

Cómo colaborar con GenBank para enlazar la información de las secuencias con las bases de datos en línea de las colecciones

Francisco Pando



Sumario

- De que estamos hablando
- Ideas y requisitos preliminares
- Enlazar desde la colección a GenBank
- Enlazar desde GenBank a la colección
 - Que son los “Linkouts”
 - Procedimiento
 - La parte técnica
 - La reparación de los ficheros

De que estamos hablando

Herbario de Criptogamia

RJB > Colecciones > Herbario > Criptogamia > Bases de datos > Búsqueda

English? Try Babelfish

Los nombres en negrita indican que se puede acceder a su información nomenclatural --haciendo clic sobre el mismo-- contenida en la CABI Bioscience FunIndex Database

Este link ha sido posible gracias a la colaboración entre los dos institutos

Nomenclatural information for names in boldface is available by clicking on them. This information is provided by CABI Bioscience FunIndex Database

This link has been made possible through a collaboration between both institutes

Información
Bases de datos
Noticias
Adquisiciones
Contactar

Por F. Pando



Resultados

Para volver al formulario de búsqueda
Back to the query form

Rhizopogon buenoi Calonge & M.P.Martín

Typus (H)

ESP. Av: Arenas de San Pedro, Hontanares
bajo Pinus sylvestris
S.Barbero
Det. F.D.Calonge & M.P.Martín
21-nov-1998

MA-Fungi 47676 GenBank Acc. No. AJ297263

Nombre aceptado en el herb. / herb. current name: *Rhizopogon buenoi* Calonge & M.P.Martín

Rhizopogon buenoi Calonge & M.P.Martín

ESP. CR: Viso del Marqués, Valle de los Perales
bajo Pinus pinea
F.D. Calonge
Det. F.D.Calonge & M. P. Martín
03-dic-2005

MA-Fungi 63381 Nombre aceptado en el herb. / herb. current name: *Rhizopogon buenoi* Calonge & M.P.Martín

NCBI Resources How To

Nucleotide Nucleotide Limits Advanced

Display Settings: GenBank Send:

Rhizopogon buenoi 5.8S rRNA gene, internal transcribed spacer 1 (ITS1) and internal transcribed spacer 2 (ITS2), isolate MA-Fungi 47676

GenBank: AJ297263.1
[FASTA](#) [Graphics](#)

Go to:

LOCUS	AJ297263	664 bp	DNA	linear	PLN 22-MAR-2004
DEFINITION	Rhizopogon buenoi 5.8S rRNA gene, internal transcribed spacer 1 (ITS1) and internal transcribed spacer 2 (ITS2), isolate MA-Fungi 47676.				
ACCESSION	AJ297263				
VERSION	AJ297263.1	GI:27368647			
KEYWORDS	5.8S ribosomal RNA; 5.8S rRNA gene; internal transcribed spacer 1; internal transcribed spacer 2; ITS1; ITS2.				
SOURCE	Rhizopogon buenoi				
ORGANISM	Rhizopogon buenoi Eukaryota; Fungi; Dikarya; Basidiomycota; Agaricomycotina; Agaricomycetes; Agaricomycetidae; Boletales; Suillineae; Rhizopogonaceae; Rhizopogon.				
REFERENCE	1				

Ideas y requisitos preliminares

- La colección tiene que estar en línea
- Cada registro de la base de datos en línea -- correspondiente a cada ejemplar de la colección-- debe tener su URL: explícita y estable
- Las URLs deben poder construirse a partir del número de los ejemplares

Enlazar desde la colección a GenBank

Se necesita:

- El “Accession number” de Genbank para cada secuencia a enlazar
- La parte constante de la URL de la base de datos del GenBank de las secuencias en cuestión

Conviene:

- Un texto de contexto (para etiquetas, para web)
- Registrar quien ha proporcionado la información

Enlazar desde la colección a GenBank: Ejemplo

161.111.170.202/herb/asp/dos.asp?conque=MA-Fungi&fnhe=35883&fadid=1

Información
Bases de datos
Noticias

Herbario de Criptogamia

RJB > Colecciones > Herbario > Criptogamia > Bases de datos > Búsqueda

texto de contexto

por F. Pando

Resultados

Volver al formulario de búsqueda

URL

Battarrea stevenii (Libosch.) Fr.

Det. S. Capello

ESP. Co: carretera Comarcal 437 km 2,6 ; 30SUG4092
S. Capello
05-oct-1994
MA-Fungi 35883 ex Herb. COF(F): 1495
GenBank Acc. No. AH008982, AF090863, AF090862
- GenBank Acc. No. AF090862 - GenBank Acc. No. AH008982 - GenBank Acc. No. AF090863 -

Accession number

CSIC Real Jardín Botánico Webmaster Página preparada por F. Pando Actualizada el 27 de abril del 2000

www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=search&db=nucleotide&term=AF090862[accn]&doptcmdl=GenBank

Enlazar desde la colección a GenBank: Ejemplo

Coleccion: **MA-Fungi** **Herbario MA** MODIFICACIÓN DE ESPECIMENES

NÚMERO DE HERBARIO: 35883 - 1

Grupo:
 GÉNERO: Battarrea HIB: ESPECIE: stevenii Nombres:
 AUT_ESPEC: (Libosch.) Fr. INFRANK: INFRA:
 AUT_INFRA: IDQUAL: DETERM: S Capello
 MESAN: ES_TIPO: CAMISA: S TAXONOMÍA:
 OBSERV: PRIM: ULTI:

Registro: 1 de 1 Sin filtro Buscar

Localidad Ligar georefs. Formato, Exsiccata, pliegos... Iconografía, links,...

Añadir / editar imagen:	35883-1-1	GenBank Acc. No. AF090862	Imagen
	35883-1-2	GenBank Acc. No. AH008982	Imagen
	35883-1-3	GenBank Acc. No. AF090862	Imagen

Registro: 1 de 3

texto de contexto (points to 'GenBank Acc. No. AF090862')

URL (points to 'http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=search&db=nucleotide&term=AF090862')

Accesion number (points to 'AF090862')

Imágenes y links (modal window)

Nherbario: 35883 - 1 Nfoto: 1

Título: GenBank Acc. No. AF090862

URL: [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=search&db=nucleotide&term=AF090862\[accn\]&doptcmdl=GenBank](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=search&db=nucleotide&term=AF090862[accn]&doptcmdl=GenBank)

Archivo:

Nota:

Observ.:

Battarrea stevenii (Libosch.) Fr.

Zoom (modal window)

[http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=search&db=nucleotide&term=AF090862\[accn\]&doptcmdl=GenBank](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=search&db=nucleotide&term=AF090862[accn]&doptcmdl=GenBank)

Accepta Cancel

Enlazar desde GenBank a la colección

Linkouts

- Enlaces externos
- Mantenedos por los proveedores externos
- Publicaciones, colecciones, CBOL

Linkouts: Procedimiento

The image shows a screenshot of the NCBI LinkOut website. The main page is titled "Information for Other Resource Providers" and is located at www.ncbi.nlm.nih.gov/books/NBK3802/. The page includes a navigation menu with options like "Welcome", "Libraries", "Full-Text Providers", and "Other Resource Providers". The main content area contains a list of frequently asked questions (FAQs) for other resource providers, such as "LinkOut Help for Other Resource Providers", "LinkOut File Validation Utility", "Provider Responsibilities", "Prerequisites for Participation", "Apply for Inclusion in LinkOut", "File Preparation: Identity File", "File Preparation: Resource File (XML)", "File Preparation: Resource File (Simple Text)", "File Transfer", and "File Maintenance". There are also links to "Technical Information", "LinkOut Provider and Journal Lists", and "Linkout-news Announcement List".

<http://www.ncbi.nlm.nih.gov/projects/linkout/>

<http://www.ncbi.nlm.nih.gov/books/NBK3802/>

Linkouts: la parte técnica

1. Contactar LinkOut Team/NCBI/NLM
linkout@ncbi.nlm.nih.gov
2. Evaluados y aprobados. recibimos credenciales para subir ficheros por FTP al repositorio de NCBI para “linkouts”, y nuestro “**providerId**”
3. Preparar los ficheros
4. Dar aviso linkout@ncbi.nlm.nih.gov de que se han actualizado ficheros
5. Vemos que todo funciona

Linkouts: La preparación de los ficheros

- Se trata de preparar tres ficheros:
 - Datos del proveedor (nosotros)
 - ficheros de recursos (“resources”)
 - Linkouts para la BBDD de Nucleótidos
 - Linkouts para la BBDD de Taxonomía

Fichero de datos del proveedor (nosotros)

http://www.ncbi.nlm.nih.gov/books/NBK3802/#nonbib.File_Preparation_Identity_File

- This file must be named **providerinfo.xml**;
- the file name is case sensitive
- This file should be composed in a text editor



<http://notepad-plus-plus.org/>

```
File Edit Search View Encoding Language Settings Macro Run Plugins Window ? X
Synchronize Horizontal Scroll
providerinfo.xml
1 <?xml version="1.0"?>
2 <!DOCTYPE Provider PUBLIC "-//NLM//DTD LinkOut 1.0//EN"
3 "http://www.ncbi.nlm.nih.gov/projects/linkout/doc/LinkOut.dtd">
4 <Provider>
5 <!-- ProviderId is assigned by NCBI -->
6 <ProviderId>4283</ProviderId>
7 <Name>Herbarium, Real >Jardin Botanico-CSIC. Madrid</Name>
8 <NameAbbr>MA Herbarium</NameAbbr>
9 <SubjectType> culture/stock collections</SubjectType>
10 <Attribute>Free</Attribute>
11 <!-- Url is used in My NCBI and in the lists of LinkOut Providers -->
12 <Url>http://www.rjb.csic.es/jardinbotanico/jardin/index.php?Cab=109
13 </Url>
14 <!-- Brief is used in My NCBI -->
15 <Brief> On-line databases of the MA Herbarium. Real Jardin
    Botanico-CSIC; Madrid, Spain </Brief>
</Provider>
```

length: 712 | Ln: 14 Col: 58 Sel: 0 UNIX ANSI INS

Linkouts para la BBDD de Nucleótidos

- Puede ser xml, o texto
 - si es xml, debe ser conforme al [LinkOut DTD](#)
 - si es texto (la modalidad que vamos a ver aquí):
 - extensión **.ft**
 - nombre del fichero sin espacios ni caracteres especiales
 - compuesto con un editor de texto tipo bloc de notas o Notepad++
 - Con una estructura simple pero rígida

Linkouts para la BBDD de Nucleótidos

Encabezamiento / "header"

"linkOuts"

Identificador del proveedor, asignado por NCBI

BBDD desde la que se hace "LinkOut"

Tipo de "LinkOut"

Parte constante de las URL de "LinkOut"

Núm. de "LinkOut"; único en el fichero para cada Linkout

"Accession number" de Genbank desde el que se hace el Linkout

Parte variable de la URL que hace referencia al ejemplar del Linkout

```
1 |-----|
2 | prid: 4283
3 | dbase: nucleotide
4 | stype: culture/stock collections
5 | !base: http://161.111.170.202/herb/asp/dos.asp?conque=MA-Fungis
6 |-----|
7 | linkid: 1
8 | query: AJ419196 [accn]
9 | base: sbase;
10 | rule: fnhe=47708
11 |-----|
12 | linkid: 2
13 | query: FN552536 [accn]
14 | base: sbase;
15 | rule: fnhe=79156
16 |-----|
17 | linkid: 3
18 | query: FN552535 [accn]
19 | base: sbase;
20 | rule: fnhe=79155
21 |-----|
```

length: 42386 Ln: 1 Col: 1 Sel: 0 Dos\Windows ANSI INS

Linkouts para la BBDD de Taxonomía

Ídem que con la de nucleótidos (texto, “.ft”,...)

Encabezamiento /"header"

linkOuts

```
1 -----
2 prid: 4283
3 dbase: taxonomy
4 stype: culture/stock collections
5 !base: http://161.111.170.202/herb/asp/dos.asp?conque=MA-Fungi
6 -----
7 linkid: 1
8 query: Coniophora arida [name]
9 base: $base;
10 rule: fnhe=47708
11 name: MA-Fungi 47708
12 -----
13 linkid: 2
14 query: Hypochnicium guineensis [name]
15 base: $base;
16 rule: fnhe=79156
17 name: MA-Fungi 79156
18 -----
19 linkid: 3
20 query: Hypochnicium michelii [name]
21 base: $base;
```

Identificador del proveedor, asignado por NCBI

BBDD desde la que se hace "LinkOut"

Tipo de "LinkOut"

Parte constante de las URL de "LinkOut"

Núm. de "LinkOut"; único en el fichero para cada Linkout

Nombre de la taxonomía de Genbank desde el que se hace el Linkout

Parte variable de la URL que hace referencia al ejemplar del Linkout

Nombre para el Linkout

length: 53451 Ln: 1 Col: 1 Sel: 0

holdings - GenBank - WinSCP

Local Mark Files Commands Session Options Remote Help

Synchronize Queue Transfer Settings Default

GenBank New Session

C: Disco local holdir Find Files

Upload Edit Properties Download Edit Properties

C:\Datos\Projects\GBIF\GBIF.ES\Formacion\2014\Colecciones y GenBank /holdings

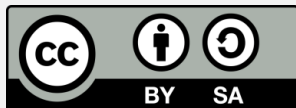
Name	Ext	Size	Type	Changed	Rights
..			Parent directory	21/03/2014 1	
~\$03-linkouts.pptx		165 B	Presentación de ...	21/03/2014 1	
01-BOL y Col.pptx		447 KiB	Presentación de ...	20/03/2014 3	
02-BBDD GenBank.pptx		1,385 KiB	Presentación de ...	21/03/2014 1	
03-linkouts.pptx		1,058 KiB	Presentación de ...	20/03/2014 4	
MA-Fungi-Accession...		12,219 B	Archivo FT	17/03/2014 1	
MA-Fungi-Accession...		53,451 B	Archivo FT	20/03/2014 1	rw-r--r-
MA-Fungi-Genbank.ft		42,386 B	Archivo FT	20/03/2014 9	rw-r--r-
providerinfo.xml		712 B	Archivo XML	17/03/2014 1	rw-rw-r

0 B of 2,997 KiB in 0 of 8 0 B of 96,329 B in 0 of 3

FTP 0:01:10

Francisco Pando

Unidad de coordinación, GBIF España
Real Jardín Botánico - CSIC
Claudio Moyano 1, 28014 Madrid, Spain
pando@gbif.es
www.gbif.es



<http://creativecommons.org/licenses/by-sa/3.0/es/>

