The Flora Mycologica Iberica Project fungi occurrence dataset

Francisco Pando¹, Margarita Dueñas¹, Carlos Lado¹, María Teresa Telleria¹

¹ Real Jardín Botánico-CSIC, Claudio Moyano 1, 28014, Madrid, Spain

Corresponding author: Francisco Pando (pando@rjb.csic.es)

Academic editor: C. Gueidan


Abstract
The dataset contains detailed distribution information on several fungal groups. The information has been revised, and in many times compiled, by expert mycologist(s) working on the monographs for the Flora Mycologica Iberica Project (FMI). Records comprise both collection and observational data, obtained from a variety of sources including field work, herbaria, and the literature. The dataset contains 59,235 records, of which 21,393 are georeferenced. These correspond to 2,445 species, grouped in 18 classes. The geographical scope of the dataset is Iberian Peninsula (Continental Portugal and Spain, and Andorra) and Balearic Islands. The complete dataset is available in Darwin Core Archive format via the Global Biodiversity Information Facility (GBIF).

Key words
Gasteromycetes, Myxomycetes, Agaricales, Boletales, Aphyllophorales, Dothideales, Gomphales, Helotiiales, Laboulbeniales, Tremellales, Ustilaginales, Cantharellaceae, Andorra, Balearic Islands, Portugal, Spain, fungi, mitosporic fungi, fungi and lichens, occurrence, observation

Copyright Francisco Pando et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
Introduction

The “Flora Mycologica Iberica Project dataset” is one of the main results produced by the “Flora Mycologica Iberica”, a research project that stood from 1988 to 2008 and involved over 30 researchers from Spain and Portugal. This dataset contains information on 2,445 species of fungi recorded from The Iberian Peninsula and Balearic Islands. As an online resource, it is a valuable source of information on fungi growing in that area, with a high reuse potential, given its coverage, taxonomic scrutiny --carried out by taxonomic experts in the different groups-- and validation processes of the associate information (location, habitat). The other major outcome of the project were the monographs of the studied groups: Aphyllophorales [p.p.] (Telleria and Melo 1995), Myxomycetes [p.p.] (Lado and Pando 1997), Gasteromycetes [p.p.] (Calonge 1998), Laboulbeniales (Santamaria 1998, 2003) and Dictyosporic Dothideales (Checa 2004). These publications provide descriptions, illustrations, identification keys and additional information on many of the species and taxa included in this dataset.

General description

Purpose: This dataset was conceived within the Flora Mycologica Iberica Project (FMI). The ultimate objective of the FMI project was to make a critical flora which enables the identification of fungi naturally growing in the Iberian Peninsula and Balearic Islands (excluding parasites of humans and other mammals).

The purpose of the dataset was:

1. To gather the available information on Iberian fungi as published in the scientific literature, and to establish baseline knowledge for the project and the project’s monographs, among other aims.

2. To incorporate the primary data produced and compiled during the project to the previously available information, and thus providing updated and validated data on Iberian fungi. These data were gathered or verified by professional researchers, many of whom also authored the project’s monographs (Calonge 1998, Checa 2004, Lado and Pando 1997, Santamaria 1998, 2003, Telleria and Melo 1995) and/or check-lists (Calonge 1990, Castro 1998, Checa 1997a, 1997b, 1998, Dueñas, 2002, Garcia-Blazquez et al. 2007, Justo and Castro 2007, 2010, Lado 1991, Melo et al. 2007, Telleria 1990). Primary data added in this way mostly came from studied herbarium specimens and field surveys. Targeted field surveys on sites poorly known and of ecological or conservation relevance were carried out within the framework of the project. Unpublished records gathered within the project, along with those found in the literature, were also made available as publications in the “Cuadernos de Trabajo de Flora Micológica” series (see References). This is reflected in the ‘AssociatedReference’ column of the dataset.
Project details


Project title: Flora Mycologica Iberica

Funding: Funded by Spanish Government under its National Science Plan through its Competitive Grant Programs. Awards:

Flora Mycologica Iberica I
DGCYT PB87-0370 (1987-1991)
Flora Mycologica Iberica II
Flora Mycologica Iberica III
DGES PB95-0129-C03 (1995-1997)
Flora Mycologica Iberica IV
PB98-0538-C04 (1998-2001)
Flora Mycologica Iberica V
REN2002-04068-C02-01 (2002-2004)
Flora Mycologica Iberica VI
CGL2006-12732-C02-01 (2005-2008)

Data capture was carried out mostly by an agreement between CSIC and National Institute of Employment of Spain (INEM). Program: Módulo de documentación Botánica INEM-CSIC.

Study area descriptions/descriptor: Terrestrial biomes and freshwater environments in the Iberian Peninsula and Balearic Islands.

Design description: The final objective of the FMI project was to make a critical flora which enables the identification of fungi naturally growing in the Iberian Peninsula and Balearic Islands (excluding parasites of humans and other mammals).

One of the pillars in pursuing this objective was the compilation of the literature-based information about Iberian fungi. This was planned as a two-step process:

The first one was to compile a bibliography of all published works recording fungi from the Iberian Peninsula (Pando et al. 1990, Cardoso and Melo 1992, Pando 1996). This in itself is a very robust resource for carrying out all kinds of studies involving fungi in the Iberian Peninsula.

The second phase was to enter chorological records (species occurrences) contained in those publications into a database. After some trial and error, we aimed to compile a dataset of all genera cited in all publications. That task was accomplished surprisingly quickly (c. 79,000 genus-in-publication entries for c. 10,000 works in less than two years, by two full time data entry persons and one part-time scientific supervisor). With
that information at hand, we could target the specific publications and occurrences needed for the check-lists and monographs to be prepared within the FMI project.

Data entry management and publication were carried out using the Bibmaster software v. 3.7 (Pando et al 2004).

Manuals and guidelines were prepared to establish a clear standard basis regarding data entry and quality control procedures (Pando 1991, Pando et al. 1999).

Data published through GBIF: http://www.gbif.es/ipt/resource?r=floramicologiaiberica

Taxonomic coverage

General taxonomic coverage description: Dataset comprising distribution records of fungal species belonging to selected groups (Agaricomycetes, Dothideomycetes, Mycetozoa, Laboulbeniomycetes, Ustilaginomycetes and aquatic Hyphomycetes) found in the Iberian Peninsula and Balearic Islands (Western Europe). Sources included literature, herbaria and field surveys.

Taxonomic ranks

The consensus classification provided by Index Fungorum in Catalogue of Life (Kirk 2016) has been followed for taxonomic categories above genus.

It is worth mentioning that recent and profound changes in fungal classification have rendered some categories used in the project such as “Gasteromycetes” or “Aphyllophorales” obsolete. These groups, now referred informally as Gasteroid and Corticioid fungi, are especially well-covered, but scattered across a number of orders (cf. Pegler et al. 1995, Kirk et al. 2008, Larsson 2007).

![Figure 1. Taxonomic distribution of the dataset (percentage of specimens per classes).](image)
Figure 2. Visual representations of taxon record abundance in the dataset.

Classes and orders included in the dataset follow:

<table>
<thead>
<tr>
<th>Classes</th>
<th>Orders</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agaricomycetes</td>
<td>Agaricales*†, Atheliales†, Auriculariales, Boletales*†, Cantharellales†, Corticales†, Geastrales*, Gloeophyllales†, Gomphales*†, Hymenochaetales†, Hysterangiales*, Phallales*, Polyporales†, Russulales†, Sebacinales, Stereaceae, Thelephorales†, Trechisporales†</td>
</tr>
<tr>
<td>Atractiellomycetes</td>
<td>Atractiellales</td>
</tr>
<tr>
<td>Classiculomycetes</td>
<td>Classiculales</td>
</tr>
<tr>
<td>Dacrymycetes</td>
<td>Dacrymycetales</td>
</tr>
<tr>
<td>Dothideomycetes</td>
<td>Capnodiales, Hysteriales, Mytilinidiales, Pleosporales</td>
</tr>
<tr>
<td>Eurotiomycetes</td>
<td>Not assigned</td>
</tr>
<tr>
<td>Exobasidiomycetes</td>
<td>Doassansiales, Entylomatales, Exobasidiales, Tilletiales</td>
</tr>
<tr>
<td>Laboulbeniomycetes</td>
<td>Laboulbeniales</td>
</tr>
<tr>
<td>Leotiomycetes</td>
<td>Helotiales, Leotiales</td>
</tr>
<tr>
<td>Microbotryomycetes</td>
<td>Microbotryales</td>
</tr>
<tr>
<td>Myxomycetes</td>
<td>Echinosteliales, Liceales, Physarales, Steronitales, Trichiales</td>
</tr>
<tr>
<td>Not assigned</td>
<td>Zoopagales</td>
</tr>
<tr>
<td>Orbiliomycetes</td>
<td>Orbiliiales</td>
</tr>
<tr>
<td>Pezizomycetes</td>
<td>Peziales</td>
</tr>
<tr>
<td>Protosteliomycetes</td>
<td>Ceratiomyxales</td>
</tr>
<tr>
<td>Pucciniomycetes</td>
<td>Helicobasidiales, Platygloeales, Septobasidiales</td>
</tr>
<tr>
<td>Sordariomycetes</td>
<td>Chaetosphaeriales, Hypocreales, Microascales, Sordariales, Xylariales</td>
</tr>
<tr>
<td>Tremellomycetes</td>
<td>Tremellales</td>
</tr>
<tr>
<td>Ustilaginomycetes</td>
<td>Urocystidales, Ustilaginales</td>
</tr>
</tbody>
</table>

* Orders containing gasteroid taxa. † Orders containing corticioid taxa.
Figure 3. Geographic distribution of the georeferenced records. The darker the color, the higher is the record density.

Figure 4. Geographic distribution. Records grouped by provinces, georeferenced only.
The Flora Mycologica Iberica Project fungi occurrence dataset

Figure 5. Geographic distribution. Records grouped by provinces, not georeferenced. Provinces where non-georeferenced records provide substantial information missing in georeferenced records are encircled in red.

Figure 6. Temporal distribution of the records.
Spatial coverage

**General spatial coverage:** Iberian Peninsula, Balearic Islands, South-Western Europe

**Coordinates:** 35°45’36”N and 44°2’60”N Latitude; 9°56’60”W and 4°54’36”E Longitude

This comprises: Continental Portugal and Spain, Andorra and Balearic Islands. No records from Gibraltar (UK) have been included. A map showing georeference records and its density is provided (Fig. 3). Two additional maps showing records aggregated by province are presented here to show how records with coordinates provide only a partial view of the actual knowledge on Iberian fungi (Figs 4, 5). This highlights the importance of retrospective georeferencing when carrying on species distribution models and other geospatial analyses.

**Temporal coverage:** January 1, 1870 - February 20, 2012

Methods

**Study extent description:** Scientific literature was the main source for fungal occurrence records. Herbarium revisions, which included published and unpublished records, supplemented literature information. Additionally, targeted field campaigns were carried out within the framework of the project to fill gaps on sites poorly known and of ecological or conservation relevance such as national parks, and other protected areas. Unpublished relevant data were published as the “Bases Corológicas series in “Cuadernos de Trabajo” and reflected in the AssociatedReference field.

**Sampling description:** These two data avenues were subjected to different methodologies, as explained below.

Data collation from literature references. Three procedures are defined in this area:

1) Identifying and obtaining relevant publications.

A set of explicit criteria to determine whether a publication was eligible to be included into the database was defined. These were published by Pando (1996: 215-217). The Library of the Real Jardín Botánico-CSIC was the main source for literature. When an eligible publication was identified and not found in this library, a copy was obtained by the usual procedures (library exchange, colleagues, etc.) and deposited in that Library.

2) Treated genus data entry.

Genus names were extracted in a systematic way from the publications and entered into the database. Publications and genus names compiled up to, and including 1995 were published in three volumes of the “Cuadernos de Trabajo” series (Pando et al. 1990, Cardoso and Melo 1992, Pando 1996). At this point, the database contained data from c. 5,000 publications, at the end of the project this gathered data pertaining to c. 10,000 publications. Although teleomorph nomenclature is used in the dataset, literature collation was made for anamorphic as well as teleomorphic genera.
3) Occurrence data entry.

A protocol in which occurrence records were entered targeting specific taxonomic groups --on the basis of the project’s priorities and the schedule for the publication of the monographs-- was implemented at the early stages of the project. An effort was made to record all information associated to each occurrence, following an established schema, as described by Pando (1991). Besides scientific name, date and locality details, habitat – including host – is provided for the 78% of the records.

These works were mostly carried out by a small team of data entry technicians and supervisors, with the support of the Project’s scientific team.

Primary data produced and compiled as part of the research conducted within the project, by researchers involved in the project, were also incorporated into the database. These come from studied specimens held in herbaria or fields surveys carried out within the framework of the project. These data, when relevant, were published as the “Bases Corológicas series in “Cuadernos de Trabajo” (14 volumes published between 1991 and 2008. See References). No species have been retrieved from molecular data.

Quality control description: Quality control and assurance comprises a number of procedures, references and tools along the data life cycle. These can be summarized as follows:

- A data-entry manual on what to capture and what not and how to capture the information was developed and published (Pando 1991).
- International standards approved or endorsed by the “Biodiversity Information Standards/TDWG). Specifically, the following were used: Brummitt and Powell (1992), Lawrence et al. (1968), and TL-2 (Stafleu and Cowan 1976–1985, Stafleu and Mennega 1992–2000, Dorr and Nicolson 2008, 2009).
- When herbarium information was available, this was recorded under column “otherCatalogNumbers” following “Index Herbariorum” (Thiers 2016) standard abbreviations.
- Additionally, the following works were used as a reference for taxonomic names: Farr et al. (1979a, b, c, 1986), and the Dictionary of the Fungi (Hawksworth et al. 1983, 1995, Kirk et al. 2001, 2008), Saccardo (1882–1931).
- The database management system used (Pando et al. 2004) had many of these standards build-in as dictionaries and controlled vocabularies.
- All newly entered records were checked against the actual publications by the supervisors as part of the database work flow.
- A final check was done by the “Bases Corológicas” authors and editors as part of the publication process.
- Dataset publication in the GBIF network include data transformation to comply with the Darwin Core specification (Wieczorek et al. 2012) and further validation procedures (geographic coordinate format, coordinates within country/provincial boundaries, absence of ASCII anomalous characters in the dataset) using DARWIN_TEST (v3.3) software (Ortega-Maqueda and Pando 2008).
Host information as well as geographic coordinates have been taken from the sources. Obvious errors and typos have been corrected, but no in-depth interpretation of these details nor retrospective georeferencing was carried out. This approach guarantees fidelity to the sources, but also results in some unavoidable heterogeneity in the information made available.

Datasets

Dataset description

Object name: Darwin Core Archive Flora Mycologica Iberica Project fungi occurrence dataset
Character encoding: UTF-8
Format name: Darwin Core Archive format
Format version: 1.0
Distribution: http://www.gbif.es/ipt/archive.do?r=floramicologicaiberica
Publication date of data: 2016-02-02
Language: Spanish
Records: 59,235, of which 21,393 are georeferenced.
Licences of use: This work is licensed under a Creative Commons Attribution (CC-BY) 4.0 License (http://creativecommons.org/licenses/by/4.0/legalcode).

External datasets

Dataset description

Object name: Flora Mycologica Iberica
Character encoding: UTF-8
Format name: Darwin Core Archive
Format version: 0
Distribution: http://www.gbif.es
Metadata language: English
Date of metadata creation: 2009-11-30
Hierarchy level: Dataset

Acknowledgments

We thank colleagues and collaborators that have been involved with the FMI project at several stages, especially we would like to mention here: Juan Carlos Hernandez, Nieves Hidalgo, Jorge Portela, Isabel Ronai, Roberto López, Félix Díaz, Sara Pérez, Ana Pradal, and Mª Jesús Corregidor.
We also thank the agencies that supported the dataset presented here: Ministry of Science and Education (Spain), and National Institute of Employment of Spain (INEM).

References


Brummitt RK, Powell CE (Ed.) (1992) Authors of plant names. A list of authors of scientific names of plants, with recommended standard form of their names including abbreviations. Royal Botanic Gardens, Kew, Richmond.


Checa J (1997b) Annotated list of the Lophiostomataceae and Mytilinidiaceae (Dothideales, Ascomycotina) reported from the Iberian Peninsula and Balearic Islands. Mycotaxon 63: 467–491.


The Flora Mycologica Iberica Project fungi occurrence dataset


Telleria MT (1990) Annotated list of the Corticiaceae, sensu lato (Aphyllophorales, Basidiomycotina), for Peninsular Spain and Balearic Islands 135: 1–152.