A newly recorded Rickettsia of the Torix group is a recent intruder and an endosymbiont in the whitefly Bemisia tabaci

Hua-Ling Wang [1,2†] Teng Lei,[1†] Xiao-Wei Wang,[1] M. N. Maruthi,[2] Dan-Tong Zhu,[1] Stephen L. Cameron,[3] Qiong Rao,[4] Hong-Wei Shan,[1] John Colvin,[2] Yin-Quan Liu [1] and Shu-Sheng Liu[1*]

[1]*The Ministry of Agriculture Key Laboratory of Molecular Biology of Crop Pathogens and Insects, Institute of Insect Sciences, Zhejiang University, 866 Yuhangtang Road, Hangzhou, 310058, China.*

[2]*Natural Resources Institute, University of Greenwich, Kent, ME4 4TB, UK.*

[3]*Department of Entomology, Purdue University, 901 West State Street, West Lafayette, IN, 479074, USA.*

[4]*School of Agriculture and Food Science, Zhejiang A & F University, Lin'an, 311300, Zhejiang, China.*

Summary

The bacterium *Rickettsia* is found widely in phytopha-gous insects and often exerts profound effects on the phenotype and fitness of its hosts. Here, we decrypt a new, independent, phylogenetically ancient Torix *Rick-ettsia* endosymbiont found constantly in a laboratory line of an economically important insect Asia II 7, a putative species of the *Bemisia tabaci* whitefly complex (Hemiptera: Aleyrodidae), and occasionally in field whitefly populations. This new *Rickettsia* distributes throughout the body of its whitefly host. Genetically, compared to *Rickettsia*_bellii_MEAM1 found earlier in whiteflies, the new *Rickettsia* species has more gene families and pathways, which may be important factors in shaping specific symbiotic relationships. We propose the name '*Candidatus* Rickettsia_Torix_*Bemisia_tabaci* (RiTBt)' for this new endosymbiont associated with whiteflies. Comparative genomic analyses indicate that RiTBi may be a relatively recent intruder in whiteflies given its low abundance in the field and relatively larger genome compared to *Rickettsia*_bellii_MEAM1.

Introduction

Plant sap-sucking insects, such as psyllids, aphids, and mealybugs, are members of the Hemiptera sub-order Sternorrhyncha and have obligatory as well as facultative associations with a diverse range of prokaryotic endo-symbionts (Baumann, 2005). Many studies have shown the vital roles of various bacteria in the biology of their hosts, such as synthesis of essential amino acids (Douglas, 2006), tolerance to high temperatures (Montllor *et al*., 2002), and facilitating virus transmission (Gottlieb *et al*., 2010).

The whitefly *Bemisia tabaci* (Gennadius) (Hemiptera: Aleyrodidae) is a cryptic species complex with at least 36 putative species (Liu *et al*., 2012; Boykin, 2014). They are distributed mostly in the tropics and subtropics, and some of them, such as Mediterranean (MED), Mid-dle East-Asia Minor 1 (MEAM1), Sub-Saharan Africa 1 (SSA1) and Asia II 1, are important agricultural pests (Liu *et al*., 2007). Whiteflies of the *B. tabaci* species complex share a long-term and intimate association with bacterial symbionts. So far, one obligate endosymbiont, '*Candidatus* Portiera aleyrodidarum' (Oceanospirillales), (hereafter *Portiera*), has been recognized and is local-ized in the bacteriome, a tissue composed of bacterio-cytes (Zchori-Fein and Brown, 2002). *Portiera* is essential to the synthesis of essential amino acids and carotenoids for its hosts and is transmitted vertically via the female ovary (Gottlieb *et al*., 2006; Luan *et al*., 2016). Seven genera of facultative endosymbionts have been found, including *Arsenophonus* spp. (Enterobacteriales) (Zchori-Fein and Brown, 2002), '*Can-didatus* Cardinium hertigii' (Bacteroidales) (Weeks *et al*., 2003), '*Candidatus* Fritschea bemisiae' (Chlamydiales) (Everett *et al*., 2005), '*Candidatus* Hamiltonella defensa' (Enterobacteriales) (Zchori-Fein and Brown, 2002), *Hemipteriphilus asiaticus*/*OLO* (Rickettsiales) (Bing *et al*., 2013), *Wolbachia* spp. (*Rickettsiales*) (Nirgianaki *et al*., 2003), and *Rickettsia* spp. (*Rickettsiales*)(Gottlieb *et al*., 2006). Although many facultative endosymbionts have been found, the functional properties of most of them are barely known.

One of these facultative endosymbionts, *Rickettsia* (class: Alpha-proteobacteria; order: Rickettsiales) contains a wide diversity of species. Phylogenomic analyses indicate the *Rickettsia* can be divided into at least four groups: ancestral group (AG), typhus group (TG), transitional group (TRG), and spotted fever group (SFG) (Gillespie *et al*., 2008). Based on partial sequence analysis, Weinert *et al*. (2009) suggested further delineations of the genus into 12 groups: Hydra, Tundra, Torix, Rhizobius, Meloidae, Bellii, Onychiurus, Adalia, Canadensis, Spotted Fever, Typhus, and Transitional, with *Orientia tsutsugamushi* as the outgroup. Up to now, over 60 *Rickettsia* genomes have been published from a variety of host systems. Most of these belong to Bellii, Canadensis, Spotted Fever, Typhus, and Transitional group, with only one to the Torix group (Pilgrim *et al*., 2017).

A notable characterization of *Rickettsia* is that some species are pathogenic to vertebrates (Weinert *et al*., 2009), such as *R. prowazekii*, *R. rickettsii*, and *R. conorii*. *Rickettsia* have also been detected in a wide variety of invertebrates and have profound effects on the biology of the hosts and vector–pathogen interactions (Pilgrim *et al*., 2017). As for *B. tabaci*, only *Rickettsia*_bellii_MEAM1 has been studied in some detail. Two localization patterns in the hosts have been recognized: 'scattered', where the bacterium distributes throughout the whitefly hemocoel, excluding the bacteriocytes; and 'confined', where the bacterium is restricted to the bacteriocytes (Caspi-Fluger *et al*., 2011). *Rickettsia*_bellii_MEAM1 with the 'scattered' distribution pattern in MEAM1 is transmitted horizontally through plants (Caspi-Fluger *et al*., 2012) and can enhance heat tolerance of the host (Brumin *et al*., 2011), increase fecundity (Himler *et al*., 2011), and promote virus transmission (Ghanim, 2016). Collectively, these properties make *Rickettsia* biologically significant for its host, yet so far, the other relevant *Rickettsia* species has been largely understudied.

Here, we firstly describe a novel Torix *Rickettsia* species that has a symbiotic relationship with species Asia II 7 of the *B. tabaci* whitefly complex. A PacBio 16S screening project identified a new 16S rRNA allied to a novel *Rickettsia* and its association with a specific host. We then investigated several aspects of the biology of this new *Rickettsia* including its field distribution, host range, *in vivo* localization, taxonomic position, and genomic properties. Finally, we sought to compare this new *Rickettsia* symbiont with the different *Rickettsia* previously described in *B. tabaci*. The results showed species-level divergence between this new *Rickettsia* and other *Rickettsia* recorded previously. We, thus, propose the name '*Candidatus* Rickettsia_Torix_*Bemisia_tabaci* (RiTBt)' for this new species of *Rickettsia*. Our analysis adds new insights into microbial community residing in the *B. tabaci* species complex and their diverse interactions with hosts.

Results

Discovery of RiTBt

In a comprehensive screening for facultative endosymbionts of the *B. tabaci* whitefly complex using PacBio 16S, a *Rickettsia*-like bacterial *16S rRNA* gene was consistently obtained from a laboratory colony of Asia II 7. Bayesian analysis based on the *16S rRNA* gene showed that this new *Rickettsia* was not clustered with the previously identified *Rickettsia*_bellii_MEAM1 but rather branched between *OLO_Bemisia-China1_JX042443* (*Hemipteriphilus asiaticus* from *B. tabaci* China 1) and the remaining *Rickettsia* spp. (Fig. 1.A). Additional analyses of the protein-coding genes *gltA* and *groEL* yielded similar phylogenetic relationships (Fig. 1.B and C). The nucleotide sequence identities of *16S rRNA*, *gltA*, and *groEL* genes between RiTBt and *Rickettsia*_bellii_MEAM1 are 96% (1442 bp), 78% (481 bp) and 78% (1646 bp) respectively. Based on these results, we propose that the new *Rickettsia* is a new species residing in *B. tabaci* and named it as '*Candidatus* Rickettsia_Torix_*Bemisia_tabaci*' (RiTBt).

Prevalence of RiTBt in the field and a laboratory whitefly population

To examine the geographical distribution of RiTBt and its association with various species of the *B. tabaci* complex, we screened 512 adults that were collected from major crop plants and vegetables such as cotton, sweet potato, pumpkin, eggplant, cucumber, and tomato, in 22 locations of 14 provinces in China (Fig. 2. and Supporting Information Table S1). As *B. tabaci* is a cryptic species complex, and more than one species may occur on the same host plant within a field block, we firstly established the species status of whiteflies by individual *mtCOI* barcoding before running the diagnostic PCR for RiTBt. Collected specimens were assigned to eight *B. tabaci* cryptic species: MED, MEAM1, China 1, China 2, Asia I, Asia II 1, Asia II 6 and Asia II 7 (Fig. 2. and Supporting Information Table S1). Five of the 512 adults examined were RiTBt-positive and were all Asia II 1 collected from plants of *Ipomoea batatas* and *Brassica chinensis* in Zhaoqing, Guangdong province (Fig. 2. and Supporting Information Table S1). In contrast, in an Asia II 7 laboratory population, all individuals tested were infected with RiTBt (46 of 46, 100%). We then conducted the test consecutively for five generations of the whitefly host, with 30 adults in each generation, and found that all individuals tested in each of the generations were infected with RiTBt. Of the eight screened

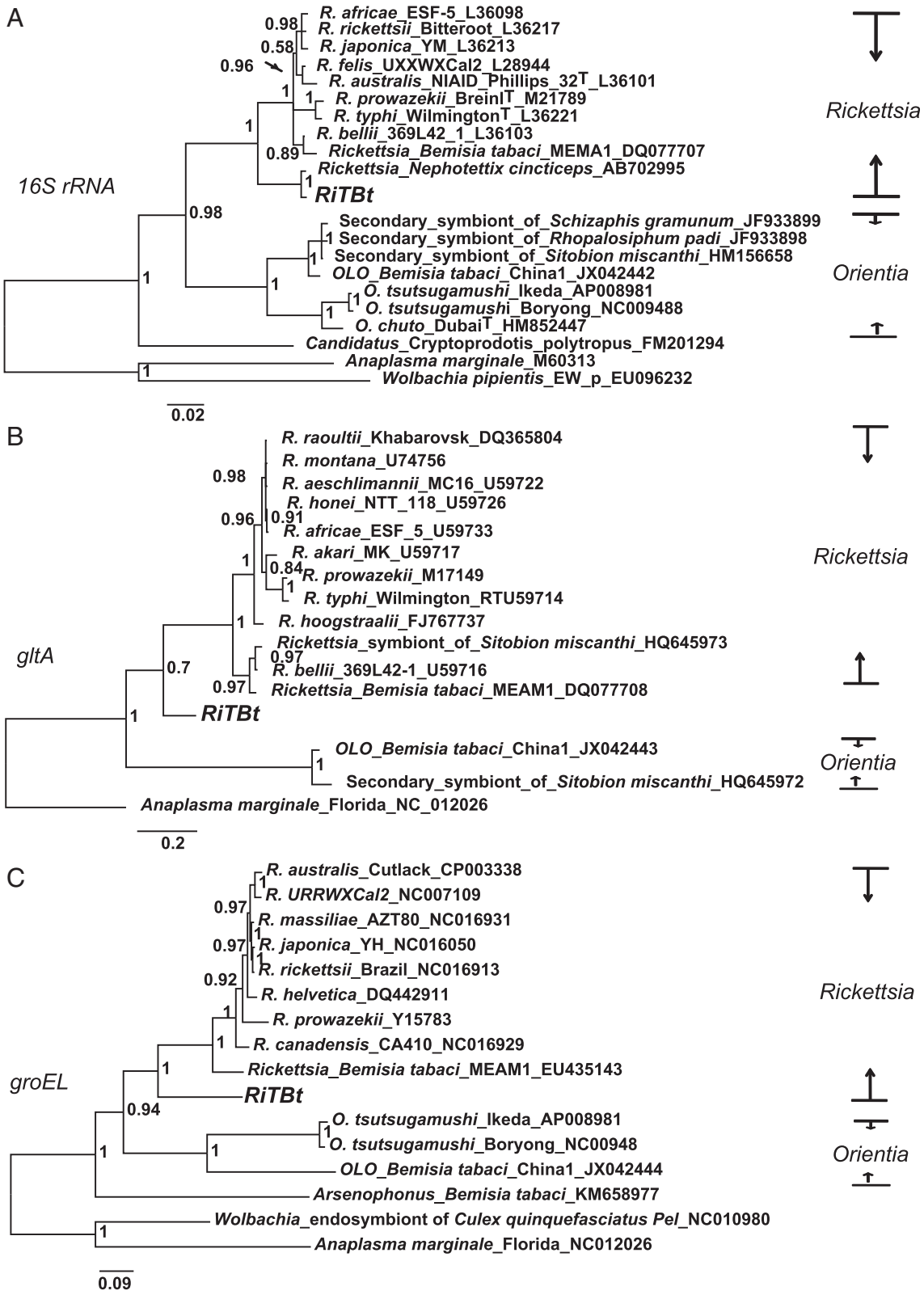


Fig. 1. Phylogenetic relationship of RiTBt with other *Rickettsia* based on bacterial *16SrRNA*(A)/*gltA*(B)/*groEL*(C) gene sequences. Posterior probabilities for the Bayesian phylogeny at 50% or higher are shown at the nodes.

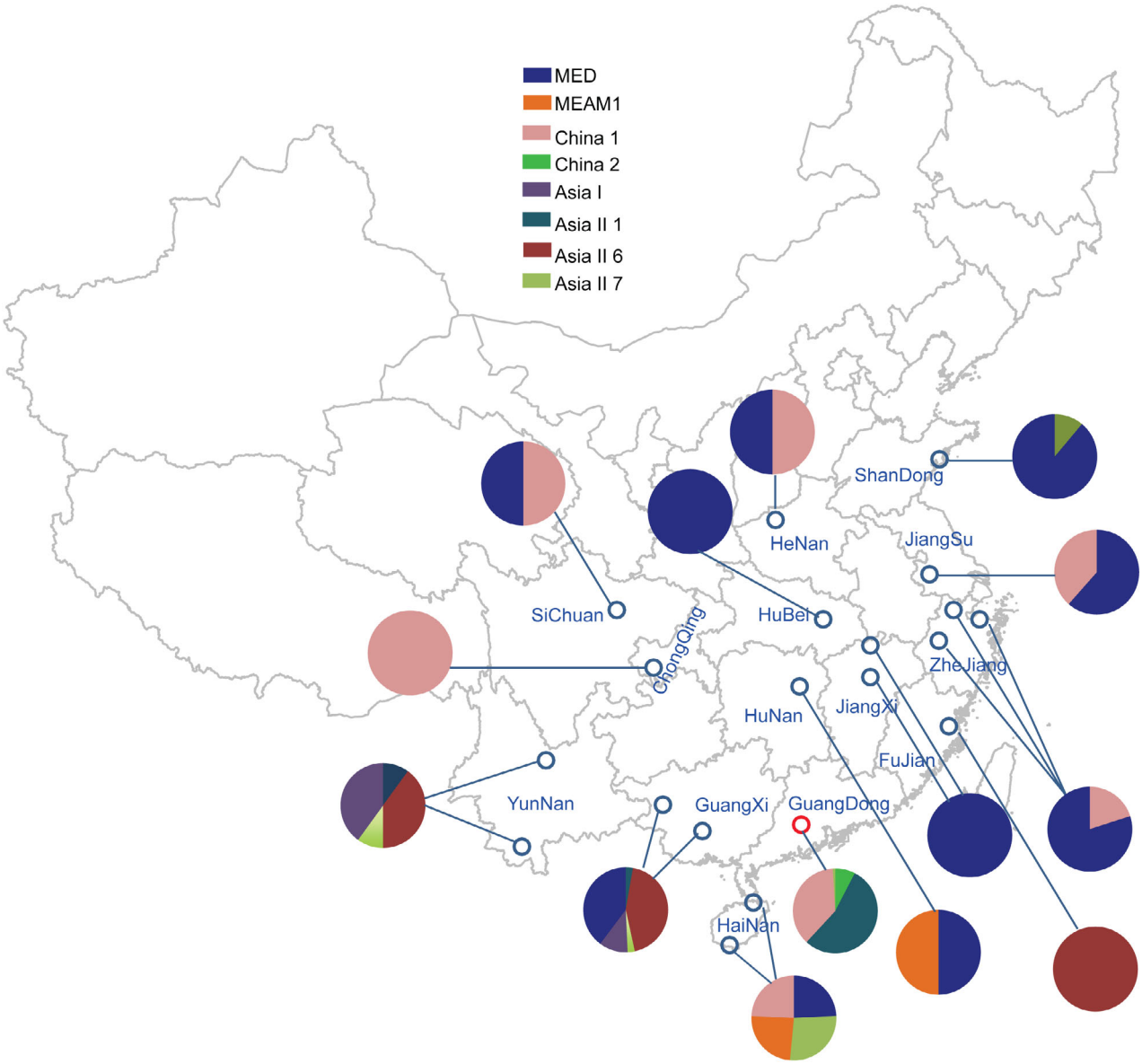


Fig. 2. Distribution of RiTBt in the field. The small red circle indicates the location where RiTBt was found. Each of the pie charts indicates the relative proportion of whitefly species in the *B. tabaci* species complex for each province.

whitefly species, only Asia II 7 and Asia II 1 harboured RiTBt. These two host species are placed in the same clade (Boykin, 2014) based on a recent extensive molecular phylogenetic analysis of the *B. tabaci* cryptic species, suggesting that the new *Rickettsia* is restricted to a specific species lineage of *B. tabaci* complex.

In vivo localization of RiTBt in whitefly host

Localizations of RiTBt were examined by fluorescence *in situ* hybridization (FISH) in oocytes, nymphs, and adults of Asia II 7 *B. tabaci*. RiTBt was detected inside the bacteriocyte of oocytes, and throughout the whole body of the fourth instar nymphs, as well as females and males of Asia II 7 (Fig. 3.), but was not detected in the negative control of MED individuals (Supporting Information Fig. S1A–D). Notably, RiTBt signals were concentrated in the anterior pole of the oocytes (Fig. 3. A), indicating that RiTBt can be transmitted to progeny via oocytes. In terms of whitefly organs, RiTBt was visualized in bacteriocytes, midguts, salivary glands, ovaries, and testes (Fig. 4.A–F) of Asia II 7 adults but not in MED (Supporting Information Fig. S1E–I). The presence of RiTBt in the above organs was confirmed by the transmission electron microscope (TEM) (Fig. 4.A5, B4, C4, D5, and E5).

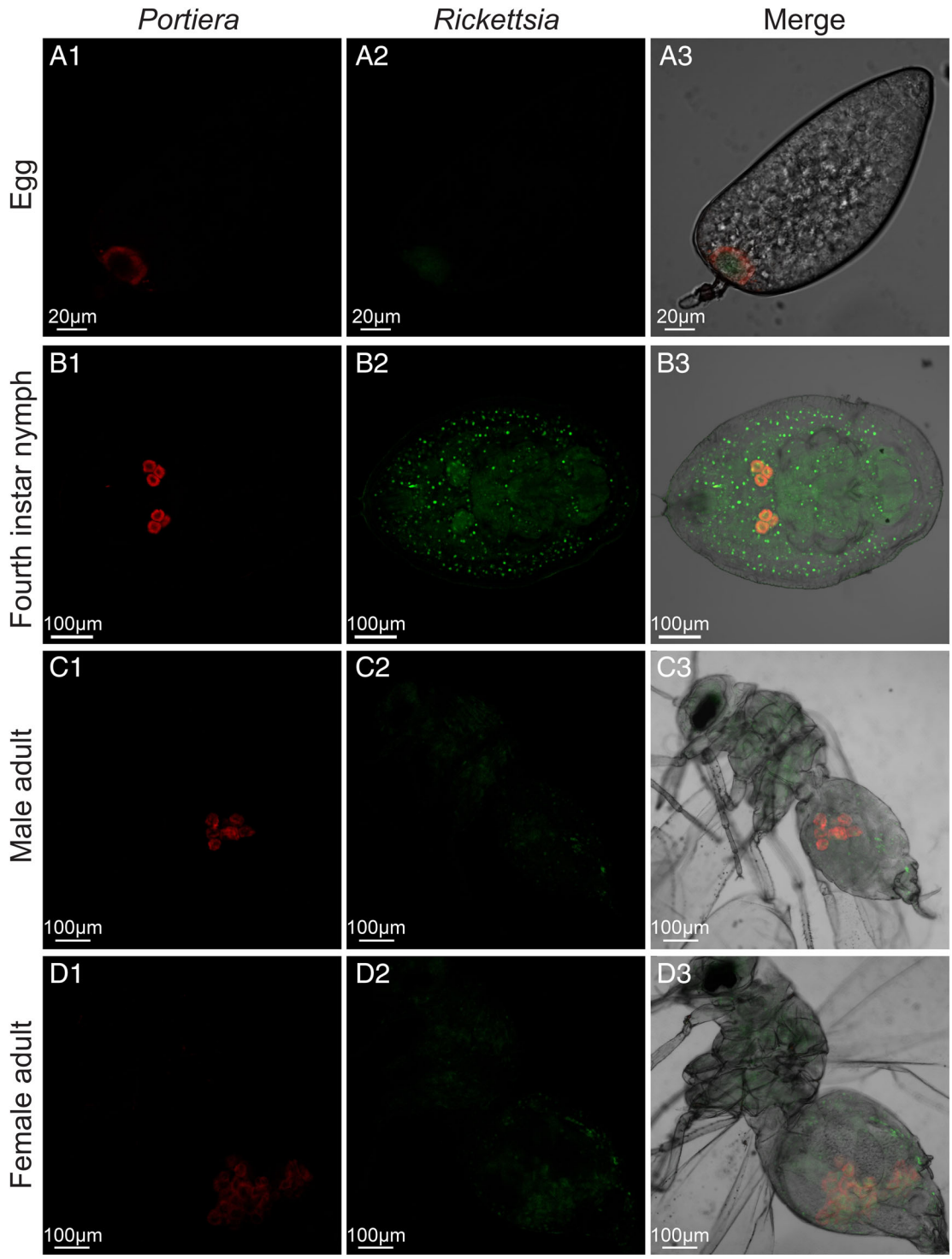


Fig. 3. FISH of RiTBt in *B. tabaci* Asia II 7. *Portiera*-specific probe (red) conjugated to cy3, RiTBt-specific probe (green) conjugated to cy5 were used. A1-A3, Egg; B1-B3, Fourth instar nymph; C1-C3, Male adult; D1-D3, Female adult. For the columns, A1, B1, C1, D1, *Portiera* Channel; A2, B2, C2, D2, 'RiTBt' channel; A3, B3, C3, D3, overlay of *Portiera* and 'RiTBt' on bright field channels.

Genome and the taxonomy of RiTBt

The draft genome of RiTBt has a length of 1,485, 811 bp (Table 1), which is slightly larger than that of Torix group member, named RiCNE (Pilgrim *et al*., 2017), previously sequenced from *Culicoides newsteadi* (1,456, 172 bp), and *Rickettsia*_bellii_MEAM1 from the whitefly MEAM1

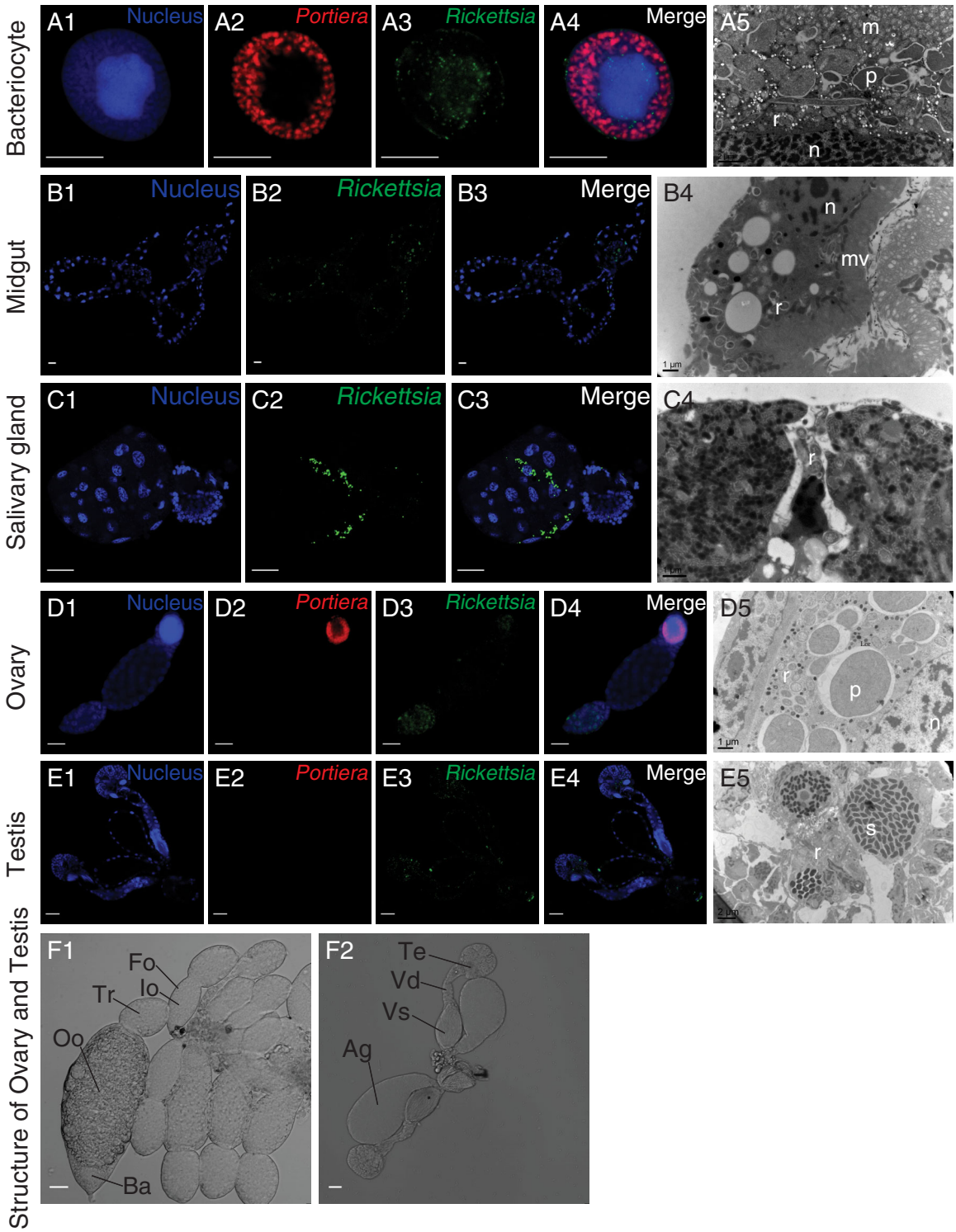


Fig. 4. FISH and TEM show the existence of RiTBt in bacteriocyte, midgut, salivary gland, ovary and testis of Asia II 7. *Portiera*-specific probe (red) conjugated to cy3 and RiTBt-specific probe (green) conjugated to cy5 were used, and nuclei were stained with DAPI (blue). FISH images: A1-A4, bacteriocyte; B1-B3, midgut; C1-C3, salivary gland; D1-D4, ovary; and E1-E4, testis. F1-F2, germinal system of *B. tabaci*. F1, the structure of ovary; F2, the structure of male reproductive system. Fo, follicular cells; Io, immature oocyte; Tr, trophocytes; Oo, oocyte; Ba, bacteriocyte; Te, testis; Vd, vas deferens; Vs, vesicular siminalis; Ag, accessory glands. TEM images: A5, bacteriocyte; B4, midgut; C4, salivary gland; D5, bacteriocyte in ovary, and E5, testis. For the columns, A1, B1, C1, D1 and E1, 'DAPI' channel; A2, D2 and E2, '*Portiera*' channel; A3, B2, C2, D3 and E3, 'RiTBt' channel; A4, B3, C3, D4 and E4, overlay of 'DAPI', '*Portiera* aleyrodidarum' and 'RiTBt' on bright field channels. Scale bar 20 μm. TEM images: A5, bacteriocyte; B4, midgut; C3, salivary gland; D5, bacteriocyte in ovary; E5, testis. In TEM images, r represents *Rickettsia*; p represents *Portiera*; m represents mitochondria; n represents the nucleus; mv represents microvilli; s represents sperm.

Table 1. General statistics and features of RiTBt.

	E-value <=1e-5
Total number of scaffolds	69
Total size (bp)	1,485, 811
N50 length (bp)	50, 092
Largest contig (bp)	157, 737
Total predicted coding sequences	1,446
Average length (bp) of coding sequences	813
Coding density	0.973 gene per kb
GC content (%)	33.3
tRNA	40
rRNA	3 (5S, 16S, 23S)

(1,378, 618 bp). The draft genome has a 33.3% GC content and comprises 69 scaffolds with an N50 size of approximately 50, 029 bp (Table 1). Benchmarking Universal Single-Copy Orthologues (BUSCO) analyses revealed 90.6% complete and/or fragmented BUSCOs in the RiTBt genome assembly [In all, there were 130 complete BUSCOs (C; 87.9%), 124 complete and single-copy BUSCOs (S; 83.8%), 6 complete and duplicated BUSCOs (D; 4.1%), 4 fragmented BUSCOs (F; 2.7%), 14 missing BUSCOs (M; 9.4%), and 148 BUSCO groups.] Totally, 1,446 genes were predicted, in which 1,219 ORFs were annotated by a homology search (Supporting Information Table S2). Of the 1,219 ORFs, 728 genes were assigned to COGs (Supporting Information Table S2 and Fig. S2), five genes were found encoding a tra conjugative DNA transfer element (Supporting Information Table S3) (Ogata *et al*., 2006), and 14 genes were annotated as coding Vir components (virB4, virB6, virB8-virB11, virD4) of P-like type IV secretion system (P-T4SS), which are highly conserved among Rickettsiales (Supporting Information Table S3). Similar to other *Rickettsia* genomes, the vir genes were dispersed into three major clusters (scaffold 4: virB4 and virB6; scaffold 10: virB4, virB8-B11, and virD4; scaffold 39: virB4 and virB6) (Gillespie *et al*., 2009).

In total, 519 core and 459 single-copy protein-coding genes (Supporting Information Fig. S3) were identified among all the 64 complete or draft *Rickettsia* genomes available (Supporting Information Table S4). Comparison of RiTBt single-copy genes with the 63 other *Rickettsia* genomes showed that RiTBt has the highest nucleotide similarity to that of the Torix group endosymbiont of RiCNE (on average: 87% nucleotide identity), but a relatively low nucleotide similarity to those of other *Rickettsia* species (on average: 66%–68% nucleotide identity) (Supporting Information Fig. S4).

To infer a robust phylogenetic relationship of RiTBt relative to other Rickettsiaceae, a Bayesian phylogenomic tree was constructed based on the concatenated alignment of 459 single-copy nuclear genes from the genomes of 64 *Rickettsia* and its sister genus *Orientia* (Supporting Information Table S5). The resulting phylogenomic tree differs from a previous one proposed by Weinert *et al*. (2009) and provides a global phylogeny of *Rickettsia* spp. with high bootstrap support for most backbone nodes. RiTBt clustered with RiCNE and was highly divergent from other groups of *Rickettsia* spp. (Fig. 5.). Moreover, RiTBt and RiCNE are situated at the most ancient lineage in all *Rickettsia* spp. genomes. Although RiTBt formed a distinct and robust monophyletic Torix clade with RiCNE, its species status with respect to RiCNE is not clear. RiCNE and RiTBt share a 99% sequence identity for the *16S rRNA* gene (1,461 bp), a widely used marker for species delimitation, but this marker does not exhibit sufficiently high resolution for detailed classification (Richter and Rosselló-Móra, 2009). To further classify its species status, average nucleotide identity, which has been widely used to determine species boundaries with cutoffs of 94%–96%, was calculated (Konstantinidis and Tiedje, 2005; Richter and Rosselló-Móra, 2009). The average nucleotide identity between RiTBt and RiCNE was 86.25%, notably lower than the 94%–96% criterion for separating species. All the evidence collected here further confirms that RiTBt is a new species under the Torix group.

Gene family expansion or contraction

Investigation of gene families (orthologous groups) across *Rickettsia* lineages showed that gene gains were scarce for most *Rickettsia* lineages, with 0–16 gene family gains at the base of each species-group. In contrast, gene losses were far more common, with 11–46 gene family decays in all groups (Fig. 5.). The most extensive gene family decays occur in the *Rickettsia_bellii*_MEAM1 (46 gene families lost), whereas RiTBt shows less than half this volume of gene family decays (22 gene families lost) (Fig. 5.). This finding confirmed that gene loss is the leading cause of genome diversification within the *Rickettsia* genus (Blanc *et al*., 2007).

Genome synteny and functional profile between 'Torix' and 'Bellii' group species

We compared genome synteny between the two Torix group species, RiTBt and RiCNE [MWZE00000000], to the two Bellii group species, RML369-c [CP000849] and MEAM1 [CP000849]. The scaffolds of RiTBt displayed little synteny with those of either the Bellii group genomes or RiCNE (Fig. 6.A).

RiTBt and RiCNE under the Torix group share similar metabolic pathways, which further support the close relationships between the two. In contrast, comparison of RiTBt with *Rickettsia*_bellii_MEAM1 showed the absence of several crucial pathways in the latter, such as

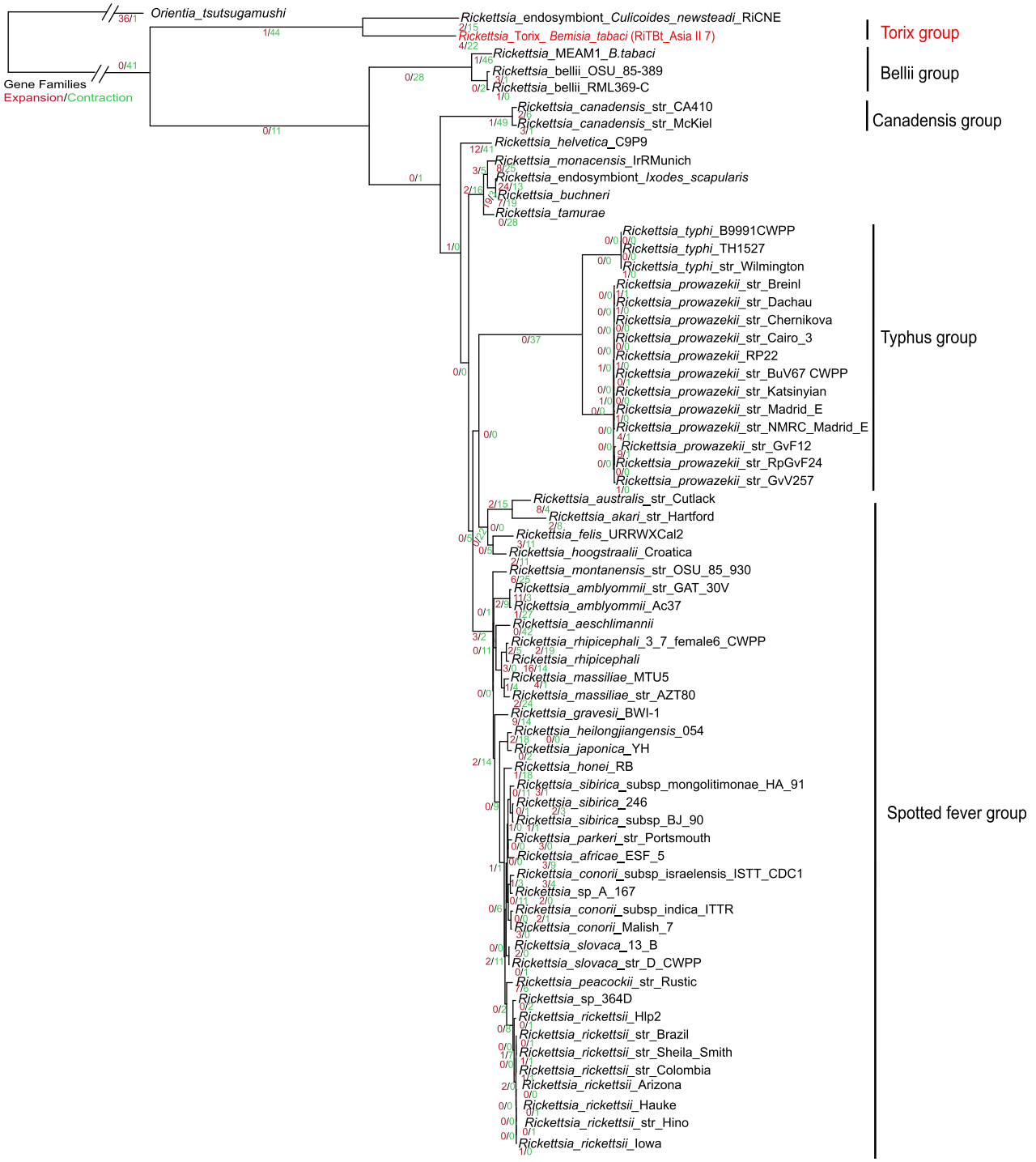


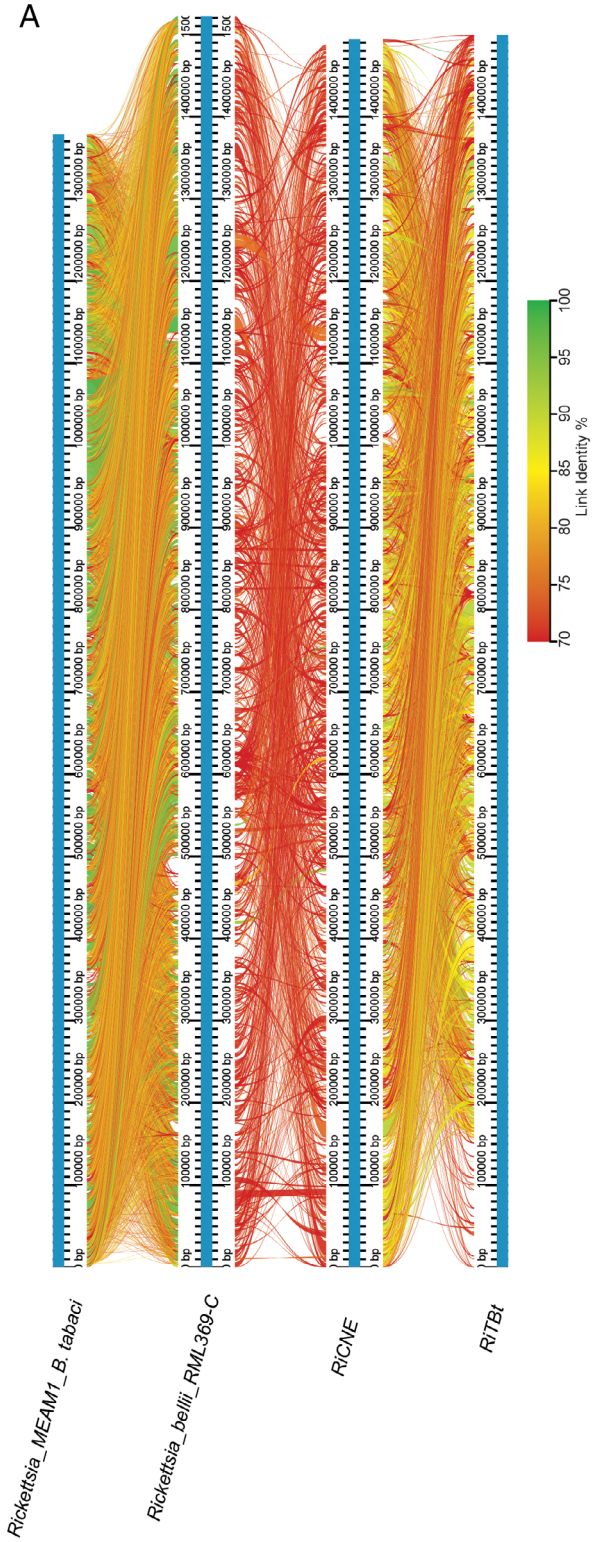
Fig. 5. Gene gain and loss of all known *Rickettsia* genomes and the phylogenetic position of RiTBt strain (red) based on 459 single-copy nuclear genes. The posterior probabilities are all over 0.85 except for the node (0.66) between the *Rickettsia_ prowazekii*_str_Breinl and the *Rickettsia_ prowazekii*_str_Dachau.

'Glycolysis (Embden-Meyerhof pathway), glucose', 'Glycolysis, core module involving three-carbon compounds", 'Methionine salvage pathway', 'Pentose phosphate pathway (Pentose phosphate cycle)' and 'Pentose phosphate pathway, non-oxidative phase, fructose 6P' (Fig. 6.B).

Discussion

Here, we report a bacterium residing in the two indigenous species Asia II 1 and Asia II 7 of the *B. tabaci* whitefly complex that has not been recorded previously.

A



B

Name	R_MEMA1	RiTBt	RiCNE
3-Hydroxypropionate bi-cycle	2	1	2
Adenine ribonucleotide biosynthesis, IMP	2	2	2
Ascorbate biosynthesis, animals, glucose-1P	1	1	1
Ascorbate biosynthesis, plants, glucose-6P	2	6	6
beta-Oxidation	0	1	0
C10-C20 isoprenoid biosynthesis	4	4	4
C1-unit interconversion	4	4	4
C4-dicarboxylic acid cycle, NAD - malic enzyme type	1	1	1
C4-dicarboxylic acid cycle, NADP - malic enzyme type	2	2	2
C5 isoprenoid biosynthesis, mevalonate pathway	2	2	1
C5 isoprenoid biosynthesis, non-mevalonate pathway	1	1	1
CAM (Crassulacean acid metabolism), dark	1	1	1
CAM (Crassulacean acid metabolism), light	2	2	2
Citrate cycle (TCA cycle, Krebs cycle)	7	7	7
Citrate cycle, first carbon oxidation, oxaloacetate	3	3	3
Citrate cycle, second carbon oxidation, 2-oxoglutarate	4	4	4
CMP-KDO biosynthesis	3	3	3
Coenzyme A biosynthesis, pantothenate	2	2	2
Cysteine biosynthesis, methionine	0	1	1
D-galactonate degradation, De Ley-Doudoroff pathway, D-galactonate	0	1	0
Dicarboxylate-hydroxybutyrate cycle	4	4	3
dTDP-L-rhamnose biosynthesis	3	1	0
Ectoine biosynthesis, aspartate	2	2	2
Ethylene biosynthesis, methionine	0	1	1
Ethylmalonyl pathway	3	3	1
Fatty acid biosynthesis, elongation	4	4	4
Fatty acid biosynthesis, initiation	2	2	2
Formaldehyde assimilation, serine pathway	2	2	2
GABA (gamma-Aminobutyrate) shunt	0	0	1
Galactose degradation, Leloir pathway, galactose	0	1	1
Gluconeogenesis, oxaloacetate	0	4	2
Glucuronate pathway (uronate pathway)	2	2	2
Glycolysis (Embden-Meyerhof pathway), glucose	0	5	4
Glycolysis, core module involving three-carbon compounds	0	3	2
Glyoxylate cycle	3	3	3
Guanine ribonucleotide biosynthesis IMP	2	2	2
Heme biosynthesis, glutamate	6	6	6
Hydroxypropionate-hydroxybutylate cycle	1	1	0
Incomplete reductive citrate cycle, acetyl-CoA	2	2	2
Inosine monophosphate biosynthesis, PRPP + glutamine	9	9	9
Inositol phosphate metabolism, Ins(1,3,4,5)P4	1	1	1
Isoleucine biosynthesis, threonine	2	2	2
Ketone body biosynthesis, acetyl-CoA	1	1	0
Leucine degradation, leucine	1	1	1
Lipopolysaccharide biosynthesis, inner core	1	1	1
Lipopolysaccharide biosynthesis, KDO2-lipid A	7	7	7
Lysine biosynthesis, acetyl-DAP pathway, aspartate	5	5	4
Lysine biosynthesis, DAP aminotransferase pathway, aspartate	5	5	4
Lysine biosynthesis, DAP dehydrogenase pathway, aspartate	4	4	4
Lysine biosynthesis, succinyl-DAP pathway, aspartate	7	7	6
Lysine degradation, lysine	16	20	16
Menaquinone biosynthesis, chorismate	1	1	1
Methionine biosynthesis, apartate	2	2	2
Methionine degradation	1	2	2
Methionine salvage pathway	0	6	6
Methylaspartate cycle	4	4	4
Ornithine biosynthesis, glutamate	1	1	0
Pantothenate biosynthesis, valine/L-aspartate	1	1	1
Pentose phosphate pathway (Pentose phosphate cycle)	0	10	10
Pentose phosphate pathway, non-oxidative phase, fructose 6P	0	4	4
Phosphatidylethanolamine (PE) biosynthesis, PA	3	3	3
Photorespiration	1	1	1
Pimeloyl-ACP biosynthesis, BioC-BioH pathway, malonyl-ACP	4	4	4
Propanoyl-CoA metabolism, propanoyl-CoA	1	1	1
PRPP biosynthesis, ribose 5P	1	1	1
Pyrimidine deoxyribonuleotide biosynthesis, CDP/CTP	6	7	6
Pyrimidine ribonucleotide biosynthesis, UMP	3	2	2
Pyruvate oxidation, pyruvate	1	1	1
Reductive acetyl-CoA pathway (Wood-Ljungdahl pathway)	1	1	1
Reductive citrate cycle (Arnon-Buchanan cycle)	26	26	26
Reductive pentose phosphate cycle (Calvin cycle)	0	4	3
Reductive pentose phosphate cycle, glyceraldehyde-3P	0	3	3
Reductive pentose phosphate cycle, ribulose-5P	0	1	0
Semi-phosphorylative Entner-Doudoroff pathway, gluconate	0	1	0
Tetrahydrofolate biosynthesis, GTP	9	11	11
Threonine biosynthesis, aspartate	2	2	2
Triacylglycerol biosynthesis	1	1	1
Ubiquinone biosynthesis, eukaryotes, 4-hydroxybenzoate	1	1	1
Ubiquinone biosynthesis, prokaryotes, chorismate	6	6	6
Valine/isoleucine biosynthesis, pyruvate	1	1	1

Fig. 6. Genome synteny and functional profile between 'Torix' and 'Bellii' group species. A. Alignment of RiTBt, *Rickettsia*_MEAM1_*B. tabaci*, RiCNE and *Rickettsia* bellii genomes. B. Assessment of the metabolic potential of RiTBt genome and comparison to the *Rickettsia*_bellii_MEAM1 and RiCNE. R_MEAM1 denotes *Rickettsia*_bellii_MEAM1. The names of the pathways that differ between RiTBt and the other two *Rickettsia* spp. are written in red, and different colours of the cells in the table indicate the varying number of genes being classified for each of the pathways: green indicates fewer and red indicates more genes.

Phylogenomic analysis based on hundreds of single-copy nuclear genes showed that this bacterium belongs to *Rickettsiae* with a close relationship with species of the Torix group. Through field survey, *in vivo* localization, and genomic comparative analysis, we named this newly found endosymbiotic bacterium in whiteflies as '*Candidatus* Rickettsia_Torix_ *Bemisia_tabaci* (RiTBt)'.

The range of hosts harbouring Rickettsia *of the Torix group*

Apart from whiteflies, Torix group *Rickettsia* have been reported previously in an array of phylogenetically distant invertebrate vectors including the rice leafhopper *Nephotettix cincticeps* (AB702995) (Noda *et al*., 2012), true bug *Macrolophus* sp. (HE583203) (Machtelinckx *et al*., 2012), water beetles *Deronectes aubei* (FM955310), *Deronectes semirufus* (FM955311), and *Deronectes delarouzei* (FM955312) (Küchler *et al*., 2009), leeches *Hemiclepsis marginata* (AB066352), *Torix tukubana* (AB113214), and *Torix tagoi* (AB066351) (Kikuchi *et al*., 2002; Kikuchi and Fukatsu, 2005), cicada *Platypleura kaempferi* (KR911839), sand fly *Phlebotomus chinensis*) (KX363668), mite *Lutzomyia apache* (EU223247) (Reeves *et al*., 2008), crane fly *Limonia chorea* (AF322442) (Perotti *et al*., 2006), ice worm *Mesenchytraeus solifugus glacier* (AB991365.1), and booklice *Cerobasis guestfalica* (DQ652595)] (Supporting Information Fig. S5). Interestingly, many of these vectors are aquatic invertebrates whose abundance has long been known to be closely tied to water quality, such as midges, water beetles, ice worms, crane flies, leeches, and sandflies. However, the remaining vectors, such as rice leafhoppers, cicadas, true bugs, and whiteflies, are sap-sucking insects. The idea that horizontal transmission occurs in nature between distant relatives is widely accepted (Caspi-Fluger *et al*., 2012; Russell *et al*., 2003). Thus, whiteflies might have captured this *Rickettsia* via plant sap, which contains this *Rickettsia* excreted by insects of the same order, such as rice leafhoppers, cicadas, or true bugs. More investigations are warranted to explore its possible transmission routes among ecologically and phylogenetically disparate hosts.

Implications of the localization of RiTBt in B. tabaci

FISH and TEM consistently showed the presence of an intracellular *Rickettsia*-like structure located in the whole-body cavity, as well as midgut, salivary gland, bacteriocyte, ovary, and testis of Asia II 7. The distribution throughout the host body implies that this *Rickettsia* has evolved to elude immune defences and survive in the body tissues of *B. tabaci* whiteflies without harming the host. The fact that the RiTBt can be detected inside whitefly eggs, nymphs, and adults indicates vertical transmission. A higher density of RiTBt was detected in the midgut, especially in the filter chamber of whitefly midgut than that of other organs/tissues, a pattern of distribution similar to that reported for *Rickettsia*_bellii_MEAM1 (Kliot *et al*., 2014). We also observed high densities of RiTBt in the salivary gland – as phloem-feeding insects whiteflies of the *B. tabaci* complex are potentially able to overwhelm plant defence mechanisms using their saliva (Will *et al*., 2013; Xu *et al*., 2019). Whether symbionts in the salivary glands of whitefly are involved in host defence to plant resistance warrants exploration.

Functions of *Rickettsia* in the biology of several phloem-feeding insects have been reported, including tolerance to heat stress, promotion of reproduction and spread, and resistance to insecticides (Ferrari *et al*., 2004; Ghanim and Kontsedalov, 2009; Brumin *et al*., 2011; Himler *et al*., 2011). The effect of RiTBt on the biology of whiteflies in the *B. tabaci* species complex is yet to be investigated, but the observations of its spatial and subcellular localization provided here may serve as clues for future effort on this topic.

Stable taxonomy based on genome data

Phylogenomic estimation is central to understand biodiversity, evolution, and ecology and provides a higher taxonomic resolution than phylogenies obtained from single marker genes (Cao and Sun, 2016). To accurately describe and accommodate RiTBt, we compared previously defined *Rickettsia* species with RiTBt to conduct a global phylogenomic analysis. RiTBt was recognized as a new, independent species under the Torix group based on both phylogenomic estimation and the recently proposed delimitation criteria of bacterial species (Konstantinidis and Tiedje, 2005; Richter and Rosselló-Móra, 2009). The results not only confirmed the evolutionary relationship of RiTBt with other species in the *Rickettsia* genus but also filled a major phylogenetic gap of species diversity in the Torix clade.

Implication of comparison between gene families and metabolic pathways

Differences in the size of gene families between organisms have generally been attributed to their adaptive evolution (Hahn *et al*., 2005). In a global view of *Rickettsia* genome family evolution, our data indicate that considerable contraction has occurred in the gene families of RiTBt, illustrating a large amount of genetic downsizing relative to *Orientia tsutsugamushi*, an obligate intracellular bacterium that also exhibits genome reduction relative to free-living bacteria (Cole *et al*., 2001). The most extensive genome degradation (37 gene family decays) reported previously was from *R. prowazekii* (Anderssen

et al., 1998). Our analysis illustrates that much higher genome degradation has occurred in *Rickettsia* genome families, as shown by the 46 gene family decays of *Rickettsia*_bellii_MEAM1.

Similar to gene families, certain essential metabolic pathways, such as 'Glycolysis (Embden-Meyerhof pathway), glucose', 'Glycolysis, core module involving three-carbon compounds', are absent in *Rickettsia*_bellii_MEAM1 compared to RiTBt. Both species of the Torix group retain a crucial pentose phosphate pathway (Pilgrim *et al*., 2017), which is associated with maintenance of carbon homoeostasis, provision of precursors for nucleotide and amino acid biosynthesis, reduction of molecules for anabolism, and decline of oxidative stress (Stincone *et al*., 2015). Similar to an uncultured pathogenic endosymbiont '*Candidatus* Liberibacter asiaticus' in psyllid (Jain *et al*., 2017), glycolysis and the pentose phosphate pathways are both absent in *Rickettsia*_bellii_MEAM1, indicating that this *Rickettsia* could be an energy scavenger and thus import ATP from its host (Jain *et al*., 2017). The absence of glycolysis and the pentose phosphate pathways in *Rickettsia*_bellii_MEAM1 could be complemented by the host (Driscoll *et al*., 2017) or other microbial symbionts in the host. The above analyses suggest that RiTBt is more recently acquired than *Rickettsia*_bellii_MEAM1 by its host and is thus less co-evolved but still shows signs of adaptation to endosymbiosis in its host strain distribution and gene family loss in comparison to the *Orientia* outgroup.

RiTBt, a new species of Rickettsia *associated with* B. tabaci

*Rickettsia*_bellii_MEAM1 is the first *Rickettsia* species described residing in *B. tabaci* whiteflies (Gottlieb *et al*., 2006; Singh *et al*., 2012). Comparison of characteristics of RiTBt with those of *Rickettsia*_bellii_MEAM1, presumably the closest relative to RiTBt, based on their being associated with the same insect host (Caspi-Fluger *et al*., 2011; Chen *et al*., 2016) shows the following differences between the two:

(i) Genome features and phylogenic classification. The genomes' differences in size, number of transposes, gene arrangement, metabolic pathways, gene families, and phylogenetic position between RiTBt and *Rickettsia*_bellii_MEAM1 are sufficient to classify RiTBt as a new species. RiTBt is more closely related to a *Rickettsia* from another host species than to the other *Rickettsia* described in whiteflies, and its genome size and content suggest that it is a more recent acquisition than *Rickettsia*_bellii_MEAM1.

(ii) Spatial and subcellular localization in host body. RiTBt resides in the midgut, bacteriocyte, ovary, male reproductive system, as well as in nucleus and muscle tissues of its host based on our observations, while *Rickettsia*_bellii_MEAM1 has two patterns of distribution – a 'scattered' pattern where the *Rickettsia* is localized throughout the whitefly hemocoel excluding bacteriome, and a 'confined' pattern where the *Rickettsia* is restricted to the bacteriocytes, both of which differ from that of RiTBt.

(iii) Host range. RiTBt has been found so far only in indigenous species of the *B. tabaci* complex, Asia II 1 and Asia II 7, while *Rickettsia*_bellii_MEAM1 has been found in both indigenous and invasive species MEAM1, Asia II 3, and China 1. Notably, their hosts do not overlap suggesting specific mechanisms of each for adaptation to its hosts, suggesting specific mechanisms of adaptation for different whitefly host species.

The evidence obtained from the comparison shows that the newly detected *Rickettsia* is a new species of the bacterium residing in whiteflies of the *B. tabaci* species complex. The differences in gene family and metabolic pathways between '*Candidatus* Rickettsia_Torix_*Bemisia_tabaci* (RiTBt)' and *Rickettsia*_bellii_MEAM1, in particular, the absence of some vital gene families in *Rickettsia*_bellii_MEAM1, indicate that the two species of *Rickettsia* may differ in their symbiosis with their hosts. Detailed analyses of these differences may reveal distinct roles of the two species of *Rickettsia* in the biology of the host whiteflies.

Materials and methods

Insect materials

A colony of *B. tabaci* Asia II 7 (identified from its mitochondrial cytochrome oxidase 1 [*mtCOI*] gene sequence; GenBank accession no. KM821541) was initially collected from the host plant *Hibiscus rosa-sinensis* (L.) in Guangzhou, Guangdong province of China, in September 2013. It was maintained on cotton [*Gossypium hirsutum* (Malvaceae) cv. Zhe-Mian 1793] under the conditions of 26 $\pm$ 1°C, 60% $\pm$ 10% relative humidity and a photoperiod of 14 h light and 10 h darkness (Liu *et al*., 2007), for subsequent genome sequencing, FISH and TEM observations.

To identify the presence of the RiTBt in field whitefly populations, 1–46 adult whiteflies were taken from each of 57 *B. tabaci* samples that were collected from 22 localities of 14 provinces of China, from June to September 2015 (Fig. 2. and Supporting Information Table S1). In total, 512 adult whiteflies were screened. These wild whitefly specimens were initially stored in 95% ethanol and subsequently kept at −80°C until DNA extraction.

DNA extraction, PCR amplification, and sequencing

Individual adult DNA extraction and the cryptic species diagnosis were done using the method of Bing *et al*. (2013). The phylogenetic relationships of the RiTBt with other symbionts were estimated with three widely used markers: a 1.4 Kb fragment of the *16S rRNA* gene with primers 16S 27F (5′-CTACGGCTACCTTGTTACGA-3′) and 16S 1494R (5′-ATTGCTTCCCTCTGTGGAT-3′) (Weisburg *et al*., 1991), a 0.5 Kb fragment of the *gltA* gene with primers *gltA* F (5′-ACATGCAGACCATGAGCAGA-3′) and *gltA* R (5′-CATTTCATTCCATTGTGCCATC-3′) (Li *et al*., 2011), and a 1.6 Kb fragment of the *groEL* gene with newly designed primers *groEL* F (5′-CACCDA AAATTACTAAAGATGG-3′) and *groEL* R (5′-TAGAART CCATWCCKCCCATWC-3′).

To detect the presence of the RiTBt in whitefly populations, specific *16S rRNA* primers, 16S F (5′-GATTAATTTAGGGCTTGCTCTG-3′) and 16S R (5′-ATTGCTTCCCTCTGTGGAT-3′), were designed by comparing a wide range of *Rickettsia* spp., and for the further confirmation, the relative amplicons were sent for sequence. *16S rRNA* sequences were downloaded from GenBank. The size of the PCR product is 1.1 Kb. Cycling condition was 94°C for 3 min, 34 cycles of 94°C for 30 s, 50–53°C for 30 s, 72°C for 2 min, and a final cycle of 72°C for 10 min. The amplicons were purified using a QIAGEN PCR purification kit (Germany) and submitted for Sanger sequencing in both directions.

Fluorescence in situ *hybridization*

Eggs, fourth instar nymphs, and adults of *B. tabaci* were subjected to FISH following the approach of Gottlieb *et al*. (2006) and Bing *et al*. (2013). As for dissected organs, FISH was conducted as previously described by Ghanim *et al*. (2009) with a slight modification. The obligate endosymbiont '*Ca*. Portiera aleyrodidarum' (*Portiera*) occurs exclusively inside bacteriocytes throughout the host life cycle of all *B. tabaci* species (Gottlieb *et al*., 2006; Bing *et al*., 2013); therefore, it provides a good control for FISH. Thus, the widely used probe BTP1-Cy3 (5′-Cy3-TGTCA GTGTCAGCCCAGAAG-3′) was chosen to target the *16S rRNA* of the *Portiera* (Gottlieb *et al*., 2006), while a new probe, R-Cy5 (5′-Cy5-TCCACGTCACCGTATTGC-3′), was developed for detection of *Ricketssia* strains using Primer3 software (http://bioinfo.ut.ee/primer3-0.4.0/primer3/) based on the 16S rRNA sequence from *Rickettsia*-like strains. The specificity of detection was confirmed using *Rickettsia*-free whitefly *B. tabaci* MED as negative controls. Fluorescence was visualized on ZEISS LSM780 at Bio-ultrastructure Analysis Laboratory of the Analysis Center of Agrobiology and Environmental Sciences, Zhejiang University.

Transmission electron microscopy

To observe the localization of RiTBt in cells or tissues, TEM was conducted on midguts, salivary glands, bacteriocytes, ovaries, and testes, dissected from adult specimens of Asia II 7. TEM analysis was performed as follows: (i) Double fixation: samples were firstly fixed with 2.5% glutaraldehyde in PBS buffer for 4 h, then rinsed three times in PBS buffer for 15 min each, and postfixed with osmium tetroxide in PBS buffer for 15 min. (ii) Dehydration: the samples were dehydrated in an ascending ethanol series for 20 min before transferring to absolute acetone. (iii) Infiltration: the samples were placed in 1:1 mixture of absolute acetone and Spurr resin overnight. (iv) Embedding and ultrathin sectioning: the samples were placed in an Eppendorf tube containing pure Spurr resin and incubated at 70°C for 9 h. In the end, samples were sectioned in a LEICA EM UC7 ultramicrotome, and sections were stained with aqueous uranyl acetate and lead citrate for 5 min and 10 min respectively. These sections were observed under a Hitachi Model H-7650 TEM at Bio-ultrastructure Analysis Laboratory of the Analysis Center of Agrobiology and Environmental Sciences, Zhejiang University.

Isolation, amplification, and sequencing of endosymbiont genome

The dissection of bacteriomes, genomic DNA isolation, library preparation, and sequencing was conducted as described in the study by Rao *et al*. (2015). Briefly, nymphs with obvious green-yellow coloured bacteriomes were collected and dissected in PBS (GIBCO®, Invitrogen, USA) using autoclaved tools. PBS containing endosymbiont cells were then divided into six separate samples and amplified by multiple displacement amplification separately using the REPLI-g UltraFast Mini Kit (Qiagen, Germany) with slight modifications. Subsequently, the yield of the six copies was mixed with equal concentrations to construct one paired-end library (2×150 bp paired reads) with an average insert size of 400–500 bp. The sample was then sequenced on Illumina HiSeq2000 (Biozeron Biotech, Shanghai), with 8 million reads generated.

Genome assembly

Raw FASTQ data were generated by the Illumina basecalling software CASAVA v1.8.2 (http://support.illumina.com/sequencing/sequencing_software/casava.ilmn).
Quality assessment and filtering of the Illumina reads were performed using FASTQC v0.10.1 (Andrews, 2010) and FASTX toolkit v0.0.13 (Gordon, 2011). Contaminating reads, containing adaptor sequences or primer

sequences were identified by SeqPrep (https://github.com/jstjohn/SeqPrep) with the parameters: '-q 20-L 75-B AGATCGGAAGAGCGTCGTGT-A AGATCGGAAGAGCACACGTC'. Sickle (https://github.com/najoshi/sickle) was applied for reads trimming with default parameters. Through the above quality control processes, clean reads were obtained for a preliminary assembly by SPAdes version 3.10 (Bankevich *et al*., 2012) with multiple k-mer sizes under 'careful' mode. The parameters for filtering redundant reads/contigs followed the approach described by Pilgrim *et al*. (2017). The final assembly was subjected to a final decontamination step to remove the contigs that had high similarity with the other endosymbionts found in *B. tabaci*. The remaining scaffolds were used for downstream analysis.

The assessment of completeness of the genome assembly, prediction of coding sequences and tRNAs identification and annotations were conducted using the method used by Pilgrim *et al*. (2017).

Orthologue identification, molecular phylogeny and evolutionary analyses of gene family

We downloaded 64 *Rickettsia* genomes and one outgroup *Orientia tsutsugamushi* genome from PATRIC (Wattam *et al*., 2014) for conducting the comparative genomic analysis. OrthoMCL (Li *et al*., 2003) was used to generate orthologous groups (OGs) of protein families from the 64 complete or nearly complete genomes. From this analysis, proteins were designated unique, core, or non-conserved based on their distribution across *Rickettsia* genomes. A set of 459 single-copy nuclear genes were selected to concatenate and were aligned in MUSCLE with default parameters (Edgar, 2004). Gblocks (Castresana, 2002) was used to mask regions of poor alignment (heterogeneous length regions). All modified alignments were then concatenated.

Phylogenetic trees for either single genes or large-scale concatenated sequences were constructed using the program MrBayes ver. 3.2 (Ronquist *et al*., 2012). PartitionFinder2 (Lanfear *et al*., 2016) was used to choose best-fit substitution models for each nucleotide and protein sequence alignment. Computational Analysis of Gene Family Evolution (CAFE) (De Bie *et al*., 2006), a tool using a stochastic birth and death process to model the evolution of gene family size over a phylogeny, was employed to study gene family gain and loss.

Comparative genomic analysis

Genome rearrangement was assessed using AliTV (Ankenbrand *et al*., 2017). Metabolic potential of the RiTBt was evaluated using the Metabolic and Physiological Evaluator (MAPLE-2.1.0) (Takami *et al*., 2016).

Acknowledgements

The authors thank Professor Myron Zalucki for critically reading an earlier version of the manuscript. The authors also thank San-Ling Wu, Zhejiang University, and Liang Zeng, Shanghai Biozeron Biotechnology Co. Ltd, for providing sequencing platform and bioinformatics support. This work was supported by Bill & Melinda Gates Foundation (African cassava whitefly project, OPP1058938), China Postdoctoral Science Special Foundation (Project no. 517000-X91609), and Chinese Postdoctoral Science Foundation (Project no. 517000-X91502).

Data accessibility

The sequences of the RiTBt bacteria genes obtained in this study are deposited in GenBank under the accession numbers MG063879 (*16S rRNA*), MG063880 (*gltA*) and MG063881 (*groEL*). The whole-genome shotgun project has been deposited at GenBank under the accession number NVXZ00000000.

References

Anderssen, S., Zomorodipour, A., and Andersson, J. (1998) The complete genome sequence of the obligate intracellular parasite *Rickettsia prowazekii*. *Nature* **396**: 133–140.

Andrews, S. (2010) FastQC: a quality control tool for high throughput sequence data. URL http://www.bioinformatics.babraham.ac.uk/projects/fastqc/'?

Ankenbrand, M.J., Hohlfeld, S., Hackl, T., and Förster, F. (2017) AliTV-interactive visualization of whole genome comparisons. *PeerJ Computer Science* **3**: e116.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., *et al*. (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* **19**: 455–477.

Baumann, P. (2005) Biology of bacteriocyte-associated endosymbionts of plant sap-sucking insects. *Annu Rev Microbiol* **59**: 155–189.

Bing, X.L., Yang, J., Zchori-Fein, E., Wang, X.W., and Liu, S. S. (2013) Characterization of a newly discovered symbiont of the whitefly *Bemisia tabaci* (Hemiptera: Aleyrodidae). *Appl Environ Microbiol* **79**: 569–575.

Blanc, G., Ogata, H., Robert, C., Audic, S., Suhre, K., Vestris, G., *et al*. (2007) Reductive genome evolution from the mother of *Rickettsia*. *PLoS Genet* **3**: e14.

Boykin, L.M. (2014) *Bemisia tabaci* nomenclature: lessons learned. *Pest Manag Sci* **70**: 1454–1459.

Brumin, M., Kontsedalov, S., and Ghanim, M. (2011) *Rickettsia* influences thermotolerance in the whitefly *Bemisia tabaci* B biotype. *Insect Sci* **18**: 57–66.

Cao, C.C., and Sun, X. (2016) Combinatorial pooled sequencing: experiment design and decoding. *Quant Biol* **4**: 36–46.

Caspi-Fluger, A., Inbar, M., Mozes-Daube, N., Mouton, L., Hunter, M.S., and Zchori-Fein, E. (2011) *Rickettsia* 'in' and 'out': two different localization patterns of a bacterial symbiont in the same insect species. *PLoS One* **6**: e21096.

Caspi-Fluger, A., Inbar, M., Mozes-Daube, N., Katzir, N., Portnoy, V., Belausov, E., *et al*. (2012) Horizontal transmission of the insect symbiont *Rickettsia* is plant-mediated. *Proc R Soc B* **279**: 1791–1796.

Castresana, J. (2002) Gblocks, v. 0.91 b. Online version.

Chen, W., Hasegawa, D.K., Kaur, N., Kliot, A., Pinheiro, P. V., Luan, J., *et al*. (2016) The draft genome of whitefly *Bemisia tabaci* MEAM1, a global crop pest, provides novel insights into virus transmission, host adaptation, and insecticide resistance. *BMC Biol* **14**: 110.

Cole, S., Eiglmeier, K., Parkhill, J., James, K., Thomson, N., Wheeler, P., *et al*. (2001) Massive gene decay in the leprosy bacillus. *Nature* **409**: 1007–1011.

De Bie, T., Cristianini, N., Demuth, J.P., and Hahn, M.W. (2006) CAFE: a computational tool for the study of gene family evolution. *Bioinformatics* **22**: 1269–1271.

Douglas, A.E. (2006) Phloem-sap feeding by animals: problems and solutions. *J Exp Bot* **57**: 747–754.

Driscoll, T.P., Verhoeve, V.I., Guillotte, M.L., Lehman, S.S., Rennoll, S.A., Beier-Sexton, M., *et al*. (2017) Wholly *Rickettsia*! Reconstructed metabolic profile of the quintessential bacterial parasite of eukaryotic cells. *MBio* **8**: e00859–e00817.

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**: 1792–1797.

Everett, K.D.E., Thao, M., Horn, M., Dyszynski, G.E., and Baumann, P. (2005) Novel chlamydiae in whiteflies and scale insects: endosymbionts '*Candidatus* Fritschea bemisiae' strain Falk and '*Candidatus* Fritschea eriococci' strain elm. *Int J Syst Evol Microbiol* **55**: 1581–1587.

Ferrari, J., Darby, A.C., Daniell, T.J., Godfray, H.C.J., and Douglas, A.E. (2004) Linking the bacterial community in pea aphids with host-plant use and natural enemy resistance. *Ecol Entomol* **29**: 60–65.

Ghanim, M. (2016) Investigating the role of *Rickettsia* in *tomato yellow leaf curl virus* interactions with its whitefly vector *Bemisia tabaci*. In *Proceedings of the Plant and Animal Genome XXIV Conference*.

Ghanim, M., Brumin, M., and Popovski, S. (2009) A simple, rapid and inexpensive method for localization of *tomato yellow leaf curl virus* and *potato leafroll virus* in plant and insect vectors. *J Virol Methods* **159**: 311–314.

Ghanim, M., and Kontsedalov, S. (2009) Susceptibility to insecticides in the Q biotype of *Bemisia tabaci* is correlated with bacterial symbiont densities. *Pest Manag Sci* **65**: 939–942.

Gillespie, J.J., Ammerman, N.C., Dreher-Lesnick, S.M., Rahman, M.S., Worley, M.J., Setubal, J.C., *et al*. (2009) An anomalous type IV secretion system in *Rickettsia* is evolutionarily conserved. *PLoS One* **4**: e4833.

Gillespie, J.J., Williams, K., Shukla, M., Snyder, E.E., Nordberg, E.K., Ceraul, S.M., *et al*. (2008) *Rickettsia* phylogenomics: unwinding the intricacies of obligate intracellular life. *PLoS One* **3**: e2018.

Gordon, A. (2011) FASTX-Toolkit. Computer program distributed by the author. http://hannonlab.cshl.edu/fastx_toolkit/.

Gottlieb, Y., Ghanim, M., Chiel, E., Gerling, D., Portnoy, V., Steinberg, S., *et al*. (2006) Identification and localization of a *Rickettsia* sp. in *Bemisia tabaci* (Homoptera: Aleyrodidae). *Appl Environ Microbiol* **72**: 3646–3652.

Gottlieb, Y., Zchori-Fein, E., Mozes-Daube, N., Kontsedalov, S., Skaljac, M., Brumin, M., *et al*. (2010) The transmission efficiency of *tomato yellow leaf curl virus* by the whitefly *Bemisia tabaci* is correlated with the presence of a specific symbiotic bacterium species. *J Virol* **84**: 9310–9317.

Hahn, M.W., De Bie, T., Stajich, J.E., Nguyen, C., and Cristianini, N. (2005) Estimating the tempo and mode of gene family evolution from comparative genomic data. *Genome Res* **15**: 1153–1160.

Himler, A.G., Adachi-Hagimori, T., Bergen, J.E., Kozuch, A., Kelly, S.E., Tabashnik, B.E., *et al*. (2011) Rapid spread of a bacterial symbiont in an invasive whitefly is driven by fitness benefits and female bias. *Science* **332**: 254–256.

Jain, M., Munoz-Bodnar, A., and Gabriel, D.W. (2017) Concomitant loss of the glyoxalase system and glycolysis makes the uncultured pathogen "*Candidatus* Liberibacter asiaticus" an energy scavenger. *Appl Environ Microbiol* **83**: e01670–e01617.

Kikuchi, Y., and Fukatsu, T. (2005) *Rickettsia* infection in natural leech populations. *Microb Ecol* **49**: 265–271.

Kikuchi, Y., Sameshima, S., Kitade, O., Kojima, J., and Fukatsu, T. (2002) Novel clade of *Rickettsia* spp. from leeches. *Appl Environ Microbiol* **68**: 999–1004.

Kliot, A., Cilia, M., Czosnek, H., and Ghanim, M. (2014) Implication of the bacterial endosymbiont *Rickettsia* spp. in interactions of the whitefly *Bemisia tabaci* with *tomato yellow leaf curl virus*. *J Virol* **88**: 5652–5660.

Konstantinidis, K.T., and Tiedje, J.M. (2005) Genomic insights that advance the species definition for prokaryotes. *PNAS* **102**: 2567–2572.

Küchler, S.M., Kehl, S., and Dettner, K. (2009) Characterization and localization of *Rickettsia* sp. in water beetles of genus *Deronectes* (Coleoptera: Dytiscidae). *FEMS Microbiol Ecol* **68**: 201–211.

Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T., and Calcott, B. (2016) PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Mol Biol Evol* **34**: 772–773.

Li, L., Stoeckert, C.J., and Roos, D.S. (2003) OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res* **13**: 2178–2189.

Li, T., Xiao, J.H., Xu, Z.H., Murphy, R.W., and Huang, D.W. (2011) Cellular tropism, population dynamics, host range and taxonomic status of an aphid secondary symbiont, SMLS (*Sitobion miscanthi* L type symbiont). *PLoS One* **6**: e21944.

Liu, S.S., Colvin, J., and De Barro, P.J. (2012) Species concepts as applied to the whitefly *Bemisia tabaci* systematics: how many species are there? *J Integr Agr* **11**: 176–186.

Liu, S.S., De Barro, P.J., Xu, J., Luan, J.B., Zang, L.S., Ruan, Y.M., and Wan, F.H. (2007) Asymmetric mating interactions drive widespread invasion and displacement in a whitefly. *Science* **318**: 1769–1772.

Luan, J.B., Shan, H.W., Isermann, P., Huang, J.H., Lammerding, J., Liu, S.S., and Douglas, A.E. (2016) Cellular and molecular remodelling of a host cell for vertical transmission of bacterial symbionts. *Proc R Soc B* **283**: 20160580.

Machtelinckx, T., Van Leeuwen, T., Van De Wiele, T., Boon, N., De Vos, W.H., Sanchez, J.A., *et al*. (2012) Microbial community of predatory bugs of the genus *Macrolophus* (Hemiptera: Miridae). *BMC Microbiol* **12**: 1.

Montllor, C.B., Maxmen, A., and Purcell, A.H. (2002) Facultative bacterial endosymbionts benefit pea aphids *Acyrthosiphon pisum* under heat stress. *Ecol Entomol* **27**: 189–195.

Nirgianaki, A., Banks, G.K., Frohlich, D.R., Veneti, Z., Braig, H.R., Miller, T.A., *et al*. (2003) *Wolbachia* infections of the whitefly *Bemisia tabaci*. *Curr Microbiol* **47**: 0093–0101.

Noda, H., Watanabe, K., Kawai, S., Yukuhiro, F., Miyoshi, T., Tomizawa, M., *et al*. (2012) Bacteriome-associated endosymbionts of the green rice leafhopper *Nephotettix cincticeps* (Hemiptera: Cicadellidae). *Appl Entomol Zool* **47**: 217–225.

Ogata, H., La Scola, B., Audic, S., Renesto, P., Blanc, G., Robert, C., *et al*. (2006) Genome sequence of *Rickettsia bellii* illuminates the role of amoebae in gene exchanges between intracellular pathogens. *PLoS Genet* **2**: e76.

Perotti, M.A., Clarke, H.K., Turner, B.D., and Braig, H.R. (2006) *Rickettsia* as obligate and mycetomic bacteria. *FASEB J* **20**: 2372–2374.

Pilgrim, J., Ander, M., Garros, C., Baylis, M., Hurst, G.D., and Siozios, S. (2017) Torix group *Rickettsia* are widespread in *Culicoides* biting midges (Diptera: Ceratopogonidae), reach high frequency and carry unique genomic features. *Environ Microbiol* **19**: 4238–4255.

Rao, Q., Rollat-Farnier, P.A., Zhu, D.T., Santos-Garcia, D., Silva, F.J., Moya, A., *et al*. (2015) Genome reduction and potential metabolic complementation of the dual endosymbionts in the whitefly *Bemisia tabaci*. *BMC Genomics* **16**: 1.

Reeves, W.K., Kato, C.Y., and Gilchriest, T. (2008) Pathogen screening and bionomics of *Lutzomyia apache* (Diptera: Psychodidae) in Wyoming, USA. *J AmMosq Control Assoc* **24**: 444–447.

Russell, J., Latorre, A., Sabater-Muñoz, B., Moya, A., and Moran, N. (2003) Side-stepping secondary symbionts: widespread horizontal transfer across and beyond the Aphidoidea. *Mol Ecol* **12**: 1061–1075.

Richter, M., and Rosselló-Móra, R. (2009) Shifting the genomic gold standard for the prokaryotic species definition. *PNAS* **106**: 19126–19131.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., *et al*. (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* **61**: 539–542.

Singh, S.T., Priya, N.G., Kumar, J., Rana, V.S., Ellango, R., Joshi, A., *et al*. (2012) Diversity and phylogenetic analysis of endosymbiotic bacteria from field caught *Bemisia tabaci* from different locations of North India based on 16S rDNA library screening. *Infect Genet Evol* **12**: 411–419.

Stincone, A., Prigione, A., Cramer, T., Wamelink, M., Campbell, K., Cheung, E., *et al*. (2015) The return of metabolism: biochemistry and physiology of the pentose phosphate pathway. *Biol Rev* **90**: 927–963.

Takami, H., Taniguchi, T., Arai, W., Takemoto, K., Moriya, Y., and Goto, S. (2016) An automated system for evaluation of the potential functionome: MAPLE version 2.1. 0. *DNA Res* **23**: 467–475.

Wattam, A.R., Abraham, D., Dalay, O., Disz, T.L., Driscoll, T., Gabbard, J.L., *et al*. (2014) PATRIC, the bacterial bioinformatics database and analysis resource. *Nucleic Acids Res* **42**: D581–d591.

Weeks, A.R., Velten, R., and Stouthamer, R. (2003) Incidence of a new sex-ratio-distorting endosymbiotic bacterium among arthropods. *Pro R Soc B* **270**: 1857–1865.

Weinert, L.A., Werren, J.H., Aebi, A., Stone, G.N., and Jiggins, F.M. (2009) Evolution and diversity of *Rickettsia* bacteria. *BMC Biol* **7**: 1.

Weisburg, W.G., Barns, S.M., Pelletier, D.A., and Lane, D.J. (1991) 16S ribosomal DNA amplification for phylogenetic study. *J Bacteriol* **173**: 697–703.

Will, T., Furch, A.C., and Zimmermann, M.R. (2013) How phloem-feeding insects face the challenge of phloem-located defenses. *Front Plant Sci* **4**: 336.

Xu, H.X., Qian, L.X., Wang, X.W., Shao, R.X., Hong, Y., Liu, S.S., and Wang, X.W. (2019) A salivary effector enables whitefly to feed on host plants byeliciting salicylic acid-signaling pathway. *PNAS* **116**: 490–495.

Zchori-Fein, E., and Brown, J.K. (2002) Diversity of prokaryotes associated with *Bemisia tabaci* (Gennadius) (Hemiptera: Aleyrodidae). *Ann Entomol Soc Am* **95**: 711–718.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Fig. S1 Negative control of FISH in different organs of MED. Fish images: (A) Egg, (B) Fourth instar nymph, (C) Male adult, (D) Female adult, (E) Bacteriocyte, (F) Midgut, (G) Salivary gland, (H) Ovary, (I) Testis. Scale bar is 100 μm for Figs. A, B, C, and 20 μm for Figs. D, E, F, G.

Fig. S2 COG-based characterization of all proteins with annotated functions in RiTBt.

Fig. S3 Venn diagram presenting the orthologous/paralogous gene clusters among a range of *Rickettsia* strains.

Fig. S4 Heat map of the protein diversity profiles of 519 core genes in 64 *Rickettsia* genomes.

Fig. S5 Vectors of Torix *Rickettsia*.

Table S1 Detection of RiTBt in various *B. tabaci* populations in China.

Table S2 Function annotations of RiTBt genome.

Table S3 P-T4SS and the *tra* conjugative DNA-transfer gene element in the RiTBt.

Table S4 Comparison of RiTBt and other published *Rickettsia* genomes.

Table S5 Single-copy nuclear genes for phylogenomic construction.