Mycoavidus sp. Strain B2-EB: Comparative Genomics Reveals Minimal Genomic Features Required by a Cultivable *Burkholderiaceae*-Related Endofungal Bacterium

Yong Guo,[a] Yusuke Takashima,[a] Yoshinori Sato,[b] Kazuhiko Narisawa,[a] Hiroyuki Ohta,[a] Tomoyasu Nishizawa[a]

[a]Ibaraki University College of Agriculture, Ami, Ibaraki, Japan

[b]Center for Conservation Science, Tokyo National Research Institute for Cultural Properties, Tokyo, Japan

Yong Guo and Yusuke Takashima contributed equally to this article. Author order was determined based on overall contributions to the data presented and writing of the manuscript.

ABSTRACT Obligate bacterial endosymbionts are critical to the existence of many eukaryotes. Such endobacteria are usually characterized by reduced genomes and metabolic dependence on the host, which may cause difficulty in isolating them in pure cultures. Family *Burkholderiaceae*-related endofungal bacteria affiliated with the *Mycoavidus-Glomeribacter* clade can be associated with the fungal subphyla Mortierellomycotina and Glomeromycotina. In this study, a cultivable endosymbiotic bacterium, *Mycoavidus* sp. strain B2-EB, present in the fungal host *Mortierella parvispora* was obtained successfully. The B2-EB genome (1.88 Mb) represents the smallest genome among the endofungal bacterium *Mycoavidus cysteinexigens* (2.64–2.80 Mb) of *Mortierella elongata* and the uncultured endosymbiont "*Candidatus* Glomeribacter gigasporarum" (1.37 to 2.36 Mb) of arbuscular mycorrhizal fungi. Despite a reduction in genome size, strain B2-EB displays a high genome completeness, suggesting a nondegenerative reduction in the B2-EB genome. Compared with a large proportion of transposable elements (TEs) in other known *Mycoavidus* genomes (7.2 to 11.5% of the total genome length), TEs accounted for only 2.4% of the B2-EB genome. This pattern, together with a high proportion of single-copy genes in the B2-EB genome, suggests that the B2-EB genome reached a state of relative evolutionary stability. These results represent the most streamlined structure among the cultivable endofungal bacteria and suggest the minimal genome features required by both an endofungal lifestyle and artificial culture. This study allows us to understand the genome evolution of *Burkholderiaceae*-related endosymbionts and to elucidate microbiological interactions.

IMPORTANCE This study attempted the isolation of a novel endobacterium, *Mycoavidus* sp. B2-EB (JCM 33615), harbored in the fungal host *Mortierella parvispora* E1425 (JCM 39028). We report the complete genome sequence of this strain, which possesses a reduced genome size with relatively high genome completeness and a streamlined genome structure. The information indicates the minimal genomic features required by both the endofungal lifestyle and artificial cultivation, which furthers our understanding of genome reduction in fungal endosymbionts and extends the culture resources for biotechnological development on engineering synthetic microbiomes.

KEYWORDS bacterial endosymbiont, *Mycoavidus*, comparative genomics, genomic reduction, minimal genome features

Heritable bacterial endosymbionts are widespread and critical to the existence of many eukaryotes (1–4) and are usually characterized by rapid molecular evolution, drastically reduced genomes, and heavy metabolic dependence on the host over

evolutionary time (5, 6). Such long-term associations lead to very strong integration of the endosymbionts, ultimately reaching a state similar to the ancestors of mitochondria and plastids (7). At the same time, degenerative evolution may fix slightly deleterious mutations in the genome, resulting in the deletion of genetic elements essential for living extracellularly (8, 9). In contrast to the essential mutualists of insects, nonessentially heritable endosymbionts of fungi display nondegenerative genome evolution associated with purging of slightly deleterious mutations from the genomes, similar to the ability of free-living bacteria (10).

Bacterial endosymbionts have been discovered in a number of fungi affiliated with the major fungal phyla Ascomycota, Basidiomycota, and Mucoromycota (11–15). Family *Burkholderiaceae*-related endofungal bacteria (BRE) associated with Mucoromycota are the most widely observed group in recent studies (16). To date, three groups of BRE have been detected, "*Candidatus* Glomeribacter gigasporarum" in the arbuscular mycorrhizal fungi (AMF) of the family *Gigasporaceae* (Glomeromycotina) (17), *Mycetohabitans rhizoxinica* (previously designated *Burkholderia rhizoxinica* or *Paraburkholderia rhizoxinica*) and *Mycetohabitans endofungorum* (previously designated *Burkholderia endofungorum* or *Paraburkholderia endofungorum*) in the phytopathogenic fungus *Rhizopus microsporus* (Mucoromycotina) (18, 19), and *Mycoavidus cysteinexigens* in the soilborne fungus *Mortierella elongata* (Mortierellomycotina) (20).

The genus *Mortierella* possesses a high species diversity, with more than 100 described species from various habitats, such as soil, plant roots, arthropod bodies, and animal dung (21). Except for *Mycoavidus* species in *Mortierella elongata* and *Mortierella minutissima* (15, 22, 23), more-diverse BRE may be present in untested *Mortierella* species. Notably, in a recent examination of the presence/absence of endosymbionts across 238 *Mortierella* isolates, more-diverse BRE were newly identified in 22 of 59 tested *Mortierella* species (4). On the other hand, once genome reduction occurs in an endosymbiont due to close host association, the difficulty in its cultivation becomes significantly greater (24). Even so, among these BREs, *Mycetohabitans* species were isolated by culturing host fungal mycelial homogenates on nutrient agar I (Serva) plates (18), whereas a fractionated "*Ca*. Glomeribacter gigasporarum" endobacterial suspension was obtained from *Gigaspora margarita* for 2 weeks of incubation using several kinds of liquid and solid media by supplementation with various vitamins and amino acids (25). A pure culture of *M. cysteinexigens* was successfully obtained after aerobic incubation of filtered host mycelial homogenates on cysteine-containing buffered charcoal-yeast extract agar (B-CYEα) plates (20), which was suggested from genomic insights into the lack of key genes involved in cysteine biosynthesis in the endobacterial genomes (26).

Genome expansion has been identified predominantly in pathogenic and environmental bacteria in the genera *Burkholderia*, *Paraburkholderia*, *Cupriavidus*, and *Ralstonia* in the family *Burkholderiaceae* (27), which possess multipartite genomes (total genome size, 5 to 12 Mb) consisting of a chromosome and other essential secondary replicons (27, 28). Likewise, some BRE also contain secondary replicons of megaplasmids (25, 29), except for the *Mycoavidus* endosymbionts, which contain only a single circular chromosome (22, 30). However, the BRE genomes are gradually constricted (chromosomal genome size, ~2.8 Mb) and lack many functional genes associated with environmental adaptation (22). Comparing with *Mycetohabitans* genomes, *Mycoavidus* genomes are smaller and lack many genes associated with host invasion, glycolytic pathways, sugar importers, and ATP/ADP antiporters, which are comparable to the genomes of uncultured "*Ca*. Glomeribacter gigasporarum" endosymbionts (22, 30). In addition, due to the intimate connection between genome reduction and mobile genetic elements (31, 32), massive expansions of transposons and prophages in *Mycoavidus* genomes revealed a possibility of discovering novel BRE with more-streamlined genomes (30). Although a few heritable BRE were successfully isolated in pure cultures, there is still a lack of a comprehensive understanding of minimal genomes containing only essential genetic elements for these cultivable endosymbionts.

The discovery of endobacteria harbored by *Mortierella* species can be expected to

provide resources to obtain novel cultivable BRE. A *Mycoavidus* bacterium dwelling in the fungus *Mortierella parvispora* was newly obtained in this study. We also characterized the morphological features of the *Mycoavidus* bacterium inside host hyphal cells using ultramicroscopic observations with transmission electron microscopy (TEM). In addition, we carried out genome sequence analysis and performed comparative genomic analyses with the previously determined genomes of *M. cysteinexigens* and "*Ca*. Glomeribacter gigasporarum" to identify the minimal genomic features associated with cultivable BRE. This study explores a more comprehensive understanding of genome evolution adapting to the fungal host metabolism, enabling an identification of the minimal genomic features required by both an endofungal lifestyle and artificial culture.

RESULTS AND DISCUSSION

A *Burkholderiaceae*-related endofungal bacterium dwelling in *Mortierella parvispora* E1425 was detected by sequencing the DNA fragment amplified by a 16S rRNA gene-targeted PCR with a template DNA extracted from the mycelia of E1425, and its intracellular feature was subsequently confirmed by fluorescence microscopic observation with a staining technique (4). Before we attempted to cultivate the BRE, strain E1425 was sustained by subculture of the mycelia on a half-strength cornmeal-malt-yeast (CMMY) agar plate once for a month. A diagnostic PCR was carried out to check for the presence of BRE at 1 week after the subculture was made. In addition, before ultramicroscopic observation of strain E1425, the bacterium-like cells inside the hyphae were also detected by fluorescence *in situ* hybridization with the universal bacterial 16S rRNA gene probe EUB338 (data not shown), as previously reported (4). These continuous positive detections suggested a stable association of the endobacterium and *M. parvispora* E1425.

Intracellular features of BRE dwelling in *Mortierella parvispora*. The intracellular features of BRE dwelling in *Mortierella parvispora* E1425 were characterized by transmission electron microscopy (TEM) (Fig. 1). Bacterium-like cells were observed inside the host-fungal hyphae; they were rod shaped and 0.3 to 0.5 by 0.8 to 1.4 μm in size, with a smooth double-layered cell envelope and a nucleoid surrounded by numerous ribosomes in the cell center. As in previous reports (15, 22), the endobacterial cells aggregated in close proximity to lipid bodies (Fig. 1A and B). In addition, our microscopic observations showed that endobacterial cells were also located near other organelles such as vacuoles and mitochondria, even directly in contact with them (Fig. 1C to E). In contrast to other *Mycoavidus*-related endofungal bacteria (4, 15), the endobacteria in *M. parvispora* were not enveloped in fungal vacuoles according to our observations. Together, these morphological features extend the understanding of intimate interactions between the BRE and their host fungus at the subcellular level.

Isolation of endofungal bacteria, genome sequencing, and general genomic features. To date, only one endofungal bacterium, *M. cysteinexigens* B1-EB[T], hosted by fungi in the genus *Mortierella* has been successfully isolated in pure cultures (20). The BRE present in *M. parvispora* E1425 was identified as the closest relative of strain B1-EB[T], based on phylogenetic analyses using an almost-full-length 16S rRNA gene (4). In order to isolate the BRE, aliquots from filtered fungal homogenates of *M. parvispora* E1425 were incubated on B-CYEα agar plates for 30 days at 23°C. Because the bacterial colonies appearing on the plate were homogeneous, a single colony from the plate was subcultured repeatedly on new B-CYEα agar plates to assess its purity. A pure culture of a Gram-negative rod-shaped bacterium was obtained and designated strain B2-EB (Fig. 2A). Although this strain shared 99.0% nucleotide identity in the 16S rRNA gene sequence with strain B1-EB[T], its colonies on B-CYEα agar plates showed a color (translucent to white) different from that of strain B1-EB[T] (white to cream) (Fig. 2B). This result implied that a novel cultivable BRE strain had been successfully isolated in pure cultures.

Taxonomic assessment of strain B2-EB based on whole-genome sequencing. Genomic DNA of strain B2-EB was separately sequenced on Illumina HiSeq2500 and
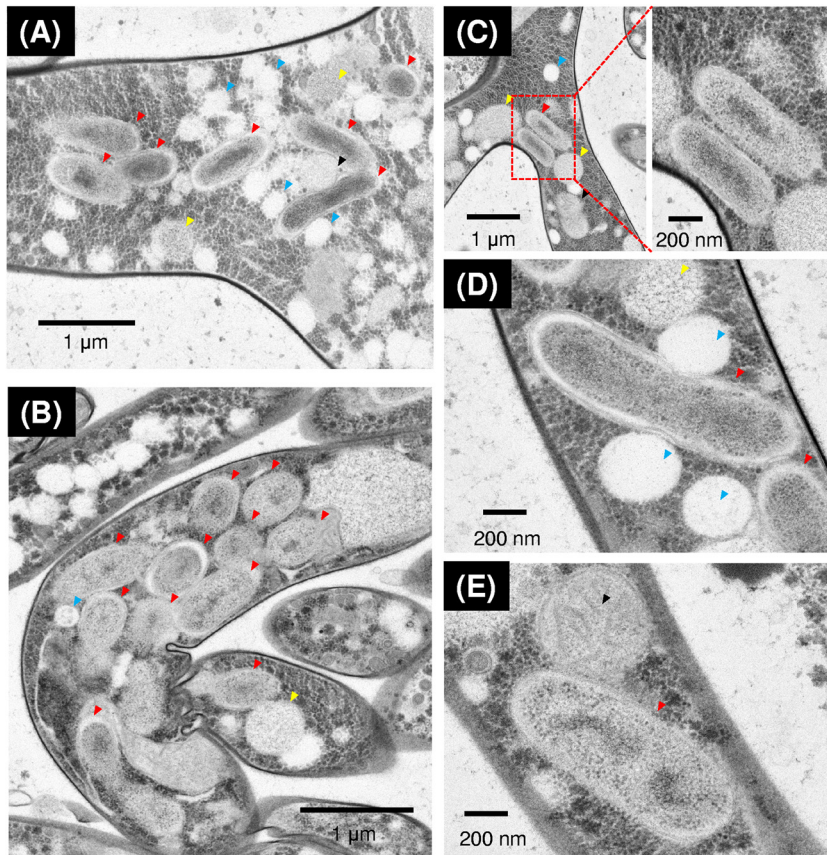
FIG 1 TEM observation of the endobacteria within the rapid-freezing and freeze-substituted hyphae. (A) Endobacterial cells aggregated in close proximity to a region filled with many lipid bodies. (B) Endobacterial cells aggregated in a region containing few lipid bodies. (C, D, and E) Endobacterial cells directly in contact with vacuoles, lipid bodies, and mitochondria, respectively. Red, yellow, blue, and black arrowheads indicate endofungal bacteria, vacuoles, lipid bodies, and mitochondria, respectively, inside the hyphal cells of *Mortierella parvispora* strain E1425.

PacBio RS II platforms, and 1.48 Gb of paired-end reads and 0.93 Gb of long reads were obtained, respectively. Hybrid assembly and read correction using SPAdes (version 3.10.1) and Pilon (version 1.22) resulted in a 1,876,900-bp circular chromosome with 1,117-fold genome coverage and 48.9 mol% GC content, which contained 2 rRNA operons (*rrn*), 42 tRNA loci for 20 amino acids, and 1,627 protein-coding sequences (CDSs) with an average size of 1,027 bp (Table 1). Of these CDSs, 1,319 were identified as single-copy CDSs accounting for 81% of total CDSs. In addition, 1,157 and 1,069 CDSs
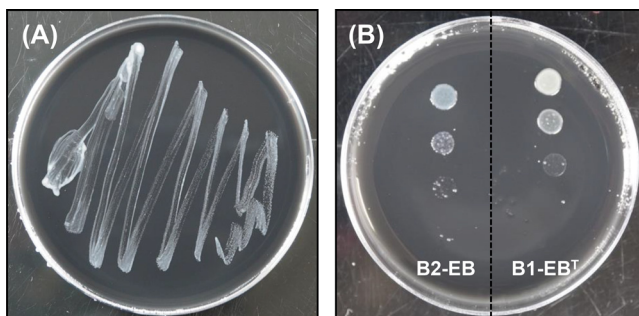

FIG 2 Photographs showing a pure culture of strain B2-EB grown at 23°C for 30 days (A) and a comparison of colonies of strains B2-EB and B1-EB^T on B-CYEα agar plates (B). In the right panel, colonies grown for 7 days by inoculating bacterial suspensions with absorbances of 0.03, 0.003, 0.0003, and 0.00003 at 600 nm are shown from top to bottom.

TABLE 1 Genomic features of *Mycoavidus* sp. strain B2-EB compared with those of other *Burkholderiaceae*-related endofungal bacteria

Feature	*Mycoavidus*			*Mycetohabitans*		"*Ca.* Glomeribacter gigasporarum"			
	B2-EB	B1-EB[T]	AG77	HKI 454[T]	HKI 456[T]	BEG34	BEG1	JA201A	IN211
Assembly level	Complete	Complete	Complete	Complete	Contig	Scaffold	Contig	Scaffold	Scaffold
No. of plasmids	0	0	0	2	No data	1	No data	No data	No data
Size (bp)	1,876,900	2,795,004	2,638,116	3,750,138	3,288,408	1,726,950	2,355,846	1,621,371	1,339,701
GC content (%)	48.9	48.9	49.0	60.7	61.9	54.8	49.7	50.7	51.1
No. of:									
CDSs	1,627	2,317	2,255	3,878	2,988	1,736	2,499	2,717	2,202
Single-copy CDSs (%)	1,319 (81)	1,476 (64)	1,493 (66)	2,204 (57)	2,172 (73)	1,280 (74)	1,293 (52)	994 (37)	764 (35)
COG orthologs	1,157	1,949	1,862	2,782	2,569	1,530	1,876	1,573	1,275
KEGG orthologs	1,069	1,174	1,338	2,030	1,903	1,089	1,288	1,249	965
rrn operons	2	2	2	3	1	1	1	No data	No data
tRNA loci	42	41	41	47	52	38	51	No data	No data
Accession or ID no.	AP021872	AP018150	224135.3[a]	GCA_000198775	GCA_002927045	GCA_000227585	GCA_001684025	GCA_001684155	GCA_001684175
Reference	This study	30	22	29	19	6	10	10	10

[a]The metagenome-assembled genome sequence of the AG77 endobacterium is available through the PATRIC website (www.patricbrc.org) under the genome identification (ID) 224135.3.
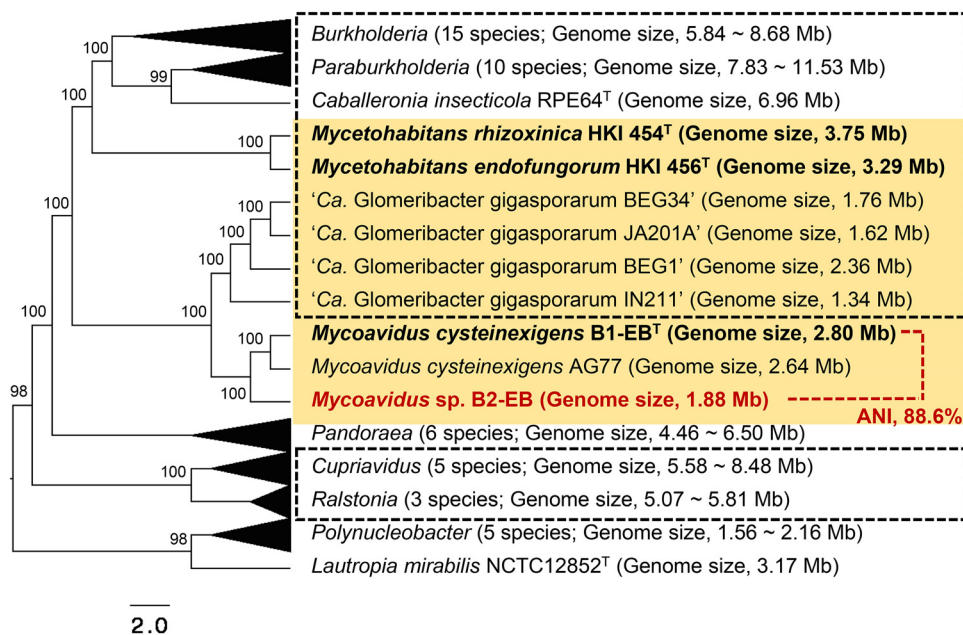
FIG 3 Phylogenetic identification of strain B2-EB using a reconstructed maximum-likelihood (ML) tree based on a concatenated amino acid sequence of 52 single-copy housekeeping genes, indicating the relative placement of the three *Burkholderiaceae*-related endofungal bacteria (yellow background) and other genera in the family *Burkholderiaceae*. The horizontal lines show genetic distances, which are supported by values estimated with 1,000 bootstrap replicates. The isolated BRE in pure cultures are shown in bold, and bacteria with multipartite genomes are outlined with dashed lines.

were assigned to COG and KEGG orthologs, respectively (Table 1); 215 CDSs matched hypothetical proteins of unknown function, and 59 of them were less than 100 amino acid residues in size (data not shown). Compared with other BRE genomes, the B2-EB genome featured the lowest GC content, the highest proportion of single-copy genes, and the smallest genome size among the known cultured BRE.

A reconstructed maximum-likelihood (ML) tree based on concatenated amino acid sequences of 52 single-copy housekeeping genes conserved in the *Burkholderiaceae* genomes showed that strain B2-EB was phylogenetically closest to the cluster of the species *M. cysteinexigens*, forming an independent clade neighboring the "*Ca.* Glomeribacter gigasporarum" clade (Fig. 3). This result was consistent with previous analyses based on the 16S rRNA gene sequence (4, 22). In addition, the whole-genome average nucleotide identity (ANI) between strains B2-EB and B1-EB[T] (88.6%) was lower than the ANI cutoff value of same species (94 to 95%) (33), suggesting that strain B2-EB can represent an undescribed species in the genus *Mycoavidus*. Moreover, our phylogenetic analysis using 49 representative complete genomes in the family *Burkholderiaceae* showed that *Mycetohabitans* endosymbionts hosted by *Rhizopus microsporus* were closer to bacteria with large and multipartite genomes in the genera *Caballeronia*, *Paraburkholderia*, and *Burkholderia* than to *Mycoavidus* and "*Ca.* Glomeribacter gigasporarum" endosymbionts in the *Mortierella* spp. and AMF, respectively (Fig. 3). These results supported that *Mycetohabitans* endosymbionts have an independent origin, sharing their most recent common ancestor with the free-living bacteria in these genera, whereas *Mycoavidus* and "*Ca.* Glomeribacter gigasporarum" endosymbionts seemed to originate from another common ancestor (16). From the cultivability, translucent colony color, and ANI analysis results, strain B2-EB could represent an undescribed species in the genus *Mycoavidus*.

Genomic comparison of *Mycoavidus* sp. B2-EB with *Mycoavidus cysteinexigens* B1-EB[T] and AG77. Although the genomic features of *Mycoavidus* sp. strain B2-EB, including GC content, number of *rrn* and number of tRNA loci, were close to those of *M. cysteinexigens* B1-EB[T], the other features of genome size, mobile genetic elements,

and proportion of single-copy genes were strikingly different between these two cultured endofungal bacteria (Table 1). Compared to the two complete genomes of *M. cysteinexigens* B1-EB^T and AG77, strain B2-EB possessed an approximately ~1.0-Mb-smaller genome with a paucity of transposable elements and a large proportion of single-copy genes (Fig. 4A; see Fig. S1 in the supplemental material). However, phylo-genetic orthology analysis using OrthoFinder showed that the B2-EB genome shared 91.3% of the orthologous groups (OGs) with the two *M. cysteinexigens* genomes, which included 1,395 core OGs plus 13 and 33 OGs shared with the B1-EB^T and AG77 genomes, respectively (Fig. 4B). These results indicated that although the genome size of strain B2-EB was reduced, its genetic elements were highly homologous with those of *M. cysteinexigens*.

Further assessment of genomic variation on a population scale showed that the ratio of nonsynonymous to synonymous sites (d_N/d_S) for individual genes between the B2-EB and *M. cysteinexigens* genomes displayed a leptokurtic distribution clustered around the mean value regardless of the d_N/d_S estimation method (Fig. 4C). This pattern was similar to those found in nonessential heritable endosymbionts such as *Wolbachia*, *Hamiltonella*/*Regiella*, and "*Ca.* Glomeribacter," as well as those of some free-living bacterial lineages such as *Burkholderia*, *Prochlorococcus*, *Bradyrhizobium*, and *Bifidobac-terium* (10). Additionally, the mean d_N/d_S ratio between B2-EB and *M. cysteinexigens* estimated by three methods was in the range 0.118 to 0.152, suggesting strong purifying selection and large population size like those detected in some microbial endosymbionts (10, 34). This tendency was even stronger than in the case of the "*Ca.* Glomeribacter gigasporarum*" endosymbionts (d_N/d_S, 0.149 to 0.198) estimated in the previous study (10). Therefore, *Mycoavidus* sp. B2-EB represents a model of nondegen-erative genome evolution similar to that for "*Ca.* Glomeribacter gigasporarum" endo-symbionts. This also created a debate on whether the B2-EB genome is reduced or the B1-EB^T and AG77 genomes are extended (35).

A syntenic comparison with *M. cysteinexigens* B1-EB^T showed that 936 CDSs related mainly to a region for prophages and transposons and accounting for 40.4% of total CDSs in the B1-EB genome were missing in the B2-EB genome (Fig. 5A). Even so, some gene loci within the translatable regions of the B1-EB^T and AG77 genomes were still found in the B2-EB genome (Fig. 5A and B).These residual genes indicate that genome reduction was likely to have resulted in the relatively small genome of strain B2-EB. Therefore, we proposed that the genome reduction occurring in strain B2-EB may be mediated by the repeated insertion-deletion of prophages and transposons, as previ-ously reported (31, 32, 36). The missing CDSs affiliated with the orthologous genes conserved in the genomes of *Burkholderiaceae*-related bacteria seemed to be reduced from the B2-EB genome over evolution and were involved mainly in the secretion system, amino acid and glycolytic metabolism, and transcription (see Fig. S2 in the supplemental material). Notably, the other specific CDSs in the B1-EB^T genome might be acquired through horizontal gene transfers, such as genes encoding insecticidal toxins close to orthologs possessed by the genus *Xenorhabdus* (*Gammaproteobacteria*) living symbiotically with soil entomopathogenic nematodes (37). However, the process of this gene gain event is still unclear.

On the other hand, 1,490 out of 1,627 CDSs in total in the B2-EB genome were well conserved in the B1-EB genome (Fig. 5A), and 60.5% of these commonly conserved CDSs shared more than 90% amino acid identity (AAI) between these two strains (see Fig. S3 in the supplemental material). Orthologous genes sharing low AAI values of <80% accounted for 21.3% of the total common orthologous genes, which encoded mainly proteins involved in RNA processing and modification, transcription, and energy production and conversion. In the case of the obligate endosymbiont *Wolbachia* associated with arthropod hosts, the specific genes for involvement in host interaction underwent elevated substitution rates that may be due in part to adaptations by *Wolbachia* to a new host environment (38). It is tempting to speculate that the genes with elevated mutation rates may represent specific gene elements in the two *My-coavidus* genomes for adapting to the different intracellular environments of the two
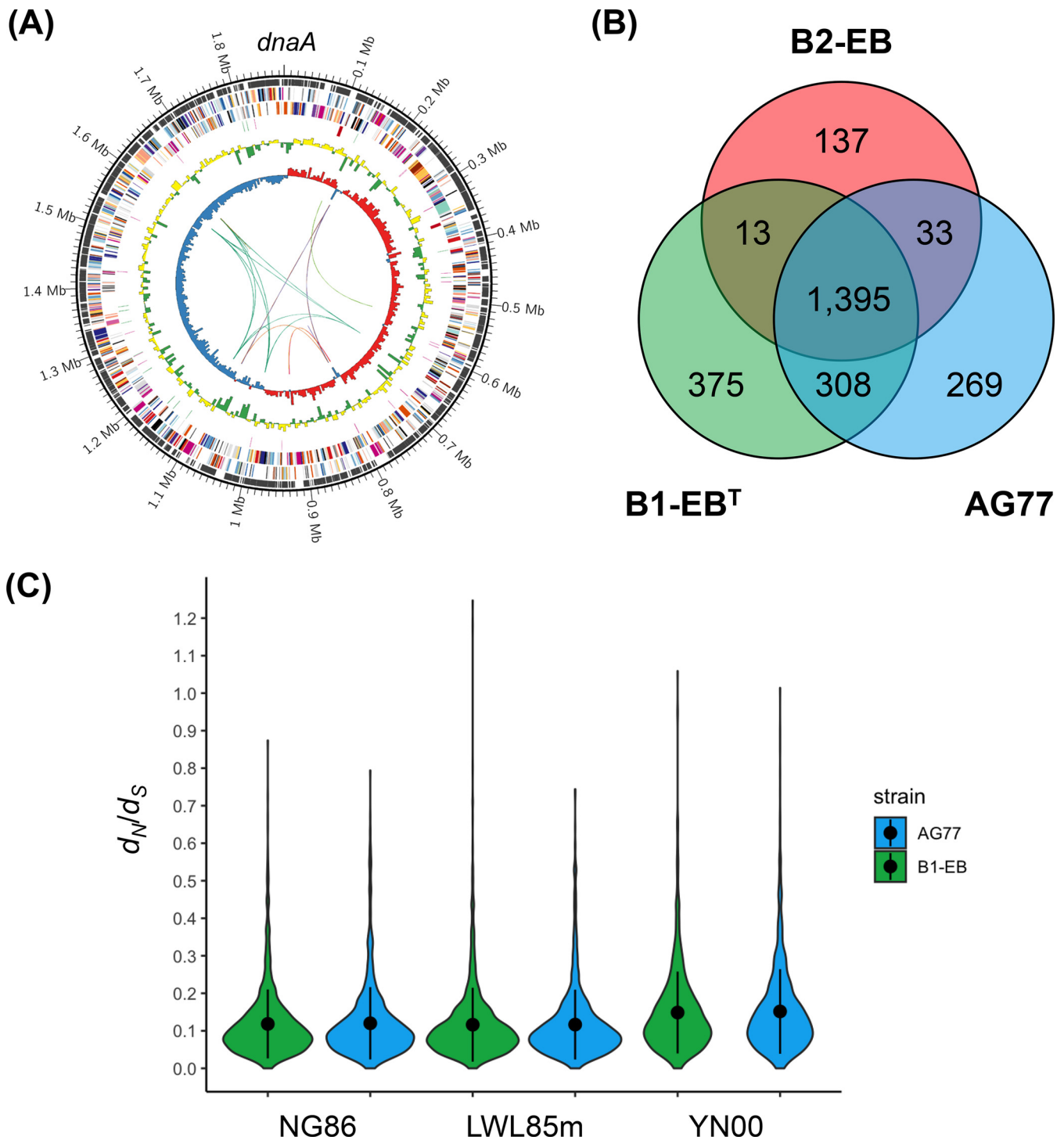
FIG 4 (A) Complete genome of *Mycoavidus* sp. B2-EB. Circles are labeled from the inside outwards: ring 1, GC skew with a 5-kb window; ring 2, GC content with a 5-kb window; ring 3, rRNA and tRNA; rings 4 and 5, predicted CDSs transcribed in counterclockwise and clockwise directions, respectively; ring 6, single-copy genes; ring 7, scale in Mb. The links in the center show the transposon positions, with different colors according to the transposon type. (B) Venn diagram of shared and unique orthologous genes among the three *Mycoavidus* genomes. (C) Violin plots show the distribution of d_N/d_S values estimated for individual genes using the NG, LWL, and YN methods. Black dots in the violin plots show the mean d_N/d_S values with standard deviation bars.

Mortierella species. Furthermore, a high degree of genomic rearrangement was identified in the 0.80- to 1.13-Mb downstream region of *dnaA*, encoding the replication initiator, showing "symmetric inversions" with endpoints that were equally distant from the origin of chromosomal replication (Fig. 5A and B). Genomic rearrangement likely
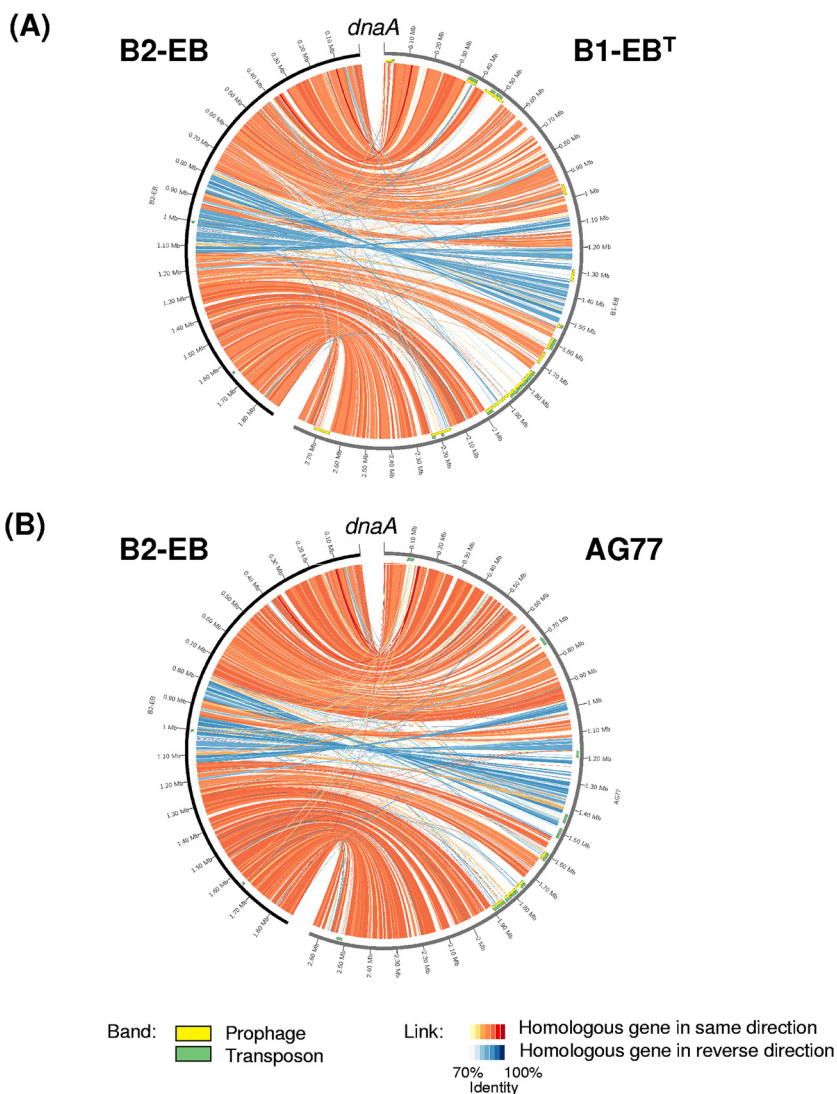
FIG 5 Syntenic comparative genomic analysis of *Mycoavidus* sp. B2-EB with *M. cysteinexigens* B1-EB[T] (A) and AG77 (B). Green and yellow bands next to the scale show the positions of transposons and prophages, respectively. Red and blue links associated with the two genomes show homologous genes in same and reverse directions, respectively. The gradient color of the link shows the nucleotide identity between the homologous genes.

impacts gene expression and results in the loss of gene function when a breakpoint occurs inside a reading frame (39, 40), which not only affects the organismal phenotype but also contributes to phylotype subdivision (41). It has been reported that extensive genome rearrangements probably triggered or enhanced the *Wolbachia* speciation event, because it could suppress homologous recombination in chromosomal regions that are not colinear (42). Here, we showed that genomic arrangements occurred between the genomes of *Mycoavidus cysteinexigens* and *Mycoavidus* sp. B2-EB (the two closest species to one another), suggesting that they may play an important role in the *Mycoavidus* speciation process.

Additionally, in spite of the high homology with the previously determined *M. cysteinexigens* genomes, the B2-EB genome also possessed 137 unique CDSs (Fig. 4B), including 89 and 48 CDSs encoding hypothetical and functional proteins, respectively. Most of the CDSs encoding functional proteins were involved in interactions with the host, mainly including 12 CDSs encoding lipopolysaccharide biosynthesis and secretion enzymes, 4 CDSs encoding secondary metabolite biosynthesis enzymes, and 4 CDSs
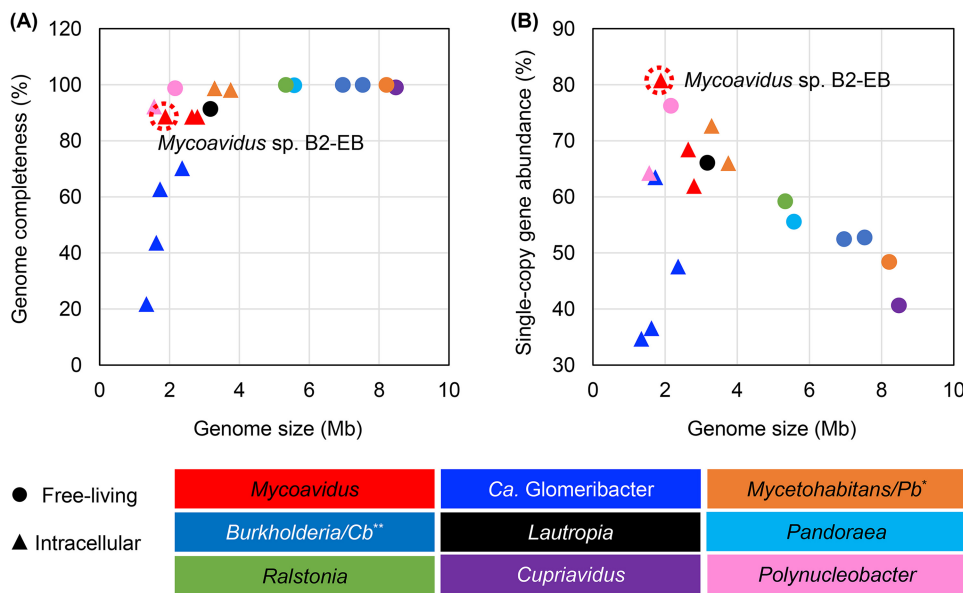
FIG 6 Genome completeness (A) and single-copy gene abundance (B) of *Burkholderiaceae* bacteria. The genome completeness was estimated by the conservation rate of 568 single-copy genes in the complete genomes of *Mycoavidus* sp. B2-EB, *Mycoavidus cysteinexigens* B1-EB[T] and AG77, *Mycetohabitans rhizoxinica* HKI 454[T], *Paraburkholderia phytofirmans* PsJN[T], *Burkholderia cepacia* ATCC 25416[T], *Caballeronia insecticola* RPE64[T], *Lautropia mirabilis* NTCT12852[T], *Pandoraea apista* DSM 16535[T], *Ralstonia pickettii* 12J[T], *Cupriavidus necator* N1[T], *Polynucleobacter asymbioticus* QLWP1DMWA1[T], and *Polynucleobacter necessarius* STIR1[T], as well as in the incomplete genomes of "*Ca*. Glomeribacter gigasporarum" BEG34, BEG1, JA201A, and IN211 and *Mycetohabitans endofungorum* HKI 456[T]. *, *Pb*, *Paraburkholderia*. **, *Cb*, *Caballeronia*.

encoding toxins or antitoxins (see Table S1 in the supplemental material). Such genes have been reported for the *Burkholderiaceae*-related endofungal bacteria in previous studies (22, 30, 43). The other functional CDSs specific to strain B2-EB were related to DNA replication and repair, nutrient metabolism, cell elongation and division, ribosome biogenesis, membrane transport, and signal transduction (Table S1). Notably, two of the unique CDSs (locus tags MPB2EB_0909 and MPB2EB_1044) were likely to be associated with bacterial survival inside the fungal cell, annotated as putative genes encoding β-lactamase and E3 ubiquitin-protein ligase, respectively. β-Lactamase can provide resistance to β-lactam antibiotics produced by fungi, such as penicillins (44), whereas the E3 ligases of some pathogenic bacteria mimic the activity of eukaryotic E3 ubiquitin ligases to benefit themselves (45). However, the roles of the two genes in the endofungal lifestyle should be examined using molecular biological approaches in the future.

The genomic features required by cultivable BRE. The smallest genomes possessed by *Buchnera aphidicola*, an obligate intracellular endosymbiont in aphids, are shaped by the evolutionary force of genetic drift, in which slightly deleterious mutations fix rapidly, causing the loss of gene functions, even including DNA repair, which further accelerates the accumulation of mutations (8). Notably, the proportion of transposable element accounted for only 2.4% of the total genome length in the B2-EB genome, versus 7.2 to 11.5% in known *Mycoavidus* genomes. In contrast, the genome reduction of heritable endofungal bacteria is caused by limited recombination in a largely clonal population, which is well supported by genomic information on "*Candidatus* Glomeribacter" (10). In this study, we estimated the genome completeness across the family *Burkholderiaceae* using the taxonomic-specific workflow of CheckM (46). Surprisingly, the results indicated that *Mycoavidus* endofungal bacteria dwelling in *Mortierella* possessed 88.6% of the checked single-copy genes (Fig. 6A), which was obviously more than those conserved in "*Ca*. Glomeribacter gigasporarum" genomes (21.8 to 70.2%). Moreover, the full gene set of DNA repair mechanisms, including base excision repair, nucleotide excision repair, mismatch repair, and homologous recom-

bination, were well conserved in the genomes of *Mycoavidus* strains B1-EB^T and B2-EB and may be responsible for purging slightly deleterious mutations from the genomes. These findings suggested that nondegenerative genome reduction occurred during the evolution of *Mycoavidus*-related endofungal bacteria adapting to the intracellular environments of fungal hosts.

The second surprising finding was that strain B2-EB represented the most stream-lined genomic architecture in the *Burkholderiaceae* bacteria with sequenced genomes, even beyond the streamlined genomes of *Polynucleobacter* species (47–49). The B2-EB genome was 1,877 kb with 1,627 CDSs and featured the highest proportion of single-copy genes in the genome (Fig. 6B). Currently, the free-living bacterium with the smallest genome (1,304 kb) is an uncultured but abundant ocean methylotroph of the *Betaproteobacteria*, HTCC2181, with 1,377 CDSs (50), followed by the hyperthermophilic bacterium *Aquifex aeolicus* with a slightly larger genome (1,591 kb and 1,613 genes) (51). The smallest photosynthetic cell, represented by the oceanic cyanobacterium *Prochlorococcus marinus*, has a 1,660-kb genome containing 1,765 genes (52). In the family *Burkholderiaceae*, the smallest genome is represented by *Polynucleobacter necessarius* endosymbionts isolated from ciliates living in aquatic ecosystems (53), with 1,560 kb and 1,279 CDSs (47). However, the single-copy genes in this genome accounted for 64.3% of the total genes (Fig. 6B), which was lower than the proportion in the B2-EB genome. Strain B2-EB has a small genome comparable to the minimal genomes of the above-mentioned free-living bacteria and possesses a radically stream-lined structure. A reduction in genome size has previously been identified in the *M. cysteinexigens* genomes (22, 30), and our analysis here showed that the proportions of single-copy genes were 61.9% and 68.5% in the B1-EB^T and AG77 genomes, respectively (Fig. 6B). Further comparisons with their free-living relatives (proportion of single-copy genes, 48.4% to 52.8%) (Fig. 6B) suggested that all three of the *Mycoavidus* genomes were streamlined, but the most streamlined was the B2-EB genome. There-fore, the genetic elements conserved in the B2-EB genome may represent the minimal genomic features required by a cultivable endofungal bacterium.

Compared with those of close relatives in another fungal hosts and free-living relatives in various habitats, the genomic features of *Mycoavidus* were characterized by the harbor-ing of more abundant genes responsible for functions including intracellular trafficking, secretion, vesicular transport, RNA processing and modification, and recombination (Fig. 7). Furthermore, the streamlined genome of strain B2-EB harbored almost complete sets responsible for encoding the 30S and 50S ribosomal proteins, tRNA aminoacylation, rRNA and tRNA modifications, DNA replication and repair, transcription, translation, and protein folding (see Tables S2 to S9 in the supplemental material). Although some genes involved in amino acid metabolism (e.g., *asnA*, *asnB*, *gltB*, and *hutH*) and secretion systems (e.g., *yscF*, *virB5*, *gspC*, *gspI*, *gspL*, *gspM*, *vgrG*, *hcp*, *impL*, *impK*, and *vasG*) were absent in the B2-EB genome (see Fig. S4 in the supplemental material), it was similar to the genomes of other BRE that possessed many of these genes (6, 22, 29, 30), suggesting that they have an important role in the endofungal lifestyle. In addition, a large number of genes involved in fatty acid metabolisms were present in the B2-EB genome, indicating that fatty acids are likely to be a major carbon source for *Mycoavidus* endofungal bacteria (Fig. S4). Together with other essential genes for life, such as those involved in tricarboxylic acid (TCA) cycling, energy generation, and various transport systems, the genetic elements harbored by strain B2-EB showed the minimal genomic features required by both an endofungal lifestyle and artificial culture.

Conclusion. BRE have been detected in several fungal lineages, including *Gigaspora*, *Racocetra*, and *Cetraspora* (Glomeromycotina), *Rhizopus* (Mucoromycotina), and *Mortierella* (Mortierellomycotina) (10, 54). The sequenced BRE genomes have revealed common features such as reduced genome size, loss of genes encoding primary metabolism, and specialization in fungal metabolite uptake. In particular, the BRE affiliated with the *Mycoavidus-Glomeribacter* clade were elucidated as having amino acid auxotrophies that have limited their artificial cultivation in pure cultures. Our
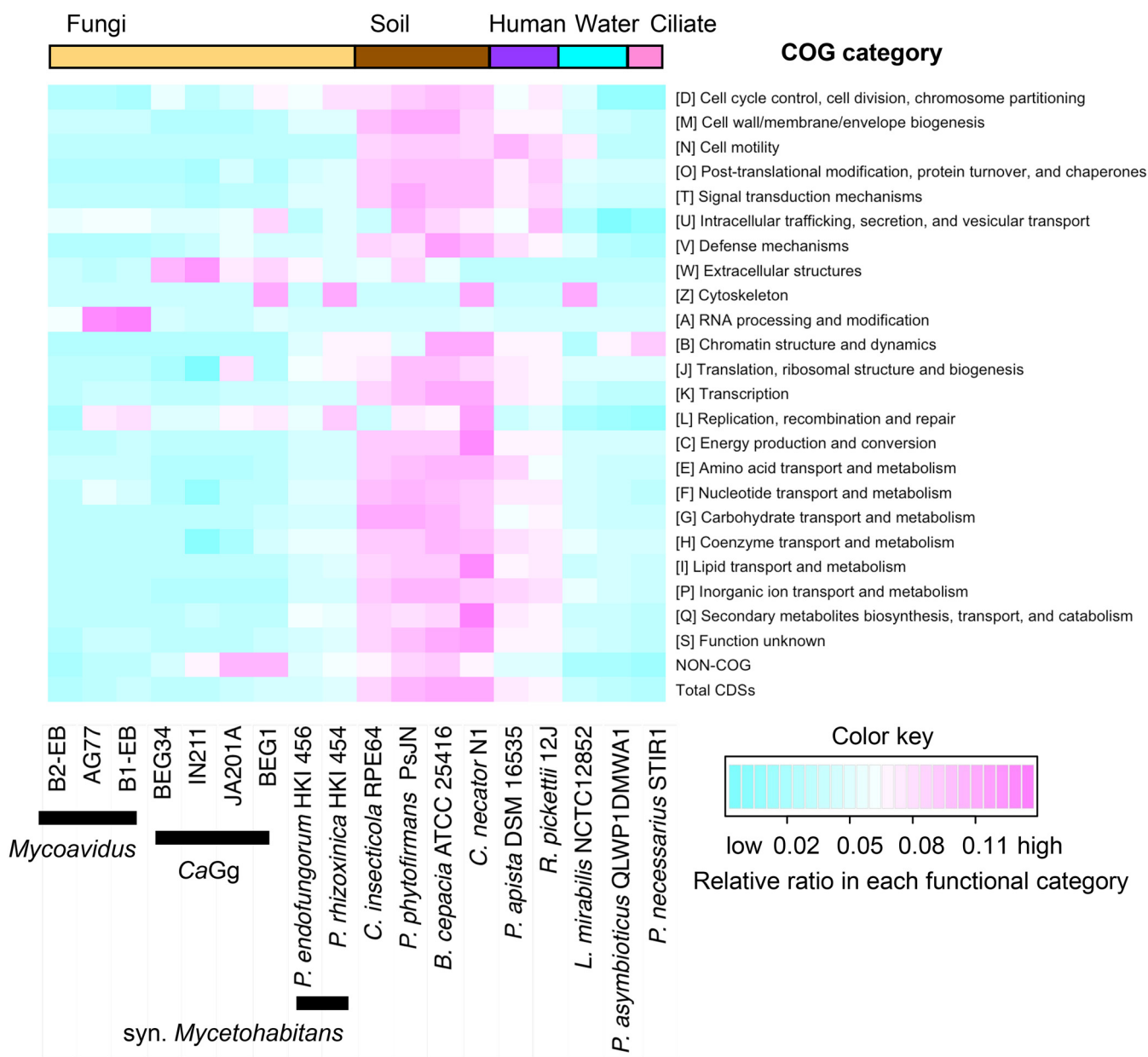
FIG 7 Heat map analysis of the genomic profiles of *Burkholderiaceae* bacteria based on the COG category.

previous efforts in reverse genomics have contributed to the isolation of *Mycoavidus cysteinexigens* from the host fungus *Mortierella elongata* by supplying exogenous cysteine to substitute for the genetic defect in bacterial cysteine biosynthesis (19, 26).

In this study, we successfully isolated an endobacterium phylogenetically close to *M. cysteinexigens*, named *Mycoavidus* sp. B2-EB, from the host fungus *Mortierella parvispora*. Moreover, the hybrid assembly of deep-sequencing data sets reconstructed a complete genome sequence of strain B2-EB, revealing a reduced genome size (1.88 Mb), which was less than those of *M. cysteinexigens* strains (2.64 to 2.79 Mb) and comparable to those of "*Ca.* Glomeribacter gigasporarum" genomes (1.34 to 2.36 Mb). Further assessment of genomic variation on a population scale indicated that *Mycoavidus* sp. B2-EB possessed a nondegenerative genome comparable to that found in "*Ca.* Glomeribacter gigasporarum" endosymbionts. In addition, comparative genomic analysis with other BRE and free-living relatives affiliated with the family *Burkholderiaceae* showed that the B2-EB genome had a relatively high genome completeness and high

proportion of single-copy genes but only a few transposable genetic elements. These results suggested that the B2-EB genome reached a state of relative evolutionary stability. The genetic elements conserved in the stabilized genome appear to have resulted from the adaption of endobacteria to the host metabolism. In our previous study (30) and this study, we found no genes involved in invasion or lysis of fungal cells in the *Mycoavidus* genome according to comparative genomic analysis between *Mycoavidus* and *Mycetohabitans*, indicating that strain B2-EB has already lost a way to get in and out of fungal cells. To the best of our knowledge, the B2-EB genome represents the smallest genome size among cultivable endofungal bacteria. Our findings are expected to fill a knowledge gap in the understanding of genome evolution of *Mycoavidus*-related endofungal bacteria, which results in a minimal genome of a cultivable endosymbiont.

The information gleaned from this genome revealed the minimal genome features required by both an endofungal lifestyle and artificial cultivation, which furthers our understanding of genome reduction in fungal endosymbionts and extends the culture resources for the biotechnological development of engineering synthetic microbiomes.

MATERIALS AND METHODS

Ultramicroscopic observation of *Burkholderiaceae*-related endobacteria living in *Mortierella parvispora*. Transmission electron microscopy (TEM) with sample preparation using rapid freezing and a freeze-substitution method was performed to observe the endohyphal features in the fungus *Mortierella parvispora* E1425 (JCM39028) at Tokai Electron Microscopy Inc., Nagoya, Japan. In brief, 7-day-old mycelia grown on sterilized cellophane sheets placed onto glycerol agar (containing, per liter of distilled water, 10 g glycerol, 0.5 g Casamino Acids, 2 g $NaNO_3$, 1 g KH_2PO_4, 0.5 g KCl, 0.5 g $MgSO_4\cdot7H_2O$, 0.01 g $FeSO_4$, and 15 g agar) at 23°C were sandwiched with copper disks and frozen in liquid propane at −175°C, followed by substitution with 2% osmium tetroxide in acetone and 2% distilled water at −80°C for 48 h, −20°C for 4 h, and 4°C for 2 h, dehydrating through anhydrous acetone twice for 30 min each and 3 changes of 100% ethanol for 30 min each at room temperature (RT), infiltrating with propylene oxide (PO) twice for 30 min each and a PO-resin mixture (7:3) for 1 h, and polymerizing in fresh 100% resin at 60°C for 48 h. The polymerized sample was ultrathin sectioned at 90 nm using an Ultracut UCT ultramicrotome (Leica, Vienna, Austria) and stained with 2% uranyl acetate at RT for 15 min and with lead stain solution (Sigma-Aldrich, Tokyo, Japan) at RT for 3 min. The grids were observed using a JEM-1400Plus transmission electron microscope with an EM-14830RUBY2 charge-coupled device (CCD) camera (JEOL, Tokyo, Japan) at an acceleration voltage of 100 kV.

Isolation of an endobacterium dwelling in *M. parvispora*. Endohyphal bacteria living in the fungus *M. parvispora* E1425 were isolated using the method reported previously (20), with a minor modification. In brief, fungal strain E1425 was cultivated for 7 days at 23°C on half-strength CMMY agar. The cultivated mycelia were homogenized using a sterilized pestle in a 1.5-ml microcentrifuge tube at maximum rotation speed for 2 min and centrifuged at 2,000 $\times$ g for 10 min. The supernatant was filtered through 8-μm- and 3-μm-pore-size membrane filters to remove fragmented hyphae and sporangiospores. One-milliliter aliquots of the filtered suspension were spread on a B-CYEα agar plate [containing, per liter of distilled water, 10 g yeast extract, 2 g charcoal powder, 0.4 g L-cysteine hydrochloride, 0.25 g ferric pyrophosphate (soluble), 10 g *N*-(2-acetamido)-2-aminoethanesulfonic acid (ACES), 1 g potassium 2-oxoglutarate, and 15 g agar, pH 6.9] (Eiken Chemical, Tokyo, Japan) and incubated for 30 days at 23°C. When the bacterial colonies grew on the plate, a single colony from the plate was subcultured repeatedly on a B-CYEα plate in order to assess its purity.

Genome sequencing, assembly, annotation, and comparative genomic analysis. Genomic DNA of endobacterial isolates grown on a cysteine-containing buffered charcoal-yeast extract agar (B-CYEα) plate was extracted using a modified lysozyme buffer method, as in our previous study (30). The obtained DNA (39.2 μg in total) was used to build a pair-end library with an insert size of ~550 bp and a 20-kb SMRT library and sequenced on the Illumina HiSeq2500 and PacBio RS II platforms, respectively, at GeneBay, Yokohama, Japan. Sequence processing and hybrid assembly were performed in accordance with our previous strategy (55). Briefly, the single-end reads were trimmed out from the HiSeq paired-end reads using sickle v1.33 with default settings. The qualified paired-end reads and PacBio long reads were hybrid-assembled using SPAdes v3.10.1 with the pipeline option of −careful. Genome polishing and finishing were performed using pilon v1.22 and GenoFinisher v2.1 with the BWA-MEM v0.7.12 algorithm.

The genome sequence was basically annotated using the DDBJ Fast Annotation and Submission Tool (DFAST v1.1.0) pipeline (56), followed by determination of the metabolic pathways using the KEGG online services BlastKOALA v2.1 (57) and KEGG mapper v3.2, identifying the COG function category using EggNOG-mapper v2.0 (58), and searching the secondary metabolic gene clusters using antiSMASH v4.0 (59). Prophage sequences and transposons were identified using the PHASTER web server (60) and the previously reported method (30), respectively. The complete genome sequences of the type species for each genus in the family *Burkholderiaceae* were obtained from the GenBank database to perform the comparative genomic analysis. Orthologous clusters across the selected genomes were grouped using OrthoFinder v2.3.3 (61), and the COG profiles were characterized using the EggNOG-mapper. Whole-genome phylogenetic analysis was performed using ezTree v0.1 (62) with the complete genome

sequences of type strains for the species in the family *Burkholderiaceae*. Genome completeness was estimated using CheckM v1.0.13 with the option of taxonomy_wf to calculate the conservation rate of the 568 single-copy housekeeping genes found in the genomes of free-living bacteria in the family *Burkholderiaceae* (46). To estimate the *dN*/*dS* ratio across the *Mycoavidus* genomes, the orthologous groups detected between the B2-EB and *M. cysteinexigens* genomes (B1-EB^T and AG77), three different methods were used, NG86 (63), LWL85m (64), and YN00 (65). The R package orthologr was used to perform the d_N/d_S calculation (66). In addition, these orthologous groups were sorted based on the physical locations and directions in the genomes and used to estimate the genome rearrangement by syntenic comparison.

Circos v0.69-6 (67) was used to generate a schematic spherical illustration of the complete B2-EB genome sequence and visualize the syntenic comparative analysis between the B2-EB and *M. cysteinexigens* genomes. The R package ggplot2 was used to generate a heat map showing the genome profiles of representative *Burkholderiaceae* genomes and violin plots of the distribution of the d_N/d_S values and the amino acid identity of individual genes in the *Mycoavidus* sp. B2-EB genome compared with those in the *M. cysteinexigens* B1-EB^T and AG77 genomes. The metabolic features of B2-EB were determined using the KEGG mapper with the option of "Reconstruct Pathway" (https://www.genome.jp/kegg/tool/map_pathway.html) and visualized using Microsoft PowerPoint.

Data availability. The complete genome sequence of *Mycoavidus* sp. B2-EB has been deposited in DDBJ/ENA/GenBank under the accession number AP021872. The raw sequence reads for the Illumina paired-end and PacBio RS II libraries are available under accession number DRA009105.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, PDF file, 0.4 MB.

SUPPLEMENTAL FILE 2, XLSX file, 0.03 MB.

ACKNOWLEDGMENTS

REFERENCES

1. Munson MA, Baumann P, Clark MA, Baumann L, Moran NA, Voegtlin DJ, Campbell BC. 1991. Evidence for the establishment of aphid-eubacterium endosymbiosis in an ancestor of four aphid families. J Bacteriol 173:6321–6324. https://doi.org/10.1128/jb.173.20.6321-6324.1991.

2. Nowack ECM, Melkonian M. 2010. Endosymbiotic associations within protists. Philos Trans R Soc Lond B Biol Sci 365:699–712. https://doi.org/10.1098/rstb.2009.0188.

3. Kawafune K, Hongoh Y, Hamaji T, Nozaki H. 2012. Molecular identification of rickettsial endosymbionts in the non-phagotrophic volvocalean green algae. PLoS One 7:e31749. https://doi.org/10.1371/journal.pone.0031749.

4. Takashima Y, Seto K, Degawa Y, Guo Y, Nishizawa T, Ohta H, Narisawa K. 2018. Prevalence and intra-family phylogenetic divergence of *Burkholderiaceae*-related endobacteria associated with species of *Mortierella*. Microbes Environ 33:417–427. https://doi.org/10.1264/jsme2.ME18081.

5. Shigenobu S, Watanabe H, Hattori M, Sakaki Y, Ishikawa H. 2000. Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. APS. Nature 407:81–86. https://doi.org/10.1038/35024074.

6. Ghignone S, Salvioli A, Anca I, Lumini E, Ortu G, Petiti L, Cruveiller S, Bianciotto V, Piffanelli P, Lanfranco L, Bonfante P. 2012. The genome of the obligate endobacterium of an AM fungus reveals an interphylum network of nutritional interactions. ISME J 6:136–145. https://doi.org/10.1038/ismej.2011.110.

7. Gray MW. 1999. Evolution of organellar genomes. Curr Opin Genet Dev 9:678–687. https://doi.org/10.1016/s0959-437x(99)00030-1.

8. McCutcheon JP, Moran NA. 2011. Extreme genome reduction in symbiotic bacteria. Nat Rev Microbiol 10:13–26. https://doi.org/10.1038/nrmicro2670.

9. Akman L, Yamashita A, Watanabe H, Oshima K, Shiba T, Hattori M, Aksoy S. 2002. Genome sequence of the endocellular obligate symbiont of tsetse flies, *Wigglesworthia glossinidia*. Nat Genet 32:402–407. https://doi.org/10.1038/ng986.

10. Mondo SJ, Salvioli A, Bonfante P, Morton JB, Pawlowska TE. 2016. Nondegenerative evolution in ancient heritable bacterial endosymbionts of fungi. Mol Biol Evol 33:2216–2231. https://doi.org/10.1093/molbev/msw086.

11. Arendt KR, Hockett KL, Araldi-Brondolo SJ, Baltrus DA, Arnold AE. 2016. Isolation of endohyphal bacteria from foliar Ascomycota and *in vitro* establishment of their symbiotic associations. Appl Environ Microbiol 82:2943–2949. https://doi.org/10.1128/AEM.00452-16.

12. Almeida C, Pereira CS, Gonzalez-Menendez V, Bills G, Pascual J, Sánchez-Hidalgo M, Kehraus S, Genilloud O. 2018. Unveiling concealed functions of endosymbiotic bacteria harbored in the ascomycete *Stachylidium bicolor*. Appl Environ Microbiol 84:e00660-18. https://doi.org/10.1128/AEM.00660-18.

13. Wang R, Dong L, Chen Y, Qu L, Wang Q, Zhang Y. 2017. *Esteya vermicola*, a nematophagous fungus attacking the pine wood nematode, harbors a bacterial endosymbiont affiliated with *Gammaproteobacteria*. Microbes Environ 32:201–209. https://doi.org/10.1264/jsme2.ME16167.

14. Sharma M, Schmid M, Rothballer M, Hause G, Zuccaro A, Imani J, Kämpfer P, Domann E, Schäfer P, Hartmann A, Kogel KH. 2008. Detection and identification of bacteria intimately associated with fungi of the order Sebacinales. Cell Microbiol 10:2235–2246. https://doi.org/10.1111/j.1462-5822.2008.01202.x.

15. Sato Y, Narisawa K, Tsuruta K, Umezu M, Nishizawa T, Tanaka K, Yamaguchi K, Komatsuzaki M, Ohta H. 2010. Detection of *Betaproteobacteria* inside the mycelium of the fungus *Mortierella elongata*. Microbes Environ 25:321–324. https://doi.org/10.1264/jsme2.me10134.

16. Bonfante P, Desirò A. 2017. Who lives in a fungus? The diversity, origins

and functions of fungal endobacteria living in Mucoromycota. ISME J 11:1727–1735. https://doi.org/10.1038/ismej.2017.21.

17. Bianciotto V, Lumini E, Bonfante P, Vandamme P. 2003. *Candidatus* Glomeribacter gigasporarum gen. nov., sp. nov., an endosymbiont of arbuscular mycorrhizal fungi. Int J Syst Evol Microbiol 53:121–124. https://doi.org/10.1099/ijs.0.02382-0.

18. Partida-Martinez LP, Groth I, Schmitt I, Richter W, Roth M, Hertweck C. 2007. *Burkholderia rhizoxinica* sp. nov. and *Burkholderia endofungorum* sp. nov., bacterial endosymbionts of the plant-pathogenic fungus *Rhizopus microsporous*. Int J Syst Evol Microbiol 57:2583–2590. https://doi.org/10.1099/ijs.0.64660-0.

19. Estrada-de los Santos P, Palmer M, Chávez-Ramírez B, Beukes C, Steenkamp ET, Briscoe L, Khan N, Maluk M, Lafos M, Humm E, Arrabit M, Crook M, Gross E, Simon MF, dos Reis Junior FB, Whitman WB, Shapiro N, Poole PS, Hirsch AM, Venter SN, James EK. 2018. Whole genome analyses suggests that Burkholderia sensu lato contains two additional novel genera (*Mycetohabitans* gen. nov., and *Trinickia* gen. nov.): implications for the evolution of diazotrophy and nodulation in the *Burkholderiaceae*. Genes 9:389. https://doi.org/10.3390/genes9080389.

20. Ohshima S, Sato Y, Fujimura R, Takashima Y, Hamada M, Nishizawa T, Narisawa K, Ohta H. 2016. *Mycoavidus cysteinexigens* gen. nov., sp. nov., an endohyphal bacterium isolated from a soil isolate of the fungus *Mortierella elongata*. Int J Syst Evol Microbiol 66:2052–2057. https://doi.org/10.1099/ijsem.0.000990.

21. Wagner L, Stielow B, Hoffmann K, Petkovits T, Papp T, Vágvölgyi C, de Hoog GS, Verkley G, Voigt K. 2013. A comprehensive molecular phylogeny of the *Mortierellales* (Mortierellomycotina) based on nuclear ribosomal DNA. Persoonia Mol Phylogeny Evol Fungi 30:77–93. https://doi.org/10.3767/003158513X666268.

22. Uehling J, Gryganskyi A, Hameed K, Tschaplinski T, Misztal PK, Wu S, Desirò A, Vande Pol N, Du Z, Zienkiewicz A, Zienkiewicz K, Morin E, Tisserant E, Splivallo R, Hainaut M, Henrissat B, Ohm R, Kuo A, Yan J, Lipzen A, Nolan M, Labutti K, Barry K, Goldstein AH, Labbé J, Schadt C, Tuskan G, Grigoriev I, Martin F, Vilgalys R, Bonito G. 2017. Comparative genomics of *Mortierella elongata* and its bacterial endosymbiont *Mycoavidus cysteinexigens*. Environ Microbiol 19:2964–2983. https://doi.org/10.1111/1462-2920.13669.

23. Desirò A, Hao Z, Liber JA, Benucci GMN, Lowry D, Roberson R, Bonito G. 2018. *Mycoplasma*-related endobacteria within Mortierellomycotina fungi: diversity, distribution and functional insights into their lifestyle. ISME J 12:1743–1757. https://doi.org/10.1038/s41396-018-0053-9.

24. Stewart EJ. 2012. Growing unculturable bacteria. J Bacteriol 194:4151–4160. https://doi.org/10.1128/JB.00345-12.

25. Jargeat P, Cosseau C, Ola'h B, Jauneau A, Bonfante P, Batut J, Bécard G. 2004. Isolation, free-living capacities, and genome structure of "*Candidatus* Glomeribacter gigasporarum," the endocellular bacterium of the mycorrhizal fungus *Gigaspora margarita*. J Bacteriol 186:6876–6884. https://doi.org/10.1128/JB.186.20.6876-6884.2004.

26. Fujimura R, Nishimura A, Ohshima S, Sato Y, Nishizawa T, Oshima K, Hattori M, Narisawa K, Ohta H. 2014. Draft genome sequence of the betaproteobacterial endosymbiont associated with the fungus *Mortierella elongata* FMR23-6. Genome Announc 2:e01272-14. https://doi.org/10.1128/genomeA.01272-14.

27. Dicenzo GC, Mengoni A, Perrin E. 2019. Chromids aid genome expansion and functional diversification in the family *Burkholderiaceae*. Mol Biol Evol 36:562–574. https://doi.org/10.1093/molbev/msy248.

28. Harrison PW, Lower RPJ, Kim NKD, Young J. 2010. Introducing the bacterial "chromid": not a chromosome, not a plasmid. Trends Microbiol 18:141–148. https://doi.org/10.1016/j.tim.2009.12.010.

29. Lackner G, Moebius N, Partida-Martinez L, Hertweck C. 2011. Complete genome sequence of *Burkholderia rhizoxinica*, an endosymbiont of *Rhizopus microsporus*. J Bacteriol 193:783–784. https://doi.org/10.1128/JB.01318-10.

30. Sharmin D, Guo Y, Nishizawa T, Ohshima S, Sato Y, Takashima Y, Narisawa K, Ohta H. 2018. Comparative genomic insights into endofungal lifestyles of two bacterial endosymbionts, *Mycoavidus cysteinexigens* and *Burkholderia rhizoxinica*. Microbes Environ 33:66–76. https://doi.org/10.1264/jsme2.ME17138.

31. Manzano-Marín A, Latorre A. 2016. Snapshots of a shrinking partner: genome reduction in *Serratia symbiotica*. Sci Rep 6:32590. https://doi.org/10.1038/srep32590.

32. Hendry TA, Freed LL, Fader D, Fenolio D, Sutton TT, Lopez JV. 2018. Ongoing transposon-mediated genome reduction in the luminous bacterial symbionts of deep-sea ceratioid anglerfishes. mBio 9:e01033-18. https://doi.org/10.1128/mBio.01033-18.

33. Konstantinidis KT, Tiedje JM. 2005. Genomic insights that advance the species definition for prokaryotes. Proc Natl Acad Sci U S A 102:2567–2572. https://doi.org/10.1073/pnas.0409727102.

34. Jarett JK, Nayfach S, Podar M, Inskeep W, Ivanova NN, Munson-McGee J, Schulz F, Young M, Jay ZJ, Beam JP, Kyrpides NC, Malmstrom RR, Stepanauskas R, Woyke T. 2018. Single-cell genomics of co-sorted *Nanoarchaeota* suggests novel putative host associations and diversification of proteins involved in symbiosis. Microbiome 6:161. https://doi.org/10.1186/s40168-018-0539-8.

35. Nakjang S, Williams TA, Heinz E, Watson AK, Foster PG, Sendra KM, Heaps SE, Hirt RP, Embley TM. 2013. Reduction and expansion in microsporidian genome reduction: new insights from comparative genomics. Genome Biol Evol 5:2285–2303. https://doi.org/10.1093/gbe/evt184.

36. Moran NA, McLaughlin HJ, Sorek R. 2009. The dynamics and time scale of ongoing genomic erosion in symbiotic bacteria. Science 323:379–382. https://doi.org/10.1126/science.1167140.

37. Ffrench-Constant RH, Bowen DJ. 2000. Novel insecticidal toxins from nematode-symbiotic bacteria. Cell Mol Life Sci 57:828–833. https://doi.org/10.1007/s000180050044.

38. Newton ILG, Clark ME, Kent BN, Bordenstein SR, Qu J, Richards S, Kelkar YD, Werren JH. 2016. Comparative genomics of two closely related *Wolbachia* with different reproductive effects on hosts. Genome Biol Evol 8:1526–1542. https://doi.org/10.1093/gbe/evw096.

39. Sousa C, De Lorenzo V, Cebolla A. 1997. Modulation of gene expression through chromosomal positioning in *Escherichia coli*. Microbiology 143:2071–2078. https://doi.org/10.1099/00221287-143-6-2071.

40. Couturier E, Rocha E. 2006. Replication-associated gene dosage effects shape the genomes of fast-growing bacteria but only for transcription and translation genes. Mol Microbiol 59:1506–1518. https://doi.org/10.1111/j.1365-2958.2006.05046.x.

41. Darling AE, Miklós I, Ragan MA. 2008. Dynamics of genome rearrangement in bacterial populations. PLoS Genet 4:e1000128. https://doi.org/10.1371/journal.pgen.1000128.

42. Ellegaard KM, Klasson L, Näslund K, Bourtzis K, Andersson S. 2013. Comparative genomics of *Wolbachia* and the bacterial species concept. PLoS Genet 9:e1003381. https://doi.org/10.1371/journal.pgen.1003381.

43. Salvioli di Fossalunga A, Lipuma J, Venice F, Dupont L, Bonfante P. 2017. The endobacterium of an arbuscular mycorrhizal fungus modulates the expression of its toxin-antitoxin systems during the life cycle of its host. ISME J 11:2394–2398. https://doi.org/10.1038/ismej.2017.84.

44. Abraham EP, Chain E. 1940. An enzyme from bacteria able to destroy penicillin. Nature 146:837–837. https://doi.org/10.1038/146837a0.

45. Hicks SW, Galán JE. 2010. Hijacking the host ubiquitin pathway: structural strategies of bacterial E3 ubiquitin ligases. Curr Opin Microbiol 13:41–46. https://doi.org/10.1016/j.mib.2009.11.008.

46. 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.

47. Boscaro V, Felletti M, Vannini C, Ackerman MS, Chain PSG, Malfatti S, Vergez LM, Shin M, Doak TG, Lynch M, Petroni G. 2013. *Polynucleobacter necessarius*, a model for genome reduction in both free-living and symbiotic bacteria. Proc Natl Acad Sci U S A 110:18590–18595. https://doi.org/10.1073/pnas.1316687110.

48. Meincke L, Copeland A, Lapidus A, Lucas S, Berry KW, del Rio TG, Hammon N, Dalin E, Tice H, Pitluck S, Richardson P, Bruce D, Goodwin L, Han C, Tapia R, Detter JC, Schmutz J, Brettin T, Larimer F, Land M, Hauser L, Kyrpides NC, Ivanova N, Göker M, Woyke T, Wu QL, Pöckl M, Hahn MW, Klenk HP. 2012. Complete genome sequence of *Polynucleobacter necessarius* subsp. *asymbioticus* type strain (QLW-P1DMWA-1 T.). Stand Genomic Sci 6:74–83. https://doi.org/10.4056/sigs.2395367.

49. Hoetzinger M, Schmidt J, Jezberová J, Koll U, Hahn MW. 2017. Microdiversification of a pelagic *Polynucleobacter* species is mainly driven by acquisition of genomic islands from a partially interspecific gene pool. Appl Environ Microbiol 83:e02266-16. https://doi.org/10.1128/AEM.02266-16.

50. Giovannoni SJ, Hayakawa DH, Tripp HJ, Stingl U, Givan SA, Cho JC, Oh HM, Kitner JB, Vergin KL, Rappé MS. 2008. The small genome of an abundant coastal ocean methylotroph. Environ Microbiol 10:1771–1782. https://doi.org/10.1111/j.1462-2920.2008.01598.x.

51. Deckert G, Warren PV, Gaasterland T, Young WG, Lenox AL, Graham DE, Overbeek R, Snead MA, Keller M, Aujay M, Huber R, Feldman RA, Short

JM, Olsen GJ, Swanson RV. 1998. The complete genome of the hyper-thermophilic bacterium *Aquifex aeolicus*. Nature 392:353–358. https://doi.org/10.1038/32831.

52. Rocap G, Larimer FW, Lamerdin J, Malfatti S, Chain P, Ahlgren NA, Arellano A, Coleman M, Hauser L, Hess WR, Johnson ZI, Land M, Lindell D, Post AF, Regala W, Shah M, Shaw SL, Steglich C, Sullivan MB, Ting CS, Tolonen A, Webb EA, Zinser ER, Chisholm SW. 2003. Genome divergence in two *Prochlorococcus* ecotypes reflects oceanic niche differentiation. Nature 424:1042–1047. https://doi.org/10.1038/nature01947.

53. Hahn MW. 2003. Isolation of strains belonging to the cosmopolitan *Polynucleobacter necessarius* cluster from freshwater habitats located in three climatic zones. Appl Environ Microbiol 69:5248–5254. https://doi.org/10.1128/AEM.69.9.5248-5254.2003.

54. Bonfante P, Venice F, Lanfranco L. 2019. The mycobiota: fungi take their place between plants and bacteria. Curr Opin Microbiol 49:18–25. https://doi.org/10.1016/j.mib.2019.08.004.

55. Guo Y, Matsuoka Y, Miura T, Nishizawa T, Ohta H, Narisawa K. 2018. Complete genome sequence of *Agrobacterium pusense* VsBac-Y9, a bacterial symbiont of the dark septate endophytic fungus *Veronaeopsis simplex* Y34with potential for improving fungal colonization in roots. J Biotechnol 284:31–36. https://doi.org/10.1016/j.jbiotec.2018.07.045.

56. Tanizawa Y, Fujisawa T, Nakamura Y. 2018. DFAST: a flexible prokaryotic genome annotation pipeline for faster genome publication. Bioinformatics 34:1037–1039. https://doi.org/10.1093/bioinformatics/btx713.

57. Kanehisa M, Sato Y, Morishima K. 2016. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. J Mol Biol 428:726–731. https://doi.org/10.1016/j.jmb.2015.11.006.

58. Huerta-Cepas J, Forslund K, Coelho LP, Szklarczyk D, Jensen LJ, Von Mering C, Bork P. 2017. Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. Mol Biol Evol 34:2115–2122. https://doi.org/10.1093/molbev/msx148.

59. Weber T, Blin K, Duddela S, Krug D, Kim HU, Bruccoleri R, Lee SY, Fischbach MA, Müller R, Wohlleben W, Breitling R, Takano E, Medema MH. 2015. AntiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. Nucleic Acids Res 43:W237–W243. https://doi.org/10.1093/nar/gkv437.

60. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. 2016. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res 44:W16–W21. https://doi.org/10.1093/nar/gkw387.

61. Emms DM, Kelly S. 2015. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. Genome Biol 16:157. https://doi.org/10.1186/s13059-015-0721-2.

62. Wu YW. 2018. ezTree: an automated pipeline for identifying phylogenetic marker genes and inferring evolutionary relationships among uncultivated prokaryotic draft genomes. BMC Genomics 19:921. https://doi.org/10.1186/s12864-017-4327-9.

63. Nei M, Gojobori T. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol Biol Evol 3:418–426. https://doi.org/10.1093/oxfordjournals.molbev.a040410.

64. Tzeng YH, Pan R, Li WH. 2004. Comparison of three methods for estimating rates of synonymous and nonsynonymous nucleotide substitutions. Mol Biol Evol 21:2290–2298. https://doi.org/10.1093/molbev/msh242.

65. Yang ZH, Nielsen R. 2000. Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. Mol Biol Evol 17:32–43. https://doi.org/10.1093/oxfordjournals.molbev.a026236.

66. Drost HG, Gabel A, Grosse I, Quint M. 2015. Evidence for active maintenance of phylotranscriptomic hourglass patterns in animal and plant embryogenesis. Mol Biol Evol 32:1221–1231. https://doi.org/10.1093/molbev/msv012.

67. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA. 2009. Circos: an information aesthetic for comparative genomics. Genome Res 19:1639–1645. https://doi.org/10.1101/gr.092759.109.