Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 July 2019 doi:10.20944/preprints201907.0222.v1

Pathogen taxonomy updates at the Comprehensive Antibiotic Resistance Database: Implications for molecular epidemiology

Kara K. Tsang, David J. Speicher, and Andrew G. McArthur*

M.G. DeGroote Institute for Infectious Disease Research, Department of Biochemistry and Biomedical Sciences, DeGroote School of Medicine, McMaster University, Hamilton, Ontario L8S 4K1, Canada

*To whom correspondence should be addressed.

Contact: mcarthua@mcmaster.ca

Abstract

With the increasing use of genome sequencing as a surveillance tool for molecular epidemiology of antimicrobial resistance, we are seeing an increased intersection of genomics, microbiology, and clinical epidemiology. Clear nomenclature for AMR gene families and pathogens is critical for communication. For CARD release version 3.0.3 (July 2019), we updated the entire CARD database to reflect the latest pathogen names. In total, we detected 48 name changes or updates, some of which reflect major changes in familiar names.

Keywords: antimicrobial resistance, biocuration, microbial nomenclature, molecular epidemiology

Introduction

With the increasing use of genome sequencing as a surveillance tool for molecular epidemiology of antimicrobial resistance (AMR) (1, 2), we are seeing an increased intersection of genomics, microbiology, and clinical epidemiology. As such, clear nomenclature for AMR gene families and pathogens is critical for communication. At the Comprehensive Antibiotic Resistance Database (CARD) (3), we provide highly curated reference data on the molecular basis of AMR. As a rule, sequences added to CARD must include peer-reviewed, published experimental evidence of antibiotic minimum inhibitory concentration (MIC) elevated above that of a control, plus availability of the DNA sequence of the corresponding gene or mutant in GenBank (4). All data is organized ontologically, most notably using the Antibiotic Resistance Ontology and an ontological mirror of GenBank's Taxonomy Database (3). Yet, since its inception CARD has not performed an update of the microbiological taxonomy and nomenclature underlying its organization of AMR sequences, predicted allelic variants, resistome annotations, and epidemiological estimates. As such, for CARD release version 3.0.3 (July 2019) we added a new software module to our CARD Quality Control tools to detect nomenclatural changes in GenBank not incorporated into CARD, which we then used to update the CARD database to reflect the latest pathogen names. In total, we detected 48 name changes or updates, some of which reflect major changes in familiar names (Table 1). These changes are outlined below with citations for convenience and clarity, using the acronym *p.k.a.* for "previously known as".

Clostridioides difficile (p.k.a. *Clostridium difficile*)

CARD recently (CARD version 3.0.3, July 2019) performed a literature review of genes and mutations conferring AMR to the Gram-positive spore-forming bacterium *C. difficile*, a causative agent of diarrheal disease and colitis via colonization of the human colon. Familiar to clinicians as *Clostridium difficile*, Yutin & Galperin (5) proposed a name change in 2013 to *Peptoclostridium difficile* based on genomic and evolutionary analyses, which proved to be unpopular due to loss of the familiar '*C. diff*' and 'CDAD' (*Clostridium difficile*-associated diarrhea) monikers. Subsequent phylogenetic analysis by Lawson *et al.* (6) confirmed the genus *Clostridium* had been poorly defined, that *C. difficile* was indeed more closely related to *Peptoclostridium* than other *Clostridium*, and supported the new genus *Clostridioides* (meaning 'Clostridia-like'). This genus contains *Clostridium mangenotii* and *Clostridium difficile* as *Clostridioides mangenotii* and *Clostridioides difficile*, respectively. This new name conveniently maintains the '*C. diff*' moniker.

Cutibacterium acnes (p.k.a. *Propionibacterium acnes*)

Propionibacterium acnes is a Gram-positive bacterium involved in the pathogenesis of acne. Originally known *Bacillus acnes*, the species was renamed *Cutibacterium acnes* by Scholz & Kilian (7) based on genome and phylogenetic analysis. In CARD, *Cutibacterium acnes* has been associated with *gyrA* and 16S rRNA mutations conferring resistance to fluoroquinolones and tetracycline, respectively.

Kitasatospora aureofaciens (p.k.a. *Streptomyces aureofaciens*)

Streptomyces aureofaciens is widely known as a source of many tetracycline antibiotics (8) and in many ways represents the archetype for antibiotic producing *Streptomyces*. However, multi-gene phylogenetic analyses (9) have illustrated that this species is a member of the genus *Kitasatospora*, hence a name change to *Kitasatospora aureofaciens* has been adopted.

Borreliella burgdorferi (p.k.a. *Borrelia burgdorferi*)

The spirochete bacteria *Borrelia burgdorferi* is a tick-borne causative agent of Lyme disease and thus a focus of considerable scrutiny for vaccine and drug development, as well as molecular epidemiology (10). Yet, phylogenomic investigation has split the genus into *Borrelia* and *Borreliella* (11), resulting in the new and validated name *Borreliella burgdorferi*.

 doi:10.20944/preprints201907.0222.v1

Klebsiella aerogenes (p.k.a. *Enterobacter aerogenes*)

The familiar ESKAPE pathogen *Enterobacter aerogenes* has a long nomenclatural history. This Gram-negative, opportunistic pathogen was originally included as *Aerobacter aerogenes* and later renamed *Enterobacter aerogenes* based on nomenclatural rules, but as early as 1971 it was realized that not all strains of *Enterobacter* were likely to be of the same genus and that *Enterobacter aerogenes* was more likely a member of *Klebsiella* **(12, 13)**. Nomenclatural rules at the time blocked this name change, but recognition of overlapping types (NCTC 10006, ATCC 13048) resulted in the short-lived moniker *Klebsiella mobilis* (i.e. *Klebsiella mobilis* and *Enterobacter aerogenes* are homotypic synonyms). Tindall *et al.* **(13)** reviewed this naming history and current nomenclatural rules, resulting in the definitive and widely adopted change to *Klebsiella aerogenes*. It is worth noting that this taxonomic revision complicates the familiar ESKAPE pathogens acronym for prevalent, AMR hospital acquired infections: *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter* species ("*Klebsiella pneumoniae"* could be replaced with *"Klebsiella* species"). From a phenotypic point of view, care should be taken for assumptions of antimicrobial susceptibilities, since *Klebsiella aerogenes* has a different AMR gene compliment than other *Klebsiella*.

Mycobacterium, Mycobacteroides, & *Mycolicibacterium*

Until recently, the diversity of species related to *Mycobacterium tuberculosis* have been placed in the same genus, e.g. *Mycobacterium abscessus*, *Mycobacterium chelonae*, *Mycobacterium fortuitum*, *Mycobacterium goodii*, *Mycobacterium smegmatis*, *Mycobacterium bovis*. Yet, recent taxonomic revisions based on comparative genomics and evolutionary studies have distributed these species among three genera with separate evolutionary histories (4, 5): *Mycobacterium tuberculosis* & *Mycobacterium tuberculosis* variant *bovis*; *Mycobacteroides abscessus* & *Mycobacteroides chelonae*; *Mycolicibacterium fortuitum*, *Mycolicibacterium goodii*, & *Mycolicibacterium smegmatis*. Other species, not yet included in CARD, have been placed in a fourth & fifth new genus, *Mycolicibacter* and *Mycolicibacillus* (14).

Conclusions

A large number of databases relating to AMR exist (see CARD's amr_curation repo and Google spreadsheet list of databases at https://github.com/arpcard/amr_curation), yet most curate a molecular sequence once, reflecting its initial report, as is done at CARD. This initial report and entry in GenBank often reflect a gene's original detection in a single pathogen, yet AMR is an evolutionary process involving both evolution of the AMR sequences themselves over time and their horizontal gene transfer among pathogens. As such, AMR databases must do more than form a collection of reference sequences, but over time must track their evolution and epidemiological distribution. To do this well, AMR databases must update molecular sequences as well as our understanding of pathogen relationships. As AMR is an evolutionary phenomena, our analyses of AMR must also have an accurate evolutionary context which is most often reflected in the names of microbes. As illustrated above, bacterial nomenclature is increasingly guided by comparative genomic, phylogenetic, and phylogenomic analyses. Ignoring these nomenclatural changes places our analyses and interpretation of AMR into incorrect context. AMR databases need to update microbial nomenclature to remain relevant. Yet, in clinical settings nomenclatural changes in academia often can appear arbitrary, as familiar monikers are important for clinical diagnostics, public health decisions, and clear lines of communication. As outlined in Dr. Audrey Schuetz's online article (15), changes in nomenclature can disrupt clinical processes and have unanticipated costs. As such, and following GenBank's lead, CARD now strives to follow the latest changes in bacterial nomenclature and as it is based on both the Chado data schema and organized exclusively using ontologies (3), is able to rapidly update names and descriptions within CARD to reflect the latest nomenclature while simultaneously supporting depreciated names as ontological synonyms, allowing users to search CARD using familiar monikers in case they are unfamiliar with recent name changes, allowing them to find the data they need.

Acknowledgements

This research was funded by the Canadian Institutes of Health Research (PJT-156214) to A.G.M., who also holds a Cisco Research Chair in Bioinformatics, supported by Cisco Systems Canada, Inc. K.K.T. was supported by an Ontario Graduate Scholarship and McMaster University's MacDATA Institute Graduate Fellowship, Michael G. DeGroote Institute for Infectious Disease Research Michael Kamin Hart Memorial Scholarship, Faculty of Health Sciences Graduate Programs Excellence Award, and Biochemistry & Biomedical Sciences Fred and Helen Knight Enrichment Award. D.J.S. was supported by the Michael G. DeGroote Initiative for Innovation in Healthcare. Computer resources were supplied by the McMaster Service Lab and Repository computing cluster, funded in part by grants to A.G.M. from the Canadian Foundation for Innovation (34531).

References

1. **McArthur AG**, **Tsang KK**. 2017. Antimicrobial resistance surveillance in the genomic age. *Ann N Y Acad Sci* **1388**:78–91.

2. **McArthur AG**, **Wright GD**. 2015. Bioinformatics of antimicrobial resistance in the age of molecular epidemiology. *Curr Opin Microbiol* **27**:45–50.

3. **Jia B**, **Raphenya AR**, **Alcock B**, **Waglechner N**, **Guo P**, **Tsang KK**, **Lago BA**, **Dave BM**, **Pereira S**, **Sharma AN**, **Doshi S**, **Courtot M**, **Lo R**, **Williams LE**, **Frye JG**, **Elsayegh T**, **Sardar D**, **Westman EL**, **Pawlowski AC**, **Johnson TA**, **Brinkman FSL**, **Wright GD**, **McArthur AG**. 2017. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Res* **45**:D566–D573.

4. **Benson DA**, **Cavanaugh M**, **Clark K**, **Karsch-Mizrachi I**, **Lipman DJ**, **Ostell J**, **Sayers EW**. 2013. GenBank. *Nucleic Acids Res* **41**:D36–42.

5. **Yutin N**, **Galperin MY**. 2013. A genomic update on clostridial phylogeny: gram-negative spore formers and other misplaced clostridia. *Environ Microbiol* **15**:2631–2641.

6. **Lawson PA**, **Citron DM**, **Tyrrell KL**, **Finegold SM**. 2016. Reclassification of *Clostridium difficile* as *Clostridioides difficile* (Hall and O'Toole 1935) Prévot 1938. *Anaerobe* **40**:95–99.

7. **Scholz CFP**, **Kilian M**. 2016. The natural history of cutaneous propionibacteria, and reclassification of selected species within the genus *Propionibacterium* to the proposed novel genera *Acidipropionibacterium* gen. nov., *Cutibacterium* gen. nov. and *Pseudopropionibacterium* gen. nov. *Int J Syst Evol Microbiol* **66**:4422–4432.

8. **DARKEN MA**, **BERENSON H**, **SHIRK RJ**, **SJOLANDER NO**. 1960. Production of tetracycline by *Streptomyces aureofaciens* in synthetic media. *Appl Microbiol* **8**:46–51.

9. **Labeda DP**, **Dunlap CA**, **Rong X**, **Huang Y**, **Doroghazi JR**, **Ju K-S**, **Metcalf WW**. 2017. Phylogenetic relationships in the family Streptomycetaceae using multi-locus sequence analysis. *Antonie Van Leeuwenhoek*, 2nd ed. **110**:563–583.

10. **Tilly K**, **Rosa PA**, **Stewart PE**. 2008. Biology of infection with *Borrelia burgdorferi*. *Infect Dis Clin North Am* **22**:217–34– v.

11. **Adeolu M**, **Gupta RS**. 2014. A phylogenomic and molecular marker based proposal for the division of the genus *Borrelia* into two genera: the emended genus *Borrelia* containing only the members of the relapsing fever *Borrelia*, and the genus *Borreliella* gen. nov. containing the members of the Lyme disease *Borrelia* (*Borrelia burgdorferi* sensu lato complex). *Antonie Van Leeuwenhoek*, 2nd ed. **105**:1049–1072.

12. **Bascomb S**, **Lapage SP**, **Willcox WR**, **Curtis MA**. 1971. Numerical classification of the tribe Klebsielleae. *J Gen Microbiol* **66**:279–295.

13. **Tindall BJ**, **Sutton G**, **Garrity GM**. 2017. *Enterobacter aerogenes* Hormaeche and Edwards 1960 (Approved Lists 1980) and *Klebsiella mobilis* Bascomb et al. 1971 (Approved Lists 1980) share the same nomenclatural type (ATCC 13048) on the Approved Lists and are homotypic synonyms, with consequences for the name *Klebsiella mobilis* Bascomb et al. 1971 (Approved Lists 1980). *Int J Syst Evol Microbiol* **67**:502–504.

14. **Gupta RS**, **Lo B**, **Son J**. 2018. Phylogenomics and comparative genomic studies robustly support division of the genus *Mycobacterium* into an emended genus *Mycobacterium* and four novel genera. *Front Microbiol* **9**:67.

15. **Schuetz, A.N.** 2017. Celebrating successes and contemplating messes in bacterial taxonomy. https://www.asm.org/Articles/2017/September/Clinical-and-Public-Health-Microbiology-(7)

 doi:10.20944/preprints201907.0222.v1

Table 1. Taxonomic updates at the Comprehensive Antibiotic Resistance Database, with associated NCBI Taxonomy ID. Bolded taxa are discussed in the text.

Old Name	New Name	NCBI
Acinetobacter baumannii ATCC 19606	*Acinetobacter baumannii* ATCC 19606 = CIP 70.34 = JCM 6841	575584
Acinetobacter sp. ADP1	*Acinetobacter baylyi* ADP1	62977
Acinetobacter genomosp. 14	*Acinetobacter colistiniresistens*	70345
Acinetobacter genomosp. 17BJ	*Acinetobacter dispersus*	70348
Acinetobacter genomosp. 15	*Acinetobacter variabilis*	70346
Streptomyces caeruleus	*Actinoalloteichus cyanogriseus*	195949
Agrobacterium tumefaciens str. C58	*Agrobacterium fabrum* str. C58	176299
Bacillus amyloliquefaciens subsp. *plantarum* str. FZB42	*Bacillus velezensis* FZB42	326423
Borrelia burgdorferi	***Borreliella burgdorferi***	**139**
Borrelia burgdorferi B31	*Borreliella burgdorferi* B31	224326
Campylobacter jejuni subsp. *jejuni* NCTC 11168	*Campylobacter jejuni* subsp. *jejuni* NCTC 11168 = ATCC 700819	192222
Epilithonimonas sp. Stok-2	*Chryseobacterium* sp. Stok-2	1620219
Clostridium difficile	***Clostridioides difficile***	**1496**
Clostridium difficile 630	*Clostridioides difficile* 630	272563
Propionibacterium acnes	***Cutibacterium acnes***	**1747**
Enterococcus faecalis HH22	*Enterococcus faecalis* EnGen0297	491075
Streptomyces aureofaciens	***Kitasatospora aureofaciens***	**1894**
Enterobacter aerogenes	***Klebsiella aerogenes***	**548**
Mycobacterium bovis	***Mycobacterium tuberculosis* variant *bovis***	**1765**
Mycobacterium bovis AF2122/97	*Mycobacterium tuberculosis* variant *bovis* AF2122/97	233413
Mycobacterium bovis BCG str. Mexico	*Mycobacterium tuberculosis* variant *bovis* BCG str. Mexico	717522
Mycobacterium bovis BCG str. Pasteur 1173P2	*Mycobacterium tuberculosis* variant *bovis* BCG str. Pasteur 1173P2	410289
Mycobacterium abscessus	***Mycobacteroides abscessus***	**36809**
Mycobacterium chelonae	***Mycobacteroides chelonae***	**1774**
Mycobacterium fortuitum	***Mycolicibacterium fortuitum***	**1766**
Mycobacterium goodii	***Mycolicibacterium goodii***	**134601**
Mycobacterium smegmatis	***Mycolicibacterium smegmatis***	**1772**
Mycobacterium smegmatis str. MC2 155	*Mycolicibacterium smegmatis* MC2 155	246196
Rhodococcus equi	*Rhodococcus hoagii*	43767
Salmonella enterica subsp. *enterica* serovar Arizona	*Salmonella enterica* subsp. *arizonae*	486989
Salmonella enterica subsp. *arizonae* str. ATCC BAA-1577	*Salmonella enterica* subsp. *arizonae* serovar 41:z4,z23:- str. ATCC BAA-1577	523833
Salmonella enterica subsp. *arizonae* serovar 62:z4,z23:--	*Salmonella enterica* subsp. *arizonae* serovar 62:z4,z23:-	41514
Salmonella enterica IIIb 50:k:z	*Salmonella enterica* subsp. *diarizonae* serovar 50:k:z	41512
Salmonella enterica subsp. *enterica* serovar 4,12;I,-	*Salmonella enterica* subsp. *enterica* serovar 4,12:i:-	353569
Salmonella enterica subsp. *enterica* serovar Amesterdam	*Salmonella enterica* subsp. *enterica* serovar Amsterdam	593904
Salmonella enterica subsp. *enterica* serovar Typhimurium var. Copenhagen	*Salmonella enterica* subsp. *enterica* serovar Copenhagen	353544
Salmonella enterica subsp. *enterica* serovar Paratyphi C strain RKS4594	*Salmonella enterica* subsp. *enterica* serovar Paratyphi C str. RKS4594	476213
Salmonella typhimurium DT104	*Salmonella enterica* subsp. *enterica* serovar Typhimurium str. DT104	85569
Salmonella typhimurium SL1344	*Salmonella enterica* subsp. *enterica* serovar Typhimurium str. SL1344	216597
Salmonella enterica VI 1,6,14,25:a:e,n,x	*Salmonella enterica* subsp. *indica* serovar 1,6,14,25:a:e,n,x	59246
Salmonella enterica VII 1,40:g,z51:--	*Salmonella enterica* subsp. *VII* serovar 1,40:g,z51:--	41520
Salmonella enterica VII 40:z4,z24:--	*Salmonella enterica* subsp. *VII* serovar 40:z4,z24:--	41521
Enterobacteria phage Fels-2	*Salmonella virus* Fels2	194701
Staphylococcus aureus ED98	*Staphylococcus aureus* subsp. *aureus* ED98	665030
Staphylococcus aureus RN4220	*Staphylococcus aureus* subsp. *aureus* RN4220	561307
Streptomyces roseosporus NRRL 11379	*Streptomyces filamentosus* NRRL 11379	457430
Streptomyces roseochromogenes subsp. *oscitans*	*Streptomyces roseochromogenus* subsp. *oscitans*	149682
Streptomyces roseochromogenes subsp. *oscitans* DS 12.976	*Streptomyces roseochromogenus* subsp. *oscitans* DS 12.976	1352936