



Counting counts: revised estimates of numbers of accepted species of flowering plants, seed plants, vascular plants and land plants with a review of other recent estimates

EIMEAR NIC LUGHADHA¹, RAFAËL GOVAERTS², IRINA BELYAEVA², NICHOLAS BLACK², HEATHER LINDON², ROBERT ALLKIN², ROBERT E. MAGILL³ & NICKY NICOLSON²

¹Conservation Science Department, Royal Botanic Gardens, Kew, Richmond, Surrey, TW9 3AE, UK; E-mail: e.niclughadha@kew.org

²Biodiversity Informatics and Spatial Analysis Department, Royal Botanic Gardens, Kew, Richmond, Surrey, TW9 3AE, UK

³Science and Conservation, Missouri Botanical Garden, P.O. Box 299, St. Louis, Missouri 63166, USA

Abstract

We present revised estimates of the numbers of accepted species of flowering plants (369,434), seed plants (370,492), vascular plants (383,671) and land plants (403,911) based on a recently de-duplicated version of the International Plant Names Index and rates of synonymy calculated from the seed plant families published in the *World checklist of selected plant families*. Alternative approaches to estimating or calculating the number of accepted plant species are discussed and differences between results are highlighted and interpreted.

Key words: global plant diversity, known species, described species estimates

Introduction

Accurate, evidence-based figures for numbers of plant species known to science are important to a wide diversity of end-users. Scientists and policy makers at national and regional level use such figures to interpret their data and prioritise their biodiversity in a global context (e.g. Forzza *et al.* 2012) and at national and international levels to evaluate progress against international targets and quantify the contribution of plant diversity to global biodiversity (e.g. Nic Lughadha *et al.* 2005, Paton *et al.* 2008). Individual plant diversity scientists use estimates of total known plant species to contextualise their own study group (e.g. Starr & Chouinard 2009).

In the absence of a single, complete and definitive list of all accepted names for plant species (i.e. those which have been formally described and are not generally considered to be synonyms), scientists have adopted three basic approaches to determine the total number of plant species known to science. Summing published estimates or counts for each taxonomic group was the basis for publications by Stebbins (1974), Thorne (2000, 2002) and, most recently, Christenhusz & Byng (2016). Bramwell (2002) focused on numbers of known species in countries rich in endemic species, and then estimated number of additional species in neighbouring countries to derive regional estimates that were combined to form a global estimate. Several authors (e.g. Govaerts 2001, Scotland & Wortley 2003) related numbers of accepted names for species to total numbers of names at species rank (synonymy rates) in taxonomic treatments (monographs, checklists) and extrapolated to reach global estimates for plant species. This last approach produced widely diverging estimates depending on the size and scope of the resource used to derive synonymy rates, with those based on small subsets of plant diversity dominated by temperate groups proving particularly unreliable (e.g. Govaerts 2003, Wortley & Scotland 2004).

Here, we present new estimates of total numbers of accepted flowering plants, seed plants, vascular plants and land plants resulting from a recent analysis of names at species rank in the *International plant names index* (IPNI, 2016) and using synonymy rates derived from the *World checklist of selected plant families* (WCSP, 2016). Our aim is to estimate the numbers of species in each group that have been validly described and are still accepted as distinct species (i.e. their names have not been relegated to synonymy). Our approach resembles that adopted in Paton *et al.* (2008) but differs in several key aspects detailed below.

Method and results

For each of the 173 families for which global checklists have been completed and published in the WCSP, we queried the live, regularly updated database and obtained totals for accepted names at species rank and all names at species rank (including synonyms, as at 1 April 2016). We plotted accepted names for species against all names for species and calculated a line of best fit for these data points, each representing a single family: $y = 0.3671x - 5.1898$ with an $R^2 = 0.9617$, indicating that the line fits the data well (Fig 1.).

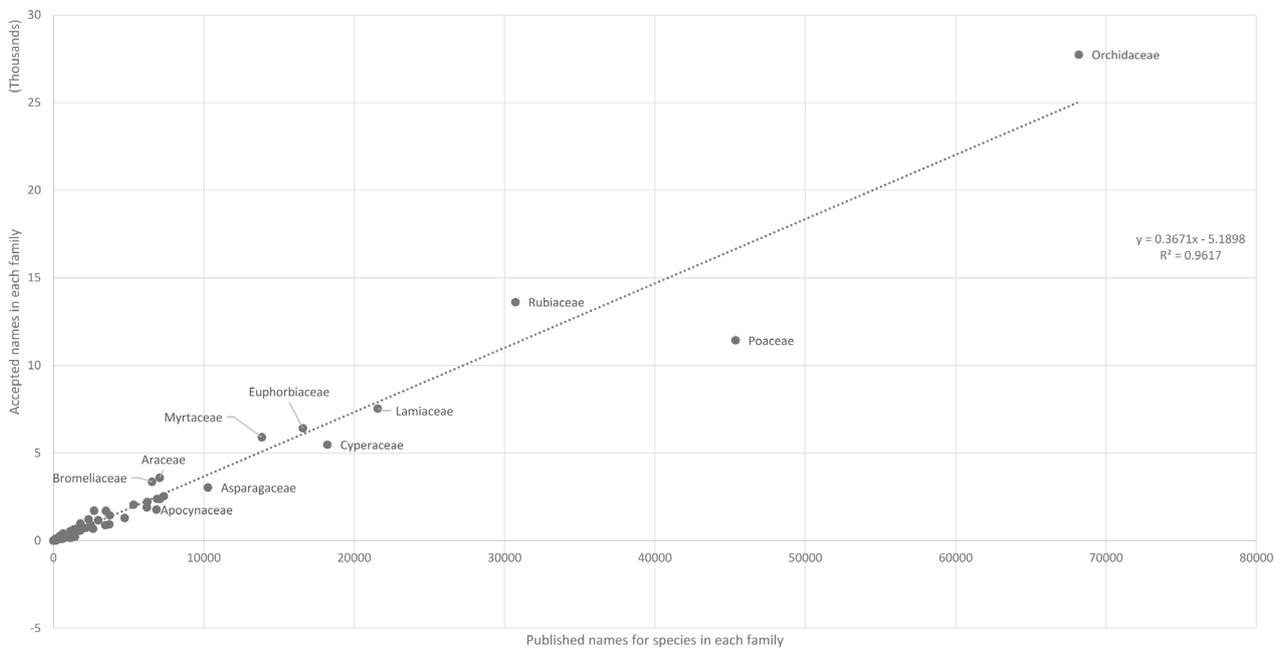


FIGURE 1. Accepted names at species rank as a function of all published species names for selected seed plant families (all published in WCSP).

In early 2016, as part of ongoing enhancements to the content of IPNI, the IPNI team completed a de-duplication process. This involved identifying duplicate IPNI entries that are records of the same bibliographic citation but derived from more than one of the three original source databases (*Index Kewensis*, the *Gray card index* and the *Australian plant names index*) that had been combined in the late 1990s to create IPNI (Croft *et al.* 1999). From the de-duplicated version of IPNI, we obtained the total number of name citations at species rank for vascular plants: 1,064,908 (as at 1 April 2016). Using family names, we mapped these citations to APG IV (2016) major groups to obtain numbers of species name citations for monocots, angiosperms and seed plants (Table 1).

TABLE 1. Number of species names for monocots, flowering plants, seed plants and vascular plants in de-duplicated IPNI and predicted numbers of accepted species based on a linear relationship between accepted names and all names.

Major group	Number of species names	Predicted number of species
Monocots	230654	84668
Flowering plants	1006371	369434
Seed plants	1011708	371393
Vascular plants	1064908	390923

We assumed that the linear relationship observed between accepted species names and all species names for seed plant families already completed for WCSP should also apply in general to seed plant families not yet published in WCSP. On this assumption, we applied the linear equation obtained above to the 1,006,371 species names citations for angiosperms in IPNI, a key source database for WCSP, and obtained an estimate of 369,434 accepted flowering plant species. Repeating this operation for the 1,011,708 seed plant species names suggested that they represent 371,393 accepted species. We also derived an alternative estimate for seed plant species by adding our flowering plant estimate

to published numbers for gymnosperms (1058 *vide* WCSP 2016) and obtained an estimate of 370,492 accepted species of seed plants.

We tested the assumption that the linear relationship observed between accepted species names and all species names for families published in WCSP would also apply to ferns and lycopods (groups that are not treated in WCSP but are indexed in IPNI). We applied the same linear equation to the 1,064,908 species name citations for vascular plants in de-duplicated IPNI and obtained an estimate of 390,923 accepted vascular plant species. We also derived an alternative estimate for vascular plant species by adding our flowering plant estimate to recently published numbers for gymnosperms (1058 *vide* WCSP 2016) and for ferns and fern allies (13269 *vide* Hassler 2016) and obtained an estimate of 383,671 accepted vascular plant species.

Finally, we estimated the total accepted species of land plants using the vascular plant estimate above and published estimates for bryophytes as follows. Crosby *et al.* (2000) recognised 12,754 species of mosses, whereas Soderstrom *et al.* (2016) listed 7486 species of hornworts and liverworts. Adding these estimates for bryophytes to our vascular plant estimate of 383,671, resulted in an estimate of 403,911 accepted species of land plants.

We tabulated evidence-based estimates of described plant species diversity published since 2000 to facilitate comparison and discussion (Table 2).

TABLE 2. Evidence-based estimates of plant species diversity published since 2000. **Preferred numbers are in bold.** [*Bracketed numbers in italics are inferred from other estimates*]. *See text for discussion of this number.

Author(s)	Year	Flowering plants	Seed Plants (includes flowering plants and gymnosperms)	Vascular Plants (includes seed plants, ferns and lycophytes)	Land Plants (includes vascular plants and bryophytes)	Method
Prance	2000	300,000 to 320,000				Based on Stebbins (1974) estimate and rates of description in the interim
Thorne	2000		257,400			Summing figures for taxonomic groups
Govaerts	2001		422,127			Ratio of accepted names in incomplete world checklist to all names
Bramwell	2002	421,968				Combining regional estimates
Thorne	2002	258,650				Summing published figures for taxonomic groups
Scotland & Wortley	2003	223,300				Proportion of accepted names to all names in some recent monographs
Wortley & Scotland	2004		346,527			Extrapolating synonymy rates from an extended set of monographs
Paton <i>et al.</i>	2008	352,282	[353,673]	[366,511]	379,881	Accepted names as a function of synonyms in WCSP extended to untreated families of angiosperms then adding published figures for non-flowering plant groups
Pimm <i>et al.</i>	2014				400,000	Proportion of accepted names to synonyms in TPL 1.0 extended to unresolved names
Pimm & Joppa	2015	400,000*				Proportion of accepted names to synonyms in TPL 1.0 extended to unresolved names
Christenhusz & Byng	2016	295,383	[296,462]	308,312	[334,000]	Summing counts from WCSP (some adapted) and from recent monographs, revisions and Flora treatments
This paper	2016	369,434	371,393	390,923		Extrapolated from line fitted to data for seed plant families in WCSP
This paper	2016		370,492	383,671	403,911	Addition to flowering plant estimate of recent published estimates for gymnosperms, and for ferns and lycophytes and for bryophytes

Assumptions and differences in methodology

Since our method has much in common with that of Paton *et al.* (2008), we outline the differences here before discussing our results in the context of theirs and those of other authors.

Firstly, Paton *et al.* (2008) estimated the number of unique species name citations in IPNI by counting the number of unique binomial strings at species rank. This produced a more conservative estimate of species name citations than our current method because it discounted not only duplicate entries of the same citation from different source databases but also (i) instances for which the same name is represented by two different citations within a single source database (e.g. in cases where a name is initially published as a *nomen nudum* and subsequently validated in a later publication) and (ii) homonyms. Paton *et al.* (2008) applied a 4.1% correction for homonyms and a 0.5% correction for orthographic and typographic variations but did not correct for having eliminated different citations of the same name from a single source database. We consider our current approach more consistent with the fact that both homonyms and subsequent citations that validate a name are regularly included in the synonymy of WCSP and therefore contribute to the synonymy rates observed for WCSP families and used in quantifying the linear relationship.

Secondly, Paton *et al.* (2008) included data points for some families not covered by WCSP (e.g. Fabaceae (Leguminosae), Gesneriaceae).

Third, Paton *et al.* derived their estimate for species of land plants by summing their new estimate for flowering plants with estimates published elsewhere for gymnosperms, ferns, lycopods and bryophytes. For our seed plant and vascular plant estimates, we compared this additive approach with that based on applying the linear model derived from seed plant data.

Discussion

Comparison with recent estimates for flowering plants

Given the similarity in methods, it is scarcely surprising that our estimate of 369,434 flowering plant species is close to that of Paton *et al.* (2008), 352,282. In fact, it could be argued that most of the difference between these two figures is due to the number of new species of flowering plants published in the years between their estimate and ours. Over the period 2007–2015 inclusive, a total of 19,237 names of newly published taxa at species rank were added to IPNI, the vast majority (c. 95%) representing new species of flowering plants. In addition to this growth in species names due to new descriptions, there are also many species names that have been pulled out of synonymy for ‘resurrected’ species. Increase in accepted species numbers due to such ‘resurrections’ is challenging to quantify because they do not involve a formal nomenclatural act, but such growth will have been offset in part by ongoing synonymization that is also nomenclaturally silent. That the two datasets represent different overall rates of synonymy is best illustrated by the difference in the slope of the line of best fit, calculated at 0.395 for the 2006–7 dataset and at 0.367 for the current dataset. The intercepts have changed too, but this difference is negligible when the model is applied to numbers in the tens and hundreds of thousands. Put simply, our analysis suggests that although new plant species continue to be published at a steady rate, the proportion of all flowering plant names at species rank that are recognised as being the correct name for an accepted species is 2.5–3.0% lower than estimated by Paton *et al.* (2008).

Our estimate of 369,434 flowering plant species is 7.6% lower than that of 400,000 recently published by Pimm & Joppa (2015), which was based on the proportions of accepted names and synonyms in Version 1 of *The plant list* (2010, hereafter TPL 1.0). A small portion of this difference is due to an apparent misreading by these authors of TPL 1.0 statistics, so that total numbers for vascular plants and bryophytes were interpreted as numbers for angiosperms. Correcting this error and applying Pimm & Joppa’s (2010) method to the angiosperm data published in TPL 1.0 results in a substantially lower estimate of 355,643 accepted species of flowering plant, whereas applying their method to the angiosperm data published in Version 1.1 of TPL (2013) yields a result of 394,065 flowering plant species. Our estimate based on WCSP and IPNI is still lower by 6.2%, a difference that may be attributable in part at least to Pimm & Joppa’s key assumption, which we consider to be unjustifiable. They assumed that the unresolved names listed by TPL would, once resolved, be found to contain the same proportions of accepted names and synonyms as is seen in the names already resolved. This assumption, which is also the basis for the Pimm *et al.* (2014) estimate for land plants, is unfounded, and, in our view, unlikely to prove correct when further names are resolved. Our experience, which includes close involvement in the development of TPL, is that the remaining unresolved names are comparatively unlikely to be accepted once resolved. A relatively low proportion of accepted names among the unresolved names is to be expected as a result of the method by which TPL was constructed. Some contributing data sets included nearly

all accepted names in the group on which they focused but with incomplete synonymy. Names not obtained from other sources were added from IPNI at the end of TPL compilation and treated as unresolved in the absence of clear evidence to the contrary. In light of this insight, we might assume, for example, that the proportion of unresolved names that will eventually be accepted is perhaps 25% lower than that seen in names already resolved (30% rather than 40–41%); the resulting estimate for species of flowering plants would be 369,332.

Our estimate of 369,434 accepted species of flowering plants is much higher than the count of 295,383 published by Christenhusz & Byng (2016) while this paper was in review. These authors report that they generally used species numbers from WCSP where available, adapting some in light of new evidence, whereas for families not treated in WCSP they consulted the most recent monographs, revisions and flora treatments. Without a detailed understanding of their methods (e.g. how they reconciled or chose between conflicting estimates, whether they added species published subsequent to a monographic treatment etc.), it is difficult to suggest evidence-based explanations for the gap between our results and theirs. We consider that the difference may be partly due to monographs and revisions being outnumbered by floristic treatments among their sources, which could lead to the omission of little-known species that nonetheless might be accepted by a botanist adopting a more comprehensive approach. This idea is borne out by the experience of those involved in compilation of families for WCSP: once a checklist has been compiled, the number of accepted species is typically found to be at least 20% greater than other species numbers previously published for the family (R. Govaerts, pers. comm.). This effect is particularly noticeable in families containing large genera, for which recent revisions are generally lacking. For example, Myrtaceae were estimated to include 4620 species (Mabberley 1997), but the compilation of a complete list with input from a dozen specialists on the family resulted in a revised figure of 5671 (Govaerts *et al.* 2008), which had risen to 5894 by the time of the analysis reported here (WCSP, 2016).

Comparison with previous estimates for seed plants

Our estimate of 371,393 species of seed plants is higher than all but one of the seed plants numbers published over the review period (Table 2). Obviously, this estimate is dominated by the flowering plant data. Since up-to-date synonymised checklists are available for all gymnosperms, the number of accepted species does not need to be estimated (1058 *fide* WCSP 2016). The difference between our flowering plant and seed plant estimates is 1959, considerably higher than the number of known species of gymnosperms, a discrepancy that reflects the relatively high synonymy rate in conifers, attributed to their predominantly temperate distribution and too much attention from horticulturalists (Govaerts 2003).

Comparison with previous estimate for vascular plants

Global estimates of accepted vascular plant species are published less frequently than those for flowering plants or seed plants, but an estimate of 366,511 can be inferred from Table 1 in Paton *et al.* (2008; by adding the species numbers for non-flowering plants to the published figure of 352,282 for flowering plants). Our figure derived from this additive approach is 383,671, and almost all of this discrepancy results from our higher estimate for flowering plants (discussed above). Our estimate resulting from extrapolation is more than 7,000 higher at 390,923. However, this extrapolated figure can be argued to be less strongly evidence-based than that derived from the additive approach because, taken in conjunction with our flowering plant figure, it would imply that there were > 21,000 accepted non-flowering vascular plant species. Since gymnosperms number c. 1,000, the balance of accepted species, over 20,000, could only be ferns and lycopods, a figure greatly in excess of any recent published estimates. Such estimates vary from 11,835 obtained by the additive method (Christenhusz & Chase 2014) to the checklist-based figure of 13,269 species of ferns and lycopods used in our analysis (Hassler 2016). IPNI holds 53,200 names at species rank for ferns and lycopods, indicating that there are 4.0–4.5 names for each species. In other words c. 25% of all species names for ferns and lycopods are accepted, indicating much higher synonymy rates than are generally considered to prevail in seed plants, where the proportion of species names that are accepted have variously been estimated as 34% (Wortley & Scotland 2004) or 39.5% (Paton *et al.* 2008), with an earlier estimate of 22% (Scotland & Wortley 2002) being discredited as based on a small and unrepresentative sample (Govaerts 2003). The exceptionally high synonymy rates in ferns and lycopods are a product of centuries of flux in fern classification with species being placed in genera by guesswork and subsequently corrected in light of new knowledge (Christenhusz & Chase 2014), generating much homotypic synonymy. The disparity between our estimates for vascular plants and the 308,312 estimate by Christenhusz & Byng (2016) is likely attributable to the factors discussed under flowering plants above.

Comparison with previous estimates for land plants

Our estimate for accepted land plant species is 24,000 more than that obtained by Paton *et al.* (2008) using a similar

method but close to that of Pimm *et al.* (2014), who used a different method. Differences between our result and that of Paton *et al.* are attributable mainly to our different estimates of known flowering plant diversity, which differ by >17,000, as discussed above. The remaining difference (c. 7,000) is attributable to the omission by Paton *et al.* of numbers for hornworts and liverworts from their bryophyte estimate, which is based only on numbers of mosses. Estimates of accepted moss species have changed relatively little over the intervening period (Magill, pers. comm.).

Conclusion

Our best estimates of species for major groups of plants are: 369,434 flowering plants, 370,942 seed plants, 383,671 vascular plants and 403,911 land plants. We consider that these estimates will be of use to those involved in monitoring progress in documenting, evaluating and conserving global plant diversity.

References

- Bramwell, D. (2002) How many plant species are there? *Plant Talk* 28: 32–34.
- APG IV. (2016) An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean Society* 181: 1–20.
<http://dx.doi.org/10.1111/boj.12385>
- Christenhusz, M.J. & Byng, J.W. (2016) The number of known plants species in the world and its annual increase. *Phytotaxa* 261 (3): 201–217.
<http://dx.doi.org/10.11646/phytotaxa.261.3.1>
- Christenhusz, M.J. & Chase, M.W. (2014) Trends and concepts in fern classification. *Annals of Botany* 113: 571–594.
<http://dx.doi.org/10.1093/aob/mct299>
- Croft, J., Cross, N., Hinchcliffe, S., Nic Lughadha, E., Stevens, P.F., West, J.G. & Whitbread, G. (1999) Plant names for the 21st century: the International Plant Names Index, a distributed data source of general accessibility. *Taxon* 48: 317–324.
<http://dx.doi.org/10.2307/1224436>
- Forzza, R.C., Baumgratz, J.F.A., Bicudo, C.E.M., Canhos, D.A., Carvalho, A.A., Coelho, M.A.N., Costa, A.F., Costa, D.P., Hopkins, M.G., Leitman, P.M., Lohmann, L.G., Nic Lughadha, E., Maia, L.C., Martinelli, G., Menezes, M., Morim, M.P., Peixoto, A.L., Pirani, J.R., Prado, J., Queiroz, L.P., Souza, S., Souza, V.C., Stehmann, J.R., Sylvestre, L.S., Walter, B.M.T. & Zappi, D.C. (2012) New Brazilian floristic list highlights conservation challenges. *BioScience* 62: 39–45.
<http://dx.doi.org/10.1525/bio.2012.62.1.8>
- Crosby, M.R., Magill, R.E., Allen, B. & He, S. (2000) A checklist of the mosses. Missouri Botanical Garden, St. Louis. Published on the Internet: <http://www.mobot.org/MOBOT/tropicos/most/checklist.shtml> (accessed 1 August 2016)
- Govaerts, R. (2001) How many species of seed plants are there? *Taxon* 50: 1085–1090.
<http://dx.doi.org/10.2307/1224723>
- Govaerts, R. (2003) How many species of seed plants are there? – a response. *Taxon* 52: 583–584.
<http://dx.doi.org/10.2307/3647457>
- Govaerts, R., Sobral, M., Ashton, P., Barrie, F., Holst, B.K., Landrum, L.L., Matsumoto, K., Mazine, F.F., Lughadha, E.N., Proença, C., Soares-Silva, L.H., Wilson, P.G. & Lucas, E. (2008) *World checklist of Myrtaceae*. Royal Botanic Gardens.
- Hassler, M. (2016) World ferns: checklist of ferns and lycophytes of the world. In: Roskov, Y., Abucay, L., Orrell, T., Nicolson, D., Kunze, T., Flann, C., Bailly, N., Kirk, P., Bourgoin, T., DeWalt, R.E., Decock, W. & De Wever, A. (Eds.) *Species 2000 & ITIS catalogue of life, 25th March 2016*. [www.catalogueoflife.org/col. Species 2000: Naturalis, Leiden, the Netherlands. ISSN 2405-8858]
- Mabberley, D.J. (1997) *The plant-book: a portable dictionary of the vascular plants*. Cambridge University Press, Cambridge, 858 pp.
- Nic Lughadha, E., Bailie, J., Barthlott, W., Brummitt, N.A., Cheek, M.R., Farjon, A., Govaerts, R., Hardwick, K.A., Hilton-Taylor, C., Meagher, T.R., Moat, J., Mutke, J., Paton, A.J., Pleasants, L.J., Savolainen, V., Schatz, G.E., Smith, P., Turner, I., Wyse-Jackson, P. & Crane, P.R. (2005) Measuring the fate of plant diversity: towards a foundation for future monitoring and opportunities for urgent action. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 360: 359–372.
<http://dx.doi.org/10.1098/rstb.2004.1596>
- Paton, A. J., Brummitt, N., Govaerts, R., Harman, K., Hinchcliffe, S., Allkin, B. & Nic Lughadha, E. (2008) Towards Target 1 of the *Global Strategy for Plant Conservation*: a working list of all known plant species—progress and prospects. *Taxon* 57: 602–611.
- Pimm, S.L., Jenkins, C.N., Abell, R., Brooks, T.M., Gittleman, J.L., Joppa, L.N., Raven, P.H., Roberts, C.M. & Sexton, J.O. (2014) The

- biodiversity of species and their rates of extinction, distribution, and protection. *Science* 344: 1246752.
<http://dx.doi.org/10.1126/science.1246752>
- Pimm, S.L. & Joppa, L.N. (2015) How many plant species are there, where are they, and at what rate are they going extinct? *Annals of the Missouri Botanical Garden* 100: 170–176.
<http://dx.doi.org/10.3417/2012018>
- Prance, G.T., Beentje, H., Dransfield, J. & Johns, R. (2000) The tropical flora remains undercollected. *Annals of the Missouri Botanical Garden* 87: 67–71.
<http://dx.doi.org/10.2307/2666209>
- Scotland, R.W. & Wortley, A.H. (2003) How many species of seed plants are there? *Taxon* 52: 101–104.
<http://dx.doi.org/10.2307/3647306>
- Söderström, L., Hagborg, A., von Konrat, M., Bartholomew-Began, S., Bell, D., Briscoe, L., Brown, E., Cargill, D.C., Costa, D.P., Crandall-Stotler, B.J. & Cooper, E.D. (2016) World checklist of hornworts and liverworts. *PhytoKeys* 59: 1–828.
<http://dx.doi.org/10.3897/phytokeys.59.6261>
- Starr, J.R., Naczi, R.F. & Chouinard, B.N. (2009) Plant DNA barcodes and species resolution in sedges (*Carex*, Cyperaceae). *Molecular Ecology Resources* 9: 151–163.
<http://dx.doi.org/10.1111/j.1755-0998.2009.02640.x>
- Stebbins, G.L. (1974) *Flowering plants: evolution above the species level*. Arnold, London, 399 pp.
<http://dx.doi.org/10.4159/harvard.9780674864856>
- The international plant names index (2016) Published on the Internet: <http://www.ipni.org> (accessed 31 March 2016)
- The plant list (2010) Version 1. Published on the Internet: <http://www.theplantlist.org/> (accessed 20 April 2016)
- The plant list (2013) Version 1.1. Published on the Internet: <http://www.theplantlist.org/> (accessed 20 April 2016)
- Thorne, R.F. (2000) The classification and geography of the flowering plants: dicotyledons of the class Angiospermae. *The Botanical Review* 66: 441–647.
<http://dx.doi.org/10.1007/BF02869011>
- Thorne, R.F. (2002) How many species of seed plants are there? *Taxon* 51: 511–512.
<http://dx.doi.org/10.2307/1554864>
- WCSP (2016) *World checklist of selected plant families*. Facilitated by the Royal Botanic Gardens, Kew. Published on the Internet: <http://apps.kew.org/wcsp/> (accessed 1 April 2016)
- Wortley, A.H. & Scotland, R.W. (2004) Synonymy, sampling and seed plant numbers. *Taxon* 53: 478–480.
<http://dx.doi.org/10.2307/4135625>