



# *Validación taxonómica*



I Taller GBIF.ES:  
Mejora de la calidad de  
datos de biodiversidad

*Cristina Ronquillo*

*Ayudante investigación*

mncn 25  museo  
nacional de  
ciencias  
naturales



**CSIC**

CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS

# 1 Taller GBIF.ES: Mejora de la calidad de datos de biodiversidad

**Cristina Ronquillo**

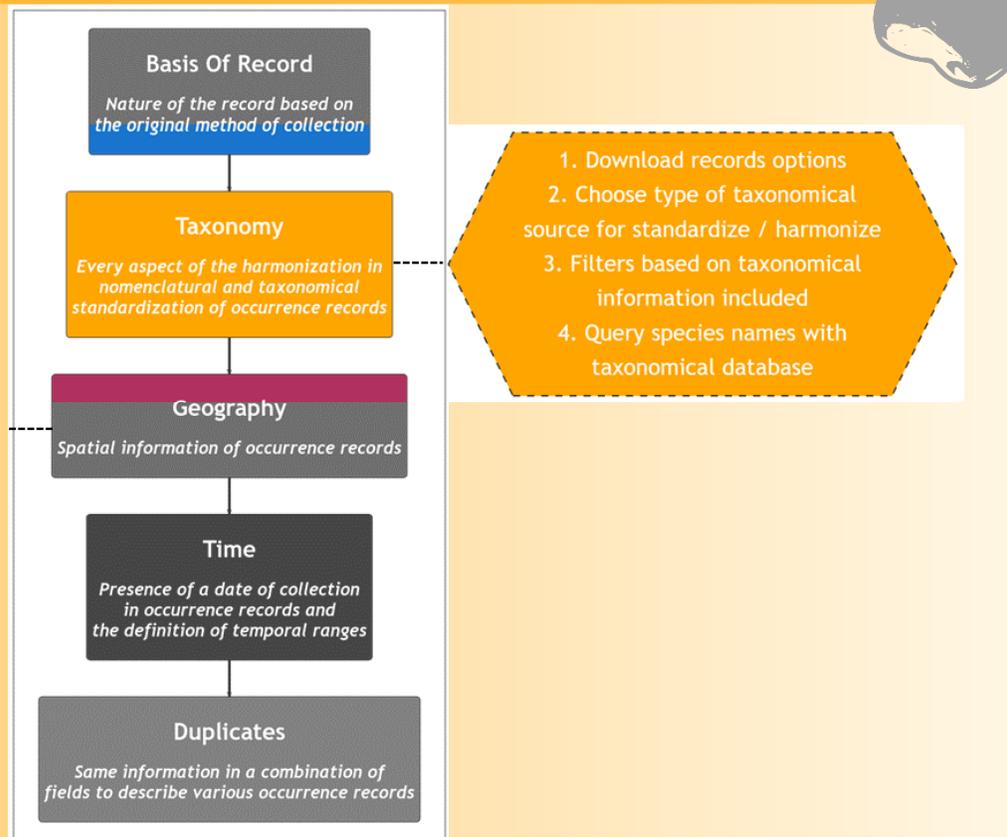
**Ayudante investigación**

mncn 25 1771  
2021 museo  
nacional de  
ciencias  
naturales



**CSIC**

CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS



## Objetivos

- ✓ ¿Qué tengo mirar?
- ✓ ¿Cómo comprobar los datos?
- ✓ ¿Qué tengo que descartar?
- ✓ ¿Qué tengo que corregir y cómo puedo hacerlo?



1 Taller GBIF.ES:  
Mejora de la calidad de  
datos de biodiversidad

**Cristina Ronquillo**

**Ayudante investigación**



***Allobates niputidea***  
(Grant, Acosta-Galvis & Rada, 2007)

A Name for the Species of *Allobates* (Anura: Dendrobatoidea: Aromobatidae)  
from the Magdalena Valley of Colombia

TARAN GRANT, ANDRÉS ACOSTA, AND MARCO RADA

Se describe una nueva especie de *Allobates* del Valle del Magdalena en Colombia. Esta especie y su especie hermana, *A. talamancense*, son las únicas dos especies de *Allobates* que ocurren al occidente de los Andes. La nueva especie difiere de *A. talamancense* por ser más pequeña, poseer una difusa franja oblicuolateral pálida, y tener el dedo III de la mano en los machos adultos fuertemente ensanchado. Además, las larvas difieren en el tamaño relativo de las papilas del labio anterior y posterior. Entre las especies cisandinas de *Allobates*, la más parecida es *A. melanolaemus*, con la cual comparte la difusa franja oblicuolateral pálida y la garganta negra en los machos adultos. La nueva especie se diferencia de *A. melanolaemus* por su menor tamaño y porque los machos adultos exhiben el dedo III de la mano fuertemente ensanchado y la pigmentación del pecho y del abdomen negra.

**Etymology.** - The specific epithet is a noun that is apposition and is the name commonly applied by Colombian herpetologists to this and other small and brown frogs of unknown identity.

**Etimología.** - El epíteto específico es un sustantivo es aposición y es el nombre comúnmente aplicado por herpetólogos Colombianos a esta y otras ranas pequeña y cafés de identidad desconocida.

mncn 25 1771 2021 museo nacional de ciencias naturales

**CSIC**  
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS



# Objetivos

✓ ¿Qué tengo mirar?

- Nombres científicos:
  - Binomiales (*Quercus ilex*)
  - Infraespecies: subespecies, variedades, subvariedades, formas y subformas.
- Deletreo: Códigos de nomenclatura
- Errores: typos, ambigüedades, sinónimos, etc.



# Errores típicos:

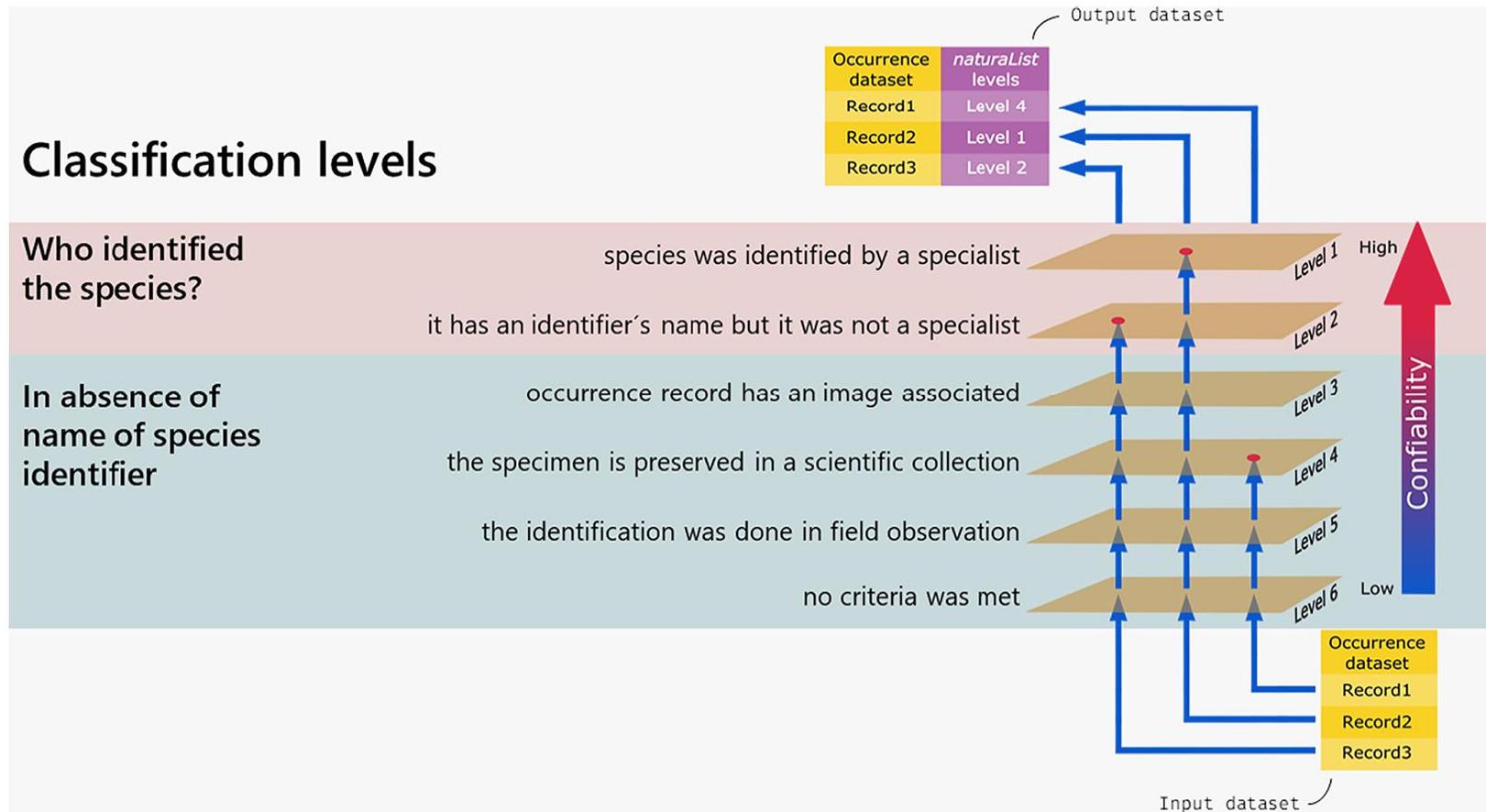
*Domain Schizophrenia* – This involves searching for fields used for purposes for which they may not have been intended. This often happens where a certainty field has not been included in the database and question marks, uncertainties such as cf., aff. are added in the same field as the species epithet, or comments added (Table 2). The nature of this ‘error’ may also depend on database design.

Family	Genus	Species
Myrtaceae	Eucalyptus	globulus?
Myrtaceae	Eucalyptus	? globulus
Myrtaceae	Eucalyptus	aff. globulus
Myrtaceae	Eucalyptus	sp. nov.
Myrtaceae	Eucalyptus	?
Myrtaceae	Eucalyptus	sp. 1
Myrtaceae	Eucalyptus	To be determined

*Table 2. Examples of Domain schizophrenia (from Chapman 2005a).*



# Identificación:



Rodrigues et al. 2022. Ecological Informatics  
<https://doi.org/10.1016/j.ecoinf.2022.101625>



# Módulo 2: Taxonomía

Home

Basis of Record

Taxonomic <

- >> 1. Download records options
- >> 2. Choose type of taxonomical source for standardize / harmonize
- >> 3. Filters based on taxonomical information included
- >> 4. Query species names with taxonomical database

Geographic <

Temporal

Duplicates

Final Report

References

Checklist Type   Spatial coverage   **Taxonomical coverage**   Matching Type

**Choose a source of taxonomic information to validate your records based on their taxonomical coverage:**

General (includes multiple taxa)

Specific (includes only study taxa)

Trade-off   **Methods**

Taxonomical coverage pros	Taxonomical coverage cons
Use in studies of specific taxa.	May not have an automated tool available for certain taxa.

Category	Value (approx. % of range)
Certainty	25
Data coverage	75

**Your selection:**

You set 0/4 steps to validate taxonomically and 1/4 to choose tax. db



**1. Opciones de descarga**

2. Escoger una fuente de información taxonómica para validar

3. Filtros de información

4. Estatus taxonómico



1. Opciones de descarga

2. Escoger una fuente de información taxonómica para validar

3. Filtros de información

4. Estatus taxonómico

Lista de especies cerrada



Descargar todo desde un nivel taxonómico superior





# 1. Opciones de descarga

# 2. Escoger una fuente de información taxonómica para validar

# 3. Filtros de información

# 4. Estatus taxonómico



## Lista de especies cerrada



The screenshot shows the GBIF Occurrences interface. The top navigation bar includes 'Get data', 'How-to', 'Tools', 'Community', and 'About'. The user 'cristinaronquillo' is logged in. The main content area shows a search for 'Panthera leo' and 'Panthera tigris' with 19,381 results. The 'TAXONOMY' tab is selected, displaying a hierarchical tree of taxonomic groups and a corresponding sunburst chart.

Taxonomic Group	Count
Animalia	19,381
Chordata	2,127,152,853
Mammalia	19,381
Carnivora	19,381
Felidae	19,381
Panthera	19,381
Panthera leo	12,796
Panthera tigris	6,585
Unknown class	2,127,133,472
Arthropoda	266,171,146
Mollusca	20,900,419



## 1. Opciones de descarga

2. Escoger una fuente de información taxonómica para validar

3. Filtros de información

4. Estatus taxonómico



# Descargar todo desde un nivel taxonómico superior



The screenshot shows the GBIF Occurrences interface. On the left, there are filter options for Occurrence status, Licence, and Scientific name. The 'Scientific name' filter is set to 'Panthera Oken, 1816'. Below this is a search bar and a 'Major groups' tree view. The main content area shows the 'TAXONOMIC DISTRIBUTION OF OCCURRENCES' with a tree view and a list of counts for each taxonomic level. A zoomed-in view of the 'Panthera' level is shown on the right.

Taxonomic Level	Count
Animalia	37,988
Chordata	37,988
Mammalia	37,988
Carnivora	37,988
Felidae	37,988
Panthera	37,988
Panthera leo	12,796
Panthera pardus	8,893
Panthera onca	8,662
Panthera tigris	6,585
Panthera spelaea	386
Panthera gombaszoegensis	63
Panthera blytheae	4
Unknown species	599



1. Opciones de descarga

**2. Escoger una fuente de  
información taxonómica  
para validar**

3. Filtros de información

4. Estatus taxonómico



1. Opciones de descarga

**2. Escoger una fuente de información taxonómica para validar**

3. Filtros de información

4. Estatus taxonómico





1. Opciones de descarga

2. Escoger una fuente de información taxonómica para validar

3. Filtros de información

4. Estatus taxonómico



*Rosmarinus officinalis*



*Salvia rosmarinus*



1. Opciones de descarga

2. Escoger una fuente de información taxonómica para validar

3. Filtros de información

4. Estatus taxonómico



Catalogue of Life



ZooBank

Welcome to the

World Spider Catalog Version 23.0

Currently **50'146** accepted species included

WFO  
The World Flora Online



Royal Botanic Gardens **Kew** | Plants of the World Online



International Plant Names Index

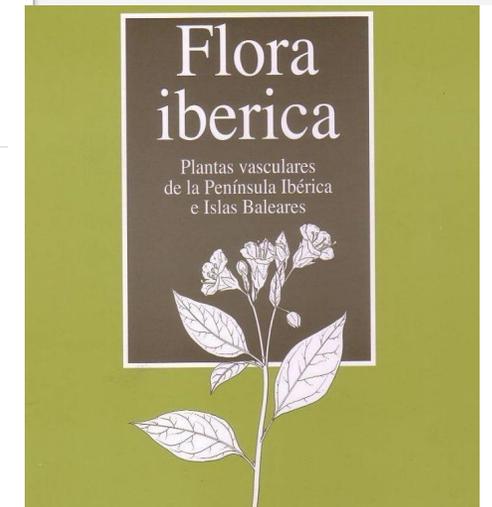
WORMS  
World Register of Marine Species



TNRS

Taxonomic Name Resolution Service

An online tool for the standardization of plant taxonomic names.



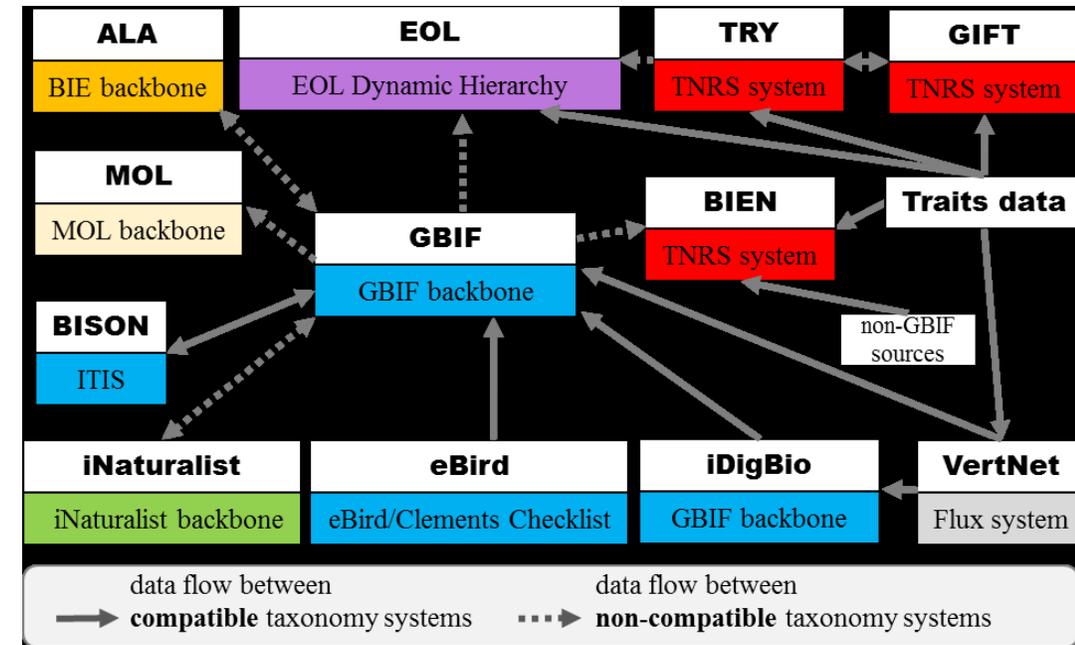
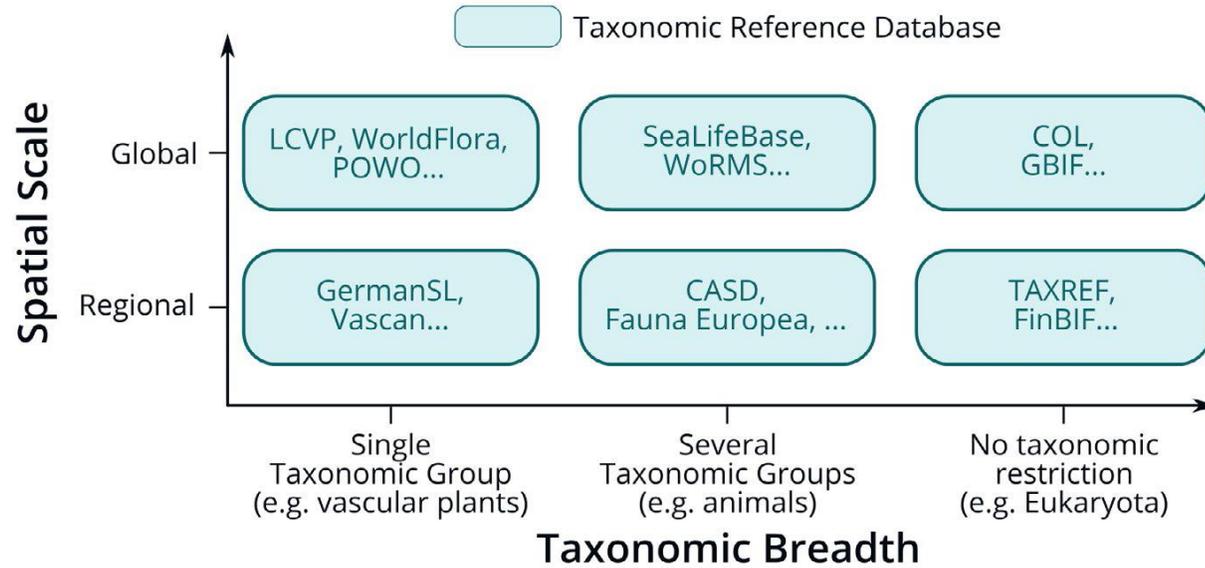


1. Opciones de descarga

2. Escoger una fuente de información taxonómica para validar

3. Filtros de información

4. Estatus taxonómico



Feng et al. *GEB* (2022)

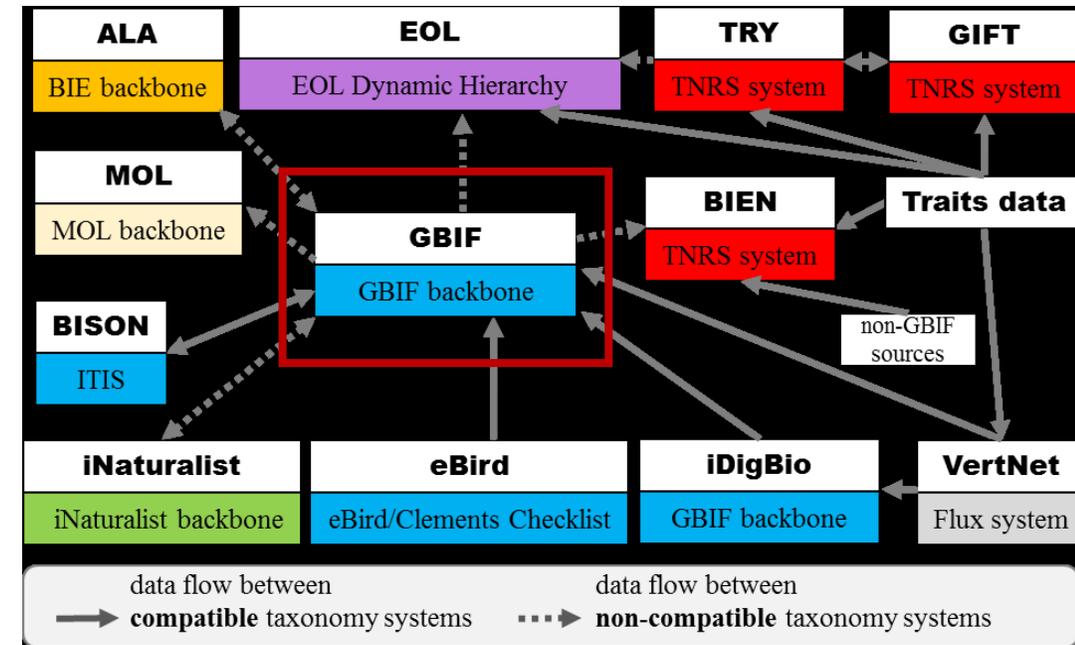
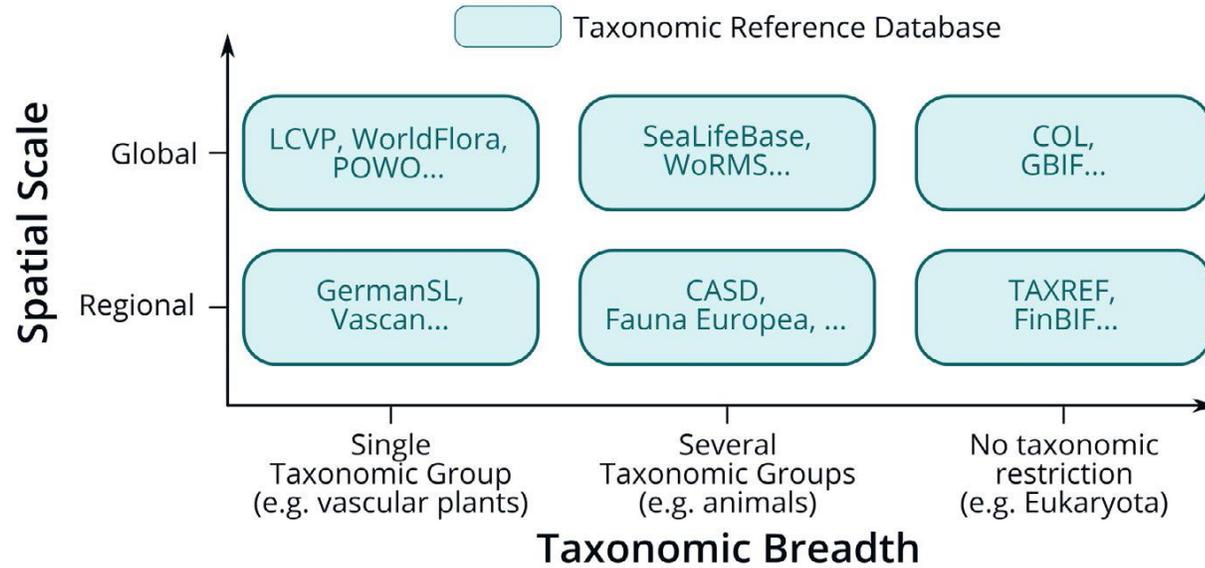


1. Opciones de descarga

2. Escoger una fuente de información taxonómica para validar

3. Filtros de información

4. Estatus taxonómico



Feng et al. *GEB* (2022)



1. Opciones de descarga

2. Escoger una fuente de información taxonómica para validar

3. Filtros de información

4. Estatus taxonómico

## ***GBIF BACKBONE***

*“The GBIF Backbone Taxonomy is a single, synthetic management classification with the goal of covering all names GBIF is dealing with. It's the taxonomic backbone that allows GBIF to integrate name based information from 105 resources (...)”*





1. Opciones de descarga

2. Escoger una fuente de información taxonómica para validar

3. Filtros de información

4. Estatus taxonómico

## Tipos

Manual

Automática



1. Opciones de descarga

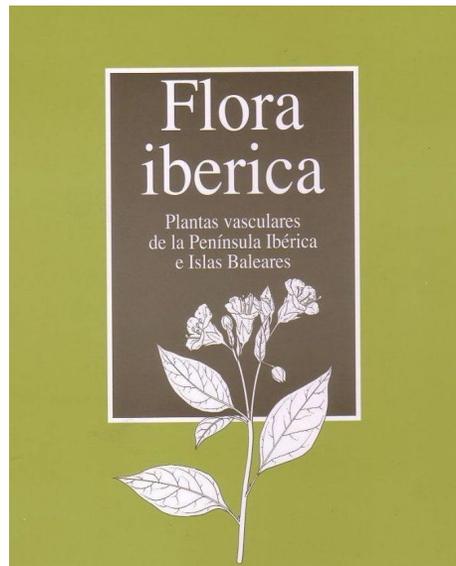
2. Escoger una fuente de información taxonómica para validar

3. Filtros de información

4. Estatus taxonómico

## Tipos

Manual

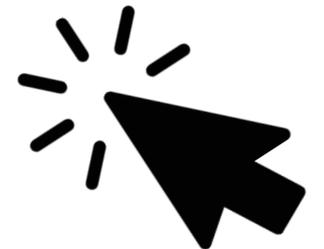


Automática



```
library(rgbif) # tax. standardization  
library(Taxonstand) # tax. standardization  
library(worldFlora) # tax. standardization
```

Taxize





1. Opciones de descarga

2. Escoger una fuente de información taxonómica para validar

3. Filtros de información

4. Estatus taxonómico

# Tipos

Manual

Automática



**WoRMS**  
World Register of Marine Species

Home About Subregisters Users Photogallery Documents LifeWatch Contribute

Quick search...

## WoRMS name details

An authoritative classification **★ *Conaea aestus* Oliveira, 1946**



<b>AphiaID</b>	350304 (urn:lsid:marinespecies.org:taxname:350304)
<b>Classification</b>	Biota > ★ Animalia (Kingdom) > ☆ Arthropoda (Phylum) > ★ Crustacea (Subphylum) > ★ Multicrustacea (Superclass) > ★ Copepoda (Class) > ★ Neocopepoda (Infraclass) > ★ Podoplea (Superorder) > ★ Cyclopoida (Order) > ★ Ergasilida (Suborder) > ★ Oncaeidae (Family) > ★ <i>Conaea</i> (Genus) > ★ <i>Conaea aestus</i> (Species)
<b>Status</b>	uncertain > taxon inquirendum
<b>Rank</b>	Species
<b>Parent</b>	★ <i>Conaea</i> Giesbrecht, 1891
<b>Original name</b>	★ <i>Conaea aestus</i> Oliveira, 1946
<b>Environment</b>	marine, brackish, fresh, terrestrial
<b>Original description</b>	Oliveira, L.P.H. de. (1946). Estudos sobre o microplâncton capturado durante a viagem do Navio Hidrográfico Lahmeyer nas baías de Ilha Grande e Sepetiba. Memórias do Instituto Oswaldo Cruz, Rio de Janeiro 44(3):441-488, figs. 1-14, pls. 1-12. (ix-1946) [details] Available for editors  [request]
<b>Descriptive notes</b>	★ <b>Taxonomy</b> <i>Conaea aestus</i> Oliveira, 1946 is regarded as species inquirenda in the family Oncaeidae based on the following grounds: in...
<b>Taxonomic citation</b>	Walter, T.C.; Boxshall, G. (2024). World of Copepods Database. <i>Conaea aestus</i> Oliveira, 1946. Accessed through: World Register of Marine Species at: <a href="https://www.marinespecies.org/aphia.php?p=taxdetails&amp;id=350304">https://www.marinespecies.org/aphia.php?p=taxdetails&amp;id=350304</a> on 2024-08-20

### News



**Global Cyperaceae Database now**  
Added on: 2024-08-14 10:08:33 by Dekeyser, Stefanie

Through the iSedge project, the Cyperaceae



1. Opciones de descarga

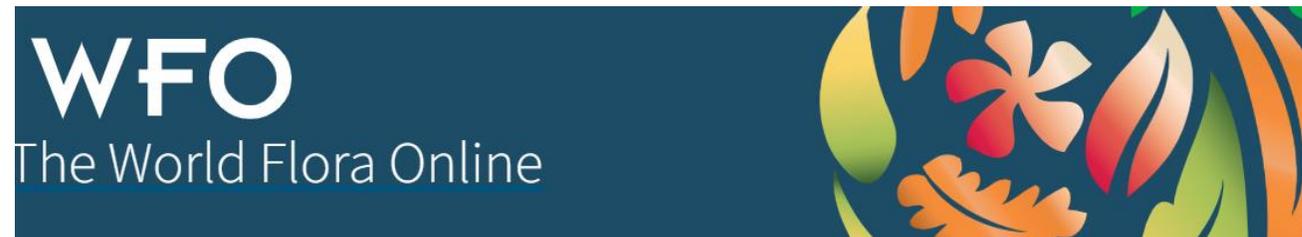
2. Escoger una fuente de información taxonómica para validar

3. Filtros de información

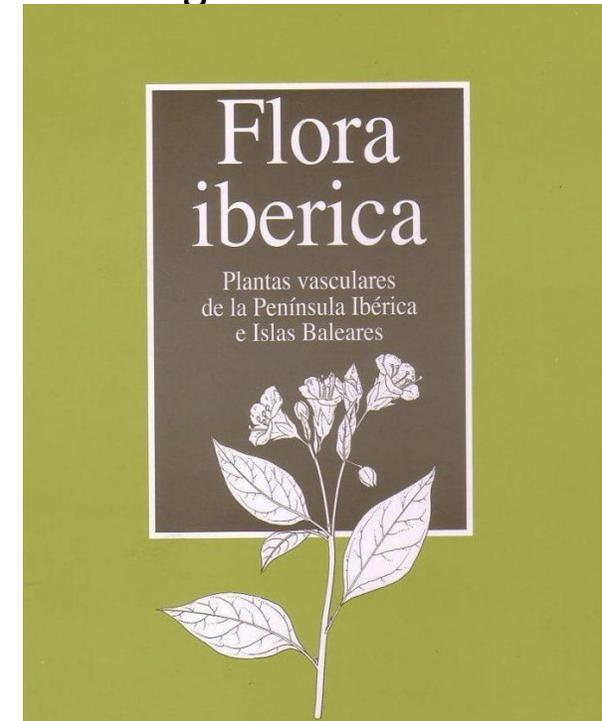
4. Estatus taxonómico

## Cobertura espacial:

Global



Regional





1. Opciones de descarga

2. Escoger una fuente de información taxonómica para validar

3. Filtros de información

4. Estatus taxonómico

## Cobertura taxonómica

General



Específico

Welcome to the

World Spider  
Catalog Version  
23.0

Currently **50'146** accepted species included



1. Opciones de descarga

2. Escoger una fuente de información taxonómica para validar

3. Filtros de información

4. Estatus taxonómico

## Tipos

Manual

Automática

### ‘Coincidencia de cadenas de texto’

- ‘Fuzzy’
- Exacta

*Similarity and levenshtein distance*

G	E	E		I
G	E	E	L	Y

G	E	E	I	
G	E	E	L	Y



1. Opciones de descarga

2. Escoger una fuente de información taxonómica para validar

3. Filtros de información

4. Estatus taxonómico

## Tipos

Manual

### ‘Coincidencia de cadenas de texto’

- ‘Fuzzy’ → Quercus ilex / Quercus ilex
- Exacta

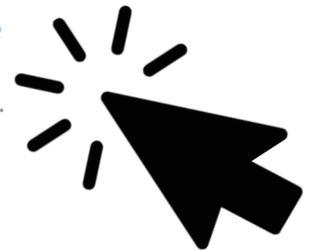
Automática



# TNRS

Taxonomic Name Resolution Service

An online tool for the standardization of plant taxonomic names.





1. Opciones de descarga

2. Escoger una fuente de información taxonómica para validar

3. Filtros de información

4. Estatus taxonómico

## Tipos

Manual

Automática

### ‘Coincidencia de cadenas de texto’

- ‘Fuzzy’ → Quercus ilex / Quercus ilex
- Exacta



**Abutilon grandifolium**



1. Opciones de descarga

2. Escoger una fuente de información taxonómica para validar

3. Filtros de información

4. Estatus taxonómico



## Abutilon grandiflorum

Change selected name

Select	Name Matched	Source	Overall Score	Author Matched	Author Score	Accepted Name	Unmatched Terms	Taxonomic Status
<input checked="" type="checkbox"/>	Abutilon grandiflorum G.Don	WFO	.99			Abutilon grandiflorum G.Don		Accepted
<input type="checkbox"/>	Abutilon grandifolium (Willd.) Sweet	WFO	.97			Abutilon grandifolium (Willd.) Sweet		Accepted

CLOSE



1. Opciones de descarga

2. Escoger una fuente de información taxonómica para validar

3. Filtros de información

4. Estatus taxonómico

Methods in Ecology and Evolution

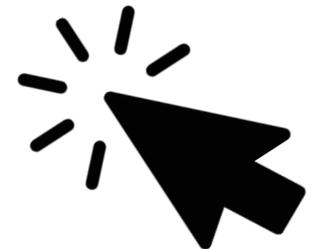


REVIEW |  Open Access |  

## Harmonizing taxon names in biodiversity data: A review of tools, databases and best practices

Matthias Grenié  Emilio Berti, Juan Carvajal-Quintero, Gala Mona Louise Dädlow, Alban Sagouis, Marten Winter

First published: 17 January 2022 | <https://doi.org/10.1111/2041-210X.13802> | Citations: 20





1. Opciones de descarga

2. Escoger una fuente de información taxonómica para validar

**3. Filtros de información**

4. Estatus taxonómico



1. Opciones de descarga

2. Escoger una fuente de información taxonómica para validar

**3. Filtros de información**

4. Estatus taxonómico

- Campos importantes:

kingdom	phylum	Class	Order	family	scientificName	genus	specificEpithet	scientificNameAuthorship



1. Opciones de descarga

2. Escoger una fuente de información taxonómica para validar

**3. Filtros de información**

4. Estatus taxonómico

- Campos importantes:

## Taxonomic Rank

scientificName	genus	specificEpithet
<i>Canis lupus</i>	<i>Canis</i>	<i>lupus</i>





1. Opciones de descarga

2. Escoger una fuente de información taxonómica para validar

3. Filtros de información

4. Estatus taxonómico

- Campos importantes:

## Taxonomic Rank

scientificName	genus	specificEpithet
<i>Canis lupus</i>	<i>Canis</i>	<i>lupus</i>

*Subspecies!*

*Canis lupus familiaris* != *Canis lupus lycaon*





1. Opciones de descarga

2. Escoger una fuente de información taxonómica para validar

**3. Filtros de información**

4. Estatus taxonómico

- Campos importantes:

kingdom	phylum	Class	Order	family	scientificName	genus	specificEpithet	scientificNameAuthorship

***Genus = 'Prunella'***



1. Opciones de descarga

2. Escoger una fuente de información taxonómica para validar

3. Filtros de información

4. Estatus taxonómico

- Campos importantes:

kingdom	phylum	Class	Order	family	scientificName	genus

***Genus = 'Prunella'***





1. Opciones de descarga

2. Escoger una fuente de información taxonómica para validar

**3. Filtros de información**

4. Estatus taxonómico

- Campos importantes:

Nombres de autores

kingdom	phylum	Class	Order	family	scientificName	genus	specificEpithet	scientificNameAuthorship



1. Opciones de descarga

2. Escoger una fuente de información taxonómica para validar

3. Filtros de información

4. Estatus taxonómico

- Campos importantes:

Nombres de autores

kingdom	phylum	Class	Order	family	scientificName	genus	specificEpithet	scientificNameAuthorship
					<i>Brassica rupestris</i> Raf.	<i>Brassica</i>	<i>rupestris</i>	
					<i>Brassica rupestris</i> Kuntze	<i>Brassica</i>	<i>rupestris</i>	



1. Opciones de descarga

2. Escoger una fuente de información taxonómica para validar

**3. Filtros de información**

4. Estatus taxonómico

- Campos importantes:

Nombres de autores

kingdom	phylum	Class	Order	family	scientificName	genus	specificEpithet	scientificNameAuthorship
					<i>Spiraea hypericifolia</i>			



 *Spiraea hypericifolia* Kar. & Kir.  
; 1842 348 1842  
**Status:**Unchecked  
**Rank:**Species  
**Family:** Rosaceae  
**Order:** Rosales

 *Spiraea hypericifolia* R.Parker  
A forest flora for the Punjab ; 1918 212 1918  
**Status:**Unchecked  
**Rank:**Species  
**Family:** Rosaceae  
**Order:** Rosales

 *Spiraea hypericifolia* Cambess.  
; 1824 355 1824  
**Status:**Unchecked  
**Rank:**Species  
**Family:** Rosaceae  
**Order:** Rosales

 *Spiraea hypericifolia* DC.  
Flore française ; 1805 645 1815  
**Status:**Unchecked  
**Rank:**Species  
**Family:** Rosaceae  
**Order:** Rosales

 *Spiraea hypericifolia* Turcz.  
Flora baicalensis-dahurica ; 1842 359 1843  
**Status:**Unchecked  
**Rank:**Species

 *Spiraea hypericifolia* DC.  
Flore française ; 1805 645 1815  
**Status:**Unchecked  
**Rank:**Species  
**Family:** Rosaceae  
**Order:** Rosales

 *Spiraea hypericifolia* Turcz.  
Flora baicalensis-dahurica ; 1842 359 1843  
**Status:**Unchecked  
**Rank:**Species  
**Family:** Rosaceae  
**Order:** Rosales

 *Spiraea hypericifolia* Nyman  
Conspectus florae europaeae ; 1878 215 1878  
**Status:**Unchecked  
**Rank:**Species  
**Family:** Rosaceae  
**Order:** Rosales

 *Spiraea hypericifolia* Sm.  
The Cyclopaedia ; 1802 *Spiraea* no.7 1816  
**Status:**Synonym of [Spiraea italica](#) Raf.  
**Rank:**Species  
**Family:** Rosaceae

 *Spiraea hypericifolia* M.Bieb.  
Fl. Taur.-Caucas. 1: 392. 1808  
**Status:**Unchecked  
**Rank:**Species  
**Family:** Rosaceae

 *Spiraea hypericifolia* Marshall  
Arbust. Amer. 146. 1785  
**Status:**Unchecked  
**Rank:**Species  
**Family:** Rosaceae  
**Order:** Rosales

 ***Spiraea hypericifolia* L.**  
Sp. Pl. : 489 (1753)  
**Status:**Accepted Name  
**Rank:**Species  
**Family:** Rosaceae  
**Order:** Rosales



1. Opciones de descarga

2. Escoger una fuente de información taxonómica para validar

3. Filtros de información

4. Estatus taxonómico

- Campos importantes:

## Nombres de autores

kingdc	phylum	Class

 SPECIES  
*Quercus ilex* Lour.  
Synonym of: *Quercus helferiana*  
A.DC.

 SPECIES  
*Quercus ilex* L.  
Common Name: Holm Oak



scientificNameAuthorship

---

---

---



1. Opciones de descarga

2. Escoger una fuente de información taxonómica para validar

3. Filtros de información

**4. Estatus taxonómico**



1. Opciones de descarga

2. Escoger una fuente de información taxonómica para validar

3. Filtros de información

4. Estatus taxonómico

 Spiraea hypericifolia Kar. & Kir. ; 1842 348 1842 <b>Status:</b> Unchecked <b>Rank:</b> Species <b>Family:</b> Rosaceae <b>Order:</b> Rosales	
 Spiraea hypericifolia R.Parker A forest flora for the Punjab ; 1918 212 1918 <b>Status:</b> Unchecked <b>Rank:</b> Species <b>Family:</b> Rosaceae <b>Order:</b> Rosales	 Spiraea hypericifolia Marshall Arbust. Amer. 146. 1785 <b>Status:</b> Unchecked <b>Rank:</b> Species <b>Family:</b> Rosaceae <b>Order:</b> Rosales
 Spiraea hypericifolia Cambess. ; 1824 355 1824 <b>Status:</b> Unchecked <b>Rank:</b> Species <b>Family:</b> Rosaceae <b>Order:</b> Rosales	 <b>Spiraea hypericifolia L.</b> Sp. Pl. : 489 (1753) <b>Status:</b> Accepted Name <b>Rank:</b> Species <b>Family:</b> Rosaceae <b>Order:</b> Rosales
 Spiraea hypericifolia DC. Flora française ; 1805 643 1815 <b>Status:</b> Unchecked <b>Rank:</b> Species <b>Family:</b> Rosaceae <b>Order:</b> Rosales	
 Spiraea hypericifolia Turcz. Flora baicalensis-dahurica ; 1842 359 1843 <b>Status:</b> Unchecked <b>Rank:</b> Species <b>Family:</b> Rosaceae <b>Order:</b> Rosales	

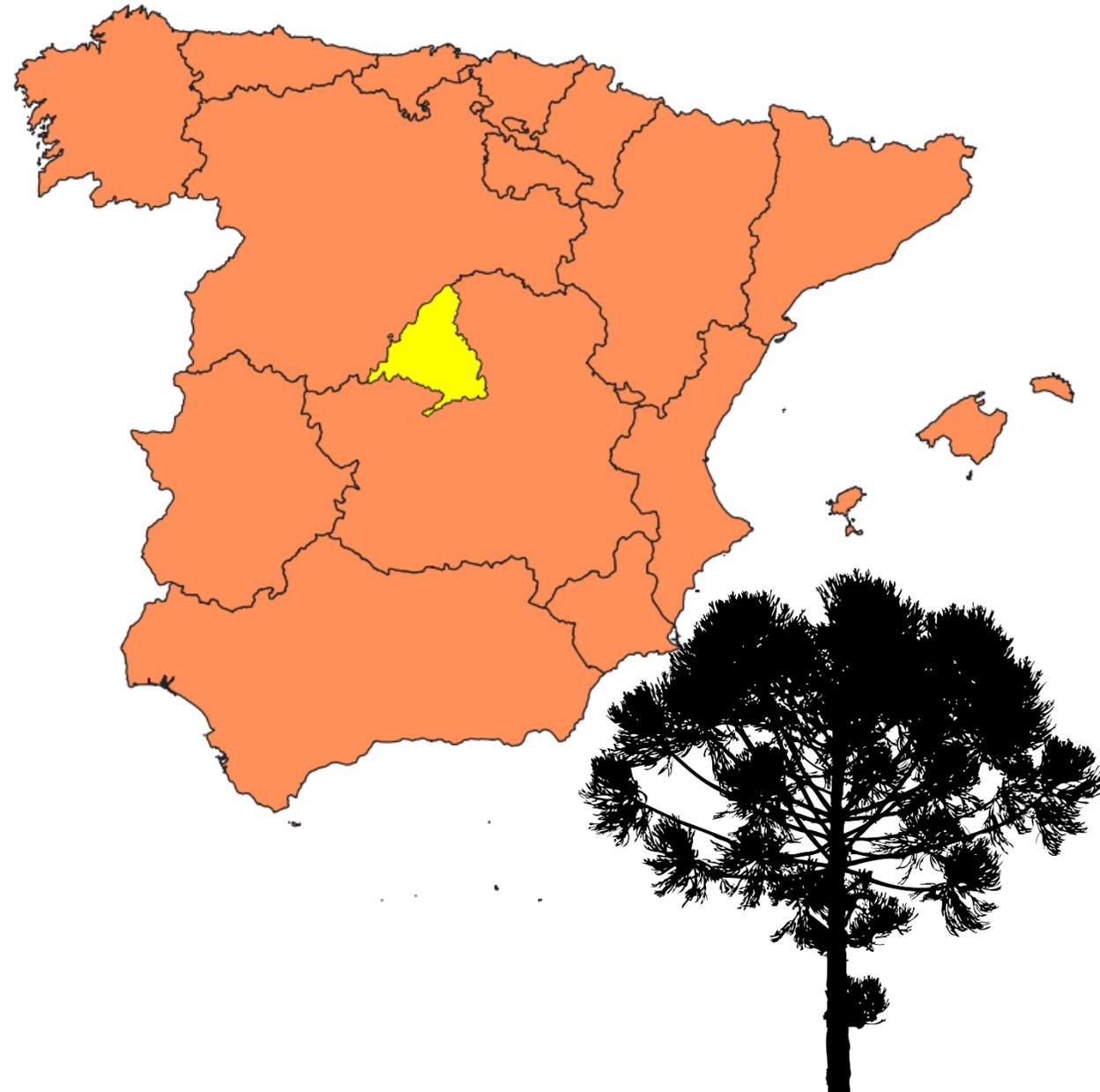
Casos:

Aceptado

Sinónimo

Sin resolver

(Unresolved / Unchecked / No opinion)





# PRÁCTICA 1

## Download Information

DOI: <https://doi.org/10.15468/dl.2h3r3q> (may take some hours b

Creation Date: 07:54:20 20 August 2024

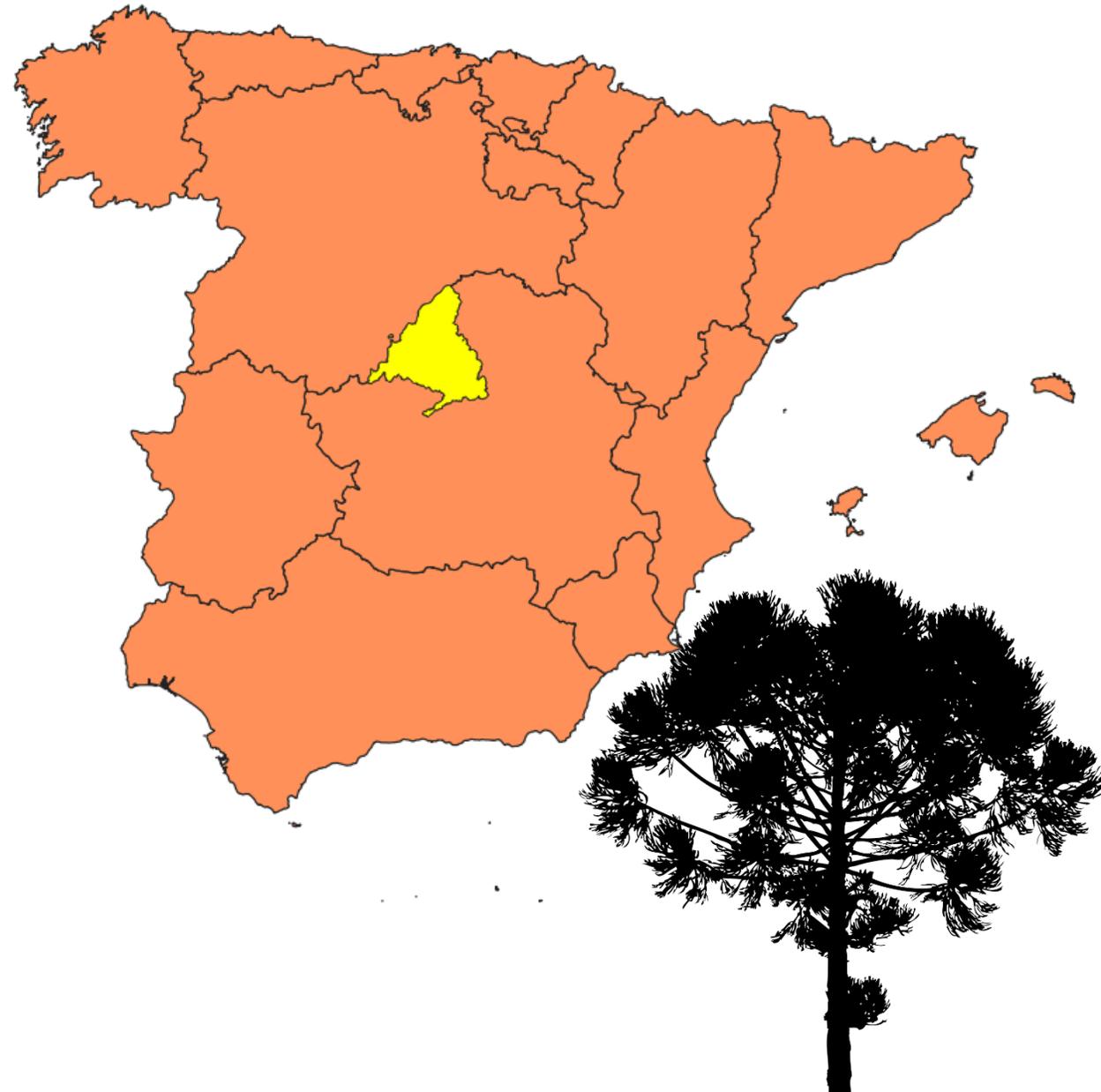
Records included: 9418 records from 37 published datasets

Compressed data size: 474.0 kB

Download format: simple tab-separated values (TSV)

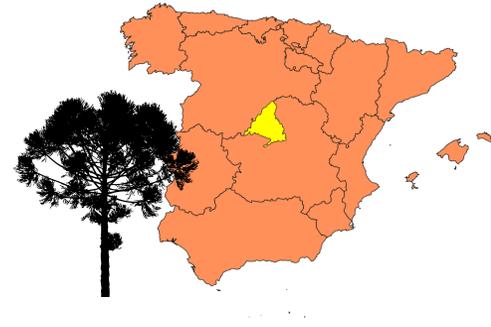
Filter used:

```
{  
  "and" : [  
    "GadmGid is ESP.8_1",  
    "TaxonKey is Pinopsida"  
  ]  
}
```





# PRÁCTICA 1



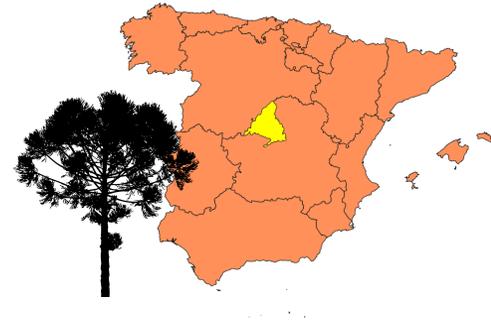
Responde a las cuestiones planteadas en **Formulario\_práctica1.pdf**

1. Abre el archivo 'occurrences\_CAM\_Pinopsida.csv' en Excel

Recordatorio: usa el tabulador como separador de campos



# PRÁCTICA 1



Responde a las cuestiones planteadas en **Formulario\_práctica1.pdf**

1. Abre el archivo 'occurrences\_CAM\_Pinopsida.csv' en Excel

Recordatorio: usa el tabulador como separador de campos

2. Crea tu lista taxonómica:

Escoge los campos necesarios para una validación taxonómica

"family", "genus", "species", "infraspecificEpithet",

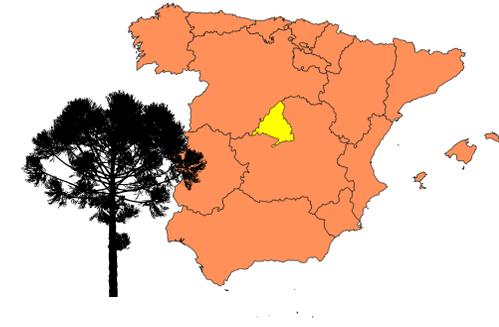
"taxonRank",

"scientificName", "verbatimScientificName", "verbatimScientificNameAuthorship",

"identifiedBy"



# PRÁCTICA 1



Ahora probamos una forma automática:



**WFO**  
The World Flora Online

