Synthetic and Systems Biotechnology xxx (2017) 1–10

Contents lists available at ScienceDirect

Synthetic and Systems Biotechnology

journal homepage: http://www.keaipublishing.com/en/journals/synthetic-and-systems-biotechnology/

On monospecific genera in prokaryotic taxonomy

Guanghong Zuo, Bailin Hao*

T-Life Research Center and Department of Physics, Fudan University, Shanghai 200433, China

ARTICLE INFO

Article history:
Received 31 July 2017
Received in revised form
28 August 2017
Accepted 28 August 2017

Keywords:
Composition Vector Tree
16S rRNA phylogeny
Prokaryotic taxonomy
Monospecific genera
Monophyly

ABSTRACT

A monospecific genus contains a single species ever since it was proposed. Though formally more than half of the known prokaryotic genera are monospecific, we pick up those which actually raise taxonomic problems by violating monophyly of the taxon within which it resides. Taking monophyly as a guiding principle, our arguments are based on simultaneous support from 16S rRNA sequence analysis and whole-genome phylogeny of prokaryotes, as provided by the LVTree Viewer and CVTree Web Server, respectively. The main purpose of this study consists in calling attention to this specific way of global taxonomic analysis. Therefore, we refrain from making formal emendations for the time being.
© 2017 The Authors. Production and hosting by Elsevier B.V. on behalf of KeAi Communications Co. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Microbiology has entered the era of Big Data. The number of known small-subunit ribosomal RNA (SSU rRNA) sequences reaches several millions, as reflected by the ARB-SILVA [1] and RDP [2] projects. The number of sequenced prokaryotic genomes now exceeds 120,000 and grows daily, see GOLD [3] for a timely summary. As of early January 2017, there were 2552 Archaea and Bacteria genus names and 14,621 species names (without counting subspecies) collected in the List of Prokaryotes with Standing in Nomenclature (LPSN) [4]. These numbers provide a kind of upper bounds as they contain some validly published but defunct names.

Traditionally, prokaryotic taxonomy has been centered around designated species or taxa. Now the availability of these Big Data adds a global angle to look at the issue. It is a common practice that most new genera when first proposed consist of a single or two species. With the advance of microbiology in exploring various ecological niches the monospecific status of many genera would change. However, some genera may remain monospecific over many years. Fig. 1 shows the number of monospecific genera listed in LPSN (July 2017) versus the year of discovery. The leftmost circle in Fig. 1 represents a species *Beggiatoa alba*; it was first described by J. P. Vaucher in 1803 and assigned to genus *Beggiatoa* by V. Trevisan in 1845. The sharp rising part of the curve in Fig. 1 would keep moving rightwards in forthcoming years and special attention should be put to the flat left part of the curve, i.e., those genera which remain monospecific for a long period of time.

In the present work, we examine some of the monospecific genera in prokaryotic taxonomy which may eventually necessitate taxonomic revisions. A search of LPSN reveals that 1320 genera, i.e. more than half of the total 2552, have only a single species listed. In other words, they seem to be monospecific genera. However, as the concept of prokaryotic species has been subject to long debate, we shall not touch on the definitive aspect of a taxon being monospecific or not. To this end an extreme example is provided by the genus *Brucella* which has been described manifestly as a monospecific genus in a "validly-published" way [5]. Whereas the recognition of this taxon being a monospecific or a multi-specific genus has finally left to individual microbiologists as preferring one or another taxonomic opinion (see related notes in LPSN [4] for details). Instead, we pick up those cases where a seemingly monospecific genus violates the monophyly of a broader genus and thus creates undoubtedly a taxonomic problem.

2. Materials and methods

Methodologically, we start from two distinct and independent kinds of phylogenetic trees, namely, the All-Species Living Tree (LVTree) based on alignment of high-quality 16S rRNA sequences [6–8] and the Composition Vector Tree (CVTree) based on

Abbreviations: LVTree, All-Species Living Tree; CVTree, Composition Vector Tree; LPSN, List of Prokaryotes with Standing in Nomenclature.

* Corresponding author.
E-mail address: hao@mail.itp.ac.cn (B. Hao).

Peer review under responsibility of KeAi Communications Co., Ltd.

http://dx.doi.org/10.1016/j.synbio.2017.08.004
2405-805X/© 2017 The Authors. Production and hosting by Elsevier B.V. on behalf of KeAi Communications Co. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Please cite this article in press as: Zuo G, Hao B, On monospecific genera in prokaryotic taxonomy, Synthetic and Systems Biotechnology (2017), http://dx.doi.org/10.1016/j.synbio.2017.08.004

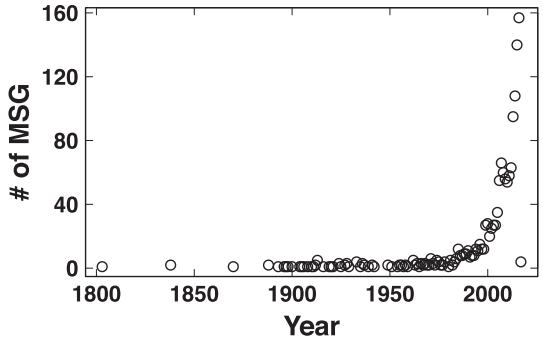

Fig. 1. Number of monospecific genera versus year of discovery according to LPSN (July 2017). The leftmost circle represents *Beggiatoa alba* discovered in 1803, which, however, no longer keeps monospecific status as a new species has been validly published quite recently.

alignment-free comparison of whole prokaryotic genomes [9–12].

CVTrees are constructed by using the latest release of the CVTree Web Server [13]. Instead of the 3000 plus *Bacteria* and *Archaea* genomes provided in the built-in dataset of the publically available CVTree Web Server, we use 8000 to 10,000 genomes with as wide as possible taxonomic distribution as a background in any study of a designated group of species. In this way the overall results appear to be extremely stable and reproducible. As the algorithm behind the CVTree method and the features of the Web Server have been described repeatedly in literature over the years, we refer the readers to the published papers [9–12] and to the online help file that comes with the CVTree3 Web Server [13].

The latest release of LVTree (as of February 2017) is built on 475 *Archaea* and 12,478 *Bacteria* 16S rRNA sequences [14]. In order to compare the branching order of LVTree at all taxonomic ranks from phyla down to species, we have transplanted most of the distinctive features of the CVTree Web Server to a LVTree Viewer [15]. Among the important and convenient features we indicate the possibility of making trial lineage modifications and automatic reporting monophyly/none-monophyly summary of taxa at all ranks of the taxonomic hierarchy. It is desirable to emphasize from the outset that CVTree and LVTree are independent in the input data and in the underlying methodology. There is no a priori reason that the two approaches should yield identical results. The fact that a cluster of leaves in one tree agrees topologically with that in another tree adds weight to the objectiveness of the results. In this study we rely on facts compatible in both kinds of trees. Since LVTree is built with emphasis on type strains while CVTree is based on available sequenced genomes, one cannot expect that one and the same taxon name in both trees corresponds to an identical organism. Nonetheless, this tolerance is understood in taxonomy as classification scheme always concerns populations designated to a taxon, not restricting to individual organisms.

In making comparison of a cluster of leaves in a tree with a taxon in a classification scheme a guiding principle is the notion of monophyly. Coined by Ernst Haeckel in 1866 and stemmed from zoology, the original definition of monophyly requires the knowledge on an ancestor and all its descendants, see, e.g., discussion of James Farris over the years [16,17]. Obviously, this requirement does not apply to prokaryotes which are overwhelmingly reproduced asexually. Therefore, we adopt a pragmatic point of view on monopoly by restricting ourselves to the input dataset with a certain reference taxonomy. If in the reference taxonomy species designated to a taxon appear only in that taxon and do not present in any other taxa, then the taxon is said to be monophyletic. If in a tree branch all leaves come exclusively from a reference taxon and no species from that taxon fall in any other branch, the branch is said to be monophyletic. In other words, monophyly is a reciprocal notion with respect to both phylogeny and taxonomy. Monophyly of a tree branch or a taxon may change when new species appear in the input dataset. Only monophyletic taxa are considered to be well-defined and acceptable in a flawless taxonomy.

We have mentioned reference taxonomy. For the time being we have in mind the combined use of the following resources:

1. The second edition of Bergey's Manual of Systematic Bacteriology [18] which had been completed in the years 2001–2012. The Bergey's Manual Trust made it clear that further editions of the Manual would be electronic. In fact, the framework of the electronic edition of Bergey's Manual of Systematics of Archaea and Bacteria, abbreviated as BMSAB, appeared in 2015 [19]. For the time being it is only a framework for an electronic edition, as the content of many chapters appears to be identical to their counterparts in Ref. [18] without essential updating.
2. The 4th edition of the multi-volume treatise Prokaryotes [20], especially, the 6 vol on prokaryotic taxonomy published in 2014. These volumes are organized around families and some entries are more updated than the Bergey's Manuals.
3. List of Prokaryote Names with Standing in Nomenclature [4]. This constantly updating list collects names and taxonomic proposals validly published in International Journal of Systematic and Evolutionary Microbiology and a few other periodicals. In using LPSN one should be aware of possible redundancy of the list, as some defunct or erroneous names may be kept until new emendations are officially published.
   Please note that all the above three resources are mainly based on 16S rRNA sequence analysis and thus naturally inherit the limitations of the latter, especially, as compared with the whole-genome-based approaches.
4. Besides the above resources we also make use of the EzBioCloud database maintained by the Chun Lab in Korea [21]. Though still based on 16S rRNA sequence analyses, EzBioCloud pay more attention to genome data. In the latest release of 11 May 2017 [22] it contains 82,605 quality-controlled genome sequences and 62,685 16S rRNA sequences. More importantly, the taxonomic assignments in EzBioCloud are not restricted by published information. In general, these assignments better reflect the actual positions in phylogenetic trees and thus, in most cases, happen to be closer to the CVTree results. This point will be demonstrated by our first example below on monospecific genera within the genus *Pseudomonas*.

3. Monospecific genera that call for taxonomic modifications

3.1. Monospecific genera related to Pseudomonas

It is instructive to commence with *Pseudomonas*. Historically, *Pseudomonas* has caused many taxonomic confusions. The species collected under the genus *Pseudomonas* in the 1st edition of the Bergey's Manual of Systematic Bacteriology [23] were subdivided into more than 10 genera in the following years [24]. Not long ago it was stated in a footnote of a well-known book [25] that members of the *Pseudomonas* genus might be assigned to several different classes (*Alpha-*, *Beta-* or *Gamma-*) within the phylum *Proteobacteria.* However, nowadays the majority of *Pseudomonas* species behaves well by forming monophyletic branches in both LVTree and CVTree mainly due to a number of well-founded taxonomic revisions, among which we make emphasis on those dealing with monospecific genera.

First of all, a “new” genus *Serpens* was suggested in 1977 with a single species *Serpens flexibilis* [26] described. In later years it has

Please cite this article in press as: Zuo G, Hao B, On monospecific genera in prokaryotic taxonomy, Synthetic and Systems Biotechnology (2017), http://dx.doi.org/10.1016/j.synbio.2017.08.004

Download English Version:

https://daneshyari.com/en/article/8426041

Download Persian Version:

https://daneshyari.com/article/8426041

Daneshyari.com