



1

2

## ABSTRACT

3

4 *Premise of the study:* Standardization of plant names is a critical step in various fields of biology  
5 including biodiversity, biogeography and vegetation research. *WorldFlora* matches lists of plant  
6 names with a static copy from World Flora Online (WFO), an ongoing global effort of completing an  
7 online flora of all known vascular plants and bryophytes by 2020.

8

9 *Methods and results:* Based on direct and fuzzy matching, *WorldFlora* inserts matching cases from  
10 the WFO to a submitted data set of with taxa. Results of selecting the expected best single matches  
11 are presented for four data sets, including a working list of commercial timber tree species, a subset  
12 from *GlobalTreeSearch* and 2 data sets used in previous comparisons of software tools for correcting  
13 taxon names. The success rate of credible matches varied from 94.7 (568 taxa) to 99.9 (1740 taxa)  
14 percent.

15

16 *Conclusions:* *WorldFlora* offers a straightforward pipeline for semi-automatic plant name checking.

17

18 *Key Words:* Biodiversity informatics; Species name matching; Online flora; Plant name identification;  
19 Spelling errors; Taxonomic databases

20

1

2

## INTRODUCTION

3 Scientific names of organisms and the higher groups in which they are classified are key identifiers of  
4 the world's biodiversity (Rees, 2014). Assigning a species identify to an organism is essential in a  
5 wide array of disciplines including ecology, conservation and forestry (Tyrell, 2019). Removing  
6 synonyms from plant species lists is needed to predict the total number of vascular, seed and  
7 flowering plant species (Lughadha et al., 2016). Taxonomic uncertainty is one of the major gaps in  
8 plant occurrence data needed in global plant ecological, biogeographical and conservation  
9 applications (Meyer et al., 2016). Misspellings can lead to failures to retrieve data from global  
10 databases that encompass millions of species (Boyle et al., 2013). Matching alternative species  
11 names and resolving synonyms was for example essential in the development of the Agroforestry  
12 Species Switchboard, a portal that currently documents the presence of 172,395 plant species across  
13 35 web-based information sources through 307,404 hyperlinks (Kindt et al., 2019). If not corrected,  
14 lack of standardization of names can lead to inflated estimates of species richness. For example, a  
15 revision of a checklist of 11,675 Amazonian tree species through a taxonomic vetting process  
16 resulted in an updated list of 10,071 species, representing a 'loss' of around 15% of taxa (ter Steege  
17 et al., 2019). Issues of misspellings and unresolved synonyms increases the risks of erroneous  
18 scientific conclusions, such as misidentification of medicinal applications in plants, with potentially  
19 serious consequences (Sharma et al., 2019).

20 Target 1 of the Global Strategy for Plant Conservation aims to complete an online flora of all known  
21 plants by 2020. To achieve this goal, the World Flora Council was formed, including 36 participating  
22 institutions (Miller and Ulate, 2017). Building on The Plant List (<http://www.theplantlist.org>) that  
23 became static in 2013, the World Flora Online's (WFO) taxonomic backbone is actively curated by  
24 taxonomic specialists and Taxonomic Expert Networks (Palese et al., 2019). Static copies of the

Kindt. WorldFlora R package. p 4

- 1 Taxonomic Backbone data are available from the WFO website
- 2 (<http://www.worldfloraonline.org/downloadData>). The WorldFlora package matches a list of taxa
- 3 with the WFO taxonomic backbone and was the only package known to do this at article submission.
- 4

1

2

## METHODS AND RESULTS

3

*WorldFlora* is an R package that matches a list of plant names with a static version of the Taxonomic

4

Backbone data that can be downloaded from the World Flora Online portal

5

(<http://www.worldfloraonline.org/downloadData>; examples presented herein used version 2019.05

6

of 17-May-2019, the most recent version available at article submission). Data sets can be imported

7

in R in a wide variety of formats, with various available through the Graphical User Interface of the

8

R-Commander (Fox, 2005). Exact matching with the `WorldFlora::WFO.match` function can be

9

undertaken for the entire plant name or simultaneously for genus, species and infraspecific levels.

10

Fuzzy matching is implemented by `WFO.match` for the full plant name, calculating the Levenshtein

11

Distance (LD) via R functions of `base::agrep` (a user can modify its argument of `max.distance` from

12

the default 0.1) and `utils::adist`. The LD measures single character substitutions, insertions or

13

deletions, each contributing a value of 1 to the total distance. The Taxamatch algorithm (Rees,

14

2014), also implemented in the Taxonomic Name Resolution Service (TNRS; Boyle et al., 2013), uses

15

the Modified Damerau-Levenshtein distance where transpositions are given lower weights (2 in the

16

example of 'vecusilosus' and 'vesiculosus') than the weights of substituting each character

17

individually (4 in the example). Users can limit the number of matches to those with the smallest LD

18

by setting argument `Fuzzy.min = TRUE`, the default setting. The `WorldFlora::WFO.one` function finds

19

single matches for each submitted plant name by filtering records by accepted names and synonym

20

status. Information about the rationale for selection is given in a separate column in the output.

21

Where multiple candidates remained, the function selects the match with the smallest WFOID field.

22

Successful matches by *WorldFlora* are limited to the scope of the World Flora Online of vascular

23

plants and bryophytes, similar to software packages that use The Plant List such as *Taxonstand*

Kindt. WorldFlora R package. p 6

1 (Cayuela *et al.*, 2012). Therefore, users ideally should not attempt to resolve names from organism  
2 groups that are not covered such as algae, fungi or lichens (Wagner, 2016).

3

4

1 **Table 1.** Results of matching plant names from four data sets. The difference between correct and  
2 credible matches is clarified in the text. Details about access, data manipulation, outputs and R  
3 scripts are available from the appendices.

4

Data set	GTS	CTTS	Wagner	SALVIAS
Taxa	1,000	1741	600	1,000
Unique direct matches	930	1513	0	694
Several direct matches	27	125	0	40
Unique fuzzy matches	41	100	491	232
Several fuzzy matches	2	3	41	22
No match	0	0	68	12
Correct matches	957	1728	500	951
Credible matches	965	1740	568	975
Appendix	Appendix S1	Appendix S2	Appendix S3	Appendix S4

5

6 Note: GTS = Global Tree Search data set; CCTS = Commercial Timber Tree Species data set; Wagner =  
7 testing data set from Wagner (2016); SALVIAS = testing data set from Boyle et al. (2013).

8

9 Four data sets were used to check the performance of *WorldFlora* and to describe some of its  
10 features. The first is a random subset of 1,000 species selected from the GlobalTreeSearch (GTS)  
11 database (Beech et al., 2017; version 1.3 accessed from  
12 [https://tools.bgci.org/global\\_tree\\_search.php](https://tools.bgci.org/global_tree_search.php)). Of these, 957 species were matched directly (Table  
13 1). The remaining 43 were matched by the fuzzy algorithm. Where several matches were retrieved,  
14 the first option of finding the single best match was by selecting the record with the smallest LD

1 between the submitted and matched Authority (selecting the best author match is the default  
2 option for WorldFlora::WFO.one if the Authority variable is declared). This option resulted in 24  
3 single matches, including 21 matches with a LD between authorities of zero (for example, the  
4 selection of *Bauhinia grandifolia* (Bong.) D.Dietr. and rejection of *Bauhinia grandifolia* Steud.; see  
5 Appendix S1). The remaining five single matches were based on not selecting a synonym name, with  
6 4 of those having a LD between authorities of zero as by selecting *Xylosma intermedia* (Seem.)  
7 Griseb. and rejecting synonym *Xylosma intermedia* (Seem.) Triana & Planch. For 957 species, the  
8 name retrieved by *WorldFlora* was exactly the same as the submitted name (classified as 'correct' in  
9 Table 1). Among the names that were not matched exactly, three were spelling variants (*Callianthe*  
10 *sylvatica* vs. *Calanthe sylvatica*, *Euodia cuspidata* vs. *Evodia cuspidata* and *Guatteria moralesii* vs.  
11 *Guatteria moralesi*). Five names would have been matched if a scientist familiar with revision of the  
12 *Acacia* genus (Brummitt, 2004) would have substituted the genera *Senegalia* or *Vachellia* by *Acacia*,  
13 resulting in a tally of 965 (957 + 3 + 5; Table 1) 'credible' matches that can be achieved by *WorldFlora*  
14 in a typical supervised workflow. Of the 35 species without credible matches, 28 were partially  
15 matched at the genus level through the default option of argument 'Fuzzy.one' that searched for  
16 matches only for the first word, in case that the full name yielded no matches (Table2 provides an  
17 overview of various function arguments).

18



1 **Table 2.** Overview of some of the arguments of WorldFlora::WFO.match.

Argument	Details
Authorship	If this variable is found in the submitted data, the result will include a column of 'Auth.dist' with the LD between the submitted and matched naming author.
acceptedNameUsageID.match	In the default setting, where the WFO includes a acceptedNameUsageID (typically indicating the accepted name for a synonym), then the matched details in the results will show the accepted name and the columns of 'Old.status', 'Old.ID' and 'Old.name' those of the first match.
Fuzzy.max	The maximum number of names matched by the fuzzy algorithm.
Fuzzy.min	In the default setting (TRUE), only the matches with the smallest LD are retained.
Fuzzy.two	'Fuzzy.two' flags (TRUE/FALSE) whether the match was obtained for only the first 2 words of the submitted name
Fuzzy.one	'Fuzzy.one' flags (TRUE/FALSE) whether the match was obtained for only the first word of the submitted name
spec.name.nonumber	'Number.detected' flags (TRUE/FALSE) whether the submitted name contained a number. If that was the case, a match was searched only for the first word of the submitted name.
spec.name.tolower	Converts all characters of the submitted name to lower case, except the first character.

Argument	Details
spec.name.nobrackets	'Brackets.detected' flags (TRUE/FALSE) whether the submitted name contained a bracket. If that was the case, the function searched only for the part of the submitted name before the bracket.

1

2 The second data set used as a case study was a working list of 1,741 commercial timber tree species  
3 (CTTS; Mark et al., 2014). Names were matched directly for 1,638 taxa and by fuzzy matching for the  
4 remaining 103 names (Table 1). Of the fuzzy matches, 88 had a LD of 1 due to a space character at  
5 the end of the submitted name. The WFO.one function used the criterion of the smallest WFOID for  
6 14 names (for example selecting wfo-0000913932: *Scottellia coriacea* A.Chev. ex Hutch. & Dalziel  
7 and rejecting wfo-0000913938: *Scottellia coriacea* A.Chev.), but in only one case were the candidate  
8 species different (selecting synonym wfo-0000416717: *Ulmus glabra* and rejecting synonym wfo-  
9 0000475766: *Planera aquatica* for the submitted *Ulmus campestris*). It was confirmed for 28 species  
10 that where the original data set had included a synonym name between brackets, that this was  
11 indeed the accepted synonym name (for example, *Dysoxylum euphlebioides* to be a synonym of  
12 *Dysoxylum alliaceum*). Among the 13 species where the submitted and retrieved name was not  
13 exactly the same (not 'correct' as in Table 1), 12 were spelling variants such as *Pinus englemannii* vs.  
14 *Pinus engelmannii* or *Sequoiadendron giganteum* vs. *Sequoiadendron giganteum*. *Populus canescens*  
15 was matched by its hybrid name *Populus ×canescens*, correctly reflecting that the species is now  
16 accepted to be a hybrid of *Populus alba* × *Populus tremula* (Tutin et al., 1993). The only species  
17 without a credible match was *Upuna borneensis* (Randi et al., 2019) that was matched by *Upuna*  
18 *borensis*.

1 The third data set that was tested was a combination of three sub data sets initially created to  
2 compare different software packages for correcting plant names ('Wagner', Wagner, 2016). The  
3 argument Fuzzy.max was increased to 2,500 because an initial run of WFO.match could not resolve  
4 many of the names submitted at genus rank. As the testing procedure involved deleting the last  
5 character from a list of species names, all the matches were fuzzy, as might be expected (Table 1).  
6 One hundred of the names were not identical to the expected names (not 'correct' in Table 1). Of  
7 these, 60 were names of algae, fungi and lichen species outside the scope of World Flora Online and  
8 its predecessor The Plant List. Five of these names were spelling variants, one was an interspecies  
9 hybrid (*Carex acuta x elata*) and two were matched as varieties rather than the submitted  
10 subspecies names (*Keckiella antirrhinoides* var. *microphylla* for *Keckiella antirrhinoides* subsp.  
11 *microphylla*, *Saxifraga adscendens* var. *oregonensis* for *Saxifraga adscendens* var. *oregonensis*).  
12 Reasons that no acceptable matches were found for the remaining 32 species included 7 names  
13 where the submitted family is not included in the WFO (Aceraceae, Najadaceae, Punicaceae,  
14 Taccaceae, Theophrastaceae, Tiliaceae and Vittariaceae). None of these families were retained in the  
15 fourth update of the Angiosperm Phylogeny Group (2016) classification of orders and families of  
16 angiosperms. Seven were names where the number of fuzzy matches was too high to retrieve the  
17 genus name, while there were nine names where a submitted hybrid was matched by a non-hybrid  
18 name. In three cases, mismatches resulted from selecting the record with the smallest ID from  
19 otherwise valid fuzzy alternatives (selecting *Lomatia* R.Br. [Proteaceae] and not *Lomatium* Raf.  
20 [Apiaceae], *Placea* instead of Poaceae and *Sabicea* instead of Sabiaceae). With the overall matching  
21 success of 87.6 percent (219 out of 250 species) for spellchecking tests across different taxonomic  
22 ranks, WorldFlora performed better than the 74 percent achieved by the Global Names Resolver  
23 (GNR), which was the best software in the comparisons undertaken by Wagner (2016). With 99.6  
24 percent correct matches in comparisons across different geographic regions (only *Leptinella*  
25 *conjuncta* was not matched from the 250 submitted names), WorldFlora also marginally

1 outperformed GNR as the best software, with GNR giving 99.2 percent correct matches in the earlier  
2 comparison (Wagner, 2016).

3 The fourth data set that was analysed, 'SALVIAS', is a list of 1,000 plant names that was used earlier  
4 in a comparison of TNRS with other online tools for automated standardization of plant names  
5 (Boyle et al., 2013). The same data set was used in a more recent evaluation study (Sharma et al.,  
6 2019). This data set offers various challenges to plant name checking, such as inclusion of names in  
7 capital letters. This particular challenge is handled by WFO.match by option  
8 `spec.name.tolower=TRUE`, whereby submitted names are converted to lower case, except for the  
9 first character. The challenge of handling semi-standardized qualifiers that are used in plant names,  
10 such as 'cf.' that indicates that not all of the diagnostic characters correspond to a given species, or  
11 'aff.' that indicates that a specimen has some affinity but is not identical to a known species, is  
12 handled by deleting these characters from the name. Taxamatch (Rees, 2014) and TNRS (Boyle et al.,  
13 2013) use a similar approach. The default list of qualifiers used by WFO.match was derived from  
14 descriptions provided by Sigovini et al. (2016), also including qualifiers such as 'sp.', 'indet.' and  
15 'nom. nov.'. As WFO.match expects that the submitted plant name does not include the naming  
16 authority, when no matches are found with the submitted name that included the authority, a  
17 search is done for the first two words of the submitted name now expected to correspond to the  
18 genus and species names (default option for argument `Fuzzy.two`). If the search still does not find a  
19 name match, then a match is attempted for the first word only (default option for argument  
20 `Fuzzy.one`). Names that contain brackets, possibly at the beginning of the authority name, are  
21 stripped from the entire part starting with the bracket (default option of `spec.name.nobrackets`).  
22 Names that contain numbers are searched only for the first word, with the remaining part of the  
23 submitted name suspected to correspond to unidentified species (default option of  
24 `spec.name.nonumber`).

1 An initial run of WorldFlora.match with SALVIAS yielded 25 names with no match. Visual inspection  
2 of these unmatched names revealed cases where a qualifier 'cf' was used instead of the standard  
3 'cf.' or 'aff' instead of 'aff.'. Simulating how an actual semi-automatic pipeline of plant name  
4 checking would work, incomplete qualifiers were replaced by the correct qualifier that is recognized  
5 in the argument of 'sub.pattern'. Likewise, argument Fuzzy.max was increased to 2,000, as the  
6 output of WorldFlora.match indicated a series of names where the number of fuzzy matches was  
7 above the default 250, with a maximum of 1,974 fuzzy matches for the submitted 'Miconia'. The  
8 second run of WorldFlora.match with the modified names and arguments resulted in 734 directly  
9 matched names and 254 names with fuzzy matches (Table 1). Directly comparing the final subset of  
10 names obtained via WorldFlora.one with those obtained by TNRS  
11 (<http://tnrs.iplantcollaborative.org/TNRSapp.html>; accessed on 12-12-2019 with default settings)  
12 showed 951 identical names ('correct' in Table 1). Among the 49 names that were not identical,  
13 eight were spelling variants such as *Commiphora laxecymigera* vs. *Commiphora laxicymigera*. For 16  
14 of the names that were not identical, WorldFlora resulted in a more credible match than TNRS. For  
15 example, the submitted Asteraceae and Fabaceae families were correctly matched by WorldFlora,  
16 but matched incorrectly by the Asteliaceae or Fagaceae by the TNRS. For example, TNRS provided a  
17 partial match of the submitted *Solanum schlehtendali* to *Solanum*, whereas WorldFlora correctly  
18 matched *Solanum schlehtendalianum*. Accepting these credible matches by WorldFlora resulted in  
19 the correct matching of 975 names, a number close to the 980 correct matches reported for TNRS  
20 (Boyle et al., 2013). The number of matches is above the 950 of Solr-Plant for the same data set,  
21 which performed second best among TNRS, Solr-Plant, Plantminer and GNResolver (Sharma et al.,  
22 2019).

1

2

## CONCLUSIONS

3 Analysis of the four data sets showed that the new package WorldFlora matched plants names  
4 correctly for the majority of submitted names, with success rates comparable to and sometimes  
5 better than alternative software packages such as the Taxonomic Name Resolution Service or Solr-  
6 Plant. Users are advised not to blindly accept results, however, as limitations in plant name checking  
7 software and spelling mistakes in reference databases can result in incorrect matches (Wagner,  
8 2016). Users are therefore especially advised to screen the output of the WorldFlora::WFO.match  
9 and WorldFlora::WFO.one functions for information that could indicate possible mismatches, such  
10 as a large LD between the submitted and matched plant names, or the justification of a single match  
11 based on the smallest WFOID. Where possible, users are further advised to compare results  
12 obtained between different applications, for example, by submitting names both to WorldFlora and  
13 TNRS, first checking where both software packages agree on the accepted plant name.

14 As World Flora Online provides a further development of The Plant List that became static in 2013,  
15 matching of plant names with *WorldFlora* is expected to provide results that are more up-to-date  
16 than software packages that rely on The Plant List, such as Taxonstand (Cayuela et al., 2012). To  
17 benefit maximally from the current state of knowledge on plant names, users should ensure that  
18 they use the most recent version of the World Flora Online backbone database from the website.

19 Since naming authorities are not part of the plant name that is checked, users ideally should submit  
20 the plant name and the authority in separate fields to reduce mismatches.

21

22

1

2

## ACKNOWLEDGMENTS

3

4 RK thanks Ian Dawson (ICRAF) for reviewing the article prior to submission, and XX anonymous  
5 reviewers for further improving this manuscript. RK thanks the CGIAR Research Programs on Forests,  
6 Trees and Agroforestry (supported by the CGIAR Trust Fund) and the Provision of Adequate Tree  
7 Seed Portfolios project (supported by the Royal Norwegian Embassy in Ethiopia and NICFI) for  
8 supporting the author's time to contribute.

9

10

11

## DATA AVAILABILITY

12

13 The *WorldFlora* package is published under the GNU General Public License (Version 2). The  
14 software and related documentation are available for free download from the Comprehensive R  
15 Archive Network (<https://cran.r-project.org/package=WorldFlora>).

16

17

18

19

1

2

3

## LITERATURE CITED (MAX. 30)

- 4 Beech, E., M. Rivers, S. Oldfield, and P. P. Smith. 2017. GlobalTreeSearch: the first complete global  
5 database of tree species and country distributions. *Journal of Sustainable Forestry* 36: 454-489.
- 6 Boyle, B., N. Hopkins, Z. Lu, J. A. R. Garay, D. Mozzherin, T. Rees, N. Matasci et al. 2013. The  
7 taxonomic name resolution service: an online tool for automated standardization of plant names.  
8 *BMC Bioinformatics* 14: 16.
- 9 Brummitt, R. K. 2004. Report of the Committee for Spermatophyta: 55. Proposal 1584 on *Acacia*.  
10 *TAXON* 53: 826–829.
- 11 Cayuela, L., I. Granzow-de la Cerda, F. S. Albuquerque, and D. J. Golicher. 2012. Taxonstand: An r  
12 package for species names standardisation in vegetation databases. *Methods in Ecology and*  
13 *Evolution* 3: 1078-1083
- 14 Fox, J. 2005. The R Commander: A Basic Statistics Graphical User Interface to R. *Journal of Statistical*  
15 *Software* 14: 1-42.
- 16 Hackett, R. A., M. W. Belitz, E. E. Gilbert, and A. K. Monfils. 2019. A database management workflow  
17 of biodiversity data from the field to data users. *Applications in Plant Sciences* 7: e11310.
- 18 Kindt, R., I. John, J. Ordonez, I. Dawson, J.-P. B. Lillesø, A. Muchugi, L. Graudal, and R. Jamnadass.  
19 2019. Agroforestry Species Switchboard: a synthesis of information sources to support tree research  
20 and development activities. Version 2.0. Website  
21 <http://www.worldagroforestry.org/products/switchboard> [accessed 1 February 2020]
- 22 Lughadha, E. N., R. Govaerts, I. Belyaeva, N. Black, H. Lindon, R. Allkin, R. E. Magill, and N. Nicolson.  
23 2016. Counting counts: revised estimates of numbers of accepted species of flowering plants, seed



- 1 plants, vascular plants and land plants with a review of other recent estimates. *Phytotaxa* 272: 82-  
2 88.
- 3 Mark, J., A. C. Newton, S. Oldfield, and M. Rivers. 2014. The International Timber Trade: A working  
4 list of commercial timber tree species. Botanic Gardens Conservation International, Richmond,  
5 United Kingdom.
- 6 Miller, C. and W. Ulate. 2017. World Flora Online Project: An online flora of all known plants.  
7 Proceedings of TDWG 1: e20529.
- 8 Palese, R., C. Boillat, and P.-A. Loizeau. 2019. World Flora Online (WFO) – Quality control workflow  
9 for an evolving taxonomic backbone. *Biodiversity Information Science and Standards* 3: e35307.
- 10 Randi, A., V. Bodos, and J. Pereira, J. 2019. *Upuna borneensis*. *The IUCN Red List of Threatened*  
11 *Species* 2019: e.T33148A68075816.
- 12 Rees, T. 2014. Taxamatch, an Algorithm for Near ('Fuzzy') Matching of Scientific Names in Taxonomic  
13 Databases. *PLOS ONE* 9: e107510.
- 14 Sharma, V., M. I. Restrepo, and I. N. Sarkar. 2019. Solr-Plant: efficient extraction of plant names from  
15 text. *BMC Bioinformatics* 20: 263
- 16 Sigovini, M., E. Keppel, and D. Tagliapietra. 2016. Open Nomenclature in the biodiversity era.  
17 *Methods in Ecology and Evolution* 7: 1217–1225.
- 18 ter Steege, H., S. M. de Oliveira, N. C. A. Pitman, D. Sabatier, A. Antonelli, J. E. G. Andino, G. A.  
19 Aymard, and R. F. Salomao. 2019. *Scientific Reports* 9: 3501.
- 20 The Angiosperm Phylogeny Group, M. W. Chase, M. J. M. Christenhusz, M. F. Fay, J. W. Byng, W. S.  
21 Judd, D. E. Soltis, et al. 2016. An update of the Angiosperm Phylogeny Group classification for the  
22 orders and families of flowering plants: APG IV, *Botanical Journal of the Linnean Society* 181: 1-20.,  
23 Tutin, T.G., N. A. Burges, A. O. Chater, J. R. Edmondson, V. H. Heywood, D. M. Moore, D. H.  
24 Valentine, et al. 1993). *Flora Europaea Volume 1: PSILOTACEAE TO PLATANACEAE*. Cambridge  
25 University Press, Cambridge, United Kingdom.

- 1 Tyrell, C. D. 2019. A method to implement continuous characters in digital identification keys that  
2 estimates the probability of an annotation. *Applications in Plant Sciences* 7: e1247  
3 Wagner, V. 2016. A review of software tools for spell-checking taxon names in vegetation databases.  
4 *Journal of Vegetation Science* 27: 1323–1327.

## 6 TABLES

7 (Tables 1 and 2 were included in the main text)

## 9 APPENDIX

Filename	Description
Appendix S1 – SuppInfo.xlsx	<b>APPENDIX S1.</b> Details on the data, scripts and WorldFlora results for a subset of the GlobalTreeSearch (GTS) database.
Appendix S2 – SuppInfo.xlsx	<b>APPENDIX S2.</b> Details on the data, scripts and WorldFlora results for a working list of Commercial Timber Tree Species (CTTS).
Appendix S3 – SuppInfo.xlsx	<b>APPENDIX S3.</b> Details on the data, scripts and WorldFlora results for the ‘Wagner’ data set.
Appendix S4 – SuppInfo.xlsx	<b>APPENDIX S4.</b> Details on the data, scripts and WorldFlora results for the SALVIAS data set.

Kindt. WorldFlora R package. p 19

Appendix S5 – SuppInfo.txt	<b>APPENDIX S5.</b> Scripts and messages obtained during analysis of the four case studies.
----------------------------	---