DEVELOPING PROCEDURES AND SOFTWARE

TO COMPARE AND UPDATE DATABASES

REPORT FOR FAPEAM

Fundação de Amparo à Pesquisa do Estado do Amazonas

7- 12

January 2013

Denis Filer
Department of Plant Sciences
University of Oxford, UK

Introduction

This document explains how you can connect a BRAHMS database to another BRAHMS database and then auto-check for botanical records in your database than also occur in the external database, based on selected criteria.

All matched records are assembled into your botanical record extract file. The origin of the data is indicated in the DBID field and the file can then be sorted to show matched record pairs together.

Matched records are then easily compared in a sorted extract file. Records from the external database can be used to selectively update and improve your own database. You can also add specimens and their determinations from the external database and then subsequently, use the newly added specimens to update determinations in your own database.

These match and update functions are especially useful for herbarium curation databases. The herbarium database can be connected to an external database and draw in improved quality data. For example, you may have duplicates from another herbarium and if that herbarium has a BRAHMS database, their data may be better than your own. Also, if you have access to a database that specializes in any taxonomic group, then you can connect to this and draw in the improved quality data.

Data may be exchanged in both directions. A real example tested in Brazil was comparing the large INPA database with several smaller regional Amazonian databases with INPA benefitting by drawing in improved botanical record and specimen data. But in turn, the smaller herbaria were able to benefit and update their own records form those in the INPA database with a higher quality.
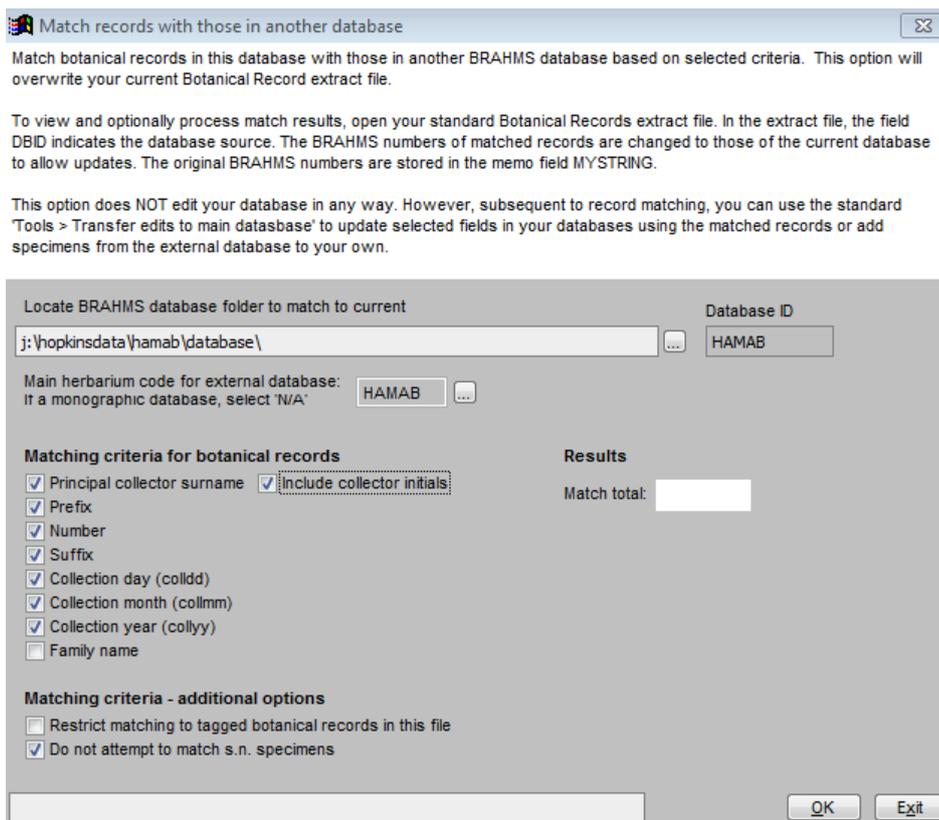
This work was undertaken at INPA in Manaus http://www.inpa.gov.br/ in collaboration with Dr. M. Hopkins and the REFLORA project, funded by FAPEAM - Fundação de Amparo à Pesquisa do Estado do Amazonas.

Matching botanical records

Matching records
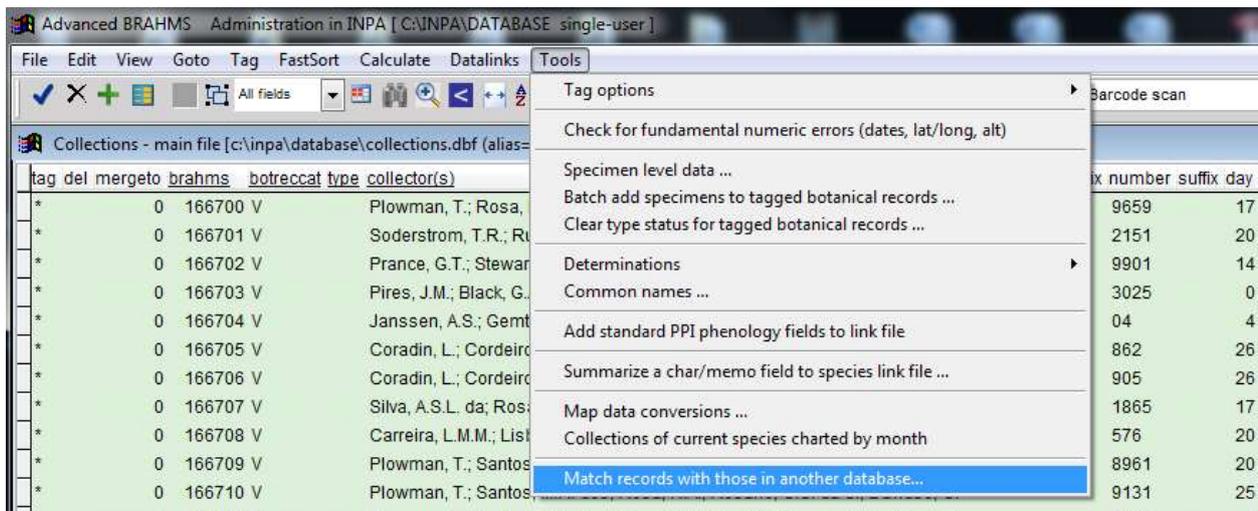
With the main Botanical Records file open, select Tools > Match records with those in another database…

The matching form.

Locate the external BRAHMS database and enter its herbarium code if relevant. NB for research databases, choose herbarium 'N/A' (not applicable). The external database can be in an older version of BRAHMS.

Choose the criteria for botanical record matching. Some experimentation may be necessary to optimise matching and alter the result you get in 'Match total'. For example, excluding family name from the match
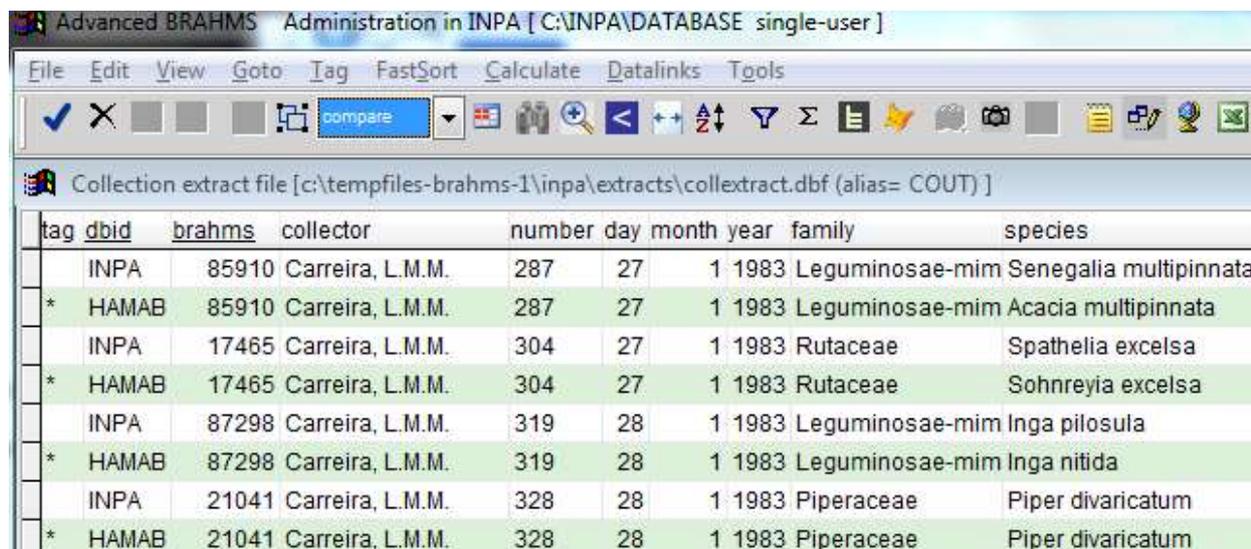
may improve matching as families may be used differently in the external database. This option does not edit your database in any way. By varying the match criteria, you can often locate interesting differences and/or potential errors in your data.

Match results in the extract file

Matched records from the external database are added to the main Botanical Records extract file together with the relevant records from your own database.

In this extract file, the field DBID indicates the database source. In the extract file the BRAHMS numbers of matched records are changed to those of the current database to allow updates. The original BRAHMS numbers are stored in the memo field MYSTRING.

To display matched records together, you must sort the file on BRAHMS number or collector name+number. Thus, for example, use **Fastsort** to sort on 'Collector+number'.


The above screen shows an example sort. The INPA database has been match with HAMAB. The HAMAB BRAHMS numbers have been edited to be the same as the matched INPA records to allow updating. NB the original HAMAB BRAHMS number is stored in the MYSTRING memo as this can then be used to draw further data including specimens from that database. You can see, even if this small example, differences are being picked up with determinations.

Processing match results

Using Update function for botanical records

As the BRAHMS numbers of the external records have been changed to those in the receiving database, it is possible to the Update function from this extract file to selectively update your database. These functions are available under **Tools > transfer edits to main database.**
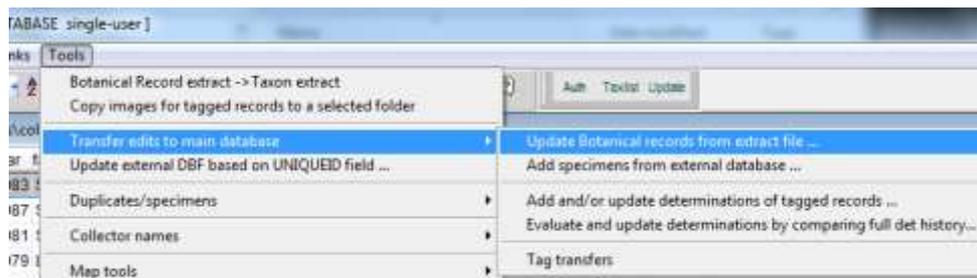

Using **Tools > transfer edits to main database > Update Botanical records from extract file ...**, you can choose categories of data to update, using the tagged records from the external database. For example, the external database may have better locality data or plant description notes and these data can be used to update your own records.
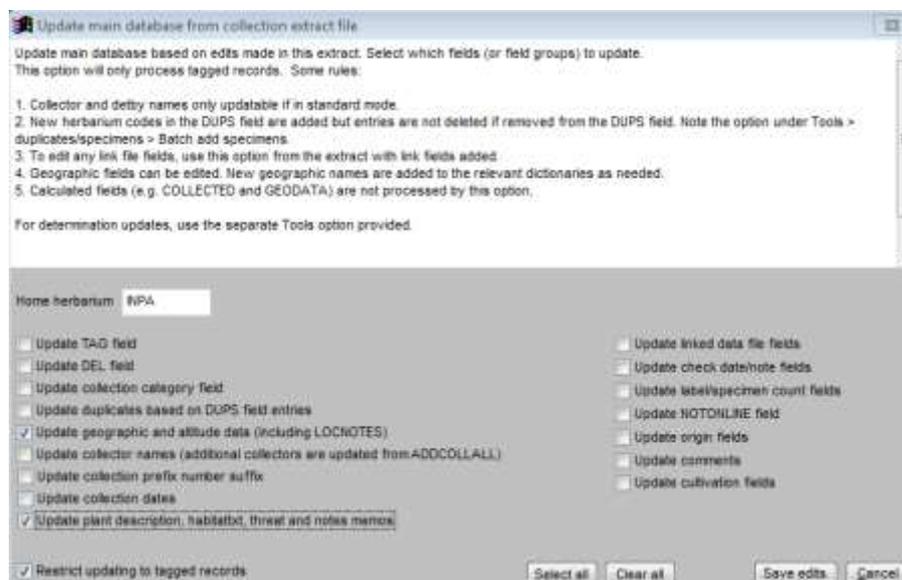
The Update form with some options selected. Updates are always restricted to tagged records.

Importing specimen data with determinations

Using the matched records in the extract file, and the option Using **Tools > transfer edits to main database > Add specimens from external database…**, you can also draw in specimens from the external database together with their determinations.
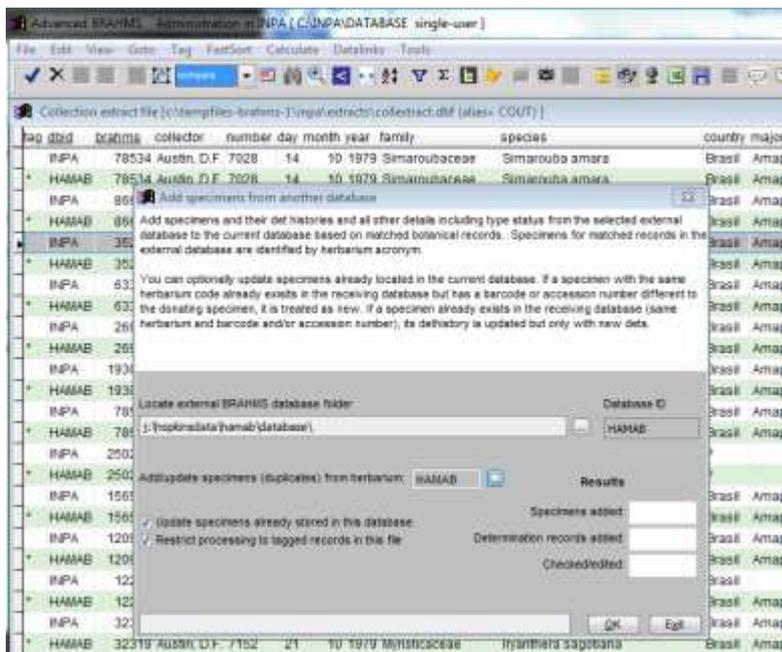
The specimen import form.

Specimens are imported for a selected herbarium ID, in the above example, 'HABAM'. The option located the specimens in the external database and adds these to the matched Botanical Records in your own database. The determination history and any type details of these specimens are also added but the current determination of the botanical record in your database is NOT edited (this being an optional and separate process discussed below).

Specimens are not added if already present in your database although you can optionally update them. If your database already has a specimen for the selected herbarium ID (in this case HAMAB), but the barcode and/or accession number in the external database are different, it will be treated as a new specimen.

Updating determinations

If you have added specimens from the external database, you can now proceed to examine all the determinations you have available and selectively update your current determinations. Clearly, this has to be done with discretion but if you have added specimens from a source likely to include more recently updated (or more reliable) determinations, you can now further improve your database.

As a first step, you must open the determination records association with the current Botanical Record extract file.
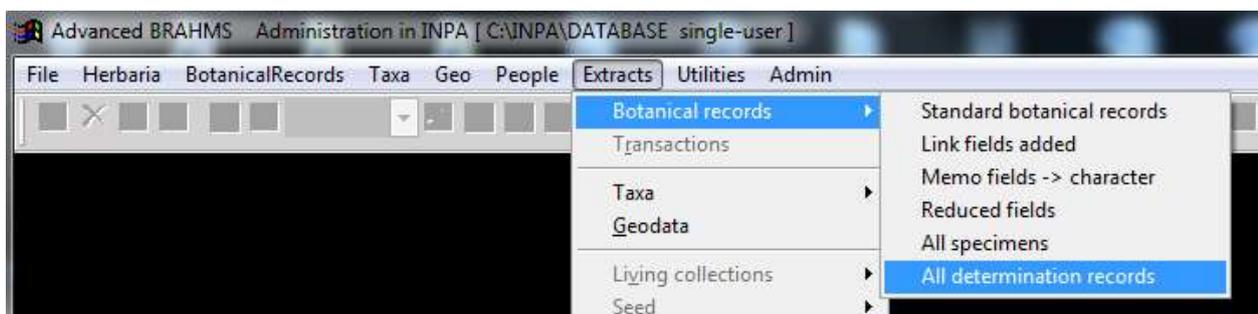

Opening determination records

In this extract file, there are various useful **Tools** options to help review the available determinations and select the best option.
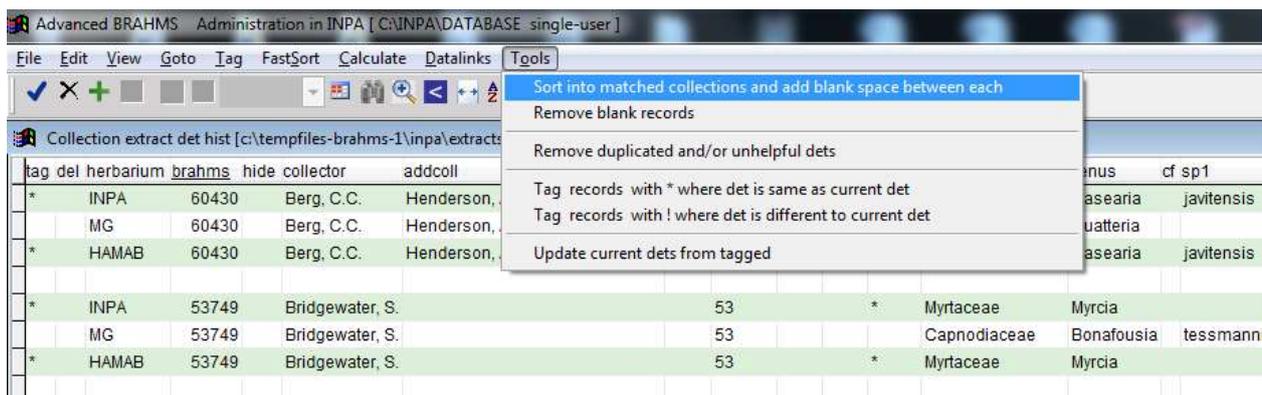
Determination file Tools options

In the above example, the option **'Sort into matched collections and add blank space between each'** has been selected.

The option **Tools > Remove duplicated and/or unhelpful dets** is especially useful in this context as you can remove, from the extract file, records with only one det (*i.e.* nothing to compare) and duplicated dets. Also, you can keep only determinations that are indicated as being a current determination in your database and for those determinations imported from an external database.
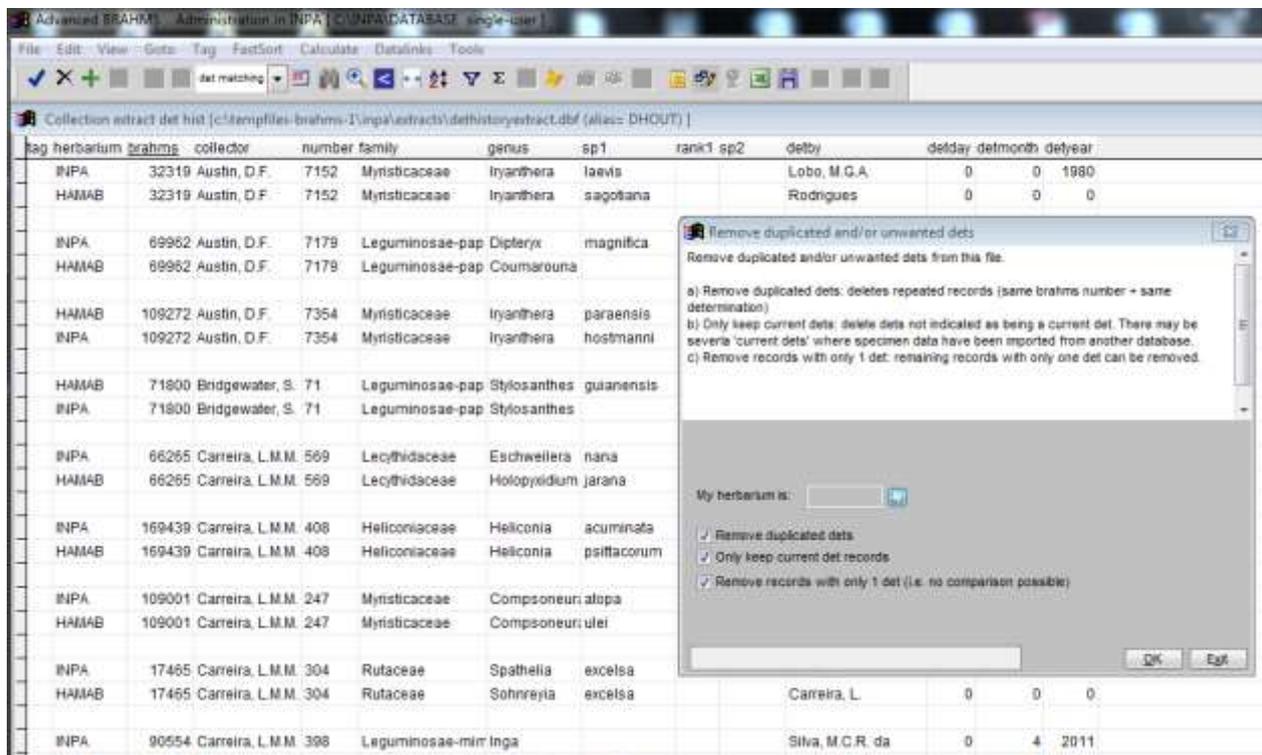
The form opened using Tools > Remove duplicated and/or unhelpful dets

Once you have tagged the records you prefer, you can then use **Tools > Update current dets from tagged**.