

EDIT - bug #10475

Removing names from homotypic group does not always work

02/17/2024 01:19 PM - Andreas Müller

Status:	New	Start date:	
Priority:	New	Due date:	
Assignee:	Andreas Müller	% Done:	0%
Category:	cdmlib	Estimated time:	0:00 hour
Target version:	Release 5.43	Found in Version:	
Severity:	normal		
Description			
<p>This has been observed in Casearia where some names showed more types than expected. Research showed that these types came from names that were removed from taxonomy (or in 1 case moved to another taxon) in the meanwhile. However, the names did still bear type information which was therefore shown in print publication (other task: shouldn't the type module select only types from names shown in the synonymy or at least give a warning if such types/names exist?).</p> <p>The following query gives a list of names which have a homotypic group used at least in 2 taxa or in 1 taxon and a name not used in a taxon at all.</p> <pre>SELECT hid, COUNT(accId) naccId, GROUP_CONCAT(accId), GROUP_CONCAT(fullName), GROUP_CONCAT(nid) nid FROM (SELECT n.homotypicalGroup_id hid, GROUP_CONCAT(n.titleCache) fullname, GROUP_CONCAT(n.id) nid, CASE WHEN tb.DTYPE = 'TAXON' THEN tb.id WHEN tb.DTYPE = 'Synonym' THEN acc.id ELSE 'xxx' END accId FROM TaxonName n LEFT JOIN TaxonBase tb ON tb.name_id = n.id LEFT JOIN TaxonBase acc ON acc.id = tb.acceptedTaxon_id -- WHERE n.homotypicalGroup_id = 814 GROUP BY n.homotypicalGroup_id, accId) AS tmp GROUP BY hid HAVING naccID > 1 ORDER BY naccId DESC, hid, fullName</pre> <p>Be aware that names used in multiple taxa are also shown here, but can be recognized by duplicates in the nid column.</p> <p>In Caryophyllales there are 690 such homotypic groups (including a larger number of names being used in >1 taxon for some reason, but still there are >>100 remaining cases).</p>			
Related issues:			
Related to EDIT - bug #10357: CDMLight fix handling for names used as synonym...		Resolved	
Related to EDIT - discussion #10403: Should syn. sec. be part of the homotypi...		New	
Related to EDIT - feature request #10490: Exclude syntypes from typification ...		New	

History

#1 - 02/17/2024 01:28 PM - Andreas Müller

- Description updated

It still needs to be decided if this is primarily a cdmlib or a TaxEditor ticket.

#2 - 02/17/2024 01:33 PM - Andreas Müller

In Casearia at least in 1 (first) case the name that was not used in a synonym was originally in a synonym that was 1x updated but never deleted according to the AUD tables. This generally can not happen in a normal save. So either the record was deleted manually in the DB (I don't remember such an action) or something went wrong during saving the AUD table.

```
SELECT ev.*, tb.*
FROM TaxonBase_AUD tb INNER JOIN AuditEvent ev ON ev.revisionnumber = tb.REV
```

WHERE tb.name_id IN (xxx)

#3 - 02/17/2024 02:13 PM - Andreas Müller

- Related to bug #10357: CDMLight fix handling for names used as synonym and accepted taxa added

#4 - 02/17/2024 02:16 PM - Andreas Müller

AM:

I did some further research on this issue and found some issues.

Generally the problem arises when the handling of the homotypic groups is not correct as I already described below. This means that in the DB there are more names in the group than can be seen in TaxEditor taxon editor (synonymy). So there are 2 things we can do on this: (1) avoid that this happens and/or (2) remove those names and their types from the homotypic group during print publication output.

Fixing this for the output should be possible in reasonable time. We already recently had a similar issue where too many names were shown in the export (<https://dev.e-taxonomy.eu/redmine/issues/10357>). This needs to be done anyway to be on the safe side.

About (1): it is still not fully clear when exactly this happens. 1 possible case is when you try to delete a name from synonymy but the name record is still used elsewhere (e.g. in a name relationship as basionym or as an earlier homonym or as nameUsedInSource, ...). In this case the name is removed from synonymy but is still in the database AND is still in the homotypic group.

This is something that needs to be discussed and maybe changed. A solution is not trivial as in theory it is not forbidden to have a name in the database which belongs to the same homotypic group as other names but is for some reason (e.g. regional focus) not mentioned in the synonymy. The only thing that is definitely forbidden by taxonomic rules is a situation where 2 names belong to 2 taxa of the same classification and they share the same homotypic group (in the database) – except if 1 name is an autonym created from the other name.

In the below example e.g. a name record for the name “Samyda tomentosa Sessé & Moc.” adds additional type information to C. dentate. This name in some point of time must have been in the same homotypic group as C. dentate. Later you correctly removed it from there but for some unclear reason it was not fully removed from the database but stayed as an orphaned name including type information.

So to summarize, you probably did not do anything completely wrong but the main issue is probably a less accurate behaviour of the software.

#5 - 02/17/2024 02:52 PM - Andreas Müller

Further issues you have to have in mind is HG handling during reuse of names: this is usually an intended situation, however the handling of the homotypic group is unclear for the user here. If you reuse a name in >1 taxon (including misapplications) for the UI perspective it is difficult to make clear what names still belong to a HG.

Also it needs to be discussed how the HG of orth. var. is to be handled should original spellings or orth. var. have their own HG in the DB or should they get the HG of the correctly spelled name? As they refer to the same type the later seems to be correct from the taxonomic point of view.

#6 - 02/17/2024 03:07 PM - Andreas Müller

In general it needs to be discussed if the concept of a HG which is only referenced by name records and where 1 name record can have only exact 1 HG makes sense in practice. Closer to what we do in the UI is a HG which is more a "homotypic synonym group". So a homotypic group of 1 or more names belongs to in the context of a given taxon. This is mostly how we currently use HGs. Even if there are caveats in theory as (1) names can not be marked as belonging to the same HG without existing in a taxon or (2) names belonging to the same HG are not recognized as such if they are not mentioned in the same synonymy (but in 2 different ones that e.g. have different geographic focus or (3) autonyms and their original names can not belong to the same HG though in reality they are based on the same types.

However, in current practice all these caveats are of no real importance and/or, if necessary, can be handled in another way. E.g. (3) autonym relations could be handled via name relationships or (2) the full group of names that might belong to a "taxonomically correct" HG can still be computed (though computation is a bit more complex). For (1) and also (2+3) we could also have a "has same type as" name relationship as we had in the Berlin Model. Though this is definitely less comfortable.

#7 - 02/17/2024 03:09 PM - Andreas Müller

The "homotypic synonym group" may also go together with the idea of a syn. sec. that is attached to a HG rather than to a single synonym. ([#10403](#))

#8 - 02/17/2024 03:09 PM - Andreas Müller

- Related to discussion #10403: Should syn. sec. be part of the homotypical group instead of single synonym? added

#9 - 04/05/2024 08:03 AM - Andreas Müller

- Related to feature request #10490: Exclude syntypes from typification output if lectotype exists added

#10 - 04/05/2024 08:04 AM - Andreas Müller

- Tags changed from delete to delete, type designation

#11 - 04/11/2024 06:25 PM - Andreas Müller

- Description updated