

## Comparative Genomic Insights into Endofungal Lifestyles of Two Bacterial Endosymbionts, *Mycoavidus cysteinexigens* and *Burkholderia rhizoxinica*

DILRUBA SHARMIN<sup>1†</sup>, YONG GUO<sup>1†</sup>, TOMOYASU NISHIZAWA<sup>1,2\*</sup>, SHOKO OHSHIMA<sup>1</sup>, YOSHINORI SATO<sup>3</sup>, YUSUKE TAKASHIMA<sup>2</sup>, KAZUHIKO NARISAWA<sup>1,2</sup>, and HIROYUKI OHTA<sup>1,2</sup>

<sup>1</sup>Ibaraki University College of Agriculture, Department of Bioresource Science, Ibaraki 300–0393, Japan; <sup>2</sup>United Graduate School of Agricultural Science, Tokyo University of Agriculture and Technology, Fuchu, Tokyo 183–8509, Japan; and <sup>3</sup>Center for Conservation and Restoration Techniques, Tokyo National Research Institute for Cultural Properties, Tokyo 110–8713, Japan

(Received September 3, 2017—Accepted December 19, 2017—Published online March 14, 2018)

Endohyphal bacteria (EHB), dwelling within fungal hyphae, markedly affect the growth and metabolic potential of their hosts. To date, two EHB belonging to the family *Burkholderiaceae* have been isolated and characterized as new taxa, *Burkholderia rhizoxinica* (HKI 454<sup>T</sup>) and *Mycoavidus cysteinexigens* (B1-EB<sup>T</sup>), in Japan. Metagenome sequencing was recently reported for *Mortierella elongata* AG77 together with its endosymbiont *M. cysteinexigens* (Mc-AG77) from a soil/litter sample in the USA. In the present study, we elucidated the complete genome sequence of B1-EB<sup>T</sup> and compared it with those of Mc-AG77 and HKI 454<sup>T</sup>. The genomes of B1-EB<sup>T</sup> and Mc-AG77 contained a higher level of prophage sequences and were markedly smaller than that of HKI 454<sup>T</sup>. Although the B1-EB<sup>T</sup> and Mc-AG77 genomes lacked the chitinolytic enzyme genes responsible for invasion into fungal cells, they contained several predicted toxin-antitoxin systems including an insecticidal toxin complex and PIN domain imposing an addiction-like mechanism essential for endohyphal growth control during host colonization. Despite the different host fungi, the alignment of amino acid sequences showed that the HKI 454<sup>T</sup> genome consisted of 1,265 (32.6%) and 1,221 (31.5%) orthologous coding sequences (CDSs) with those of B1-EB<sup>T</sup> and Mc-AG77, respectively. This comparative study of three phylogenetically associated endosymbionts has provided insights into their origin and evolution, and suggests the later bacterial invasion and adaptation of B1-EB<sup>T</sup> to its host metabolism.

**Key words:** genome, endohyphal symbiont, interaction, *M. cysteinexigens*, *B. rhizoxinica*.

‘The swollen hypha with two characteristic filiform appendages at its tip’ was the first report of the presence of an endobacterium in the cytosol of Glomeromycotina spores (44). Further morphological confirmation was obtained in subsequent studies (6, 7), and the presence of endobacteria then became widely evident in fungal lineages (14, 60). Recent studies revealed that bacteria and fungi interact at different levels in the bacterial-fungal interface, such as the bacterial slack attachment with the hyphal surface or metabolic complementation with associates and alterations to partners of various forms ranging from antibiosis and signaling to genetic exchange (20, 25, 57). However, endohyphal symbionts residing inside fungal cells exemplify the most intimate illustration of congenial interactions. Despite their genetic and functional diversities, bacterial-fungal interactions have become a new interest in microbiology with significance for microbial ecology (20).

Bacterial endohyphal symbionts are best characterized by the family *Burkholderiaceae* (*Burkholderia*-related) in the class *Betaproteobacteria* and frequently occur in the fungal phyla of Mucoromycota (63). Ascomycota and Basidiomycota are also known to host gammaproteobacteria (2) and alphaproteobacteria (61), respectively. Two endohyphal bacterium (EHB) belonging to the family *Burkholderiaceae* were isolated as new taxa, *Burkholderia rhizoxinica* (HKI 454<sup>T</sup>) and *B. endofungorum* (HKI 456<sup>T</sup>); however, only *B. rhizoxinica*

HKI 454<sup>T</sup> has been characterized in detail (53). An interesting multilateral microbial interaction involving *B. rhizoxinica* HKI 454<sup>T</sup> has been shown to cause rice seedling blight and the causative agent is the Mucoromycotinan fungus *Rhizopus microsporus*. Phytopathogenic *R. microsporus* used an anti-mitotic agent, rhizoxin, as a virulence factor (59, 67). It later became evident that this virulence factor was not produced by *R. microsporus* itself, but by the EHB, *B. rhizoxinica* HKI 454<sup>T</sup> (53, 54). *B. rhizoxinica* controls the vegetative reproduction of the host (31). A previous study isolated *B. rhizoxinica* HKI 454<sup>T</sup> together with its host strain *R. microsporus* var. *chinensis* (ATCC 62417) from rice seedlings in Japan (54). Recent findings have characterized several strains of the non-pathogenic fungus *Mortierella elongata* (Mortierellomycotina) hosting a *Burkholderia*-related obligate endohyphal symbiont named B1-EB<sup>T</sup> (58), which was recently characterized as *Mycoavidus cysteinexigens*, gen. nov., sp. nov., based on its phenotypic, chemotaxonomic, and phylogenetic features (49). The basic function and characteristics of B1-EB<sup>T</sup> as a fungal endosymbiont currently remain unknown. An isolate of *M. elongata* (AG77) from a soil/litter sample close to a poplar tree in Duke forest, the USA was revealed to host *M. cysteinexigens* and it was attained through metagenome sequencing of the host *M. elongata* AG77 together with the endobacterium *M. cysteinexigens* (67). A large and growing endobacterial population of *M. cysteinexigens* AG77 has been observed under nitrogen limiting conditions, whereas *M. elongata* AG77 exhibited decreased growth (38).

Bacterial endosymbionts of fungi are widespread, and extremely different microbial populations may co-exist in the

\* Corresponding author. E-mail: tomoyasu.nishizawa.agr@vc.ibaraki.ac.jp;  
Tel: +81–29–888–8687; Fax: +81–29–888–8525.

† These authors contributed equally to this work.

same host fungal strain. For example, rod-shaped Gram-negative ‘*Candidatus Glomeribacter gigasporarum*’ (*CaGg*) and coccoid Gram-positive *Mollicutes*-related endobacteria (*Mre*) co-exist in a single spore of the arbuscular mycorrhizal fungus *Gigaspora margarita* (13). Diversity in microbes is influenced by many different factors such as vertical and horizontal gene transfer, the activity of mobile genetic elements, the recombination machinery, and mutation rate (46, 47, 66). *B. rhizoxinica* HKI 454<sup>T</sup> is the rhizoxin-producing EHB of the phytopathogen *R. microsporus* and causes blight symptoms in rice seedlings, whereas *M. cysteinexigens* B1-EB<sup>T</sup> is the endohyphal symbiont of the non-pathogenic fungus *M. elongata*. Despite diverse host specificity and modes of interaction with respective hosts, *B. rhizoxinica* and *M. cysteinexigens* have a close phylogenetic relationship (49, 67). Although *CaGg* is the most phylogenetically related to *M. cysteinexigens*, the *CaGg* genome was incompletely sequenced, which may have been due to approximately 7% DNA contamination by the fungal host *G. margarita* (22). In addition, the coexisting endofungal bacteria *Mre* and other *CaGg* species were potentially contained in the spore lysates of *G. margarita* used to concentrate the endobacterial fraction (13, 42), which produced a number of contigs or induced big gaps in the subsequent genome assembly. Although the draft genome of *M. cysteinexigens* B1-EB<sup>T</sup> was sequenced in our previous study (21), information regarding structure, mobile genetic elements, rRNA operons, and tRNAs was still inconclusive. Therefore, in the present study, we elucidated the complete genome sequences of *M. cysteinexigens* B1-EB<sup>T</sup> by sequencing the DNA extract from culture collection and comparing it with the complete genome of the uncultured endosymbiont of *M. elongata* AG77 (named here as Mc-AG77), and also with that of the cultured *B. rhizoxinica* HKI 454<sup>T</sup> as a reference. Comparative genomic analyses of the three phylogenetically associated endosymbionts; *M. cysteinexigens* B1-EB<sup>T</sup>, Mc-AG77, and *B. rhizoxinica* HKI 454<sup>T</sup>, in two different fungal hosts may provide insights into their origin and evolution as well as potential host-bacterium interactions.

## Materials and Methods

### DNA extraction and genome sequencing

Strain B1-EB<sup>T</sup> was incubated on cysteine-containing buffered charcoal-yeast extract agar (B-CYE $\alpha$ ) as described in our previous study (49). Colonies were suspended in sterile water and recovered by centrifuging at 12,000 $\times$ g for 2 min. The DNA of strain B1-EB<sup>T</sup> was extracted using the lysozyme buffer method (69). Genomic DNA (21.8  $\mu$ g) was used to prepare a sequencing library with a SMRTbell Template Prep Kit (Pacific Biosciences of California, Menlo Park, CA, USA). Genome sequencing was performed by Macrogen Korea using PacBio RSII single-molecule real-time (SMRT) sequencing technology (Pacific Biosciences of California).

### PacBio read assembly and genome completing

The long single pass reads (122,714 reads with 1,069,198,690 bases in total) obtained from 1 SMRT cell were assembled using a PacBio *de novo* assembler optimized for the speed Hierarchical Genome Assembly Process (HGAP, v3.0) (9). Three credible contigs were generated: contig 1, 1,355,844 bp with 282-fold coverage; contig 2, 1,037,856 bp with 293-fold coverage; and contig 3, 398,059 bp with 102-fold coverage. These three contigs were combined with the previous draft genome contigs sequenced by the Roche 454 Titanium system (21) by aligning them using the sequence assembler ATGC

ver. 4.2 (Genetyx Co., Tokyo, Japan), resulting in a completely circular genome sequence. The positions of 454 contigs in the complete genome were visualized using a CIRCOS circular visualization data tool (ver. 0.69) (30).

### Multilocus sequence analysis

Five housekeeping genes (*atpD*, *gyrB*, *lepA*, *recA*, and *rpoB*) and the 16S rRNA gene sequence were selected for a multilocus sequence analysis (MLSA) as described previously (17), but with slight modifications. In the present study, nine major members (*Cupriavidus necator* N-1<sup>T</sup>, *C. pinatubonensis* JMP134, *Pandoraea apista* DSM 16535, *P. pulmonicola* DSM 16583, *B. diffusa* RF2-non-BP9, *B. cepacia* ATCC 25416, *B. pseudomallei* K96243, and *B. thailandensis* E264<sup>T</sup>) belonging to the family *Burkholderiaceae* with completely sequenced genomes, and two species (*Ralstonia pickettii* 12D<sup>T</sup> and *R. solanacearum* GMI1000) within the family *Ralstoniaceae*, were selected as the interspecific references and *Polynucleobacter necessarius* subsp. *necessarius* STIR1 as outgroup, respectively, in the MLSA. Briefly, multiple sequence alignments were inferred for each dataset using MAFFT (ver. 7.222). MEGA (ver. 7) was used to modify the alignment scale and change the PHYLIP multiple sequence alignment format. The multiple alignment file and module selection of nucleotide substitution were combined using Kakusan version 4. Among-site rate variations were modeled by a GAMMA distribution with four rate categories. Tree searches were initiated using Randomized Axelerated Maximum Likelihood (RAxML) based on the concatenated dataset (12,492 and 11,023 positions in total with/without 16S rRNA, respectively) of the five housekeeping genes and 16S rRNA. Only strains with a full set of sequenced genes were compared in this analysis. Multilocus phylogenetic trees were visualized using the program FigTree (ver. 1.4.3) (49).

### Genome annotation and general comparison

The complete genome sequence of strain B1-EB<sup>T</sup> was annotated automatically using Microbial Genome Annotation Pipeline (MiGAP) including MetaGeneAnnotator (23) for identifying protein coding sequences (CDSs), tRNAscan-SE (39) for tRNAs, and RNAMmer (35) for rRNAs. GC-skew and GC-content were calculated using a custom Perl script. In order to confirm the annotation of CDSs, the predicted protein sequences were annotated manually using BLASTP with the NCBI-nr database and KEGG Orthology and Links Annotation (KOALA) (26). PHAGE Search Tool Enhanced Release (PHASTER) was used to identify prophage sequences within the genome (3, 71). Detailed pan-genome analyses were performed at the protein level using a bidirectional best-hit (BDBH) algorithm implemented in GET\_HOMOLOGUES (ver. 2.0.23), imposing a minimum pairwise alignment coverage of 75.0% with default parameters (11) and visualized using the CIRCOS circular visualization data tool (ver. 0.69) (30).

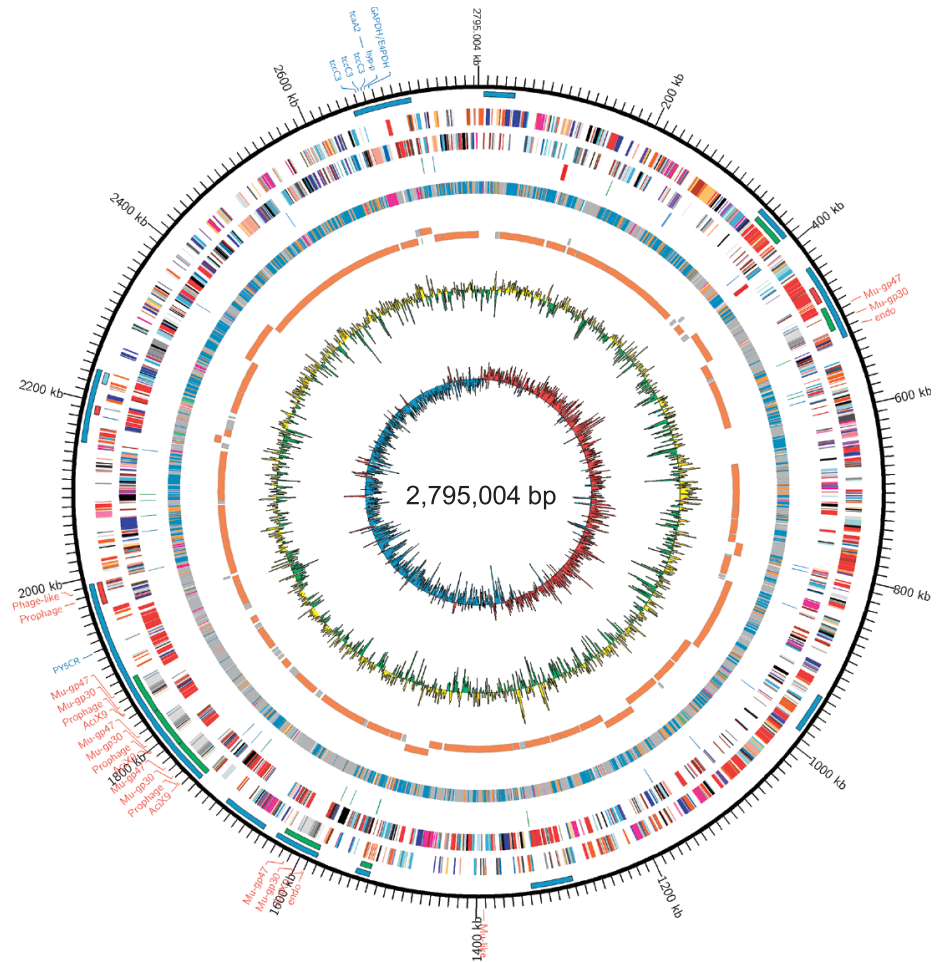
### Nucleotide sequence accession number

The genome sequence of *M. cysteinexigens* B1-EB<sup>T</sup> has been deposited in the DDBJ/EMBL/GenBank databases under the accession number AP018150. The complete genome sequences of *B. rhizoxinica* HKI 454<sup>T</sup> and *M. cysteinexigens* AG77 were obtained from NCBI RefSeq and GenBank (Supplemental Table 7, Reference 68), respectively.

## Results and Discussion

### General characteristics of the complete *M. cysteinexigens* B1-EB<sup>T</sup> genome

The complete B1-EB<sup>T</sup> genome consisted of a 2,795,004-bp circular chromosome with a G+C content of 48.9%, which contained two sets of ribosomal RNA (*rrn*) operons, 41 tRNAs, and 2,317 predicted CDSs (Fig. 1). This result was approximately in accordance with our studied draft genome of strain



Ring 1: GC-skew  $> 0$ ,  $< 0$  Ring 2: GC-content  $> 50\%$ ,  $< 50\%$  Ring 3: contig  $\geq 5$  kb,  $< 5$  kb  
 Ring 4: relative organisms of predicted CDSs.  
 Ring 5: predicted rRNA and tRNA.  
 Ring 6 and 7: COG functional classifications  
 Ring 8: intact, questionable, and incomplete prophage sequences are indicated in blue, green, and red solid bars, respectively; (ring 9) transposons; and (ring 10) scale in kb. Red and blue text outside shows the positions of phage-related genes and insecticidal toxin proteins, respectively.

■ Actinobacteria    ■ Alphaproteobacteria    ■ Betaproteobacteria    ■ Burkholderia    ■ Chlorobi  
■ Cyanobacteria    ■ Deinococcus    ■ Deltaproteobacteria    ■ Gammaproteobacteria    ■ No data  
■ rRNA    ■ tRNA in clockwise direction    ■ tRNA in counterclockwise direction  
■ RNA processing and modification    ■ Chromatin structure    ■ Energy production/ conversion    ■ Chaperones  
■ Cell division/ chromosome partitioning    ■ Amino acid transport/ metabolism    ■ Inorganic ion metabolism  
■ Carbohydrate transport/ metabolism    ■ Coenzyme transport/ metabolism    ■ Lipid transport/ metabolism  
■ Translation/ ribosomal structure    ■ Transcription    ■ Replication/ recombination/ repair    ■ Cell motility  
■ Cell wall/ membrane/ envelope biogenesis    ■ Nucleotide transport/ metabolism    ■ Signal transduction  
■ Secondary metabolite biosynthesis/ transport/ catabolism    ■ General function prediction    ■ Cytoskeleton  
■ Intracellular trafficking/ secretion    ■ Nuclear structure    ■ Defense mechanisms    ■ Extracellular structures  
■ Function unknown    ■ No COG category

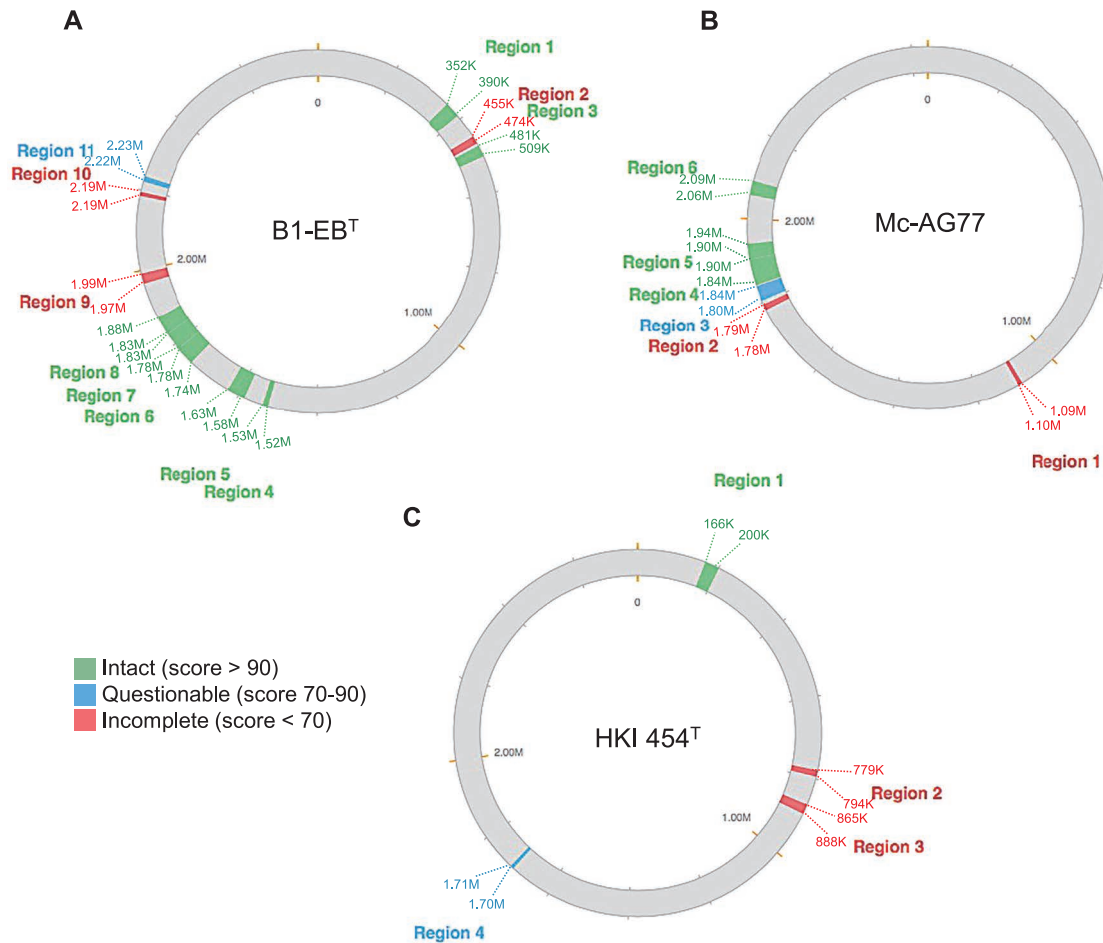
**Fig. 1.** Schematic spherical illustration of the B1-EB<sup>T</sup> genome. Circles display from the inside to the outside as follows: (ring 1) GC-skew (G-C/G+C ratio) using a 1-kb window; (ring 2) GC-content using a 1-kb window; (ring 3) 454-contigs; (ring 4) relative bacteria of predicted CDSs; (ring 5) rRNAs and tRNAs; (rings 6 and 7) predicted CDSs transcribed in a counterclockwise/clockwise direction; (ring 8) intact, questionable, and incomplete prophage sequences are indicated in blue, green, and red solid bars, respectively; (ring 9) transposons; and (ring 10) scale in kb. Red and blue text outside shows the positions of phage-related genes and insecticidal toxin proteins, respectively.

B1-EB<sup>T</sup> (21), revealing a genome size of 2,653,566 bp with a G+C content of 46.1%, 1 set of *rrn* operons, 39 tRNAs, and 2,390 CDSs. Most of the CDSs were identified as well conserved with CDSs in bacteria within the genus *Burkholderia* (Fig. 1). Of note, the CDSs encoding insecticidal toxin proteins, which are of potential benefit to the fungal host, were close to those conserved by *Xenorhabdus* (*Gammaproteobacteria*), a genus living symbiotically with soil entomopathogenic nematodes.

Nevertheless, the annotation of the complete genome indi-

cated that B1-EB<sup>T</sup> lacks cysteine biosynthetic pathways, as reported previously (21). In addition, the B1-EB<sup>T</sup> genome had conserved 11 transposons, 7 intact prophages, and 1 questionable and 3 incomplete prophages. Six of these transposons overlapped with the 11 prophage sequences (Fig. 1) encoding prophage-related proteins such as phage integrase and phage transposase (Fig. 2A and S1), suggesting that these regions were acquired by prophage insertion. Besides transposons, there were several retron-type reverse transcriptase





**Fig. 2.** Prophage sequences within B1-EB<sup>T</sup> (A), Mc-AG77 (B), and HKI 454<sup>T</sup> (C) genomes. The details of each region are shown in Fig. S1. Position-based identification of prophage region is shown in colors. The scale bar is in bp.

and phage proteins. The genetic evidence supporting the introduction to foreign DNA contrasts endosymbionts such as *Buchnera* sp. APS, which was harbored by *Acyrtosiphon pisum* (pea aphid) (43, 62). The B1-EB<sup>T</sup> genome also contained several predicted toxin-antitoxin systems including the insecticidal toxin complex RelE/StbE, HipA, MazF, and PIN domain proteins encoding genes. These features imposed an addiction-like mechanism on the prokaryotic host genome and arbitrated the stability of mobile elements (1). In addition to their genetic addiction, these factors have been anticipated to facilitate growth retardation under stress conditions (51). Therefore, toxin-antitoxin systems may be essential for endohyphal growth control during host colonization. The B1-EB<sup>T</sup> genome also included 358 hypothetical proteins. We hypothesized that some strain-specific genomic components of unknown function may correlate with the evolution of the endofungal lifestyle.

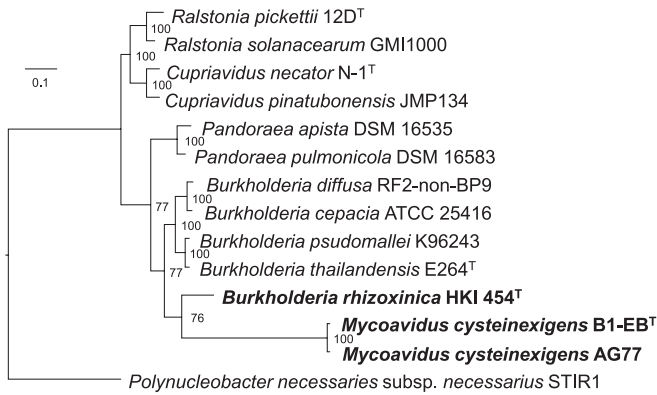
#### Phylogenetic analysis of *M. cysteinexigens* and related taxa within the family Burkholderiaceae

In the present study, we generated a multilocus phylogenetic tree of 14 reference strains (including 4 type strains) of Burkholderiaceae using the *atpD*, *gyrB*, *lepA*, *recA*, and *rpoB* genes in combination with the 16S rRNA gene sequence and employed RAxML with support values estimated with 100 bootstrap replicates for the evolution of the genus *Mycoavidus*

as a sister clade to *Burkholderia* (Fig. 3 and S2). The output of this analysis also showed a distinct clade for *B. rhizoxinica*, which was positioned within the genus *Burkholderia*. Based on phylogenomics and the separation of fungal hosts, two alternative scenarios, ‘early’ and ‘late’ bacterial invasion, were hypothesized (5). Regarding *Burkholderia*-related endobacteria, early bacterial invasion estimated the existence of an ancestral free-living *Burkholderia*-related endobacteria that invaded the common ancestor of Glomeromycotina and Mortierellomycotina (5, 33). As discussed by Bonfante & Desirò (2017) (5), late bacterial invasion may have occurred when the evolutionary lines leading to Glomeromycotina and Mortierellomycotina had already separated. However, *B. rhizoxinica* appeared to have an independent origin shared with the free-living *Burkholderia*-related endobacteria that diverged from the ancestral lineage (34). This hypothetical scenario may lead to further discussions on whether specific bacterial-fungal interactions occurred in the common ancestor only (not separated) and not in separated Glomeromycotina and Mortierellomycotina in which obligate endobacteria essentially lack the ability to grow independently of their fungal hosts.

#### Comparative genomics of EHB belonging to *M. cysteinexigens* and *B. rhizoxinica*

In order to obtain insights into the origin and evolution of gene families characteristic of the phylogenetically associated



**Fig. 3.** Randomized accelerated maximum likelihood (RAxML) tree based on concatenated sequences (11,023 positions in total) of five housekeeping genes (*atpD*, *gyrB*, *lepA*, *recA*, and *rpoB*) indicating the relative placement of the three endofungal bacteria (bold) and other genera in the families *Burkholderiaceae* and *Ralstoniaceae*. Pairwise distance between B1-EB<sup>T</sup> and Mc-AG77 is 0.013. The horizontal lines show genetic distances, which are supported by values estimated with 100 bootstrap replicates. The scale bar indicates the number of substitutions per nucleotide position.

endothelial symbionts in diverse fungal hosts, we conducted comparative genome analyses on *M. cysteinexigens* B1-EB<sup>T</sup>, Mc-AG77, and *B. rhizoxinica* HKI 454<sup>T</sup> by the computation of pairwise homologous gene clusters. We excluded *CaGg* from this comparative analysis because of its controversial genome sequence, which we clarified earlier in this study. A previous study reported that the total genome of *B. rhizoxinica* HKI 454<sup>T</sup> is 3,750,139 bp, consisting of the chromosome and two plasmids, pBRH01 and pBRH02 (33). The entire genome comprises 3,878 CDSs, 2,437 of which were estimated to have biological functions, 1,441 had no known function, 897 appeared to be similar to other database entries, and 544 remained without function and similarity (34). The genome of Mc-AG77 is 2,638,116 bp and contained 2,255 CDSs (68). Both bacterial genomes were markedly smaller than that of *B. rhizoxinica* HKI 454<sup>T</sup> and harbored fewer transcriptional regulator genes (Table 1). Pan-genome GET\_HOMOLOGUES analyses at the protein level using the BDBH algorithm resulted in the extraction of 1,266 and 1,221 ortholog clusters with shared homology to the HKI 454<sup>T</sup> genome based on the % identity of amino acid pairs with B1-EB<sup>T</sup> and Mc-AG77, respectively, whereas 1,788 homologous regions with high % identity were predicted between B1-EB<sup>T</sup> and Mc-AG77. We herein summarized the distribution (% shared proteomes) of ortholog clusters based on the % identity among the proteomes of *Mycoavidus* and *B. rhizoxinica* genomes (Fig. 4A). Pairwise comparative mapping of the B1-EB<sup>T</sup>, Mc-AG77,

and HKI 454<sup>T</sup> genomes displayed regions of shared homology (Fig. 4B). Although the MLSA of 14 reference strains of *Burkholderiales* using five housekeeping genes and the 16S rRNA gene revealed that *M. cysteinexigens* B1-EB<sup>T</sup> and Mc-AG77 both had a similar distance with *B. rhizoxinica* HKI 454<sup>T</sup> (Fig. 3 and S2), pairwise comparative mapping displayed the high homology of Mc-AG77 with *B. rhizoxinica* HKI 454<sup>T</sup>. The B1-EB<sup>T</sup> genome was disrupted by a large amount of mobile genetic elements, with 146 CDSs (6.3%) encoding proteins with similarities to transposases or inactivated derivatives, plasmid stability proteins, reverse transcriptases, and phage proteins, which was similar to HKI 454<sup>T</sup> (Table S1); however, it is currently unclear how many of these were active insertion elements. We also predicted 271 (11.6%) characterized proteins, 137 hypothetical proteins, and 51 transposase and inactivated derivatives in B1-EB<sup>T</sup>, which were non-homologous with Mc-AG77 and HKI 454<sup>T</sup> (Table S2). Since the annotations of Mc-AG77 CDSs are not yet known, we attempted to predict their functions from homology shared with the B1-EB<sup>T</sup> genome.

#### *Prophage sequences, horizontal gene transfer, and genome reductions in M. cysteinexigens and B. rhizoxinica.*

Genomes of obligate intracellular bacteria are often categorized by high levels of genetic drift, genome streamlining, and a lack of mobile DNA (plasmid, transposon, and prophage), which emerge with prominent degrees of stability and gene synteny (8, 36). However, a number of studies have showed that the smaller genomes of obligate intracellular species comprise a level of prophage sequences (e.g., the prophage sequence of *Wolbachia pipientis* is less than 2.0% of the total genome) similar to the lower end of the range found in facultative intracellular species (8). We herein identified the prophage sequence in the genomes of EHB *M. cysteinexigens* B1-EB<sup>T</sup>, Mc-AG77, and *B. rhizoxinica* HKI 454<sup>T</sup> by PHASTER. Of note, the chromosomes of B1-EB<sup>T</sup>, Mc-AG77 comprised a markedly higher level of prophage sequences than those in the HKI 454<sup>T</sup> chromosome. The prophages and their remnants in B1-EB<sup>T</sup> and Mc-AG77 accounted for 11.5% (11 regions with 267 phage-related CDSs) and 7.2% (6 regions with 133 phage-related CDSs) of the total lengths of their chromosomes, respectively (Fig. 2A and B), whereas that of HKI 454<sup>T</sup> was only 2.9% (4 regions with 79 phage-related CDSs, Fig. 2C). A longer association with a host resulted in the endobacterial genomes undergoing a mutational bias toward a lower GC content (12, 40, 43). The GC contents of prophage sequences in B1-EB<sup>T</sup> and Mc-AG77 chromosomes (prophage GC contents, 49.2% and 48.4%, respectively) were closer to those of their chromosomes (Table 2), whereas an apparently lower GC

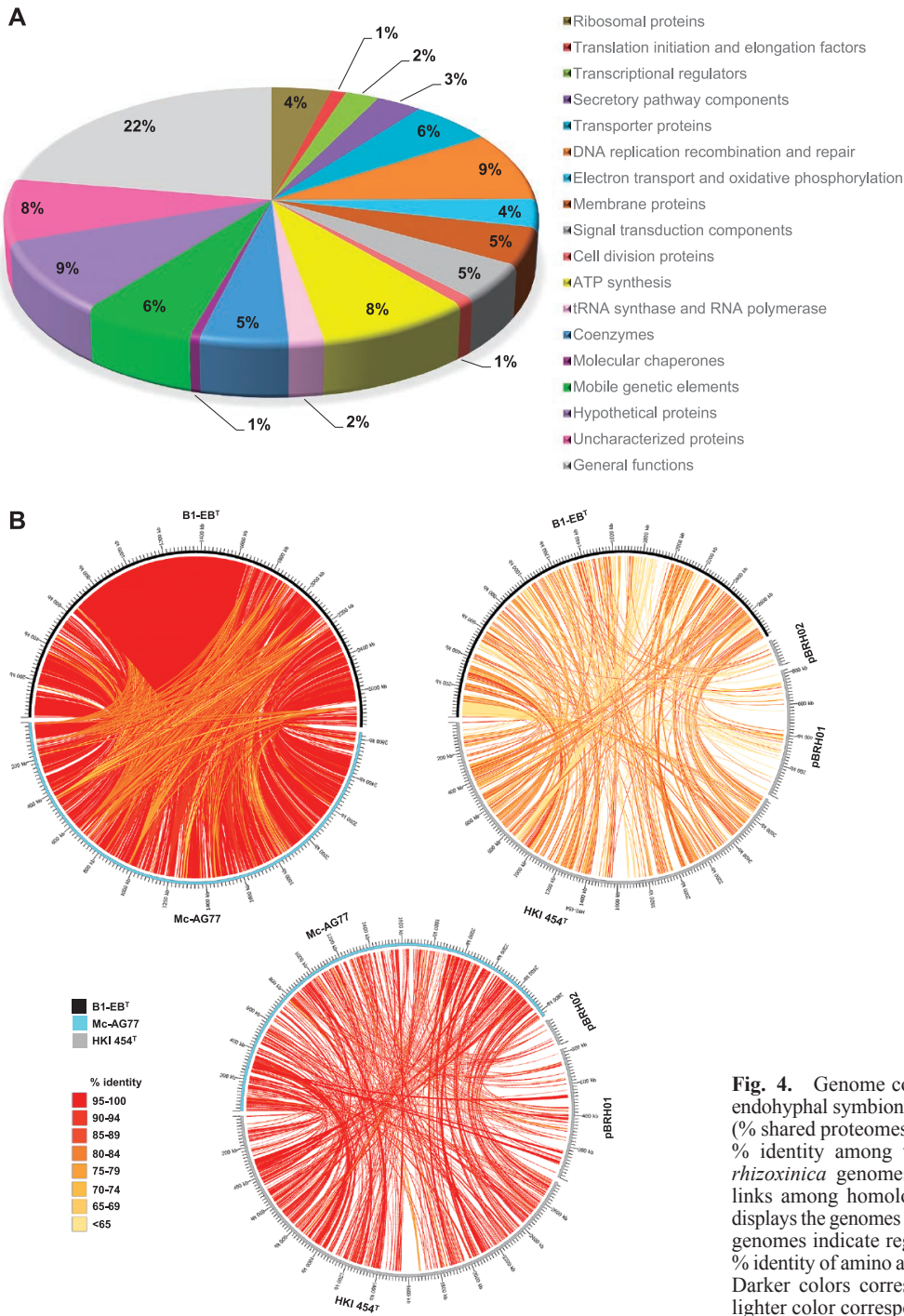
**Table 1.** Comparative genome analysis of *M. cysteinexigens* B1-EB<sup>T</sup>, Mc-AG77, and *B. rhizoxinica* HKI 454<sup>T</sup>.

Organism	Genome size (bp)	G+C content (%)	Coding sequences (CDSs)	Mobile genetic elements (% for all CDSs)	Transcriptional regulators (% for all CDSs)	References
<i>M. cysteinexigens</i> B1-EB <sup>T</sup>	2,795,004	48.9	2,317	146 (6.3)	39 (1.68)	This research
<i>M. cysteinexigens</i> AG77	2,638,116	49.0	2,255	71 (3.14)**	39 (1.72)***	Uehling <i>et al.</i> , 2017 (68)
<i>B. rhizoxinica</i> HKI 454 <sup>T</sup>	3,750,139*	60.7	3,870	255 (6.6)	193 (5.0)	Lackner <i>et al.</i> , 2011 (34)

\* Includes a 2,755,309-bp chromosome, 822,304-bp mega plasmid, and 172,525-bp plasmid.

\*\* Genes coding for mobile genetic elements are predicted from the homologous sequence of the B1-EB<sup>T</sup> genome.

\*\*\* Genes coding for transcriptional regulators are predicted from the homologous sequence of the B1-EB<sup>T</sup> genome.



**Fig. 4.** Genome comparison of three beta-proteobacterial endohyphal symbionts. (A) Pie chart showing the distribution (% shared proteomes) of ortholog clusters (BDBH) based on % identity among the proteomes of *Mycoavidus* and *B. rhizoxinica* genomes. (B) Comparative mapping showing links among homologous sequences. The circular exterior displays the genomes of organisms clockwise. Ribbons between genomes indicate regions of shared homology based on the % identity of amino acid pairs between each pair of organisms. Darker colors correspond to higher identity links and the lighter color corresponds to a lower identity (see legends).

content (58.6%) was observed for the prophage sequences in the HKI 454<sup>T</sup> chromosome (chromosome GC content, 61.2%) (Fig. 2C). This result suggests a relatively longer association of prophages with *M. cysteinexigens* hosts than those with *B. rhizoxinica* HKI 454<sup>T</sup>.

As an important means for gene-gain events, prophages horizontally transfer genes from one bacterial genome to another. More than 70% of phage-related CDSs in the genomes of *M. cysteinexigens* were close to those CDSs of the bacteriophages infecting *Gammaproteobacteria* such as *Aggregatibacter actinomycetemcomitans*, *Edwardsiella tarda*, and *Escherichia coli*, followed by those of phage-infecting

bacteria of *Betaproteobacteria*, *Alphaproteobacteria*, *Firmicutes*, *Cyanobacteria*, *Bacteroidetes*, *Actinobacteria*, and *Tenericutes* (Mollicutes) (Table 2 and S3). In contrast, most phage-related CDSs in the *B. rhizoxinica* HKI 454<sup>T</sup> chromosome (72.2%) were similar to those found in the phages infecting betaproteobacteria, particularly *B. cenocepacia*, *B. cepacia*, and *B. thailandensis*, followed by the CDSs related to the phages infecting bacteria of *Gammaproteobacteria* (22.8%) and *Firmicutes* (5.0%) (Table 2 and S3). This phage originated by distinction, suggesting a difference in horizontal gene transformation between the two lineages. However, many researchers have shown that the smaller genomes of obligate intracellular



**Table 2.** Phage-related CDSs in genomes of *M. cysteinexigens* B1-EB<sup>T</sup>, Mc-AG77, and *B. rhizoxinica* HKI 454<sup>T</sup>.

Bacterial host of relative phage/prophage in	B1-EB <sup>T</sup>		Mc-AG77		HKI 454 <sup>T</sup>	
	CDS count	Percentage	CDS count	Percentage	CDS count	Percentage
<i>Actinobacteria</i>	2	0.75	1	0.75	0	—
<i>Bacteroidetes</i>	4	1.50	2	1.50	0	—
<i>Cyanobacteria</i>	2	0.75	4	3.01	0	—
<i>Firmicutes</i>	9	3.37	7	5.26	4	5.06
<i>Alphaproteobacteria</i>	13	4.87	3	2.26	0	—
<i>Betaproteobacteria</i>	39	14.61	21	15.79	57	72.15
<i>Gammaproteobacteria</i>	197	73.78	95	71.43	18	22.78
<i>Tenericutes</i> (Mollicutes)	1	0.37	0	—	0	—
<b>Total</b>	<b>267</b>	<b>100.00</b>	<b>133</b>	<b>100.00</b>	<b>79</b>	<b>100.00</b>

bacteria preferred the inactivation and deletion of prophage sequences over their insertion and duplication (8, 36). An analysis using PHASTER showed that only one prophage region in the *B. rhizoxinica* HKI 454<sup>T</sup> chromosome was intact, which was consistent with the deletion of the prophage in the intracellular bacterial genome. Nevertheless, *M. cysteinexigens* had more intact prophages in their chromosomes (7 in B1-EB<sup>T</sup>, and 3 in Mc-AG77), indicating that prophages in *M. cysteinexigens* genomes were undergoing deletion and inactivation processes. In consideration of the relatively longer association of prophages with *M. cysteinexigens*, the deletion rate of prophages in the genomes of *M. cysteinexigens* appeared to be slower than those proceeding in the genome of *B. rhizoxinica*, suggesting that the prophages still possess physiological functions in *M. cysteinexigens*. However, the lengths of non-prophage regions in the chromosomes of B1-EB<sup>T</sup>, Mc-AG77, and HKI 454<sup>T</sup> were 2,469,125 bp, 2,447,528 bp, and 2,674,052 bp, respectively. Therefore, the primordial genetic elements in the *M. cysteinexigens* genome were presumed to be more reduced than those in the HKI 454<sup>T</sup> chromosome if the two endohyphal bacteria were derived from a common ancestor bacterium.

#### Metabolic potential and nutrient uptake in *M. cysteinexigens* and *B. rhizoxinica*

There has always been a primary interest in the metabolic potential of EHB, whether they are nutrient producers or consumers and how they exchange nutrients and cofactors with their fungal hosts. We compared the primary metabolism, cofactor biosynthesis, and membrane transport of B1-EB<sup>T</sup> and Mc-AG77 with HKI 454<sup>T</sup> (Table 3) and created a model of metabolic processes in *M. cysteinexigens* (Fig. 5). The deduction of primary metabolism and transporters suggested that *M. cysteinexigens* (B1-EB<sup>T</sup> and Mc-AG77) possess genes for glucose/maltose/*N*-acetylglucosamine- and mannose/fructose-specific phosphotransferase system components, but do not utilize glucose, arabinose, mannose, mannitol, *N*-acetyl-D-glucosamine, maltose, or gluconate. It was already known that HKI 454<sup>T</sup> prefers glycerol over glucose as a carbon source (53). HKI 454<sup>T</sup> may consume organic acids such as citrate, malate, glycerol, and ethanol from *Rhizopus* (41), whereas no genes encoding ethanol assimilation or organic acid importers were detected in the B1-EB<sup>T</sup> genome (Table 3 and S1), suggesting that B1-EB<sup>T</sup> cannot utilize these organic acids from its host.

*M. cysteinexigens* and *B. rhizoxinica* both harbor genes responsible for importing branched-chain and aromatic amino

acids, histidine, acidic amino acids, and glycine; the biosynthetic routes for all proteinogenic amino acids and efflux systems for basic amino acids (Table S4) (68). Additionally, both endosymbionts had genes for dipeptide and oligopeptide import systems. However, B1-EB<sup>T</sup> and Mc-AG77 do not possess efflux systems for cysteine. B1-EB<sup>T</sup> possesses genes encoding biosynthesis enzymes for essential cofactors such as pyridoxine, heme, flavin, biotin, and thiamine, which were similar to those in HKI 454<sup>T</sup>. HKI 454<sup>T</sup> may provide some amino acids and cofactors to its host (34), which suggests that amino acid metabolism and transportation systems adapted during endobacterium-host symbiosis in B1-EB<sup>T</sup> and HKI 454<sup>T</sup>.

#### Secondary metabolites, lipopolysaccharides, exopolysaccharides, and efflux pumps in *M. cysteinexigens* and *B. rhizoxinica*

The genomic locus for antimetabolic rhizoxin biosynthesis in HKI 454<sup>T</sup> revealed a hybrid system of non-ribosomal peptide synthetase (NRPS) and polyketide synthase (PKS) (52). The NRPS/PKS system architecture in HKI 454<sup>T</sup> possessed a phosphopantetheinyl transferase to activate the acyl and peptidyl carrier protein domains, 14 NRPS multimodular enzymes, and macrolide-specific ABC-transporters for toxin delivery to the host fungus (34). Since no peptides corresponding to these gene clusters were isolated from HKI 454<sup>T</sup>, it was hypothesized that some NRPS products may function as siderophores or antibiotics (34). *M. cysteinexigens* B1-EB<sup>T</sup> contained a NRPS module (MCB1EB\_1072), which was an ortholog of the adenylation domain in the first NRPS module of bacillibactin present in pBRH01 (RBRH\_00588) in *B. rhizoxinica* HKI 454<sup>T</sup>. On the other hand, this NRPS module in B1-EB<sup>T</sup> shared strong identity (97%) with *Glomeribacter* sp. 1016415. However, its function and biosynthetic potential remain unclear. Virulence factors help bacteria to invade their hosts, which were completely functional in EHB HKI 454<sup>T</sup>. Therefore, the mechanisms by which B1-EB<sup>T</sup> invaded its host fungus without a complete virulence system have not yet been elucidated. Although bacterial intrusion often depends on genetic characteristics, the invasion mechanism strictly differed with the environment encountered by the bacterium (56) and with different life history policies such as avirulent microorganisms, latent pathogens, and virulent pathogens in the early stages of infection or at a certain phase of their development (28). A single mutation was shown to result in the loss of the virulence factor, transforming the pathogenic fungus, *Colletotrichum magna*, into an endophyte (19). Therefore, virulence-related genes may be lost during bacterial genome reduction in B1-EB<sup>T</sup>. We speculated how *M. cysteinexigens*

**Table 3.** Comparison of metabolic potentials, nutrient uptake and defense responses in B1-EB<sup>T</sup>, Mc-AG77, and HKI 454<sup>T</sup>.

Attributes	B1-EB <sup>T</sup>	Mc-AG77	HKI 454 <sup>T</sup>
<b>Primary metabolism</b>			
Gluconeogenesis*	–	–	+
Ethanol assimilation*	–	–	+
Import of organic acids	–	–	+
Amino acid metabolism	+	+	+
Biosynthetic routes for proteinogenic amino acids*	+	+	+
Dipeptide and oligopeptide import system*	+	+	+
<b>Cofactor biosynthesis</b>			
Biosynthesis of pyridoxine*, heme*, flavin*, biotin, and thiamine*	+	+	+
<b>Membrane transport</b>			
Putative cobalamin transport system*	+	+	+
Fe <sup>2+</sup> , Mg <sup>2+</sup> , Co <sup>2+</sup> , Zn <sup>2+</sup> , K <sup>+</sup> uptake system	+	+	+
Na <sup>+</sup> /Ca <sup>2+</sup> antiporter*	+	+	+
Na <sup>+</sup> /citrate symporter*	+	+	+
NRPS/PKS** system for rhizoxin biosynthesis	–	–	+
Lantibiotic biosynthesis*	–	–	+
Chitinolytic enzymes and chitin-binding proteins*	–	–	+
<b>Amino acid metabolism</b>			
Branched-chain amino acid	+	+	+
Aromatic amino acid	+	+	+
Histidine	+	+	+
Glutamate/aspartate	+	+	+
Glycine	+	+	+
Citrate	+	+	+
Biosynthetic routes for proteinogenic amino acids	+	+	+
Efflux systems for arginine, histidine, lysine	+	+	+
Efflux systems for cysteine	–	–	+
<b>Polar lipids</b>			
Phosphatidylethanolamine*	+	+	–
Phosphatidylglycerol*	+	+	–
Diphosphatidylglycerol*	+	+	–
Aminophospholipid*	+	+	–
Aminolipids*	+	+	–
<b>Fatty acid synthesis</b>			
Acetyl-CoA carboxylase components	+	+	+
Saturated/unsaturated synthase components*	+	+	+
Peripheral enzymes (ACP synthase and biotin ligase) *	+	+	+
Glycerophospholipid metabolism*	+	+	+
<b>Lipopolysaccharide gene cluster</b>			
Lipopolysaccharide	–	–	+
O-antigen*	–	–	+
Inner membrane ABC transporter, MsbA*	+	+	+
Transport complex (LptA, LptAB, LptC, LptD, LptE, LptG) *	–	–	+
<b>Exopolysaccharide/capsular polysaccharide biosynthesis</b>			
Antigen flippase and polymerase*	–	–	+
Trans-envelope transport complex (Wza/Wzb/Wzc) *	–	–	+
Glycosyltransferases gene*	+	+	+
Mannosyltransferase gene*	–	–	+

\* Genes listed in Mc-AG77 are predicted from the homologous sequence of the B1-EB<sup>T</sup> genome.

\*\* Non-ribosomal peptide synthetase/polyketide synthase

may invade host fungal tissues, such as when fungal hyphae are damaged by other fungi or by grazing soil fauna. Subsequent invasion occurred through breaks in the fungal cell wall. In contrast, *B. rhizoxinica* HKI 454<sup>T</sup> invades the host, even under laboratory conditions (34).

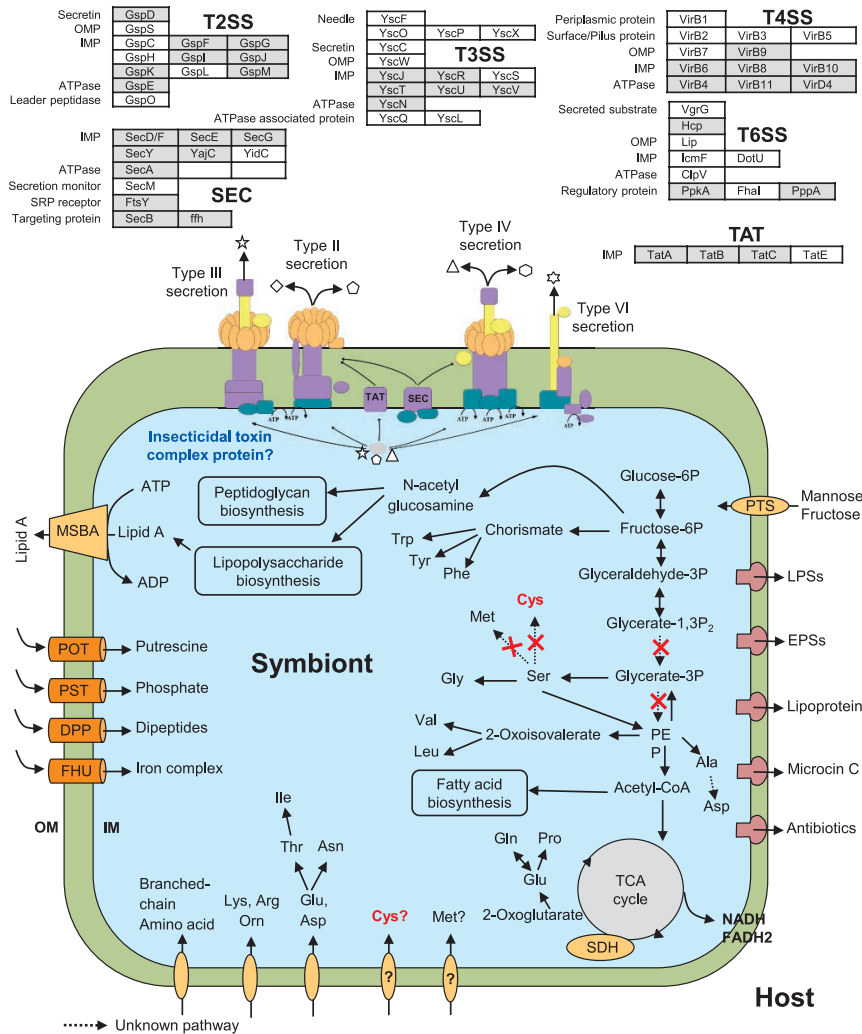
Lipopolysaccharides (LPSs) and exopolysaccharides (EPSs) play key roles in the recognition and stability of the symbiosis of *B. rhizoxinica* HKI 454<sup>T</sup> within its host and a gene cluster involved in LPS biosynthesis in HKI 454<sup>T</sup> (37). Besides LPS biosynthesis, genome sequencing revealed genes that enabled HKI 454<sup>T</sup> to also produce EPSs (34). The biosynthesis of the

homopolymeric O-antigen LPS depends on an ABC transporter (70). B1-EB<sup>T</sup> and Mc-AG77 genomes identified an inner membrane ABC transporter, MsbA, similar to the macrolide-specific ABC transporter in HKI 454<sup>T</sup> and a glycosyltransferase gene (Table S1), which indicates that only these two components of the carbohydrate-derived putative LPS/EPS system of the recognition process are present in *M. cysteinexigens*.

*Secretion systems and insecticidal toxin complexes in M. cysteinexigens and B. rhizoxinica*

Secretory proteins pass through the cell membrane in bacteria,





**Fig. 5.** Model representing metabolic processes and the secretion system assumed from the B1-EB<sup>T</sup> genome. Putative type II, III, IV, and VI secretory proteins are highlighted in grey. POT, polyamine transport system; PST, phosphate-specific transport; DPP, dipeptide permease; FHU, ferrichrome-uptake; EPSs, exopolysaccharides; LPSs, lipopolysaccharides; IM, inner membrane; OM, outer membrane; OMP, outer membrane protein; IMP, inner membrane protein; SEC, general secretion pathway; TAT, twin arginine translocation pathway.

either via the general secretion pathway (4, 15) or twin arginine translocation pathway (50), and regulate host control mechanisms. Therefore, in the present study, we analyzed the genes coding for putative protein secretion systems. The Mc-AG77 genome acquired genes for type II (T2SS), III (T3SS), and IV (T4SS) secretion systems, and these contributed to the exchange of proteins and DNA with its host (68). Additionally, the B1-EB<sup>T</sup> genome possesses a few components of a type VI secretion system (T6SS).

T2SS is known to be involved in the secretion of toxins and lytic proteins (*e.g.*, chitinolytic enzymes, chitin-binding proteins) outside the cell in various bacteria (10, 18). In addition, microbial chitinases are presumed to degrade hyphal walls and allow bacteria to invade fungi. The betaproteobacterium *Nitrosomonas europaea* appears to have significant, but incomplete T2SS core components (10). *B. rhizoxinica* HKI 454<sup>T</sup> possessed a complete T2SS, whereas B1-EB<sup>T</sup> possessed a partial T2SS constituent, which is homologous with that in HKI 454<sup>T</sup> (Table S5). It currently remains unclear whether T2SS homologs in diverse bacteria are expressed and functional. *B. rhizoxinica* and other free-living, pathogenic

*Burkholderia* species (*e.g.* *B. thailandensis* and *B. pseudomallei*) contain *hrp* T3SS gene clusters and these are known to play a significant role in symbiosis through the translocation of effector proteins with their hosts (32, 64). B1-EB<sup>T</sup> possessed a putative T3SS constituent with partial homology with the HKI 454<sup>T</sup> T3SS pathway (Table S6). In contrast, the T3SS components of Mc-AG77 shared strong homology with *Gamma*proteobacteria such as *Salmonella* and *Yersinia* (68).

The T4SS system was shown to be involved in the plant pathogenicity of *B. cenocepacia* (16). A partial T4SS was found in the B1-EB<sup>T</sup> genome and in the plasmid pBRH02 of HKI 454<sup>T</sup> (Table S1). In the B1-EB<sup>T</sup> genome, we recognized a hypothetical protein homolog type 4 major prepilin protein, pilus assembly pathway ATPase Pule/Tfp homolog type 4 pilus biogenesis protein, a prepilin signal peptidase, and ATPase pilus assembly distributed throughout the chromosome in HKI 454<sup>T</sup> (Table S1). Adhesive structures, particularly type 4 bacterial pili, are generally major adhesion factors for host surfaces in Gram-negative bacteria (27). However, the pilus structure in B1-EB<sup>T</sup> lacked several key genes *e.g.*, secretin, motor protein, and other accessory proteins. Moreover,

hemolysin gene products are known to be involved in plant pathogenicity (24), and the hemagglutinin-like proteins, putative hemolysin/phospholipase/lecithinase, were also found in the B1-EB<sup>T</sup> genome and had sequence similarities with RBRH\_02330, RBRH\_00133, and RBRH\_01303 of HKI 454<sup>T</sup>. Additionally, a gene cluster coding for an insecticidal toxin protein (coding genes: CSV86\_11535, *tccC3*, PCH70\_41490, and *tcaA2*) was found in the homologs of *Xenorhabdus bovienii* SS-2004 and *Photorhabdus asymbiotica* (45, 55). The insecticidal toxin proteins present in the HKI 454<sup>T</sup> genome were completely different from those present in the B1-EB<sup>T</sup> genome.

## Conclusions

Compared with animals and plants, the huge genetic diversity of the bacterial biosphere makes it an outstanding resource for studying genetics and evolution. The bacteria that dwell inside eukaryotes contain some of the smallest, most stable, highly repetitive, and recombined bacterial genomes sequences currently known (40, 48). Although bacterial-fungal symbiosis has important implications for agriculture, biotechnology, and food safety, this interaction has long been overlooked (29, 65). The genome of B1-EB<sup>T</sup> offers an extraordinary picture of an obligate symbiont that has most likely lost its capacity to grow outside of its fungal host (*i.e.*, cannot grow in conventional nutrient media) (12, 21, 49). In the present study, we found that the B1-EB<sup>T</sup> and Mc-AG77 genomes exhibited amino acid similarities between ~20–84% and ~18–100%, respectively, with HKI 454<sup>T</sup> (Table S1). Obligate endosymbiotic bacteria were anticipated to have lost many essential genes as a result of genome reduction during evolution (51). In contrast to other fungal endosymbionts in the family *Burkholderiaceae*, both strains of *M. cysteinexigens* lack many genes associated with host invasion such as chitinolytic enzymes (chitinase, chitinase), chitin-binding proteins, and several T2SS components, combined with sugar importers and the ATP/ADP antiporter, during bacterial genome reduction, suggesting later bacterial invasion and adaptation to its host metabolism.

## Acknowledgements

We thank Drs. Reiko Fujimura and Kenshiro Oshima from The University of Tokyo for their help with the data analysis. We are grateful to Hikaru Nabatame and Nanami Konuma for their technical assistance. This research was supported by the Institute for Fermentation, Osaka, JSPS KAKENHI Grant Number 17K07695, and the Council for Science, Technology and Innovation (CSTI), Cross-ministerial Strategic Innovation Promotion Program (SIP), “Technologies for creating next-generation agriculture, forestry and fisheries” (funding agency: Bio-oriented Technology Research Advancement Institution, NARO).

## References

- Arcus, V.L., P.B. Rainey, and S.J. Turner. 2005. The PIN-domain toxin-antitoxin array in mycobacteria. *Trends Microbiol.* 13:360–365.
- Arendt, K.R., K.L. Hockett, S.J. Araldi-Brondolo, D.A. Baltrus, and A.E. Arnold. 2016. Isolation of endohyphal bacteria from foliar Ascomycota and *in vitro* establishment of their symbiotic associations. *Appl. Environ. Microbiol.* 82:2943–2949.
- Arndt, D., J.R. Grant, A. Marcu, T. Sajed, A. Pon, Y. Liang, and D.S. Wishart. 2016. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res.* 44:W16–W21.
- Berks, B.C., T. Palmer, and F. Sargent. 2005. Protein targeting by the bacterial twin-arginine translocation (Tat) pathway. *Curr. Opin. Microbiol.* 8:174–181.
- Bonfante, P., and A. Desirò. 2017. Who lives in a fungus? The diversity, origins and functions of fungal endobacteria living in Mucoromycota. *ISME J.* 11:1727–1735.
- Bonfante-Fasolo, P., and S. Scannerini. 1976. The ultrastructure of the zygospore in *Endogone flammicorona* Trappe and Gerdemann. *Mycopathologia* 59:117–123.
- Bonfante-Fasolo, P., and S. Scannerini. 1977. Cytological observations on the mycorrhiza *Endogone flammicorona*-*Pinus strobus*. *Allionia* 22:23–34.
- Bordenstein, S.R., and W.S. Reznikoff. 2005. Mobile DNA in obligate intracellular bacteria. *Nat. Rev. Microbiol.* 3:688–699.
- Chin, C.-S., D.H. Alexander, P. Marks, *et al.* 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat. Methods* 10:563–569.
- Cianciotto, N.P. 2005. Type II secretion: a protein secretion system for all seasons. *Trends Microbiol.* 13:581–588.
- Contreras-Moreira, B., and P. Vinuesa. 2013. GET\_HOMOLOGUES, a versatile software package for scalable and robust microbial pangenome analysis. *Appl. Environ. Microbiol.* 79:7696–7701.
- Dale, C., and N.A. Moran. 2006. Molecular interactions between bacterial symbionts and their hosts. *Cell* 126:453–465.
- Desirò, A., A. Salvioli, E.L. Ngonkeu, *et al.* 2014. Detection of a novel intracellular microbiome hosted in arbuscular mycorrhizal fungi. *ISME J.* 8:257–270.
- Desirò, A., A. Faccio, A. Kaech, M.I. Bidartondo, and P. Bonfante. 2015. *Endogone*, one of the oldest plant-associated fungi, host unique *Mollicutes*-related endobacteria. *New Phytol.* 205:1464–1472.
- Driessen, A.J.M., E.H. Manting, and C. van der Does. 2001. The structural basis of protein targeting and translocation in bacteria. *Nat. Struct. Biol.* 8:492–498.
- Engledow, A.S., E.G. Medrano, J.J. Lipuma, C.F. Gonzalez, and E. Mahenthiralingam. 2004. Involvement of a plasmid-encoded type IV secretion system in the plant tissue watersoaking phenotype of *Burkholderia cenocepacia*. *J. Bacteriol.* 186:6015–6024.
- Estrada-De Los Santos, P., P. Vinuesa, L. Martínez-Aguilar, A.M. Hirsch, and J. Caballero-Mellado. 2013. Phylogenetic analysis of *Burkholderia* species by multilocus sequence analysis. *Curr. Microbiol.* 67:51–60.
- Francetic, O., D. Belin, C. Badaut, and A.P. Pugsley. 2000. Expression of the endogenous type II secretion pathway in *Escherichia coli* leads to chitinase secretion. *EMBO J.* 19:6697–6703.
- Freeman, S., and R.J. Rodriguez. 1993. Genetic conversion of a fungal plant pathogen to a non-pathogenic, endophytic mutualist. *Science* 260:75–78.
- Frey-Klett, P., P. Burlinson, A. Deveau, M. Barret, M. Tarkka, and A. Sarniguet. 2011. Bacterial-fungal interactions: hyphens between agricultural, clinical, environmental, and food microbiologists. *Microbiol. Mol. Biol. Rev.* 75:583–609.
- Fujimura, R., A. Nishimura, S. Ohshima, Y. Sato, T. Nishizawa, K. Oshima, M. Hattori, K. Narisawa, and H. Ohta. 2014. Draft genome sequence of the betaproteobacterial endosymbiont associated with the fungus *Mortierella elongata* FMR23-6. *Genome Announc.* 2:e01272-14.
- Ghignone, S., A. Salvioli, I. Anca, *et al.* 2012. The genome of the obligate endobacterium of an AM fungus reveals an interphylum network of nutritional interactions. *ISME J.* 6:136–145.
- Gordon, D., C. Desmarais, and P. Green. 2001. Automated finishing with autofinish. *Genome Res.* 11:614–625.
- Guilhbert, M.R., and B.C. Kirkpatrick. 2005. Identification of *Xylella fastidiosa* antivirulence genes: hemagglutinin adhesins contribute a biofilm maturation to *X. fastidiosa* and colonization and attenuate virulence. *Mol. Plant Microbe Interact.* 18:857–868.
- Hoffman, M.T., and A.E. Arnold. 2010. Diverse bacteria inhabit living hyphae of phylogenetically diverse fungal endophytes. *Appl. Environ. Microbiol.* 76:4063–4075.
- Kanehisa, M., Y. Sato, and K. Morishima. 2016. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J. Mol. Biol.* 428:726–731.
- Kline, K.A., S. Fälker, S. Normark, S. Normark, and B. Henriques-Normark. 2009. Bacterial adhesins in host-microbe interactions. *Cell Host Microbe* 5:580–592.

28. Kobayashi, D.Y., and J.D. Palumbo. 2000. Bacterial endophytes and their effects on plants and uses in agriculture, p. 199–236. *In* C.W. Bacon, J.F. White (ed.), *Microbial Endophytes*. Marcel Dekker, New York.
29. Kobayashi, D.Y., and J.A. Crouch. 2009. Bacterial/fungal interactions: from pathogens to mutualistic endosymbionts. *Annu. Rev. Phytopathol.* 47:63–82.
30. Krzywinski, M., J. Schein, I. Birol, J. Connors, R. Gascoyne, D. Horsman, S.J. Jones, and M.A. Marra. 2009. Circos: an information aesthetic for comparative genomics. *Genome Res.* 19:1639–1645.
31. Lackner, G., and C. Hertweck. 2011. Impact of endofungal bacteria on infection biology, food safety, and drug development. *PLoS Pathog.* 7:5–8.
32. Lackner, G., N. Moebius, and C. Hertweck. 2011. Endofungal bacterium controls its host by an hrp type III secretion system. *ISME J.* 5:252–261.
33. Lackner, G., N. Moebius, L. Partida-Martinez, and C. Hertweck. 2011. Complete genome sequence of *Burkholderia rhizoxinica*, an endosymbiont of *Rhizopus microsporus*. *J. Bacteriol.* 193:783–784.
34. Lackner, G., N. Moebius, L.P. Partida-Martinez, S. Boland, and C. Hertweck. 2011. Evolution of an endofungal lifestyle: deductions from the *Burkholderia rhizoxinica* genome. *BMC Genomics* 12:210.
35. Lagesen, K., P. Hallin, E.A. Rødland, H.H. Stærfeldt, T. Rognes, and D.W. Ussery. 2007. RNAMmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* 35:3100–3108.
36. Lawrence, J.G., R.W. Hendrix, and S. Casjens. 2001. Where are the pseudogenes in bacterial. *Trends Microbiol.* 9:535–540.
37. Leone, M.R., G. Lackner, A. Silipo, R. Lanzetta, A. Molinaro, and C. Hertweck. 2010. An unusual galactofuranose lipopolysaccharide that ensures the intracellular survival of toxin-producing bacteria in their fungal host. *Angew. Chem., Int. Ed.* 49:7476–7480.
38. Li, Z., Q. Yao, S.P. Dearth, *et al.* 2017. Integrated proteomics and metabolomics suggests symbiotic metabolism and multimodal regulation in a fungal-endobacterial system. *Environ. Microbiol.* 19:1041–1053.
39. Lowe, T.M., and S.R. Eddy. 1996. TRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25:955–964.
40. McCutcheon, J.P., B.R. McDonald, and N.A. Moran. 2009. Origin of an alternative genetic code in the extremely small and GC-rich genome of a bacterial symbiont. *PLoS Genet.* 5:e1000565.
41. Millati, R., L. Edebo, and M.J. Taherzadeh. 2005. Performance of *Rhizopus*, *Rhizomucor*, and *Mucor* in ethanol production from glucose, xylose, and wood hydrolyzates. *Enzyme Microb. Technol.* 36:294–300.
42. Mondo, S.J., K.H. Toomer, J.B. Morton, Y. Lekberg, and T.E. Pawlowska. 2012. Evolutionary stability in a 400-million-year-old heritable facultative mutualism. *Evolution* 66:2564–2576.
43. Moran, N.A., J.P. McCutcheon, and A. Nakabachi. 2008. Genomics and evolution of heritable bacterial symbionts. *Annu. Rev. Genet.* 42:165–190.
44. Mosse, B. 1970. Honey-coloured, sessile *Endogone* spores. *Arch. Microbiol.* 70:167–175.
45. Mulley, G., M.L. Beeton, P. Wilkinson, *et al.* 2015. From insect to man: photorhabdus sheds light on the emergence of human pathogenicity. *PLoS One* 10:e0144937.
46. Naito, M., and T.E. Pawlowska. 2015. The role of mobile genetic elements in evolutionary longevity of heritable endobacteria. *Mobile Genet. Elem.* 6:e1136375.
47. Naito, M., and T.E. Pawlowska. 2016. Defying muller's ratchet: ancient heritable endobacteria escape extinction through retention of recombination and genome plasticity. *mBio* 7:1–8.
48. Nakabachi, A., A. Yamashita, H. Toh, H. Ishikawa, H.E. Dunbar, N.A. Moran, and M. Hattori. 2008. The 169-kilobase genome of the bacterial endosymbiont *Carsonella*. *Science* 314:267.
49. Ohshima, S., Y. Sato, R. Fujimura, Y. Takashima, M. Hamada, T. Nishizawa, K. Narisawa, and H. Ohta. 2016. *Mycosporidium 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.
50. Osborne, A.R., T.A. Rapoport, and B. van den Berg. 2005. Protein translocation by the Sec61/Secy channel. *Annu. Rev. Cell Dev. Biol.* 21:529–550.
51. Pandey, D.P., and K. Gerdes. 2005. Toxin-antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. *Nucleic Acids Res.* 33:966–976.
52. Partida-Martinez, L.P., C.F. de Looß, K. Ishida, M. Roth, K. Buder, and C. Hertweck. 2007. Rhizonin, the first mycotoxin isolated from the zygomycota, is not a fungal metabolite but is produced by bacterial endosymbionts. *Appl. Environ. Microbiol.* 73:793–797.
53. Partida-Martinez, L.P., I. Groth, I. Schmitt, W. Richter, M. Roth, and C. Hertweck. 2007. *Burkholderia rhizoxinica* sp. nov. and *Burkholderia endofungorum* sp. nov., bacterial endosymbionts of the plant-pathogenic fungus *Rhizopus microsporus*. *Int. J. Syst. Evol. Microbiol.* 57:2583–2590.
54. Partida-Martinez, L.P., and C.A. Hertweck. 2007. A gene cluster encoding rhizoxin biosynthesis in “*Burkholderia rhizoxina*”, the bacterial endosymbiont of the fungus *Rhizopus microsporus*. *ChemBioChem* 8:41–45.
55. Proschak, A., Q. Zhou, T. Schöner, A. Thanwisai, D. Kresovic, A. Dowling, E. Proschak, and H.B. Bode. 2014. Biosynthesis of the insecticidal xenoclyolins in *Xenorhabdus bovienii*. *ChemBioChem* 15:369–372.
56. Ribet, D., and P. Cossart. 2015. How bacterial pathogens colonize their hosts and invade deeper tissues. *Microbes Infect.* 17:173–183.
57. Ruiz-Herrera, J., C. León-Ramírez, A. Vera-Núñez, A. Sánchez-Arreguín, R. Ruiz-Medrano, H. Salgado-Lugo, L. Sánchez-Segura, and J. José Peña-Cabriles. 2015. A novel intracellular nitrogen-fixing symbiosis made by *Ustilago maydis* and *Bacillus* spp. *New Phytol.* 207:769–777.
58. Sato, Y., K. Narisawa, K. Tsuruta, M. Umezui, T. Nishizawa, K. Tanaka, K. Yamaguchi, M. Komatsuzaki, and H. Ohta. 2010. Detection of betaproteobacteria inside the mycelium of the fungus *Mortierella elongata*. *Microbes Environ.* 25:321–324.
59. Sato, Z., T. Noda, I. Matsuda, S. Iwasaki, H. Kobayashi, J. Furukawa, and S. Okuda. 1983. Studies on rhizoxin, a phytotoxin produced by *Rhizopus chinensis* causing rice seedling blight. *Ann. Phytopathol. Soc. Jpn.* 49:128.
60. Scannnerini, S., and P. Bonfante-Fasolo. 1991. Bacteria and bacteria-like objects in endomycorrhizal fungi (*Glomeceae*), p 273–287. *In* L. Margulis and R. Fester (ed.), *Symbiosis as source of evolutionary innovation: speciation and morphogenesis*. The MIT Press, London.
61. Sharma, M., M. Schmid, M. Rothballer, *et al.* 2008. Detection and identification of bacteria intimately associated with fungi of the order *Sebacinales*. *Cell Microbiol.* 10:2235–2246.
62. Shigenobu, S., H. Watanabe, M. Hattori, Y. Sakaki, and H. Ishikawa. 2000. Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. *APS. Nature* 407:81–86.
63. Spatafora, J.W., Y. Chang, G.L. Benny, *et al.* 2016. A phylum-level phylogenetic classification of zygomycete fungi based on genome-scale data. *Mycologia* 108:1028–1046.
64. Stevens, M.P., A. Haque, T. Atkins, *et al.* 2004. Attenuated virulence and protective efficacy of a *Burkholderia pseudomallei* bsa type III secretion mutant in murine models of melioidosis. *Microbiology* 150:2669–2676.
65. Tarkka, M.T., A. Sarniguet, and P. Frey-Klett. 2009. Inter-kingdom encounters: recent advances in molecular bacterium-fungus interactions. *Curr. Genet.* 55:233–243.
66. Toomer, K.H., X. Chen, M. Naito, S.J. Mondo, H.C. Bakker, N.W. VanKuren, Y. Lekberg, J.B. Morton, and T.E. Pawlowska. 2015. Molecular evolution patterns reveal life history features of mycoplasma-related endobacteria associated with arbuscular mycorrhizal fungi. *Mol. Ecol.* 24:3485–3500.
67. Tsuruo, T., T. Oh-hara, H. Iida, *et al.* 1986. Rhizoxin, a macrocyclic lactone antibiotic, as a new antitumor agent against human and murine tumor cells and their vincristine-resistant sublines. *Cancer Res.* 46:381–385.
68. Uehling, J., A. Gryganskyi, K. Hameed, *et al.* 2017. Comparative genomics of *Mortierella elongata* and its bacterial endosymbiont *Mycosporidium cysteinexigens*. *Environ. Microbiol.* 19:2964–2983.
69. Wang, G.C.Y., and Y. Wang. 1997. Frequency of formation of chimeric molecules as a consequence of PCR coamplification of 16S rRNA genes from mixed bacterial genomes. *Appl. Environ. Microbiol.* 63:4645–4650.
70. Whitfield, C. 2006. Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli*. *Annu. Rev. Biochem.* 75:39–68.
71. Zhou, Y., Y. Liang, K.H. Lynch, J.J. Dennis, and D.S. Wishart. 2011. PHAST: a fast phage search tool. *Nucleic Acids Res.* 39:347–352.