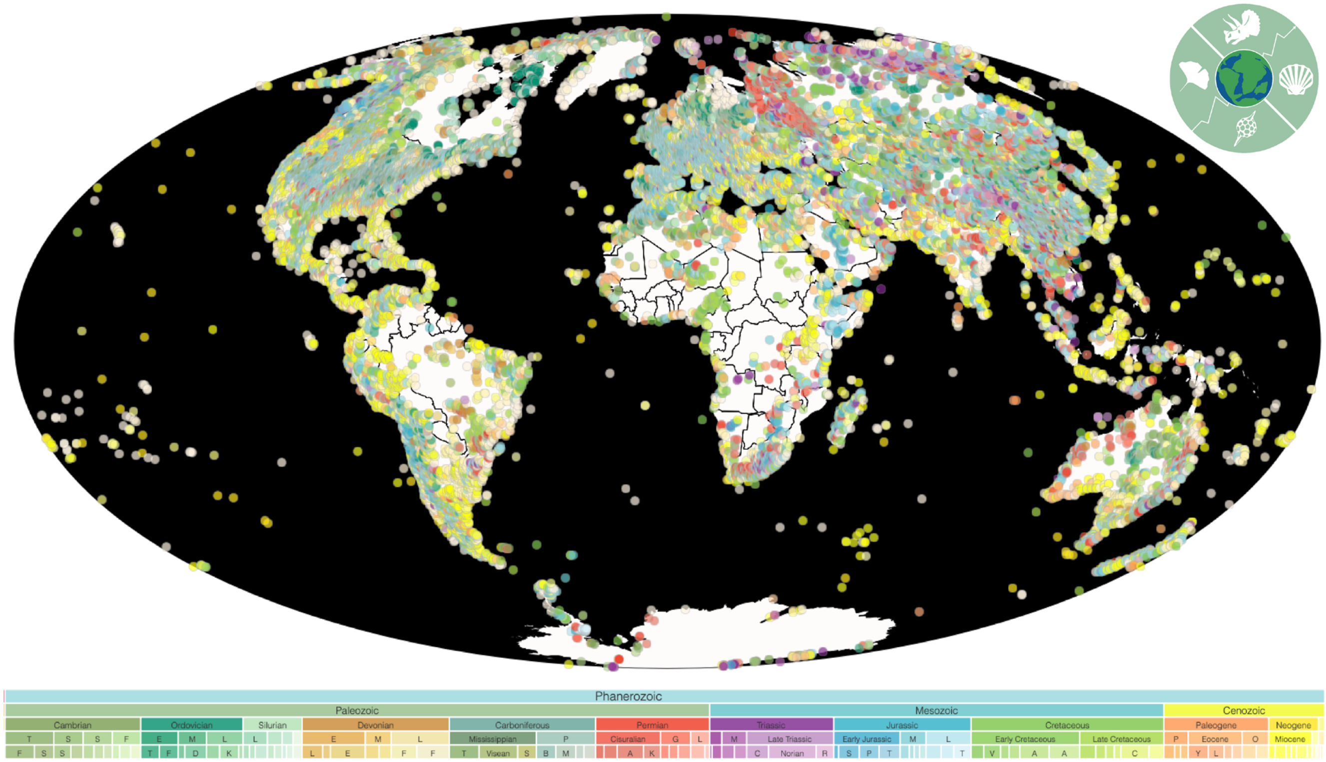
*PaleoBios* 40(11): 1-56, August 2023

# *PaleoBios*

OFFICIAL PUBLICATION OF THE UNIVERSITY OF CALIFORNIA MUSEUM OF PALEONTOLOGY



# Mark D. UHEN, Bethany ALLEN, Noushin BEHBOUDI, Matthew E. CLAPHAM, Emma DUNNE, Austin HENDY, Patricia A. HOLROYD, Melanie HOPKINS, Philip MANNION, Phil NOVACK-GOTTSHALL, Catalina PIMIENTO, and Peter WAGNER (2023). Paleobiology Database User Guide Version 1.0

**Cover:** Paleobiology Database interface with localities color-coded by geologic age.

**Citation:** Uhen, M.D., B. Allen, N. Behboudi, M. E. Clapham, E. Dunne, A. Hendy, P. A. Holroyd, M. Hopkins, P. Mannion, P. Novack-Gottshall, C. Pimiento, and P. Wagner. 2023. Paleobiology Database User Guide Version 1.0. *PaleoBios* 40(11): 1-56.

**DOI:** https://doi.org/10.5070/P9401160531

# Paleobiology Database User Guide Version 1.0

Mark D. Uhen[1], Bethany Allen[2], Noushin Behboudi[1], Matthew E. Clapham[3], Emma Dunne[4], Austin Hendy[5],
Patricia A. Holroyd[6], Melanie Hopkins[7], Philip Mannion[8], Phil Novack-Gottshall[9], Catalina Pimiento[10-12],
and Peter Wagner[13]

[1]Department of Atmospheric, Oceanic, and Earth Sciences, George Mason University, Fairfax, VA 22030 USA; <muhen@gmu.edu>
[2]Department of Biosystems Science and Engineering, ETH Zurich, Basel; Computational Evolution Group, Swiss Institute of Bioinformatics, Lausanne; <bethany.allen@bsse.ethz.ch>
[3]Department of Earth and Planetary Sciences, University of California, Santa Cruz, CA 95064 USA; <mclapham@ucsc.edu>
[4]GeoZentrum Nordbayern, Department of Geography and Geosciences, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), 91054 Erlangern, Germany; <dunne.emma.m@gmail.com>
[5]Natural History Museum of Los Angeles County, Los Angeles, CA 90007 USA; <ahendy@nhm.org>
[6]Museum of Paleontology, University of California, Berkeley CA 94720 USA; <pholroyd@berkeley.edu>
[7]Division of Paleontology, American Museum of Natural History, New York, NY 10024; <mhopkins@amnh.org>
[8]Department of Earth Sciences, University College London, London, WC1E 6BT, UK; <philipdmannion@gmail.com>
[9]Biological Sciences, Benedictine University, Lisle, IL 60532 USA; <pnovack-gottshall@ben.edu>
[10]Paleontological Institute and Museum, University of Zurich, 8006 Zurich, Switzerland; <catalina.pimientohernandez@pim.uzh.ch>
[11]Department of Biosciences, Swansea University, SA28PP Swansea, UK
[12]Smithsonian Tropical Research Institute, Box 207 Balboa, Panama
[13]School of Biological Sciences, University of Nebraska, Lincoln, NB 68182 USA; <peterjwagner@unl.edu>

The Paleobiology Database is an online, non-governmental, non-profit public resource for paleontological data. It is organized and operated by a multi-disciplinary, multi-institutional, international group of paleobiological researchers. This volume is designed to be a comprehensive guide for Paleobiology Database users, both General and Contributory. It covers most database uses from data retrieval and mapping to data contribution of all types. It contains numerous examples to illustrate database use as well as definitions of terms and additional links to numerous other sources. We hope that this user guide will help all users access the great volume of data in the Paleobiology Database and lead others to start and continue to add data to the system.

## 1. Introduction

### 1.1 What is the Paleobiology Database?

The **Paleobiology Database** (PBDB, paleobiodb.org) is an online, non-governmental, non-profit public resource for paleontological data. It is organized and operated by a multi-disciplinary, multi-institutional, international group of paleobiological researchers. Its purpose is to provide global, collection-based occurrence and taxonomic data for organisms of all geological ages, as well as data services to facilitate access to data for independent development of analytical tools, visualization software, and applications of all types. The PBDB's broader goal is to encourage and enable data-driven collaborative efforts that address large-scale paleobiological questions (Uhen et al. 2013).

**Citation:** Uhen, M.D., B. Allen, N. Behboudi, M. E. Clapham, E. Dunne, A. Hendy, P. A. Holroyd, M. Hopkins, P. Mannion, P. Novack-Gottshall, C. Pimiento, and P. Wagner. 2023. Paleobiology Database User Guide Version 1.0. *PaleoBios* 40(11): 1-56.

**DOI:** https://doi.org/10.5070/P9401160531

### 1.2 History of the Paleobiology Database

The history of the Paleobiology Database began in the US-based National Center for Ecological Analysis and Synthesis (NCEAS)-funded Phanerozoic Marine Paleofaunal Database initiative, which operated from August 1998 through August 2000.

The original working group included: John Alroy, Richard Bambach, Karl Flessa, Mike Foote, Steven Holland, Scott Lidgard, David Jablonski, Charles Marshall, Michael McKinney, Arnold Miller, Mark Patzkowsky, David Raup, Kaustuv Roy, Jack Sepkoski, and Peter Wagner.

The PBDB was supported from 2000 to 2008 and from 2010 to 2013 by the US National Science Foundation (NSF) and has continued to receive funding from various sources within the NSF over the years. It has also received funding from the Australian Research Council and the Research Data Alliance. In addition, many science-driven projects have been funded by grants to individuals and research groups around the world, primarily to enter data into the PBDB for research purposes. The PBDB has also benefited from science education grants from the NSF to develop systems to leverage data in the system for educational purposes. Ongoing funding information can be found here.

John Alroy was particularly instrumental in the development of the PBDB in several different capacities. John was one of the original founders of the PBDB (along with Charles Marshall) and served as the Database Administrator (later renamed Database Director) from 2000 to September 2013.

### 1.3 Organizational Structure

An **Executive Committee (ExCom)** runs the Paleobiology Database. It consists of 12 people from the paleobiology community and spans a wide range of career stages, expertise, and countries of origin and backgrounds. These individuals serve 3-year terms and one third of the ExCom is up for rotation or re-election every year. The Authorizer level users (see 1.8 *PBDB User types*) of the PBDB are asked to vote on those who wish to join the ExCom. The ExCom is led by a Chair and a Secretary who are chosen from the members of the ExCom. The Executive Committee approves requests for Authorizer user status, provides assistance to new authorizers and data enterers, and also sets policies for the PBDB.

### 1.4 Database Structure

The PBDB is a relational database that includes dozens of tables to hold data of many different types. These data fall into four primary classes: bibliographic, taxonomic, geologic & geographic, and occurrences. These data types, along with their interrelationships, are shown in

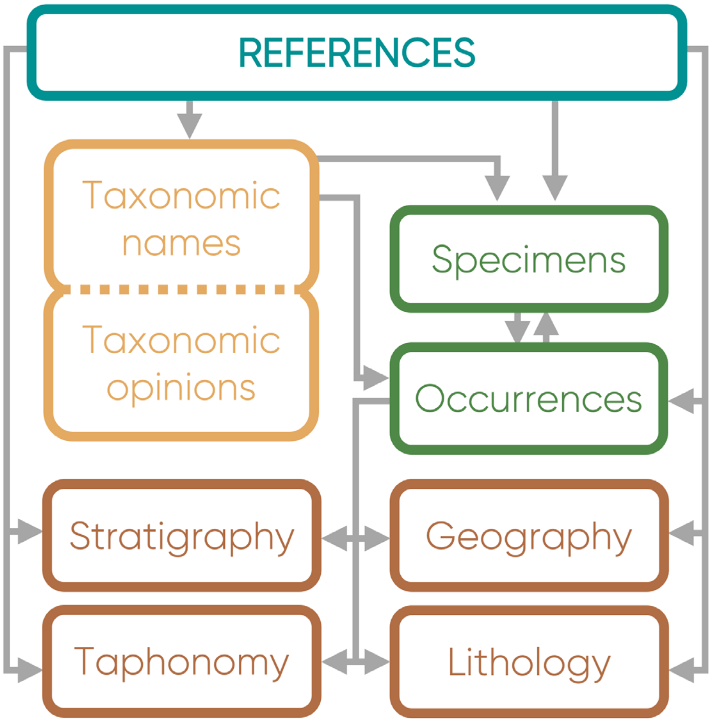


**Figure 1.** Basic data types in the PBDB and their interrelationships

Figure 1. All taxonomic, geologic, geographic and occurrence data in the PBDB are linked back to References (bibliographic data), the sources from which these data were derived. In this way, any perceived inconsistencies or errors in the data can be checked by going back to the References.

In addition to these primary data types, the PBDB also includes information describing the collection process (i.e., who, when, purpose, repository information, etc.), taphonomic information (i.e., size classes, degree of disarticulation, preservational mode, etc.), and metadata associated with the data entry process itself (i.e., who entered or modified the data, data entry/modification date and time, etc.) to track contributions to the system.

### 1.5 Database Policies

The PBDB policies as described here apply at the time of the drafting of this User Guide. For the most up to date PBDB policies, see Database Policies under *About* on the Front Page of the PBDB website.

### 1.5.1 Data Integrity

If you think you have identified an error in the PBDB, the first step is to make sure that it is indeed an error. There are several reasons why information might be inaccurate or appear to be inaccurate, including transcription errors during data entry, inaccurate or outdated information in the original Reference, or professional disagreement in the literature over taxonomy, stratigraphy, age, or other attributes.

Data in the PBDB are primarily entered from References by people, who occasionally make transcription errors. Please refer back to the original source Reference to check if the data were accurately entered into the PBDB or if it is indeed a transcription error upon entry into the PBDB. Transcription errors can be corrected by any Contributory User.

If the data were accurately entered from the Reference in question, it is NOT considered an error, even though you may be aware of newer information or you disagree with the information as presented in the Reference. PBDB contributors try to update records when newer information becomes available, but the scale of the database makes this challenging. Also, please remember that Collections may have data derived from multiple References. There is always a Primary Reference, but Collections can have an unlimited number of Secondary References from which data can be entered. Information that was accurate at the time of the Primary Reference, but that is now outdated, can be updated in the database from newer data in Secondary References. It is always best practice to make a note in the Comments section of the original data and which Secondary reference is the source of the change. Even if the data in the Reference are truly erroneous, we still enter it into the PBDB to accurately reflect what was stated by the author(s) of the Reference. Our practice is to compile all data and opinions as stated in the References.

Disagreements among different References (e.g., different stratigraphic age assignments or identification of specimens in taxonomic occurrence lists) can also be documented by entering additional information from other references into the PBDB later. Again, for Collections, this additional information comes from Secondary References that are added to the Collection record. Taxonomic disagreements among different References (for example, synonymies or assignments to higher level taxonomy) are documented by entering additional taxonomic opinions from other References. Taxonomic names and opinions are dynamically linked to taxonomic occurrence lists but not modified by information entered into Collections records. When there is unresolved disagreement in the literature, the PBDB will typically reflect the most recently published work.

### 1.5.2 Data Errors and Updates

If a genuine data entry error has occurred, any Contributory User can update the data entry to correct the error. Any Contributory User can also add Secondary References to update information in the database. For General Users, your first point of contact should be the person who entered the data. The best way to find out who this is, is to click on the link for the publication that the data were entered from (e.g., the "Osborn 1905" link next to *Tyrannosaurus rex* on this page: https://paleobiodb.org/classic/basicTaxonInfo?taxon_no=54833). Following the reference, you will see 2 or 3 names, e.g., [M. Carrano/M. Carrano/M. Carrano]. The first name is the original Authorizer of the entry, the second name is the Enterer of the original data (these are the same person when the data were entered by the Authorizer). The third name, if present, is the last person who most recently edited the data. To try to get a data entry error fixed, please contact the Authorizer first (the first name), as they are ultimately responsible for the entry. The names and affiliations of Authorizers are listed on our website. If you have trouble contacting the person who authorized the data entry, you can also send a message to info@paleobiodb.org.

### 1.5.3 Bug Reporting

The PBDB interface wraps around a complex set of software, and as such, there are undoubtedly errors in the code and new ones will inevitably be introduced as new code is written. We track these errors or bugs in the code via the PBDB GitHub site. You must be a GitHub user to use this system and be logged in to GitHub. If you find any problems with the website, Navigator, or the Application Programming Interface (API), please let us know:

Report Website Bugs
Report Navigator Bugs
Report API Bugs

Before you add a new bug report, it is a good idea to look through the list of previously reported bugs to see if the one you discovered is already on the list. You can add additional information to previously reported bugs, which can be helpful in tracking them down and correcting them. If you cannot use GitHub to report bugs, please email a detailed description of the bug to info@paleobiodb.org. Be sure to include all relevant information like Reference number, Collection number, Taxon ID number, etc.

### 1.5.4 Paleobiology Database License

PBDB data and visualizations are available to the public based on a Creative Commons license that requires no permission nor attribution for use (a CC-0 international License).

### 1.5.5 Data Embargo

Data embargoes can be applied to PBDB collections and occurrences by the Enterer for up to five years while working on a project. This data will be irretrievable to General Users for the duration of the embargo (access to other Contributory Users can also be limited, if preferred). No other data types may be embargoed.

### 1.5.6 PBDB Citation

You can cite the PBDB in several ways. First, you can cite a specific Data Archive that you created or used (see *Data Archives* below). Second, you can refer to a specific download from the PBDB, giving the date and parameters of the download. Here is an example:

The data were downloaded from the Paleobiology Database on May 3, 2022, with the following parameters: time intervals = Eocene through Oligocene, continent = Europe, environment = marine, taxon = Cetacea.

Alternatively, one could simply give the URL created by the PBDB API to retrieve the data, along with the date:

https://paleobiodb.org/data1.2/occs/list.csv?datainfo&rowcount&base_name=Cetacea&interval=Eocene,Oligocene&cc=EUR&envtype=marine Downloaded from the Paleobiology Database on 3 May, 2022.

It is important to include the date of data download because the PBDB is a dynamic system, with data being added every day. Thus, including the download date in your citation allows others to effectively compare your data download with others. Furthermore, because it is possible to download only data that was entered or modified before a specific date, reporting the date of download ensures replicability.

It is recommended to also cite or acknowledge the people who contributed the bulk of the data to the dataset used. We also strongly encourage that you create a secondary bibliography for work that uses large data sets. You can easily download all of the references that contributed to a dataset for this secondary bibliography. Many journals will accept a secondary data bibliography as a supplementary material file, and this type of citation helps ensure that the hard work performed by members of our community collecting and describing fossil organisms gets acknowledged properly (Payne et al. 2012).

To get a list of all of the people who have entered data in your data set, be sure to check enterer names under Choose output options. This will allow you assess who has entered the bulk of the data in your data set and allow you to give credit to acknowledge them appropriately.

## 1.6 Data Standards

### 1.6.1 Paleogeographic Coordinates

Collections in the PBDB are assigned present-day geographic coordinates (latitudes and longitudes) and geological time intervals. These two pieces of information are combined with paleogeographic rotation models to obtain reconstructed paleogeographic positions at the time of deposition of the rocks of a given Collection.

We use two different plate rotation models. The first are the rotations from the PALEOMAP project (Scotese et al. 1979; Scotese and Sager 1988), provided to us by Christopher Scotese. The second source is GPlates (Wright et al. 2013). GPlates is the default rotation model used in the newer parts of the PBDB, including in the Navigator application, whereas the Scotese rotations are shown in older parts of the PBDB, such as the pages that display Collection data. GPlates rotation models are obtained by dynamic calls to web data services managed by the GPlates team, and so will continue to evolve as the GPlates model evolves. GPlates plate ID numbers can be found here. Paleocoordinates for both rotations are stored in the PBDB for all collections and are available for download.

### 1.6.2 Geologic Time Scale

Age data can be entered into the PBDB using one of the many time scales included in the PBDB. Currently, most downloaded data are then binned using the Gradstein 2012 geologic time scale (Gradstein et al. 2012). Jurassic and Cretaceous substages (e.g., late Bajocian, early Turonian) use the 2004 geologic time scale (Gradstein et al. 2004), and some other intervals use different timescales from different published papers (e.g., selected Devonian substages from Becker et al. 2004). In some instances, the age of boundaries shifted between the earlier time scales and 2012, affecting binning for searches and downloads depending on the time rule selected. See section 2.2.2 Download for more information on time rules, and exercise caution to ensure the download contains all of the desired data.

## 1.7 Data Archives

Data Archives are thematic data sets created by many contributory users. If you have entered or uploaded a coherent set of data (for example, a particular taxonomic group, time period, or geographic area) and you are ready to publish your study, you can create a Data Archive by checking the “Choose Archive Options” option on the Download Records page. Be sure to include a description of the scope of the data set, and a list of contributors, along with a title for the data set. PBDB will then assign a DOI and post an entry on our Data Archives list that will reference the data set. If you have any questions on this procedure, email info@paleobiodb.org.

## 1.8 PBDB User Types

### 1.8.1 General Users

The PBDB is an open access resource that can be accessed by anyone in the world via the Internet. **General Users** can explore the available data via the Navigator and download data. The PBDB also has a special user type called **Guest Users**, which allows them to contribute Resources to the system. These Resources help people make use of the PBDB, and includes video tutorials, web & phone apps, and lesson plans. The only thing that General Users cannot do is contribute data.

### 1.8.2 Contributory Users

The PBDB has three user types that are allowed to contribute data to the database. These are outlined here from the user type with the most responsibility, to the least responsibility.

#### Authorizers

**Authorizer** status is open to all paleontological experts who agree to abide by PBDB policies and contribute data. Authorizer status is approved by the Executive Committee. Most Authorizers have doctoral degrees; however experienced researchers without graduate-level training also may be considered. If you don't fit these criteria but want to contribute, consider entering data under the direction of another PBDB Authorizer. See our list of Authorizers for contact information and get in touch with them directly. If you do wish to join as an Authorizer, please go here and select *Create Guest Account*. Fill in the appropriate fields to first obtain a Guest account, and then follow the instructions to request the level of access appropriate to your needs.

#### Enterers

**Enterer** status is primarily for those working with Authorizers to enter data into the PBDB. Authorizers can add Enterers to the system without approval from the Executive Committee. Enterer status is often used for graduate students working with an advisor who is an Authorizer. Enterers can enter any data like an Authorizer, but their Authorizer is expected to supervise this data entry and edit their work freely.

#### Students

**Student** status is given by Authorizers to students (usually undergraduates) who are working on class projects or research projects using the PBDB. Student users are prohibited from entering taxonomic data but are free to enter any other types of data, including occurrence data. Students can be added to the system by any Authorizer without approval of the Executive Committee. Again, Authorizers are expected to supervise and edit this Student data entry.

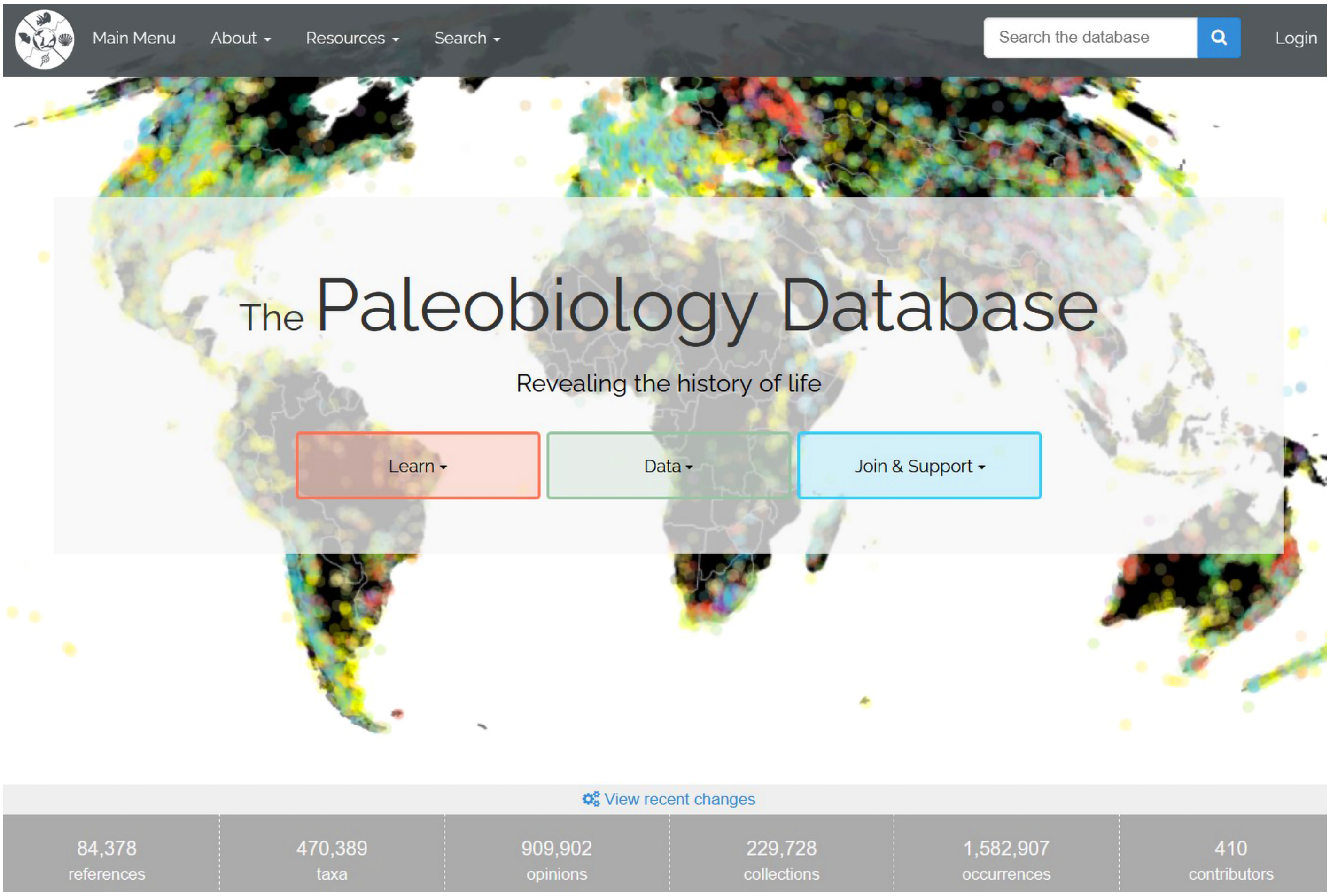
## 1.9 About this User Guide

This User Guide was drafted by the Executive Committee of the Paleobiology Database under the leadership of Mark D. Uhen. Some of the text was derived from the tip sheets that have been present on various data entry pages of the PBDB, originally written by John Alroy. Some of the tip sheets regarding geology and stratigraphy were also written by Steve Holland. The plant organ tip sheet was drafted by the original Paleobotany Working Group. We thank them for their early work to help guide user data entry. Our goal in drafting this guide is to help facilitate clear data entry for both novice and long-time users of the PBDB. We also hope to update and re-issue volumes of this guide as the PBDB continues to grow and evolve. If you have any questions regarding what you read here or the PBDB itself, please email info@paleobiodb.org for more information.

## 2. General User Guide

### 2.1 The Front Page of the PBDB

The Front Page of the Paleobiology Database is where you first land when you visit the PBDB URL, paleobiodb.org. From here, visitors to the site can navigate to all of the publicly accessible features of the database. Let's explore the Front Page shown in Figure 2. Across the top of the page is a dark gray (or sometimes black) navigation bar. In the upper left-hand corner, you can see the *PBDB logo*. This logo will appear here on virtually all PBDB pages. If you ever want to return to the Front Page, just click this logo. In the upper right-hand corner is the *Login* button. The *Login* button is for Contributory users, and these actions are covered below in the **Contributory User Guide** section 3. To the left of Login, you can see a **Quick Search** field. Here you can enter a quick search term and the PBDB will try to match it to a taxon name, stratigraphic unit name, or the name of a collection. Any potential matches will be displayed as you type in a drop-down menu, and you can click on them as they are displayed. If you don't see what you are looking for, see the **Search** section (2.2.1) for more advanced search options. (Note that the index in the Quick Search is rebuilt at the end of each day so will not include information just added, whereas Search searches in real time.) At the bottom of the Front Page are some real-time statistics about how much data have been entered into the PBDB. Additional features of the Front Page are discussed below.



| 84,378 | 470,389 | 909,902 | 229,728 | 1,582,907 | 410 |
|---|---|---|---|---|---|
| references | taxa | opinions | collections | occurrences | contributors |

**Figure 2.** The Front Page of the Paleobiology Database.

### 2.2 Main Menu

Clicking on **Main Menu** in the navigation bar takes you to a text-based interface that presents all of the actions

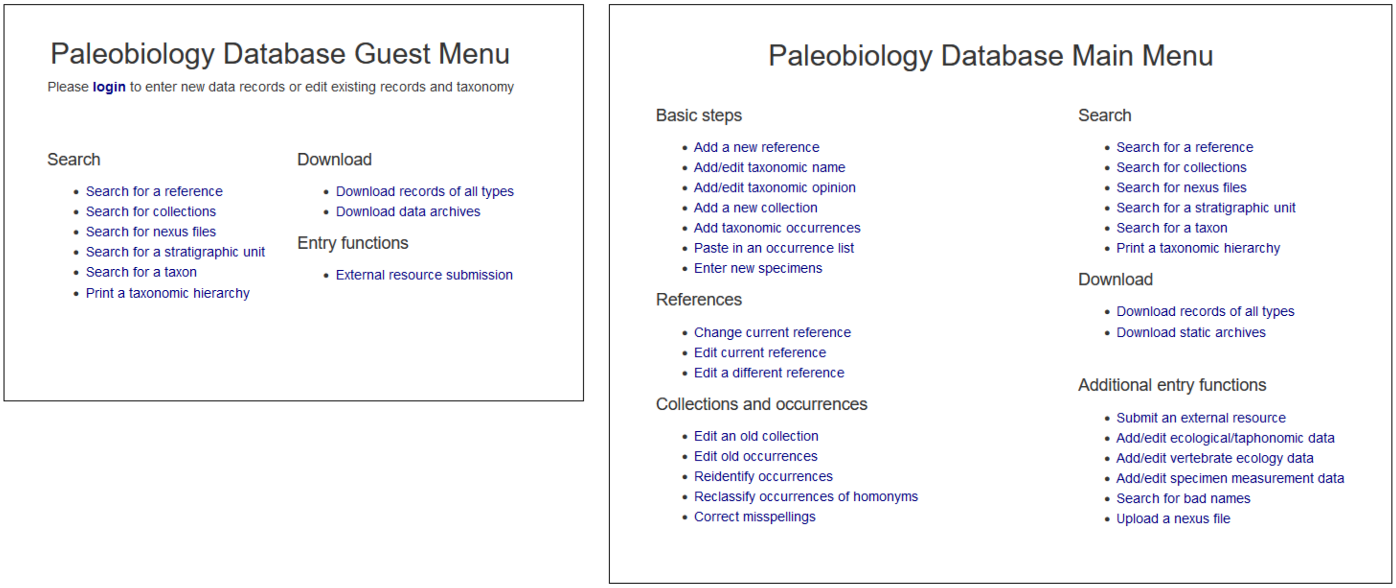


**Figure 3.** The differing functionality of the Main Menu for General (left) and Contributory (right) Users.

available for General Users (see Figure 3). These actions are grouped into three categories: Search, Download, and Entry functions. If logged in as a Contributory User, this is also where you can access a variety of data entry options.

### 2.2.1 Search

You can search for data using the various search forms listed under **Search** on the Main Menu. You can search for a reference, collection, NEXUS file (used for storing phylogenetic data), stratigraphic unit, or taxon. Just click on the search function you want, enter your search parameters, and click Submit. The PBDB will return a list of any matches for your search parameters.

The option that is different from the other search options is *Print a taxonomic hierarchy*. If you click this, you are asked for search parameters. The simplest one is a **taxon name**. Enter any taxon name, and the PBDB will show you any taxa classified within that name, given the current prevailing taxonomic opinions. For instance, if you type Basilosauridae, it will show all of the genera in that family (including junior synonyms, *nomen dubia*, and other invalid names). Clicking show all expands the hierarchy to show all of the species in those genera. In this view, all taxon names are clickable for more information. Hide all reverses the action of the show all button.

### 2.2.2 Download

The **Download** functions allow you to extract data from the PBDB for any use. You can either *Download records of all types* or *Download data archives*.

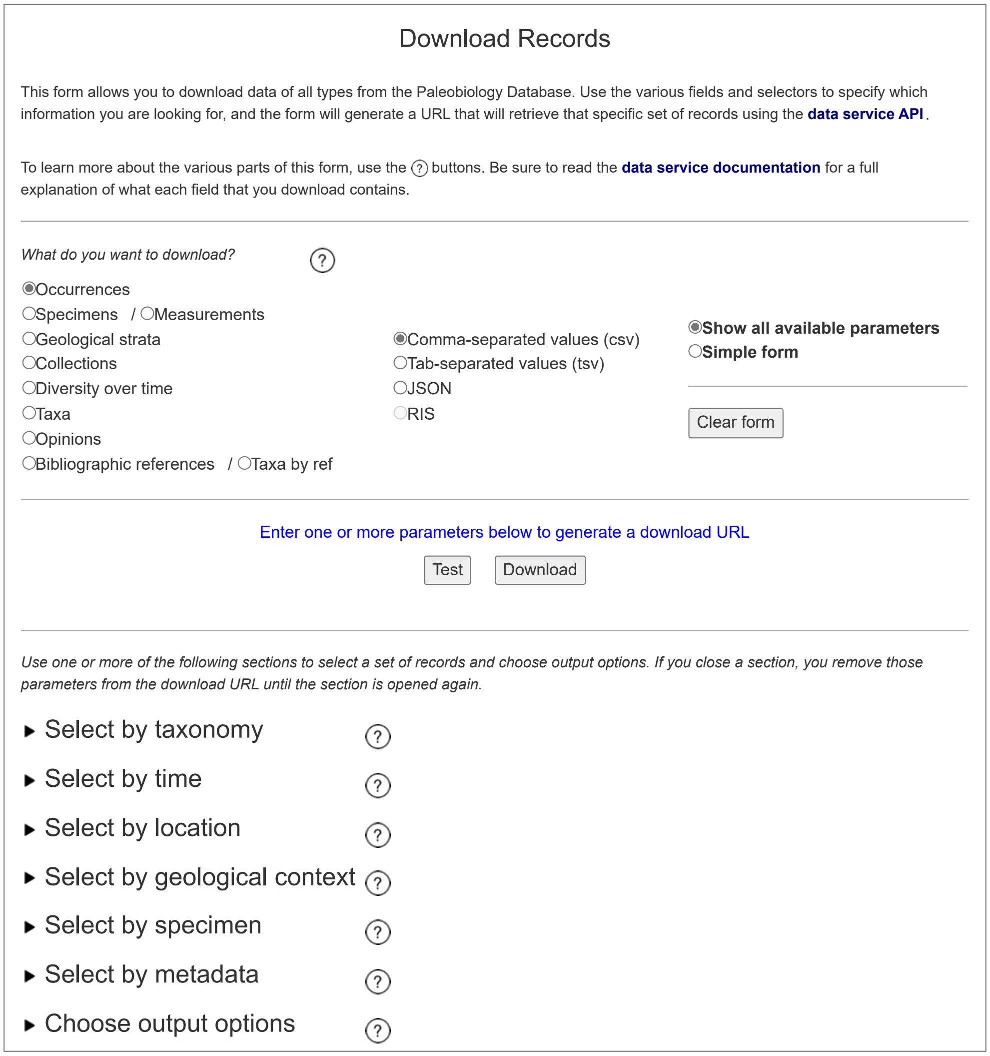
### Download records of all types

*Download records of all types* allows users to extract virtually any and all data from the PBDB using a variety of search parameters. There are many types of records one can download, and many parameters one can use to delimit the data for download. Because this activity can be complicated, let's first look at an overview of the **Download Records** form, and then explore each of the sections one at a time. See Figure 4 for an overview of the form.

The first part of the forms helps you to determine what type of data you want to download and what format (file type) the download in. Some formats are only available for certain data types. For instance, RIS format is only for bibliographic data. Here we will explore the complete form (all available parameters), but there is also an option to show a simpler version of this form, which restricts some of your options. On this form, you will see several question marks in circles. Clicking these reveals more information on the parameter that the symbol is next to, which can be very helpful if you are uncertain what to choose.

*Data Type*

Your first choice is what type of data you want to download. Your choices include *occurrences*, *specimens*, *measurements*, *geological strata*, *collections*, *diversity over time*, *taxa*, *opinions*, *bibliographic references*, and *taxa by ref*(erence). Your intended purpose for the data determines which type of data you want to download,



Download Records

This form allows you to download data of all types from the Paleobiology Database. Use the various fields and selectors to specify which information you are looking for, and the form will generate a URL that will retrieve that specific set of records using the **data service API**.

To learn more about the various parts of this form, use the ? buttons. Be sure to read the **data service documentation** for a full explanation of what each field that you download contains.

*What do you want to download?* ?

● Occurrences
○ Specimens / ○ Measurements
○ Geological strata
○ Collections
○ Diversity over time
○ Taxa
○ Opinions
○ Bibliographic references / ○ Taxa by ref

● Comma-separated values (csv)
○ Tab-separated values (tsv)
○ JSON
○ RIS

● **Show all available parameters**
○ **Simple form**

Enter one or more parameters below to generate a download URL

*Use one or more of the following sections to select a set of records and choose output options. If you close a section, you remove those parameters from the download URL until the section is opened again.*

▸ Select by taxonomy ?
▸ Select by time ?
▸ Select by location ?
▸ Select by geological context ?
▸ Select by specimen ?
▸ Select by metadata ?
▸ Choose output options ?

**Figure 4.** PBDB download form

so give it some thought before you proceed. Often, you will want to download multiple types of data based on the same search parameters, which you can do by performing one download, then selecting a different data type and performing another download without changing any search parameters. These search parameters are discussed below. Be aware that the options you can specify for the “Select by” Data Download Parameters (see below) will sometimes vary according to the data type selected at the top of the download page. (For example, there are 15 taxonomic rank categories you can choose among when downloading taxa, but only six when downloading occurrences.)

## Data Format

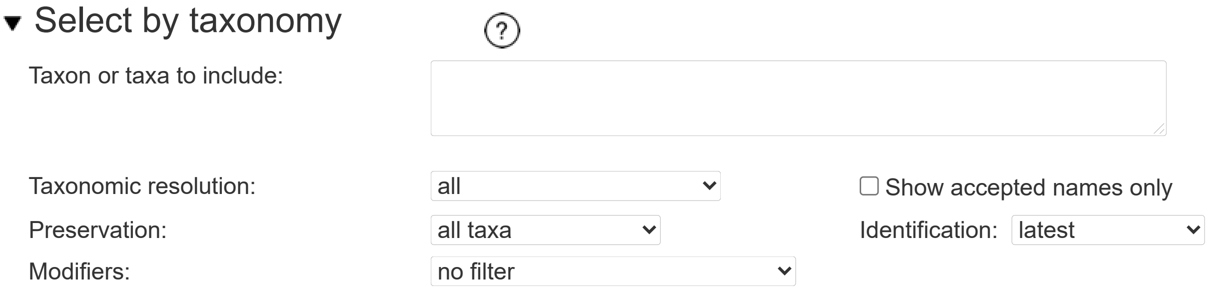
You can select from three basic data formats for your download, plus one more exclusively for bibliographic data. These three are: *comma-separated* values (csv), *tab-separated values* (tsv), and *JSON* (JavaScript Object Notation). CSV and TSV are good formats if you plan to use the data in a spreadsheet or statistical software. JSON is better if you plan to send the data to another software tool that can interpret the JSON format. Text fields are output with UTF-8 encoding, so letters with diacritical marks and other special characters may not appear correctly in spreadsheets or in objects in R unless you specify the encoding yourself.

*RIS* (Research Information Systems) is a format for bibliographic data and is used by bibliographic reference management software. If you plan to publish a data bibliography for your PBDB data, you can import this file directly into your reference management software to help craft a well-formatted bibliography.

## Data Download Parameters

As you select different parameters, notice that a URL is generated in real-time (the blue text in Fig. 4). This URL is how the Application Programming Interface (API) directly interfaces with the database infrastructure in order to download your data. Advanced users and third-party apps (such as Rockd and the R packages paleobioDB and divDyn) can directly download desired data bypassing this download page. (See the PBDB Data Service Documentation for examples and the correct format to download using the URL alone.)

**Select by taxonomy** allows you to limit your download by taxon and specify the taxonomic resolution of the downloaded data. See Figure 5. For instance, if you type "Dinosauria" (without quotation marks) in the *Taxon or taxa to include* box, your download will include all taxa in the clade Dinosauria. You can also exclude parts of a taxonomic group using the hat, or caret, character. So, let's say you want Dinosauria, but don't want to include Aves, you would enter "Dinosauria^Aves". This will give you all taxa that are included in Dinosauria that are not included in Aves (i.e., non-avian dinosaurs).



▾ Select by taxonomy

Taxon or taxa to include:

Taxonomic resolution: all

☐ Show accepted names only

Preservation: all taxa

Identification: latest

Modifiers: no filter

**Figure 5.** Taxonomy download parameters (when downloading Occurrences).

If you want multiple taxa in a download, you can string multiple taxonomic names together separated by commas. For instance, entering "Sirenia, Cetacea, Pinnipedimorpha" will download the data for each of these groups into one file.

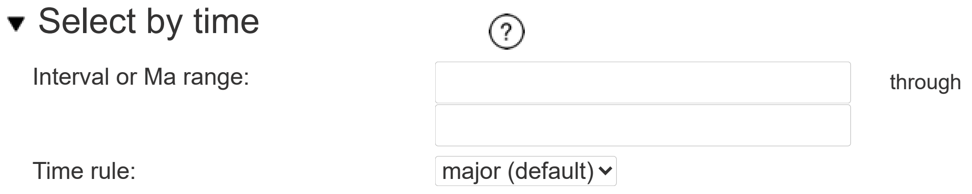
Beyond this basic delimitation by taxon name, you can use this form to specify much more about the data to be returned. The *Taxonomic resolution* field allows you to specify the taxonomic level at which the data are returned. Using the pull-down, you can select *species*, *genus*, or *family*. For example, if *genus* is selected when downloading occurrence data, then the download will include all occurrences identified at genus or species level, but no occurrences identified to coarser (less precise) taxonomic levels. If downloading taxa, then selecting *genera* will download only genus names.

The *Preservation* field will allow you to restrict your download to *regular taxa*, *form taxa*, *ichnotaxa*, *form & ichnotaxa*, or *all* preservation forms. It is helpful to select *regular taxa* if you want to exclude trace fossils like footprints or bite marks, or *all* if you want to include them. (See 3.6 *A note on trace fossils (ichnotaxa) and form taxa* below for recommendations when handling form taxa and ichnotaxa.) See the Glossary for definitions of preservation forms.

The *Modifiers* field allows you to include or exclude occurrences that are marked with aff., cf., ?, etc. which indicate uncertainty of varying forms in identifications of taxa. This is helpful for determining whether you want to specify only those occurrences that have been confidently assigned to your taxon of interest or consider all occurrences that could be allied to the taxon. Older occurrences that use "wastebasket" genus names, such as *Nautilus* sp. for a nautiloid or *Terebratula* sp. for a terebratulid brachiopod, have often been entered with a question mark or in quotation marks, so it is typically good to exclude uncertain genus occurrences and/or vet the data following download.

The *Identification* field allows you to specify *latest* (default), *original*, *reidentifications*, or *all*. Change this from the default value (*latest*) if you want to retrieve data about any fossil that has ever been or has previously been identified as your taxon of interest. Checking the *Show accepted names only* will provide occurrences with their updated taxonomic identifications (according to the opinions currently accepted by the PBDB) as opposed to the obsolete names under which they were originally entered.

**Select by time** (Figure 6) allows you to restrict your data download to a single interval of geologic time. You can enter the name of a single geologic interval or specify a range by providing both a minimum and maximum interval to include. For instance, if you enter Paleocene through Oligocene, all occurrences in either of those



▾ Select by time

Interval or Ma range: through

Time rule: major (default)

**Figure 6.** Temporal download parameters (when downloading Occurrences).

time intervals will be included, as well as in the Eocene between them, as per the *Time rule*. You can also specify a time interval by entering a number of millions of years in these fields as well.

The *Time rule* parameter lets the PBDB determine whether or not to include a particular occurrence in a time interval or not. Selecting a different time rule will significantly affect the download, so this parameter should be adjusted with care. The **contain** time rule is the strictest. It selects occurrences whose temporal range is entirely contained in the specified interval. The **major** time rule (default) is more permissive and selects occurrences for which a majority (over 50%) of their temporal range overlaps the specified interval. The **buffer** time rule selects occurrences whose temporal range overlaps the specified interval and is contained inside a buffer zone around the interval. The default buffer is 5 Ma outside of the interval for the Cenozoic, 12 Ma for the earlier eras. The **overlap** time rule is the most permissive and selects all occurrences whose temporal range overlaps the specified interval by any amount.

You may be asking yourself, why does PBDB need this *Time rule* algorithm? The most common situation in which this is necessary is when collections are designated to a large time interval which may only be partially included in your search interval, either due to considerable uncertainty in age or due to the compilation of fossil occurrences across multiple time intervals. In addition,

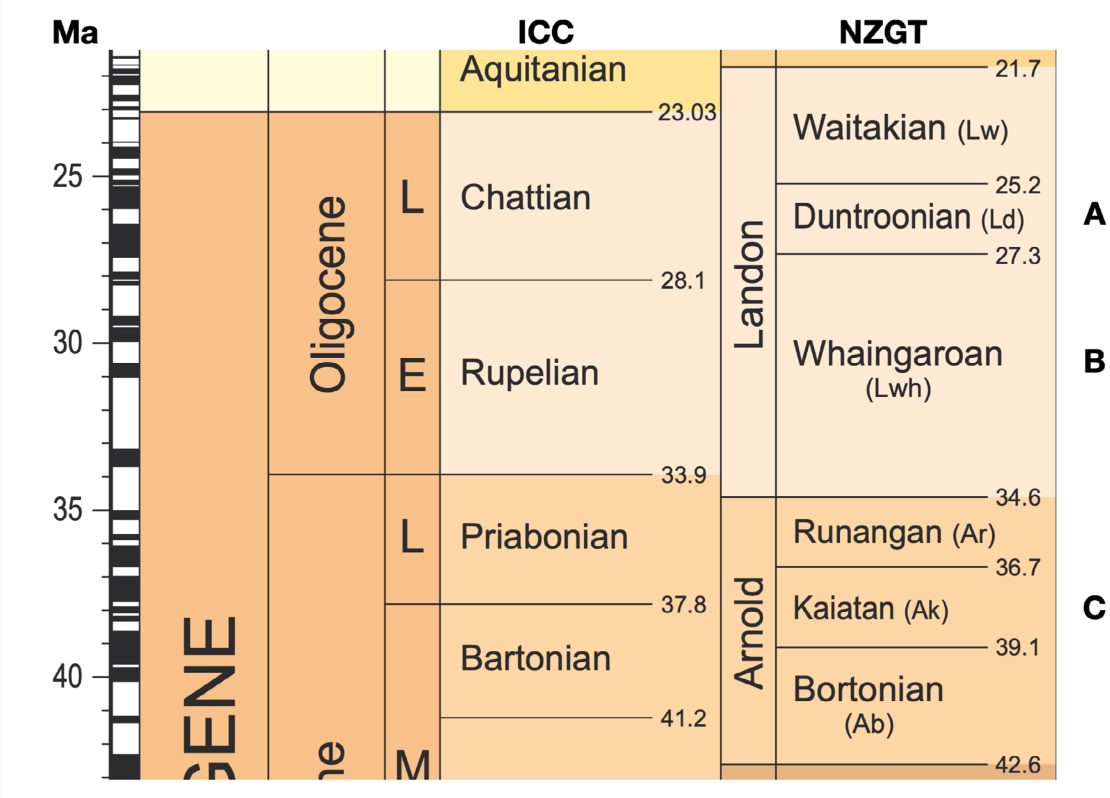
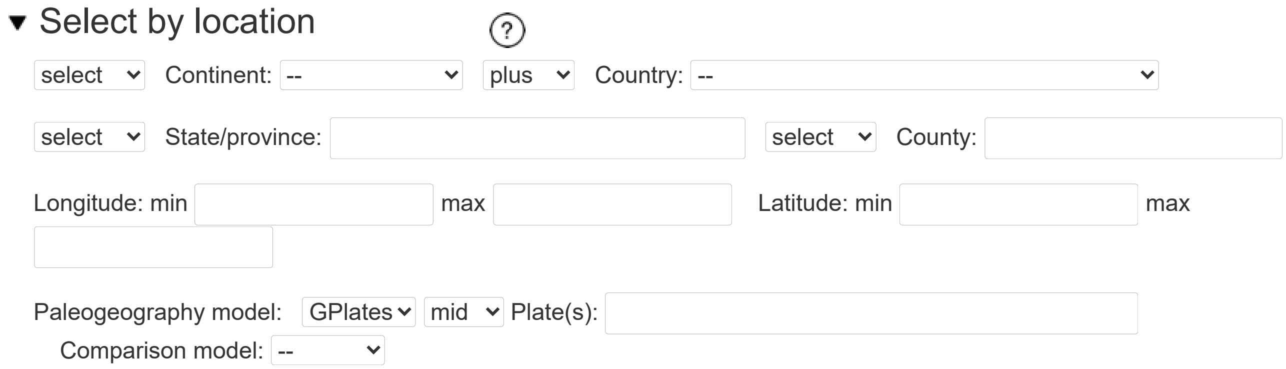


**Figure 7.** Examples of how the PBDB time rules assign occurrences to time bins. This timescale is abstracted from the New Zealand Geological Timescale (Raine et al. 2015). See the text for a full explanation of these cases.

the PBDB allows users to enter data from different time scales and needs to try to make these data compatible with one another for translation into a single, unified time scale. Note that the PBDB will only output the assignment to a global stage by checking the time binning box in the output options.

Figure 7 illustrates how differences in the *Time rule* work in practice. This shows a segment of the Cenozoic time scale from New Zealand (Raine et al. 2015), compared to the global geologic time scale (Cohen et al. 2013). If one were using PBDB to do any kind of global analysis during the Cenozoic, one would probably use the global stages to group by time, because the majority of fossil Occurrences (grouped into PBDB Collections) are reported using this time scale. However, Occurrences from the Cenozoic of New Zealand are often reported using these New Zealand stage names, and if we want to include them as well, we need to specify a *Time rule* to help us translate those data to the global stage scale. Let's consider an Occurrence at A, in the Duntroonian. The Duntroonian is entirely within the temporal bounds of the Chattian, so all of the time rules would count Occurrences reported as Duntroonian as being in the Chattian. Now consider an Occurrence from B, from somewhere within the Whaingaroan. The Whaingaroan is NOT entirely contained within the Rupelian, and the fossil might instead be Chattian or Priabonian, so the **contain** rule would not include this Occurrence, because you cannot be sure the Occurrence is Rupelian. The majority of the Whaingaroan overlaps the Rupelian, so the **major** time rule and all of the other time rules would include it in the Rupelian. In addition, the **overlap** time rule would count the occurrence in the Chattian and Priabonian as well, because the Whaingaroan overlaps all three global stages. Finally, consider an Occurrence in the Kaiatan at C. The majority of the Kaiatan is in the Bartonian, so the **major** time rule would assign it to the Bartonian. The **buffer** time rule could assign it to both the Bartonian and the Priabonian, depending on the size of the buffer, and the **overlap** time rule would assign it to both stages. In both B and C, the **contain** time rule would not count the Occurrence in any stage because neither the Whaingaroan nor the Kaiatan are contained within a global stage.


▼ Select by location ?

select Continent: -- plus Country: --

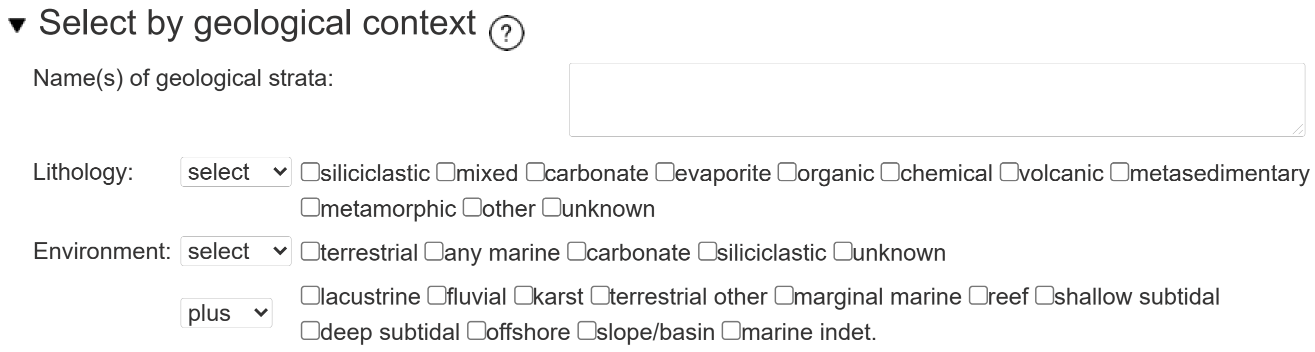
select State/province: select County:

Longitude: min max Latitude: min max

Paleogeography model: GPlates mid Plate(s):

Comparison model: --

**Figure 8.** Geographic location download parameters (when downloading Occurrences).


▼ Select by geological context ?

Name(s) of geological strata:

Lithology: select ☐siliciclastic ☐mixed ☐carbonate ☐evaporite ☐organic ☐chemical ☐volcanic ☐metasedimentary ☐metamorphic ☐other ☐unknown

Environment: select ☐terrestrial ☐any marine ☐carbonate ☐siliciclastic ☐unknown

plus ☐lacustrine ☐fluvial ☐karst ☐terrestrial other ☐marginal marine ☐reef ☐shallow subtidal ☐deep subtidal ☐offshore ☐slope/basin ☐marine indet.

**Figure 9.** Geologic context download parameters (when downloading Occurrences).

**Select by location** (Figure 8) allows you to specify data for download based on modern geographic location. You can specify by *continent*, *country*, *state/province*, and *county*. You can also choose a *tectonic plate* (either GPlates or Scotese model) by entering the corresponding plate number. Finally, you can specify a rectangular bounding box using *min & max longitude and min & max latitude*. If no selections are made, the default downloads global data.


▾ Select by specimen ?

Abundance: no filter Minimum:

**Figure 10.** Specimen abundance download parameters (when downloading Occurrences).

**Select by geologic context** (Figure 9) allows you to specify one or more named stratigraphic units. You can also include or exclude stratigraphic units by *lithology* or by *environment*. You can use the check boxes to mix and match varieties of environments and lithologies. It is worth noting that environments apply to the collection information, and not necessarily the taxon; for example, a terrestrial mammal preserved in lacustrine sediments will be included in a "lacustrine" dataset.

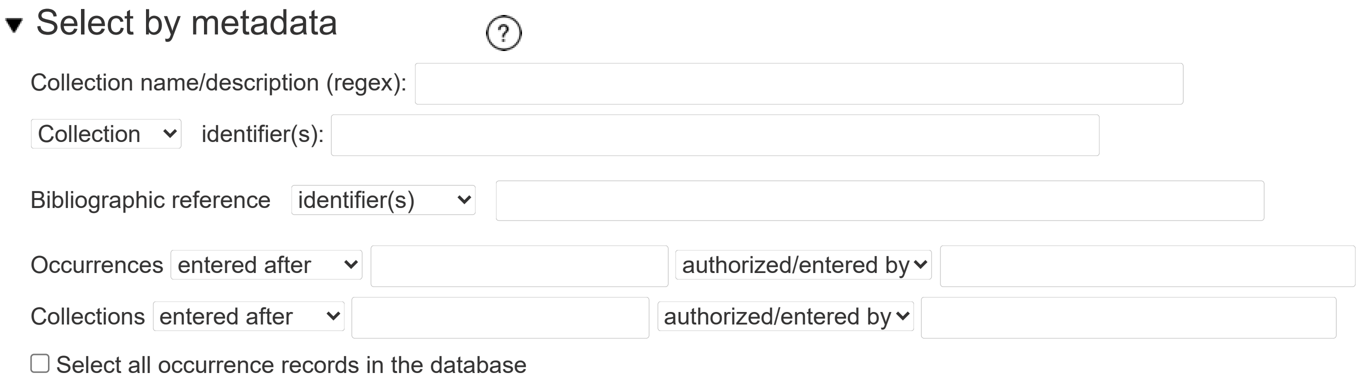
**Select by specimen** (Figure 10) allows you to specify occurrences based on *abundance* of specimens. You can choose count, coverage, or any abundance information. Not all occurrences include abundance data, so be aware that specifying a minimum abundance will exclude any below that minimum value as well as those lacking abundance data. As a result, the output will contain some occurrences from a collection but will omit other occurrences from that collection.

**Select by metadata** (Figure 11) allows you to specify a data return based on *metadata*, or data about the data. You can specify a variety of parameters including bibliographic reference, dates entered, and Contributory User who entered the data. This can be helpful to track your own data entry, or data entry for those who work with you, including your students. It also allows you to compare how the quantity of uploaded data has changed over time.

## Output Options

**Output Options** (Figure 12) allows you to select exactly which data fields you want included in your download. (Note that the output fields will change depending on the type of data being downloaded.) You can use the check boxes to select individual fields (or sets of related fields), or you can select the most frequently used ones by checking *Include all output blocks whose names are boldfaced below*. We **strongly recommend** checking *include all output blocks whose names are boldfaced below* because if you forget to check one of the checkboxes, it could be very difficult to go back and re-create your download to get that one field you missed. It's better to download all of the fields and ignore the ones you are not interested in. You can also use the pulldown menus to determine the order in which your records are listed in the download. You can *limit the number of records*, but this is not recommended because you will be excluding some records that match your download parameters.

Checking *Include metadata at the beginning of the output* appends rows to the start of your download file that specifies all the search parameters you chose, as well as the date, time, direct API URL, and other metadata for your download. This is very useful for a record of what you downloaded but will cause trouble if you intend to import your data directly into subsequent software that expects the first row to include column names.


▾ Select by metadata ?

Collection name/description (regex):

Collection identifier(s):

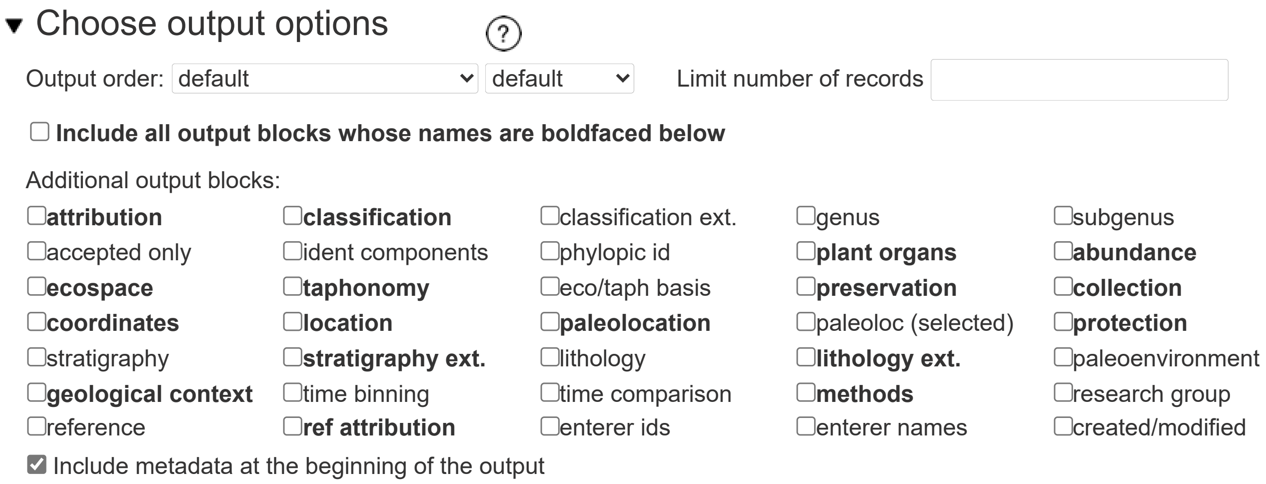
Bibliographic reference identifier(s)

Occurrences entered after authorized/entered by

Collections entered after authorized/entered by

☐ Select all occurrence records in the database

**Figure 11.** Metadata download parameters (when downloading Occurrences).



Choose output options

Output order: default default Limit number of records

☐ **Include all output blocks whose names are boldfaced below**

Additional output blocks:

☐**attribution**
☐accepted only
☐**ecospace**
☐**coordinates**
☐stratigraphy
☐**geological context**
☐reference
☐**classification**
☐ident components
☐**taphonomy**
☐**location**
☐**stratigraphy ext.**
☐time binning
☐**ref attribution**
☐classification ext.
☐phylopic id
☐eco/taph basis
☐**paleolocation**
☐lithology
☐time comparison
☐enterer ids
☐genus
☐**plant organs**
☐**preservation**
☐paleoloc (selected)
☐**lithology ext.**
☐**methods**
☐enterer names
☐subgenus
☐**abundance**
☐**collection**
☐**protection**
☐paleoenvironment
☐research group
☐created/modified

☑ Include metadata at the beginning of the output

**Figure 12.** PBDB download field options (when downloading Occurrences).

### Data archives

**Choose archive options** allows you to create a citable data archive for a particular data download. We strongly recommend that you use this option judiciously, and only create a data archive when you are sure the data download that you are crafting is exactly what you want to archive (such as when you intend to use the data in a publication that others may want to replicate).

When you are ready to create a Data Archive, prepare your download using all of the appropriate parameters outlined in section 2.2.2, Download. Then, choose a short but descriptive title for your archive, list the authors for the data set (those who have contributed effort and you want to identify as creators of the Data Archive), and a brief description of what is in the Data Archive.

When you Download after filling in these options, the PBDB will create a draft archive that will then be approved by a member of the PBDB ExCom. They will acquire a DOI for the data archive and post it to the PBDB Data Archive site. You can then use this DOI in publications as a permanent repository for your data.

### 2.2.3 Download results

The result of most downloads is a tabular data set in CSV or TSV format. The tabular data comes in two blocks. The first block contains metadata about the data set itself (if selected during the download process). For more in-depth explanations of the data fields returned, see the PBDB Documentation that explains all of the column headers in the data download. Users can also choose to download a JSON file, or RIS if the data are Bibliographic references, depending on the choice of the user as noted above.

### 2.2.4 Entry functions

General Users must register for a Guest account to have access to the *External resource submission function*. This allows PBDB to attribute authorship to submitted resources. If you click on *External resource submission* and you do not have a guest user account, PBDB directs you to the Account page where you can register for a Guest account. Just fill in the fields, choose a password, and the PBDB will guide you to the Resource submission page. Here you can enter some metadata about your resource, including a title, description, a link to the resource itself, an icon to represent your resource, type of resource, intended audience, author information, and keywords. Once you submit your resource, a PBDB volunteer will approve your submission and post it to the Resource page. Any resource that deals with the PBDB is welcome, including videos, lesson plans, phone, and web apps, etc.

## 2.3 About pages

The **About** pages can be accessed via the navigation bar at the top of the Front Page of the PBDB. These include information regarding the PBDB, much of which is covered in more detail in this guide but is likely to be more up to date. Clicking on About reveals a pull-down menu with the following options: *FAQ (Frequently Asked Questions), Recent Changes, People, Funding, Project Development, Official Publications, Partner Organizations, Database Policies, Grant Writing Assistance,* and *Contact Us*. Each of these is described below.

### 2.3.1 FAQ (Frequently Asked Questions)

*FAQ* or Frequently Asked Questions includes the questions that we get asked often, by all types of users, and answers to those questions. These get updated from time to time as users ask different questions and as we update the system.

### 2.3.2 Recent Changes

*Recent Changes* outlines recent features added to the PBDB and changes to how older features work.

### 2.3.3 People

*People* lists the current members of both the Executive Committee and the Tech Team working on database programming.

### 2.3.4 Funding

*Funding* lists current and past grants that support and have supported the PBDB over the years.

### 2.3.5 Project Development

*Project Development* shows some statistics regarding the PBDB and how those statistics have changed over the past 20 years.

### 2.3.6 Official Publications

Researchers using PBDB data in their publications can request an official PBDB publication number to include in their publications. *Official Publications* includes a list of these official publications, and instructions on how to request an official publication number. This page also includes a link to the PBDB Google Scholar page, which lists all of the publications that mention the PBDB, whether they are “official” or not.

### 2.3.7 Partner Organizations

PBDB partners with other organizations on various database and data accessibility efforts. These are listed under *Partner Organizations*.

### 2.3.8 Database Policies

*Database Policies* includes the policies that govern the operation of the PBDB. These are listed in this guide under 1.5 Database Policies, but the most up-to-date policies will be listed on this page.

### 2.3.9 Grant Writing Assistance

Researchers who want to write a grant to fund research projects using the PBDB are encouraged to read the *Grant Writing Assistance* page. This will help researchers writing their grant, particularly with the Data Management Plan. It also guides researchers through how to include funds to help support the PBDB if using the PBDB for their data repository.

### 2.3.10 Contact Us

*Contact Us* includes email addresses to contact the PBDB for any additional questions or concerns. The most general address for questions is info@paleobiodb.org.

## 2.4 Resources

Clicking on **Resources** in the navigation bar opens a pull-down menu to various resources outside of the PBDB, that are either about the PBDB itself or use PBDB data in some form. The list of resource types includes *Web Apps, Mobile Apps, Lesson Plans and Activities, Introductory Tutorials, Data Entry Tutorials, R Packages and Scripts, Presentations, API Resources,* and *Other Resources*. All registered user types can upload metadata regarding resources that they want to contribute by clicking on Contribute Resources. The Video Resources are an excellent source of information regarding using the database, and many of these are hosted at the PBDB YouTube channel. Take some time to explore these if you are struggling with data entry or download. In addition, there are many Resources that help with data analysis and data exploration.

## 2.5 Search

The **Search** pull-down in the navigation bar on the Front Page takes you directly to the Search functions outlined in section 2.2, Main Menu.

## 2.6 Navigator

The PBDB **Navigator** (Peters 2014) can be accessed by clicking on Revealing the history of life or the Visualize option in the green Data drop-down menu on the Front Page. Navigator is a powerful, map-based tool that can be used to visualize and explore data in the PBDB. Figure 13 shows the landing page when you first open Navigator. Navigator has five primary parts: a search field in the upper right, a toolbox on the left, the map in the center, the taxon display on the right, and the geologic time scale on the bottom.

### 2.6.1 Navigator Search

In the **Navigator Search field**, there are the four types of data that you can enter into the gray search box in the upper right to display on the map: *time, taxa, authorizer, stratigraphy*. When you type anything in this field, the PBDB will display potential matches in a list below the search field. Once you see what you want to search for, you can click on that item in the list. Then, the PBDB will display a dot for each collection that matches those search criteria on the map. If you don’t see what you are looking for, check the spelling. If it is spelled correctly, and no matches appear, that data is not currently in the PBDB.

Let's explore each of these types of data you can search for in turn. Entering the name of a time period (or clicking on it in the timescale at the bottom) will call up all collections from that time interval. The taxon display changes to reflect the taxa in the time period that you searched for. If you double-click on a time interval it will enlarge and expand the time scale at the bottom of the page so you can see finer intervals of time.

Entering the name of a taxon will call up any matches on that name. For instance, if you type *Dorudon*, it will call up the three matches for names currently in use (*Dorudon*, *Dorudon atrox*, and *Dorudon serratus*), followed by names no longer in use (Dorudontidae, *Dorudon intermedius* etc.). Clicking on a name currently in use will show collections including that taxon on the map. Clicking on a name no longer in use will show collections including the synonyms of that taxon. See 3.5.3 *Add/edit a taxonomic opinion* below for more information on synonyms in the PBDB. The bottom of the taxon display shows the numbers of collections and occurrences represented on the map.

Entering the name of an Authorizer will show all collections authorized by that Authorizer on the map.

Entering the name of a stratigraphic unit will show all collections attributed to that stratigraphic unit on the map. These will be in a limited geographic area (unless that name has been used for multiple different geologic units).

Filters, including search terms, can be removed by clicking the 'x' next to their names in the bottom left.

### 2.6.2 Navigator Toolbox

The **Navigator Toolbox** allows you to alter the map, create other graphic representations of the data, and download the data on your map. First, the + and – buttons zoom in (+) and out (-) on the map. When you are zoomed in, you can click and drag the map to see different parts of the world. The icon showing South America and Africa together allows you to reconstruct the map of the world at different times in the past. The reconstructions are based on the GPlates model outlined under 1.6.1, *Paleogeographic Coordinates*. To reconstruct the world at a particular time, just click on a time period in the geologic time scale at the bottom of the window, then click the plate reconstruction icon. To return to the present-day map, just click the icon again. You can choose both taxa in the search field and a time period in the geologic time scale. You cannot use the zoom in and zoom out tools on a paleogeographic map.

The next tool in the toolbox is the taxon browser, shown as a bug icon. This tool allows you to type in the name of a taxon, and it will pull up a brief summary of what is included in that taxon.

The diversity curve tool is shown by the graph icon. It will show a basic diversity curve for the data shown on the map. If your map includes the entire world and you have no particular taxon selected, it will show you a global diversity curve for all of life. If you have zoomed in on the map, it will only include those taxa represented in collections shown on the map. You can show a diversity curve for a particular taxon by selecting that taxon in the search field first. This window also includes an advanced diversity curve generator with more options.

The download data tool is represented by a down arrow icon, and it allows you to download the data shown on the map. After you click on it you can choose what file format you want (CSV, TSV, JSON or RIS (references only)), and what type of data you want (occurrences, references, diversity summary, or full diversity data). These data types are more fully explained in the Download section of this guide. Once you have made your choices, click *Download*.

The last tool in the toolbox is the Show Examples tool, indicated with the pointed finger. This tool calls up three data sets you can use to explore the use of the Navigator. These include Cambrian Trilobites of Eastern China, Jurassic Dinosaurs of Western USA, and Permian Plants Paleogeography. Click on any of these buttons to explore these data sets further.

### 2.6.3 Navigator Map

The **Map** displays what is currently being searched for. Dots on the map are color coded to the geologic time scale, with the size of the dots scaled to the number collections represented by the dots. When you initially view the map, before you make any search choices, most of the dots are either light blue or black. This is because light blue represents the Phanerozoic, and those dots represent multiple collections from across the Phanerozoic. The black dots represent multiple collections from across the entire geologic time scale. These dots resolve into smaller dots, with more specific time periods, as you zoom in on the map. You can then click on any small dot to call up a summary of either the collections included in the dot, or a short summary of the collection information and the taxonomic occurrences in those collections. From these summaries, you can navigate directly to Collection records by clicking on the collection number or further

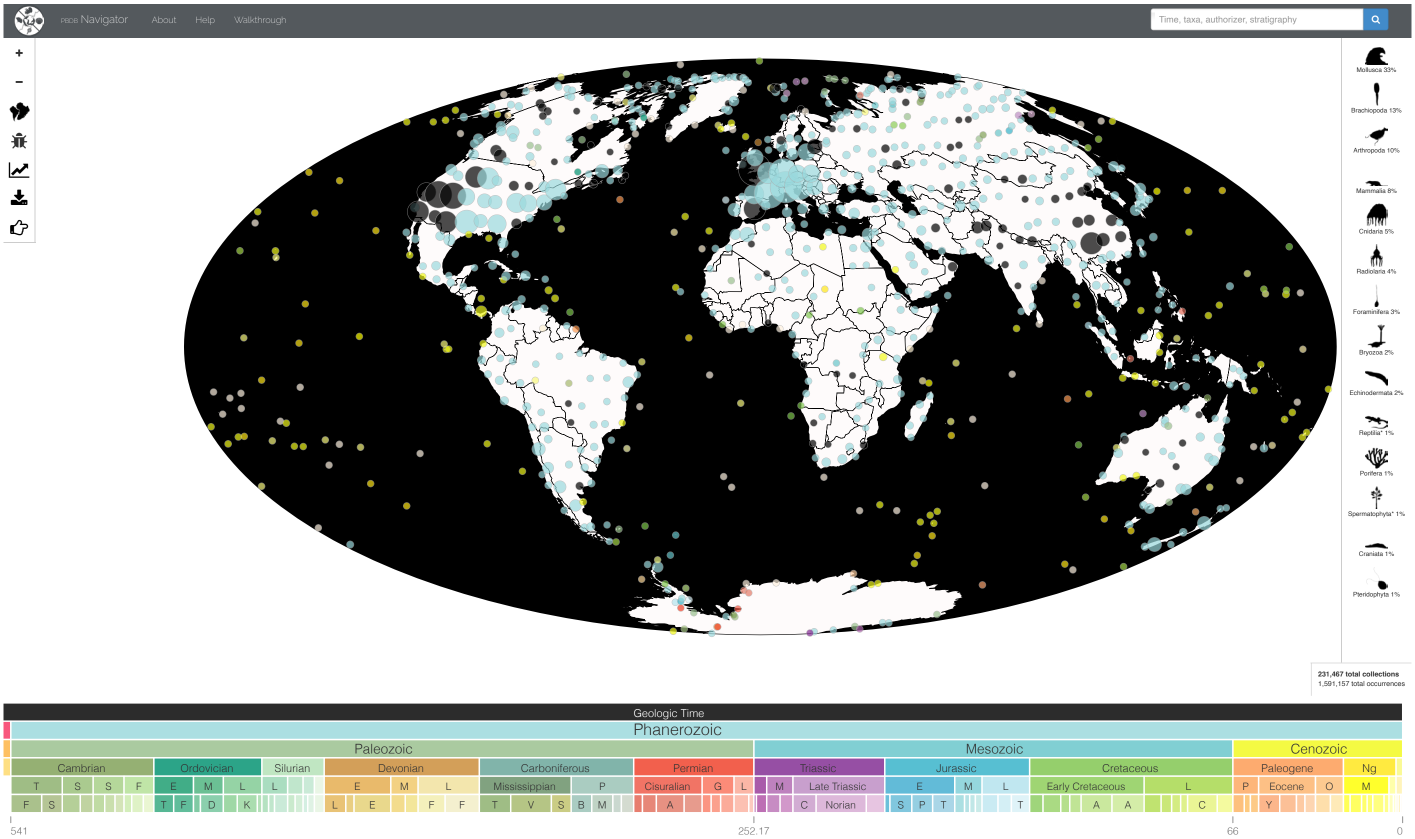


**Figure 13.** PBDB Navigator landing page.

filter the map results for specific taxa by clicking on taxon names in the occurrences window.

### 2.6.4 Navigator Taxon Display

The **Taxon Display** on the right always shows the seven most common subtaxa that are currently displayed on the map, from most common to least common, and tells you their relative frequencies. It changes every time what is shown on the map changes. Names marked with an asterisk * are broader categories that include one or more of the other listed names. Taxon silhouettes are sourced from PhyloPic.

### 2.6.5 Navigator Geologic Time Scale

The **Geologic Time Scale** is a graphic representation of geologic time, color coded to the official colors (according to the International Commission on Stratigraphy) for each segment of geologic time. You can expand and contract the geologic time scale by double clicking on the time period you want to fill up the display. For instance, if you want to expand the Cenozoic, double click on it, and it will expand to fill up the full width of the window and allow you to see subunits of the Cenozoic better. To return to the default display of the Phanerozoic, double click on Phanerozoic. You can display all of geologic time by double clicking on Geologic Time to see both the Phanerozoic and the pre-Cambrian.

The geologic time scale can also be used as a tool to display data from a particular segment of geologic time. Just single click on a time interval, and PBDB will show only collections from that time interval on the map. You can only display data from one interval at a time.

### 2.7 Quick Links in center of PBDB Front Page

In the center of the front page are three colored drop-down menus that provide quick links to the most commonly used functions of those described above.

### 2.7.1 Learn

The red **Learn** drop-down menu leads you to the Educational Resources also linked under Resources, as well as a page with news items from the PBDB.

### 2.7.2 Data

The green **Data** drop-down menu with the following items: Enter, Visualize, Download, and Browse. These take you directly to the Main Menu, Navigator, Download form, and Collections Search form, respectively. You will

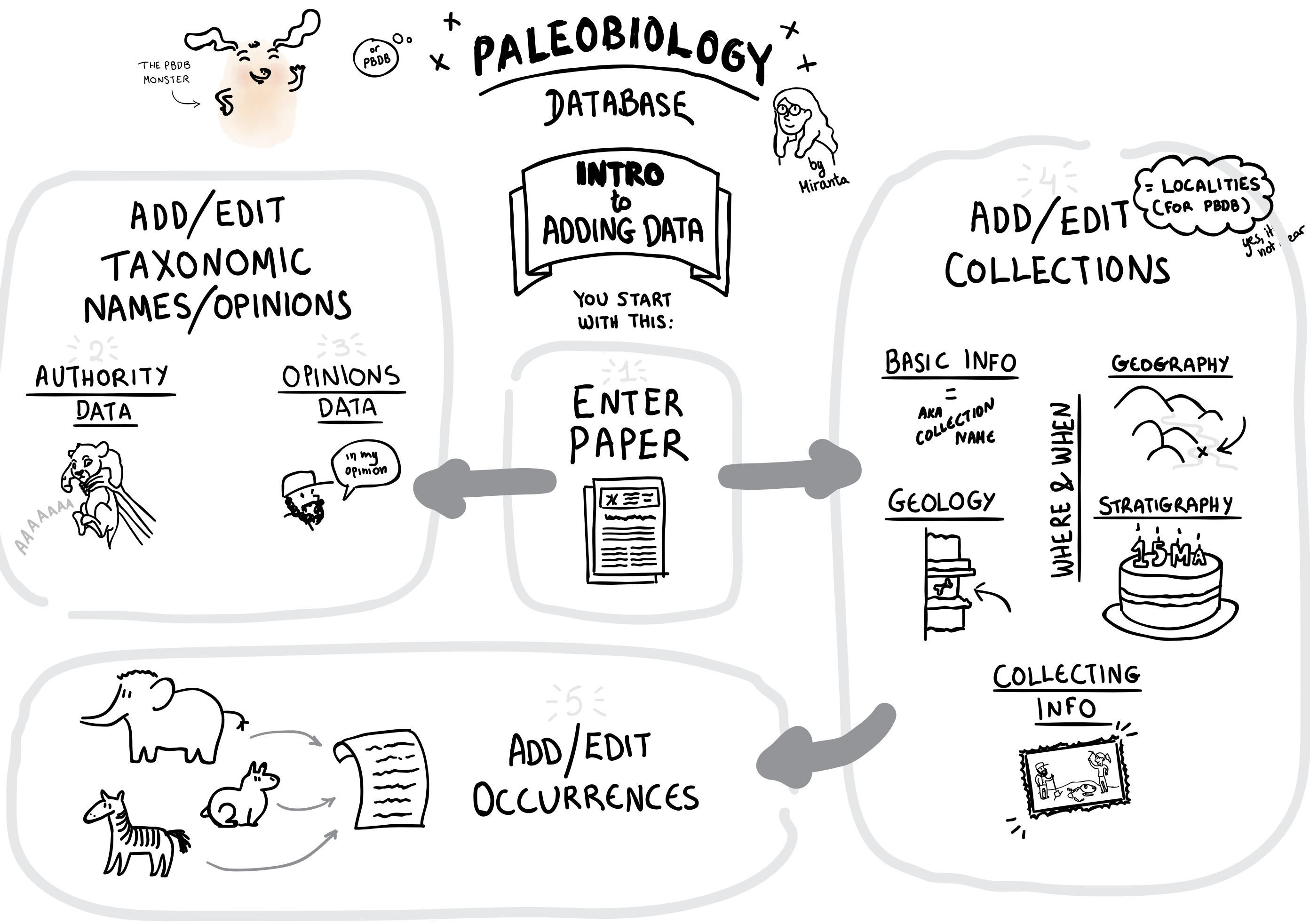


**Figure 14.** Comic showing common workflow for PBDB data-entry (Courtesy of Miranta Kouvari).

only be able to enter data from the Main Menu if you are logged in as a Contributory User.

### 2.7.3 Join & Support

The blue **Join & Support** drop-down menu includes three items: Join the PBDB, Support the PBDB and PBDB Merch. Join the PBDB allows you to obtain a Guest Account and/or upgrade your account to one of the Contributory User types. Support the PBDB allows you to make either a one-time donation or an ongoing donation to help PBDB maintain and grow the database for the future. PBDB Merch allows you to buy merchandise like t-shirts, coffee mugs, and tote bags with the PBDB logo. A portion of every purchase helps to support the PBDB as well. Be the first in your department to have a PBDB mug. All the cool kids are doing it!

## 3. Contributory User Guide

### 3.1 Introduction

Thank you for becoming part of the team of hundreds of researchers who enter data into the Paleobiology Database. It is through our combined efforts that we will eventually achieve the goal of databasing every published fossil occurrence on the planet. Despite the vast functionality of the PBDB, the basic workflow for uploading data is relatively straightforward (Figure 14). All of the following instructions apply to all contributory users (Authorizers, Enterers, and Students) except for the instructions regarding taxonomic data, which Student users are not permitted to enter.

Another tool to help you learn how to enter PBDB data is found in Appendix 2. This is a checklist that will help ensure that you enter all necessary and appropriate data from each Reference as you learn how to enter data. It also frames an efficient workflow which will be good to follow until you develop a data entry workflow of your own that works for you.

As you proceed, please note the following. In all of the data entry forms, blue text represents links that bring up additional information that could be helpful. Also, any field names in red indicate that these fields are required to be entered. Some data are entered with a pull-down menu, others with radio buttons or checkboxes. With radio buttons, you must pick one of the choices. With check boxes, you may pick one or several of the choices.

## 3.2 Account Settings

### 3.2.1 Your PBDB Account

The very top of the Main Menu page presents a black navigation bar that offers shortcuts to other items on the Main Menu page itself. In the upper right-hand corner, you should see your name with a pull-down arrow to the right. This area is where you can manage your PBDB account and how you interact with the PBDB.

**Account Settings**

Account settings are where you can update your name, home institution, email address, ORCID, etc. This is basically metadata about you, the PBDB User.

**Manage Resources**

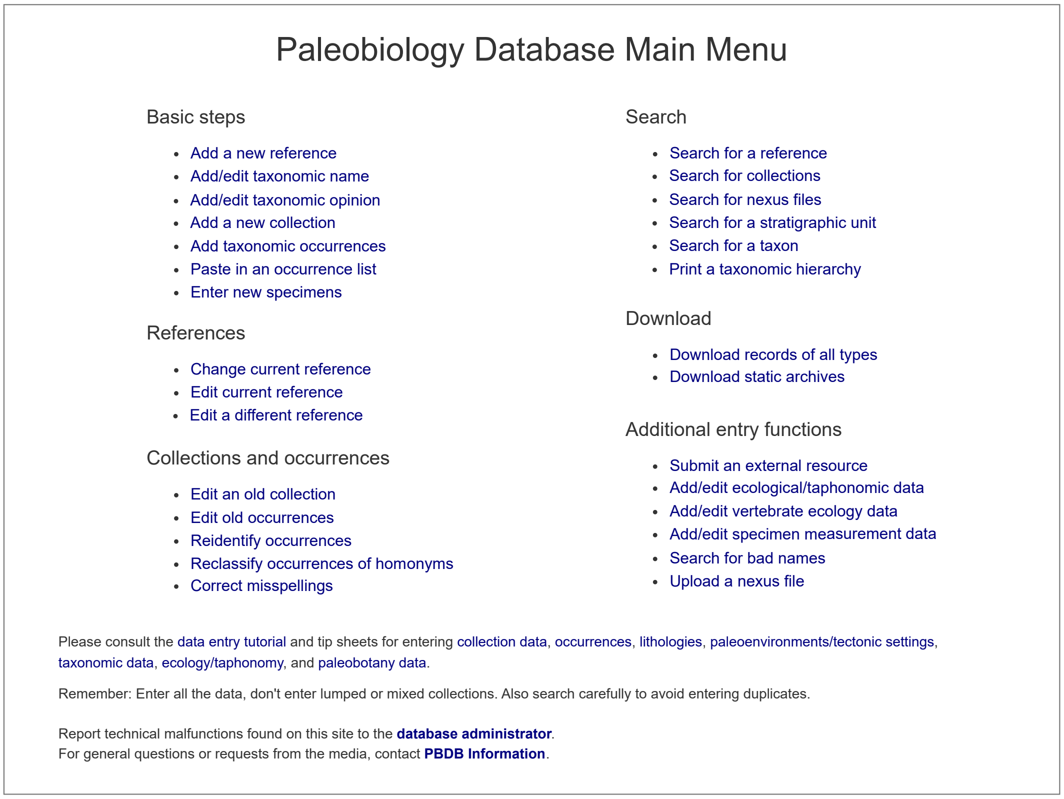
Manage Resources allows you to update metadata regarding any Resources that you have uploaded to the PBDB. Resources are discussed under section 2.4, Resources. (This will only be visible if you have submitted any Resources.)

**Manage Archives**

Manage Archives allows you to update metadata regarding and Data Archives that you have created. Data Archives are discussed under section 1.7, Data Archives. (This will only be visible if you have submitted any Archives.)

**Manage My Enterers/Students**

Manage My Enterers/Students allows you (an Authorizer) to manage those Enterer or Student users who work under you. To add new Enterers or Students, these



Paleobiology Database Main Menu

Basic steps

- Add a new reference
- Add/edit taxonomic name
- Add/edit taxonomic opinion
- Add a new collection
- Add taxonomic occurrences
- Paste in an occurrence list
- Enter new specimens

References

- Change current reference
- Edit current reference
- Edit a different reference

Collections and occurrences

- Edit an old collection
- Edit old occurrences
- Reidentify occurrences
- Reclassify occurrences of homonyms
- Correct misspellings

Search

- Search for a reference
- Search for collections
- Search for nexus files
- Search for a stratigraphic unit
- Search for a taxon
- Print a taxonomic hierarchy

Download

- Download records of all types
- Download static archives

Additional entry functions

- Submit an external resource
- Add/edit ecological/taphonomic data
- Add/edit vertebrate ecology data
- Add/edit specimen measurement data
- Search for bad names
- Upload a nexus file

Please consult the data entry tutorial and tip sheets for entering collection data, occurrences, lithologies, paleoenvironments/tectonic settings, taxonomic data, ecology/taphonomy, and paleobotany data.

Remember: Enter all the data, don't enter lumped or mixed collections. Also search carefully to avoid entering duplicates.

Report technical malfunctions found on this site to the **database administrator**.
For general questions or requests from the media, contact **PBDB Information**.

**Figure 15.** Paleobiology Database Main Menu

users need to first create a Guest account and then you can identify them as your Enterers or Students using this tool. Once Enterers or Students are no longer working for you, you can and should remove them from your lists of Enterers or Students.

### Data Entry Settings

Data Entry Settings allows you to personalize how PBDB presents some of the data entry forms to you. This primarily affects the Collections data entry form. If you are entering many collections with similar data, you can also tell PBDB to pre-fill some of the fields for you. Just be sure to turn those settings off when you are done with that data set!

### Editing Permissions List

Editing Permissions List used to be where Authorizers could tell PBDB which other Authorizers could edit their Collections data entry. In 2021, PBDB changed this policy so that all Authorizers can edit Collections data, so this function is no longer in use.

### Log Out

Log out safely takes you out of the PBDB. Otherwise, the system will keep you logged in for several days even if you close your web browser or browser tab with PBDB open.

## 3.3 Main Menu

### 3.3.1 General data entry recommendations

Before proceeding to the Main Menu to enter data, here are a few tips that apply throughout the PBDB data entry system. First, to the best of your ability please enter data EXACTLY as it is presented in the references. Your thoughts on these data are important, but your goal here is to enter the ideas of the reference authors, not yours. This includes using characters that are not standard in English (e.g., ø, ñ, é, etc.) wherever they are used in the reference. It also includes spelling mistakes in the title or names of taxa or geologic units. Second, do your best to enter all data from the reference. If you are only interested in the vertebrate taxa, still be sure to enter all of the invertebrate taxa too. Third, try to be sure not to duplicate data entry. This will save you time if data are already in the system. More tips on how to avoid duplicate entry in section 3.4.1 on Search for a Reference below.

### 3.3.2 Main Menu Page

The **Main Menu** (Figure 15) is the navigation hub from which all other functions can be accessed. Functions are grouped into six categories: **Basic Steps**, **References**, **Collections and Occurrences**, **Search**, **Download**, and **Additional Entry Functions** (see Figure 13). Given that all data entered into the PBDB come from References, we will start our instructions with those involving references. As we move forward, we are going to group these functions a little differently than on the **Main Menu**, but you can always refer back to this figure if you get lost.

## 3.4 Reference Functions

Reference-based functions can be found under **Basic Steps**, **Search**, and **References**.

### 3.4.1 Search for a reference

Sometimes you need to check and see if a given reference has already been added to the PBDB. To do so, click *Search for a reference* under **Search***.* On the **Reference search form** which opens, type in some information (but not too much) to try and find the reference you are looking for. We recommend entering the first author's last name and year in which the reference was published, then click Search. All potential matches will be listed on the resulting page. If there are no matches, PBDB will indicate that nothing matches your search parameters and take you back to the **Reference search form**.

You might think that entering more data will result in a more exact match, which might be true, if you enter EXACTLY what was entered into the PBDB. For instance, if you enter the title, any deviation from what was entered (extra spaces, added diacritical marks, etc.) will make the match fail. The same is true of the first author's last name, particularly with diacritical marks (accents; for example, "ö" may be entered as "oe"), so please be sure to try variations of the spelling, or perhaps enter a keyword from the title and the year instead of the author name and year.

Another issue to be aware of is the year of publication. Sometimes, references are posted online in one year, and then are published in print subsequently, and they take the later year as the year of publication. Since 2012, publications with nomenclatural acts can take the date of online posting only if there is evidence of registration in Zoobank given in the text, so to determine the correct publication date you will need to search for a Zoobank statement or LSID numbers in the document. Thus, if the reference was entered when it was published online, the reference may have a different (earlier) date than the finally published reference. A simple way to double check is to just use the author's last name, which will then show all entered publications, sorted by first author name and then by year.

A related consideration concerns monographs authored by multiple individuals. For example, some volumes of the *Treatise of Invertebrate Paleontology* are entered as a single reference whereas others are entered as individual chapters. The same is often true for older monographs, where different paleontologists may have authored sections for different taxonomic groups. Searching for "Treatise" or a related term in the title or book name usually offers a simple way to confirm.

### 3.4.2 Add a new reference

Adding a new reference is usually the first step in any data entry procedure (Figures 14 and 15): all other data types need to be linked to the Reference, so it needs to be entered first. Please refer to the Data Entry Checklist, in section 7.1, Appendix 1, I. Reference. When you click *Add a new reference*, you are taken to the **Duplicate reference search form**. The same guidelines given in section 3.4.1, *Search for a reference* also apply here. You are doing the same search, just for a different purpose. So, as above, try the first author's last name + year of publication, and/or perhaps a title keyword and click Search.

If potential matches are shown, look at them carefully and see if any truly are a match to the one you want to enter. If yes, then click on the matching reference and confirm that the user who entered the reference has also entered all of the data from the reference (see subsequent sections on *Taxonomic Functions* (3.5), *Collections Functions* (3.7), and *Occurrence Functions* (3.8)). If not, please select the reference and continue to enter the missing data. If none of the references match the one you wish to enter, select "Add Reference" and continue.

**Figure 16.** Navigation bar at top showing no reference selected and at bottom showing Uhen 2004 as the selected reference.

If there are no potential matches, you will be taken directly to the **New reference form** where you can enter the bibliographic data for your reference. You will want to select the **Publication type** first, as different fields appear and disappear depending on the Publication type. English is the default language, so be sure to select an alternate language if your reference is not in English. Any relevant data that does not fit into the pre-defined fields should be entered in the **Comments** field. (Please ignore the **Project Code** and **Taxonomy** fields as these are old data types no longer in use.) When you have all of the data filled in, click Submit to enter the data.

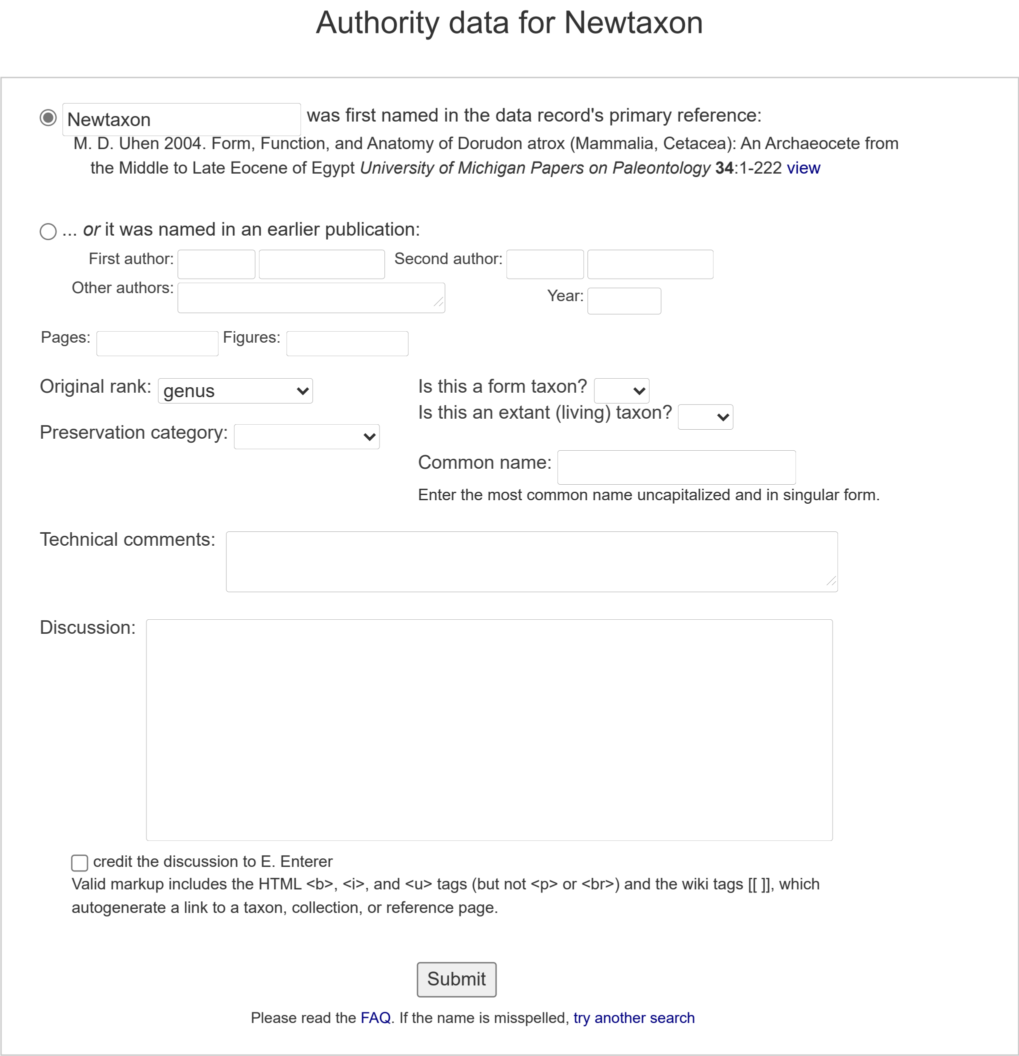
If you are at all concerned about which data go in which field, use *Search for a reference* to find a similar one that has been entered before that you can use as a guide. Remember, you can always open the PBDB in several browser windows and look at one filled-in form while adding data into another. Also, be sure that you enter the full titles of journals and do not use abbreviations of journal titles.

Once you have either selected a reference or added a new reference, you will notice a change in the black navigation bar at the top of your browser screen (see Figure 16 below). In A, no reference is selected. In B, Uhen 2004 is the selected reference. This means that all data entered from this point (until another reference is selected) will be associated with the Uhen 2004 reference

As you enter names, opinions, and collections from this reference (using the instructions below), it can be informative (and rewarding) to click on the reference link in the navigation bar. It will show you the names, opinions, and collections that you (and perhaps others) have entered. (Finding a great reference, one you know well and that has already been entered to the PBDB, and then looking at the opinions, names, and collections that have already been entered is a great way to get a feel for how others have entered data.)

### 3.4.3 Change current reference

If you have an active reference and you want to enter data from a different reference, go to the **Main Menu** and select *Change current reference* in the **References** section. Now, follow the procedure outlined under section 3.4.1, *Search for a reference* above and select the one you



Authority data for Newtaxon

● Newtaxon was first named in the data record's primary reference:
M. D. Uhen 2004. Form, Function, and Anatomy of Dorudon atrox (Mammalia, Cetacea): An Archaeocete from the Middle to Late Eocene of Egypt *University of Michigan Papers on Paleontology* **34**:1-222 view

○ ... *or* it was named in an earlier publication:
First author: Second author:
Other authors: Year:
Pages: Figures:

Original rank: genus
Is this a form taxon?
Is this an extant (living) taxon?
Preservation category:
Common name:
Enter the most common name uncapitalized and in singular form.

Technical comments:

Discussion:

☐ credit the discussion to E. Enterer
Valid markup includes the HTML <b>, <i>, and <u> tags (but not <p> or <br>) and the wiki tags [[ ]], which autogenerate a link to a taxon, collection, or reference page.

Please read the FAQ. If the name is misspelled, try another search

**Figure 17.** Authority data entry form.

want to enter data from.

Once you have an active reference, as indicated in the navigation bar, all data entered after that point will be associated with that reference. If your navigation bar shows no active reference, you can't enter any data other than reference data. If you try to enter any other data, the PBDB will prompt you to select or enter a reference first. It's good practice to get in the habit of selecting **Clear reference** in the **Actions** drop-down once you have finished entering data from a reference, to reduce the chances of accidentally entering the next reference's information under the previous one.

### 3.4.4 Edit current reference

If you have the correct reference active in your navigation bar, but you notice that there is an error in the bibliographic data for the reference itself, go to the **Main Menu** and click *Edit current reference*. This will pull up the *Reference entry form* with the reference data filled in. Make any needed changes and then click Submit. You can also just click on the reference in the navigation bar, then click edit.

### 3.4.5 Edit a different reference

If you want to edit a reference other than the active reference, go to the **Main Menu** and click *Edit a different reference*. This will pull up the *Reference search form*. Search for the reference you want to edit by following the procedure outlined under section 3.4.1, *Search for a reference*. Selecting the reference you want to edit will pull up the *Reference entry form* with the reference data filled in. Make any needed changes and then click Submit.

### 3.5 Taxonomic Functions

Taxonomic functions in the PBDB are divided into two types. The first lets you add or edit taxonomic names; the other lets you enter a particular author's opinion regarding the relationship between taxonomic names. Many data enterers find it convenient to enter taxonomic names and opinions (i.e., the Systematic Paleontology section) before entering other information from a reference (Figure 17), but this is a personal preference.

#### 3.5.1 Add/edit a standard taxonomic name

The *Add/edit a taxonomic name* function lets you add or edit a taxonomic name of any (or no) rank. Please refer to the Data Entry Checklist, in section 7.1, Appendix 1, II. Taxonomic Names. When you click on this option, you are first prompted to enter the name itself, along with optional metadata about the taxon. You must enter a taxonomic name, but the other metadata are not needed at this step. Note that the taxonomic name is entered with the exact spelling used in the original paper (with few exceptions for non-standard characters, outlined below). When you enter a new species, you will need to know the original genus in which the species was placed. Once you enter the name, click Search. The PBDB will search to see if the name has already been entered. If it has, it will bring up the **Authority data** entry form (Figure 15) where you can add or edit information regarding the taxon. If you change any previously entered data, please be absolutely sure what you are replacing was incorrectly entered by the previous PBDB user (as opposed to the information in the active Reference being "incorrect" or "out-of-date"), and that you are replacing it with the correct information. It may also give you some already-entered taxa which are spelled closely to what you entered, so please check carefully that these options definitely do not match the taxon you are trying to enter (for example, either your spelling or the name already entered could be a mistake).

Given the vast number of named fossil taxa in paleontological history, it is inevitable that many taxa will have homonyms. Adding opinions for the wrong name (i.e., the homonym) is an unfortunate error in data entry that can take significant effort to undo. It is worth spending a few extra seconds to ensure you are working with the correct name! When you search for a name, it will usually include the taxon you are expecting, but it might also include genuine homonyms (such as the bivalve *Donax* Linnaeus 1758 and the plant *Donax* Loureiro 1790) or outdated (junior) homonyms that have been subsequently replaced with a different name to avoid disallowed homonyms. The PBDB usually lists the name of a higher taxon in brackets to the right of each name. If the authority and the higher taxon match what you are expecting, then you can be confident you are choosing the correct name. Another easy way to confirm what a name means is to type the name in the Search bar at top right, which may give alternative higher classifications for the name and allow you to open its display page in a different window tab to get more information.

If you enter the taxonomic name, click Search and if the name has not been previously entered into the PBDB, it will bring up a blank **Authority data** entry form for you to fill in. At the top of the form are two radio buttons. Clicking the first will associate the name and the metadata you enter with the active reference shown in your black navigation bar and repeated next to the button. If your active reference says someone else named the taxon, you can click the second radio button and enter the metadata about the authorship of the taxon. For a species, the authors' last names and year of publication will be needed, but higher taxonomic ranks can be entered without this information. (When Jack Sepkoski's (2002) genus compendium was uploaded to the PBDB, the names were entered without authority information. If you notice a species or genus lacking an author and year next to it, your efforts to edit the name to add the authority, if you know it, will be appreciated. This also reduces the likelihood of homonym errors for future users.)

Use the additional pull-down fields to enter further metadata about the taxon, including original taxonomic rank, preservation category ("not a trace" is a legacy category that is no longer allowed), whether it is a form taxon or not (see section 3.6 below for how to handle ichnofossils and form taxa), and whether it is an extant taxon or not. If you are not sure whether the taxon includes living representatives, we recommend searching using a taxonomic authority database to see if you can fill in that field accurately. Extinct taxa are marked with a dagger (†) in display views. (The Integrated Taxonomic Information System [ITIS] and the World Register of Marine Species [WoRMS] can be really useful for this purpose). Enter a common name for the taxon if it is known. This is rare, except for extant species and genera. The database propagates common names down the taxonomic hierarchy, so it is not necessary to enter "snail" for every gastropod species, because the common name "snail" has been assigned to the taxon Gastropoda and everything within it (unless overridden by a different common name at a lower taxonomic level). Use the Technical Comments field to optionally add any additional data regarding the taxon that is not captured in the other fields.

When you are done entering the taxon and its associated metadata, click <u>Submit</u> at the bottom of the form. After you submit the new taxonomic name, if missing required data are missing, a yellow box will appear and list that data that need to be completed. If the taxon and associated data were entered properly, a follow up menu will appear with many options. When the new taxon is at the genus rank or higher (or is unranked), you will want to click *Add an opinion about Newtaxon*, where Newtaxon is the new taxonomic name you entered. This will take you to the same place as clicking on *Add/edit a taxonomic opinion* in the **Main Menu** without needing to enter the Newtaxon name again. When entering a new subgenus (make sure the subgenus is not already entered as a genus, which would require an opinion rather than a name), species, or subspecies, the opinion will automatically be created.

### 3.5.2 Adding a taxonomic name with non-standard naming

Some older taxonomic names do not follow the modern rules of taxonomic nomenclature and PBDB will reject those names that are improperly formatted. The PBDB rules are based on the International Code of Zoological Nomenclature (ICZN 1999), so some allowable practices under other codes (such as hyphens in names) are currently not permitted in the PBDB. Some of the issues with these older names include but are not limited to inclusion of numbers in the name, inclusion of punctuation (such as hyphens) in the name, inclusion of diacritical marks in the name, inclusion of spaces in the name. Here we list recommendations as to how to deal with these issues. It is also a good idea to see how subsequent taxonomists have dealt with these issues for the name(s) in question and to follow their lead. Many of these issues and more are covered in ICZN article 32.5. which deals with original spellings. Please refer to the code for any additional concerns. Whatever you do, please use the Technical Comments field to note the original spelling of the original name.

#### Numbers

Early taxonomists occasionally included numerical digits in the names of taxa. To enter the name, translate the number into Latin and make that Latin word part of the name. For example, Linnaeus named the ladybird *Coccinella 7-punctata*, which was entered into the database as *Coccinella septempunctata*.

#### Hyphens

If a name includes a hyphen, simply eliminate the hyphen and combine the prefix with the main portion of the name. For example, the modern striped dolphin *Stenella coeruleoalba* (Meyen, 1833) was originally named *Delphinus coeruleo-albus* Meyen, 1833. To correct this, the hyphen is simply eliminated, and the gender of the specific epithet was also corrected when the species was moved from the genus *Delphinus* (male, -*us*) to *Stenella* (female, -*a*).

#### Diacritical marks

If a name includes letters with diacritical marks, just drop the diacritical except in the following instances.

| Letter | Name derived from | Replacement |
|---|---|---|
| ä | German | ae |
| ö | German | oe |
| ü | German | ue |
| ß | German | ss |
| å | Norwegian/Danish | aa |

If the name is not derived from those languages (e.g., nordenskiöldi, derived from a Swedish name and not German), the replacement character should not be inserted (e.g., nordenskioldi and not nordenskioeldi). Authors do not always follow these rules correctly, sometimes omitting the replacement character and at other times inserting one incorrectly. Such misspellings should be recorded faithfully, but noted as a misspelling, when entering taxonomic opinions (described more below).

#### Spaces

If the species epithet was originally written as two or more words (and not intended to be a species and subspecies), combine them into a single word. For example, the Cretaceous oyster originally spelled *Ostrea Sanctæ Crucis* should be entered as *Ostrea sanctaecrucis*.

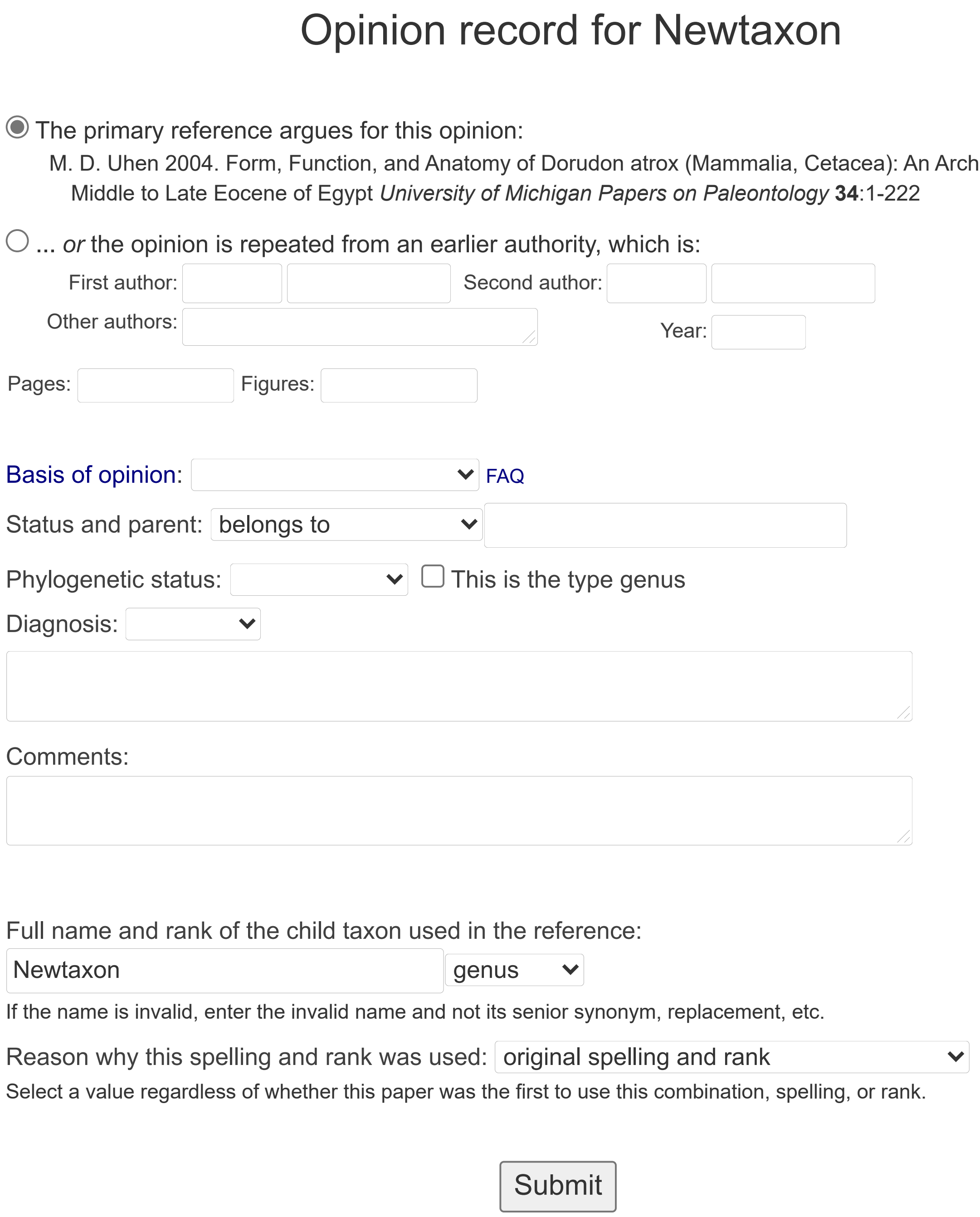
#### Capitalized specific epithet

If the species epithet was originally capitalized, for example in reference to a person or place, replace the capital letter with a small letter. For example, the Ordovician trilobite originally spelled *Calymene Blumenbachii* should be entered as *Calymene blumenbachii*.

### 3.5.3 Add/edit a taxonomic opinion

**Taxonomic opinions** are claims of a relationship between two different names. Please refer to the Data Entry Checklist, in section 7.1, Appendix 1, III, Taxonomic Opinions. For example, *Tyrannosaurus rex* is a member of the genus *Tyrannosaurus*, which belongs to family Tyrannosauridae. These opinions are critical for providing taxonomic and phylogenetic structure to the taxa in the PBDB. Many users start their data entry with the Systematic Paleontology section of systematics paper, which contains many such opinions, of varying degrees of strength. Most new (and even advanced) data enterers find entering opinions correctly among the most challenging of tasks. Luckily, most opinions are simple and straightforward. Here, we provide guidance on how to do it correctly every time. (And like many tasks in the PBDB discussed here, it can be informative to find a great reference you know well that has already been entered to the PBDB, and then looking at the opinions, and names and collections, that have already been entered.)

These opinions are typically found in the systematic or discussion portions of a paper or in tables. Presentation of a phylogenetic tree may express taxonomic opinions but should only be entered if the author expressly uses them as the basis for classification. The *Add/edit a taxonomic opinion* function lets you enter a particular reference's opinion regarding a named taxon. When you click on this function, the PBDB will ask for the taxon



Opinion record for Newtaxon

● The primary reference argues for this opinion:

M. D. Uhen 2004. Form, Function, and Anatomy of Dorudon atrox (Mammalia, Cetacea): An Archaeocete from the Middle to Late Eocene of Egypt *University of Michigan Papers on Paleontology* **34**:1-222

○ ... *or* the opinion is repeated from an earlier authority, which is:

First author: Second author:

Other authors: Year:

Pages: Figures:

Basis of opinion: FAQ

Status and parent: belongs to

Phylogenetic status: ☐ This is the type genus

Diagnosis:

Comments:

Full name and rank of the child taxon used in the reference:

Newtaxon genus

If the name is invalid, enter the invalid name and not its senior synonym, replacement, etc.

Reason why this spelling and rank was used: original spelling and rank

Select a value regardless of whether this paper was the first to use this combination, spelling, or rank.

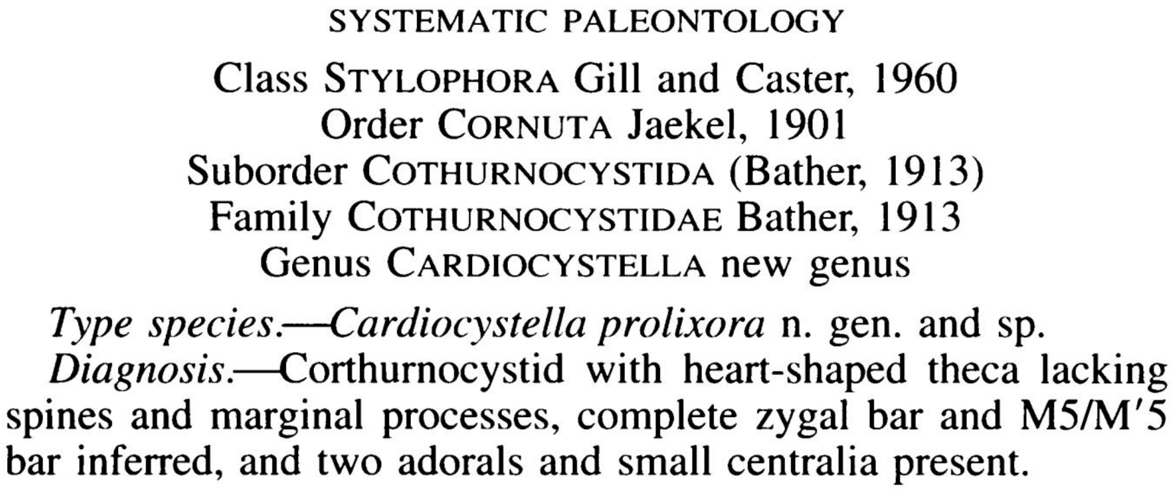
**Figure 18.** Opinion record form.

name, along with optional metadata about the taxon. You only need to enter the taxon name and click *Search* (and heeding the same recommendations above when searching for a taxon name). The PBDB will then display the list of taxonomic opinions that have been entered for this taxon, along with a clickable link that says *Create a new opinion record*. Before entering a new opinion about the taxon, be certain that it is the correct taxon and not a homonym. If the database contains homonymous names, you will be prompted to choose the pertinent one before proceeding, but you may be dealing with a homonym that has not previously been entered. To check, compare the current reference's opinion to the placement of the taxon in previous opinions. If the placements are substantially different (i.e., belonging to different orders, classes, or phyla), you should investigate whether your reference is dealing with a homonym that is not yet in the database. You will notice that one of these opinions is in **boldface type**. This is the one that the PBDB is currently using to assign this taxon to a higher taxon. Click on one of the previously entered opinions to edit (this is only necessary if your current reference corresponds to that previous opinion and you want to formally link the two) or click *Create a new opinion record* to create a new opinion.

As when searching for a name, if there is more than one taxon of the same name, the PBDB will ask you first to choose "Which taxon do you mean?" If there are no matches, it will offer a list of similar names to consider. As above, it is worth spending a few seconds to confirm that you are adding opinions to the correct name. (To be sure, you can open the opinions for each matched name on a second tab/window to confirm that the opinions are consistent with what you expect for a given name.)

The *Opinion record* form (Figure 18) is divided into three boxes. In the first box at the top of the form are two radio buttons. Clicking the first will associate the opinion and the metadata you enter with the active reference shown in your black navigation bar and repeated next to the button. If you want to enter an opinion from an earlier reference that is stated in your active reference, you can click the second radio button and enter the metadata about the opinion. (These secondary opinions are most commonly used when adding synonyms or genus recombinations for a particular taxon.)

The second box is where the opinion itself is expressed. This can be quite tricky and can have profound downstream consequences in the database, so please proceed carefully! For this reason, we do not allow Student users to enter taxa or opinions. Only Authorizers and their Enterers are allowed to enter these data. See Figure 19 for an example of how these opinions are often expressed in paleontological publications. In this section, it's handy to have some terminology to describe the taxa whose opinions we are entering. We will use the terms **child taxon** and **parent taxon**. The child taxon is the taxon we are entering opinions about. The parent taxon is the taxon that we are placing the child taxon within. For example, if we have the genus *Dorudon* as our child



SYSTEMATIC PALEONTOLOGY

Class STYLOPHORA Gill and Caster, 1960
Order CORNUTA Jaekel, 1901
Suborder COTHURNOCYSTIDA (Bather, 1913)
Family COTHURNOCYSTIDAE Bather, 1913
Genus CARDIOCYSTELLA new genus

*Type species.*—*Cardiocystella prolixora* n. gen. and sp.

*Diagnosis.*—Corthurnocystid with heart-shaped theca lacking spines and marginal processes, complete zygal bar and M5/M′5 bar inferred, and two adorals and small centralia present.

**Figure 19.** Typical Systematic Paleontology section of a primary journal reference (Sumrall et al. 2009, PBDB ref. 65844). In this listing, the opinion linking species *Cardiocystella prolixora* to genus *Cardiocystella* and that linking *Cardiocystella* to Cothurnocystidae are "with evidence" whereas the other opinions are "without evidence."

taxon, one opinion that can be expressed about it is that the genus *Dorudon* (the child taxon) belongs to the family Basilosauridae (the parent taxon). One thing to remember is that to place a child taxon into a parent taxon, the parent taxon must already exist in the PBDB. If it does not, enter the parent taxon first. If you forget, the PBDB will insist that you do so. (It is usually easiest to open the "Add/Edit Names" link in a new window tab to add this, so that you don't have to spend time re-entering your initial opinion.)

To enter an opinion, you must fill in the first three fields in the second box. The others are optional. The first field is the *Basis of opinion*. This allows you to indicate the weight that should be put behind the opinion. The pull-down has four choices, in order from greatest weight to least weight: *stated with evidence, stated without evidence, implied,* and *second hand*. (Technically, there is also a blank category, but if you have a reference, there is no reason to leave this blank.) For the most part, these categories are self-explanatory, as long as one understands what "evidence" is. Evidence is some sort of reason that the author said this, such as a phylogenetic analysis or demonstration that a particular taxon has the identifying characteristics of a group. Stated without evidence is most common in lists of "included taxa" (when the individual taxa are not formally evaluated or described). Although there is some disagreement among data enterers, comprehensive monographs written by expert authorities (such as the *Treatise on Invertebrate Paleontology*) are typically treated as "with evidence," but only when the assignment of parent and child taxa include formal diagnoses or phylogenetic analyses. References that simply list subtaxa without such information, or that summarize other references, should be treated as "without evidence" or "implied," even when you trust the opinions or the expertise of the authors.

An explanation of why these categories are important is informative. When the database compares opinions for a taxon, it will always use the most recent opinion stated with evidence as the opinion of record (both for displaying information on a taxon and when downloading data). If no opinions exist with evidence, then it searches downward for the most recent opinion without evidence, then implied, etc. until it finds some opinion to use. (Never change the weight simply because you disagree with a reference! If you think the reference's opinion is outdated or wrong, and the database is treating it as the opinion of record, then the only scientifically honest solution is to add a more recent reference that expresses a newer opinion with evidence.)

The *status* and *parent* fields work together. *Status* is a pull-down menu where the choices delimit the nature of the relationship between the parent taxon and the child taxon, while *parent* is where you enter the name of the parent taxon. The items on the status list include *belongs to, subjective synonym of, objective synonym of, replaced by, misspelling of, invalid subgroup of, nomen dubium, nomen nudum, nomen oblitum* and *nomen vanum*. These all have very specific meanings that are distinctly different from one another, so please be sure you are selecting the correct relationship.

*Belongs to* is by far the most common entry in this field. As noted, it means that the child taxon is a member of the parent taxon e.g., a species is included in a genus, or a family is included in an order. While we are speaking in terms of Linnaean ranks here, don't worry about rank just yet. *Belongs to* works for unranked names just as well.

The next two choices deal with synonymy. The PBDB includes a powerful algorithm that works in the background to match occurrences that were entered under different names but are actually the same taxon because the names have since been determined to be synonyms of one another. A taxon that is biologically the same as another taxon that was named previously, even though the names are different, are called **synonyms**. By the rules of nomenclature, the oldest name must be used (i.e., has "priority" and is "senior" to the other name), except in special circumstances. The PBDB keeps track of this for you, determines the senior synonym where applicable, and updates all taxonomic occurrences accordingly. All you need to do is enter opinions regarding synonymy carefully and correctly. In particular, it is crucial to determine whether a paper is synonymizing a species or whether the paper instead feels that a previous work identified specimens incorrectly (i.e., the species is valid, but the previous paper misidentified their material, which would be treated as a reidentification of the relevant occurrences in the PBDB). If the synonymy list does not distinguish the author(s) of the species from the author(s) of the previous paper, it can be difficult to distinguish synonymy from misidentification, so be careful.

To enter an opinion regarding synonymy, just enter the name of the senior synonym in the parent field and choose *subjective synonym of* or *objective synonym of* in the status field. These synonyms must be of the same rank. So, for example, a species can be a synonym of another species but not of a genus. A synonym is objective in the rare case that the same type specimen defines both species. For example, Cope (1882) named A.M. 3038 *Taeniolabis sulcatus*, and Cope (1884) renamed it *Taeniolabis scalper*. So, *Taeniolabis scalper* is a junior objective synonym of *Taeniolabis sulcatus*. If the author is declaring that two different taxa with two different type specimens are actually the same species, this is a subjective synonym instead.

**Homonyms** describe multiple taxa that have been given the exact same name. Often this is because, under the nomenclatural rules, there is nothing stopping animal taxonomists from using a name that has already been given to a plant taxon, and vice versa. Occasionally, authors make mistakes and use the same name for a new taxon that had been previously used for a completely different taxon, within animals or within plants, which is not permitted under nomenclatural rules. The later taxon usually needs to have a new name applied to avoid confusion. For example, *Mesocetus* was originally named by Van Beneden (1880) for a mysticete whale from Belgium. Later, Moreno (1892) named a sperm whale from Argentina *Mesocetus*, and Fraas (1904) also named an archaeocete from Egypt *Mesocetus*. The rules of nomenclature say that *Mesocetus* Van Beneden (1880) has priority because it was named first, so the Van Beneden genus gets to be called *Mesocetus*, and *Mesocetus* Moreno 1892 and *Mesocetus* Fraas 1904 needed to have replacement names. To enter one of these *replaced by* opinions, you would first choose one of the taxa that needs the replacement name (for instance *Mesocetus* Moreno 1892) and then on the **Opinion record** form, choose the status *replaced by*, and enter the replacement name in the parent field. The new name needs to have already been added to the PBDB before this relationship can be applied. In this case, *Mesocetus* Moreno 1892 was replaced by *Diaphorocetus* Hay 1902. Also, *Mesocetus* Frass 1904 was replaced by *Eocetus* Fraas 1904 in a subsequent publication in the same year by the same author when he realized his mistake. It can be informative for new data enterers, and a good refresher for seasoned ones, to view how opinions like these have been entered to the PBDB. Homonyms present special difficulties for many data enterers, so it is worth heeding the recommendations above on identifying homonyms to confirm that the names you are intending to add opinions to are actually the names you intend. Long-time PBDB data enterers can recount, with some horror and with sympathies to the programmers who had to go through the source code, the difficulties of disentangling opinions that were linked to the wrong homonym.

Occasionally, authors make spelling mistakes, and these should also be recorded in the PBDB so that incorrect names can be correctly linked to their proper names. How you enter these misspellings depends on whether the reference is noting a historical misspelling or whether you recognize that the reference has a misspelled name (i.e., an unintentional typo). In the first case, simply search for the misspelled name and add an opinion for it. Choose misspelling of in the status pulldown, make sure the parent is the correct spelling of the taxon, and choose misspelling in the "Reason why this spelling and rank was used." In the second case, search instead for the correctly spelled name (i.e., not the misspelled name in the publication), choose "belongs to" in the status pulldown, enter the parent name, and choose misspelling in the "Reason why this spelling and rank was used." If the misspelled name has not been previously entered, you will be prompted (as a warning) to allow the PBDB to automatically add the new name to its database of names. In both cases, the child taxon should be the misspelled name of the taxon; if this is not the case, and the child name is instead a correction of the parent name, see the Opinion record section below on corrections.

You can choose *invalid subgroup of* when authors enter opinions regarding some, but not all, members of a taxonomic group. This almost always applies to taxa at the family-group rank or higher (or unranked). For instance, X is an invalid subgroup of Y if the author thinks all of X's genera do belong to Y, but they do not form a coherent group. X and Y could be, say, a subfamily and family. This is also used if authors argue for discontinuing the use of subfamilies within a family, for example.

The last four choices in the status pull-down take the form of *nomen* ____. These indicate that a taxon is recommended to be dissolved, because it is unclassifiable or was erroneously or inappropriately named and is not merely a synonym. For species, the difference usually is that a synonym's type is specifically determinate (you know which species it really is), whereas the types of invalid names cannot be equated with a particular species or cannot be defined as a coherent whole. The terms can be defined this way:

- *nomen dubium* means that the type is poor (its identity is currently unclear, and it deserves further study) or that a taxon is incoherent (often it is also indicated that its constituent specimens can be transferred to other taxa);
- *nomen nudum* means that the taxon cannot be formally recognized because the author did not follow the rules of nomenclature when naming it (usually, because the type specimen was not designated, illustrated, or described);

- *nomen oblitum* means that the taxon is probably identical and senior to another one, but has not been used as a valid name in the literature since 1899, and the younger name has been used in at least 50 publications, so the younger name is conserved;
- *nomen vanum* means that the type specimen is so poor that it is certainly indeterminate at the species level. Typically, this is used when there is little or no potential that additional material will clarify the identity of the taxon.

See Smith (1945) (and the FAQ pop-up) for additional information on these and other taxonomic terms. For all of these, enter the most specific higher taxon possible in the parent field. For example, if *Equus validus* is a *nomen dubium* but still definitely a horse, you can enter Equidae in the parent field. It is possible for a species to be a *nomen dubium* falling within a recognizable genus.

Continuing on in box two, there are three optional fields. The first is *phylogenetic status* where you can choose from: monophyletic, paraphyletic, polyphyletic, or unknown. For any family level taxon or lower rank, there is a checkbox that states: *This is the type [rank].* Be sure to check this if you are entering an opinion about a type species (for a genus) or type genus (for a family-group taxon), but it is not required for other ranks. When specifying a type genus, a subsequent page will let you choose all of the family-group ranks (tribe, subfamily, family, superfamily) for which that genus is the type, but only if opinions connecting those family-group ranks have been entered from the same Reference. The *Diagnosis* field is required if the opinion is *belongs to*. The *Diagnosis* choices are *none*, *new*, *emended*, *repeated*. Quite often authors do not include a diagnosis, in which case choose *none*. If the taxon is new, then the diagnosis must be new, so choose *new*. If the diagnosis is a revision or based on a previous diagnosis, but not repeated exactly as the previous author, choose *emended*. If the diagnosis is quoted directly without changes from a previous author, or the diagnosis simply cites a prior reference, choose *repeated*. The text box under the *Diagnosis* pull-down is where you can type or cut-and-paste the diagnosis itself. Finally, for box 2, use the comments field for anything else regarding the opinion that is not captured in the other fields.

The last part of the **Opinion record** form is in box three. Here you can specify the spelling and rank of the child taxon, along with the reason that spelling and rank are being used. The *full name* field is for the name of the child taxon spelled *as it is in the reference*. The *rank* pull-down is for the Linnaean rank of the taxon, if given. If no rank is listed in the reference, choose *unranked clade*.

Finally, you must choose a *Reason why this spelling and rank was used* using the final pull-down. The four choices are: *original spelling and rank, correction of* [*child taxon*]*, misspelling,* and *rank changed from original rank of* [*original rank*] (for a species, this last option will be *recombination or rank change of [child taxon]*). The *original spelling and rank* option is for when the name is exactly as it was when it was first named, with no changes. The *correction of* [*child taxon*] option is for when the author of your reference is fixing a problem with the original spelling of the name regarding number, gender, tense, etc. in the original Latin derivation of the name. (Examples include correcting an animal superfamily from ending in -acea to ending in -oidea, or when the crinoid *Roveacrinus alata* was corrected to *Roveacrinus alatus*.)

The misspelling option is used if the reference corrects or inadvertently uses a misspelled name. (See discussion of how to enter misspellings above.) In this instance, put the incorrect spelling in the full name field in box three and choose "misspelling" in the reason.

The *rank changed from original rank of* [*original rank*] option is for when a taxon changes rank, which may also include a spelling change or a recombination. For instance, the spelling of a family (Scolytidae) changes if it is downranked to a subfamily (Scolytinae). The same is true when a species moves from one genus to another (listed as *recombination or rank change of [child taxon]*): for instance, *Prozeuglodon atrox* was moved to the genus *Dorudon*, so the species name is now *Dorudon atrox*. Similarly, the new combination or spelling is the name that should be entered in this box. Pay careful attention to changes in the spelling of the species epithet, for example to match the gender of the genus name, but enter the name as spelled in the reference.

Finally, for opinions about subgenera, there is a fifth option: *reassigned from the original genus "..."*. This is used when authors argue that the taxon should still be considered as a subgenus but wish to move it into a different parent genus.

When all of the fields are appropriately filled in, click Submit at the bottom of the form. PBDB will do a check to see if there are any problems with the data entered. If all is well, PBDB will return with a summary of the actions indicated by the form, and a new menu to allow you to continue with various taxonomic data entry tasks. (Clicking "Get information about Newtaxon" on the confirmation page is usually a good way to confirm everything looks as intended, but see below.)

If there is a problem, the PBDB will return an error box in yellow with red text alerting you to the error. See Appendix 1. In the case of a recombination, the PBDB will always return an error box because the new species name (e.g., *Dorudon atrox* from the above example) is also new to PBDB. The purpose of this error box is to require the user to confirm the new name before the changes are made to the database, because these changes are difficult to undo. This is also a good opportunity to double-check the spelling of the name to ensure that you remembered to update the ending of the species epithet if it changed in the recombination. Please read through these boxes very carefully, and only proceed with caution. If you don't fully understand what these boxes are trying to convey, do not proceed. Consult with another Authorizer or with someone from the PBDB ExCom to help you figure out how to proceed.

When you enter new misspellings or new binomial recombinations (i.e., a species is now placed in a new genus), you will be prompted (as a warning) to allow the PBDB to automatically add the new name to its database of names. This allows other references that have used this new name (even if misspelled) to correctly link the alternative names to the same biological name. If you agree that the new name was used in your reference, then click *Submit* to proceed. Always be sure that the recombination is a true recombination, and not just a collection of specimens previously assigned to one species being reassigned to a different or new species. Instances of the latter are accommodated by reidentifying occurrences in Collections records (see more in section 3.7 below).

If you make a small mistake, you can probably correct it yourself. If you make a serious mistake, for instance, combining two taxa that should not have been combined or adding opinions for the wrong homonym, you will need to contact PBDB support for assistance. For example, if you get a giant error box that states that "Taxon already exists in the database with opinions classifying it," it is best to stop, do not submit the page, and contact PBDB support. These can sometimes be carefully undone by data enterers with deep familiarity with how the algorithms work. Sometimes it requires knowledgeable people to change code manually. We are all human, and the authors of this guidebook have all been humbled by making such mistakes. Rather than ignore a mistake, please contact the PBDB for assistance.

Please be sure to enter ALL opinions about all taxa in a given reference, even if a similar opinion has already been entered into the PBDB. It is essential that we capture all opinions about all taxa for future reference. You also never know when a taxon is moved from one parent to another, and then moved back into the original parent. Capturing all opinions records the history of each taxon as they have changed over time.

Be aware that if you add a new name or opinion (especially if the parent name is new), it will not show up in the navigation bar search tool (nor correctly in the Classification tab of the display page) until the next day, after the database re-builds its namespace and parent-child relationships overnight. If you immediately want to confirm it will display as intended, it is best to click the "Get general information about Newtaxon" link on the page after being entered (or to search for it on the Search > Fossil organisms" page). To confirm the Classification tab of a newly entered taxon name shows correctly, you will need to wait until the next day.

### 3.6 A note on trace fossils (ichnotaxa) and form taxa

Trace and form fossils constitute an important and common part of the fossil record, and so it is incumbent to include them in the PBDB. However, these types of fossils require special considerations during data entry, downloading, and carrying out subsequent analyses. A form taxon (specifically a morphotaxon) is "a fossil taxon which, for nomenclatural purposes, comprises only the parts, life-history stages, or preservational states represented by the nomenclatural type" (McNeill et al. 2006: 4)." Examples include different plant organs (e.g., seeds, pollen, bark, branches, fruit, and leaves of the same plant having different names for each organ), beetles named from isolated elytra, caddisfly cases, rhyncholites (cephalopod beaks), shark fin-spine taxa, and occasionally isolated decapod claws. Although not typically treated as form taxa *sensu strictu*, trace fossils (ichnotaxa) should also qualify as such under this definition. An ichnotaxon is a formal name given to a trace fossil, "a morphologically recurrent structure resulting from the life activity of an individual organism (or homotypic organisms) modifying the substrate" (Bertling et al. 2006: 266). Examples include trackways, resting or feeding traces, burrows, borings and bioclaustrations, root traces, and sometimes even coprolites, gastroliths, nests, and the agglutinated tests of certain foraminifera. Complicating matters further, in both types of fossils, the same biological source artifact can be assigned different names depending on the nature of preservation, facies, or depositional environment.

The basic concern for the PBDB is that, unlike "regular" body fossils, these other fossil taxa are not always diagnostic of a single biological unit (i.e., they are polyphyletic). In most cases, a single organism can yield multiple distinct form and trace fossils, whereas a diverse number of unrelated organisms can produce similar form and trace fossils (Bengtson 1988; Bertling et al. 2006; Bateman and Hilton 2009; Buatois and Mangano 2011). A further complication is that it is difficult, if not impossible, to identify the producer of these fossils with confidence, although there are some exceptions. In part because of these challenges, there is also a lack of a clear consensus among taxonomists working with such fossils on how best to name, taxonomically classify, and assign them to higher taxa (Buatois and Mangano 2011, especially their chapter two). For example, the ICZN has long provided rules for "works of animals," which includes trace fossils, whereas paleobotanists more recently (in the 2000 and 2008 International Code of Botanical Nomenclature) diminished the standing of paleobotanical form taxa (Bateman and Hilton 2009).

Here, we offer the following suggestions when dealing with form taxa and ichnotaxa (trace fossils). Ultimately, however, data enterers should be respectful of the conventions the source reference uses when entering these fossils.

1. When adding new names, **make ichnotaxa "extinct" by default**. Ichnologists (and the ICZN formally) exclude naming ichnotaxa based on living organisms. In this sense, they are treated as past relics, even when the likely tracemaker is living (Bertling et al. 2006). An exception might be able to be made for exceptional forms (e.g., *Paleodictyon*, Rona et al. 2009; Hsieh et al. 2019). There is less of a consensus on this issue for most form taxa, and so it is best to follow the convention of the references being added for form taxa.
2. When adding new names, **classify ichnotaxa as both "form" taxa and as "preservation = trace."**
3. When adding new names, **classify form taxa as "form" taxa**.

Within the realm of PBDB form taxa and ichnotaxa, some paleontologists often assign ranks and affinities to their specimens (e.g., *Grallator* and six other genera in Grallatoridae "belongs to" Saurischia), whereas other paleontologists typically treat them as phylogenetically unassignable (i.e., *Cruziana* "belongs to" to the unranked ichnotaxon Ichnofossils instead of Trilobita *sensu strictu*). These differences can have substantial consequences on downstream analyses. For example, if creating a genus diversity curve for "all" PBDB genera for the clade Life, saurischians will include such trace fossils (and potentially double count the tracemaker), but entirely exclude the more than 300 ichnogenera like *Cruziana* and *Skolithos*. The best way to address these discrepancies is to be aware they exist and to account for them. One simple way to treat these cases consistently is to take advantage of the "regular" versus "form and ichnotaxa" preservation categories on the Download page (see section 2.2.2); hence the importance of heeding recommendations #2 and #3 above). For data that has already been downloaded, taxonomic and occurrence downloads have a field named "Flag" that can be used to identify these categories; it uses the letters B, I, F, or combinations (e.g., IF, BF) to identify taxa that are body (=regular), ichnotaxa, and form taxa, respectively. (The occurrences download also uses the Flag "R" for taxa that have been reidentified; see 3.8.4 *Reidentify occurrences* below.)

4. Given the lack of consistency across disciplines, we recommend **adding ichnotaxa opinions that "belong to" a biological parent taxon when the probable tracemaker is confidently identifiable** and constitutes a single and biologically cohesive group. **When such an affinity is unknown or the affinity is confidently polyphyletic, then it is best to have them "belong to" parent Ichnofossils**, and to then add a note about the possible identity of the tracemaker(s) to the "Discussion" field on the taxon's "edit name" page so that users can understand its possible affinity.

### 3.7 Collections Functions

A **collection** in the PBDB is the basic time/space/geologic unit that contains occurrences of fossils. A good way to think of a collection is that it refers to a group of fossils reported from a single stratigraphic horizon at a single geographic place, often collected at a single time. Depending on the information given in the publication, the stratigraphic horizon can be precise (a single rock layer) or vague (unspecified and perhaps multiple rock layers at the place). The geographic place can also be precise (a small collection at one clearly demarcated site such as a roadcut) or vague (only referring to a nearby town or feature). If the taxonomic list is for an entire formation or basin, and clearly represents multiple locations grouped together, you should avoid entering that data. What defines a collection can be subjective, but your goal is to determine the most precise way of dividing the fossils into collections.

It can be challenging to decide whether occurrences from a reference should be added to an existing collection or should form the basis of a new collection. The definition of a collection is fuzzy and depends on the context. We offer some tips below as to how to determine whether fossils come from the same collection.

If you have any doubts about whether fossils come from the same collection, add a new collection, even if it has the same basic time/space/geologic unit data.

If the fossils come from a specific numbered museum collection site (e.g., UCMP A-3210, PIN 1701, etc.), new occurrences should be added to that collection.  Do not combine museum locality numbers under a single collection unless this is specifically indicated in specific References. Although legacy instances of such collections occur in the database, it is not desirable.

If an existing collection name is the same or similar to the description in the reference, click on the collection to examine its geographic and stratigraphic information in the comment fields. If the collection has specific and detailed information about the location and the stratigraphic horizon, compare with the data in your reference to make sure everything matches. If all information matches exactly (beware of fossils collected from different stratigraphic horizons at the same place), the new occurrences should be added to the existing collection. It may be necessary to check the original collection reference if the geography and stratigraphy comment fields do not contain specific details.

If the publication only provides vague geographic and/or stratigraphic data, and existing database collections have generally matching but also vague information, it can be difficult or impossible to determine whether the specimens came from the same outcrop and bed. For example, some publications only report the name of a nearby town and only give the geological formation for a stratigraphic unit. In these situations, here are some points to consider. Are some of the authors of your current paper the same as authors of other papers in the collection? If the collector is given, is it the same collector as for the new occurrences? If the collector isn't known, are the specimens at the same museum? Agreement in some or all of those areas would lean towards adding the occurrences to the existing collection rather than creating a new one. Conversely, if the authors are different, the collectors or repositories differ, and/or the publications dates are far apart (e.g., one paper from the 1880s and the other from the 1970s), it is more likely that the available outcrops have changed over the intervening time. In that situation, it may be better to create a new collection. For example, Schimper (1853), Bill (1914), and Förster (1967) all reported material from the Grès à Voltzia exposed in unspecified quarries in the vicinity of Wasselonne (Wasselnheim). Because the specimens were collected by different people, some are stored in different museums, and because the papers span more than a century, each was made a separate collection.

Some literature data are especially vague, particularly in older publications and/or for material from famous Lagerstätten where specimens were often purchased without context (e.g., Solnhofen, Baltic amber, etc.). Rather than placing thousands of species in a single collection, these Lagerstätten have largely been split into separate collections on the basis of the repository. For example, there are separate collections "Solnhofen (BMNH collection)", "Solnhofen (MNHN coll)", "Solnhofen (SMNS collection)", and so forth. Those vaguely defined collections exist alongside collections that have more specific details, such as the collection "Blumenberg quarry, Eichstätt, Solnhofen (BMNH collection)."

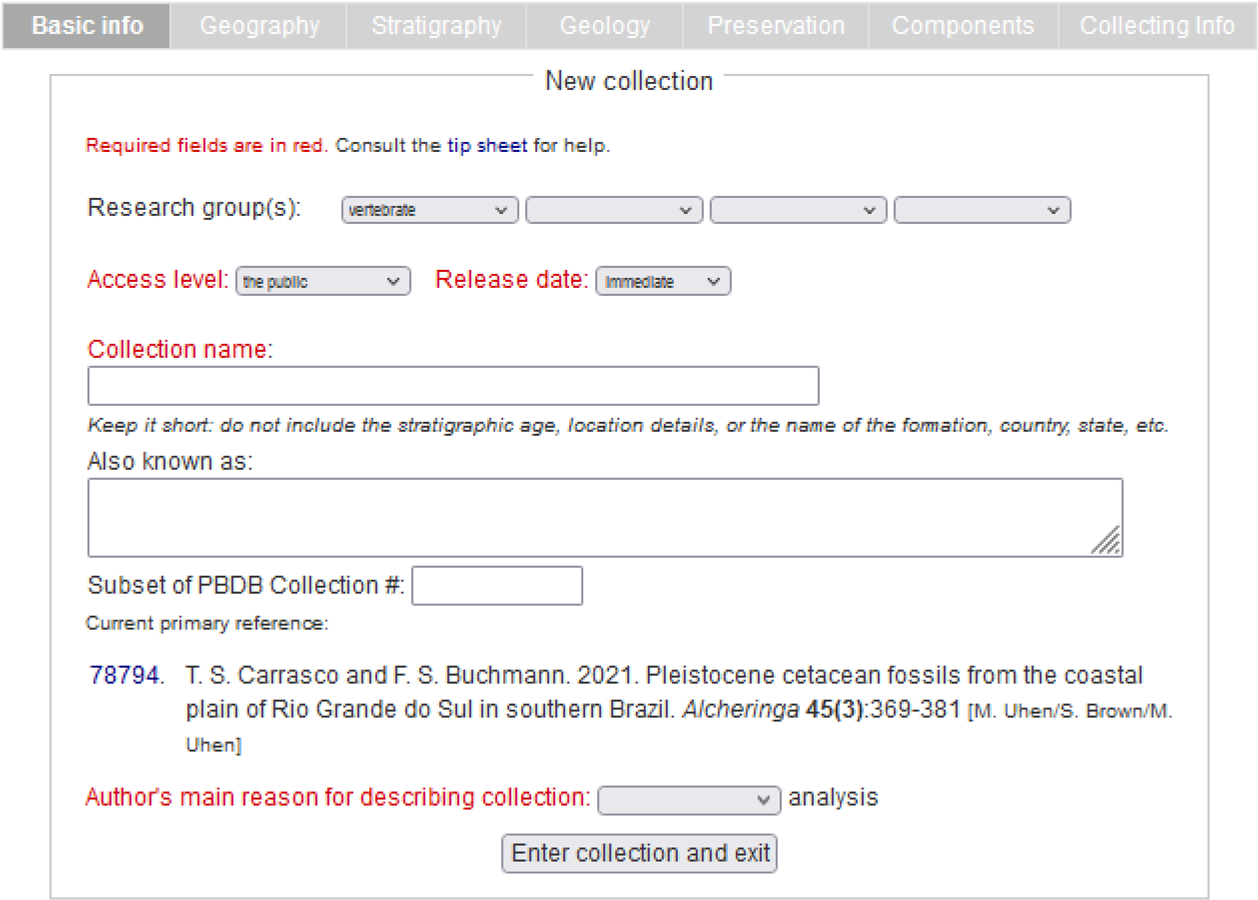
In general, it is advisable to examine any existing collections that might be similar to the material described in a new reference, becoming familiar with the practices that have been used by previous data contributors.

Here we will show you how to enter a collection into the PBDB, and how to edit collections that have already been entered into the PBDB by previous users.

### 3.7.1 Add a new collection

To add a new collection, click on *Add a new collection* in the Main Menu under Basic Steps. Please refer to the Data Entry Checklist, in section 7.1, Appendix 1, IV, Collection. This will bring up the *Duplicate collection search form*. Filling in this form accurately will help you determine whether the collection you want to enter has already been entered previously. First, use the *Period* pull-down to select the time period that includes your collection. If you do not see the name of the time period of your collection, consult the geologic time scale to see what period on the list includes the relevant age. Next, enter the *latitude* and *longitude* of your collection. You can either enter degrees, minutes, and seconds OR degrees + decimal degrees. Be sure to also use the *Direction* pull-down to select the correct quadrant of the earth's surface.

Hopefully, your reference includes geographic coordinates for your collection, but often they do not. We highly recommend GEOLocate (Rios and Bart 2010) to help you find the coordinates of your collection if you do not have them listed in the reference. After you enter a *Period* and *Latitude* and *Longitude* (including *Directions*),



Required fields are in red. Consult the tip sheet for help.

Research group(s): vertebrate

Access level: the public Release date: immediate

Collection name:

*Keep it short: do not include the stratigraphic age, location details, or the name of the formation, country, state, etc.*

Also known as:

Subset of PBDB Collection #:

Current primary reference:

78794. T. S. Carrasco and F. S. Buchmann. 2021. Pleistocene cetacean fossils from the coastal plain of Rio Grande do Sul in southern Brazil. *Alcheringa* **45(3)**:369-381 [M. Uhen/S. Brown/M. Uhen]

Author's main reason for describing collection: analysis

**Figure 20.** Basic info tab of the collection entry form.

look carefully at the map to see that your coordinates are in the correct place on the map. If not, check your *Latitude*, *Longitude* and *Directions* to make sure they are correct. Then, click <u>Search collections</u> to see if there are any potential matches in the PBDB.

A list of potential matches within 100 kilometers of your coordinates is then displayed, sorted by proximity. If there are none, you will be taken directly to the Collections entry form. If there is a list, look carefully at them to see if any are a match to your collection. If so, you can click on the collection number on the left to look more closely, and potentially add your fossil occurrences to the previously entered collection (see 3.8 *Occurrence Functions* below). Because individual collections are the fundamental unit of the PBDB, you should only add to a collection if you are confident your reference's fossils were collected at the same locality, stratigraphic unit, and facies (depositional environment). If there is a list of potential matches, but none of them actually match as the same collection, click *Add a new collection* at the bottom of the list.

Figure 20 shows the first page of the Collection entry form. The form is long and divided into separate tabbed pages listed in gray across the top: *Basic info*, *Geography, Stratigraphy, Geology, Preservation, Components,* and *Collecting Info.* While we encourage you to enter as much data as possible, at as fine a resolution as possible, in this form, remember that only the fields in red are required. (Many of the tabs also have hyperlinks to pop-up tip sheets that can offer hints or additional information to guide your data entry.) Let's look at each of the tabs in order, starting with *Basic info*.

### Basic info

On the **Basic info** tab, the first set of fields you can enter are the Research group(s) fields. You can use these pull-downs to select up to four research groups that you think might be interested in this data, or that you are a member of.

The second set of fields lets you set the *Access level* and the *Release date* for the data you are about to enter. You can use the Access level pull-down to set the access level to the *public, database members, group members,* or *authorizer only*. You can also embargo your data for up to five years using the *Release date* pull-down. After the release date, the data are automatically released to the public. We ask that you select "public" for *Access level* and "immediate" for *Release date* unless you have strong reasons to keep the data more restricted. If you need to embargo data, it is best to choose "database members," otherwise the data will be completely invisible to other contributors, potentially leading to duplication of collections.

The *Collection name* field is where you can give your collection a name. Try to use something geographic and descriptive, not something about the fossil occurrences themselves. For instance, "Thompson Farm" is a good choice as it is the name of a place. "Mastodon fossil site" is not a good choice as it is about the occurrences, not the place the fossils were found. Sometimes fossil collection sites have numbers or codes as well. You could use this number either in the *Collection name* field or you can give it a name and put the number in the *Also known as* field. This second field can also be useful for alternative spellings, place names in alternative languages, or to type the location name without diacritical marks (for easier searching).

If this new collection is a spatiotemporal subset of a more general collection, enter the collection ID number of this larger collection in the Subset of PBDB collection number field. For example, if you have a new bulk sample from a roadcut while an older, previously entered paper already described the general fauna from the roadcut, you can create a new collection for the bulk sample and enter the ID number of the older, general faunal collection here. However, we strongly prefer to avoid entering both general collections and subset collections. So, try to avoid creating both general and subset collections if they largely represent the same physical acts of collection. In other words, don't make two collections if the taxonomic lists pertain to mostly or entirely the same actual specimens, or if a new reference describes specimens that were part of the original collection but not described in the original paper. In those cases, record the small-scale collection and just record some comments mentioning any stray occurrences you can't nail down to a specific hole in the ground.

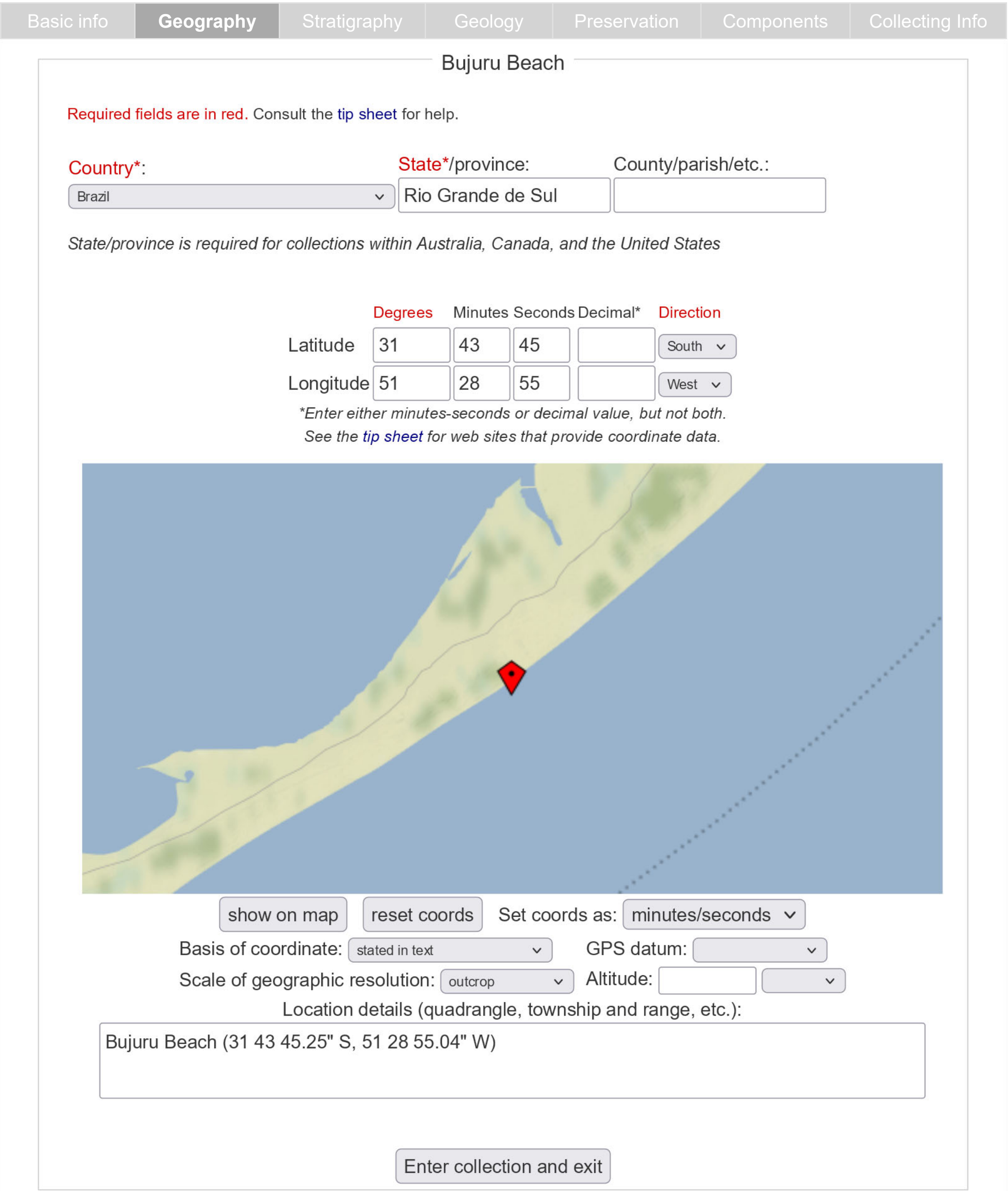
The *References* that should be associated with a collection are computed for you by the system. Whenever you add occurrences to a taxonomic list while using a reference, that reference becomes associated with the collection. If you edit a previously entered collection, you need to manually add your reference to the list by checking the checkbox next to your reference at the bottom of the Basic info form. We track one reference as the *primary reference*, while others will be identified as *secondary references*. The *primary reference* is the first to be published that (1) appears in a book or peer-reviewed journal (if earlier references were in gray literature); (2) gives at least one taxonomic occurrence; and (3) gives key data including the geographic location and geological time interval of the collection. When more than one reference is associated with a collection, you can change the one that is designated as the primary reference by clicking the radio button.

The *Author's main reason for describing collection* is a required field, describing the intent or subject area for which the collection was studied. Be careful because different parts of one paper might have been written for different purposes. When in doubt, use the category that best describes the whole paper. Your choices on this pull-down include:

- *biostratigraphic*: described for the purpose of temporal correlation, or in the course of a geological study (e.g., lists given within descriptions of measured sections).
- *paleoecologic*: described for the purpose of paleoecological analysis, for studying the functional morphology of fossils, or for inferring paleoenvironments, including taxonomic lists that are given without descriptions of fossils and without discussions of temporal correlation.
- *taphonomic*: described for the purpose of investigating taphonomy in the section.
- *taxonomic*: a systematic paleontology treatment, including simple descriptions of fossils. Use this if all the taxa are included in a systematic paleontology section, or if the taxonomic list is given only in support of an otherwise taxonomic account of one or a few taxa.
- *general faunal/floral*: there is no detailed discussion of the age, environment, ecology, preservation, or morphology of the fossils, but a taxonomic list is given. Use this category only if the other ones are too specific.

## Geography

The **Geography** tab (Figure 21) includes information on the modern geographic location of the collection on



Bujuru Beach

Required fields are in red. Consult the tip sheet for help.

Country*: Brazil
State*/province: Rio Grande de Sul
County/parish/etc.:

*State/province is required for collections within Australia, Canada, and the United States*

|  | Degrees | Minutes | Seconds | Decimal* | Direction |
|---|---|---|---|---|---|
| Latitude | 31 | 43 | 45 |  | South |
| Longitude | 51 | 28 | 55 |  | West |

**Enter either minutes-seconds or decimal value, but not both.*
*See the tip sheet for web sites that provide coordinate data.*

Set coords as: minutes/seconds
Basis of coordinate: stated in text
GPS datum:
Scale of geographic resolution: outcrop
Altitude:
Location details (quadrangle, township and range, etc.):
Bujuru Beach (31 43 45.25" S, 51 28 55.04" W)

**Figure 21.** Geography tab of the collection entry form.

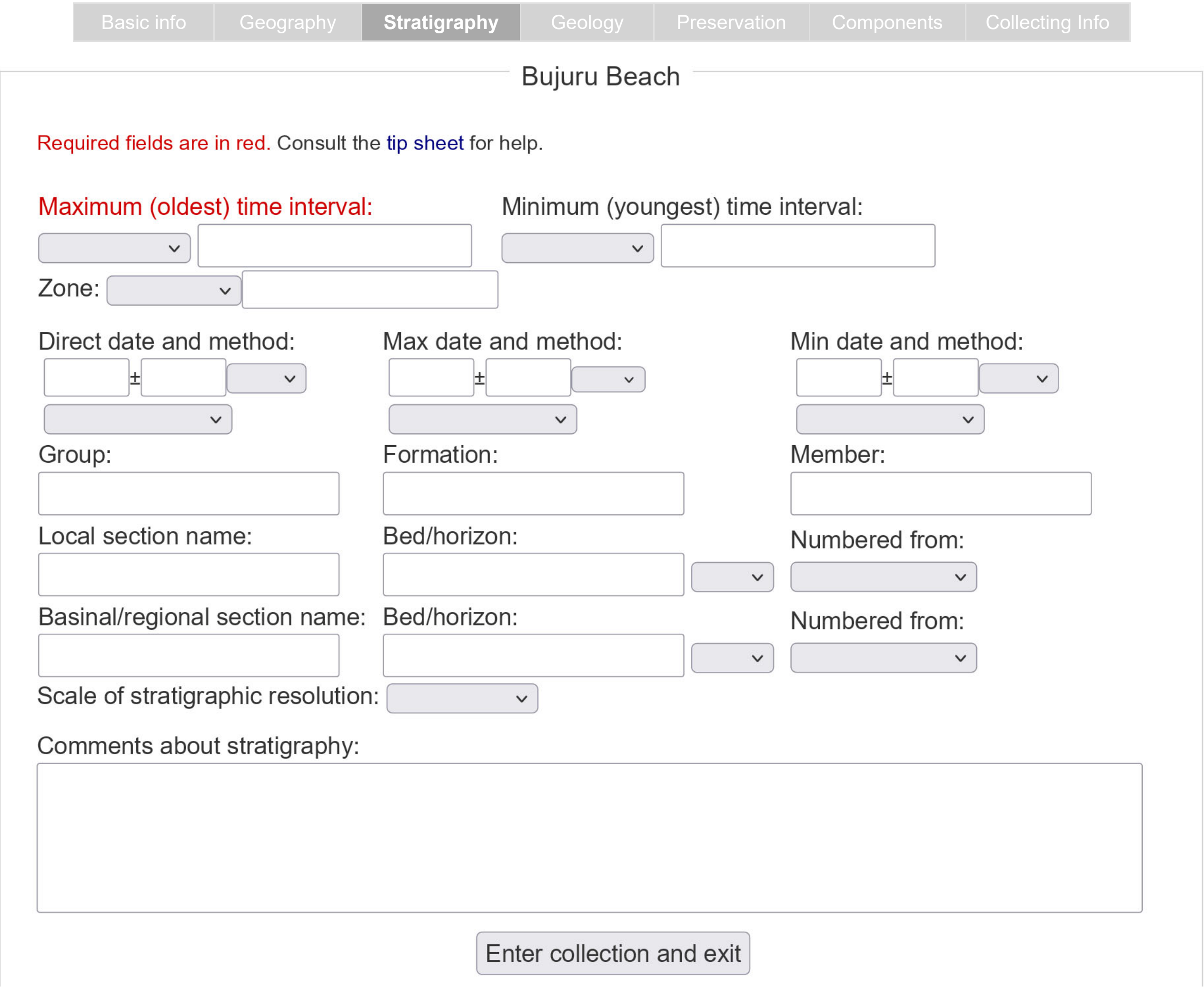
the surface of the planet. The *Latitude* and *Longitude* will already be filled in from when you searched using them previously. You are required to enter the *Country* using the pull-down. *State/province* is also required for Australia, Canada, and the United States. *County/parish* is also available, but not required, for the next largest geographic region. When entering old publications, you may need to do a bit of sleuthing to update geographic place names. Many countries, states, prefectures, counties, etc. have changed their names and/or boundaries since the time of publication. It's always good practice to try to enter the best data on modern geography, and then put the verbatim description of the place the fossils were collected in the Locality Details field.

Below the map are a series of optional fields. It is always good to fill in the *Basis of coordinate* field to indicate how you derived the coordinates for the collection. Your options on this pull-down include:

- *stated in text*: the latitude and longitude are provided explicitly in the text
- *based on nearby landmark*: the coordinate is based on that of a town, mountaintop, river, etc. that is described as being near to the fossil locality
- *based on political unit*: based on the midpoint, or any other point falling within, a county, state, province, country, etc.
- *estimated from map*: a coordinate estimated from a point on a locality map within the reference, or read from a map that includes latitude/longitude markings, including coordinates based on township and range data
- *unpublished field data*: the coordinate was taken from a museum database, unpublished field notes, or personal communication with the author(s)

Use the *Scale of geographic resolution* field to select the resolution that best characterizes the size of area from which the fossils in the collection were found. Note that this is NOT the extent of uncertainty in the placement of the coordinate (although these two things can be related to each other). It's best to err on the side of conservatism: if you know the collection is from sections along a road but are unsure as to whether this is from a single road cutting or several, it is better to use "local area" rather than "outcrop". If no information is given, do not make a selection. Choices on this pull-down include:

- *small collection*: from an area < 10 x 10 m; e.g., a single fossil, an individual bulk sample, a sampled bed within a section, or a small quarry
- *outcrop*: from an area < 1 x 1 km; e.g., a large quarry, road cut, or wash



Required fields are in red. Consult the tip sheet for help.

Maximum (oldest) time interval:
Minimum (youngest) time interval:
Zone:
Direct date and method: ±
Max date and method: ±
Min date and method: ±
Group:
Formation:
Member:
Local section name:
Bed/horizon:
Numbered from:
Basinal/regional section name:
Bed/horizon:
Numbered from:
Scale of stratigraphic resolution:
Comments about stratigraphy:

**Figure 22.** Stratigraphy tab of the collection entry form.

- *local area*: from an area < 100 x 100 km; e.g., approximately a one-by-one degree lat/long rectangle or the size of most individual counties in the USA
- *basin*: from any area > 100 x 100 km

Note that multiple collections may be made from a single locality. For example, if the author has sampled several beds at an outcrop or along a stratigraphic section, each sample may be considered a collection and each should be entered as a separate Collection record, even though the geographic information is the same for each. This situation is common in the biostratigraphic literature. In such cases, the PBDB has tools for stream-lining the entry of such collections. Specifically, at the bottom of any given Collection record, there is an option to "Add a collection copied from this one". Selecting this will populate a new collection record with all of the information from the previous record, and all the user needs to do is make any modifications that are specific to the new collection (e.g., meterage in section, lithology, taphonomic details, etc). Note that this option will NOT copy the taxonomic occurrence list. Another way to streamline data entry is to modify your Data Entry Settings in your profile (see section 3.12 below for more information).

Fill in the GPS datum field if it is stated in text. If calculated by selecting a point using the PBDB map interface, GEOLocate, or other digital maps, the datum will be WGS84. If the stated datum is not in the dropdown menu, make note of it in the location details.

The *Altitude* field is optional, but it is strongly recommended for late Quaternary data. Enter Altitude value

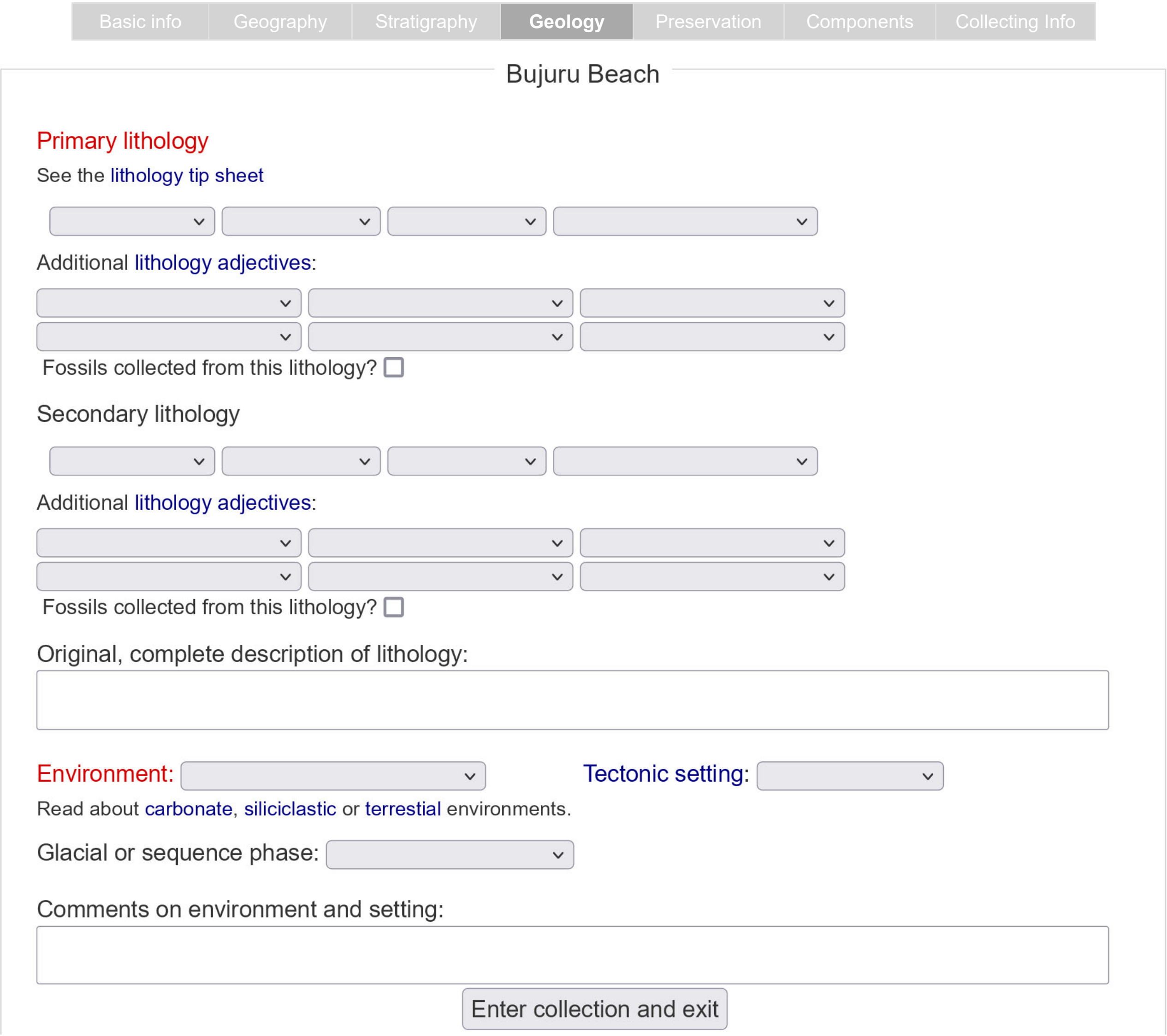


**Figure 23.** The Geology tab of the collection entry form.

and use the pull-down to select the units of measurement, meters, or feet.

It is also best practice to fill in the exact text from the reference describing the geographic location from in the *Location details* field. If you have made any assumptions or choices about where to place the coordinate (for example, because the information provided in the reference wasn't clear) it can also be useful to briefly mention this here.

### Stratigraphy

The **Stratigraphy** tab (Figure 22) includes fields about the age of the collection and the rock unit(s) the fossils are extracted from. You are required to enter at least one time period at the top of the form. If there is only one, enter it in the *Maximum* field. If there is a range, put the older time period in the *Maximum* field and the younger time period in the *Minimum* field. Both the *Maximum* and *Minimum* fields have a modifier pull-down to the left that further restricts the time period (only applicable to periods, Cenozoic epochs, and Jurassic, Cretaceous, and some Devonian stages). For instance, if your collection is from the Early Miocene and cannot be resolved to stage level, enter Miocene in the *Maximum* field, and use the pull-down on the left to choose Early/Lower. Terms from a variety of different timescales can be used. The database is populated with a list of authenticated time intervals it expects to find in these fields and will autofill the name as you type. When you submit your collection, it is not uncommon to get a warning that the time interval provided was not recognized by the database. In this case, write the actual interval reported in the reference in the Comments field and try using a correlated time interval instead (such as one found in Gradstein 2012).

Next, you can enter a biozone in the *Zone* fields if it is listed in your reference. Enter the name of the biozone in the blank part of the field, and the nature of the biozone using the pull-down to the left. If you have more complex data (for example, more than one biozone designation) than is allowed by the fields presented, be sure to enter it into the Comments about stratigraphy field.

The next set of fields allow you to enter numerical dates, if available. Remember that this should be the date of the unit in question, so if a different unit has been used to infer the age of the unit which the fossils came from, this should be explained in the Comments section. Use the *Direct date* fields if you have a single date from the fossiliferous bed, and the *Max* and *Min* date fields if you have a range with a maximum and minimum age from underlying and/or overlying beds. Be sure to choose the temporal units and method of dating if you use these fields.

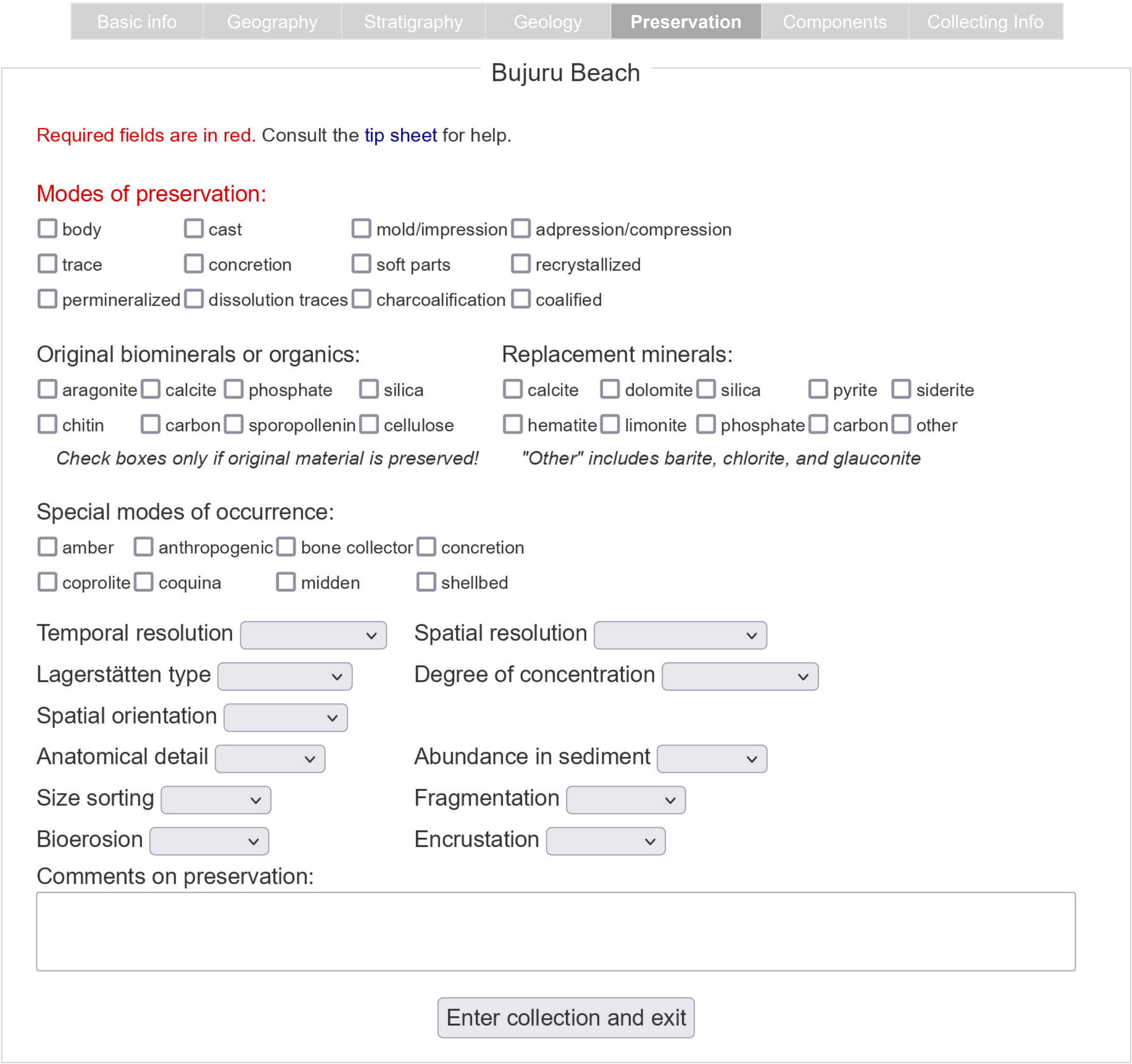
The next set of fields deal with the named stratigraphic units that the collection is from. *Group*, *Formation*, and *Member* are for formal stratigraphic names for rock units. Do not include the words "Formation," "Member," etc. in the unit name (e.g., the Hamilton Group should be entered just as "Hamilton" in the Group field). You can write the name of the rock type if it is part of the stratigraphic unit's name (e.g., Wheeler Shale). Use any and all of these that apply to your collection. If your reference includes information on *Local section* or *Basinal/regional* sections, use the next set of fields to enter these data as appropriate. *Local section* refers to collections made from individual beds within a measured stratigraphic section or core, and you can enter either a bed number or height/depth for *bed/horizon*. *Basinal/regional section* is very rare but can be used for units within a sequence stratigraphic scheme, or for intervals within a long-distance correlational scheme based on geochemical matching of ash fall horizons.

Use the *Scale of stratigraphic resolution* to indicate the scale of the stratigraphic unit that includes the fossils. Again, this is NOT uncertainty in the stratigraphic placement of the fossils. For example, if the collection is from more than one member of a formation, then pick "formation"; if it is from throughout a measured section that spans the exposure of a full member, then pick "member," but if it is from just part of the member then pick "group of beds"; if it is from two thin beds that are immediately superposed in an outcrop, pick "group of beds" and not "bed." If the fossils are described as coming from a single bed but which bed in the member is not stated, then pick "bed", not "member". If you are unsure, don't make a selection.

If you are entering older literature where you believe the stratigraphic data may be outdated, enter the data as presented in the paper. Then find an updated reference that can be entered as a secondary reference (see below) and update the record as appropriate. Updates to stratigraphy can be found by use of the National Geologic Map Database, Macrostrat, or other regional or country based geological surveys and literature searches.

### Geology

The **Geology** tab (Figure 23) includes information on the rocks and environment of deposition for the collection. The *Primary lithology* fields are for the dominant rock type of the sediments for the collection. Each collection can have two rock types, so these same fields



Bujuru Beach

Required fields are in red. Consult the tip sheet for help.

Modes of preservation:

☐ body ☐ cast ☐ mold/impression ☐ adpression/compression
☐ trace ☐ concretion ☐ soft parts ☐ recrystallized
☐ permineralized ☐ dissolution traces ☐ charcoalification ☐ coalified

Original biominerals or organics:

☐ aragonite ☐ calcite ☐ phosphate ☐ silica
☐ chitin ☐ carbon ☐ sporopollenin ☐ cellulose

*Check boxes only if original material is preserved!*

Replacement minerals:

☐ calcite ☐ dolomite ☐ silica ☐ pyrite ☐ siderite
☐ hematite ☐ limonite ☐ phosphate ☐ carbon ☐ other

*"Other" includes barite, chlorite, and glauconite*

Special modes of occurrence:

☐ amber ☐ anthropogenic ☐ bone collector ☐ concretion
☐ coprolite ☐ coquina ☐ midden ☐ shellbed

Temporal resolution
Spatial resolution
Lagerstätten type
Degree of concentration
Spatial orientation
Anatomical detail
Abundance in sediment
Size sorting
Fragmentation
Bioerosion
Encrustation
Comments on preservation:

**Figure 24.** Preservation tab of the collection entry form.

are duplicated below under *Secondary lithology*, in case there is a secondary rock type. The easiest way to fill in these fields is to start in the upper right corner and use the pull-down to select the primary rock type. This pull-down includes a long list of rock names in several genetic groups. If you see the one from your reference, just select it using the pull-down. If you do not see your rock type name, use the lithologies tip sheet to look for synonyms of your name. If you are still not sure, use reliable sources to try and find synonyms of the rock name that you have. Try to be as specific as you can, based on what is reported in the reference; *grainstone* is more informative than *limestone*, which is better than *carbonate*. Often, your reference will not include any information regarding the rocks your collection is from. If this is the case, select the first item on the list, *not reported*.

To the left of this rock type field are three fields that include adjectives that describe the primary rock type further. Underneath are six *Additional lithology adjectives* pull-downs that allow you to specify even more details about the rocks. Review the lithology adjectives carefully before making selections; they include several different classes of information.

Beneath the *Additional lithology adjectives* is a check box stating: *Fossils collected from this lithology?* Check this if there is specific information stating that the fossils in this collection came from this lithology. If both lithologies definitely include fossils, check both boxes.

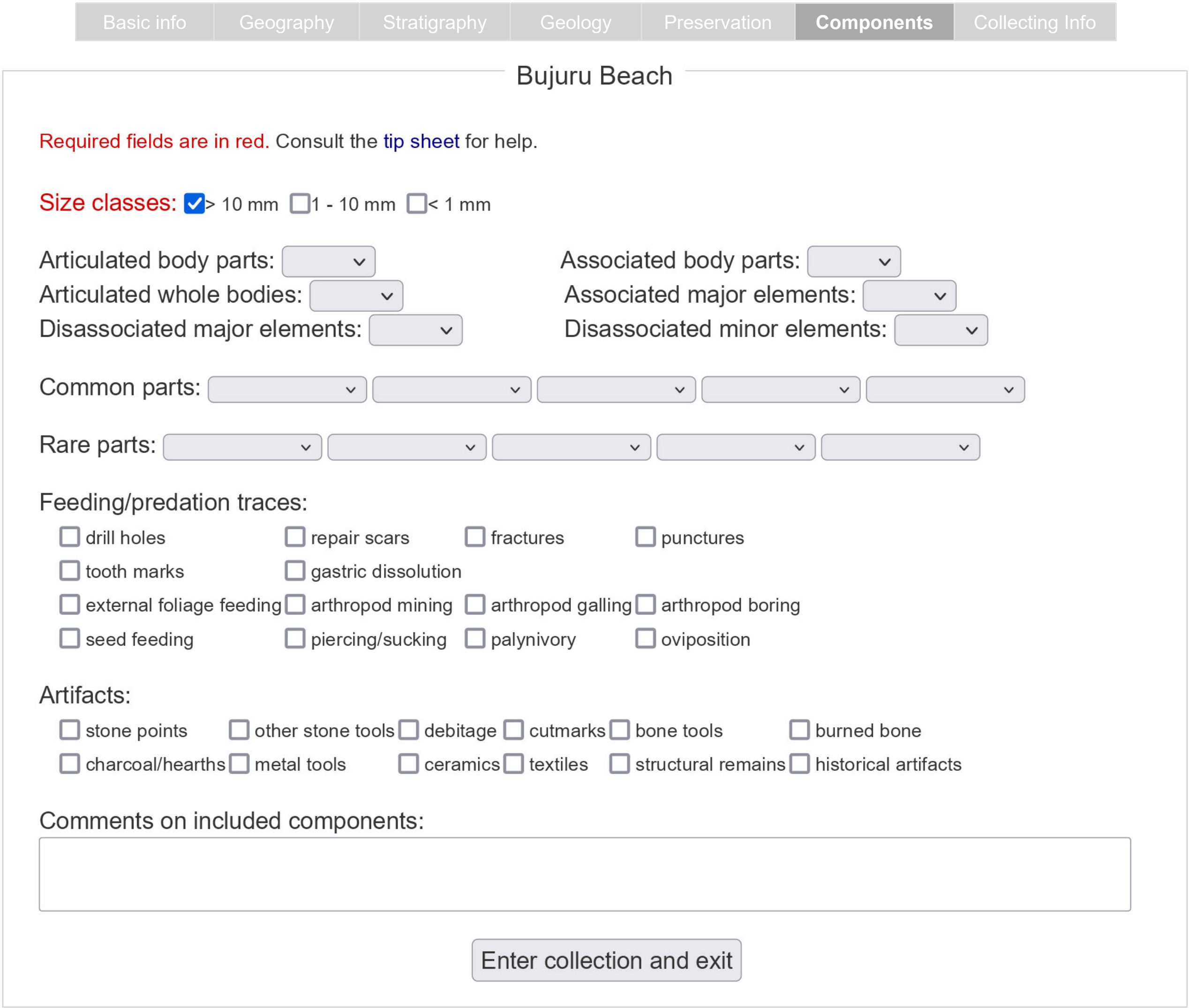
Use the *Original, complete description of lithology* field to quote (or paraphrase, if very long) the entire description of the lithology.

The *Environment* pull-down is used to select the most precise environmental category you can attribute for the collection. Each section of the pulldown includes some generalized "indet." terms that you should use only if a more precise determination is not possible. The most general environment you must choose is either; *marine indet.* or *terrestrial indet*; remember that *terrestrial indet.* includes fluvial and lacustrine environments, so fossils from aqueous settings should not necessarily be assumed to be marine. There are some environments that are more specific but still encompass several categories, such as *carbonate indet.* (any carbonate environment), *coastal indet.* (siliciclastic marine from continental shelf environments, but excluding deep-water, submarine fan, or slope/basin settings) and *fluvial-lacustine indet.* (any river or lake environment). Finally, there are the most specific categories of environment, such as *shoreface* (siliciclastic above fair-weather wave base), *open shallow subtidal* (carbonate above fair-weather wave base), or *pond*. Papers focusing on systematic paleontology often do not provide details about environment, and it is only sometimes possible to infer environment from the lithology description.

Use the *Tectonic setting* pull-down to indicate the tectonic setting of the collection if it is stated in the reference.

Use the *Glacial or sequence phase* field to indicate the phase of a glacial cycle (for terrestrial, generally Pleistocene collections) or sequence stratigraphic cycle (for marine collections) if stated in the reference.

The hyperlinked pop-ups for lithology tip sheet,



Bujuru Beach

Required fields are in red. Consult the tip sheet for help.

Size classes: ☑ > 10 mm ☐ 1 - 10 mm ☐ < 1 mm

Articulated body parts:
Articulated whole bodies:
Disassociated major elements:
Associated body parts:
Associated major elements:
Disassociated minor elements:

Common parts:

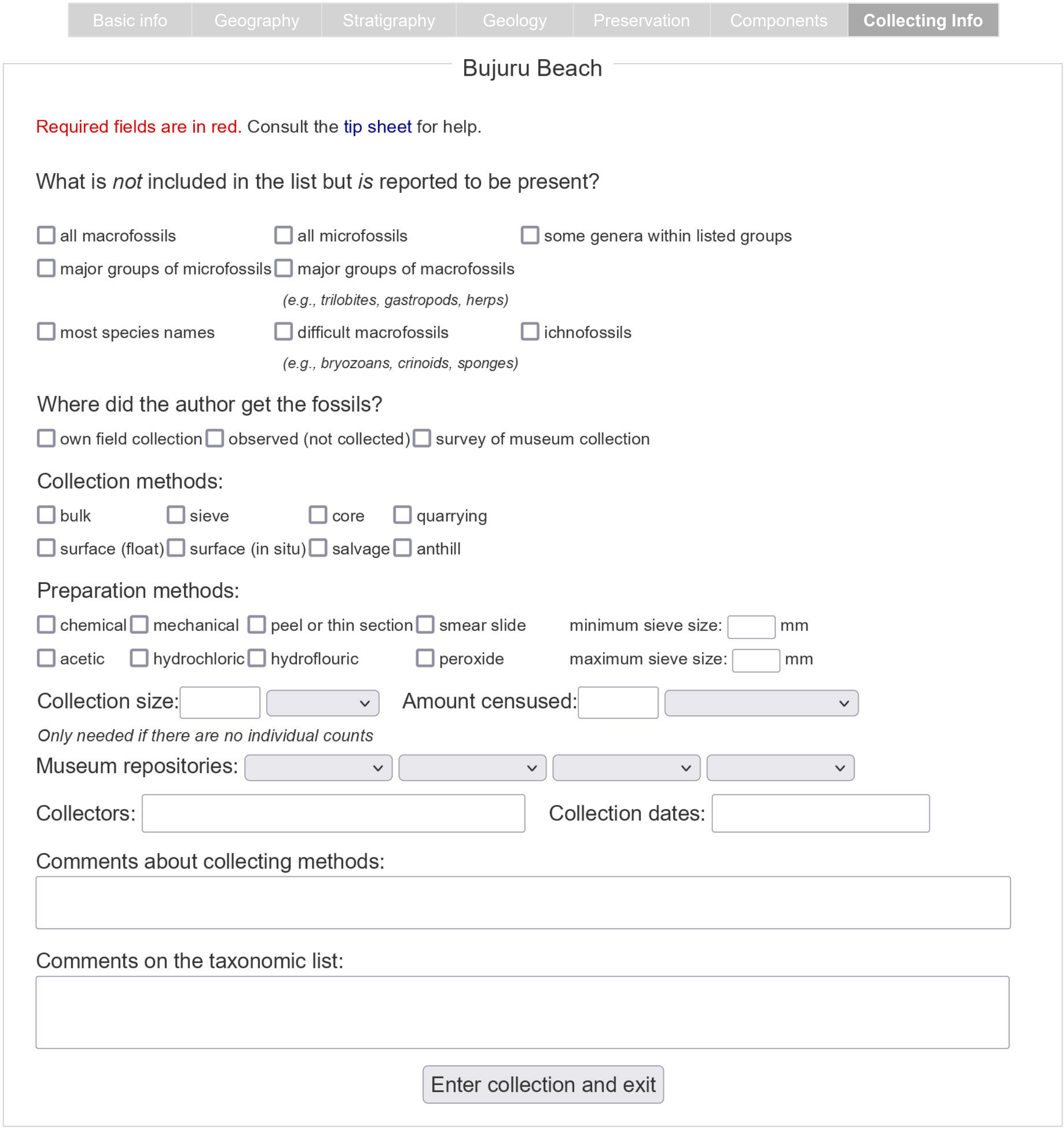
Rare parts:

Feeding/predation traces:
☐ drill holes ☐ repair scars ☐ fractures ☐ punctures
☐ tooth marks ☐ gastric dissolution
☐ external foliage feeding ☐ arthropod mining ☐ arthropod galling ☐ arthropod boring
☐ seed feeding ☐ piercing/sucking ☐ palynivory ☐ oviposition

Artifacts:
☐ stone points ☐ other stone tools ☐ debitage ☐ cutmarks ☐ bone tools ☐ burned bone
☐ charcoal/hearths ☐ metal tools ☐ ceramics ☐ textiles ☐ structural remains ☐ historical artifacts

Comments on included components:

**Figure 25.** The Components tab of the collection entry form.



Bujuru Beach

Required fields are in red. Consult the tip sheet for help.

What is *not* included in the list but *is* reported to be present?

☐ all macrofossils ☐ all microfossils ☐ some genera within listed groups
☐ major groups of microfossils ☐ major groups of macrofossils *(e.g., trilobites, gastropods, herps)*
☐ most species names ☐ difficult macrofossils *(e.g., bryozoans, crinoids, sponges)* ☐ ichnofossils

Where did the author get the fossils?

☐ own field collection ☐ observed (not collected) ☐ survey of museum collection

Collection methods:

☐ bulk ☐ sieve ☐ core ☐ quarrying
☐ surface (float) ☐ surface (in situ) ☐ salvage ☐ anthill

Preparation methods:

☐ chemical ☐ mechanical ☐ peel or thin section ☐ smear slide minimum sieve size: ☐ mm
☐ acetic ☐ hydrochloric ☐ hydroflouric ☐ peroxide maximum sieve size: ☐ mm

Collection size: Amount censused:

*Only needed if there are no individual counts*

Museum repositories:

Collectors: Collection dates:

Comments about collecting methods:

Comments on the taxonomic list:

**Figure 26.** Collecting Info tab of the collection entry form.

lithology adjectives, and depositional environments can offer informative hints and context when entering these fields.

Use the *Comments about environment and setting* field to provide any additional information, including extensive quotes from the paper, that you used to determine these values.

## Preservation

The **Preservation** tab (Figure 24) records how the fossils in the collection have been preserved.

*Modes of preservation* is the only required field on this tab, so at least one of the boxes must be checked in this group of checkboxes. The boxes record the kind of fossils that are present (body, cast, etc.) and the mineral composition of the fossils. Check as many of them as apply to any of the fossils in the collection. Impressions, casts, and molds are fossils with none of the original biomineralized or organic tissue preserved.

A mold is an impression of a fossil in the rock or sediment matrix that encased the original organism. Often the term “mold” is used for 3D impressions and “impression” is used for 2D impressions (e.g., leaf impression). The adpression/compression checkbox can be used for flattened fossils that retain some original or altered organic material; these are most common in fossil plants and insects. A cast is a 3D fossil that is a positive image of the original, often formed by the infilling of a shell. This object is also called a steinkern. In the PBDB, these and other terms refer to the natural outcomes of preservational processes, rather than artificial copies of fossils.

Use the additional sets of checkboxes for *Original biominerals or organics, Replacement minerals,* and *Special modes of occurrence* to capture additional preservation information, if reported (the concretion mode appears in two separate checkboxes but are stored in one field, so you only need to choose one).

Use the additional pull-downs at the bottom of the form to add additional information regarding the suite of fossils preserved in the collection.

Finally, *Comments on preservation* is an open field to include any taphonomically relevant information and/or interpretations of the author(s) concerning the processes responsible for assemblage formation. Quote extensively because many preservational categories are qualitative, so the values you enter may need to be reassessed later.

### Components

The **Components** tab (Figure 25) is used to capture additional information regarding the fossil elements in the collection. The only required field is the first one: *Size classes*. Check all that apply to the fossils in the collection.

Use the first set of six pull-downs to describe the degree of articulation and association of the parts of individual organisms in the collection.

Use the *Common parts* and *Rare parts* pull-downs to list the common and rare parts of individual organisms in the collection.

Use the *Feeding/predation traces* check boxes to indicate the presence of these trace fossils. Check all that apply to specimens within the collection based on either the original reference or other unambiguous source.

Some Pleistocene and Holocene collections are associated with human activity and are also associated with various *Artifacts*. An artifact is any object found in the same horizon as the fossils that was produced by humans. Use these check boxes to identify what artifacts are included in the collection along with the fossils.

Finally, use the *Comments on components* field to enter key additional information not captured in the other fields.

### Collecting Info

The **Collecting Info** tab (Figure 26) is used to capture information on who collected the fossils, when the fossils were collected, and where the fossils have been reposited. None of these fields are required.

Use *What is not included in the list but is reported to be present?* to indicate that some broad taxa were described as present in the outcrop or collections but weren't listed in more detail. The options are: all macrofossils, all microfossils, some genera within listed groups, major groups of microfossils, major groups of macrofossils, most species names, difficult macrofossils, ichnofossils. Check any or all that apply.

*Where did the author get the fossils?* indicates how the author of the reference acquired access to the fossils in the collection. The options are: own field collection, observed (not collected), and survey of museum collections. Check any or all that apply.

Use *Collection methods* to indicate how the fossils were originally collected. The options are: bulk, sieve, core, quarrying, surface (float), surface (*in situ*), salvage, and anthill. Check any or all that apply.

Use *Preparation methods* to indicate how the fossils were prepared in the laboratory. The options are: chemical, mechanical, peel or thin section, smear slide, acetic, hydrochloric, hydrofluoric, peroxide, and minimum and maximum sieve size. Check any or all that apply.

The *Collection size* and *Amount censused* fields are used to record the total number of specimens or individuals for the entire fossil collection (if known) or the amount of space surveyed for fossils.

The *Museum repositories* pull-downs allow you to indicate where the fossils are reposited (if known). If your museum is not listed, or you don’t recognize the acronym, put this information in the *Comments about collecting methods* field.

Use *Collectors* to list the names of the people who collected the fossils, including the authors if appropriate. Separate names by commas and do not write out "and." Example: J. Doe, J. Smith, J. Brown.

Use *Collection dates* to enter the year or years during which the fossils were taken from the field, as in 1940, 1940 - 1949, or 1940s.

Use the *Comments about collecting methods* field to provide as much additional information as possible.

Use the *Comments on the taxonomic list* field to give any information about the identifications, or inclusion and exclusion of taxa, that should not be entered as a comment about an individual occurrence.

### Finish Collection data entry

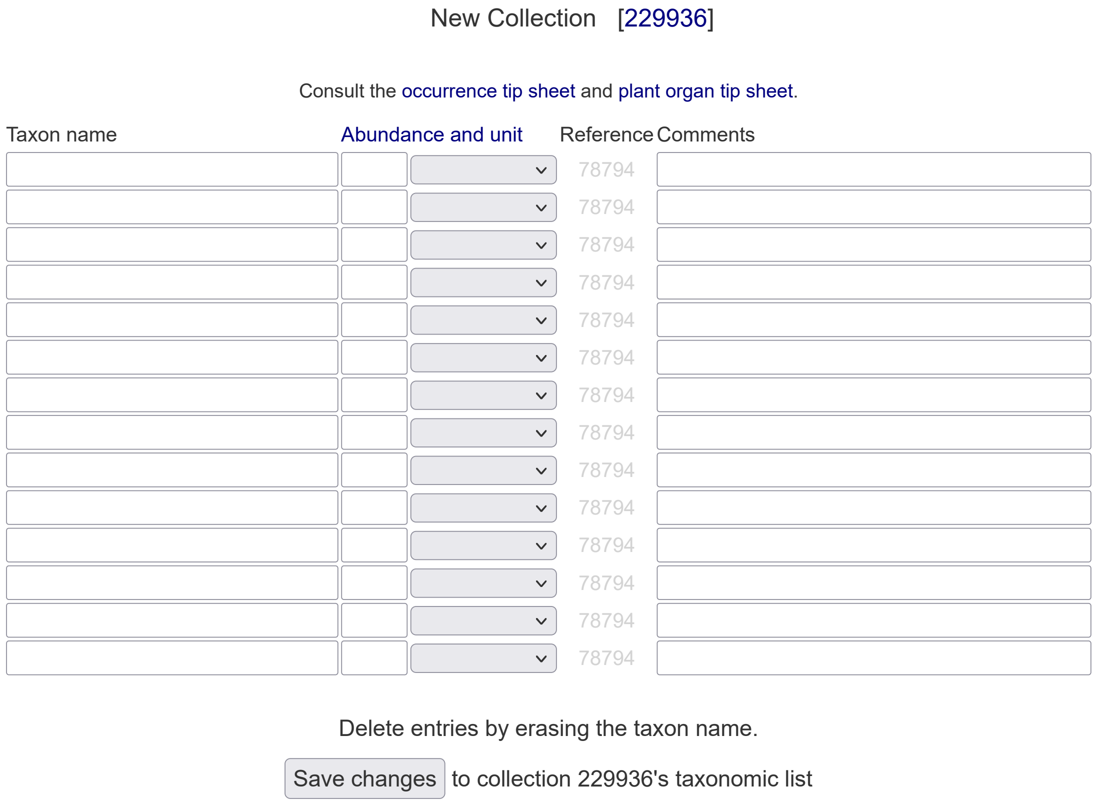
After you have filled in data for all of the tabs, it is a good idea to go back and click on each tab to proofread your data entry. Make sure all red fields are filled in, and that you have filled in as much on the Stratigraphy and Geology tabs as possible, including any content reported in the reference that is not required. When you are done, click Enter collection and exit at the bottom of any of the pages. The PBDB will either alert you to any missing required data or display a short menu of choices showing where you are likely to want to go next. (A common warning is that the listed time interval is not one of the authenticated ones used by the PBDB binning system; in this case, go back to the Stratigraphy tab and try an alternative one that correlates to the age listed in the reference. Make sure to write the interval reported in the reference in the stratigraphy comments field.) So far, we have only entered data on the collection itself but have not entered any data on the occurrences of fossil taxa in this collection, which is the whole point of this exercise. Thus, you will most likely want to click on Edit taxonomic list to do this next.

### 3.7.2 Edit an old collection

**Edit an old collection** lets you edit a collection that has been previously entered by another user. To use this function, go to the **Main Menu** and click on **Edit an old collection** under the **Collections and occurrences menu**. This will take you to the *Collection search* form where you can enter parameters to find the old collection. If you have the collection number, this is the best option because it is unique to the collection. If not, you can use almost any parameter including the name, geography, time, taxa in the occurrence list, etc.

After you enter your search parameter(s), click Search collections at the bottom of the page. If there are multiple matches, you will be shown a list of matches from which you can pick the collection that you want to edit by clicking on the collection number. If there is one match, you will be taken directly to the Collection entry screen, which will be filled in. Here, you can step through the tabs to alter the data previously entered or add new data to the collection.

If you are changing what had previously been entered



New Collection [229936]

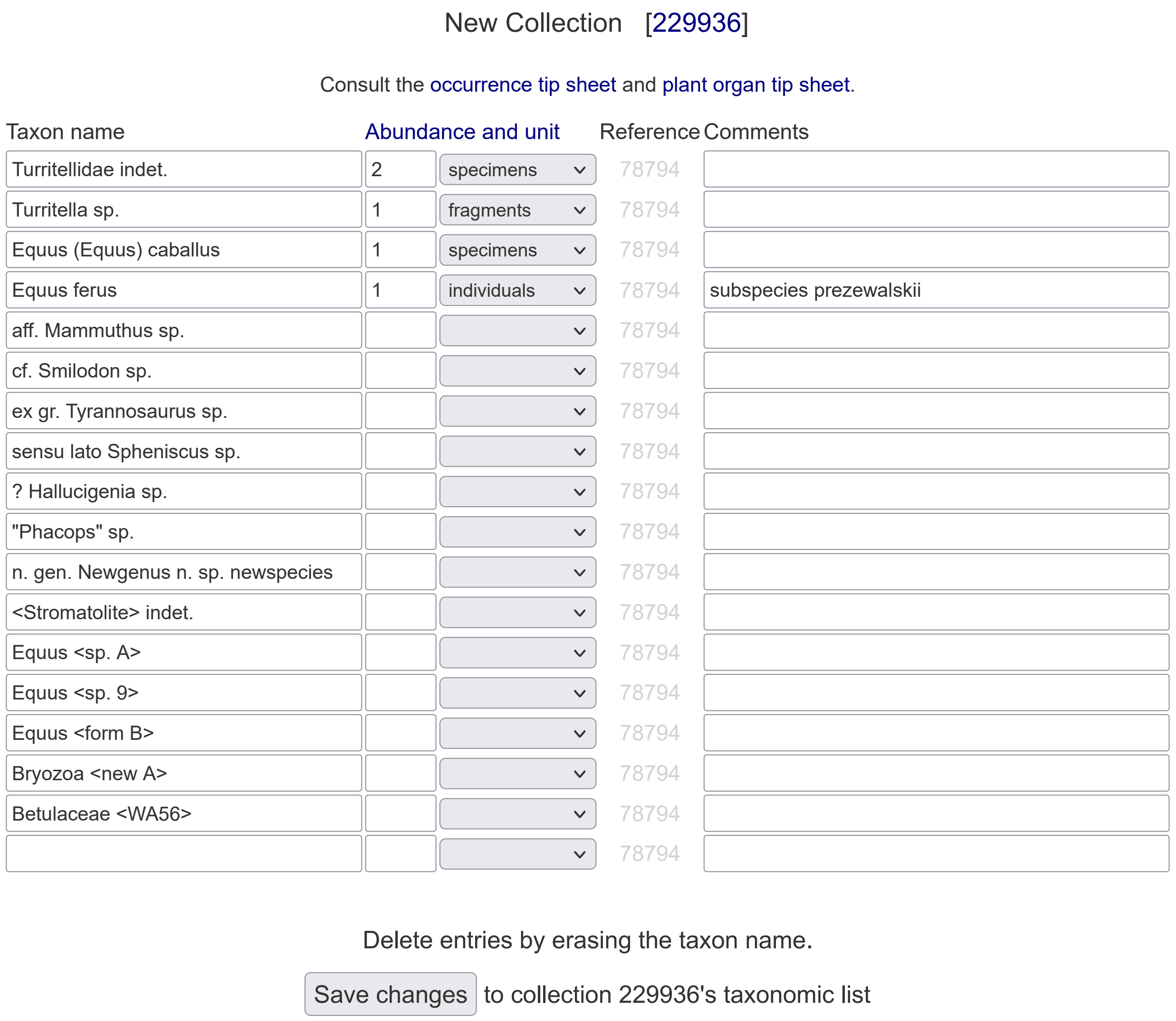
Consult the occurrence tip sheet and plant organ tip sheet.

| Taxon name | Abundance and unit | | Reference | Comments |
|---|---|---|---|---|
| | | | 78794 | |
| | | | 78794 | |
| | | | 78794 | |
| | | | 78794 | |
| | | | 78794 | |
| | | | 78794 | |
| | | | 78794 | |
| | | | 78794 | |
| | | | 78794 | |
| | | | 78794 | |
| | | | 78794 | |
| | | | 78794 | |
| | | | 78794 | |
| | | | 78794 | |

Delete entries by erasing the taxon name.

to collection 229936's taxonomic list

**Figure 27.** Occurrences entry form example. Yours may look different depending on your data entry preferences.



New Collection [229936]

Consult the occurrence tip sheet and plant organ tip sheet.

| Taxon name | Abundance and unit | | Reference | Comments |
|---|---|---|---|---|
| Turritellidae indet. | 2 | specimens | 78794 | |
| Turritella sp. | 1 | fragments | 78794 | |
| Equus (Equus) caballus | 1 | specimens | 78794 | |
| Equus ferus | 1 | individuals | 78794 | subspecies prezewalskii |
| aff. Mammuthus sp. | | | 78794 | |
| cf. Smilodon sp. | | | 78794 | |
| ex gr. Tyrannosaurus sp. | | | 78794 | |
| sensu lato Spheniscus sp. | | | 78794 | |
| ? Hallucigenia sp. | | | 78794 | |
| "Phacops" sp. | | | 78794 | |
| n. gen. Newgenus n. sp. newspecies | | | 78794 | |
| <Stromatolite> indet. | | | 78794 | |
| Equus <sp. A> | | | 78794 | |
| Equus <sp. 9> | | | 78794 | |
| Equus <form B> | | | 78794 | |
| Bryozoa <new A> | | | 78794 | |
| Betulaceae <WA56> | | | 78794 | |
| | | | 78794 | |

Delete entries by erasing the taxon name.

to collection 229936's taxonomic list

**Figure 28.** Occurrence entry form with examples of various forms of taxonomic names.

because of a data entry error, just go ahead and correct the mistake. If you are adding new information from a different reference, make sure to first choose that Secondary Reference using the Search functions, so that this reference will be linked to the collection. If you are updating something like the names of the stratigraphic units, the age of the collection, or the depositional environment because of new data, it's generally a good idea to note the old data in the comments fields found on the bottom of each page and replace the old data with the new data in the primary fields. That way, other users can follow when collections have been updated from their original data.

When you are done, click Edit collection and exit at the bottom of the page. If you are editing an older collection "reason for describing collection", "environment", "preservation mode" and "size class" may not have been entered, and you will receive a pop-up alert. Complete all of the fields listed in the pop-up and then click Edit collection and exit again.

### 3.8 Occurrence Functions

**Occurrences** of fossil taxa are perhaps the most fundamental data in the Paleobiology Database. These occurrences identify what taxa are found in what collections, which in turn link to information on when, where, and in what rock units they were found. Together, these data create the pattern of diversity of life over geologic time on earth.

To enter these occurrences, you can click on Edit taxonomic list on the brief menu displayed after you enter a collection, or you can click on Add taxonomic occurrences under **Basic Steps** OR Edit old occurrences under **Collections and occurrences** on the **Main Menu**. If you take the first path, you will be adding occurrences to the collection you just entered. If you take either path from the **Main Menu**, you will need to search for the correct collection to add your occurrences to. Any of these paths will take you to the **Occurrence entry** form where you can **Add taxonomic occurrences**.

### 3.8.1 Add taxonomic occurrences

Please refer to the Data Entry Checklist, in section 7.1, Appendix 1, V, Occurrences. Once you arrive at the **Occurrence entry form**, you will see the name and number of the collection at the top of the page. Figure 27 shows the **Occurrence entry form** for the collection called Newcollection, number 220130. Your **Occurrence entry form** might look different based on your Data Entry Settings. On this version, you see rows of fields from left to right: *Taxon name*, *Abundance*, *Unit*, *Reference*, and *Comments*. Each field is discussed in turn below. You may see other fields listed based on differences in your PBDB account settings.

### Taxon name

The **Taxon name** field is the heart of the *Occurrence entry form*. You just need to type or paste in the name of the taxon in this field. You should not use italics when entering these data. You can enter names of any rank or no rank. When you enter these names, please follow these guidelines:

- Enter the full species, genus or higher taxon name for every taxon tied to the collection, including algae, microfossils, trace fossils, etc.
- Always enter generically indeterminate taxa such as families, orders, and classes followed by the qualifier indet. (e.g., “Trilobita indet.”)
- If you enter a genus but the species is not specified, enter the genus name followed by the qualifier sp. (e.g., “Tyrannosaurus sp.”)
- Always capitalize genus or higher order names.
- Enter subgenus names in parentheses, as in Equus (Equus) caballus.
- If this collection is the type locality for a species, tag the name with “n. sp.” either before or after the species epithet.
- If the species is the type for its genus, tag the name with “n. gen.” either before the genus name or after the species epithet (only if used as part of “n. gen. n. sp.”).
- Put subspecies or variety names in the *Comments* field.

Examples of all of these types of data entry are shown in Figure 28.

Occurrences are often identified with taxon names that include various qualifiers that convey uncertainty or affinity. Any of the following qualifiers of genus and species names may be used by adding it before the taxon name: aff., cf., ex gr., sensu lato, ?, " ". Although the authors may use these qualifiers following a name, the database is standardized to expect them before the taxon name. See Glossary for definitions of qualifiers.

When you enter the occurrence of a taxon that represents the type species of a new genus, add the qualifier **n. gen**. Do not enter n. gen. if a species is new but it is not the type species of the genus. When you enter the occurrence of a taxon that represents the type species of a new subgenus, add the qualifier **n. subgen**. When you enter the occurrence of a taxon that represents the holotype occurrence of a new species, add the qualifier **n. sp**. Only enter n. sp. if the type specimen comes from this particular collection, not just any collection mentioned in the same paper. After the occurrence of the type specimen for a species is entered into a taxonomic list associated with a Collection record, the Collection record number will be automatically added to the Authority Form for that species name.

You may use only one qualifier each for the genus, subgenus, and species names, so if there is more than one, pick the most important. Below are several examples of taxon names with properly formatted qualifiers:

n. gen. Equus n. sp. caballus

aff. Equus ? caballus

sensu lato Equus sp.

cf. Equidae indet.

"Equus" sp.

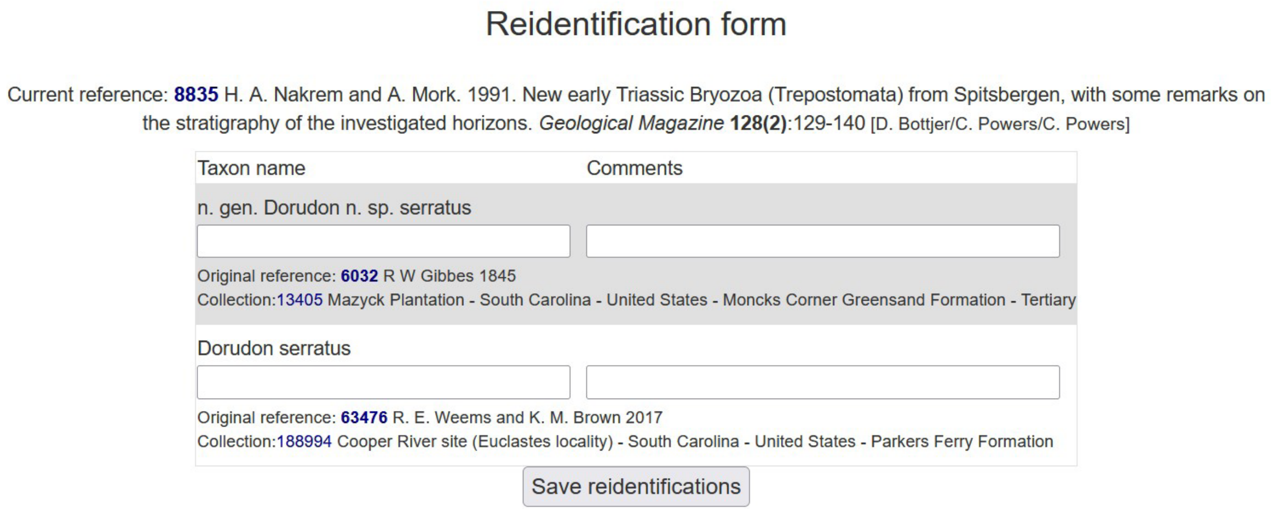
If an identification says "Xus cf. X. yus," ignore the "X." part. Repeating an abbreviation for the genus name is grammatically correct but unnecessary in our system.

Informal names (i.e., names that are not properly formed according to Linnean rules) may still be entered. Put them in < > symbols like this:

<Stromatolite> indet.

Equus <sp. A>

You can enter multiple unnamed but lettered or numbered species like this:



Reidentification form

Current reference: **8835** H. A. Nakrem and A. Mork. 1991. New early Triassic Bryozoa (Trepostomata) from Spitsbergen, with some remarks on the stratigraphy of the investigated horizons. *Geological Magazine* **128(2)**:129-140 [D. Bottjer/C. Powers/C. Powers]

Taxon name Comments

n. gen. Dorudon n. sp. serratus

Original reference: **6032** R W Gibbes 1845
Collection:13405 Mazyck Plantation - South Carolina - United States - Moncks Corner Greensand Formation - Tertiary

Dorudon serratus

Original reference: **63476** R. E. Weems and K. M. Brown 2017
Collection:188994 Cooper River site (Euclastes locality) - South Carolina - United States - Parkers Ferry Formation

**Figure 29.** Reidentification form example.

Equus <sp. A>

Equus <sp. 9>

Equus <form B>

Avoid entering informal English names for higher taxa. Translate names like "horse", "bryozoan", "coral", "crinoid", "shark", "fish" and "grass" into Linnean names like Equidae indet., Bryozoa indet., Anthozoa indet., Crinoidea indet., Chondrichthyes indet., Actinopterygii indet., and Poales indet.

Whenever possible, enter the Linnean name of the higher taxon and make the species name informal, not the other way around. For example, if the author listed "new bryozoan A," enter:

Bryozoa <new A>

Morphotype codes such as "WA56" should be entered along with higher taxon names in a similar way:

Betulaceae <WA56>

### Abundance

*Abundance* is how many of the taxon are reported present and can be given as counts, estimated number of individuals, categories (“Abundant”, “Common”, “Rare”, etc.), or percentages based on counts or other measures. The number of representations of the taxon should be typed in the *Abundance* field itself.

### Unit

The representation of things counted are in the *Unit* pull-down and can be specimens, individuals, fragments, etc.

### Reference

The *Reference* field is pre-filled with the reference number of the active reference for new data entry. Previously entered occurrence data will have the reference filled in, but it is also editable. It is best practice to change this to assign the type specimen of a taxon to the correct reference or to re-assign occurrences to a different reference.

### Comments

The *Comments* field is for any information about a particular occurrence that is not captured in any of the other fields but might be useful. For example, this could include specimen numbers if provided, subspecies name, anatomical parts included within the specimen(s), or further information on the abundance units.

### 3.8.2 Paste in an occurrence list

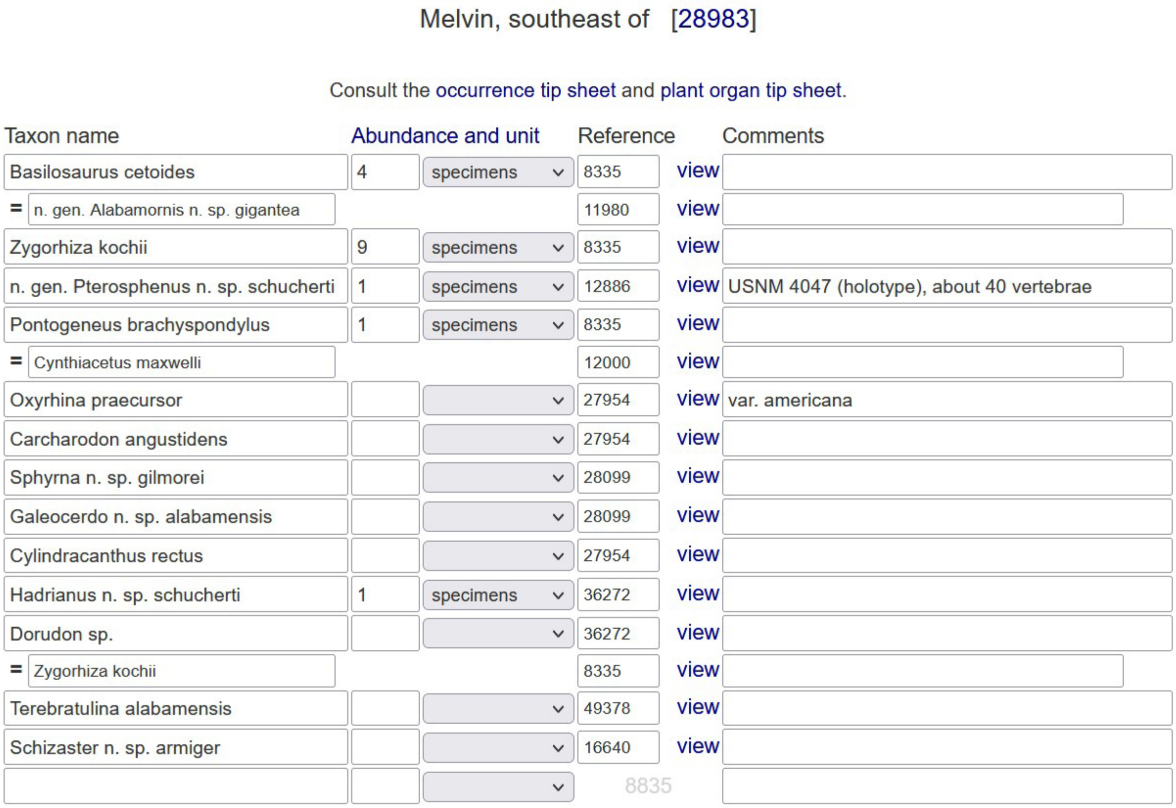
Another way to enter occurrences is to *Paste in an occurrence list*. All taxonomic names plus qualifiers must be formatted perfectly as outlined above for PBDB to accept a pasted occurrence list.

### 3.8.3 Edit old occurrences

As noted above, clicking on Edit old occurrences will take you to the *Occurrence entry form* after you search for a collection to add the occurrences to.

### 3.8.4 Reidentify occurrences

*Reidentify occurrences* is used to indicate that a particular occurrence of a fossil was previously identified



Melvin, southeast of [28983]

Consult the occurrence tip sheet and plant organ tip sheet.

| Taxon name | Abundance and unit | | Reference | | Comments |
|---|---|---|---|---|---|
| Basilosaurus cetoides | 4 | specimens | 8335 | view | |
| = n. gen. Alabamornis n. sp. gigantea | | | 11980 | view | |
| Zygorhiza kochii | 9 | specimens | 8335 | view | |
| n. gen. Pterosphenus n. sp. schucherti | 1 | specimens | 12886 | view | USNM 4047 (holotype), about 40 vertebrae |
| Pontogeneus brachyspondylus | 1 | specimens | 8335 | view | |
| = Cynthiacetus maxwelli | | | 12000 | view | |
| Oxyrhina praecursor | | | 27954 | view | var. americana |
| Carcharodon angustidens | | | 27954 | view | |
| Sphyrna n. sp. gilmorei | | | 28099 | view | |
| Galeocerdo n. sp. alabamensis | | | 28099 | view | |
| Cylindracanthus rectus | | | 27954 | view | |
| Hadrianus n. sp. schucherti | 1 | specimens | 36272 | view | |
| Dorudon sp. | | | 36272 | view | |
| = Zygorhiza kochii | | | 8335 | view | |
| Terebratulina alabamensis | | | 49378 | view | |
| Schizaster n. sp. armiger | | | 16640 | view | |
| | | | 8835 | | |

**Figure 30.** Re-identification results example.

as one taxon but has since been identified as a different taxon by a different reference (see Figure 29). This is not a taxonomic act (like synonymy) because it does not apply to the entire taxon (for which you would instead enter this as an "opinion"), but only to a specific occurrence. For instance, let's say that for a particular collection, a specimen was identified by the original reference as *Dorudon serratus*, but a subsequent author identified it as *Zygorhiza kochii*. You can use the Reidentify occurrences tool to enter this information into the PBDB.

When you click on *Reidentify occurrences*, the **Reidentifications search form** is called up. Use this form to locate the specific occurrence that you want to reidentify. You can *Search by taxon*, or *Search by collection*. After you enter your search terms, a list of occurrences that match those search terms will be called up on the **Re-identification form**. An example of this form is shown in Figure 29. It shows two different occurrences of the species *Dorudon serratus* in two different collections (13405 & 188994). Below each of these names are spaces where you can enter a re-identification. Using the same example from above, it is here you would type *Zygorhiza kochii* in the relevant blank line, indicating that the occurrence originally called *Dorudon serratus* is now *Zygorhiza kochii.* This reidentification will be attributed to the currently active reference, which is also noted on the **Reidentification form**. After you type in the re-identification, click <u>Save reidentifications</u> at the bottom of the form.

Note in the above example that the occurrence of *Dorudon serratus* from collection 13405 is the type of both the genus *Dorudon* and the species *Dorudon serratus*. If you find you are about to re-identify the holotype specimen, you should consider whether the author has performed a synonymy instead. Be sure you know the difference between these two actions before you perform either one. Similarly, if only some of the specimens previously assigned to a particular species were re-identified, do not use the Reidentification form to enter this information. Instead, simply add a new taxonomic occurrence from the active reference, perhaps noting in the comments section that the occurrence is based on some, but not all, of the specimens previously assigned to the other taxon. If abundance data are available, this should be updated accordingly for both taxa.

Figure 30 shows an example of the results of re-identification. Specimens from this collection that had previously been assigned to *Basilosaurus cetoides*, *Pontogeneus brachyspondylus* and Dorudon sp. were all re-identified in a later publication. The specimen(s) of *Basilosaurus cetoides* was redescribed in reference 11980 as *Alabamornis giganteus*, a new genus and a new species. The specimen(s) assigned to *Dorudon* sp. was re-identified as *Zygorhiza kochii* in reference 8835.

### 3.8.5 Reclassify occurrences of homonyms

When names that are homonyms get added to the PBDB, all occurrences must be identified as one of the taxa in question. Using the example outlined earlier in this guide regarding entry of homonyms, let's assume that the genus *Mesocetus* Van Beneden (1880) was entered into the PBDB first. Then, *Mesocetus* Moreno (1892) was added to the PBDB. At this point, the user is prompted to *Reclassify occurrences of homonyms*. The same would happen when the name *Mesocetus* Fraas (1904) was entered as well. One can also initiate this process from the **Main Menu** by clicking *Reclassify occurrences of homonyms* under **Collections and occurrences**. The Reclassification search form is identical to the **Reidentifications search form** because both are looking for particular occurrences to potentially modify. Enter search parameters to find the occurrences that need reclassification. On the resulting form, use the pull-down lists to assign each occurrence to the correct homonymic taxon. If you are uncertain, then choose the option "leave unclassified" so that future users more familiar with the taxa can choose the correct one.

## 3.9 Search Functions

PBDB search functions are fully available to both General users and Contributory users. These functions are outlined in section 1.8.1, General users.

## 3.10 Download Functions

PBDB download functions are fully available to both General users and Contributory users. These functions are outlined in section 1.8.1, General users.

## 3.11 Additional entry functions

Several *Additional entry functions* are available from the Main Menu. See below for a discussion of each function.

### 3.11.1 Enter new specimens

*Enter new specimens*, found under **Basic steps** in the **Main Menu**, allows you to enter measurement data on specimens and link those specimens to specific occurrences in the PBDB.

### 3.11.2 Correct misspellings

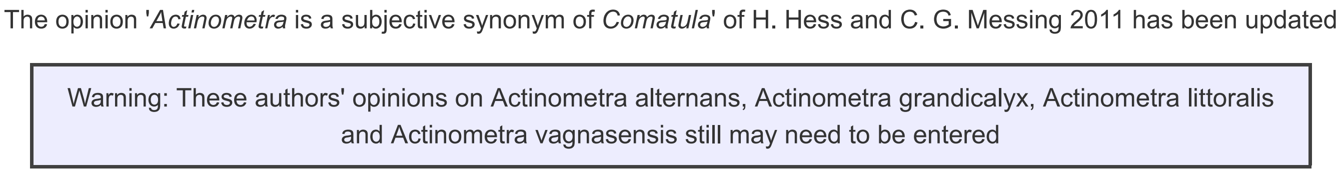
*Correct misspellings* can be found under **Collections and occurrences** in the **Main Menu**. Clicking calls up the **Possibly misspelled/unclassified occurrences search form**, which is identical to the **Reidentifications search form** because both are looking for particular occurrences to potentially modify. Type in the taxon name you want to look for misspellings of. The PBDB will search for occurrences with taxa with spellings close to, but not identical to, the one you are searching for. If any are found, you can look at them in more detail to determine if they are indeed misspellings of your taxon of interest. Be very careful to consult the original reference for each occurrence before correcting the misspellings to make sure that is indeed what they are; some correctly spelled names of different taxa can have spellings that are a single letter different to your taxon of interest.

### 3.11.3 Submit an external resource

*Submit an external resource* is available to both Contributory users and Guest users and is covered under section 2.4, Resources.

### 3.11.4 Add/edit ecological/taphonomic data

*Add/edit ecological/taphonomic data* allows you to enter taxon-specific ecological and taphonomic data associated with that taxon. When you click on this, you need to enter the name of the taxon for which you would like to add this data. If you enter this data for a species, it only applies to that species. If you enter it for a higher taxon, it applies to all lower taxa included in that higher taxon,



The opinion '*Actinometra* is a subjective synonym of *Comatula*' of H. Hess and C. G. Messing 2011 has been updated

Warning: These authors' opinions on Actinometra alternans, Actinometra grandicalyx, Actinometra littoralis and Actinometra vagnasensis still may need to be entered

**Figure 31.** Sample of reidentification warning that displays when adding synonym opinions.

EXCEPT if additional data has been entered for lower-level taxa within that higher level taxon. For instance, if you enter carnivore for the Carnivora, you can also add herbivore for *Aliuropoda melanoleuca* (giant panda). Otherwise, all Carnivora will be assigned a carnivorous diet.

These ecological and taphonomic data are grouped into three categories: *Morphology and development, Adult size,* and *Life history and behavior*, along with a Comments field for any additional relevant ecological or taphonomic data not covered in the other fields.

### 3.11.5 Add/edit vertebrate ecology data

The Add/edit vertebrate ecology data form is identical to the *Add/edit ecological/taphonomic data* form but only includes *Adult size,* and *Life history and behavior*. You can enter these data in either form for vertebrate taxa.

### 3.11.6 Add/edit specimen measurement data

The *Add/edit specimen measurement data* form was used to enter specimen measurements but has been superseded by the **Enter new specimens form**, so please use that instead. This form remains active to edit old data entered in the old system.

### 3.11.7 Search for bad names

The taxonomic hierarchy used by PBDB is constantly being dynamically re-assessed based on taxonomic opinions entered into the system. Because of this, some taxa are actively assigned to other taxa that may not currently be in use in the taxonomic hierarchy. The *Search for bad names* function is used to find these instances to help you determine what, if anything, can or should be done to re-assign them to taxa currently in use. When you click on *Search for bad names* you are asked to enter a taxon name plus some optional metadata about the names to be assessed. Doing so for Cetacea with no additional metadata recalls 14 "bad names". One of these returns looks like this:

> species *Plesiocetus giganteus*
> currently assigned to *Plesiocetus* based on Abel 1901
> *Plesiocetus* is a *nomen dubium* belonging to Mysticeti

This is alerting us to the fact that the species *Plesiocetus giganteus* is currently assigned to a genus (*Plesiocetus*) that is a *nomen dubium* AND it is implied that there is no active opinion that places the species "*Plesiocetus*" *giganteus* in any other taxon. Thus, the binomen *Plesiocetus giganteus* is a "bad name". There is nothing the PBDB can do about this given the current opinions entered into the system, but it suggests that an opinion regarding the disposition of "*Plesiocetus*" *giganteus* would be helpful if one could be found and entered. This tool can be very helpful to taxonomists trying to clean up the nomenclatural problems within a given taxon.

This same search is provided to you when you add an opinion linking two taxa by synonymy. For example, Figure 31 shows a warning that displays when two crinoid genera were synomized. It notes that several species in the PBDB belonging to the junior synonym exist (either currently or when originally added) in the database but may lack opinions recombining them into the senior synonym. If the reference notes the genus these species should be placed in, it is helpful to add these opinions. Unfortunately, it is common for papers to synonymize genera without discussing the placement of all species within those genera. (And if the reference does not mention them, but you know where they should be placed, it is a good opportunity to add that reference, if it hasn't already been entered.)

### 3.11.8 NEXUS file

*Upload a NEXUS file* can be used to do just that. These should be a published phylogenetic data matrix in the NEXUS file format. To do so, use the Choose file button to select the NEXUS file to upload and add any notes in the *Optional description* field. Click Upload at the bottom of the form to complete the process. This feature is being phased out soon as PBDB will be linking to NEXUS files in Morphobank.

## 3.12 Data Entry Settings

The PBDB allows users to specify some display options and pre-filled fields. These can be found by clicking on your name in the black navigation bar and selecting *Data Entry Settings*. **Display form selections** allows you to choose how you want to see the forms noted. Checking them all gives you the maximum access to data and data entry fields. For example, *Plant organs in occurrence form* is only needed if you are entering plant data.

Prefilled checkboxes, pulldowns, and text fields allow you to let the PBDB pre-fill select fields for you, so you don't need to enter those data over and over. If you are entering a suite of data that have the same values for these fields, it can be handy to use these. If not, it is better to leave them blank, otherwise you risk accidentally copying old data into new entries without realizing.

Once you have made all of your selections, click *Submit preferences* at the bottom of the form and these preferences will immediately be activated.

## 4. Paleobiology Database Application Programming Interface (API)

An Application Programming Interface (API) is a service that forms a connection between computer programs. It is a type of software interface that provides a service to other pieces of software. The PBDB API (Peters and McClennen 2015) includes a written standard that describes how to build or use the service. In contrast to the user interface, an API connects pieces of software to each other. The API is made up of different parts that act as tools or services that are available to the programmer. The API specification defines these calls, meaning that it explains how to use or implement them. The purpose of APIs is to expose only those parts a user or external program will find useful and keep them consistent, even if the internal details later change.

The PBDB API is used by various parts of the PBDB web interface, and it is also available for use by external resources such as web apps, phone apps, R scripts, etc. Many of the Resources use the PBDB API to query the PBDB automatically or dynamically for data. You, too, can craft scripts and write programs that use the PBDB API, or even just calls to the API which you can paste into your web browser to make individual data calls. (See section 3.11.3, *Submit an external resource* for instructions how to submit such programs for formal recognition.)

If you are new to using the PBDB API, a good place to start is with the Download Records form. As you enter download parameters into the form, it builds the API query for you in the blue linked text in the middle of the page (see Figure 3). As you add download parameters, the API query changes. You can use this form to build your API query, copy it, and then add additional details that are unavailable on the Download Records form by following the PBDB API documentation.

Currently, the PBDB API is only available for data output, but we are building a parallel system for data input that will allow other authorized systems to add data to the PBDB as well.

## 5. Glossary

‘ ’ — The use of single quotes or sometimes double quotes (see below) around a taxon name indicates that the name is obsolete in the immediate context, for instance, paired with a specific epithet (Bengston 1988).

" " — The use of double quotes or sometimes single quotes (see above) around a taxon name indicates that the name is obsolete in the immediate context, for instance, paired with a specific epithet (Bengston 1988). In phylogenetic nomenclature, quotation marks have been used to indicate a genus is not monophyletic (Cantino et al. 1999), and may indicate either paraphyly or polyphyly.

? — A question mark in front of a taxonomic name is a more informal way to indicate uncertain determination akin to cf. (Bengston 1988). See below.

aff. — This prefix in front of a taxonomic name indicates that the taxon in question is likely a new taxon with affinity to a known species. Often it is used when there is insufficient material available for formal description (Bengston 1988; Sigovini et al. 2016) or where it differs very slightly from the taxon “such that I might have doubts about referring it” (Estes 1987, p101).

cf. — This modifier is an abbreviation of the Latin *conferre*, meaning to bring together or compare. It may be used in multiple ways. Often, the use of cf. in front of a taxonomic name indicates that the determination is uncertain, perhaps due to poor preservation of the material in question (Bengston 1988; Sigovini et al. 2016). Alternatively, it may be used to mean “to be compared with” and indicate a conservative identification (Estes 1987).

ex gr. —This is abbreviation of *ex grege* (‘of the group including’)’ It has been applied to microfaunal groups with fossil records such as Ostracoda (Sigovini et al. 2016) and is often used for Pleistocene vertebrates to indicate a specimen can be assigned to a species group within a genus, but not identified to any one individual species.

form taxon — A group of specimens, organisms, or subtaxa which share a common overall form. Form taxa are common in botany where similar parts of organisms have been grouped together. For example, the form taxon *Stigmaria* refers to specimens of root structures of lycopsid trees. More casually, this term also refers to groups that share similarities but are known or believed to be non-monophyletic. The name may be retained for practical purposes (e.g., ‘lizards’) or because the true taxonomic affinities remain uncertain (e.g., many Ediacaran groups).

ichnotaxon —A taxon based on the fossilized work of an organism, including fossilized trails, tracks or burrows (trace fossils) made by an animal (ICZN 1999). The taxonomy of ichnotaxa is based on types of behavior and the morphology of the traces, but the trace maker(s) may or may not be known and may or may not be related to one another. Ichnotaxa are recognized by the International Commission on Zoological Nomenclature and follow the same naming rules as other animal taxa.

indet. —This designation has been used to indicate both the fact of being undetermined and the inability to be determined (for example, due to poor preservation). For genera, Sigovini et al. (2016) recommend the use of 'sp.' for the former situation and 'indet.' for the second situation. However, in practice, they may be used interchangeably.

regular taxon — A group of specimens, organisms, or subtaxa which are more closely related to each other than they are to organisms or subtaxa outside the group.

s.l. or *sensu lato* — 'in a broad sense'. Refers to a conception of a taxon that encompasses broad variation of form and/or a larger number of subtaxa than another conception.

s.s., s. str. or *sensu strictu* – 'in a strict sense'. Refers to a conception of a taxon that restricts the variation of form and/or includes a smaller number of subtaxa than another conception. This term is NOT available in Paleobiology Database.

sp. — The use of sp. after a genus or subgenus name indicates that the species has not been determined but does not preclude the possibility that species determination is possible. See also discussion under 'indet.'

Sigovini et al. (2016) provide a helpful flow-chart (their fig. 1) for the best practice of open nomenclature.

## ACKNOWLEDGEMENTS

We would like to thank the hundreds of people who have contributed their time and effort to entering data into the Paleobiology Database. Without you, it would just be an empty shell. We also want to thank all of the Paleobiology Database leaders, Executive Committee members, Principal Investigators and others who have guided the progress of the project since its inception. Specifically, we want to thank John Alroy. Without his leadership and efforts on many fronts, Paleobiology Database might not have survived.

We also thank the funding agencies who have supported our efforts over many years including: the National Center for Ecological Analysis and Synthesis (NCEAS), the US National Science Foundation (NSF, including EarthCube), Research Data Alliance (RDA), Swiss National Science Foundation (SNSF), ETH Zurich, and others.

## LITERATURE CITED

Bateman, R.M., and J. Hilton. 2009. Palaeobotanical systematics for the phylogenetic age: applying organ-species, form-species and phylogenetic species concepts in a framework of reconstructed fossil and extant whole-plants. *Taxon* 58(4):1254–1280.

Becker, T.R., U. Jansen, G. Plodowski, E. Schindler, S.Z. Aboussalam, and K. Weddige. 2004. Devonian litho- and biostratigraphy of the Dra Valley area: an overview. Pp. 3-18 in A. El Hassani (ed.), *Devonian Neritic-Pelagic Correlation and Events in the Dra Valley (Western Anti Atlas, Morocco)*. Institut Scientifique, University Mohammed V, Rabat, Morocco.

Bengston, P. 1988. Open nomenclature. *Palaeontology* 31(1):223-227.

Bertling, M., S.J. Braddy, R.G. Bromley,., G.R. Demathieu, J. Genise, R., Mikuláš, J.K. Nielsen, K.S. Nielsen, A.K. Rindsberg, M. Schlirf and A. Uchman. 2006. Names for trace fossils: a uniform approach. *Lethaia* 39(3):265-286. https://doi.org/10.1080/00241160600787890

Buatois, L. and M.G. Mangano. 2011. *Ichnology: Organism-Substrate Interactions in Space and Time*. Cambridge University Press; Cambridge xii + 358 pp.

Cantino, P.D., H.N. Bryant, K. de Queiroz, M.J. Donoghue, T. Eriksson, D.M. Hillis, M.S.Y. Lee. 1999. Species names in phylogenetic nomenclature. *Systematic Biology* 48(4):790-807. https://doi.org/10.1080/106351599260012

Cohen, K.M., S.C. Finney, P.L. Gibbard, and J.X. Fan. 2013. The ICS international chronostratigraphic chart. *Episodes Journal of International Geoscience* 36(3):199-204. https://doi.org/10.18814/epiiugs/2013/v36i3/002

Estes, R. 1987. Lucas and Zidek on taxonomic syntax. *Journal of Vertebrate Paleontology* 7(1):101. https://doi.org/10.1080/02724634.1987.10011642

Gradstein, F.M., J.G. Ogg, and A.G. Smith. 2004. *A Geologic Time Scale 2004*. Cambridge University Press; Cambridge, UK. xix+589 pp.

Gradstein, F.M., J.G. Ogg, M.D. Schmitz, and G.M. Ogg. 2012. *The Geologic Time Scale 2012*. Elsevier; Oxford, UK. xviii+1144 pp. https://doi.org/10.1016/C2011-1-08249-8

Greuter, W., ., J. McNeill, F.R. Barrie, H.-M. Burdet, V. Demoulin, T.S. Filgueiras, D.H. Nicolson, P.C. Silva, J.E. Skog, P. Trehane, N.J. Turland, and D.L. Hawksworth. 2000. *International Code of Botanical Nomenclature (Saint Louis Code): Adopted by the Sixteenth International Botanical Congress; St. Louis, Missouri, July-August 1999*. Regnum Vegetabile 138. Koeltz Scientific Books, Königstein.

International Commission on Zoological Nomenclature. 1999. International Code of Zoological Nomenclature. Fourth edition. London, U.K. International Trust for Zoological

Nomenclature. Online at https://www.iczn.org/the-code/the-code-online/

McNeill, J., F.R. Barrie, H.-M. Burdet, V. Demoulin, D.L. Hawksworth, J. Marhold, D.H. Nicolson, J. Prado, P.C. Silva, J.E. Skog, J.H. Wiersema, N.J. Turland. 2006. *International Code of Botanical Nomenclature (Vienna Code): Adopted by the Seventeenth International Botanical Congress; Vienna, Austria, July 2005*. Regnum Vegetabile 146. Ruggell: Gantne .

Payne, J.L., FF. A. Smith, M. Kowalewski, R. A. Krause Jr, A. G. Boyer, C. R. McClain, S. Finnegan, P. M. Novack-Gottshall, and L. Sheble. 2012. A lack of attribution: closing the citation gap through a reform of citation and indexing practices. *Taxon* 61(6):1349-1354. https://doi.org/10.1002/tax.616030

Peters, S.E. 2014. The Paleobiology Database releases PBDB Navigator: *Priscum* 21(1):1-2.

Peters, S.E. 2018. Macrostrat: a platform for geological data integration and deep-time earth crust research. *Geochemistry, Geophysics, Geosystems* 19(4):1393-1409.

Peters, S.E., and M. McClennen. 2015. The Paleobiology Database application programming interface. *Paleobiology* 42(1):1-7. https://doi.org/10.1017/pab.2015.39

Raine, J.I., A.G. Beu, A.F. Boyes, H.J. Campbell, R.A. Cooper, J.S. Crampton, M.P. Crundwell, C.J. Hollis, H.E. Morgans, and N. Mortimer. 2015. New Zealand geological timescale NZGT 2015/1. New Zealand *Journal of Geology and Geophysics* 58(4):398-403. https://doi.org/10.1080/00288306.2015.1086391

Rios, N.E., and H.L. Bart. 2010. GEOLocate (Version 3.22) [Computer software]. Belle Chasse, LA: Tulane University Museum of Natural History.

Scotese, C.R., R.K. Bambach, C. Barton, R. Van der Voo, and A. M. Ziegler. 1979. Paleozoic base maps. *The Journal of Geology* 87(3):217–277. https://www.jstor.org/stable/30060818Scotese, C. R., and W.W. Sager. 1988. Mesozoic and Cenozoic Plate Tectonic Reconstructions. *Tectonophysics* 155:27-48.

Sigovini, M., E. Keppel and D. Tagliapietra. 2016. Open Nomenclature in the biodiversity era. *Methods in Ecology and Evolution* 7(10):1217-1225. https://doi.org/10.1111/2041-210X.12594Smith, H. M. 1945. Categories of species names in zoology. Science 102(2643): 185-189. https://doi.org/10.1126/science.102.2643.185

Soller, D.R. 2023. The National Geologic Map Database: An accessible archive of the nation's standardized geoscience information. Geostrata 27(2):56-63. https://doi.org/10.1061/geosek.0000476

Sumrall, C.D., J. Sprinkle, S. Pruss, and S. Finnegan. 2009. *Cardiocystella*, a new cornute stylophoran from the Upper Cambrian Whipple Cave Formation, eastern Nevada, USA. *Journal of Paleontology* 83(2):307-312.

Uhen, M. D., A.D. Barnosky, B. Bills, J. Blois, M.A. Carrasco, M.T. Carrano, G.M. Erickson, J.T. Eronen, M. Fortelius, R.W. Graham, E.C. Grimm, M.A. O'Leary, A. Mast, W.H. Piel, P.D. Polly, and L.K. Säilä. 2013. From card catalogs to computers: databases in vertebrate paleontology. *Journal of Vertebrate Paleontology* 33(1):13-28. https://doi.org/10.1080/02724634.2012.716114

Wright, N., S. Zahirovic, R.D. Müller and M. Seton. 2013. Towards community-driven paleogeographic reconstructions: integrating open-access paleogeographic and paleobiology data with plate tectonics. *Biogeosciences* 10(3):1529-1541. https://doi.org/10.5194/bg-10-1529-20137.

## 7.1 Appendix 1

### Paleobiology Database Basic Data Entry Checklist

Go to paleobiodb.org and log in.

I. Reference
  A. Click Main Menu
  B. Click Add a new reference
  C. Enter the first author's last name and the year of the publication and click Search
  D. If a match is displayed, check to see if the matching publication is the one you are trying to enter. If it is, then STOP.
  E. If the PBDB opens up a reference entry screen, or it does NOT match, then proceed.
  F. Enter author names using instructions on the screen and the Tip Sheet.
  G. Choose the Publication Type.
  H. Make sure the Publication Year is filled in correctly.
  I. If the reference is in a language other than English, use the pull down to enter the correct language. If the language isn't on the list, choose Other and type the language in the Comments field below.
  J. Enter the Title and Serial name for journal articles.
  K. Enter the Volume and Number (also sometimes called Issue or may not exist).
  L. Enter page numbers.
  M. Enter the DOI for journal articles, if available.
  N. Proofread your entry and then click Submit.

II. Taxonomic Names
  A. Ensure you have the correct reference active by looking at the Ref: in the black bar at the top of the screen.
  B. Click Main Menu
  C. Click Add/edit taxonomic name
  D. Enter the taxonomic name into the Taxon name field
  E. Click Search.
  F. Enter the Authority data for your taxon. Be sure to use the pull-down fields to answer the questions about the taxon. The Preservation category pull-down refers to the holotype specimen of the taxon.
  G. Click Submit.

III. Taxonomic Opinions
  A. Ensure you have the correct reference active by checking Ref: in the black bar at the top of the screen.
  B. Click Main Menu
  C. Click Add/edit taxonomic opinion
  D. Enter the taxonomic name for which you want to enter an opinion into the Taxon name field
  E. Click Search.
  F. Either click on an opinion to edit or Create a new opinion record
  G. Enter as much data as possible regarding the opinion. Note at the top you can indicate whether the opinion is from the active reference OR if it is listed in the reference but is actually the opinion of an earlier authority.
  H. Be careful about using the exact spelling listed in the reference and the reason why this spelling and rank is used by this authority.
  I. Click Submit.

## Paleobiology Database Basic Data Entry Checklist, continued

IV. Collection
  A. Ensure you have the correct reference active by checking Ref: in the black bar at the top of the screen.
  B. Click Main Menu
  C. Click Add a new collection
  D. Enter the time interval of the collection.
  E. Enter the coordinates of the collection. Make sure to choose the correct North/South and East/West directions for your coordinates. If you don't have coordinates, try GEOLocate to help you find them.
  F. All collections from the time period you indicated within 100 km are now displayed, from nearest to farthest away. Check to see if any of them seem to match your collection. If all collections parameters match, and you are SURE your fossils are from the same hole in the ground, then STOP and go to Occurrences below.
  G. If none match, scroll to the bottom of the page and click Add a new collection.
  H. Enter Collection data here, making sure to click on all 7 gray tabs at the top of the form, starting with those on the left, and moving to the right. Make sure you enter data in all fields marked in Red. Consult the Tip Sheet if you need additional help.
  I. Make sure Access Level is set to the public, and release date is set to immediate.
  J. Make sure when you enter coordinates you EITHER enter degrees, minutes, seconds OR degrees + a decimal fraction of degrees, but not both.
  K. Look at the map. Make sure the red kite shape is in the right place. If not, check your coordinates, including the direction to see what went wrong.
  L. Enter the most specific time period you have. The Maximum and Minimum fields will autopopulate with approved time period names. If you have one that is NOT autopopulating, Google it to see what it might be equivalent to in other more standard time scales. If you only have one time period, enter it in the Maximum field.
  M. While stratigraphic names (Group, Formation, Member) are not required, they are very desirable. Enter them if you have them, which is very likely.
  N. You must enter a Primary Lithology. If the rocks are not described at all, choose unkown from the pull down.
  O. You must enter an environment. The most general environments are marine indet. or terrestrial indet. Hopefully you can figure out one of those or you can choose something more specific.
  P. Under Preservation, you must choose a Mode of preservation. Do your best, but if you don't see anything specific, choose body.
  Q. Under Components you must choose a Size class. These should be easy to pick based on the sizes of your fossils. You can pick more than one.
  R. When you are done entering collection data, click Enter collection and exit at the bottom of the page.
  S. You will be presented with a summary page for your Collection. Click through the gray tabs and proofread your entry.
  T. When you are satisfied, click Edit taxonomic list at the bottom of the page and move on to Occurrences below, at Step D.

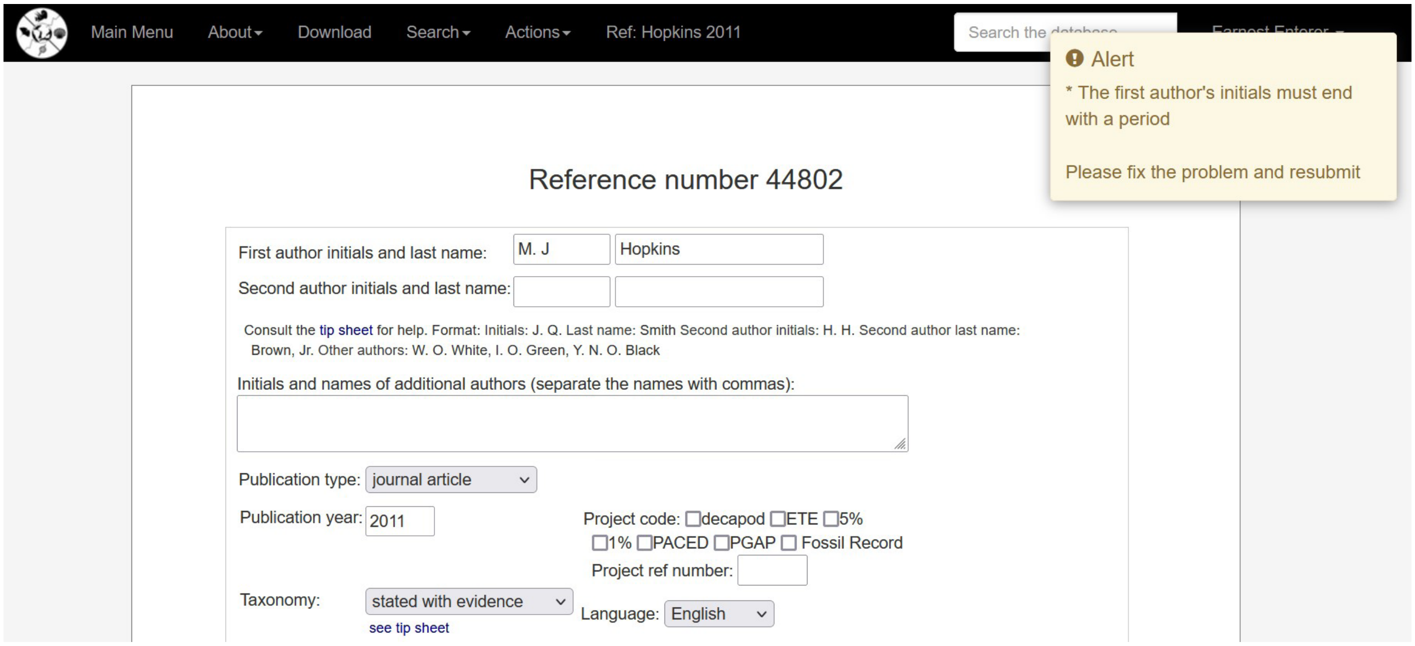
**Paleobiology Database Basic Data Entry Checklist, continued**

V. Occurrences
    A. Click Main Menu
    B. Click Add Taxonomic Occurrences
    C. Enter the name or number of the collection for which you want to enter occurrences
    D. Enter the taxonomic names at whatever level they are listed in your reference.
    E. If it is a genus without a particular species, enter as Genus sp.
    F. If it is a taxon higher than genus, enter it as Taxon indet.
    G. If it is a new genus, put n. gen. in front of the genus name.
    H. If it is a new species, put n. sp. in front of the specific epithet.
    J. When you are done entering taxa into the list, click save changes at the bottom of the page.
    K. Inspect the summary list for any spelling errors.

## 7.2 Appendix 2

### Examples of PBDB warning messages

Alerts regarding misformatted data are usually presented in the form of yellow pop-up alert boxes. To clear the alert, fix the problem and re-submit the form.



Alert

* The first author's initials must end with a period

Please fix the problem and resubmit

Reference number 44802

First author initials and last name: M. J Hopkins

Second author initials and last name:

Consult the tip sheet for help. Format: Initials: J. Q. Last name: Smith Second author initials: H. H. Second author last name: Brown, Jr. Other authors: W. O. White, I. O. Green, Y. N. O. Black

Initials and names of additional authors (separate the names with commas):

Publication type: journal article

Publication year: 2011

Project code: ☐decapod ☐ETE ☐5% ☐1% ☐PACED ☐PGAP ☐ Fossil Record

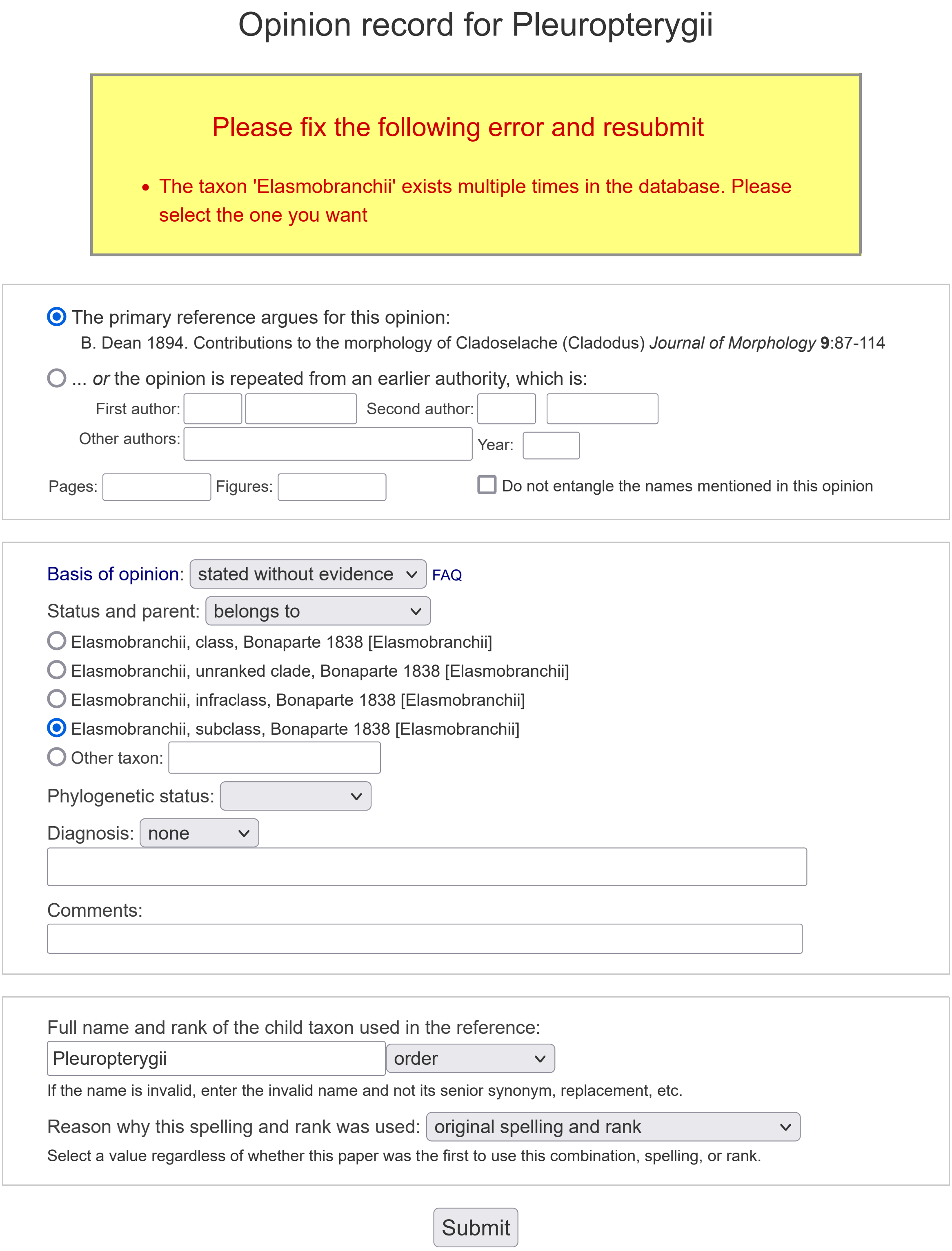
Project ref number:

Taxonomy: stated with evidence

see tip sheet

Language: English

When you enter a taxonomic opinion, sometimes the PBDB needs additional information to ensure the opinion is accurate. In the case below, the opinion Pleuropterygii belongs to Elasmobranchii is accurate, but insufficient because there are three different versions of Elasmobranchii are in the system with different ranks. Here, the author assigned Pleuropterygii to the Subclass Elasmobranchii, so that radio button was selected, and the opinion submitted successfully.



Opinion record for Pleuropterygii

Please fix the following error and resubmit

- The taxon 'Elasmobranchii' exists multiple times in the database. Please select the one you want

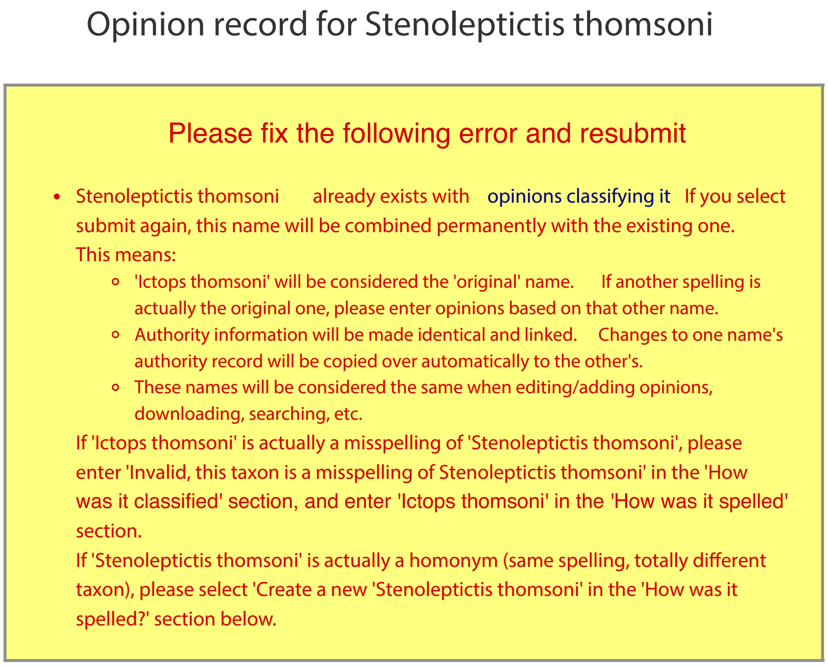
● The primary reference argues for this opinion:
B. Dean 1894. Contributions to the morphology of Cladoselache (Cladodus) *Journal of Morphology* **9**:87-114

○ ... *or* the opinion is repeated from an earlier authority, which is:
First author: Second author:
Other authors: Year:
Pages: Figures: ☐ Do not entangle the names mentioned in this opinion

Basis of opinion: stated without evidence
Status and parent: belongs to
○ Elasmobranchii, class, Bonaparte 1838 [Elasmobranchii]
○ Elasmobranchii, unranked clade, Bonaparte 1838 [Elasmobranchii]
○ Elasmobranchii, infraclass, Bonaparte 1838 [Elasmobranchii]
● Elasmobranchii, subclass, Bonaparte 1838 [Elasmobranchii]
○ Other taxon:
Phylogenetic status:
Diagnosis: none
Comments:

Full name and rank of the child taxon used in the reference: Pleuropterygii order
If the name is invalid, enter the invalid name and not its senior synonym, replacement, etc.
Reason why this spelling and rank was used: original spelling and rank
Select a value regardless of whether this paper was the first to use this combination, spelling, or rank.

If previous contributors have not followed correct procedure when entering taxonomic names, the same species may exist under two (or more) combinations in the PBDB. Recombination of a species name may also create a secondary homonym with a different species already in the PBDB. In either of those situations, you may encounter this error message when entering taxonomic opinions. If so, stop and do not proceed. Combining the names triggers a bug where the names are not truly combined and where the taxonomic opinion algorithm often does not use the most recent opinion. The option to create a new taxon in the case of a homonym also does not work, and the names will be combined and entangled incorrectly. If you encounter this error message, contact info@paleobiodb.org for assistance.



Opinion record for Stenoleptictis thomsoni

Please fix the following error and resubmit

- Stenoleptictis thomsoni already exists with opinions classifying it If you select submit again, this name will be combined permanently with the existing one.
  This means:
  - 'Ictops thomsoni' will be considered the 'original' name. If another spelling is actually the original one, please enter opinions based on that other name.
  - Authority information will be made identical and linked. Changes to one name's authority record will be copied over automatically to the other's.
  - These names will be considered the same when editing/adding opinions, downloading, searching, etc.

  If 'Ictops thomsoni' is actually a misspelling of 'Stenoleptictis thomsoni', please enter 'Invalid, this taxon is a misspelling of Stenoleptictis thomsoni' in the 'How was it classified' section, and enter 'Ictops thomsoni' in the 'How was it spelled' section.

  If 'Stenoleptictis thomsoni' is actually a homonym (same spelling, totally different taxon), please select 'Create a new 'Stenoleptictis thomsoni' in the 'How was it spelled?' section below.