Belgian Science Policy Office

belspo

GORA-MOLCOL

Ensuring Belgian Natural History Collections for the 22nd Century, towards a joint tissue and DNA Collection

JEMU

Joint Experimental Molecular Unit

Africa TERVUREN

museum

KATHOLIEKE UNIVERSITEIT LEUVEN

Section 1 : Molecular systematics - Exploring, testing, and establishing collection conservation protocols for future molecular systematic research

---

Section 2 : Bioinformatics - Developing, testing and improving an institutionally overarching joint collection data management, retrieval and accessibility system

Section 1 : Molecular systematics - Exploring, testing, and establishing collection conservation protocols for future molecular systematic research

WP 1.1 : Testing for optimal retrieval of genetic material

WP 1.2 : Development of protocols for field practices for molecular systematic collections

WP 1.3 : Incorporation of tissues and DNA material into a standardized collection management system

WP 1.4 : Development of guidelines on compatibility and interchangeability of field and specimen data sheets

WP 1.5 : Testing of samples for NGS

WP 1.6 : Exchange and incorporation of international standards

Section 1 : Molecular systematics - Exploring, testing, and establishing collection conservation protocols for future molecular systematic research

WP 1.1 : Testing for optimal retrieval of genetic material

WP 1.2 : Development of protocols for field practices for molecular systematic collections

WP 1.3 : Incorporation of tissues and DNA material into a standardized collection management system

WP 1.4 : Development of guidelines on compatibility and interchangeability of field and specimen data sheets

WP 1.5 : Testing of samples for NGS

WP 1.6 : Exchange and incorporation of international standards

Testing for optimal genetic retrieval

- DNA extraction on different type of museum specimens
  - Elien Aerts : Testcase on 8 fish samples stocked in formol (*Labeobarbus c.*), 1920-2006

Species	Collection number	Date	Sample type
Labeobarbus Cardozoi	MRAC 3238	1920	Fin clip
Labeobarbus Cardozoi	MRAC 3238	1920	Muscle biopsy
Labeobarbus Cardozoi	73-2-P-1523	1973	Fin clip
Labeobarbus Cardozoi	73-2-P-1523	1973	Muscle biopsy
Labeobarbus Cardozoi	73-2-P-1523	1973	Gills
Labeobarbus Cardozoi	85-19-P-2-3	1985	Fin clip
Labeobarbus Cardozoi	A7-009-P-0822-0825	2006	Fin clip
Labeobarbus Cardozoi	A7-009-P-0822-0825	2006	Muscle biopsy

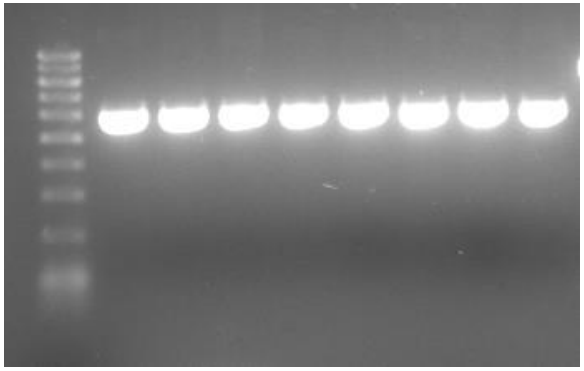
- Different kits were used but no successful results were obtained

➔ No more testing were planned (time and cost consuming)

Testing for optimal genetic retrieval

- DNA extraction on different type of museum specimens
  - DNA extraction of plant samples (dry leaves, fresh <2 years) with NucleoSpin® Plant II Columns kit

Adaptations from manufacturer's protocol :
*Cut the leave with scalpel, put in the tube, cut with scissors
*Add 200 µl lysis buffer first, crush little pieces of plant, then add 200 µl lysis buffer to rinse the crushing tool
*Add 10 µl RNAse for 1 hour then add 4 µl RNAse for 30 minutes in the thermomixer



Amplification rbcL (Nathalie Smitz)

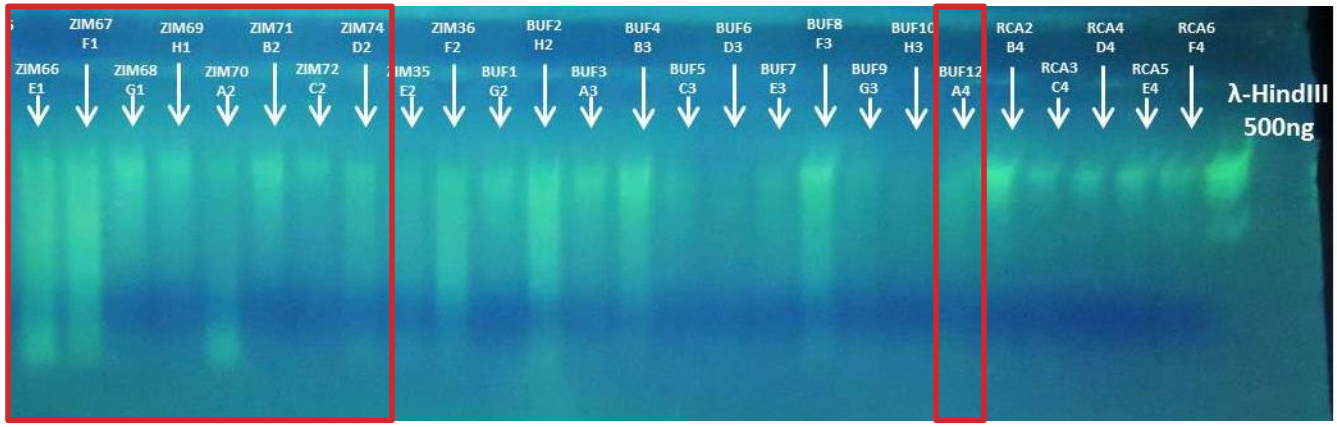
➔ Potential range of applications using museum plant specimens

Testing for optimal genetic retrieval

- DNA extraction on different type of museum specimens
  - DNA extraction of buffalo tanned skins (stored dried or in ethanol, 2011) with QIAGEN Blood & tissue kit

Adaptations from manufacturer's protocol :
*Cut a piece of tissue and leave it soak in absolute ethanol o/n
*Add 180 µl Buffer T1 and 40 µl Proteinase K Solution. Vortex to mix.
*Elution in 80µ-100µl of pre warmed elution buffer

Tanned skin

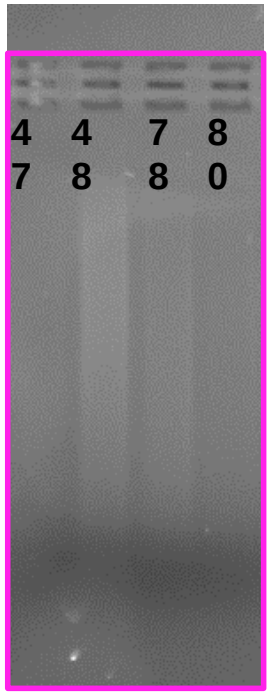


Agarose gel electrophoresis to evaluate quality of whole DNA extracted from buffalo samples tanned skins and hairs

Testing for optimal genetic retrieval

- DNA extraction on different type of museum specimens
  - DNA extraction of buffalo teeth (1947-1980) with the The QIAamp DNA Investigator Kit

Year Sample	ng/ul
1947	0.09
1948	22.5
1978	7.97
1980	1.73



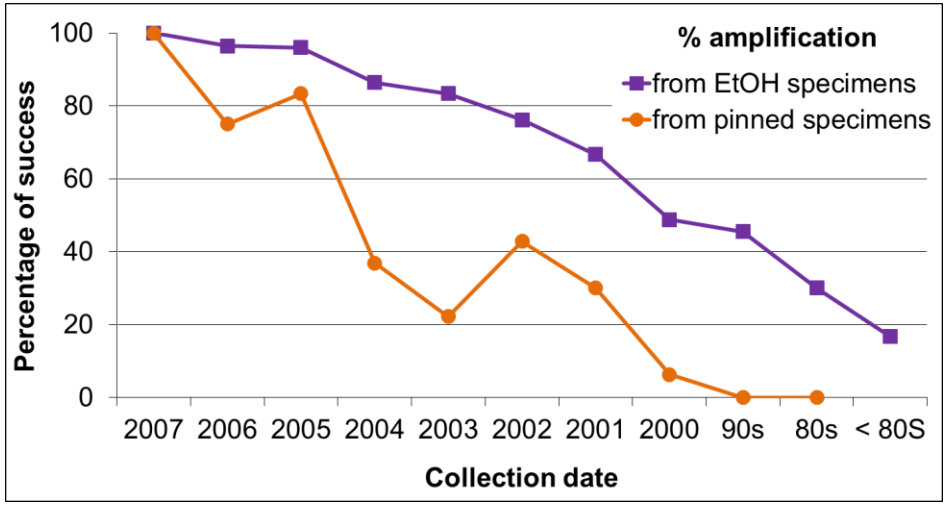
3 ul DNA loaded

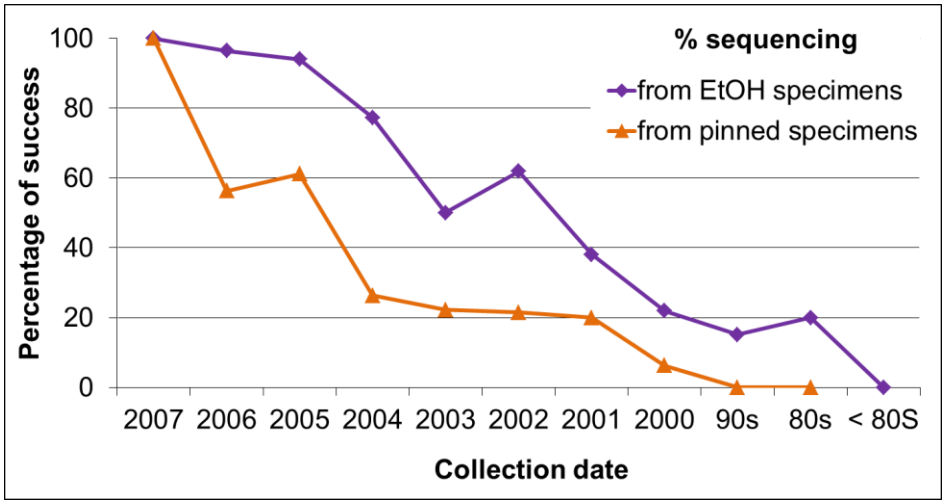
➔Could be useful for population genetics when fresh samples are not easily available

Testing for optimal genetic retrieval

- Effect of the age on DNA extraction ?
  - DNA extraction from dipteran specimens aged between 1980's-2007

Evaluation of the differences in terms of success of DNA amplification and sequencing when extracting DNA from pinned specimens (n=138) vs ethanol preserved specimens (n=256) (Massimiliano Virgilio)

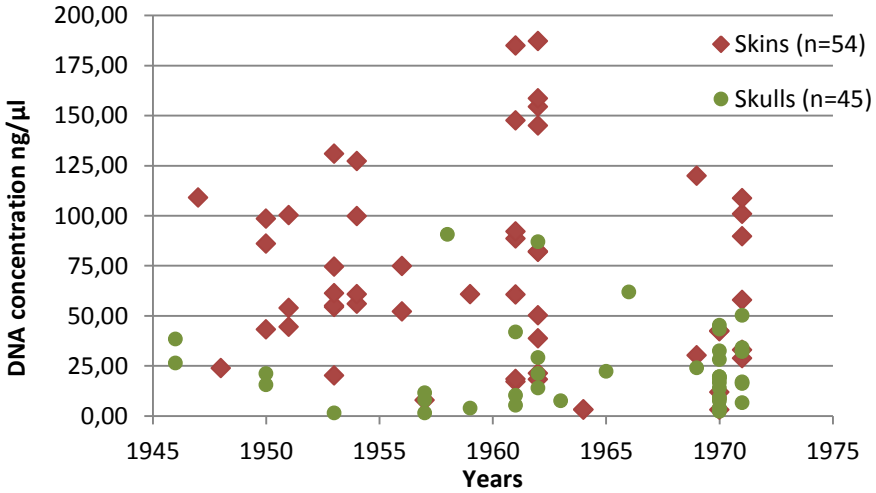


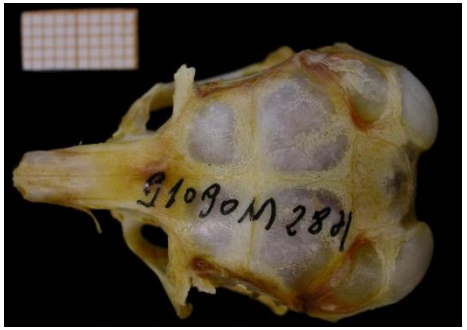


Testing for optimal genetic retrieval

- Effect of the age on DNA extraction ?
  - Comparing two kind of dry samples : skins vs skulls, but unknown drying procedures

Evaluation of the differences in terms of DNA concentrations when extracting DNA from micro pieces of soft tissues either from skulls (n=45) or skins samples (n=54) of Jaculus specimens (unpublished data provided by Gontran Sonet)

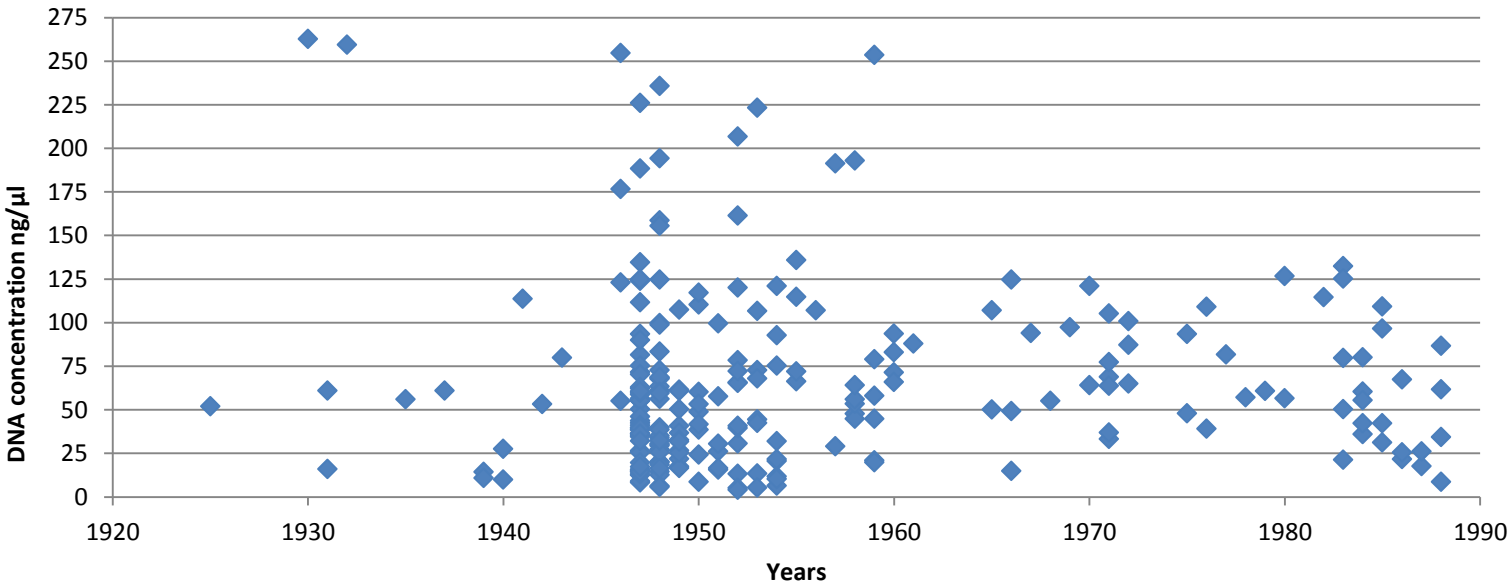




Testing for optimal genetic retrieval

- Effect of the age on DNA extraction ?
  - DNA extraction from Toe pads aged between 1920-1988

Evaluation of the differences in terms of DNA concentrations when extracting DNA from toe pads (n=186) of birds from Africa using DNA easy kit QIAGEN, digestion O/N and elution in 70 µl (Floris labbook)



Section 1 : Molecular systematics - Exploring, testing, and establishing collection conservation protocols for future molecular systematic research

WP 1.1 : Testing for optimal retrieval of genetic material

WP 1.2 : Development of protocols for field practices for molecular systematic collections

WP 1.3 : Incorporation of tissues and DNA material into a standardized collection management system

WP 1.4 : Development of guidelines on compatibility and interchangeability of field and specimen data sheets

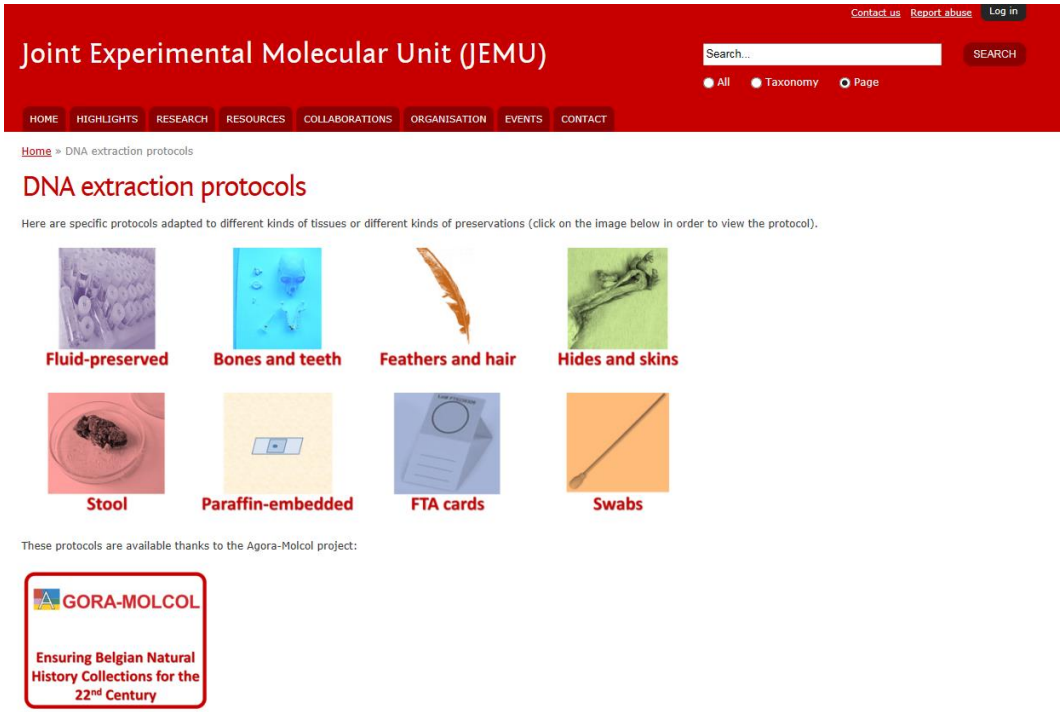
WP 1.5 : Testing of samples for NGS

WP 1.6 : Exchange and incorporation of international standards

Development of protocols for molecular systematic collections

- Development of lab protocols (collaboration with JEMU)
  - Development of general lab protocols for the RMCA and of specific DNA extraction protocols

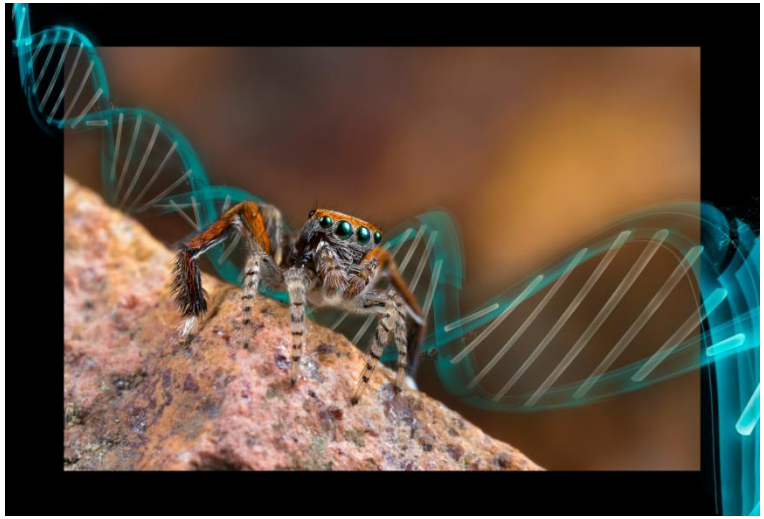
➔ All protocols are available on JEMU website : ***http://jemu.myspecies.info/***



Development of protocols for molecular systematic collections

- Development of protocols for field practices
  - Redaction of a proposal for an ABC Taxa Volume in collaboration with Kurt Jordaens (RMCA) and Zoltan Nagy (RBINS) - http://www.abctaxa.be/

The use of anesthetics, fixatives and preservatives on invertebrates dedicated to molecular analyses : a manual of best practice.



©Photo : Arnaud Henrard, RMCA

Section 1 : Molecular systematics - Exploring, testing, and establishing collection conservation protocols for future molecular systematic research

WP 1.1 : Testing for optimal retrieval of genetic material

WP 1.2 : Development of protocols for field practices for molecular systematic collections

WP 1.3 : Incorporation of tissues and DNA material into a standardized collection management system

WP 1.4 : Development of guidelines on compatibility and interchangeability of field and specimen data sheets

WP 1.5 : Testing of samples for NGS

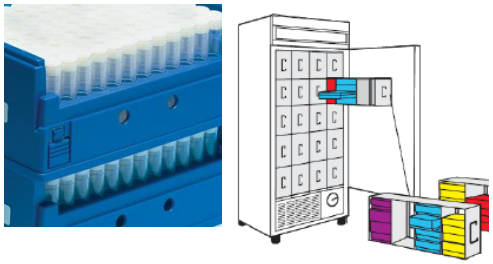
WP 1.6 : Exchange and incorporation of international standards

Incorporation of tissues and DNA material into a standardized collection management system

- Common system for DNA storage facilitates exchange between RBINS and RMCA : **2D Barcode tubes**

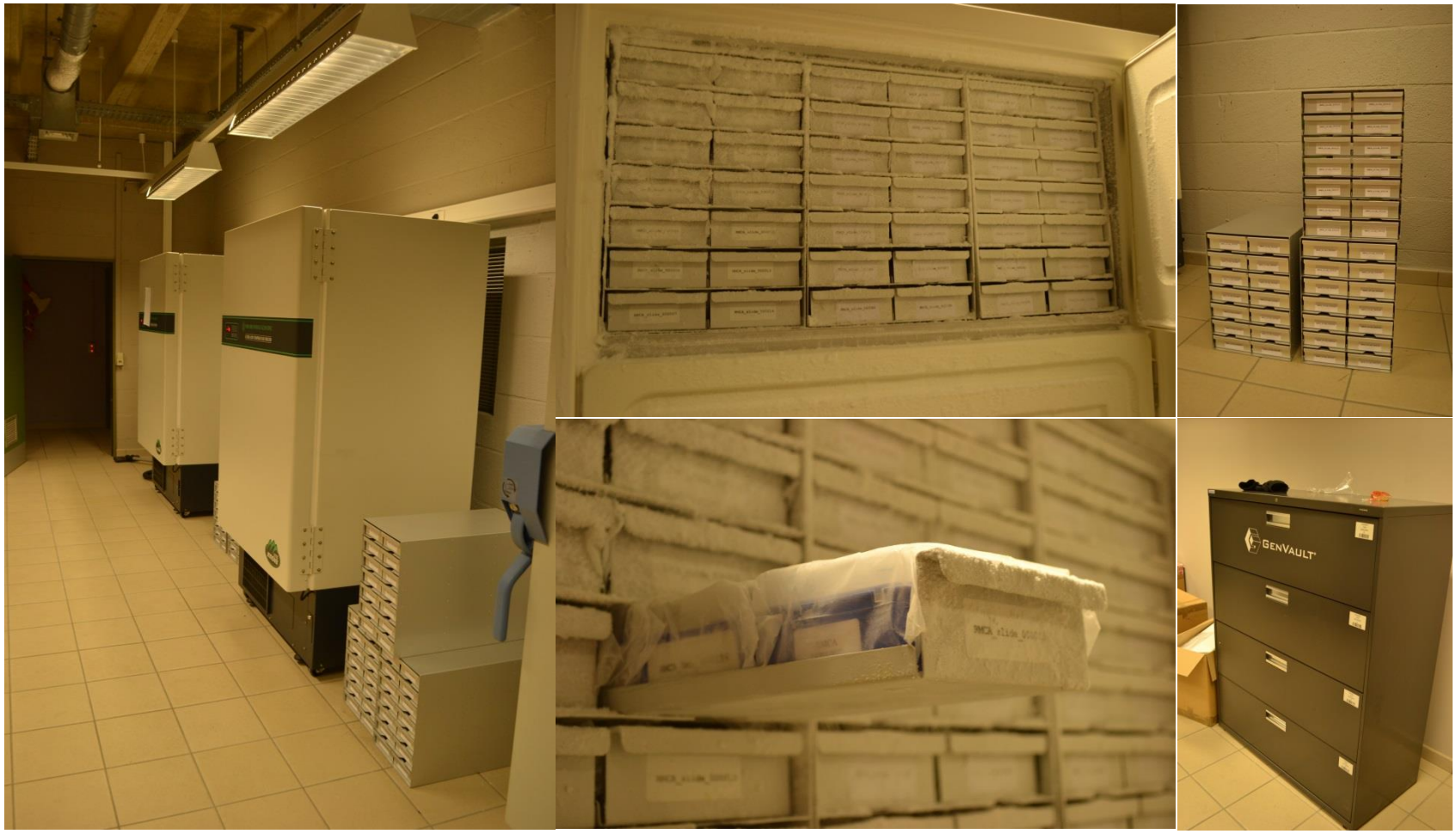
Storage under dry conditions at room temperature

Storage in ultra-low temperature freezers



Incorporation of tissues and DNA material into a standardized collection management system

- DNA storage facility in Tervuren



Section 1 : Molecular systematics - Exploring, testing, and establishing collection conservation protocols for future molecular systematic research

WP 1.1 : Testing for optimal retrieval of genetic material

WP 1.2 : Development of protocols for field practices for molecular systematic collections

WP 1.3 : Incorporation of tissues and DNA material into a standardized collection management system

WP 1.4 : **Development of guidelines on compatibility and interchangeability of field and specimen data sheets**

WP 1.5 : Testing of samples for NGS

WP 1.6 : Exchange and incorporation of international standards

Development of guidelines on compatibility and interchangeability of field and specimen data sheets

- Collaboration with Marielle Adam (RBINS) : development of field data sheets

Unique ID | Project or expedition name
Collector(s):

DATE
Start date (yyyy-mm-dd) | Time (HH:MM) | End date (yyyy-mm-dd) | Time (HH:MM)

LOCATION
Ocean | Sea | Continent
Country | State or province | Region or district
Municipality | Exact site
Field number | Station number | Sampling code
Altitude (m) | Depth (m) | Ecology
Latitude | Longitude

SITE MEASUREMENTS
Parameter | Value | Parameter | Value

COLLECTION STRATEGY
Sampling method:
Fixatives or anesthesics
Notes

NOTES

SPECIMEN MEASUREMENTS OR FACTS
General information
Tissue/part of organism | Sex | Stage | Social status
Total number | Number of males | Number of females | Number of unknown sex
Custom values
Parameter | Value | Parameter | Value

HOST
Class | Order | Family
Genus | Species | Author, year
Notes

CLASSIFICATION
Taxon name
Species | Subspecies | Author, year
Complete taxonomy
Phylum | Class | Order
Super family | Family | Sub family
Genus | Sub genus
Additional information
Variety or form | Common name
Identification method
Identified by | Identified on
Notes

Section 1 : Molecular systematics - Exploring, testing, and establishing collection conservation protocols for future molecular systematic research

WP 1.1 : Testing for optimal retrieval of genetic material

WP 1.2 : Development of protocols for field practices for molecular systematic collections

WP 1.3 : Incorporation of tissues and DNA material into a standardized collection management system

WP 1.4 : Development of guidelines on compatibility and interchangeability of field and specimen data sheets

WP 1.5 : Testing of samples for NGS

WP 1.6 : Exchange and incorporation of international standards

Testing samples for NGS

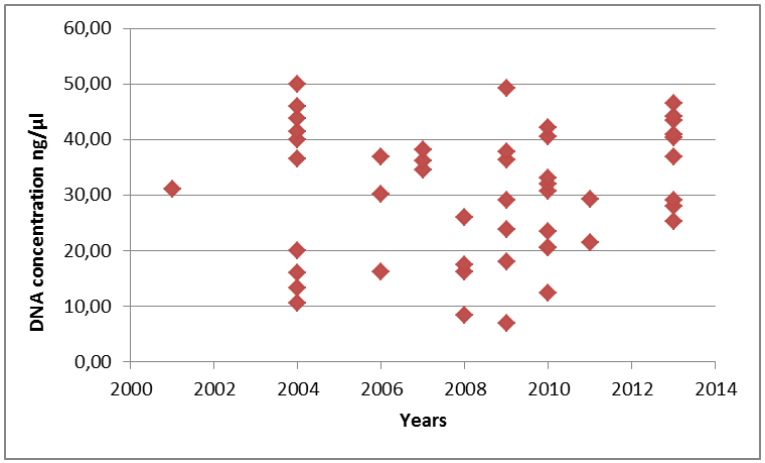
- Using **RAD-Sequencing** to resolve a small tephritid complex (FAR Complex in partnership with Massimiliano Virgilio)
- Using **GBS** (Cornell University, Genomic Diversity Facility) to study the population structure/phylogeny of C. capitata (in partnership with Raul Ruiz – APHIS lab and Massimiliano Virgilio – JEMU)
- Using **GBS** (Cornell University, Genomic Diversity Facility) to study the population structure/phylogeny of Buffalo (in partnership with Nathalie Smitz – JEMU)

Testing samples for NGS – RAD Sequencing

- Using **RAD-Sequencing** to resolve a small tephritid complex (FAR Complex) in partnership with Massimiliano Virgilio (JEMU)

Selection of 16 samples (2004-2013)

- 48 tephritid specimens aged between 2004 et 2013
- Non-destructive extraction, whole specimen in buffer overnight. Macherey-nagel tissue kit. 50µl elution buffer
- DNA concentration ranged from 6,92 to 49,90 ng/µl



Testing samples for NGS – RAD Sequencing

- Using **RAD-Sequencing** to resolve a small tephritid complex (FAR Complex) in partnership with Massimiliano Virgilio (JEMU)

Selection of 16 samples (2004-2013)

Library preparation

Protocol based on Fuentes-Utrilla (unpublished), which is based on Etter et al (2011 Methods Mol Biol 772:157-78) and on Baird et al (2008 PloS ONE 3:e3376).

Sending to a facility

edinburgh genomics
design execution analysis

Data processing (Massi & Gontran)

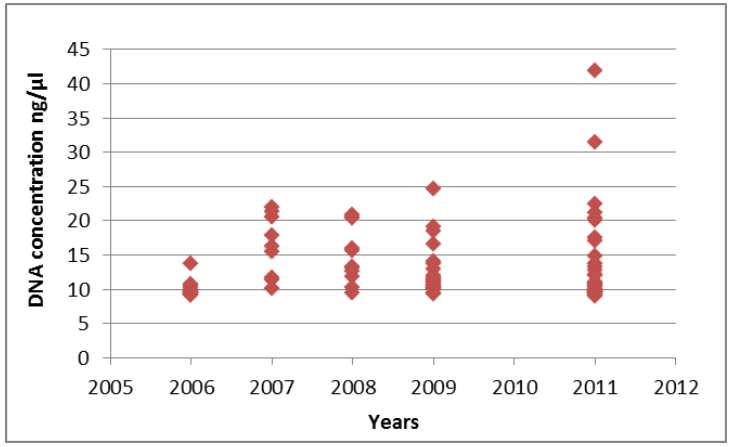
Poster presentation “Get off on the right foot. The importance of initial steps in NGS data analysis” – Symposium « Entomology in Belgium » 2014

Testing samples for NGS – GBS

- Using **GBS** (Cornell University, Genomic Diversity Facility) to study the population structure/phylogeny of C. capitata (in partnership with Raul Ruiz – APHIS lab and Massimiliano Virgilio – JEMU)

Selection of 95 samples (2006-2011)

- 95 tephritid specimens aged between 2006 and 2011
- Non-destructive extraction, whole specimen in buffer overnight. QIAGEN 100 µl elution buffer
- DNA concentration ranged from 9 to 42 ng/µl



Testing samples for NGS – GBS

- Using **GBS** (Cornell University, Genomic Diversity Facility) to study the population structure/phylogeny of C. capitata (in partnership with Raul Ruiz – APHIS lab and Massimiliano Virgilio – JEMU

Selection of 95 samples (2006-2011)

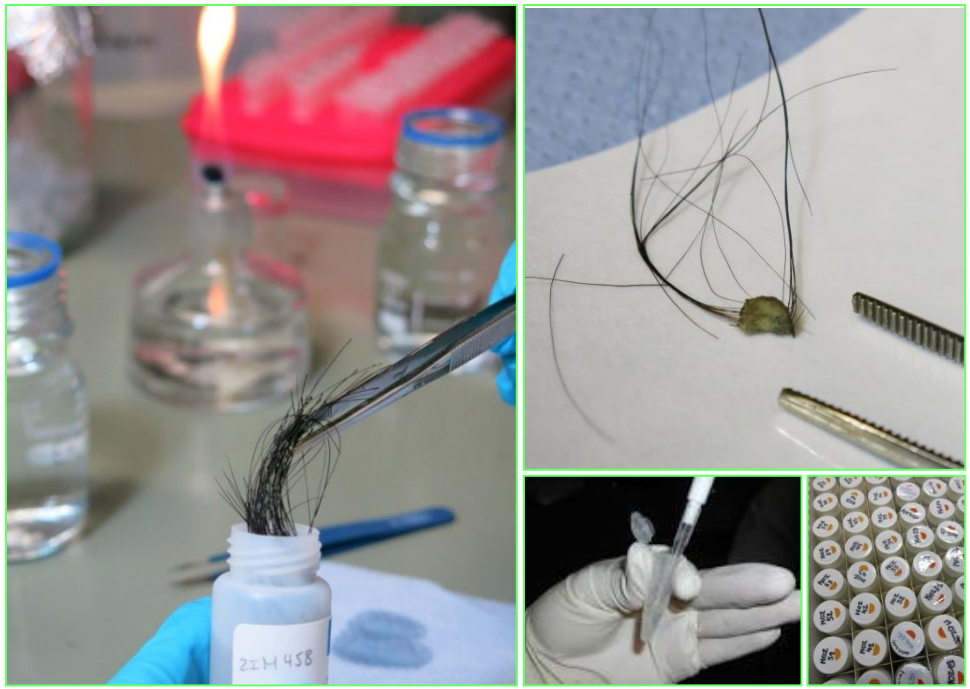
Restriction
Trial digestions using an inexpensive restriction enzyme (6-base-cutter, HindIII / EcoRI) + DNA quality check was performed

Sending to a facility

Testing samples for NGS – RAD Sequencing

- Using **GBS** (Cornell University, Genomic Diversity Facility) to study the population structure/phylogeny of Buffalo (in partnership with Nathalie Smitz – JEMU)

➔ DNA extraction on hairs, skins, tissues



Section 1 : Molecular systematics - Exploring, testing, and establishing collection conservation protocols for future molecular systematic research

WP 1.1 : Testing for optimal retrieval of genetic material

WP 1.2 : Development of protocols for field practices for molecular systematic collections

WP 1.3 : Incorporation of tissues and DNA material into a standardized collection management system

WP 1.4 : Development of guidelines on compatibility and interchangeability of field and specimen data sheets

WP 1.5 : Testing of samples for NGS

WP 1.6 : Exchange and incorporation of international standards

Exchange and incorporation of international standards

- Collaboration with Larissa Smirnova with the Global Genome Biodiversity Network Data Standard included in SYNTHESYS NA2, task 2.3 “Develop protocols for data collection from DNA extraction” and task 2.4 (GGBN)

SYNTHESYS
Synthesis of systematic resources

Valorisation

- Participation to national/international conferences

  - First GGBN International Conference on Biodiversity Biobanking, London - June 30-July 2 2014. **Poster presentation**
  - Collection management training organized by Didier Van den Spiegel. **Oral presentation** “From specimens to DNA” – 19/11/2014
  - Meeting Synthesys3, WP 2.4: Establish network of DNA and tissue banks – 24-26 November 2014, Vienna, Austria.
  - Symposium “Entomology in Belgium”, 28th November 2014, Brussels, Belgium. **Poster Presentation**.
  - Zoology conference, 12-13th December 2014, Liege, Belgium. **Poster presentation**.

Valorisation

- Participation to national/international conferences
- Production of internal guidelines and protocols (lab protocols, field protocol, field sheet, …)
- Dissemination of guidelines and protocol : JEMU website
- Publishing a volume for the ABC Taxa journal ?

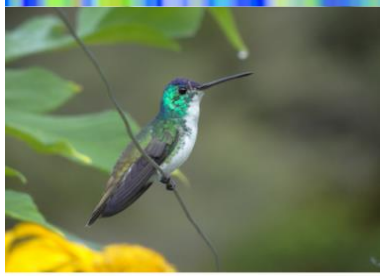
Perspectives

- The physical conditions where our minus 80 freezers are stocked are not optimal ➔ cooling infrastructure to be considered ??

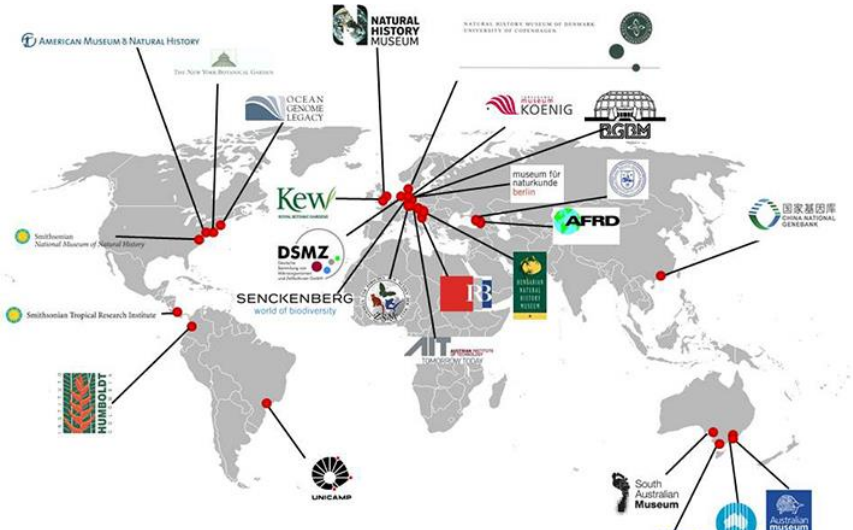
- Ensure the traceability of tissue/DNA samples in our collections with DaRWIN

- Membership of GGBN ? Once DaRWIN is implemented, maybe we can consider it as an option….

The Global Genome Biodiversity Network (GGBN)

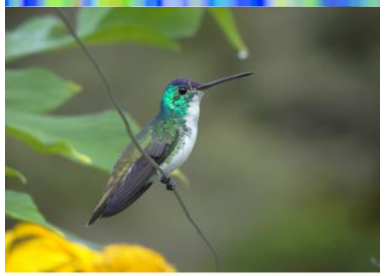


The Global Genome Biodiversity Network (GGBN) is a global network formed in 2011 that seeks to bring together the world's leading collections of genome quality specimens from across the Tree of Life in order to help accelerate research and collaboration between repositories of molecular biodiversity - http://ggbn.org/



26 Biorepository Members & Collaborators

The Global Genome Biodiversity Network (GGBN)



Membership is open to natural history museums and other organizations or individual involved in non-human genomic biodiversity.

Requirements :

- Applied specimen databases have to be **GBIF compliant** so that they can be accessed via GBIF and the GGBN Data Portal.
- **DNA metadata** (contact details, terms of use, database description) and the **ABCDDNA extension** are needed for mapping DNA databases.
- **Every DNA dataset has to be referenced to its respective specimen dataset through GBIF**. That reference consists of the record identifiers UnitID, CollectionCode, and InstitutionCode. These ABCD or DarwinCore elements together with the WrapperUrl unambiguously define a specimen dataset

For more information : http://www.ggbn.org/ggbnMembership.html

Section 1 : Molecular systematics - Exploring, testing, and establishing collection conservation protocols for future molecular systematic research

---

Section 2 : Bioinformatics - Developing, testing and improving an institutionally overarching joint collection data management, retrieval and accessibility system

Main objective 2 : Bioinformatics

"Developing, testing and improving an institutionally overarching joint collection data management, retrieval and accessibility system."

- Transfer of all major datasets from both institutions into DaRWIN, by means of excel-templates following the ABCD international standard
- Development of a standardised template following the ABCD format, for new data entry
- Further development and improvement of DaRWIN, as internal database system, to allow the import of ABCD formatted data

At the last meeting…

Workflow

WP2.1 - 'Further development and testing of the DaRWIN database module for tissue and DNA collections'

WP2.2 - 'Exploring means for data transfer from other platforms'

WP2.3 - 'New data entry and retrieval'

-> SPECIMEN-sheet
-> SAMPLE-sheet
-> DNA-sheet

DNA template

Specific templates

General template + fieldsheet

STANDARDIZATION

DaRWIN adaptations (WP2.4)

Development of an import tool for XML ABCD formatted files

New version of DaRWIN to make the system more efficient and more user-friendly

Additional templates and new tools…

In April 2014, production versions of the templates were released. After numerous tests made within the IT, among which import of a part of the Poppe collection, some adaptations appeared to be necessary:

1) Import itself is too long when big files are being imported

➔ Improvement of the efficiency in imports through revision of the different « system » steps for an import

2) Dealing with GTU within the import doesn't correspond the "field reality"

➔ Revision of this part through restructuring the GTU creation or matching for an imported specimen

3. Risk of errors through the automatic creation of taxon tool within the specimens import

Decision to separate the creation of taxonomy from the specimens import itself in order to

⇒ avoid the erroneous creation of taxa in the taxonomic dictionary

⇒ promote an in-depth checking of the taxonomy before specimens import

What is different?

- The automatic creation of taxa within the specimens import is not available anymore.
- The checking of the taxonomy during specimens import is based on one unique field "taxonFullName" field through an exact match on this field (not on the rest of the taxonomical tree)...

How?

✓ Development of a template and an import module in DaRWIN for taxonomy import

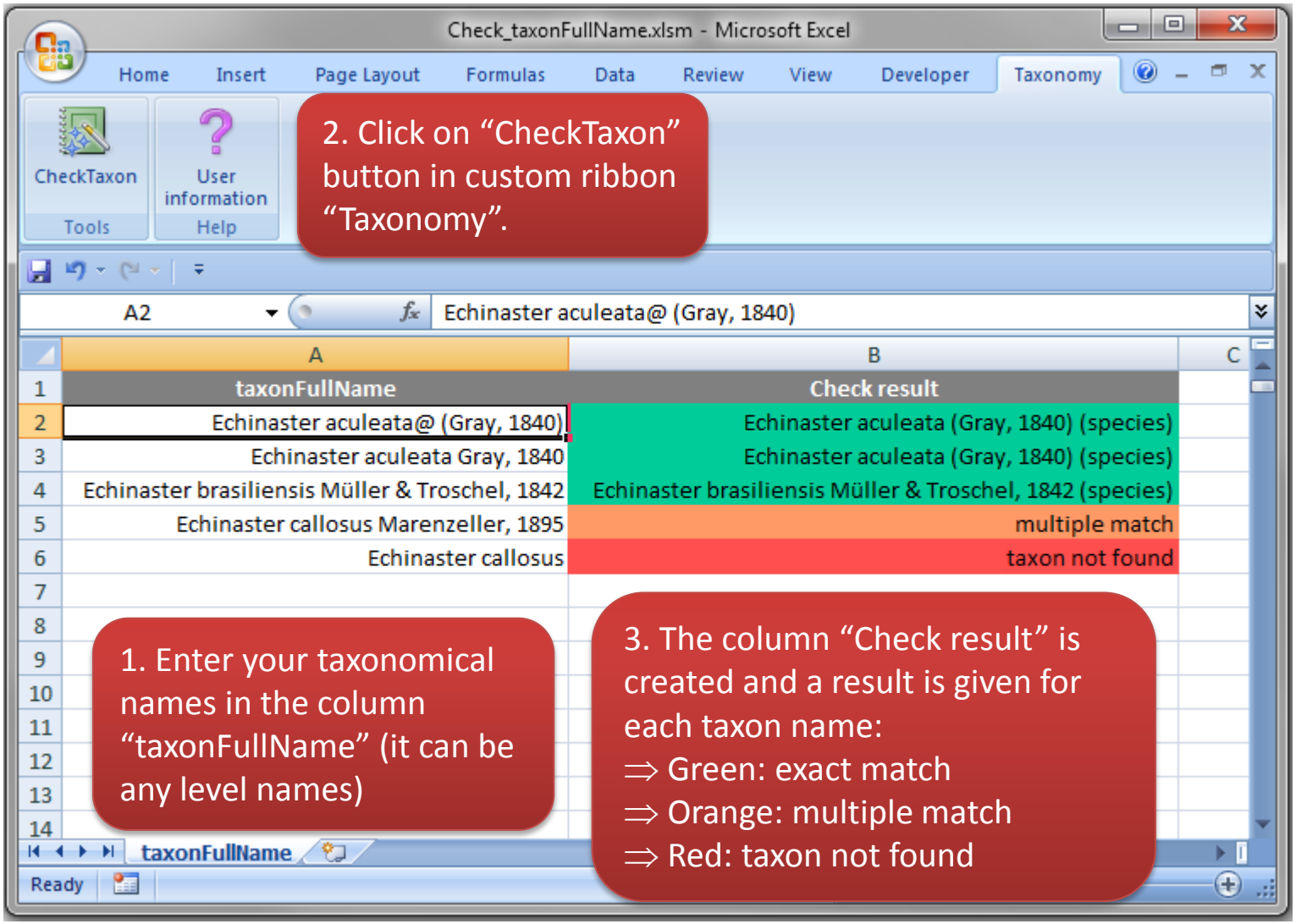
✓ Development of a taxonomy verification tool (Excel macro that allows the user to check whether a taxonomic name is present in DaRWIN or not)

✓ Development of additional templates for specimens import including only taxonomic name (and no taxonomic tree), to be used with the taxonomy verification tool and the taxonomy import module (adapted user manuals available)

⇒ Large reduction in the number of errors to correct at the moment of the specimens import

The old templates can still be used even if the preparation of taxonomy prior to specimens import is strongly recommended.

Check_taxonFullName.xlsm



2. Click on “CheckTaxon” button in custom ribbon “Taxonomy”.

1. Enter your taxonomical names in the column “taxonFullName” (it can be any level names)

3. The column “Check result” is created and a result is given for each taxon name:
⇒ Green: exact match
⇒ Orange: multiple match
⇒ Red: taxon not found

DaRWIN Help Menu

=> ABCD Templates (classical and taxonFullName versions), user manuals and additional tools for taxonomy available for download

My Preferences | Searches | Add | Administration | Help

Darwin Guide :

Manuel d'utilisation

ABCD Templates* :

*Only available in english

How to import a XML file in Darwin?

Section	Templates	User manuals
General	ABCDImport2DaRWIN_General.xlsm ABCDImport2DaRWIN_General_taxonFullName.xlsm	ABCDImport2DaRWIN_General_usermanual.pdf
DNA / JEMU	ABCDImport2DaRWIN_DNA.xlsm ABCDImport2DaRWIN_DNA_taxonFullName.xlsm	ABCDImport2DaRWIN_DNA_usermanual.pdf
Invertebrates	ABCDImport2DaRWIN_InvEntomo.xlsm ABCDImport2DaRWIN_InvEntomo_taxonFullName.xlsm	ABCDImport2DaRWIN_InvEntomo_usermanual.pdf
Entomology	ABCDImport2DaRWIN_Entomo.xlsm ABCDImport2DaRWIN_Entomo_taxonFullName.xlsm	ABCDImport2DaRWIN_Entomo_usermanual.pdf
Vertebrates	ABCDImport2DaRWIN_Vert.xlsm ABCDImport2DaRWIN_Vert_taxonFullName.xlsm	ABCDImport2DaRWIN_Vert_usermanual.pdf
Paleontology	ABCDImport2DaRWIN_Paleontology.xlsm ABCDImport2DaRWIN_Paleontology_taxonFullName.xlsm	ABCDImport2DaRWIN_Paleontology_usermanual.pdf
Geology	ABCDImport2DaRWIN_Litho.xlsm ABCDImport2DaRWIN_Mineralo.xlsm Template_Litho_localities.xlsx	ABCDImport2DaRWIN_Geol_usermanual.pdf

Additional tools for import :

Taxonomy import	Taxon full name checker
Template: TaxonomyImport.xlsm User manual: TaxonomyImport2DaRWIN_usermanual.pdf	Check_taxonFullName.xlsm

WP2.4 - 'Implementation of an overarching, web based portal as an unique access point to the data of the collections in both institutions (RBINS & RMCA)'

- Implementation of DaRWIN as collection management system
  - Importing 2 collections (Coelenthera and Arachnomorpha)
  - Two more collections being migrated (Vermes and Ichtyology) in 2015
  - Additional developments
    - External tool to check name statuses by using GBIF (before batch importation). Output is annotation of the taxonomic fields of the ABCD template
    - Test developments
      - Labels
      - XML outputs of names
      - Display of historical changes
      - Searching discrete series of label numbers
      - …

- RBINS DaRWIN-team is available for technical issues
- (Common validation procedures, test, repositories being currently discussed)

Acknowledgements

AGORA-MOLCOL

Adam Marielle
Backeljau Thierry
De Meyer Marc

JEMU
Joint Experimental Molecular Unit

Nathalie Smitz
Sonet Gontran
Tamas Nagy Zoltan

Africa TERVUREN

Cooleman Stijn
Decru Eva
De Prins Jurate
Gilissen Emmanuel
Henrard Arnaud
Huyse Tine
Jordaens Kurt
Snoeks Jos
Van den Spiegel Didier
Van Steenberge Maarten
Virgilio Massimiliano
Vreven Emmanuel

museum

De Gelas Koen
Hendrickx Frederik
Kok Philippe
Leponce Maurice
Schön Isabelle
Verheyen Erik

KATHOLIEKE UNIVERSITEIT LEUVEN

Belgian Science Policy Office
belspo