



**HAL**  
open science

## **MIxS-SA: a MIxS extension defining the minimum information standard for sequence data from symbiont-associated micro-organisms**

Fátima Jorge, Jaelle Brealey, Paul Brindley, Marie Buysse, Cinzia Cantacessi, Olivier Duron, Raina Fichorova, Connor Fitzpatrick, Megan Hahn, Christopher Hunter, et al.

### ► To cite this version:

Fátima Jorge, Jaelle Brealey, Paul Brindley, Marie Buysse, Cinzia Cantacessi, et al.. MIxS-SA: a MIxS extension defining the minimum information standard for sequence data from symbiont-associated micro-organisms. *ISME Communications*, 2022, 2 (1), pp.9. 10.1038/s43705-022-00092-w . hal-03701954

**HAL Id: hal-03701954**

**<https://hal.science/hal-03701954>**

Submitted on 5 Jan 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



**HAL**  
open science

## **MIxS-SA: a MIxS extension defining the minimum information standard for sequence data from symbiont-associated micro-organisms**

Fátima Jorge, Jaelle Brealey, Paul Brindley, Marie Buysse, Cinzia Cantacessi, Olivier Duron, Raina Fichorova, Connor Fitzpatrick, Megan Hahn, Christopher Hunter, et al.

### ► To cite this version:

Fátima Jorge, Jaelle Brealey, Paul Brindley, Marie Buysse, Cinzia Cantacessi, et al.. MIxS-SA: a MIxS extension defining the minimum information standard for sequence data from symbiont-associated micro-organisms. *ISME Communications*, 2022, 2 (1), pp.9. 10.1038/s43705-022-00092-w . hal-03796265

**HAL Id: hal-03796265**

**<https://hal.archives-ouvertes.fr/hal-03796265>**

Submitted on 10 Oct 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## BRIEF COMMUNICATION OPEN



# MlxS-SA: a MlxS extension defining the minimum information standard for sequence data from symbiont-associated micro-organisms

Fátima Jorge <sup>1</sup>✉, Jaelle C. Brealey <sup>2</sup>, Paul J. Brindley<sup>3</sup>, Marie Buysse <sup>4,5</sup>, Cinzia Cantacessi <sup>6</sup>, Olivier Duron<sup>4,5</sup>, Raina Fichorova<sup>7</sup>, Connor R. Fitzpatrick <sup>8</sup>, Megan Hahn<sup>9</sup>, Christopher Hunter <sup>10</sup>, Vincent Hervé <sup>11</sup>, Laura J. Knoll<sup>12</sup>, Kevin D. Kohl<sup>13</sup>, Marco Lalle <sup>14</sup>, Julius Lukeš <sup>15</sup>, Joaquín Martínez Martínez<sup>16</sup>, Susan L. Perkins<sup>17</sup>, Robert Poulin <sup>1</sup>, Karyna Rosario<sup>18</sup>, Adam C. Schneider <sup>19</sup>, Lynn M. Schriml<sup>20</sup>, Luke R. Thompson <sup>21,22</sup>, Ramona L. Walls<sup>23</sup> and Nolwenn M. Dheilly <sup>24,25</sup>✉

© The Author(s) 2022

The symbiont-associated (SA) environmental package is a new extension to the minimum information about any (x) sequence (MlxS) standards, established by the Parasite Microbiome Project (PMP) consortium, in collaboration with the Genomics Standard Consortium. The SA was built upon the host-associated MlxS standard, but reflects the nestedness of symbiont-associated microbiota within and across host-symbiont-microbe interactions. This package is designed to facilitate the collection and reporting of a broad range of metadata information that apply to symbionts such as life history traits, association with one or multiple host organisms, or the nature of host-symbiont interactions along the mutualism-parasitism continuum. To better reflect the inherent nestedness of all biological systems, we present a novel feature that allows users to co-localize samples, to nest a package within another package, and to identify replicates. Adoption of the MlxS-SA and of the new terms will facilitate reports of complex sampling design from a myriad of environments.

ISME Communications; <https://doi.org/10.1038/s43705-022-00092-w>

## INTRODUCTION

Interspecific interactions are ubiquitous across the Tree of Life. With the realization that eukaryotic organisms can harbor rich microbial communities, came also the view that these smaller partners may in fact play important roles in mediating host-symbiont associations, thus adding a further layer to this complex set of nested interactions, *i.e.* host-symbiont-microbe [1–8]. As the number of studies exploring the microorganisms associated with symbiotic organisms increases, likewise does the need for compliant standardized metadata that provides contextual information associated with each study and sample. Standardized

metadata allows for the integration of data across organisms, resources, and within data repositories. Here, we present the symbiont-associated (SA) environmental package as a new extension to the minimum information about any (x) sequence (MlxS) standards [9], which will be included in MlxS version 6. Whilst the MlxS-SA expands upon the MlxS host-associated environmental package [9], it reflects the need for a new standard that takes into account the distinct life history traits of symbionts, their association with one or multiple host organisms, the complex nature of host-symbiont interactions along the mutualism-parasitism continuum, and the nestedness of symbiont-associated

<sup>1</sup>Department of Zoology, University of Otago, Dunedin, New Zealand. <sup>2</sup>Department of Natural History, NTNU University Museum, Norwegian University of Science and Technology, Trondheim, Norway. <sup>3</sup>Department of Microbiology, Immunology and Tropical Medicine, School of Medicine & Health Sciences, George Washington University, Washington, D.C 20037, USA. <sup>4</sup>MIVEGEC (Maladies Infectieuses et Vecteurs: Ecologie, Génétique, Evolution et Contrôle), Centre National de la Recherche Scientifique (CNRS)—Institut pour la Recherche et le Développement (IRD) - Université de Montpellier (UM), Montpellier, France. <sup>5</sup>CREES (Centre de Recherche en Écologie et Évolution de la Santé), Montpellier, France. <sup>6</sup>Department of Veterinary Medicine, University of Cambridge, Cambridge CB3 0ES, UK. <sup>7</sup>Department of Obstetrics, Gynecology and Reproductive Biology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA. <sup>8</sup>Biology Department, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. <sup>9</sup>New York City Department of Health and Mental Hygiene, Long Island City, NY, USA. <sup>10</sup>GigaDB Director, GigaScience Press, BGI-Hong Kong, Hong Kong, China. <sup>11</sup>Institut de Recherche sur la Biologie de l'Insecte, UMR 7261, CNRS—Université de Tours, Avenue Monge, Parc Grandmont, 37200 Tours, France. <sup>12</sup>Department of Medical Microbiology and Immunology, University of Wisconsin-Madison, Madison, WI, USA. <sup>13</sup>Department of Biological Sciences, University of Pittsburgh, Pittsburgh, PA 15260, USA. <sup>14</sup>Department of Infectious Diseases, Foodborne and Neglected Parasitic Diseases; European Union Reference Laboratory for Parasite, Istituto Superiore di Sanità, Rome, Italy. <sup>15</sup>Institute of Parasitology, Biology Centre, Czech Academy of Sciences, and Faculty of Sciences, University of South Bohemia, České Budějovice (Budweis), České Budějovice, Czech Republic. <sup>16</sup>Bigelow Laboratory for Ocean Sciences, East Boothbay, ME, USA. <sup>17</sup>Department of Biology, The City College of New York, New York, NY 10031, USA. <sup>18</sup>College of Marine Science, University of South Florida, Saint Petersburg, FL, USA. <sup>19</sup>Biology and Health Sciences Department, Hendrix College, Conway, AR, USA. <sup>20</sup>Institute for Genome Sciences, Department of Epidemiology and Public Health, University of Maryland School of Medicine, Baltimore, MD 21201, USA. <sup>21</sup>Northern Gulf Institute, Mississippi State University, Mississippi State, MS 39762, USA. <sup>22</sup>Ocean Chemistry and Ecosystems Division, Atlantic Oceanographic and Meteorological Laboratory, National Oceanic and Atmospheric Administration, Miami, FL 33149, USA. <sup>23</sup>Data Collaboration Center, Critical Path Institute, Tucson, AZ, USA. <sup>24</sup>ANSES, Agence Nationale de Sécurité Sanitaire de l'Alimentation, de l'Environnement et du Travail - Laboratoire de Ploufragan-Plouzané, Unité Génétique Virale de Biosécurité, Ploufragan, France. <sup>25</sup>ANSES, INRAE, Ecole Nationale Vétérinaire d'Alfort, UMR VIROLOGIE, Laboratoire de Santé Animale, 94700 Maisons-Alfort, France. ✉email: fatima.esperancajorge@otago.ac.nz; nolwenn.dheilly@anses.fr

Received: 28 September 2021 Revised: 26 December 2021 Accepted: 10 January 2022

Published online: 01 February 2022

microbiota. We also propose adding the term ‘relationship to other packages’ to all environmental packages across the domains of life, to allow for integrated analysis of symbiont and host microbiota by linking metadata elements across environmental packages. This will allow users to nest a package within another package, and to identify replicates. This added feature is pivotal for the study of the microbiome of symbionts that are themselves nested within a host, reflects the inherent nestedness of all ecosystems and will facilitate reports of complex sampling design from a myriad of environments.

Collecting relevant metadata (data describing data) is now widely recognized as critical to contextualize samples and increase their reusability and reproducibility [9–11]. The Genomic Standards Consortium (GSC, <https://gensc.org>) has developed and maintains a suite of minimal information metadata standards for describing sequence metadata (checklists) for genome (MIGS), metagenome (MIMS), marker gene sequences (MIMARKS), simple amplified genome (MISAG), metagenome-assembled genome (MIMAG), virus genomes (MIUViG) and environmental packages for describing habitat-specific contextual data of the sampling environment [9, 10, 12, 13], collectively referred to as the Minimum Information about any (x) Sequence (MlxS) standard (ref. [9], <https://gensc.org/mixs/>).

The MlxS standards are used broadly across the microbiome research communities. These standards have been integrated into large scale microbiome projects (e.g. Human Microbiome Project, <https://www.hmpdacc.org/>), Earth Microbiome Project (<https://earthmicrobiome.org/>), Microbiology of the Built Environment (MoBE, <https://www.microbe.net>), microbiome bioinformatics platforms (e.g., QIIME, Qiita, mothur, JGI GOLD, MG-RAST, EBI, NCBI) and are now required upon manuscript submission. A primary advantage of the MlxS standards is the collation of large aggregates of associated metadata that can be harnessed to uncover, and eventually comprehend, patterns of microbial diversity and ecology.

The MlxS-SA package was initially drafted during the 1st Parasite Microbiome project workshop that involved the contribution of members of the GSC in addition to microbial ecologists, parasitologists, pathologists and marine biologists [14]. Participants rapidly identified the need to incorporate information on the nestedness of symbiont-associated systems, and the absence within the MlxS host-associated package of descriptors of complex life histories of mutualistic and parasitic symbionts. Until now, researchers have either omitted this information or added research-specific symbiont-associated annotations, limiting significantly the potential to compare, combine and/or reuse data from different systems and studies. Whereas the MlxS-SA package was initially designed for the study of parasite-microbes interaction, the scope of the package was expanded to include non-parasitic symbionts. This addition is a necessary expansion due to the context-dependent nature of symbiotic interactions and the ability of a given symbiont to interact differently with different organisms. Notably, the resulting MlxS-SA package reduces the need to develop additional highly similar packages for different types of symbionts.

Symbiotic associations are generally classified as mutualistic (mutually beneficial association), commensal (beneficial association to one of the partners, but not harmful to the other), or parasitic (detrimental association to one of the partners) [15]. In the context of the symbiont-associated package, the term symbiont applies to macro and microorganisms that can establish a physical interaction with at least one other organism at some stage of their life cycle regardless of the nature and dependence of the interaction. As such, this definition also covers symbiotic organisms that establish facultative and accidental associations (e.g., dead-end hosts), not requiring evolutionary processes to explain their association, but excludes free-living organisms that establish a symbiotic relationship with another free-living

organism (e.g., flowers and bees). The MlxS-SA package presented herein has gone through an open and iterative review process engaging the GSC community and experts studying symbiotic organisms across various symbiont and host taxa.

Here, we present the selected list of metadata descriptors for symbiont-associated microbiota studies, including a subset of mandatory (M) terms that underpin metadata compliance (Table 1; Supplementary Information SI-1 contains all MlxS-SA items). In order to allow comparative studies of the microbiota of, sometimes closely related, free-living and symbiotic organisms, the MlxS-SA includes terms already found in the MlxS host-associated package. Thus, in MlxS-SA, the term “host” (when used alone) refers to the host of the biological sample which is the symbiotic organism. New terms were created to characterise the “host of the symbiotic host”. We provide symbiont-associated package specific “Expected values” and “Examples”. Changes to the package (addition of terms, modification etc.) can be proposed by the community by creating a ticket on the MlxS GitHub page (<https://github.com/GenomicsStandardsConsortium/mixs>).

Given the diversity of symbiotic interactions and that the nature and dependence of such interactions can be context-dependent rather than a fixed trait, it was crucial to define terms and provide value syntax that were inclusive for diverse types of symbioses and also across the symbiont life histories and transmission processes. For example, the term “host dependence” (a mandatory item) and “type of symbiosis” (a conditional item) are discrete but complementary items. While “host dependence” aims to provide a general characterization of the known type of host dependence for the symbiotic organism (e.g., facultative), “type of symbiosis” was specifically designed to further characterize the type of biological interaction established between the symbiotic organism and its respective host at the moment the biological sample was taken (e.g., mutualistic). As a result, the MlxS-SA package features mandatory and conditionally mandatory, and optional features that enable flexibility according to the knowledge of the study system at the time of sampling. Two examples of MlxS-SA-compliant metadata are provided in Supplementary information (SI-2), and the respective study designs are presented in Fig. 1. The examples refer to 16S rRNA gene studies of (a) the bacterial communities of the parasite *Coitocaecum parvum*, a trematode, across four of its life stages: the sporocyst, the metacercaria and the adult, as well as the free-living cercaria [16], and (b) of the leaves and roots of the parasitic plant *Orobanche hederæ* and its ivy host, *Hedera* spp. [17].

While identical terms are often used in several of the 17 environmental packages currently available (<https://gensc.org/mixs/>), here we introduce three additional new terms: one is shared by several relevant MlxS environmental packages, and the two others will feature within the core MlxS package. The new term “observed host symbionts” provides a more comprehensive descriptor for the subject organism associations with smaller symbionts and it has been added to the host-associated, human-associated, plant-associated, human-vaginal, human-skin, human-oral and human-gut packages. The term “biotic relationship” has been added to the core package as a conditional descriptor of the relationship between the subject organism and other larger host organism(s). Finally, it appears necessary to include in the MlxS core a new term that takes into account the nested feature of most associations found in nature, such as host-symbiont-microorganism, in which multiple packages are necessary to describe the samples of the study (e.g., water, sediment, host-associated, and symbiont-associated). The proposed term “relationship to other samples” indicates the direct relationship between two samples from the same Bioproject, that are described in different environmental package(s). This proposed feature, still under development, will allow for integrated analyses of the microbiota of symbiotic organisms and their direct environment, even in the context of co-infections (e.g.,

**Table 1.** MixS symbiont-associated environmental package representative terms, along with requirement status, description and MIXS IDs.

MixS Package	Metadata category	Package item	Req	Definition	MIXS ID
Symbiont-associated	Symbiont-specific descriptors	host dependence	M	Type of host dependence for the symbiotic host organism to its host.	0001315
		type of symbiosis	C	Type of biological interaction established between the symbiotic host organism being sampled and its respective host	0001307
		symbiotic host organism life cycle type	M	Type of life cycle of the symbiotic host species (the thing being sampled). Simple life cycles occur within a single host, complex ones within multiple different hosts over the course of their normal life cycle.	0001300
		host life stage	M	Description of life stage of host.	0000251
		mode of transmission	C	The process through which the symbiotic host organism entered the host from which it was sampled.	0001312
		route of transmission	O	Description of path taken by the symbiotic host organism being sampled in order to establish a symbiotic relationship with the host (with which it was observed at the time of sampling) via a mode of transmission (specified in mode_transmission).	0001316
		host number individual	O	Number of symbiotic host individuals pooled at the time of collection.	0001305
	Symbiont – host relationship descriptors	observed host symbionts	O	The taxonomic name of the organism(s) found living in mutualistic, commensalistic, or parasitic symbiosis with the specific host. For cases when the specific host of the sample is a symbiont this field should refer to other organisms it is associated with: e.g.: hyperparasite species X (parasite of the parasite).	0001309
		host specificity	C	Level of specificity of symbiont-host interaction: e.g. generalist (symbiont able to establish associations with distantly related hosts) or species-specific.	0001308
		host of the symbiont role	C	Role of the host in the life cycle of the symbiotic organism.	0001303
		host cellular location	C	The localization of the symbiotic host organism within the host from which it was sampled: e.g., intracellular if the symbiotic host organism is localized within the cells or extracellular if the symbiotic host organism is localized outside of cells.	0001313
		duration of association with the host	O	Time spent in host of the symbiotic organism at the time of sampling; relevant scale depends on symbiotic organism and study.	0001299
		observed coinfecting organisms in host of host	O	The taxonomic name of any coinfecting organism observed in a symbiotic relationship with the host of the sampled host organism: e.g. where a sample collected from a host trematode species (A) which was collected from a host of host fish (B) that was also infected with a nematode (C), the value here would be (C) the nematode (species name) or (common name). Multiple coinfecting species may be added in a comma-separated list. For listing symbiotic organisms associated with the host (A) use the term Observed host symbiont.	0001310
	Host of the symbiont descriptors	host of the symbiotic host common name	O	Common name of the host of the symbiotic host organism.	0001324
		host of the symbiotic host local environmental context	O	For a symbiotic host organism the local anatomical environment within its host may have causal influences. Report the anatomical entity(s) which are in the direct environment of the symbiotic host organism being sampled and which you believe have significant causal influences on your sample or specimen. For example, if the symbiotic host organism being sampled is an intestinal worm, its local environmental context will be the term for intestine from UBERON ( <a href="https://uberon.github.io/">https://uberon.github.io/</a> ).	0001325
		host of the symbiotic host environmental medium	O	Report the environmental material(s) immediately surrounding the symbiotic host organism at the time of sampling. This usually will be a tissue or substance type from the host, but may be another material if the symbiont is	0001326

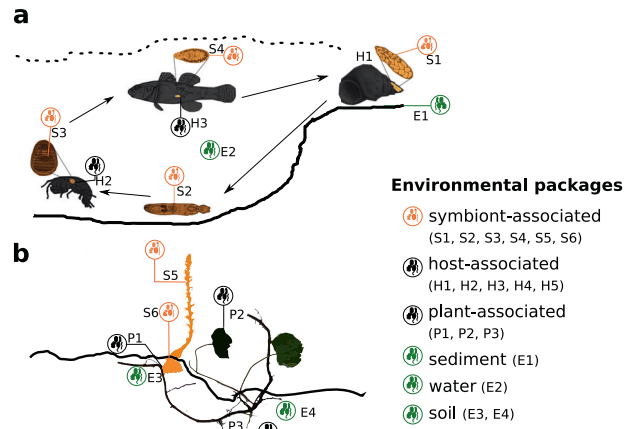
Table 1. continued

MiXs Package	Metadata category	Package item	Req	Definition	MiXs ID
		host of the symbiotic host taxon id	O	external to the host. We recommend using classes from the UBERON ontology, but subclasses of 'environmental material' ( <a href="http://purl.obolibrary.org/obo/ENVO_00010483">http://purl.obolibrary.org/obo/ENVO_00010483</a> ) may also be used. EnvO documentation about how to use the field: <a href="https://github.com/EnvironmentOntology/envo/wiki/Using-ENVO-with-MiXs">https://github.com/EnvironmentOntology/envo/wiki/Using-ENVO-with-MiXs</a> . Terms from other OBO ontologies are permissible as long as they reference mass/volume nouns (e.g., air, water, blood) and not discrete, countable entities (e.g., intestines, heart).	0001306
		host of the symbiotic host subject id	O	A unique identifier by which each host of the symbiotic host organism subject can be referred to, de-identified, e.g. #H14.	0001327
Core		Relationship to other samples <sup>a</sup>	C	indicates the direct relationship with another sample from the same Bioproject. Accepted terms are: 'technical replicate of', 'after', 'before', 'next to', 'within', and 'contains'. Can be repeated to reveal relationship to many samples.	TBD
		Biotic relationship	C	Description of relationship(s) between the subject organism and other organism(s) it is associated with. E.g., parasite on species X; mutualist with species Y. The target organism is the subject of the relationship, and the other organism(s) is the object.	0000028
Other <sup>b</sup>		Observed host symbionts	O	The taxonomic name of the organism(s) found living in mutualistic, commensalistic, or parasitic symbiosis with the specific host.	0001309

Req requirements: O optional, C conditional, M mandatory, TBD to be determined.

<sup>a</sup>Term may be slightly amended when included to all packages.

<sup>b</sup>Term added to host-associated, human-associated, plant-associated, human-vaginal, human-skin, human-oral, and human-gut packages.



**Fig. 1** Examples of study design for the sampling of microbes of symbiotic organisms, their hosts and environment. **a** Trematode *Coitocaeum parvum* different life stages (S1, S2, S3, S4) are reported with the MiXs-SA package. The microbiome of infected snail (H1), amphipod (H2) and fish (H3) hosts are reported with the MiXs host-associated package. In addition, the microbiome of environmental sediment (E1) and water (E2) from which these organisms were collected can be reported with MiXs-sediment and MiXs-water, respectively. The following relationships are reported: S1 “within” H1, S2 “within” E2, S3 “within” H2, S4 “within” H3, H1 “next to” E1, H1 “within” E2, H2 “within” E2, H3 “within” E2, E1 “next to” E2. **b** Angiosperm *Orobanchae hederiae* (S5, S6) parasitizing a host plant (P1, P2, P3) is reported using the included MiXs-SA and MiXs-PA (plant-associated) packages respectively. In addition, the MiXs-soil package is used to report corresponding soil samples. The following relationships are reported: S6 ‘within’ P1, S5 ‘next to’ S6, P1 ‘next to’ P2, P1 ‘next to’ P3, P2 ‘next to’ P3, P1 ‘within’ E3, P3 ‘within’ E4, E3 ‘next to’ E4.

symbiont-associated SA1234 is “within” host-associated sample HA8974, “next to” symbiont-associated sample SA7890). This feature will also benefit other studies by providing ecologically-relevant contextual information (e.g., host-associated HA2567 is “within” environmental water sample W1234, “next to” host-associated sample HA5679, ‘next’ to environmental soil sample S5897). In conclusion, it is our hope that the MiXs-SA, together with the new terms, will enable researchers to better conduct integrated analyses of multi-level biological systems with the ultimate goal of better understanding the role of microbes associated with symbionts.

REFERENCES

- Dheilly NM, Poulin R, Thomas F. Biological warfare: microorganisms as drivers of host-parasite interactions. *Infect Genet Evol.* 2015;34:251–9.
- Dheilly NM, Ewald PW, Brindley PJ, Fichorova RN, Thomas F. Parasite-microbe-host interactions and cancer risk. *PLoS Pathog.* 2019;15:e1007912.
- Bass D, Stentiford GD, Wang HC, Koskella B, Tyler CR. The pathobiome in animal and plant diseases. *Trends Ecol Evol.* 2019;34:996–1008.
- Brinker P, Fontaine MC, Beukeboom LW, Salles JF. Host, symbionts, and the microbiome: the missing tripartite interaction. *Trends Microbiol.* 2019;27:480–8.
- Barrow P, Dujardin JC, Fasel N, Greenwood AD, Osterrieder K, Lomonosoff G, et al. Viruses of protozoan parasites and viral therapy: Is the time now right? *Virology* 2020;17:142.
- Husnik F, Tashyreva D, Boscaro V, George EE, Lukeš J, Keeling PJ. Bacterial and archaeal symbioses with protists. *Curr. Biol.* 2021;31:R862–R877.
- Hahn MA, Dheilly NM. Experimental models to study the role of microbes in host-parasite interactions. *Front. Microbiol.* 2016;7:1300.
- Robinson AJ, House GL, Morales DP, Kelliher JM, Gallegos-Graves LV, LeBrun ES, et al. Widespread bacterial diversity within the bacteriome of fungi. *Commun Biol.* 2021;4:1168.
- Yilmaz P, Kottmann R, Field D, Knight R, Cole JR, Amaral-Zettler L, et al. Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MiXs) specifications. *Nat Biotechnol.* 2011;29:415–20.

10. Bowers R, Kyrpides NC, Stepanauskas R, Harmon-Smith M, Doud D, Reddy TBK, et al. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat Biotechnol.* 2017;35:725–31.
11. Schriml LM, Chuvochina M, Davies N, Eloë-Fadrosch EA, Finn RD, Hugenholtz P, et al. COVID-19 pandemic reveals the peril of ignoring metadata standards. *Sci Data.* 2020;7:188.
12. Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, Tatusova T, et al. The minimum information about a genome sequence (MIGS) specification. *Nat Biotechnol.* 2008;26:541–7.
13. Roux S, Adriaenssens EM, Dutilh BE, Koonin EV, Kropinski AM, Krupovic M, et al. Minimum Information about an Uncultivated Virus Genome (MIUViG). *Nat Biotechnol.* 2019;37:29–37.
14. Dheilly NM, Martínez Martínez J, Rosario K, Brindley PJ, Fichorova RN, Kaye JZ, et al. Parasite microbiome project: grand challenges. *PLoS Pathog.* 2019;15: e1008028.
15. Leung TLF, Poulin R. Parasitism, commensalism, and mutualism: exploring the many shades of symbioses. *Vie Milieu.* 2008;58:107–15.
16. Jorge F, Dheilly NM, Poulin R. Persistence of a core microbiome through the ontogeny of a multi-host parasite. *Front Microbiol.* 2020;11:954.
17. Fitzpatrick CR, Schneider AC. Unique bacterial assembly, composition, and interactions in a parasitic plant and its host. *J Exp Bot.* 2020;71:2198–209.

## ACKNOWLEDGEMENTS

This contribution is part of the Parasite Microbiome Project. The authors sincerely thank the Gordon and Betty Moore Foundation for fully sponsoring the 1st Parasite Microbiome project workshop, and for supporting Open Access publishing. FJ was funded by a grant from the Marsden Fund (Royal Society of New Zealand) to RP (Principal Investigator) and NMD (Associate Investigator). The National Institutes of Health grants R01CA164719 (PJB), and R01AI144016-01(LJK). VH was supported by the FEDER grant InFoBioS n°EX011185 (région Val de Loire). JMM acknowledges support from award OCE-1933285, National Science Foundation, USA. ML is part of the PARADISE consortium, supported by funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 773830: One Health European Joint Programme. JL was supported by the ERD funds (16\_019/0000759).

## AUTHOR CONTRIBUTIONS

NMD conceptualized the MixS package. All authors contributed to the writing and design of the MixS-SA package. FJ and NMD prepared the manuscript. All authors agreed with the final version of the package and manuscript.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s43705-022-00092-w>.

**Correspondence** and requests for materials should be addressed to Fátima Jorge or Nolwenn M. Dheilly.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022