DNA sequences as taxon ‘barcodes’.”

Hebert et al. proposed a new system of species **identification** and **discovery** using a short section of DNA from a **standardized** region of the genome.

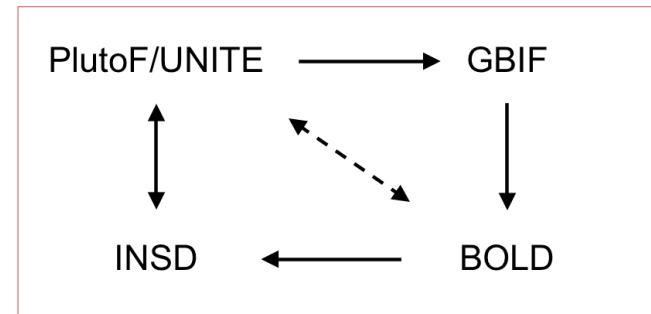
Features of a good DNA barcode:

- Easy to amplify from a wide spectrum of organisms (conserved primer sites for “universal” primers)
- Present in multiple copies (easier to amplify old type material)
- Preferably easy to align
- Enough variation to discriminate between closely related taxa
- Enough conservation to group together individuals of the same taxon

Official DNA barcode markers: COI/coxI (Animals), ITS (Fungi), MatK+rbcL (Plants)

Barcoding projects have the following components:

- Specimens
 - Expeditions, collecting specimens
 - Collections and herbaria
- Laboratory protocols
 - DNA extraction
 - PCR amplification
 - Sanger sequencing
- Data management and publishing
 - Data & metadata submission (NB! Links between voucher specimen carrying taxon name and DNA barcode)
 - Primary data storage (in an efficient way in terms of future updates)
 - Construction of public reference libraries
 - Example portals and their integration =>
- Data analysis and usage
 - Analysis software development
 - Identification of new records (researchers, applications, and general public)



Number of barcodes in BOLD



Number of species level taxa in INSD

Ranks:	higher taxa	genus	species	lower taxa	total
Archaea	179	160	610	0	949
Bacteria	1711	3157	16007	864	21739
Eukaryota	22379	75954	354894	25936	479163
Fungi	1625	5289	34505	1179	42598
Metazoa	15983	52255	180189	12984	261411
Viridiplantae	3027	15333	129708	11473	159541
Viruses	686	534	2174	0	3394
All taxa	24987	79812	373711	26800	505310

With the estimated 8.7 million eukaryotic species on our planet (Mora et al. 2011), and with 1.6 million being identified and classified (Catalogue of Life, 2016),

“Thousands of new species are discovered each year, but it will still take hundreds of years to find the rest.”



Barcoding projects in PlutoF

DNA barcoding sequences and linked data (specimen data, images, trace files, taxonomy) can be managed as projects in PlutoF.

Data can be further directly submitted to UNITE, prepared for INSD submission and published in GBIF.

Next steps – Case 2 hands on