



Specify 6 Synonym Cleanup Tool

(The Cleanup Tool is located in Specify's top-level menu:
System/System Setup/Configuration, then "Cleanup Synonyms".)

Background

Specify has rich support for managing and visualizing taxonomic names associated with Collection Objects. A Specify collection database contains both preferred taxon names and synonyms. Synonyms in Specify 6 are stored differently from the way they were in Specify 5. During conversion of Specify 5 databases to Specify 6, some synonyms created in Specify 5 could not be automatically positioned and displayed in the Specify 6 Taxon Tree. The Specify 6 Synonym Cleanup Tool, released with Specify 6.5, addresses that issue for Specify 5-to-6 database conversions. The tool has no effect on collection databases created for the first time in Specify 6, or on Specify 6 databases converted from other types of legacy systems. If your collection's data were not converted from Specify 5-to-6, the Synonym Cleanup Tool will not change or improve your Specify 6 database. If that is the case, there is no need to run it. For sites with Specify 6 databases converted from Specify 5, read this documentation to consider the utility and outputs of the tool for your database.

Key Concepts for Taxon Tree Data Management in Specify

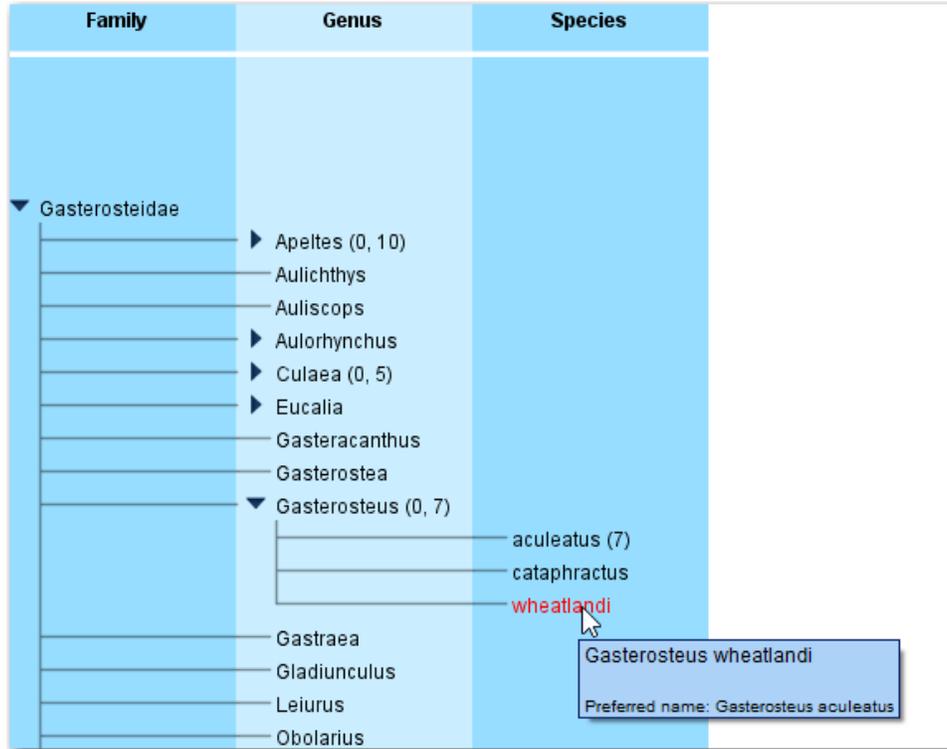
Except for the base (or root), every entry in a Taxon Tree has one parent and may have one or more children—all of which are collectively referred to as tree "nodes". Nodes are linked to one another following the traditional taxonomic hierarchy. A family node is a parent of one or more genera, a genus is a parent of one or more species nodes, a species is a child node to a genus, etc. In addition to adding, editing and deleting nodes or taxa, Specify 6 supports relocating nodes with drag-and-drop support for 'merging' nodes together and for the 'synonymization' of nodes. "Merging" nodes replaces one name (i.e. node) with another, moves all of the children of the first node into the second, and then deletes the first node. The Merge function is intended for correcting data entry mistakes, such as eliminating unwanted duplicate entries in the Taxon Tree of the same taxon resulting from a misspelling.

The Taxon Tree window also supports "Synonymizing" one node with another at the same taxonomic rank, for example synonymizing one species with another, and in some cases synonymizing taxa of different ranks, for example synonymizing a variety to a subspecies. "Synonymizing" in the Specify Taxon Tree has no automated connection to a formal nomenclatural action and may not be congruent

with the published literature. The function is primarily intended to allow you to designate which name you prefer to use for cataloging your Collection Objects. For example, two or more valid names for a taxon might be available, but for consistency you wish to use only one for cataloging. Alternatively you may prefer one taxon name and concept over some other for a particular set of collection objects. Synonymizing, in a Specify sense, is the identification of a preferred name for shelving or cataloging purposes-for whatever practical, taxonomic or nomenclature reason. Unlike the Merge function which removes a misspelled name and replaces it with a correctly spelled one, synonymization in Specify is different in three ways. 1) It keeps all the taxon names (tree nodes) involved in the Taxon Tree-nothing is deleted, 2) the Taxon Tree window shows synonyms in a different color and makes synonymic relationships visible in mouse-over, popup windows, and 3) synonymization makes a change in the Determination data table for any associated Collection Objects with annotations which use the synonymized name.

When one Taxon Tree node is synonymized with another, any existing Determination records for Collection Objects which previously used the synonymized name are automatically updated to show it and the 'Preferred Taxon' in the Determination record. Original Determinations are not changed or deleted; historical determinations for Collection Objects are always preserved. What does change in the database is that a data field "Preferred Taxon" in the Determination table is automatically filled with the name you prefer to use for Collection Objects which were previously determined to names now designated as synonyms in the Taxon Tree. The original Taxon (name) field for a Determination (and the Determiner and date) for all historical Determinations related to a Collection Object stay intact in the database, and along with the Preferred Name they can be accessed and used in searches, queries, reports, and labels. If a name is synonymized in the Taxon Tree has no associated Collection Objects, then no Determination records are annotated with a 'Preferred Taxon'. If a preferred cataloging classification with synonyms is created in a Specify Taxon Tree prior to specimen computerization, new Collection Object data entry will record each historical Determination in the Taxon field and then automatically fill the Preferred Taxon field in the Determination record for those Collection Objects.

The figure below shows a Taxon Tree display and the resulting Determination data form after a synonymization action. The species *G. wheatlandi* was made a synonym of *G. aculeatus*. The synonym then displays in a different color and moving the mouse cursor over the synonym name shows the preferred name in the popup window. The resulting Determination record in the data form shows the original Taxon Determination and Determiner, and the new Preferred Taxon name.



Trees Reports Interactions Statistics Query Workbench SGR Plugins Lifemapper Attachments

Collection Object

Cat #: 1713 Accession #: 1996-IC-044 Voucher Nu... KU 29782

Cataloger: Bentley, Andrew C Cat Date: Voucher: 29782

Determinations

Taxon: Gasterosteus wheatlandi

Preferred Taxon: Gasterosteus aculeatus

Determiner: Hartel, Karsten E Date: Type Status: None

Remarks:

Field No.:Locality: KEH 96-1: 05/10/1996: USA, Massachusetts, Barnstable: Head of Trunk River, Falmouth, 41.5408325195, -70.6250000000

Preparations

In addition to maintaining the hierarchical relationships between nodes for preferred taxa, Specify also keeps track of the parent nodes of synonyms. For example, if you 'synonymize' a species name in the Taxon Tree, the data record for that species name always includes the name of original generic node (parent) for that species. If a species is 'synonymized' to a species in the same genus, then the parent node of both names at the genus rank will be the same. If a species is synonymized to a species in different genus, the species synonym's parent will be its original genus not the genus of the Preferred Taxon. These parent-child relationships for synonyms are not visually represented in the Taxon Tree display. But examining the Taxon record in a data form for any synonym will show the parent node.

The Problem: Synonyms Converted from Specify 5-to-6

Collection databases converted from Specify 5-to-6 may have inconsistencies with the parent-child relationships and placement of synonym names in the Taxon Tree. Because Specify 5 stored species synonyms unlinked from their parent taxa in its internal database structure, there was no unambiguous way to migrate some synonyms from Specify 5-to-6 so that they would be correctly linked to their original parent at the next higher rank. As a result, some synonyms in Specify 5-to-6 conversions were incorrectly attached to the synonymized parent. Note that none of the issues with parent-child relationships for converted synonyms affected the integrity of non-synonymic names in the Specify 5-to-6 conversions. **No determination records were changed, no Preferred Taxon names in Determinations records were modified, and all synonyms created in Specify 5 are present in Specify 6.**

This issue only applies to the parent node linkages for synonyms created in Specify 5. The Synonym Cleanup Tool identifies and attempts to fix *only misplaced parent relationships for synonyms.*

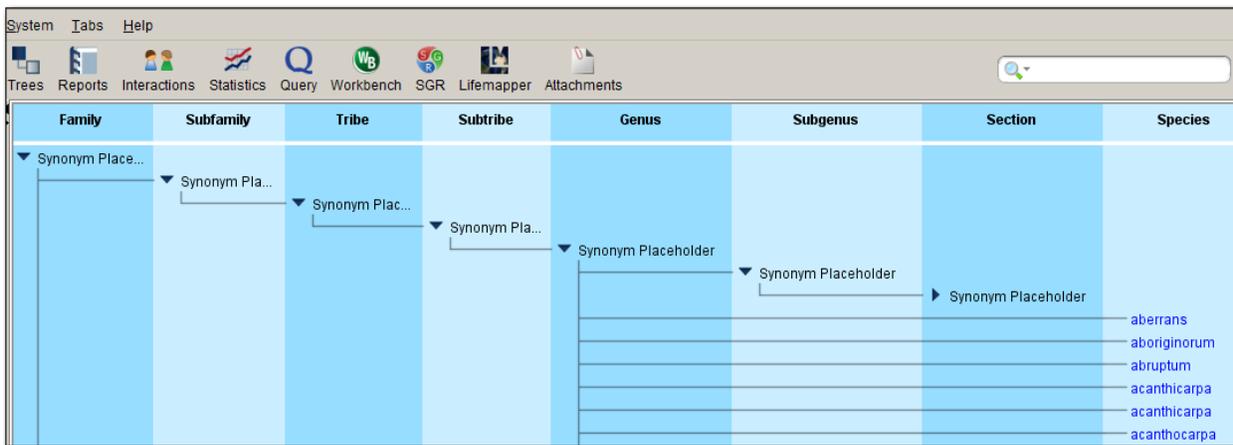
Although synonyms could be created at any rank in Specify 5, the Synonym Cleanup Tool only acts on synonyms at the level of species and subspecies. Although we use species in our examples below, the tool applies corrections to synonyms at both ranks. Reconstructing parent node linkages for taxa synonymized at higher or lower ranks is a semi-automated process, the Cleanup Tool will identify orphaned or misplaced synonyms of higher taxa and move them to the appropriate placeholder location in the tree, but they will need to be re-parented if desired by re-synonymizing them with drag-and-drop actions in the Taxon Tree or be deleted.

In the conversion from Specify 5-to-6, there are three possible outcomes for species and subspecies synonyms created in Specify 5 with regard to parent node links in Specify 6.

1. If in Specify 5 a species was synonymized to another species in the same genus, then the synonym was parented correctly under the original genus in Specify 6. These synonyms already appear in the correct place in the Specify 6 Taxon Tree, no cleanup or correction is needed.
2. If in Specify 5 a species was synonymized to a species in a different genus – during the conversion process the synonym was incorrectly re-parented in the tree to the second genus. (Again, Preferred Taxon name parent relationships were unaffected.) These species synonyms are not parented to their original genus as they should be. They will be automatically corrected by the Cleanup Tool.
3. If in Specify 5 a species was synonymized to a species in a second genus, AND that first genus was then deleted in the Taxon Tree, then like the above case, the synonym will be incorrectly parented to the second genus. The synonymic name needs to be moved back to its original genus, but the first genus is no longer in the database. The Synonym Cleanup Tool cannot correct that because it has no way of determining which family the original genus belongs to because the genus node for the parent of the original species name is gone. In this case the tool puts the species synonym in a newly-created “Synonym Placeholder” tree at the species rank. The original genus for such names can be seen in the Taxon table record for those synonyms. If one wants to retain the synonym in the database and have it placed correctly, the synonym

must be manually dragged and dropped to a new genus after the new genus (and any parents) is re-created in the Taxon Tree.

The Synonym Placeholder Tree or subtree functions as a holding position for synonyms which cannot be automatically re-parented. The Cleanup Tool creates the tree and adds it to your existing Taxon Tree display. The Synonym Cleanup Tool only works for species and sub-species level synonyms. The tool will identify un-parented synonyms at the ranks above genus and below subspecies, but it will not attempt to parent them in Specify 6. For those taxa, the Cleanup Tool moves the 'orphaned' synonyms to the appropriate rank in Synonym Placeholder Tree. The figure below shows a Synonym Placeholder Subtree with several species epithets which could not be linked to an existing genus in the tree and temporary assigned to a "Synonym Placeholder" genus entry.



Synonym Cleanup Report

The Synonym Cleanup Report is a list of the synonyms which are either incorrectly parented or which have no parent nodes in the Specify 6 database at the species or subspecies level. The figure below shows a portion of the first page of the report.

Orphan Synonyms for KUFishtissue						
Specify 5 Taxon Synonyms Found at Target Ranks:		258				
Synonyms used in Determinations:		206 79%				
Synonyms Correctly Parented:		34 13%				
Synonyms With Incorrect or No Parent Link:		224 86%				
Synonyms to be Linked to a New Parent:		222 86%				
Synonyms to be Moved to Synonym Placeholder Subtree:		2 0%				
Synonym Records in Error:		0 0%				
Update Errors:		0 0%				

	Orphan Synonym	Current Parent	Current Family	Proposed Parent	Proposed Family	Catalog Numbers Determined to Synonym
1	Acanthorhodeus macropterus	Acheilognathus	Cyprinidae	Acanthorhodeus	Cyprinidae	21437
2	Adioryx vexillarius	Sargocentron	Holocentridae	Adioryx	Holocentridae	21571
3	Adioryx xantherythrus	Sargocentron	Holocentridae	Adioryx	Holocentridae	18232, 18263
4	Agonus acipenserinus	Podothecus	Agonidae	Agonus	Agonidae	27358, 28411, 28040, 32415
5	Ammocrypta asprella	Crystallaria	Percidae	Ammocrypta	Percidae	29857, 2414, 4941, 9440, 9466, 9478, 9508, 9566, 10280, 10296, 15324, 23644, 31726, 31453, 29847, 31551
6	Anisitsia	Hemiodus	Hemiodontidae	Anisitsia	Hemiodontidae	5055, 5056

The table at the top is a statistical summary of the status of synonyms created in Specify 5 at the species and subspecies level (not for any subsequent synonyms created in Specify 6). In this example, 258 synonyms were found, of which 34 are correctly parented and 224 either have an incorrect parent or no parent assigned. 222 names can be corrected and re-assigned to their original parent node, but for 2 names the original parent does not exist in the Specify 6 tree, and they will be moved to the Synonym Placeholder Tree. The names relegated to the Synonym Placeholder Tree will stay there until they are re-parented or deleted. If there are no Determinations for those names, i.e. they have never been applied to Collection Objects, then they can be deleted without harm. If you want to keep unused synonyms for future reference and potential use, e.g. to identify in advance which names should or should not be applied to Collection Objects in your Collection, then they should be moved with the drag-and-drop 'move' capability in the Edit Taxon Tree window and be placed under the correct parent in your main tree.

The second table lists the Specify 5-created 'orphan synonyms' in the database, their current parent, the proposed parent and the catalog numbers of any collection objects which had previous determinations with that name. In the six synonyms shown, one can see that they are currently erroneously parented to the genus of the name to which they have been placed in synonymy. In each case shown the original parent of the synonym still exists in the database, and the Synonym Cleanup Tool will simply re-parent the synonym to its appropriate parent node.

After running the report, if there are Specify 5 synonyms identified which need cleaning up, it would be an excellent idea to save or print a copy of the report before proceeding to the actual cleanup operation.

Synonym Cleanup Tool

Choosing the “Cleanup Synonyms” button will make changes to your database as described above.

Before running this tool, backup your database. The Synonym Cleanup Tool could make hundreds or thousands of changes to synonyms in your database from Specify 5, and in the event of any kind of failure, a backup copy will be essential for recovery to protect months of work.

The Cleanup function will take a few minutes to run, and at the end it will automatically produce another report with the same information as before to provide one last chance to see which records were acted upon. If you have not already saved and/or printed a copy of the HTML report do so before going forward it will be useful for managing synonyms in the Placeholder Tree. Finally, if you close the report and choose the ‘Create Report’ button one more time, Specify will create a new report with the final results of the cleanup, and the new status of Specify 5 synonyms in your database. Ideally there should be no synonyms listed in with incorrect or no parent links, and zero synonyms identified to be moved to new parents or to the Placeholder Subtree, as shown in the table below.

Specify 5 Taxon Synonyms Found at Target Ranks:	256	
Synonyms used in Determinations:	204	79%
Synonyms Correctly Parented:	256	
Synonyms With Incorrect or No Parent Link:	0	0%
Synonyms to be Linked to a New Parent:	0	0%
Synonyms to be Moved to Synonym Placeholder Subtree:	0	0%
Synonym Records in Error:	0	0%
Update Errors:	0	0%

All of the changes the Synonym Cleanup Tool can make have been completed. The eventual disposition of any orphaned synonyms in the Synonym Placeholder Tree is dependent on your cataloging priorities. Please contact the Specify Software Project with any questions about moving or deleting synonyms from the Synonym Placeholder Tree.