NOMENCLATURE

Edited by Jefferson Prado, James Lendemer & Erin Tripp

Citation of a taxon name identifier issued by the *ICN*-recognized registration repositories instead of taxon name author citation

Gerhard Rambold,[1] Konstanze Bensch,[2,3] Paul M. Kirk,[4] Yi-Jian Yao,[5] Vincent Robert,[3] Veronica Sanz[2] & Dagmar Triebel[2]

1 *Department of Mycology, University of Bayreuth, Universitätsstraße 3, 95440 Bayreuth, Germany*
2 *SNSB IT-Center, Botanische Staatssammlung München, Menzinger Straße 67, 80638 München, Germany*
3 *Westerdijk Fungal Biodiversity Institute, Uppsalalaan 8, 3534 CT Utrecht, The Netherlands*
4 *Biodiversity Informatics & Spatial Analysis, Jodrell Laboratory, Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3DS, U.K.*
5 *State Key Laboratory of Mycology, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China*
Author for correspondence: *Dagmar Triebel, triebel@bsm.mwn.de*
ORCID DT, http://orcid.org/0000-0003-1980-3148

DOI https://doi.org/10.12705/665.12

Abstract Problems with taxon name author citation in the running text of scientific articles are outlined. Particularly, the length of author names strings has significantly increased in the last decades, which could be demonstrated by analysis of names in the LIAS database. In addition, ambiguity of author name citation has increased because of variations in author name abbreviation and incomplete or incorrect citation of authors' names. It is therefore suggested that in non-taxonomic scientific publications, author names should be replaced by taxon name identifiers as issued by registration repositories. The advantage of such approach especially in context with linked open data and semantic network challenges is described. An additional recommendation in Article 46 of the *ICN* with an example is suggested.

Keywords *ICN* Art. 46; identifier; name authors; registration repositories; scientific names; taxon names

INTRODUCTION

A name in the sense of the *International Code of Nomenclature for algae, fungi, and plants* (*ICN*) (McNeill & al., 2012) is a name for an algal, plant or fungal taxon (see Art. 6.3). These names are regarded as correctly cited when the author(s) of the taxon names (= author names or their abbreviations) are given, consisting of the author name(s) for the name of a new taxon or a replacement name and, in the case of a new combination or name at new rank, the basionym author name(s) in brackets, followed by the combining author name(s). This type of citation has a long tradition, for making uninomial, binomial or trinomial names unambiguous. Therefore, the *ICN* gives in Art. 46.1 the following statement: "In publications, particularly those dealing with taxonomy and nomenclature, it may be desirable, even when no bibliographic reference to the protologue is made, to cite the author(s) of the name concerned (see also Art. 22.1 and Art. 26.1)."

The author(s) of scientific names together with the citation of the protologue, along with the concept of typification, in turn are applied as a "code" for referencing a nomenclatural act. By including the basionym author(s) for the taxon names, the nomenclatural history of the systematic and phylogenetic assignment of a taxon is provided to some extent. Author names citation concern scientific names at species level as well as at supraspecific and infraspecific level, and are not necessarily identical with the citation of author(s) of the publication in which the names appear, and are not considered a part of a scientific name citation unless bibliographic data are appended. However, the author(s) of scientific names provide a kind of link to the original publication.

The use of nomenclatural acts as kind of an impact criterion in taxonomy has a long tradition. Already in early obituaries, lists of described taxa or of recombined names were presented with the purpose of indicating the productivity of an honoured person. In recent years, however, the multiplication of published author names for scientific names has reached quite an excessive extent, correlated not only with an increase in multi-author articles, but also a kind of misuse since authors of the paper and authors of the names not necessarily need to be identical.

LENGTH OF AUTHORS CITATION STRING IN TAXON NAMES

For showing the historical development of taxon name author citations, the LIAS names dataset (Triebel & Bensch,

Received: 2 May 2017 | returned for (first) revision: 8 Jun 2017 | (last) revision received: 26 Jun 2017 | accepted: 27 Jun 2017 || publication date(s): online fast track, 14 Sep 2017; in print and online issues, 24 Oct 2017 || © International Association for Plant Taxonomy (IAPT) 2017

2005–) was studied for each year from 1850 until recently. This was done by comparing the author(s) name string, e.g., "Fink ex J. Hedrick" for *Acarospora immersa* Fink ex J. Hedrick and the names of the string of basionym+combining author(s) "(Fink ex J. Hedrick) K. Knudsen & L. Arcadia" for *Caeruleum immersum* (Fink ex J. Hedrick) K. Knudsen & L. Arcadia, by annual average length over time. This analysis demonstrates a striking tendency of an increase of the length of author(s) names strings in fungal names over time (Fig. 1).

One of the longest author name strings for a combination that has been found in LIAS names, *Massjukiella kaernefeltii* (S.Y. Kondr., D.J. Galloway & Goward) S.Y. Kondr., Fedorenko, S. Stenroos, Kärnefelt, Elix, Hur & A. Thell, was published in recent years and comprises 108 letters (incl. blanks), which is even three times the length of a Universal Unique Identifier (version 4) (https://en.wikipedia.org/wiki/Universally_unique_identifier). An UUID comprises 36 digits (incl. four dashes), but in contrast comprises enormous information storage capacity. In the dataset of MycoBank an author name string with even 220 (!) letters (incl. blanks) could be found, *Carlosrosaea vrieseae* (M.F. Landell, L.R. Brandão, S.V. Safar, F.C. Gomes, C.R. Félix, A.R. Santos, D.M. Pagani, J.P. Ramos, L. Broetto, T. Mott, M.H. Vainstein, P. Valente & C.A. Rosa) A.M. Yurkov, X.Z. Liu, F.Y. Bai, M. Groenew. & Boekhout (http://www.mycobank.org/MB/814757).

The length of author names strings is already addressed in the *Code* by Rec. 46C.2. There, the option to shorten a citation of three or more authors to the first author followed by "et al." or "& al." is suggested. The result, however, is a loss of information.

The length of the author names string together with the lack of precision and the lack of standardization in this code might be the two reasons, why nowadays most biological journals demand that authors of papers either cite the author(s) of names only once when the name is cited for the first time in the text or that authors of taxon names should be omitted in general, unless nomenclatural novelties and taxonomic treatments are concerned.

For authors of non-taxonomic scientific publications, the citation of taxon names with correct taxon name author(s) might become a time-consuming challenge while preparing a manuscript. They have to check the correct author name spellings and recommended abbreviations according, e.g., to the IPNI Authors of Plant Names live database (http://www.ipni.org/ipni/authorsearchpage.do). This database continuously updates the TDWG (Biodiversity Information Standards) 2007 prior standard "Authors of plant names" (see http://www.tdwg.org/standards/) based on Brummitt & Powell (1992). However, it is not complete and mycologists, e.g., prefer to look for author names and abbreviations in the fungal nomenclatural repositories as recognised 2012 (Redhead & Norvell, 2012, 2013; Redhead & al., 2014). These repositories partly issue author names using the same TDWG prior standard. In addition, they create citations for new authors and – even among the three current fungal name registration repositories – the author citations for one and the same author may differ. In future, authors of manuscripts will have the option to query author name spellings in a number of recognized registration sites. The sites devoted by the *ICN* (Barkworth & al., 2016) focus on various organism groups and will include an overlap in author names. This will entail the problem that even more abbreviation variants for one and the same author of a scientific name will exist in parallel.

■ TAXON NAME IDENTIFIERS AND RECOMMENDATION FOR *ICN* ART. 46

Since 2012, registration of new names of fungi has been compulsory under the *ICN*. A consortium for the registration of fungal names, comprising the three repositories Fungal Names, Index Fungorum and MycoBank, is in place. This consortium is organising registration of new fungal names and provides

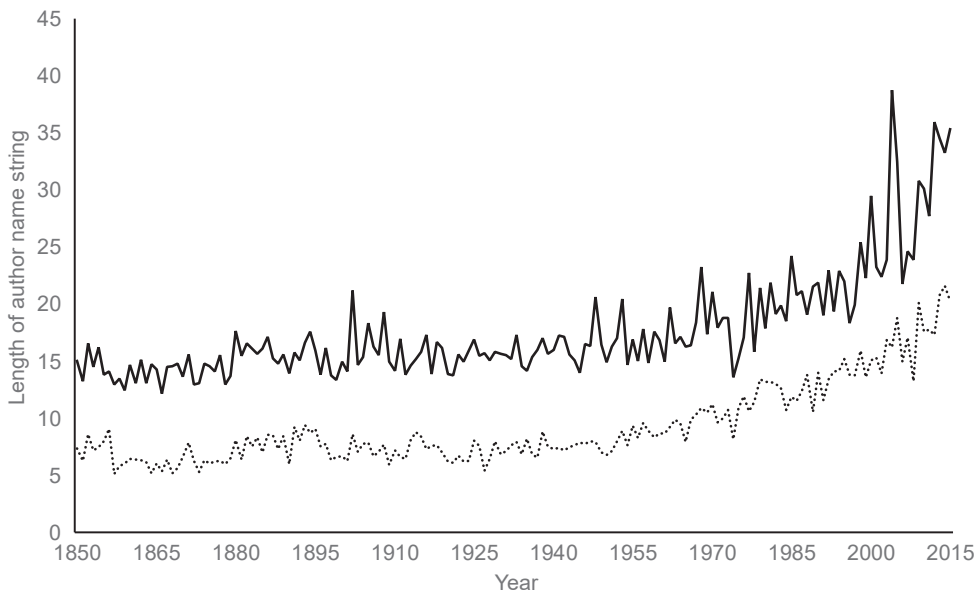


Fig. 1. Progression from 1850 of the annual average length of author names strings for new scientific taxon names or replacement names (dotted line) and new combinations, where the string is the combination of author names and basionym author names that are in brackets (solid line). Values are based on author strings for 42,065 taxon name records (27,090 new taxon names or replacement names and 14,975 new combinations) from LIAS names (http://liasnames.lias.net; version of 30 Nov 2016).

identifying codes (i.e., taxon name identifiers) that are assigned to newly published taxon names. The numbers used by each partner in the consortium come from separate defined number ranges.

These strings of digits that are taxon name identifiers could also function to replace author name strings. With only six digits, they are rather short so far, and are therefore easily human-readable. A taxon name like "*Lecanora varia* (Hoffm.) Ach." would therefore read as "*Lecanora varia* [#389546]". Even with abbreviated genus name ("*L. varia* [#389546]") this kind of taxon name identifier fulfils all requirements with regard to unambiguity of the scientific name and easy human and machine readability. The combination of the taxon name as the recognizable element and the taxon name identifier, issued by one of the registration repositories, is a comfortable way to connect information in non-taxonomic scientific publications. The entire string (scientific taxon names+taxon name identifier) is unique. Such an option of replacement of author names or author name abbreviations with a code will particularly support automated hypertext linking to the respective data record. This will be facilitated by the current use of registration numbers in two of the fungal taxon names repositories, as, for example: http://www.indexfungorum.org/names/NamesRecord.asp?RecordID=389546 and http://www.mycobank.org/mb/389546.

Proofreading the correctness of such taxon name identifiers together with the scientific taxon names could be done automatically by appropriate IT tools before submitting a paper or as part of the submission process guided by the editorial manager functions of the publishers. This would be significantly less prone to error and less time-consuming than the proofreading of the author names or its abbreviations.

For mapping names, name elements and name-related contents in the digital world, there is a strong need for web services and other tools. Such tools frequently are not capable to properly parse taxon names and author names (due to misspellings and varying abbreviation concepts), and therefore skip the latter ones, which produces ambiguity of homonymous names. The proposed kind of name citation will support more reliable name services.

For the reasons mentioned above we suggest a new **Recommendation to Article 46** along the following lines

"*46E.1.* Once a stable taxon name identifier issued by an *ICN*-recognized registration repository (see also Art. 22.1 and 26.1) is available, it might be comfortable and desirable to cite this identifier in the running text of a paper instead of the author names of the scientific taxon name (but not replacing the scientific taxon name itself)."

Furthermore, we suggest a new **example under the new Recommendation 46E** along the following lines:

"*Ex. 1. Astrothelium meristosporoides* [#816706] or *Lecanora varia* [#389546] or *Talaromyces mangshanicus* [#570288]."

Both suggestions were proposed in a simplified wording as amendment from the floor at the Nomenclature Section at the XIX International Botanical Congress (Shenzhen, China, July 2017), but were rejected.

■ OPTIONS AND IMPLICATIONS

At this Nomenclature Section 2017, a permanent Registration Committee was approved (proposal of Barkworth & al., 2016, accepted). However, registration repositories and the registration of new names in groups covered by the *ICN* other than fungi were not mandated. In parallel, starting with 2017, proposals on matters related solely to names of organisms treated as fungi will be dealt with at Nomenclature Sessions of International Mycological Congresses (IMC) (proposal of May & al., 2016, accepted). With this new situation, the recommendation presented here has to be reworded in near future and best confined to fungi where registration became mandatory already in 2013. We intend to submit a newly adapted proposal for the upcoming IMC next year. The situation for fungi provides a concrete example of the possibilities of using stable taxon name identifiers for registered and published names. For fungal names, two of the three existing repositories, Index Fungorum and MycoBank, have already achieved this by implementing Uniform Resource Locators (URLs) for HTML landing pages that are unique for each taxon name and include the registration number. Index Fungorum, in addition, offers life science identifiers according to the Life Science Identifier (LSID) system, resolving LSIDs as URLs and UUIDs, in parallel. These efforts are steps forwards to create a stable framework of unique identifiers appropriate for linking data and related network information (Triebel & al., 2012; Hardisty & al., 2013).

The new kind of citation of taxon names in the running text as proposed here relies on already existing operations of the name repositories such as web services (see http://www.indexfungorum.org/ixfwebservice/fungus.asmx and http://www.mycobank.org/Services/Generic/Help.aspx?s=searchservice). It is not intended to provide a new class of identifiers or to compete with existing identifier systems. It is also not meant to change author citation in monographs and taxonomic revisions. Instead, the proposal to recommend the citation of functional combination of the scientific name with a stable taxon name identifier as an alternative to existing methods of citation is just meant to provide a compact mode for citation of names mainly in non-taxonomic papers.

In any case, common acceptance of this kind of citation will have several additional advantages. It will allow journal publishers, data publishers and biodiversity platforms for successfully (automated) parsing and tagging any published as well as still unpublished digital documents (such as submitted papers under review) with regard to cited scientific taxon name and taxon name identifier, enabling checking for consistency with the entries in the name registries. The new citation approach could also strengthen workflows between publishers and registration repositories as described for the common registration-to-publication automated pipeline envisaged by Penev & al. (2016).

Furthermore, the validated taxon name identifiers could be used by publishers (taking advantage from web services offered by the registration repositories and others) to infer digital object identifiers (DOIs), URLs or LSIDs for creating hypertext link

addresses. These link addresses could be made functional in digital formats of published articles such as PDF and HTML. "*Lecanora* varia [#389546]" (as seen in the analogue or printed version) could include one or more link addresses, URLs, in the digital form. In the case of Index Fungorum and MycoBank it would be very easy for publishers to infer the corresponding URLs from the taxon name identifier of a given taxon because the existing URLs include a standard prefix text followed by the taxon name identifier.

The concept of attaching taxon name identifiers to scientific names will also be a step forwards to develop and enhance mechanisms for semi-automatically providing credits to name authors. The stable taxon name identifiers in a digital document combined with the scientific name could be easily parsed and enriched with standardised information on name authors delivered by web services of the registration repositories. This parsing of codes of names might jointly be achieved by publishers and external service providers like Google Scholar, ResearchGate and ORCID. In consequence, taxon name author-related credit criteria could be created based on the citation frequency of taxon names.

Finally, the proposed suggestions for amendment of *ICN* Art. 46 – if accepted and accordingly applied – will support the approaches of linked open data (LOD) activities in general and will facilitate the setup of a semantic ontology network for scientific information retrieval. (For similar approaches for actionable object identifiers see Güntsch & al., 2017.)

■ ACKNOWLEDGEMENT

We thank Dr. Nicholas Turland (Botanischer Garten und Botanisches Museum, Berlin) for discussion of an earlier version of this paper and two anonymous reviewers for very useful comments and improvements. Work on best practice for using identifiers in science has been supported by the LIS infrastructure programme of the German Research Foundation (DFG) with the projects MOD-CO (RA 731/16-1, TR 290/8-1) and GFBio (TR 290/7-2).

■ LITERATURE CITED

Barkworth, M.E., Watson, M., Barrie, F.R., Belyaeva, I.V., Chung, R.C.K., Dašková, J., Davidse, G., Dönmez, A.A., Doweld, A.B., Dressler, S., Flann, C., Gandhi, K., Geltman, D., Glen, H.F., Greuter, W., Head, M.J., Jahn, R., Janarthanam, M.K., Katinas, L., Kirk, P.M., Klazenga, N., Kusber, W.-H., Kvaček, J., Malécot, V., Mann, D.G., Marhold, K., Nagamasu, H., Nicolson, N., Paton, A., Patterson, D.J., Price, M.J., Prud'homme van Reine, W.F., Schneider, C.W., Sennikov, A., Smith, G.F., Stevens, P.F., Yang, Z.-L., Zhang, X.-C. & Zuccarello, G.C. 2016. Report of the Special Committee on Registration of Algal and Plant Names (Including Fossils). *Taxon* 65: 670–672. https://doi.org/10.12705/653.43

Brummitt, R.K. & Powell, C.E. 1992. *Authors of plant names: A list of authors of scientific names of plants, with recommended standard forms of their names, including abbreviations.* [London]: Royal Botanic Gardens, Kew.

Güntsch, A., Hyam, R., Hagedorn, G., Chagnoux, S., Röpert, D., Casino A., Droege, G., Glöckler, F., Gödderz, K., Groom, Q., Hoffmann, J., Holleman, A., Kempa, M., Koivula, H., Marhold, K., Nicolson, N., Smith, V.S. & Triebel, D. 2017. Actionable, long-term stable and semantic web compatible identifiers for access to biological collection objects. *Database* 2017: bax003. https://doi.org/10.1093/database/bax003

Hardisty, A., Roberts, D. & The Biodiversity Informatics Community 2013. A decadal view of biodiversity informatics: Challenges and priorities. *B. M. C. Ecol.* 13: 16. https://doi.org/10.1186/1472-6785-13-16

May, T.W., De Beer, W.Z., Crous, P.W., Hawksworth, D.L., Liu, X.Z., Norvell, L.L., Pennycook, S.R., Redhead, S.A. & Seifert, K.A. 2016. (362–363) Proposals to amend the *Code* to modify its governance with respect to names of organisms treated as fungi. *Taxon* 65: 918–920. https://doi.org/10.12705/654.47

McNeill, J., Barrie, F.R., Buck, W.R., Demoulin, V., Greuter, W., Hawksworth, D.L., Herendeen, P.S., Knapp, S., Marhold, K., Prado, J., Prud'homme van Reine, W.F., Smith, G.F., Wiersema, J.H. & Turland, N. (eds. & comps.) 2012. *International Code of Nomenclature for algae, fungi, and plants (Melbourne Code): Adopted by the Eighteenth International Botanical Congress Melbourne, Australia, July 2011*. Königstein: Koeltz Scientific Books.

Penev, L., Paton, A., Nicolson, N., Kirk, P., Pyle, R.L., Whitton, R., Georgiev, T.A., Barker, C., Hopkins, C., Robert, V., Bisserkov, J. & Stoev, P. 2016. A common registration-to-publication automated pipeline for nomenclatural acts for higher plants (International Plant Names Index, IPNI), fungi (Index Fungorum, MycoBank) and animals (ZooBank). *ZooKeys* 550: 233–246. https://doi.org/10.3897/zookeys.550.9551

Redhead, S.A. & Norvell, L.L. 2012. MycoBank, Index Fungorum, and Fungal Names recommended as official nomenclatural repositories. *IMA Fungus* 3: 44. http://www.imafungus.org/Issue/32/03.pdf

Redhead, S.A. & Norvell, L.L. 2013. Report of the Nomenclature Committee for Fungi 19: Official repositories for fungal names. *Taxon* 62: 173–174. http://www.ingentaconnect.com/contentone/iapt/tax/2013/00000062/00000001/art00014

Redhead, S.A., Demoulin, V., Hawksworth, D.L., Seifert, K.A. & Turland, N.J. 2014. Fungal Nomenclature at IMC10: Report of the Nomenclature Sessions. *IMA Fungus* 5: 449–462. https://doi.org/10.5598/imafungus.2014.05.02.09

Triebel, D. & Bensch, K. 2005–. LIAS names – A database with names of lichens, lichenicolous fungi and non-lichenized ascomycetes. http://liasnames.lias.net/

Triebel, D., Hagedorn, G. & Rambold, G. 2012. An appraisal of mega-science platforms for biodiversity information. *MycoKeys* 5: 45–63. https://doi.org/10.3897/mycokeys.5.4302

ICN registration sites/registration repositories

Fungal Names registration http://fungalinfo.im.ac.cn/fungalname/fungalname.html
Index Fungorum http://www.indexfungorum.org/
MycoBank http://www.mycobank.org/