



Indexing for Life

D8.1 CoL integrated interface of EU BOLD/ECBOL.

Work package 8

Vincent Robert and Alastair Culham

31 October 2013

Capacities Programme of Framework 7: EC e-Infrastructure Programme – Virtual Research Communities - INFRA-2010-2

Grant Agreement No:	261555
Project Co-ordinator:	Dr Alastair Culham
Project Homepage:	http://www.i4Life.eu
Duration of Project:	36 months
Start Date:	1 November 2010
End Date:	31 st October 2013



D8.1 CoL integrated interface of BOLD/ECBOL

Introduction

This deliverable reports the integration of the Catalogue of Life (CoL) taxonomic index into the ECBOL/EU BOLD system and the piping of names to the CoL as part of the integrated workflow described in D2.7.

There have been substantial challenges with the integration of CoL into the ECBOL/EU BOLD platform due to the changes in the over-arching BOLD platform and these were discussed in the year 2 review meeting. The agreed alternative for the integration of the CoL with ECBOL/EU BOLD was to use Mycobank and its contained barcoding portal to set up and demonstrate the processes needed such that ECBOL/EU BOLD could integrate the CoL once the new BOLD data portal was available from iBol in North America.

Integration with the Mycobank portal

The processes to integrate current versions of CoL with Mycobank have been used as a proxy for the development of links with the ECBOL/EU BOLD portal. The Catalogue of Life is now integrated into the Mycobank portal as a Link Out on the results page of any search.


This operates at all taxon ranks that cross map between the CoL and Mycobank. Mycobank is incorporated into the i4Life workflow in the same way as any other GBP and so both provides names and consumes them via the piping tool. Substantial additional names for fungi were provided by Mycobank and these were incorporated into the CoL fungal GSD as a Workpackage 5 activity. This was the single largest individual GSD name placement exercise in WP5 and has resulted in a very close match between the GBP (Mycobank) and the GSD (Species Fungorum) that illustrates the strength of the new i4Life workflow system. The new names were made available via the CoL in the 15th August Dynamic edition of the CoL.

The screenshot shows the MycoBank website interface. At the top, there are flags for various countries and the MycoBank logo with the URL www.mycobank.org. The International Mycological Association (IMA) logo is also present. The main header features a photograph of a red mushroom and the text 'Fungal Databases Nomenclature and Species Banks'. Below this is a navigation menu with links: Home, Search, Login, Register new..., Identifications, Tools, News, Forum, About, Help, and Unknown user. A search bar is located below the menu, with the text 'Search on : Mycobank'. The search results for 'Russula emetica' are displayed, showing various links categorized under 'Other fungal links', 'Bibliography links', 'General links', 'Molecular links', and 'Specimens and strains links'. The 'Catalogue of Life (CoL)' link is highlighted with a red box. At the bottom, there is a section for 'Files' with 'Associated files' and a small image of a mushroom.

Figure 1 - The Mycobank portal

This demonstrates the completed cycles of name piping from the GBP to the GSD, the GSD to the CoL and the feedback from GSD to GBP. To help monitor developments in content of CoL relative to the GBPs Mycobank was included in the i4Life search preliminary demonstrator (<http://www.i4life.eu/i4search/>; see Figure 2).


Species Search Service



CoL ENA GBIF IUCN EoL MycoBank Help

Try again

MycoBank matched 39 results on *Russula emetica*



Result 1:	<i>Russula emetica</i>
Summary:	<i>Russula emetica</i> (Schaeff.) Pers., <i>Observationes mycologicae</i> 1: 100 (1796) [MB#191650]
Synonymy:	≡ <i>Agaricus russula</i> Scop., <i>Flora carniolica</i> , 2: 435, no 1502, 1772 [MB#375256] ≡ <i>Agaricus emeticus</i> Schaeff., <i>Fungorum qui in Bavaria et Palatinatu circa Ratisbonam nascuntur Icones</i> , 4: 9, t. 15, 16, 1774 [MB#226395] ≡ <i>Amanita rubra</i> Lam., <i>Encyclopédie Méthodique, Botanique</i> , 1-1: 105, 1783 [MB#465251] ≡ <i>Agaricus ruber</i> (Lam.) DC., <i>Flore française</i> , 2: 140, 1805 [MB#246168] ≡ <i>Agaricus linnaei</i> var. <i>emeticus</i> (Schaeff.) Fr., <i>Observationes mycologicae</i> , 1: 67, 1815 [MB#497906] ≡ <i>Russula rubra</i> (Lam.) Fr., <i>Epicrisis Systematis Mycologici</i> : 354, 1838 [MB#204768] ≡ <i>Melanoleuca russula</i> (Scop.) Murrill, <i>North American Flora</i> , 10 (1): 22, 1914 [MB#100414]
Rank:	sp.
Authors:	(Schaeffer) Persoon
Classification:	Fungi, Basidiomycota, Agaricomycotina, Agaricomycetes, Russulales, Russulaceae, <i>Russula</i>
URL:	http://www.mycobank.org/BioloMICS.aspx?Link=T&Table=MycoBank&Rec=122860&Fields=All
MycoBank ID:	191650
<hr/>	
Result 2:	<i>Russula emetica</i> f. <i>silvestris</i>
Summary:	<i>Russula emetica</i> f. <i>silvestris</i> (Singer), <i>Beihefte zum Botanischen Zentralblatt</i> 49 (2): 305 (1932) [MB#263961]
Synonymy:	≡ <i>Russula emetica</i> var. <i>silvestris</i> Singer, <i>Zeitschrift für Pilzkunde</i> , 3 (6), 1924 [MB#444423] ≡ <i>Russula silvestris</i> (Singer) Reumaux, <i>Russules rares ou méconnues</i> : 289, 1996 [MB#445434]
Rank:	f.
Authors:	(Singer)
Classification:	Fungi, Basidiomycota, Agaricomycotina, Agaricomycetes, Russulales, Russulaceae, <i>Russula</i> , <i>Russula emetica</i>
URL:	http://www.mycobank.org/BioloMICS.aspx?Link=T&Table=MycoBank&Rec=123584&Fields=All
MycoBank ID:	263961

Figure 2 - the i4Life Species Search Service data portal consuming web services developed by GBP (here MycoBank as example)

This work is additional to the original project plan and is consuming new and dynamic web services that were created within the framework of the i4Life project. An output of this dynamic web service system available from the MycoBank website (example:

<http://www.mycobank.org/Services/Generic/SearchService.svc/rest/xml?layout=1468261600000161&filter=Name%20CONTAINS%2022Russula%20emetica%22>) is provided in Figure

3.

Additionally, another dynamic tool was created allowing any curator of BioloMICS (software underlying MycoBank and ECBOL/EU BOLD) to export data to Darwin Core format in order to export data to Col. The latter system can be used as an alternative exportation/importation channel to the web services mentioned above. It has been used by Col to import data from MycoBank recently.

This XML file does not appear to have any style information associated with it. The document tree is shown below.

```

- <Results>
- <Taxon>
  <Name>Russula emetica</Name>
- <E3787>
  Russula emetica (Schaeff.) Pers., Observaciones mycologicae 1: 100 (1796) [MB#191650]
  <E3787>
- <E4060>
  =Agaricus russula Scop., Flora carniolica, 2: 435, no 1502, 1772 [MB#375256]<br />=Agaricus emeticus Schaeff., Fungorum qui in Bavaria et Palatinatu circa
  Ratisbonam nascuntur Icones, 4: 9, t. 15,16, 1774 [MB#226395]<br />=Amanita rubra Lam., Encyclopédie Méthodique, Botanique, 1-1: 105, 1783
  [MB#465251]<br />=Agaricus ruber (Lam.) DC., Flore française, 2: 140, 1805 [MB#246168]<br />=Agaricus linnaei var. emeticus (Schaeff.) Fr., Observaciones
  mycologicae, 1: 67, 1815 [MB#497906]<br />=Russula rubra (Lam.) Fr., Epicrisis Systematis Mycologici: 354, 1838 [MB#204768]<br />=Melanoleuca russula
  (Scop.) Murrill, North American Flora, 10 (1): 22, 1914 [MB#100414]
  <E4060>
  <MycoBankNr_>191650</MycoBankNr_>
  <Epithet_>emetica</Epithet_>
- <Rank_Pt_>
  <TargetRecord><Id>20</Id><Name>sp.</Name></TargetRecord>
  </Rank_Pt_>
  <OrthVariantOf_Pt_>
  <Authors_>(Schaeffer) Persoon</Authors_>
  <AuthorsAbbrev_>(Schaeff.) Pers.</AuthorsAbbrev_>
- <Literature_Pt_>
  <TargetRecord><Id>1112</Id><Name>Persoon C.H. 1796 Observaciones mycologicae 1</Name></TargetRecord>
  </Literature_Pt_>
  <LiteraturePageNr_>100</LiteraturePageNr_>
  <LiteratureJournalBook_>
  <NameYear_>1796</NameYear_>
  <NameType_>Combination</NameType_>
  <Gender_>Feminine</Gender_>
  <DatePublic_>
  <NameStatus_>Legitimate</NameStatus_>
  <NameStatusExplanation_>
  <Remarks_>
  <SanctioningRef_>
  <SanctionedBy_>Fr.</SanctionedBy_>
- <SanctioningName_Pt_>
  <TargetRecord><Id>241547</Id><Name>Agaricus emeticus</Name></TargetRecord>
  </SanctioningName_Pt_>
  <ValidatedBy_Pt_>
  <Specimen_Pt_>
  <Epitype_Pt_>
  <Iconotype_Pt_>
  <Lectotype_Pt_>
  <Neotype_Pt_>
- <CurrentName_Pt_>
  <TargetRecord><Id>122860</Id><Name>Russula emetica</Name></TargetRecord>
  </CurrentName_Pt_>
- <Classification_>

```

Figure 3 - the MycoBank dynamic web service output used by Species Search Service data portal

<http://www.mycobank.org/Services/Generic/SearchService.svc/rest/xml?layout=1468261600000161&filter=Name%20CONTAINS%20%22Russula%20emetica%22>

Integration with ECBOL and the EU BOLD Systems Portal

Despite justifiable concern at the year 2 review meeting that it might be impossible to link Catalogue of Life to the ECBOL/EU BOLD platform due to reasons outside the i4Life project, the restructuring of BOLD software, systems were in place to allow this to happen very rapidly should the portal be launched in time. We are very pleased to report that this did happen and the link is in place (Figures 4 & 5).

EU BOLD
EUBOLD - European Barcoding of Life Database

Home page Specimens Taxonomy Identification Register News Contact us Help Unknown user

Specimens
Make a search on any specimen in the EUBOLD database.

Identifications
Perform a pairwise identification on the sequence data in the EUBOLD database.

Taxonomy
Search for any taxonomy in the EUBOLD database.

Collections
Search for collections in the EUBOLD database.

**EU BOLD Mirror
DNA Barcoding Made Easy**

- Search on specimen
- Search on collection
- Search on taxonomy
- Pairwise sequence alignment
- Polyphasic identification

Join us
Create your own database and associated website and be visible using BioMICS. Contact us for more details.

Statistics

Barcode records: 926.282
Number of samples: 898.509

Number of species: 97.965

Number of collections: 463.496
Number of countries: 236

Project information

The Barcode of Life Database is designed to support the generation and application of DNA Barcode data. Simultaneous pairwise sequence alignments against several Plant, Animal and Fungal databases are now possible thanks to the new search engine. Polyphasic identifications of fungi and yeasts against curated references databases are also proposed.

Figure 4 - the EU BOLD data portal

The links out operate within the 'Taxonomy' search option at all taxonomic levels (Figure 4). This gives access to species lists from the Col to users of the ECBOL/EU BOLD database.

The screenshot shows the EU BOLD portal interface. At the top, there are logos for EU BOLD, CBS, Naturalls Biodiversity Center, Government of the Netherlands, the European Union flag, and the Seventh Framework Programme. A navigation menu includes Home page, Specimens, Taxonomy, Identification, Register, News, Contact us, Help, and Unknown user. The search bar is set to 'Search on : Specimens' with options for 'Add condition', 'Match on : All conditions', 'Reset base condition(s)', and 'Switch to: Advanced Search'. The search conditions are expanded to show five criteria, all set to 'Contains': Taxon name (C_1) with 'Russula', Museum Id (C_2), Voucher type (C_3), Voucher description (C_4), and Any text field (C_5). Below the search conditions is an 'Export data' button. The search results for 'Crocidura russula' are displayed in a table-like format:

Taxon rank :	Species
Taxonomy :	Chordata , Mammalia , Soricomorpha , Soricidae , Crocidurinae , Crocidura
Associated records:	None
Catalog of Life link out :	Crocidura russula
Associated specimens :	1. BM389-04 2. RM389-04

Figure 5 - EU BOLD portal showing the CoL link

Now that the structure is in place, further developments can take place outside the i4Life project to make fuller use of the CoL resource in Mycobank and ECBOL.