



eCOMMONS

Loyola University Chicago
Loyola eCommons

Bioinformatics Faculty Publications

Faculty Publications and Other Works by
Department

9-3-2020

Best Practices for Successfully Writing and Publishing a Genome Announcement in Microbiology Resource Announcements

Julie C. Dunning Hotopp
University of Maryland at Baltimore

David A. Baltrus
University of Arizona

Vincent M. Bruno
University of Maryland at Baltimore

John J. Dennehy
Queens College

Steven R. Gill
University of Rochester

Follow this and additional works at: https://ecommons.luc.edu/bioinformatics_facpub
See next page for additional authors

Recommended Citation

Dunning Hotopp, Julie C.; Baltrus, David A.; Bruno, Vincent M.; Dennehy, John J.; Gill, Steven R.; Maresca, Julia A.; Matthijnssens, Jelle; Newton, Irene L. G.; Putonti, Catherine; Rasko, David A.; Rokas, Antonis; Roux, Simon; Stajich, Jason E.; Stedman, Kenneth M.; Stewart, Frank J.; and Thrash, J. Cameron. Best Practices for Successfully Writing and Publishing a Genome Announcement in Microbiology Resource Announcements. *Microbiology Resource Announcements*, , : , 2020. Retrieved from Loyola eCommons, Bioinformatics Faculty Publications, <http://dx.doi.org/10.1128/MRA.00763-20>

This Article is brought to you for free and open access by the Faculty Publications and Other Works by Department at Loyola eCommons. It has been accepted for inclusion in Bioinformatics Faculty Publications by an authorized administrator of Loyola eCommons. For more information, please contact ecommons@luc.edu.



















This work is licensed under a [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/).
© 2020 Dunning Hotopp et al.

Authors

Julie C. Dunning Hotopp, David A. Baltrus, Vincent M. Bruno, John J. Dennehy, Steven R. Gill, Julia A. Maresca, Jelle Matthijnssens, Irene L. G. Newton, Catherine Putonti, David A. Rasko, Antonis Rokas, Simon Roux, Jason E. Stajich, Kenneth M. Stedman, Frank J. Stewart, and J. Cameron Thrash



Best Practices for Successfully Writing and Publishing a Genome Announcement in *Microbiology Resource Announcements*

 Julie C. Dunning Hotopp,^{a,b,c}  David A. Baltrus,^{d,e}  Vincent M. Bruno,^{a,b}  John J. Dennehy,^{f,g}  Steven R. Gill,^h
 Julia A. Maresca,ⁱ  Jelle Matthijssens,^j  Irene L. G. Newton,^k  Catherine Putonti,^{l,m,n}  David A. Rasko,^{a,b}
 Antonis Rokas,^o  Simon Roux,^p  Jason E. Stajich,^q  Kenneth M. Stedman,^{r,s}  Frank J. Stewart,^{t,u}  J. Cameron Thrash^v

^aInstitute for Genome Sciences, University of Maryland School of Medicine, Baltimore, Maryland, USA

^bDepartment of Microbiology and Immunology, University of Maryland School of Medicine, Baltimore, Maryland, USA

^cGreenebaum Cancer Center, University of Maryland School of Medicine, Baltimore, Maryland, USA

^dSchool of Plant Sciences, University of Arizona, Tucson, Arizona, USA

^eSchool of Animal and Comparative Biomedical Sciences, University of Arizona, Tucson, Arizona, USA

^fBiology Department, Queens College, City University of New York, Flushing, New York, USA

^gThe Graduate Center, City University of New York, New York, New York, USA

^hDepartment of Microbiology and Immunology, University of Rochester School of Medicine and Dentistry, Rochester, New York, USA

ⁱDepartment of Civil and Environmental Engineering, University of Delaware, Newark, Delaware, USA

^jUniversity of Leuven, Department of Microbiology, Immunology, and Transplantation, Rega Institute, Leuven, Belgium

^kDepartment of Biology, Indiana University, Bloomington, Indiana, USA

^lBioinformatics Program, Loyola University Chicago, Chicago, Illinois, USA

^mDepartment of Biology, Loyola University Chicago, Chicago, Illinois, USA

ⁿDepartment of Microbiology and Immunology, Stritch School of Medicine, Loyola University Chicago, Maywood, Illinois, USA

^oDepartment of Biological Sciences, Vanderbilt University, Nashville, Tennessee, USA

^pDOE Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, California, USA

^qDepartment of Microbiology and Plant Pathology, University of California, Riverside, Riverside, California, USA

^rCenter for Life in Extreme Environments, Portland State University, Portland, Oregon, USA

^sBiology Department, Portland State University, Portland, Oregon, USA

^tDepartment of Microbiology and Immunology, Montana State University, Bozeman, Montana, USA

^uSchool of Biological Sciences, Georgia Institute of Technology, Atlanta, Georgia, USA

^vDepartment of Biological Sciences, University of Southern California, Los Angeles, California, USA

ABSTRACT *Microbiology Resource Announcements* (MRA) provides peer-reviewed announcements of scientific resources for the microbial research community. We describe the best practices for writing an announcement that ensures that these publications are truly useful resources. Adhering to these best practices can lead to successful publication without the need for extensive revisions.

M*icrobiology Resource Announcements* (MRA) provides peer-reviewed announcements of scientific resources for the microbial research community. Such resources include genomes, transcriptomes, amplicon sequence data sets, other sequence collections, culture collections, mutant libraries, and software (Table 1). The MRA editors have established minimum requirements needed to ensure that the laboratory and analytical methods can be replicated by any other research group, making these publications truly useful resources for the community. These minimum requirements are described in checklists for each resource type (https://mra.asm.org/sites/default/files/additional-assets/thumbs/MRA_Author_Checklist.pdf). Adhering to the relevant checklist streamlines the review process for both authors and reviewers. In this guide, we take the most common resource, the genome announcement, and provide

Citation Dunning Hotopp JC, Baltrus DA, Bruno VM, Dennehy JJ, Gill SR, Maresca JA, Matthijssens J, Newton ILG, Putonti C, Rasko DA, Rokas A, Roux S, Stajich JE, Stedman KM, Stewart FJ, Thrash JC. 2020. Best practices for successfully writing and publishing a genome announcement in *Microbiology Resource Announcements*. *Microbiol Resour Announc* 9:e00763-20. <https://doi.org/10.1128/MRA.00763-20>.

Copyright © 2020 Dunning Hotopp et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Julie C. Dunning Hotopp, jhotopp@som.umaryland.edu.

Published 3 September 2020

TABLE 1 Example published resource announcements in MRA

Resource announcement type ^a	Example(s)
Complete circular bacterial genome	Complete genome sequence of <i>Pseudomonas coronafaciens</i> pv. <i>oryzae</i> 1_6 (11); Complete genome sequence of <i>Luteibacter pinisoli</i> MAH-14 (12)
Draft bacterial genome with Illumina data only	Draft genome sequence of <i>Lactobacillus jensenii</i> strain UMB7766, isolated from the female bladder (13)
Bacterial genome with PacBio data only	Draft genome sequence of a fish pathogen, <i>Edwardsiella piscicida</i> isolate CK41 (14)
Bacterial genome with ONT data only	Complete genome sequence of halophilic deep-sea bacterium <i>Halomonas axialensis</i> strain Althf1 (15)
Obligate host-associate bacterial genome	Complete genome sequence of wAna, the <i>Wolbachia</i> endosymbiont of <i>Drosophila ananassae</i> (2)
Many genomes from a large bacterial culture collection	Eleven high-quality reference genome sequences and 360 draft assemblies of Shiga toxin-producing <i>Escherichia coli</i> isolates from human, food, animal, and environmental sources in Canada (16)
Phage double-stranded DNA genome	Complete genome sequence of <i>Escherichia coli</i> phage Paul (17); Complete genome sequence of <i>Escherichia coli</i> siphophage Snoke (18)
Linear viral RNA genome	Complete genome sequence of an American isolate of pepino mosaic virus (19)
Small circular viral single-stranded DNA genome	Genome sequences of three cruciviruses found in the Willamette Valley (Oregon) (20)
Draft microbial eukaryotic genome with Illumina data only	Draft genome sequence of the griseofulvin-producing fungus <i>Xylaria flabelliformis</i> strain G536 (21)
Draft microbial eukaryotic genome with PacBio data	Genome sequence of the extremely acidophilic fungus <i>Acidomyces richmondensis</i> FRIK2901 (22); Genome sequence of a California isolate of <i>Fusarium oxysporum</i> f. sp. <i>lycopersici</i> race 3, a fungus causing wilt disease on tomato (23)
Complete microbial eukaryotic genome	Complete genome sequences for two <i>Talaromyces marneffeii</i> clinical isolates from northern and southern Vietnam (24)
Nearly complete microbial eukaryotic genome	Nearly complete genome sequence of <i>Brugia malayi</i> strain FR3 (25)
16S amplicon study	16S rRNA amplicon sequencing of sediment bacterial communities in an oyster farm in Rhode Island (26)
Transcriptome	Multispecies transcriptomics data set of <i>Brugia malayi</i> , its <i>Wolbachia</i> endosymbiont wBm, and <i>Aedes aegypti</i> across the <i>B. malayi</i> life cycle (27)
Culture collections/mutant libraries	Gateway entry vector library of <i>Wolbachia pipientis</i> candidate effectors from strain wMel (28)
Metagenome	Microbiota of the Hickey Run Tributary of the Anacostia River (29); Metagenomic assembly and prokaryotic metagenome-assembled genome sequences from the northern Gulf of Mexico "dead zone" (30)
Metatranscriptome	Metatranscriptomic sequencing of a cyanobacterial soil-surface consortium with and without a diverse underlying soil microbiome (31)
Software	TwinBLAST: when two is better than one (32)
Organelle genome	Complete chloroplast genome sequence of a white spruce (<i>Picea glauca</i> , genotype WS77111) from eastern Canada (33); Mitochondrial genome sequences of <i>Diorhabda carinata</i> and <i>Diorhabda carinulata</i> , two beetle species introduced to North America for biological control (34)
Proteome	Proteome of a <i>Moraxella catarrhalis</i> strain under iron-restricted conditions (35)

^a PacBio, Pacific Biosciences; ONT, Oxford Nanopore Technologies.

detailed recommendations for the three main sections of a genome announcement. Use of these best practices can lead to successful publication without the need for extensive revisions.

INTRODUCTION AND RATIONALE FOR SEQUENCING

The first section of a genome announcement should provide a brief introduction that focuses on the rationale for or significance of sequencing. This introduction should reference appropriate literature but, due to space constraints, should not be an exhaustive review. A greater emphasis should be put on introducing the characteristics of the isolate(s) and providing a description of the provenance of the organism(s) in a manner that supports using the genome as a resource. The best practice is to comply with the Genomic Standards Consortium Minimum Information about any (x) Sequence (MIxS) checklist (1) (<https://gensc.org/mixs>), ensuring consistency with the same information available in the biosample accession record or equivalent. Type strains can be noted if the strain is listed as such in a type strain repository, such as the American Type Culture Collection (ATCC), or on a reference website, such as <https://bacterio.net>. We do

not allow claims of priority (e.g., first or novel); all genomes are new in their own unique ways without such claims.

Although an MRA genome announcement requires a taxonomic designation for the organism at the genus level, journal policy does not allow formal descriptions of sequenced organisms as new species, proposals for new taxonomy, or proposals for taxonomic reorganization. Such designations typically require much greater evidentiary support than can be reasonably included in a genome announcement. A formal taxonomic description should be directed to the *International Journal of Systematic and Evolutionary Microbiology* or *Archives of Virology*. Therefore, for genome announcements, the taxonomic nomenclature should have been described previously and should be consistent with established nomenclature rules (e.g., listed by <https://bacterio.net>). If authors would like to submit a resource describing the genome of an organism that does not have a species designation, the authors can designate the organism as a member of an existing genus (e.g., *Wolbachia* sp. strain *wAna* [2]) and note the organism's similarity to, or difference from, its closest relative(s) using appropriate techniques and analyses that are fully described.

The genome announcement should contain a description of how the isolate was acquired. This may include an accession number for a public culture collection, like the ATCC, BEI Resources, the ARS Culture Collection (NRRL), the German Collection of Microorganisms and Cell Cultures (DSMZ), or the Japanese Collection of Microorganisms. Adding information about how the strain was acquired and maintained can be helpful to others trying to interpret the sequencing results. It can be important to understand whether the specimen was acquired directly from a culture collection or from another scientist, along with a relative time frame and method for storing and/or passaging. If a new isolate is described, the genome announcement should include a description of when, where, and how the organism was isolated. A brief description can be followed by a citation to a peer-reviewed manuscript for the full isolation procedures. For new environmental isolates, the latitude and longitude or GPS coordinates of the sample site should be included. For clinical isolates, the best practice is to adhere to the metadata standards for human pathogen/vector genomic sequences (3). Authors of manuscripts describing research involving human or animal subjects must include a statement documenting the approval number and name of their institutional review board (IRB) or institutional animal care and use committee (IACUC). In all cases, the authors should make the sequenced isolate available to the community upon request; the best practice is to deposit the isolate in a culture repository (4).

While MRA does not allow for the inclusion of extensive experimental results, the rationale for sequencing may include a figure or table illustrating a specific trait, such as a strain's ability to produce fungicides or an antibiotic resistance profile. Those methods must be fully described in the main text or the figure legend with sufficient detail and/or references to be reproduced. If new antibiotic resistance profiles are included, the Clinical and Laboratory Standards Institute (CLSI) standards and methods used to determine antibiotic resistance should be fully described. Elaborate phenotypic results should not be included, as these often require extensive experimental validation. MRA does not allow references to unpublished results or personal communications.

SEQUENCING, ASSEMBLY, AND OTHER BIOINFORMATICS METHODS

Typically, the second section of a genome announcement includes a description of the methods and related outcomes. The paragraph must describe the methods for organism cultivation/acquisition, any taxonomic identification, DNA/RNA isolation, sequencing library preparation, and sequence generation, including platform specifications. The goal of this paragraph is to ensure that the sequencing procedure can be fully replicated and to enable full data reuse; therefore, details such as manufacturer, kit identifiers, and/or modifications to published protocols are essential. For Illumina sequencing, the platform should be described, along with read pairing status, the length of the reads, the number of raw reads in total and/or the sequencing depth, and

the methods for quality control and trimming, if applicable. For Pacific Biosciences (PacBio) sequencing, the platform should be described, preferably with the chemistry, along with the library construction method, whether and how DNA was sheared, whether and how DNA was size selected, the read N_{50} , the number of raw reads, and, if applicable, a description of read quality control, error correction, and adapter trimming. For Oxford Nanopore Technologies sequencing, the device and flow cell should be described, along with the library construction methods, the read N_{50} and number of raw reads, the base caller, read quality control, error correction, and, if applicable, adapter trimming. For libraries constructed with the ligation method (i.e., not RAD/RAPID libraries), whether and how DNA was sheared and size selected should be described. For capillary sequencing, which is still frequently used to sequence some viral genomes, the primers used to amplify and sequence the genome should be provided, which is often best accomplished in a table. The sequencing instrument should be specified, along with the length distribution of the reads, how the reads overlapped, and the Phred quality score threshold used for read trimming.

After describing the sequencing and preassembly quality control methods, the genome assembly and annotation methods should be fully described. Annotation methods vary widely between taxa, such that it is important to follow the best practices in the field, but a few principles are common across all taxa. All software should be cited and a version number included, even for common software like PGAP (5). Settings or options used to run the software should be provided, and we encourage including a statement such as “default parameters were used, except where otherwise noted,” if appropriate. Custom scripts must be made publicly available, and a link with a permanent DOI should be provided for the scripts (e.g., a GitHub repository with an assigned DOI using the data-archiving tool Zenodo). The annotation described in the genome announcement should be consistent with the data that are publicly available with the genome accession number(s) listed in the data availability section.

DESCRIBING THE RESULTS OF THE GENOME SEQUENCING

The final section describes the results, including the complete size, GC content, and final sequencing coverage of the genome. For genome announcements that include the genomes of more than one strain or organism, it is helpful to include a table with this information. That table should also include hyperlinked accession numbers for the genome and the raw data. For draft genomes, the announcement should include the relevant statistics for the assembly, including the number of contigs and the contig N_{50} value, as well as any method for ordering and orienting contigs. It should be clear what criteria, if any, were used for removal of contigs due to size or contamination screening. For complete linear genomes, this section describes how the ends of the chromosomes/genome were determined to be complete. For circular genomes, the method for identifying the overlap on the contig ends, trimming, and rotating, if applicable, should be specified.

Genome quality assessment using a conserved set of markers can be a useful and important metric for evaluating whether the genome is complete, particularly for large genomes. Assessment with tools like BUSCO (6), CEGMA (7), or CheckM (8) and/or whole-genome reference alignments with tools like MUMmer (9) or PacBio Quiver can be helpful for reporting genome completeness and duplication metrics but may not be universally appropriate. Visualization of assembly graphs can be helpful for assessing completeness using tools like Bandage (10).

Any remaining space is typically dedicated to describing the results in the context of the rationale for sequencing described in the introduction. Authors should avoid claims that imply that a particular gene or operon is functional in a sequenced organism. Functions of genes should be presumed and noted as “putative” or “potential,” unless they have been functionally characterized previously and a reference is provided.

A main figure can include a phylogeny to present the relationship of the new genome resource(s) to other isolates or species. When a phylogeny is included, the

methods should be rigorous and fully described in the main text or the figure legend, including the procedures used to select the included sequences, accession numbers for the included sequences, multiple sequence alignment, alignment curation, model selection, tree building with statistical support, and tree visualization.

DATA AVAILABILITY STATEMENT

All American Society for Microbiology (ASM) journals require the inclusion of a sufficient amount of publicly available data so that others can reproduce analyses and results from the published manuscripts. A list of acceptable databases is available (<https://journals.asm.org/content/list-data-repositories>). For MRA, the underlying data largely consist of sequencing reads and genome assemblies. Accession numbers hyper-linked directly to raw sequencing reads (e.g., SRX, SRR, or SRP accession numbers), from one of the acceptable repositories (International Nucleotide Sequence Database Collaboration [INSDC] [<http://www.insdc.org>], e.g., SRA or ENA), should be listed in the data availability section. Although it is acceptable to deposit reads that have been modified (for instance, through quality trimming), we ask that authors consider depositing the least modified reads and specify any types of modifications in the data availability statement. Data sets that include human reads should be placed in an appropriate repository, like dbGaP. For large collections of genomes, the data availability section can refer to a table in which the read accession numbers are provided for each isolate.

CONCLUSIONS

If these elements are all included and correct and no problems are identified with the rationale or conclusions, a manuscript could be designated by the editors a unicorn—a paper that is accepted on first submission without revisions. Interested in making your submission the next unicorn? Use the checklist and the examples (Table 1) to guide you in constructing a solid genome announcement. And remember, you can have one figure and one table to support your genome announcement.

ACKNOWLEDGMENTS

J.C.D.H., D.A.R., and V.M.B. were supported by federal funds from the National Institutes of Health, NIAID, under grant U19AI110820. J.C.D.H. was also supported by a National Institutes of Health Director's Transformative Research Award (grant 1-R01-CA206188). J.J.D. was supported by funding from the National Institutes of Health, NIGMS (grant R01-GM124446-01). I.L.G.N. was supported by funding from the National Institutes of Health, NIAID (grant R01-AI144430-02). K.M.S. was supported by funding from NASA (grant 80NSSC17K0301) and the National Science Foundation (grant MCB1929273). The work conducted by the U.S. Department of Energy Joint Genome Institute was supported by the Office of Science of the U.S. Department of Energy under contract DE-AC02-05CH11231 (S.R.). C.P. was supported by funding from the National Science Foundation (grant 1661357). S.R.G. was supported by federal funds from the Department of Defense (grant PR181406P1) and the NIAID (grant U19-AI117673). J.E.S. is CIFAR Fellow in the program Fungal Kingdom: Threats and Opportunities and was supported by NSF grants DEB 1441715 and 1557110.

REFERENCES

1. Yilmaz P, Kottmann R, Field D, Knight R, Cole JR, Amaral-Zettler L, Gilbert JA, Karsch-Mizrachi I, Johnston A, Cochrane G, Vaughan R, Hunter C, Park J, Morrison N, Rocca-Serra P, Sterk P, Arumugam M, Bailey M, Baumgartner L, Birren BW, Blaser MJ, Bonazzi V, Booth T, Bork P, Bushman FD, Buttigieg PL, Chain PSG, Charlson E, Costello EK, Huot-Creasy H, Dawyndt P, DeSantis T, Fierer N, Fuhrman JA, Gallery RE, Gevers D, Gibbs RA, San Gil I, Gonzalez A, Gordon JI, Guralnick R, Hankeln W, Highlander S, Hugenholtz P, Jansson J, Kau AL, Kelley ST, Kennedy J, Knights D, Koren O, Kuczynski J, Kyrpides N, Larsen R, Lauber CL, Legg T, Ley RE, Lozupone CA, Ludwig W, Lyons D, Maguire E, Methé BA, Meyer F, Muegge B, Nakielný S, Nelson KE, Nemergut D, Neufeld JD, Newbold LK, Oliver AE, Pace NR, Palanisamy G, Peplies J, Petrosino J, Proctor L, Pruesse E, Quast C, Raes J, Ratnasingham S, Ravel J, Relman DA, Assunta-Sansone S, Schloss PD, Schriml L, Sinha R, Smith MI, Sodergren E, Spo A, Stombaugh J, Tiedje JM, Ward DV, Weinstock GM, Wendel D, White O, Whiteley A, Wilke A, Wortman JR, Yatsunenko T, Glöckner FO. 2011. Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIxS) specifications. *Nat Biotechnol* 29:415–420. <https://doi.org/10.1038/nbt.1823>.
2. Gasser MT, Chung M, Bromley RE, Nadendla S, Dunning Hotopp JC. 2019. Complete genome sequence of *w*Ana, the *Wolbachia* endosymbiont of *Drosophila ananassae*. *Microbiol Resour Announc* 8:e01136-19. <https://doi.org/10.1128/MRA.01136-19>.
3. Dugan VG, Emrich SJ, Giraldo-Calderón GI, Harb OS, Newman RM, Pickett BE, Schriml LM, Stockwell TB, Stoeckert CJ, Sullivan DE, Singh I, Ward DV, Yao A, Zheng J, Barrett T, Birren B, Brinkac L, Bruno VM, Caler E, Chapman

- S, Collins FH, Cuomo CA, Di Francesco V, Durkin S, Eppinger M, Feldgarden M, Fraser C, Fricke WF, Giovanni M, Henn MR, Hine E, Dunning Hotopp J, Karsch-Mizrachi I, Kissinger JC, Lee EM, Mathur P, Mongodin EF, Murphy CI, Myers G, Neafsey DE, Nelson KE, Nierman WC, Puzak J, Rasko D, Roos DS, Sadzewicz L, Silva JC, Sobral B, Squires RB, Stevens RL, Tallon L, Tettelin H, Wentworth D, White O, Will R, Wortman J, Zhang Y, Scheuermann RH. 2014. Standardized metadata for human pathogen/vector genomic sequences. *PLoS One* 9:e99979. <https://doi.org/10.1371/journal.pone.0099979>.
4. Boundy-Mills K, Hess M, Bennett AR, Ryan M, Kang S, Nobles D, Eisen JA, Inderbitzin P, Sitepu IR, Torok T, Brown DR, Cho J, Wertz JE, Mukherjee S, Cady SL, McCluskey K. 2015. The United States Culture Collection Network (USCCN): enhancing microbial genomics research through living microbe culture collections. *Appl Environ Microbiol* 81:5671–5674. <https://doi.org/10.1128/AEM.01176-15>.
 5. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. *Nucleic Acids Res* 44: 6614–6624. <https://doi.org/10.1093/nar/gkw569>.
 6. Peppy M, Manni M, Zdobnov EM. 2019. BUSCO: assessing genome assembly and annotation completeness. *Methods Mol Biol* 1962: 227–245. https://doi.org/10.1007/978-1-4939-9173-0_14.
 7. Parra G, Bradnam K, Korf I. 2007. CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics* 23: 1061–1067. <https://doi.org/10.1093/bioinformatics/btm071>.
 8. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25:1043–1055. <https://doi.org/10.1101/gr.186072.114>.
 9. Delcher AL, Salzberg SL, Phillippy AM. 2003. Using MUMmer to identify similar regions in large sequence sets. *Curr Protoc Bioinformatics Chapter 10:Unit 10.3*.
 10. Wick RR, Schultz MB, Zobel J, Holt KE. 2015. Bandage: interactive visualization of de novo genome assemblies. *Bioinformatics* 31:3350–3352. <https://doi.org/10.1093/bioinformatics/btv383>.
 11. Baltrus DA, Clark M. 2020. Complete genome sequence of *Pseudomonas coronafaciens* pv. *oryzae* 1_6. *Microbiol Resour Annot* 9:e01564-19. <https://doi.org/10.1128/MRA.01564-19>.
 12. Baltrus DA, Clark M, Inderbitzin P, Pignatta D, Knight-Connoni V, Arnold AE. 2019. Complete genome sequence of *Luteibacter pinisoli* MAH-14. *Microbiol Resour Annot* 8:e00774-19. <https://doi.org/10.1128/MRA.00774-19>.
 13. Miller-Ensminger T, Wolfe AJ, Putonti C. 2020. Draft genome sequence of *Lactobacillus jensenii* strain UMB7766, isolated from the female bladder. *Microbiol Resour Annot* 9:e00392-20. <https://doi.org/10.1128/MRA.00392-20>.
 14. Baek SW, Hwang S, Kang HY, Bang WY, Moon KH. 2020. Draft genome sequence of a fish pathogen, *Edwardsiella piscicida* isolate CK41. *Microbiol Resour Annot* 9:e00061-20. <https://doi.org/10.1128/MRA.00061-20>.
 15. Tsurumaki M, Deno S, Galipon J, Arakawa K. 2019. Complete genome sequence of halophilic deep-sea bacterium *Halomonas axialensis* strain Althf1. *Microbiol Resour Annot* 8:e00839-19. <https://doi.org/10.1128/MRA.00839-19>.
 16. Tyson S, Peterson CL, Olson A, Tyler S, Knox N, Griffiths E, Dooley D, Hsiao W, Cabral J, Johnson RP, Laing C, Gannon V, Lynch T, Van Domselaar G, Brinkman F, Graham M. 2019. Eleven high-quality reference genome sequences and 360 draft assemblies of Shiga toxin-producing *Escherichia coli* isolates from human, food, animal, and environmental sources in Canada. *Microbiol Resour Annot* 8:e00625-19. <https://doi.org/10.1128/MRA.00625-19>.
 17. Holt A, Saldana R, Moreland R, Gill JJ, Liu M, Ramsey J. 2019. Complete genome sequence of *Escherichia coli* phage Paul. *Microbiol Resour Annot* 8:e01093-19. <https://doi.org/10.1128/MRA.01093-19>.
 18. Corban JE, Gramer J, Moreland R, Liu M, Ramsey J. 2019. Complete genome sequence of *Escherichia coli* siphophage Snoko. *Microbiol Resour Annot* 8:e01051-19. <https://doi.org/10.1128/MRA.01051-19>.
 19. Abrahamian P, Hammond J, Hammond RW. 2020. Complete genome sequence of an American isolate of pepino mosaic virus. *Microbiol Resour Annot* 9:e01124-19. <https://doi.org/10.1128/MRA.01124-19>.
 20. de la Higuera I, Torrance EL, Pratt AA, Kasun GW, Maluenda A, Stedman KM. 2019. Genome sequences of three cruciviruses found in the Willamette Valley (Oregon). *Microbiol Resour Annot* 8:e00447-19. <https://doi.org/10.1128/MRA.00447-19>.
 21. Mead ME, Raja HA, Steenwyk JL, Knowles SL, Oberlies NH, Rokas A. 2019. Draft genome sequence of the griseofulvin-producing fungus *Xylaria flabelliformis* strain G536. *Microbiol Resour Annot* 8:e00890-19. <https://doi.org/10.1128/MRA.00890-19>.
 22. Rosienski MD, Lee MK, Yu JH, Kaspar CW, Gibbons JG. 2018. Genome sequence of the extremely acidophilic fungus *Acidomyces richmondensis* FRIK2901. *Microbiol Resour Annot* 7:e01314-18. <https://doi.org/10.1128/MRA.01314-18>.
 23. Henry PM, Stueven M, Li S, Miyao EM, Gordon TR, Davis RM, Doan HK. 2019. Genome sequence of a California isolate of *Fusarium oxysporum* f. sp. *lycopersici* race 3, a fungus causing wilt disease on tomato. *Microbiol Resour Annot* 8:e01713-18. <https://doi.org/10.1128/MRA.01713-18>.
 24. Cuomo CA, Shea T, Nguyen T, Ashton P, Perfect J, Le T. 2020. Complete genome sequences for two *Talaromyces marneffei* clinical isolates from northern and southern Vietnam. *Microbiol Resour Annot* 9:e01367-19. <https://doi.org/10.1128/MRA.01367-19>.
 25. Tracey A, Foster JM, Paulini M, Grote A, Mattick J, Tsai Y-C, Chung M, Cotton JA, Clark TA, Geber A, Holroyd N, Korlach J, Libro S, Lustigman S, Michalski ML, Rogers MB, Twaddle A, Dunning Hotopp JC, Berriman M, Ghedin E. 2020. Nearly complete genome sequence of *Brugia malayi* strain FR3. *Microbiol Resour Annot* 9:e00154-20. <https://doi.org/10.1128/MRA.00154-20>.
 26. Stevens JTE, Fulweiler RW, Roy Chowdhury P. 2019. 16S rRNA amplicon sequencing of sediment bacterial communities in an oyster farm in Rhode Island. *Microbiol Resour Annot* 8:e01074-19. <https://doi.org/10.1128/MRA.01074-19>.
 27. Chung M, Teigen L, Libro S, Bromley RE, Kumar N, Sadzewicz L, Tallon LJ, Foster JM, Michalski ML, Dunning Hotopp JC. 2018. Multispecies transcriptomics data set of *Brugia malayi*, its *Wolbachia* endosymbiont wBm, and *Aedes aegypti* across the *B. malayi* life cycle. *Microbiol Resour Annot* 7:e01306-18. <https://doi.org/10.1128/MRA.01306-18>.
 28. Newton ILG, Sheehan KB. 2018. Gateway entry vector library of *Wolbachia pipientis* candidate effectors from strain wMel. *Microbiol Resour Annot* 7:e00806-18. <https://doi.org/10.1128/MRA.00806-18>.
 29. Cagle R, Ramachandran P, Reed E, Commichaux S, Mammel MK, Lacher DW, Miller D, Lindley S, Pop M, Bazaco MC, Wolpert B, Strain EA, Aker S, Feely J, Ottesen A. 2019. Microbiota of the Hickey Run Tributary of the Anacostia River. *Microbiol Resour Annot* 8:e00123-19. <https://doi.org/10.1128/MRA.00123-19>.
 30. Thrash JC, Baker BJ, Seitz KW, Temperton B, Campbell LG, Rabalais NN, Henrissat B, Mason OU. 2018. Metagenomic assembly and prokaryotic metagenome-assembled genome sequences from the northern Gulf of Mexico “dead zone.” *Microbiol Resour Annot* 7:e01033-18. <https://doi.org/10.1128/MRA.01033-18>.
 31. Bell TH, Trexler RV, Peng X, Huntemann M, Clum A, Foster B, Foster B, Roux S, Palaniappan K, Varghese N, Mukherjee S, Reddy TBK, Daum C, Copeland A, Ivanova NN, Kyrpidis NC, Pennacchio C, Eloe-Fadrosh EA, Bruns MA. 2020. Metatranscriptomic sequencing of a cyanobacterial soil-surface consortium with and without a diverse underlying soil microbiome. *Microbiol Resour Annot* 9:e01361-19. <https://doi.org/10.1128/MRA.01361-19>.
 32. Dunning Hotopp JC, Matsumura J, Bromley RE, Riley DR, Agrawal S, Sparklin B, Mattick J, Crabtree J, Mahurkar A. 2019. TwinBLAST: when two is better than one. *Microbiol Resour Annot* 8:e00842-19. <https://doi.org/10.1128/MRA.00842-19>.
 33. Lin D, Coombe L, Jackman SD, GagaloVA KK, Warren RL, Hammond SA, Kirk H, Pandoh P, Zhao Y, Moore RA, Mungall AJ, Ritland C, Jaquish B, Isabel N, Bousquet J, Jones SJM, Bohlmann J, Birol I. 2019. Complete chloroplast genome sequence of a white spruce (*Picea glauca*, genotype WS77111) from eastern Canada. *Microbiol Resour Annot* 8:e00381-19. <https://doi.org/10.1128/MRA.00381-19>.
 34. Stahlke AR, Ozsoy AZ, Bean DW, Hohenlohe PA. 2019. Mitochondrial genome sequences of *Diorhabda carinata* and *Diorhabda carinulata*, two beetle species introduced to North America for biological control. *Microbiol Resour Annot* 8:e00690-19. <https://doi.org/10.1128/MRA.00690-19>.
 35. Blakeway LV, Tan A, Peak IR, Atack JM, Seib KL. 2020. Proteome of a *Moraxella catarrhalis* strain under iron-restricted conditions. *Microbiol Resour Annot* 9:e00064-20. <https://doi.org/10.1128/MRA.00064-20>.