

Genomic Analysis of an Ascomycete Fungus from the Rice Planthopper Reveals How It Adapts to an Endosymbiotic Lifestyle

Hai-Wei Fan¹, Hiroaki Noda², Hong-Qing Xie³, Yoshitaka Suetsugu², Qian-Hua Zhu³, and Chuan-Xi Zhang^{1,*}

¹State Key Laboratory of Rice Biology and Ministry of Agriculture Key Laboratory of Agricultural Entomology, Institute of Insect Science, Zhejiang University, Hangzhou, China

²National Institute of Agrobiological Sciences, Tsukuba, Japan

³BGI-Shenzhen, Shenzhen, Guangdong, China

*Corresponding author: E-mail: chxzhang@zju.edu.cn.

Accepted: August 26, 2015

Abstract

A number of sap-sucking insects harbor endosymbionts, which are thought to play an important role in the development of their hosts. One of the most important rice pests, the brown planthopper (BPH), *Nilaparvata lugens* (Stål), harbors an obligatory yeast-like symbiont (YLS) that cannot be cultured in vitro. Genomic information on this YLS would be useful to better understand its evolution. In this study, we performed genome sequencing of the YLS using both 454 and Illumina approaches, generating a draft genome that shows a slightly smaller genome size and relatively higher GC content than most ascomycete fungi. A phylogenomic analysis of the YLS supported its close relationship with insect pathogens. We analyzed YLS-specific genes and the categories of genes that are likely to have changed in the YLS during its evolution. The loss of mating type locus demonstrated in the YLS sheds light on the evolution of eukaryotic symbionts. This information about the YLS genome provides a helpful guide for further understanding endosymbiotic associations in hemiptera and the symbiotic replacement of ancient bacteria with a multifunctional YLS seems to have been a successful change.

Key words: *Nilaparvata lugens*, yeast-like symbiont, *Ascomycete*, endosymbiosis, *Entomomyces delphacidicola*, evolution.

Introduction

The association between insects and intracellular endosymbionts is quite prevalent in nature, and more than 10% of insect species in several taxonomic orders depend on obligate microorganisms, which are bacteria in most cases, for fitness within their habitats (Wernegreen 2002; Gibson and Hunter 2010). The highly specialized and imbalanced diets of many insect hosts suggest that these endosymbionts play a role in nutrition. The establishment of symbiotic associations between insects and symbionts has most likely been one of the key factors in the evolutionary success of insects, as it may have allowed access to novel ecological niches by utilizing imbalanced food sources, such as plant sap (Hayashi and Chino 1990; Moran et al. 2008). Owing to the potentially important role that endosymbionts play in insects, much research has been conducted on these microbes to reveal new insights into the mechanism of their symbiosis. Phylogenetic analyses have shown cospeciation between the host and symbiont, which

is consistent with the stable transmission of endosymbionts through their host's lineage (Moran et al. 1993; Thao et al. 2000). Simultaneously, genome-level studies on many bacterial endosymbionts in insects have been conducted to reveal the forces that shape the evolution of these bacterial associates and the genetic basis of their specialization to an intracellular lifestyle (Wernegreen 2002; McCutcheon and Moran 2012).

The brown planthopper (BPH), *Nilaparvata lugens* (Stål), is a typical vascular feeder that sucks rice phloem sap, and it is one of the most destructive pests of rice in Asia. As a common species in the Hemiptera, *N. lugens* establishes an intimate symbiotic relationship with yeast-like symbionts (YLS), which are found in the mycetocytes formed from the fat body cells of the abdomen (Buchner 1965). The symbionts are maternally transmitted to the eggs by transovarial infection and are found at every developmental stages of the host (Cheng and Hou 2001). Their nutritional roles, supporting sterol biosynthesis (Noda and Koizumi 2003) and nitrogen recycling for the

host (Sasaki et al. 1996), has been reported. However, whole picture of YLS roles in the nutrition and metabolism of host planthopper remains unclear. As the YLS cannot yet be cultured *in vitro*, genome sequencing of this obligate endosymbiont would provide a powerful tool to better understand it.

In this article, we provide the first detailed description of the characteristics of the YLS genome in the BPH. Phylogenomic analysis and divergence time estimation were performed to trace the establishment of this mutualistic association. We then compared the YLS genome with those of its free-living relatives, the filamentous ascomycetes, and described whether any genomic changes have occurred to support its particular symbiotic lifestyle.

Materials and Methods

Insect Strain and Maintenance

The *N. lugens* population used in this study was also used for the genome sequencing project (Xue et al. 2014). The insects were reared on fresh rice seedlings cultured in nutrient solution at 27 °C with 70% humidity and a 16/8 h (light/dark) photoperiod. Adults, from which the YLS was isolated, were collected and stored at –80 °C.

Isolation of YLS and Genomic Preparation

The YLS were isolated from BPH following the method of Noda and Omura (1992). Several grams of planthoppers were homogenized in 0.85% NaCl solution. The homogenate was filtered through cotton cloths and centrifuged for 5 min at 100 × g. The pellet was resuspended and the centrifugation was repeated twice. The suspension of the final pellet, mixed with 4 vol of Percoll (GE healthcare bio-sciences) containing 0.25 M sucrose, was centrifuged for 30 min at 82,000 × g in a Hitachi ultracentrifuge (Himac CP80MX) using a P40ST horizontal rotor. Genomic DNA was extracted using the Yeast Smash and Grab DNA miniprep method described by Rose et al. (1990). Isolated YLS was dissolved in 200 μl of lysis buffer (100 mM NaCl, 1 mM ethylenediaminetetraacetic acid [EDTA] [pH 8.0], and 10 mM Tris-HCl [pH 8.0], 1% sodium dodecyl sulfate [SDS], 2% Triton X-100), and treated with lyticase (Sigma) at 37 °C for 30 min, then mixed with the same volume of phenol/chloroform (1:1), and vortexed vigorously with glass beads (425–600 μm, Sigma). Aqueous phase was recovered by ethanol precipitation and the pellet was dissolved in Tris-EDTA solution.

Genome Sequencing and Assembly

The genome of YLS was shotgun sequenced on a Roche 454 GS FLX at the Chinese National Human Genome Center at Shanghai. A total of 573,847 reads were generated and used to construct contigs using the Newbler software (Quinn et al. 2008) with default parameters. Additionally, DNA libraries of 300 bp and 3 kb insert-sizes were constructed and sequenced

using an Illumina HiSeq 2000. A total of 10,550,032 Illumina paired-end reads and 7,731,568 mate-pair reads were produced, respectively, and used for scaffolding using the SSPACE software with parameters: -m 30 -o 20 -k 5 -a 0.7 -x 1. A total of 28.4 Mb genome sequences were assembled. In addition, sequencing of genomic fosmid libraries of a whole female BPH body generated 20.9 Mb YLS Illumina sequences. In order to exclude contaminated planthopper sequences, whole-genome shotgun (WGS) data of BPH was mapped to the genome to plot GC-depth distribution diagram. According to different characteristics of GC content distribution in different species genomes, the scaffolds that located in a patch of separated region were filtered out and were submitted to BLAST against nr database with a cutoff *E* value of 1e-5. Finally, with all data combined, the YLS genome was assembled into 26.8 Mb. The YLS genome assemblies have been deposited at GenBank under accession number JRM100000000.

Genome Annotation

Annotations of the genomic sequences of YLS were performed with AUGUSTUS and the annotated information of *Metarhizium acridum* was incorporated as a reference (<http://www.ncbi.nlm.nih.gov/genome/?term=%20Metarhizium%20acridum>). Annotation of gene function was performed by aligning protein sequences to both the SwissProt and TrEMBL databases. Then we aligned annotated genes to Kyoto Encyclopedia of Genes and Genomes (KEGG) databases and constructed the corresponding KEGG pathway map. Whole-genome protein families were classified by InterProScan analysis (<http://www.ebi.ac.uk/interpro/>) in combination with the Treefam methodology that defines a protein family as a group of genes descended from a common ancestor. The whole gene set was also submitted for identifying secretory proteins using the program SignalP under default parameters (Petersen et al. 2011). Eukaryotic Orthologous Group (KOG) annotations were performed using Blastall software against clusters of orthologous groups database for eukaryotic complete genomes.

Genome repetitive elements were analyzed by BLAST against the RepeatMasker library (<http://www.repeatmasker.org/>) and with the Tandem Repeat Finder (<http://tandem.bu.edu/trf/trf.html>). The transposases/retrotransposases were classified by BLASTp analysis against the Repbase (<http://www.girinst.org/repbase/>).

Ortholog Prediction

We included major fungal lineages with publicly available genome sequences for ortholog prediction and phylogenetic reconstruction. Protein sequences from 17 species were used to construct gene families. We identified gene families using TreeFam (Li et al. 2006; Ruan et al. 2008) following these steps: 1) BLASTp was used to compare all protein sequences

from the 17 selected species, including the YLS of BPH, with an E value of less than $1e^{-7}$; 2) segments of high-scoring segment pairs were concatenated using Solar (Ruan et al. 2008), followed by the identification of homology among protein sequences based on the bit scores and the identities of homologous gene pairs; and 3) gene families were identified by clustering using Hcluster_sg (0.5.0, build April 4, 2007), whose algorithm was similar to that of regular hierarchical clustering. The gene family expansion and contraction were identified by using CAFE (De Bie et al. 2006).

Comparative Genomic Analysis

A BLAST score ratio (BSR) test (Rasko et al. 2005) was conducted to compare the differences between the YLS and its relatives' genomes. The BSR index for each reference protein was calculated by dividing the query score by the reference score and normalizing the result to a score between 0 and 1. A score of 1 indicated a perfect match, whereas a score of 0 indicated no BLAST match for the query protein in the reference proteome. The normalized pairs of BSR indices were then plotted using Excel 2010.

Phylogenetic Reconstruction and Divergence Time Estimation

Single-copy gene families were used to reconstruct phylogenetic relationships. Coding sequences from each single-copy family were concatenated to form one supergene for each species. All of the nucleotides at codon position 2 of these concatenated genes were extracted and used to construct a phylogenetic tree using PhyML (Guindon and Gascuel 2003; Guindon et al. 2010), with a gamma distribution across sites and an HKY85 substitution model.

The same set of sequences at codon position 2 was used to estimate divergence times among lineages. Fossil calibrations were set according to previous papers (Lücking et al. 2009). The PAML mcmctree (V4.5; Yang and Rannala 2006; Rannala and Yang 2007; Yang 2007) program was used to compute split times using the approximate likelihood calculation algorithm. Tracer (V1.5.0; Rambaut and Drummond 2009) was applied to examine convergence, and two independent runs were performed for confirmation.

Results

Genome Sequencing and General Features

The assembly of YLS genome resulted in 582 scaffolds (N50 size 310.7 kb, max 873.3 kb) containing 5,839 contigs with a total size of 26.8 Mb, which is 33% larger than the size estimated from preliminary karyotyping studies (17.3 Mb) (Noda and Kawahara 1995). The genome size and GC content were close to those of the YLS of in the palm aphid *Cerataphis brasiliensis* (Vogel and Moran 2013). Compared with several other sequenced ascomycete fungal genomes, the YLS

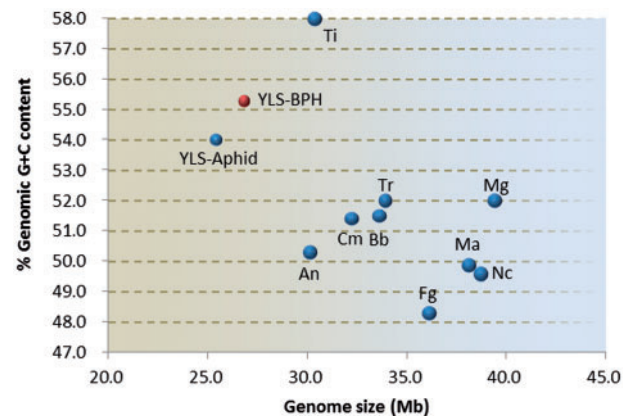


Fig. 1.—Comparison of genome sizes and %G+C content of several selected ascomycete fungi. YLS (BPH), YLS of brown planthopper; YLS (PA), YLS of palm aphid (*Cerataphis brasiliensis*); An, *Aspergillus nidulans*; Bb, *Beauveria bassiana*; Cm, *Cordyceps militaris*; Fg, *Fusarium graminearum*; Ma, *Metarhizium acridum*; Mg, *Magnaporthe grisea*; Nc, *Neurospora crassa*; Ti, *Tolypocladium inflatum*; Tr, *Trichoderma reesei*.

genome size is smaller but GC content is higher in some degree (fig. 1 and [supplementary table S1, Supplementary Material](#) online). However, a relatively larger number of repetitive sequences were observed in the YLS genome. In total, 3.35% of the YLS genome is repetitive, which is larger than the percentage found in four insect fungal pathogens (Xiao et al. 2012), including *Cordyceps militaris* (3.04%), *Beauveria bassiana* (2.03%), *Metarhizium anisopliae* (0.98%), and *M. acridum* (1.52%) ([supplementary table S2, Supplementary Material](#) online). The YLS genome contained a large number of transposable elements (TEs), accounting for approximately 2.93% of the whole genome, including DNA repeats (0.88%), long interspersed nuclear elements (0.99%), long terminal repeats (1.08%), short interspersed nuclear elements (0.002%), and unknown types of repeats (0.005%) ([supplementary table S3, Supplementary Material](#) online). The accumulation of mobile elements, representing a source of chromosomal rearrangements and gene inactivation, seems to have an important role in the first stage of genome reduction in endosymbionts (Ogata et al. 2001; Andersson et al. 2002).

Gene Prediction and Annotation

A total of 7,155 protein-coding genes were predicted (table 1), of which 6,970 genes encoded more than 100 amino acids. Genes cover at least 54.4% of the genome sequence, with an average gene density of 267 genes per Mb, which is similar to the palm aphid YLS (274), *Co. militaris* (257), *M. anisopliae* (271), and *M. acridum* (259). The average gene length is 2.04 kb, with an average of 3.3 exons per gene. Moreover, the distribution of gene length in the YLS, including the coding sequences, exons, and introns, is generally similar

Table 1

YLS Genome Features

Feature	Value
General	
Size (bp)	26,812,190
N50 (bp)	310,669
G+C content (%)	55.3
Protein-coding genes	7,155
tRNA genes	98
Percentage repeat rate (%)	15.5
Gene density (genes per Mb)	267
Exons per genes	3.3
Average gene size (bp)	2,040 (680 amino acids)
Predicted protein-coding sequences	
Identified by similarity to known sequences	2,893 (40.4%)
Conserved hypothetical proteins	3,559 (49.7%)
Predicted proteins (no similarity to known sequences)	703 (9.8%)

to that observed in *Aspergillus nidulans*, *Neurospora crassa*, and *Gibberella zeae*, whereas the distribution in *Saccharomyces cerevisiae* showed difference from the other species (supplementary fig. S1, Supplementary Material online). This result may reflect the fact that YLS of *N. lugens* belongs to the filamentous ascomycetes (subphylum Pezizomycotina), not to the true yeasts (subphylum Saccharomycotina) (Suh et al. 2001).

To annotate these predicted genes, we used BLASTX to search the sequences against the nonredundant (nr) NCBI protein database with a cutoff *E* value of 10^{-5} . As a result, 2,893 (40.4%) predicted genes were identified by similarity to known sequences. Furthermore, a total of 703 (9.8%) YLS protein-coding genes lack significant matches to known proteins from the nr database, and the other 3,559 (49.7%) genes are annotated as conserved hypothetical proteins (table 1), reflecting that the functions of many genes in YLS (59.5%) are still unknown and have yet to be described. Moreover, the genome has many contracted gene families, yet we detect expansion of a few families. Compared with other ascomycetes, the YLS of BPH has the lowest ratio of expanded gene families to contracted ones, with 78 expanded families and 368 contracted families (fig. 2).

We also used KEGG ontology (KO) assignments to classify the functions of the predicted YLS genes. Of all the predicted protein-coding genes, approximately 52.7% (3,769) could be annotated in the KO based on sequence homology. Among the KO categories, the cluster for "replication and repair" represented the largest group, followed by "folding, sorting, and degradation" and "transcription." The clusters for "carbohydrate metabolism" and "amino acid metabolism" were also relatively large (supplementary fig. S2, Supplementary Material online), which might be attributable to the

carbohydrate utilization and essential amino acid biosynthesis performed by the YLS for *N. lugens*.

Phylogenomic Analysis and Divergence Time Estimation

Phylogeny of fungal endosymbionts has been mainly analyzed based on partial sequences of ribosomal RNA genes (rDNA). Previous phylogenetic analyses implied that the YLS of planthoppers and that of the aphid *Tuberaphis styraci* both belonged to the class Sordariomycetes in the subphylum Pezizomycotina (Noda and Kawahara 1995; Fukatsu and Ishikawa 1996). More recently, by combining small-subunit and large-subunit rDNA sequences, phylogenetic analyses have placed the YLS of *N. lugens* in a more exact phylogenetic position: in the family Clavicipitaceae, within the filamentous ascomycetes (Suh et al. 2001), to which many entomopathogenic fungi also belong (e.g., *Metarhizium* and *Beauveria*).

In this study, a BSR test along with phylogenomic analysis was conducted. After classifying all putative peptides within three genomes (YLS of BPH, *M. acridum* and *Tolypocladium inflatum*, or YLS of BPH, YLS of palm aphid and *T. inflatum*) using a measure of similarity based on the ratio of BLAST scores, the output of the BSR analysis enables global visualization of the degree of proteome similarity between all three genomes. Using a BSR threshold value of 0.4, which was empirically determined to represent approximately 30% amino acid identity over approximately 30% of the peptide length (Rasko et al. 2005), the analysis revealed that 49.7% (3,560) of YLS genes were conserved when compared with those from *M. acridum* and *T. inflatum*. It also shows high degree of proteome similarity between *M. acridum* and *T. inflatum*, that is, had BSR values close to 1 (fig. 3a). Meanwhile, a different pattern was observed when comparing another three genomes. It showed that YLS of BPH is more closely related to YLS of palm aphid than *T. inflatum* as expected (fig. 3b). That is to say, these endosymbionts in insects have evolved individually owing to their special niches, different from other free-living ascomycete fungi such as *M. acridum* and *T. inflatum*. To get a better understanding of these results and pinpoint the lineage from which the endosymbionts were derived, a phylogenomic tree was constructed using 598 single-copy orthologous genes across 17 fungi (supplementary table S4, Supplementary Material online). The YLS of BPH, together with YLS of palm aphid, clustered with *T. inflatum* (asexual ascomycete fungus), a pathogen of beetle larvae in its sexual *Elaphocordyceps subsessilis* state. The endophyte *Epichloë festucae* and the arthropod pathogens *Metarhizium* spp. formed a monophyletic group, indicating that the split between the YLS and *Metarhizium* lineages occurred before *Metarhizium* diverged from the endophyte *Epichloë* lineage (fig. 3c). According to fossil calibrations documented previously (Lücking et al. 2009), split times were computed using the approximate likelihood calculation algorithm (fig. 3c) and demonstrated that the YLS diverged approximately

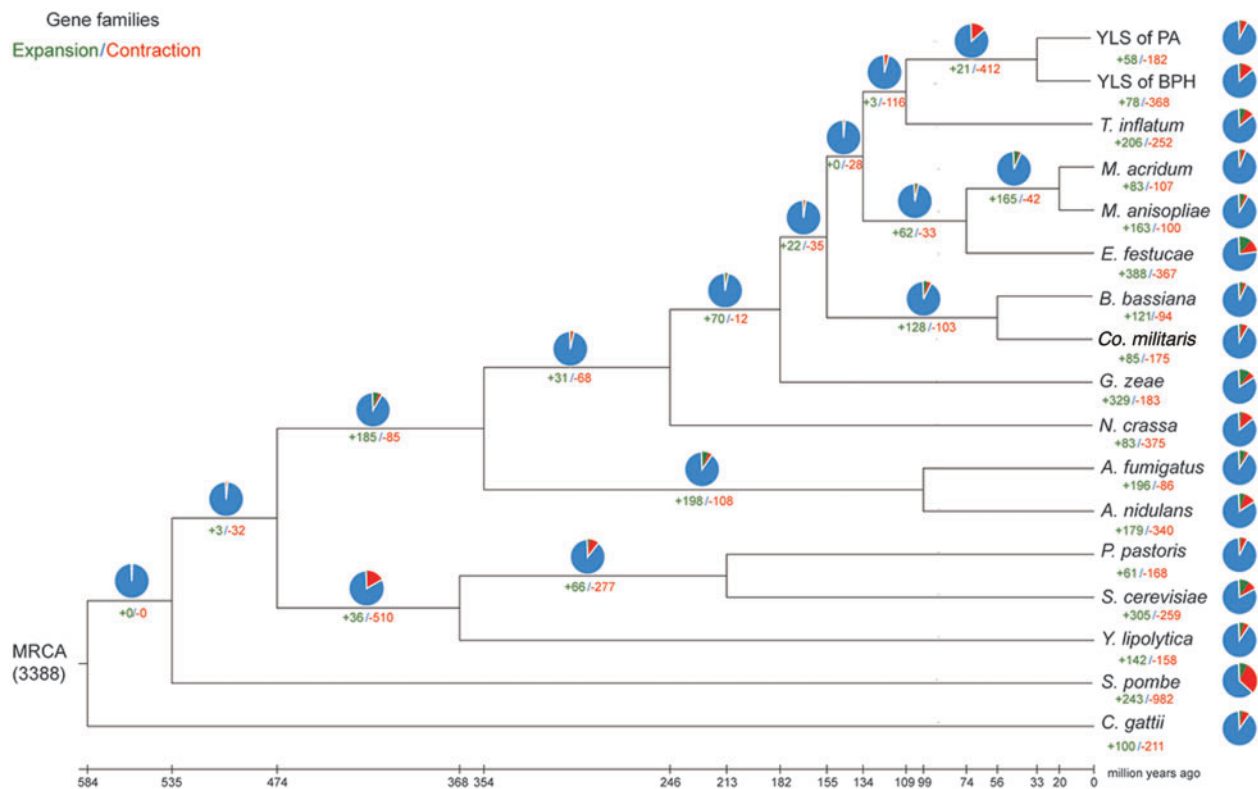


Fig. 2.—Gene family expansions and contractions in the YLS of BPH compared with other fungal genomes. Numbers for expanded (green) and contracted (red) gene families are shown below branches or taxon names with percentages indicated by pie charts.

77.4–149.5 Ma, which is as ancient as the association between obligate bacteria and insects (40–280 Ma) (Gibson and Hunter 2010).

Gene Orthology

Orthologous gene sets are important among species for tracing genomic changes along their evolution. We compared the annotated genes of the YLS of the BPH to those of 16 fungal species whose genome sequences are publicly available, and identified 6,388 genes belonging to 5,208 Treefam-method-defined gene families (supplementary table S5, Supplementary Material online). Among them, we further identified 1,126 single-copy orthologs and 361 multiple-copy orthologs belonging to 158 gene families from the YLS genome that were conserved among the 17 genomes examined (fig. 4a and supplementary table S4, Supplementary Material online). Meanwhile, it is noteworthy that the YLS in the BPH genome possesses more unique paralogs (127 genes belonging to 44 gene families) than the genomes of several closely related species, including *T. inflatum* (95 genes belonging to 42 gene families), *Co. militaris* (16 genes belonging to 7 gene families), *M. anisopliae* (48 genes belonging to 21 gene families), and *M. acridum* (33 genes belonging to 15 gene

families) (fig. 4a and supplementary table S4, Supplementary Material online). Furthermore, this barchart showed almost the same gene distribution in two YLS species. Considering the distinguishing biological roles and close relationship of these two YLS species, we also identified 253 gene families (including 560 genes in the YLS of the BPH and 396 genes in the YLS of the palm aphid) that were uniquely present in these two YLS species (supplementary table S4, Supplementary Material online). How these plenteous unique paralogs and YLS-conserved-only genes correlate with the YLS' biological functions is an exciting question.

For a better understanding of the lineage relationship between the YLS and insect pathogens, another independent orthologous analysis among five species was performed. The genes of the YLS of the BPH, together with the YLS of the palm aphid, were compared with three closely related insect pathogens and a Venn diagram of gene families was drawn. There are 3,761 conserved gene families shared among the five fungal genomes. Meanwhile, there are 780 gene families shared by the other three insect pathogens but absent in the YLS, which is very likely attributable to gene loss (fig. 4b). As a result of orthologous analysis, we also identified unclustered genes and unique paralogs in the YLS of the BPH, which were totally nominated as YLS-specific genes and, may contribute

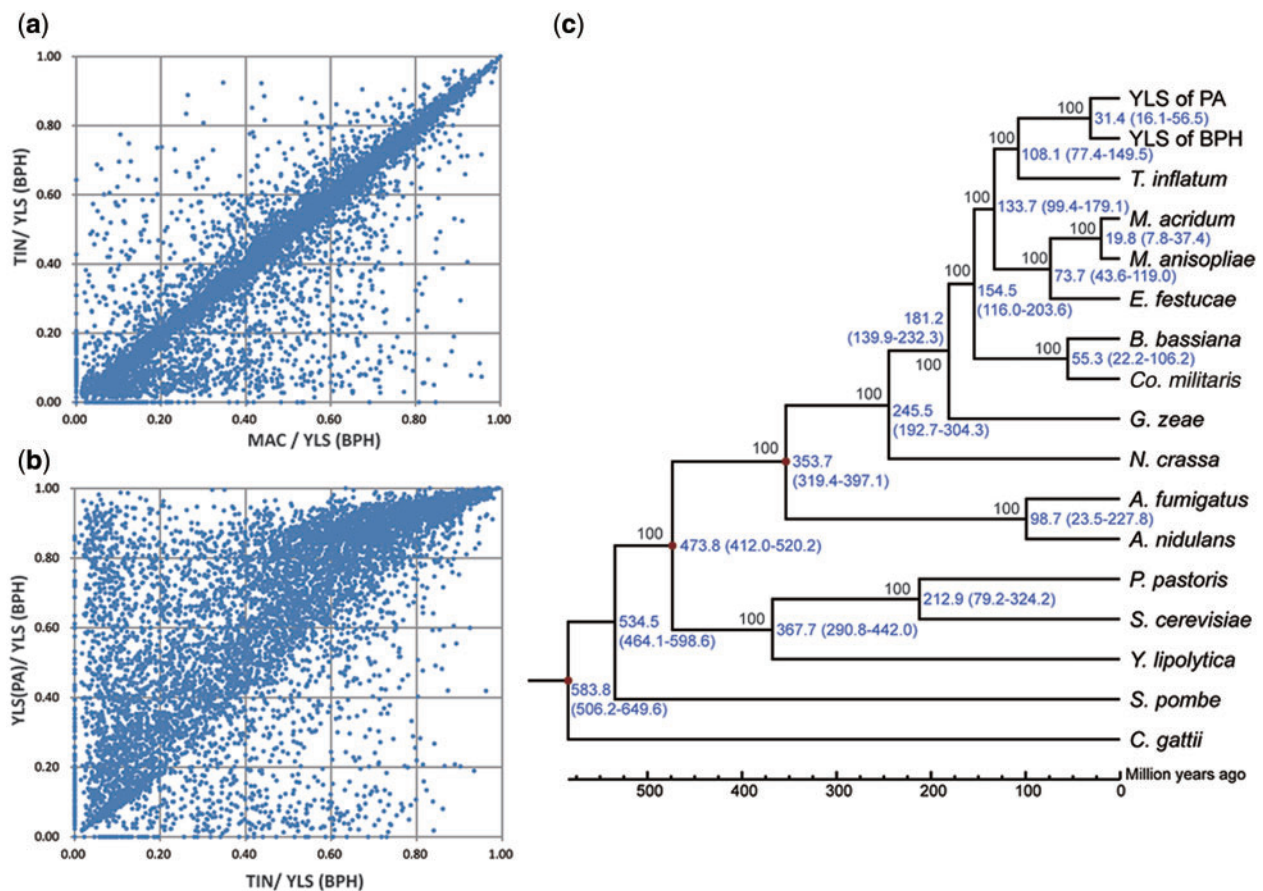


Fig. 3.—Scatter plots and phylogeny analysis. Scatter plots of BSR analysis of (a) *T. inflatum* (TIN), *M. acridum* (MAC) and YLS of BPH genomes, and (b) YLS of palm aphid, *T. inflatum* (TIN) and YLS of BPH genomes. (c) Estimation of divergence time and substitution rate by PAML mcmctree based on protein sequences of 17 species. The blue numbers on the nodes are the divergence time from present (Ma) and the black ones are the support values. The calibration times *Trichocomaceae*–*Sordariomycetes* divergence (320–400 Ma), *Pezizomycotina*–*Saccharomycotina* divergence (400–520 Ma), Ascomycota–Basidiomycota divergence (500–650 Ma) were derived from previously published paper (Lücking et al. 2009) and labeled with red dots. The genome sequences are identified with the following accession numbers: *Aspergillus fumigatus* (AAHF000000000.1), *Aspergillus nidulans* (AACD000000000), *Gibberella zeae* (AACM000000000.1), *Neurospora crassa* (AABX000000000), *Pichia pastoris* (PRJNA39439), *Saccharomyces cerevisiae* (PRJNA128), *Schizosaccharomyces pombe* (PRJNA127), *Yarrowia lipolytica* (PRJNA12414), YLS of BPH (JRM100000000), YLS of PA (AOPF000000000), *Tolypocladium inflatum* (AOHE000000000.1), *Cordyceps militaris* (AEVU000000000.1), *Metarhizium acridum* (ADNI000000000), *Metarhizium anisopliae* (ADNJ000000000.1), *Beauveria bassiana* (ADAH000000000.1), *Epichloë festucae* (AFRX000000000.1), and *Cryptococcus gattii* (PRJNA62089).

to its specific biological functions. A total of 777 YLS-specific genes were identified among these 17 fungi (26.6% of the whole gene set, including 127 duplicated genes and 650 unclustered genes), of which 369 genes (47.49%) were annotated by searching several databases (supplementary table S6, Supplementary Material online). Using the specific genes identified from the YLS of the BPH genome, an analysis of gene annotation in the KEGG pathway was performed. The YLS-specific genes were annotated mainly for metabolic processes (starch and sucrose metabolism, purine metabolism and cyanoamino acid metabolism, etc.) and degradation (polycyclic aromatic hydrocarbon degradation, limonene and pinene degradation, etc.; supplementary table S7, Supplementary Material online). The nutritional composition

of plant phloem sap is imbalanced; it is deficient in essential amino acids and nitrogen resources but is high and variable in carbohydrate content (Hayashi and Chino 1990). The YLS of *N. lugens* may play an important role in the complex and dynamic metabolism of carbohydrates. Genes related to purine metabolism may be involved in the processes of nitrogen recycling, which are the YLS's vital role in *N. lugens* nutrition (Xue et al. 2014).

Genes Lost and Retained

Genomic studies of bacterial endosymbionts have revealed that, during the evolution of endosymbiont-host associations, the bacterial endosymbiont genome becomes smaller owing

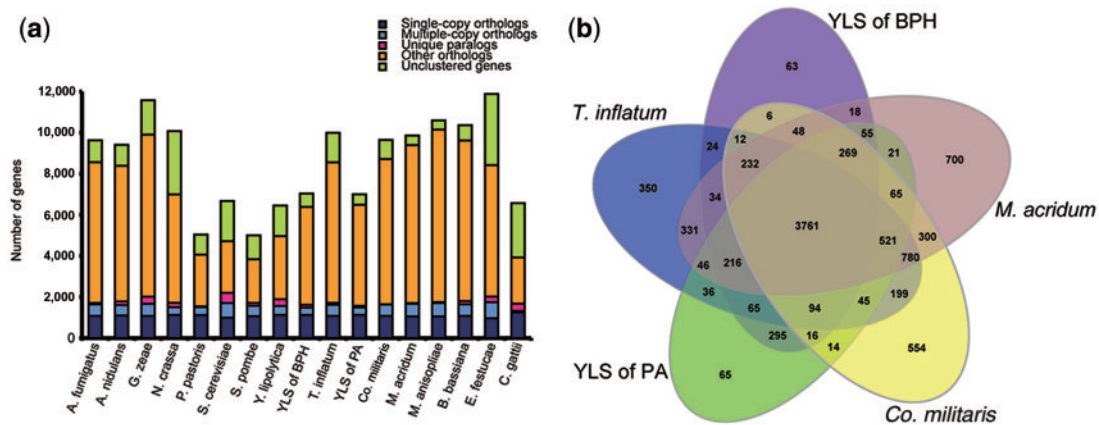


FIG. 4.—Comparative genomic analysis of several different species of fungi. (a) Protein orthology comparison among genomes of *A. fumigatus*, *A. nidulans*, *Gibberella zeae*, *Neurospora crassa*, *P. pastoris*, *S. cerevisiae*, *S. pombe*, *Y. lipolytica*, YLS of BPH, *T. inflatum*, YLS of PA, *Co. militaris*, *M. acridum*, *M. anisopliae*, *B. bassiana*, *Epichloë festucae*, and *C. gattii*. (b) Venn diagram showing shared orthologous groups among genomes of YLS of BPH, YLS of PA, *T. inflatum*, *Co. militaris*, and *M. acridum*.

to the loss of genes that are unnecessary in the new environment; however, a symbiotic organism must still retain a number of functions that allow its continued survival in the host (Moya et al. 2008). The full genome sequence of the YLS provides comprehensive information about the specific functions that are lost or retained as a result of intracellular associations. To detail the genomic changes in the YLS, we searched all predicted genes involved in Eukaryotic Orthologous Group (KOG) classifications, together with four other related free-living fungi. Based on the proportions of each gene cluster among these compared genomes, we found that genes involved in information storage and processing have been retained, including those that cluster under “replication, recombination, and repair” (193, 2.7%); “transcription” (212, 2.9%); and “translation, ribosomal structure, and biogenesis” (279, 3.9%). In contrast, the proportions of the genes clustered under “amino acid transport and metabolism” (196, 2.7%); “carbohydrate transport and metabolism” (152, 2.1%); and “secondary metabolite biosynthesis, transport, and catabolism” (116, 1.6%) are reduced, suggesting some of the genes in these categories have been lost (fig. 5 and [supplementary table S8, Supplementary Material](#) online).

Entomopathogens such as *Metarhizium* spp. need to penetrate the protein- and chitin-rich insect cuticle and solubilize host tissues for nutrition. Therefore, they would be expected to secrete large numbers of degradative enzymes (Gao et al. 2011). Because of the closeness of the YLS and entomopathogen lineages, we specifically selected gene families related to cuticle degradation to observe the gene changes that occurred in the YLS. There was a large difference in the number of genes in the trypsin family of proteases between YLS and *Metarhizium* spp., with only one gene in the YLS

compared with 32 in *M. anisopliae* and 17 in *M. acridum* (table 2). However, this difference may be the result of an increased number of trypsin genes in *Metarhizium* spp. because other fungi (including plant pathogens) did not have many genes of the trypsin family. Additionally, subtilisins, which assist in infection processes by degrading host cuticles and disabling antimicrobial peptides (Bagga et al. 2004), are almost 4–5 times more abundant in *M. anisopliae* (55) and *M. acridum* (43) than in the YLS (11). Moreover, the sizes of the gene families of other proteases (including serine proteases, aspartic proteases, threonine proteases, and cysteine proteases) are small in the YLS compared with *Metarhizium* spp. (table 2). Though the differences between the YLS and the entomopathogens are clear, they are not specific features in the YLS when compared with other fungi. Table 2 illustrates the apparent reduction in the members observed in Zn2/Cys6 transcription factors when compared with ten other ascomycetes. Zn2/Cys6 transcription factors are unique to fungi and are involved in different regulatory functions (Zhao et al. 2011; Lu et al. 2014). The reduction of these fungi-specific transcription factors in the YLS could be regarded as a consequence of the accumulation of deleterious mutations and gene loss through genetic drift. The stable intracellular niche in host insects allows the loss of several regulatory functions of endosymbionts.

The YLS of the BPH appears to propagate by budding in the mycetocytes. Unlike other ascomycetes, its sexual reproduction has not been identified by histological observation (Noda 1977; Cheng and Hou 2001). Genome alignment and annotation showed that there are no mating type (*MAT*) genes present in the sequenced YLS genome, while synteny analysis with *B. bassiana*, *Co. militaris*, and *Metarhizium robertsii* showed that, except for the idiomorphic regions, the genes

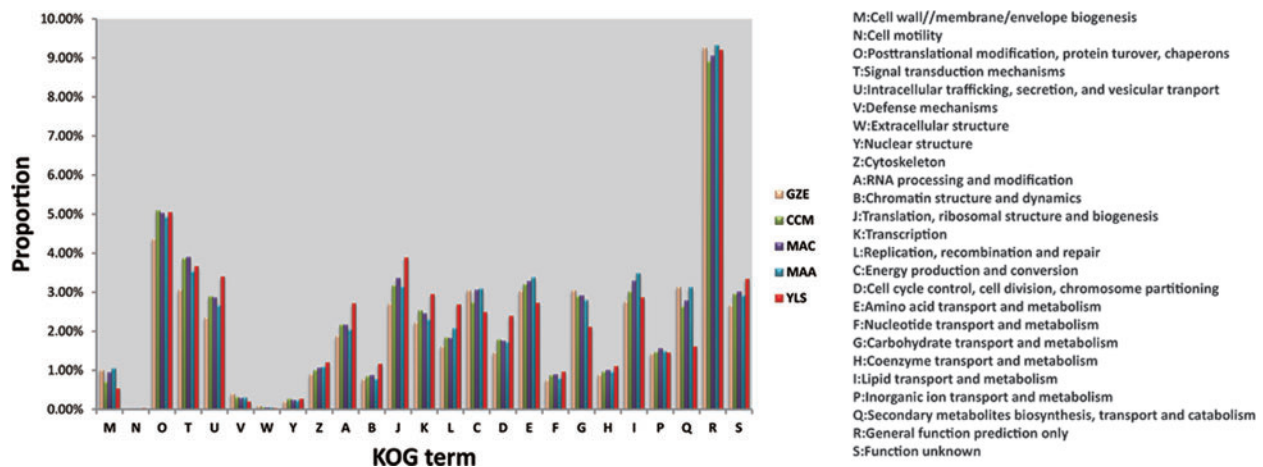


FIG. 5.—Comparative analysis of genes by functional KOG categories, from the genomes of *G. zeae* (GZE), *Co. militaris* (CCM), *M. acridum* (MAC), *M. anisopliae* (MAA), and YLS.

Table 2

Sizes of Selected Gene Families in YLS and Other Ascomycetes

Protein Family	YLS	MAA	MAC	EF	FG	MO	BC	SS	NC	AN	AF
Fungal specific transcription factor	93	114	94	68	192	95	118	90	89	209	169
C2H2 zinc finger transcription factor	42	86	53	45	85	58	48	54	63	58	51
Zn2/Cys6 transcription factor	39	202	174	107	376	155	142	108	110	307	230
Major facilitator superfamily	110	269	236	121	274	198	225	167	110	279	232
Cytochrome P450	52	123	100	57	112	137	129	93	40	116	74
Pth11-like G-protein coupled receptor	39	54	40	20	51	60	22	23	28	39	15
Protein kinase	110	151	183	158	129	129	124	164	111	127	131
Histidine kinase	10	10	9	13	20	6	3	5	8	12	6
Heterokaryon incompatibility	6	36	25	6	88	41	59	34	45	7	8
Serine protease	15/43	83/182	60/154	38/109	60/150	56/91	19/34	20/33	32/74	53/136	29/46
Subtilisin	7/11	41/55	28/43	13/18	16/24	26/29	4/7	4/6	6/10	3/4	3/7
Trypsin	0/1	27/32	9/17	1/4	2/3	3/3	1/1	1/1	0/2	1/2	0/0
Carboxypeptidase	5/7	15/25	13/26	7/7	12/21	7/8	7/9	8/11	6/9	5/12	14/15
Aspartic protease	4/16	24/33	19/25	9/12	15/18	14/19	11/14	9/21	15/19	7/16	7/9
Threonine protease	1/11	4/23	3/21	1/21	3/18	2/18	2/13	2/13	2/20	0/20	1/17
Cysteine protease	1/28	4/46	4/47	6/49	5/57	4/31	3/24	1/27	4/41	6/57	3/31
Metalloprotease	6/47	17/95	18/87	15/83	32/111	38/91	6/50	7/48	21/81	22/105	20/77
Lipase	3/18	12/32	5/23	7/20	4/31	2/23	3/28	2/25	0/16	2/27	3/25
Esterase/thioesterase	64	66	61	59	70	64	70	58	42	63	52
Glycoside hydrolase related	90	156	140	73	159	198	120	126	137	200	165
Transposase	97	148	20	1	17	15	73	426	15	15	109
Cutinase	1	2	2	5	12	18	11	8	3	4	5
Pectin lyase	8	7	7	6	25	9	25	20	5	24	27

NOTE.—Fungal species: MAA, *Metarhizium anisopliae*; MAC, *M. acridum*; EF, *Epichloë festucae*; FG, *Fusarium graminearum*; MO, *Magnaporthe oryzae*; BC, *Botrytis cinerea*; SS, *Sclerotinia sclerotiorum*; NC, *Neurospora crassa*; AN, *Aspergillus nidulans*; and AF, *Aspergillus fumigatus*. The top of fractions is the number of total proteins in each family that are secreted and the bottom is the total number of genes in each family. Secretory proteins were identified by using the program SignalP under default parameters (Petersen et al. 2011).

flanking the *MAT* locus are highly conserved (fig. 6). The *MAT* locus is replaced with a gene coding for a retrotransposable element Tf2 protein-like protein (RTF); simultaneously, the flanking region was inverted, strongly indicating the loss of

sexual reproduction-related genes through transposition and degradation during evolution. Additionally, the strict asexuality of this endosymbiont lineage exacerbates the effects of genetic drift in their small populations by preventing the

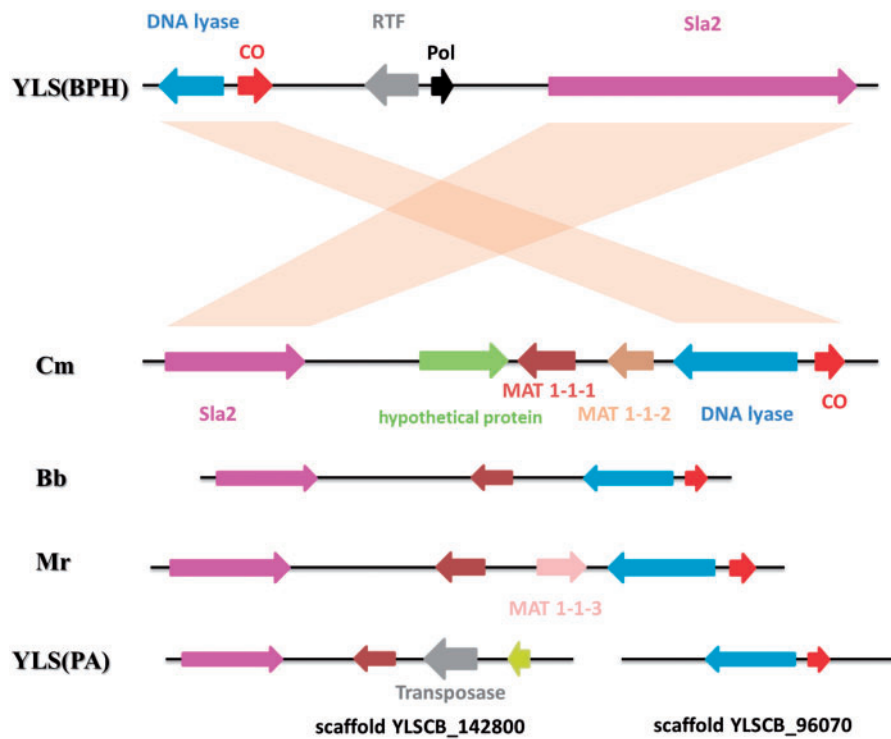


Fig. 6.—Comparative analysis and syntenic relationship of the *MAT* loci and their flanking regions between YLS of BPH, *B. bassiana* (Bb), *Co. militaris* (Cm), *M. robertsii* (Mr), and YLS of PA. Loci in the same color show orthologous relationships.

recovery of wide-type genotypes through recombination (Wernegreen 2002), resulting in the accumulation of deleterious mutations and eventual gene loss. Moreover, we analyzed the genome structure correlated with the *MAT* in YLS of palm aphid. The mating-type elements were separated into two scaffolds, although *MAT1-1-1* gene was present in the genome (fig. 6). A gene coding for a transposase was also found adjoining the *MAT* loci, which was probably in charge of the transposition of *MAT* elements. A similar change in genetic elements by transposon insertion was found in both YLS species.

Discussion

Here, we report a genomic analysis of the YLS in *N. lugens*, providing the fully sequenced genome of a eukaryotic insect endosymbiont. This study revealed some aspects that seem to be related to an endosymbiotic lifestyle and based on a long evolutionary history. Sexual reproduction-related genes, for example, were not found in the planthopper YLS (fig. 6). Genetic exchange of the planthopper YLS is prevented, which seems to have caused mutations to accumulate and promoted faster evolution of the YLS (supplementary fig. S3, Supplementary Material online). This study suggests that obligate fungal endosymbionts in insects undergo a different pattern of DNA sequence evolution than obligate bacterial

endosymbionts when establishing a symbiotic relationship. Obligate bacterial endosymbionts usually show the following main features including genome size reduction, low GC content, and limited metabolic abilities due to the accumulation of deleterious changes by genetic drift, a mutational bias toward AT base pairs, and specialization in a symbiotic relationship, respectively (Wernegreen 2002; Moran et al. 2008; Moya et al. 2008).

The YLS genome size (26.8Mb) was smaller than some ascomycete fungi (fig. 1) and the reduction in size was not as pronounced in bacterial endosymbionts, which possess a fraction of the genome of free-living bacterial relatives (Shigenobu et al. 2000; McCutcheon and Moran 2012). Bacterial genome reduction is considered to be caused by the early spread of TE activity resulting in inactivation of non-essential genes and large deletions, including pseudogene erosion (Ogata et al. 2001; Silva et al. 2001; Andersson et al. 2002; Mira and Moran 2002; Vogel and Moran 2013). Accordingly, the genomes of the present obligatory bacterial endosymbionts possess fewer TEs and an extremely slimmed down genome. Two main generally accepted trends in genome size, which work in opposition, seem to affect insect YLS. One is genome reduction similar to obligatory bacterial endosymbionts, as many nonessential genes required for free-living are lost. The other is genome expansion, which is observed in eukaryotic species with small effective population

sizes as they can tolerate the slightly deleterious accumulation of extra DNA in the form of TEs, multiple introns, and gene duplications (Lynch and Conery 2003; Lynch 2006). Genome reduction was not so obvious (as mentioned above), and TEs were rich in the obligatory planthopper YLS (supplementary table S3, Supplementary Material online), accounting for 3.35% of the whole genome. Then, did the genome expansion occur in the planthopper YLS? Vogel and Moran (2013) did detect the proliferation of introns and accelerated evolutionary rates in the YLS genome of *Ce. brasiliensis*, which appeared to fit the patterns of evolution suggested by Kelkar and Ochman (2011) for Pezizomycotina experiencing genetic drift. However, the genome of YLS of the palm aphid showed smaller intergenic spacers and higher gene density than two comparison fungi genomes, and its genome size was 25.4 Mb, which is close to that of the planthopper YLS. The authors suggested that the relatively young association between the YLS and hosts may not have permitted sufficient time for the type of genome expansion observed in many eukaryotes. Actually, the YLS is considered to be a successor of ancient bacterial endosymbionts; the aphid YLS appear to have replaced *Buchnera*, which is the primary bacterial endosymbiont of aphids. It has been identified that some genes have been lost during the evolution of the planthopper YLS, but the reduction in genome size was not large. Signs of genome expansion are not clear in YLSs. Scarce data for fungus-insect associations prevent the development of a comprehensive perspective of the evolution of the YLS genome.

GC content of the planthopper YLS was 55%, and that of *Ce. brasiliensis* YLS was 54%. Though these values are higher than those of many ascomycete fungi, some fungi show much higher GC content (fig. 1), which suggests that GC content is within the range of ordinal ascomycete fungi. However, obligatory bacterial symbionts of insects usually have extremely AT rich genomes (McCutcheon and Moran 2012). This is another large difference between obligatory bacterial symbionts and fungal symbionts. In a large range of eukaryotes, a mechanism termed GC-biased gene conversion that follows meiotic recombination accounts for large-scale variations in GC content (Eyre-Walker and Hurst 2001). However, for organisms belonging to ancient asexual lineages, such as the pathogenic fungi *Candida albicans* and *Candida dubliniensis*, two models including mitotic recombination and replicative errors were also proposed to explain GC content evolution (Marsolier and Marie 2013). In this study, we found that YLSs do not possess MAT genes and do not have sexual phase or even the parasexual cycle that *Ca. albicans* and *Ca. dubliniensis* have. Although we cannot yet explain the GC content variation in the planthopper YLS, the YLS symbiont can become a new model for further the understanding of GC content evolution in eukaryotes.

Symbiosis is a dynamic process in which the symbiont experiences many genotypic and phenotypic changes compared with its free-living relatives. The categories of planthopper YLS

genes were compared with those of some fungi. Given that the same situation is encountered in symbiotic bacteria (Gil et al. 2003; Moran 2003), it is very likely that genes involved in essential functions—such as those related to DNA replication, transcription and translation—are retained in the symbionts, whereas genes that are involved in nutritional metabolism but are not required for host survival are lost due to the stable nutritional support of the host. Some YLS-specific genes were mainly involved in metabolism, degradation, and biosynthesis, reflecting the special role of the YLS plays in providing the host with metabolic capabilities. Metabolic features, which have been mainly discussed in the planthopper YLS, are sterol biosynthesis (Noda and Koizumi 2003), nitrogen recycling (Sasaki et al. 1996; Hongoh et al. 2000), and amino acid biosynthesis (Xue et al. 2014).

YLSs synthesize ergosta-5,7,24(28)-trienol (Wetzel et al. 1992), which is converted into 24-methylenecholesterol following cholesterol synthesis by host planthoppers (Noda and Koizumi 2003; Xue et al. 2014). YLSs can supply sterols for insect hosts but bacterial symbionts cannot, because bacteria do not have a sterol biosynthetic pathway. Moreover, planthoppers feed on nitrogen-poor food, and recycling the nitrogen is beneficial for their survival. Both BPH and YLS encode the uricase gene and the nitrogenous waste product can be used for nitrogen recycling by planthoppers with the aid of YLSs (Hongoh et al. 2000; Xue et al. 2014). Nitrogen recycling has also been reported in the intracellular bacterial symbiont of cockroaches, *Blattabacterium* (Sabree et al. 2009). Amino acid synthesis and supply by symbionts are well known phenomena in hemipteran insects. The primary symbionts of the aphid *Buchnera* synthesize essential amino acids (Shigenobu et al. 2000). In many hemipteran species, endosymbionts make amino acids utilizable for the host insects (Wu et al. 2006; McCutcheon and Moran 2007). Two bacterial symbionts often work together for synthesizing all essential amino acids (McCutcheon et al. 2009; McCutcheon and von Dohlen 2011). In BPH, YLS can synthesize all essential amino acids (Xue et al. 2014), revealing that YLS is a capacious partner in supplying nutrients for host insects.

Bacterial endosymbionts are reported to be in many planthoppers. *Sulcia muelleri* and *Vidania fulgoroideae* are two ancient bacterial endosymbionts in planthoppers (Urban and Cryan 2012). *Purcellia pentastirorum* is also present in some planthopper species (Bressan et al. 2009; Gonella et al. 2011). However, economically important species of planthoppers, for example, the corn planthopper *Peregrinus maidis* and rice planthoppers, *N. lugens*, *Sogatella furcifera*, and *Laodelphax striatellus*, which are members of the family Delphacidae, harbor YLS. Facultative bacterial endosymbionts, for example, *Wolbachia* and *Cardinium*, are often found in delphacid planthoppers (Nakamura et al. 2009; Nakamura et al. 2012). However, those ancient bacterial endosymbionts, such as *Sulcia* and *Vidania*, have not been reported in these delphacid planthoppers. The most reasonable idea for the

symbiosis between planthoppers and the YLS is that a fungus species had an association in the time of the ancestral species of Delphacidae and the obligatory bacterial symbionts diminished after symbiosis was established between the planthopper and the fungus. The most important role of the obligatory bacterial endosymbionts seems to be as a nutritional supply for the host (Moran et al. 2008; McCutcheon et al. 2009; McCutcheon and von Dohlen 2011). The YLS genome shows that YLS has the abilities to synthesize essential amino acids and recycle nitrogen compounds, which are also found in bacterial endosymbionts. In addition, YLSs provide sterols, which bacterial symbionts cannot contribute directly to synthesize (Gibson and Hunter 2010). Phylogenetic analyses suggest that the YLS was closely related to several entomopathogens (Jones et al. 1999). Divergence time estimation revealed that the YLS diverged from the ancestors of *Cordyceps*, *Beauveria*, and *Metarhizium* approximately 99–203 Ma. Though the lineage from which the YLS is derived is still unclear, the closest relative is *T. inflatum*, which is widely distributed in soils and also a pathogen of beetles. Fungal endosymbionts have also been reported in diverse insect hosts (Buchner 1965; Gibson and Hunter 2010), such as scale insects (Vashishtha et al. 2011) and several families of beetles (Coleoptera: Anobiidae, Cerambycidae, and Scolytidae) (Jones et al. 1999). The bacterial symbionts had been extremely specialized, which led to a lower capacity for variation. Association with YLS might have provided a new opportunity for different evolution.

Based on the sufficient genetic information in this research, this fungal endosymbiont in planthoppers should be properly named, which would be helpful for further identification and communication. We have looked over the rules of fungal nomenclature and consulted several experts (Hawksworth et al. 2011). Finally, we have decided to adopt the Latin name "*Entomomyces delphacidicola*." The prefix "Entomo-" is short for "entomological," the suffix "-myces" means fungi and "-cola" is from Latin noun *incola* which means inhabitant. *Entomomyces delphacidicola* can be adopted to name the fungal endosymbionts that present in delphacidae insects. Symbionts from different host species are considered strains of the symbiont species, such as YLS from *N. lugens* being recognized as *E. delphacidicola* str. NLU.

Supplementary Material

Supplementary figures S1–S3 and tables S1–S8 are available at *Genome Biology and Evolution* online (<http://www.gbe.oxfordjournals.org/>).

Acknowledgments

This work was supported by grants from the National Science Foundation of China (31272374 and 31471765) and the National Basic Research Program of China (973 Program,

No. 2010CB126205). The authors thank Prof. Bo Huang in Anhui Agricultural University and Prof. Cheng-shu Wang in Institute of Plant Physiology and Ecology, Chinese Academy of Sciences for putting forward valuable comments on fungal nomenclature.

Literature Cited

- Andersson SG, et al. 2002. Comparative genomics of microbial pathogens and symbionts. *Bioinformatics* 18:517–517.
- Bagga S, Hu G, Screen SE, St Leger RJ. 2004. Reconstructing the diversification of subtilisins in the pathogenic fungus *Metarhizium anisopliae*. *Gene* 324:159–169.
- Bressan A, Arneodo J, Simonato M, Haines WP, Boudon-Padieu E. 2009. Characterization and evolution of two bacteriome-inhabiting symbionts in cixiid planthoppers (Hemiptera: Fulgoromorpha: Pentastirini). *Environ Microbiol.* 11:3265–3279.
- Buchner P. 1965. Endosymbiosis of animals with plant microorganisms. New York: Interscience Inc.
- Cheng DJ, Hou RF. 2001. Histological observations on transovarial transmission of a yeast-like symbiote in *Nilaparvata lugens* Stal (Homoptera, Delphacidae). *Tissue Cell* 33:273–279.
- De Bie T, Cristianini N, Demuth JP, Hahn MW. 2006. CAFE: a computational tool for the study of gene family evolution. *Bioinformatics* 22:1269–1271.
- Eyre-Walker A, Hurst LD. 2001. The evolution of isochores. *Nat Rev Genet.* 2:549–555.
- Fukatsu T, Ishikawa H. 1996. Phylogenetic position of yeast-like symbiont of *Hamiltonaphis styraci* (homoptera, aphididae) based on 18S rDNA sequence. *Insect Biochem Mol Biol.* 26:383–388.
- Gao Q, et al. 2011. Genome sequencing and comparative transcriptomics of the model entomopathogenic fungi *Metarhizium anisopliae* and *M. acridum*. *PLoS Genet.* 7:e1001264.
- Gibson CM, Hunter MS. 2010. Extraordinarily widespread and fantastically complex: comparative biology of endosymbiotic bacterial and fungal mutualists of insects. *Ecol Lett.* 13:223–234.
- Gil R, et al. 2003. The genome sequence of *Blochmannia floridanus*: comparative analysis of reduced genomes. *Proc Natl Acad Sci U S A.* 100:9388–9393.
- Gonella E, et al. 2011. Bacterial endosymbiont localization in *Hyalesthes obsoletus*, the insect vector of Bois noir in *Vitis vinifera*. *Appl Environ Microbiol.* 77:1423–1435.
- Guindon S, et al. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol.* 59:307–321.
- Guindon S, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol.* 52:696–704.
- Hawksworth DL, et al. 2011. The amsterdam declaration on fungal nomenclature. *IMA Fungus* 2:105–112.
- Hayashi H, Chino M. 1990. Chemical composition of phloem sap from the uppermost internode of the rice plant. *Plant Cell Physiol.* 31:247–251.
- Hongoh Y, Sasaki T, Ishikawa H. 2000. Cloning, sequence analysis and expression in *Escherichia coli* of the gene encoding a uricase from the yeast-like symbiont of the brown planthopper, *Nilaparvata lugens*. *Insect Biochem Mol Biol.* 30:173–182.
- Jones KG, Dowd PF, Blackwell M. 1999. Polyphyletic origins of yeast-like endocytobionts from anobiid and cerambycid beetles. *Mycol Res.* 103:542–546.
- Kelkar YD, Ochman H. 2011. Causes and consequences of genome expansion in fungi. *Genome Biol Evol.* 4:13–23.
- Li H, et al. 2006. TreeFam: a curated database of phylogenetic trees of animal gene families. *Nucleic Acids Res.* 34:D572–D580.

- Lu J, Cao H, Zhang L, Huang P, Lin F. 2014. Systematic analysis of Zn(2)Cys(6) transcription factors required for development and pathogenicity by high-throughput gene knockout in the rice blast fungus. *PLoS Pathog.* 10:e1004432.
- Lücking R, Huhndorf S, Pfister DH, Plata ER, Lumbsch HT. 2009. Fungi evolved right on track. *Mycologia* 101:810-822.
- Lynch M. 2006. The origins of eukaryotic gene structure. *Mol Biol Evol.* 23:450-468.
- Lynch M, Conery JS. 2003. The origins of genome complexity. *Science* 302:1401-1404.
- Marsolier K, Marie C. 2013. Models for the evolution of GC content in asexual fungi *Candida albicans* and *C. dubliniensis*. *Genome Biol Evol.* 5:2205-2216.
- McCutcheon JP, McDonald BR, Moran NA. 2009. Convergent evolution of metabolic roles in bacterial co-symbionts of insects. *Proc Natl Acad Sci U S A.* 106:15394-15399.
- McCutcheon JP, Moran NA. 2007. Parallel genomic evolution and metabolic interdependence in an ancient symbiosis. *Proc Natl Acad Sci U S A.* 104:19392-19397.
- McCutcheon JP, Moran NA. 2012. Extreme genome reduction in symbiotic bacteria. *Nat Rev Microbiol.* 10:13-26.
- McCutcheon JP, von Dohlen CD. 2011. An interdependent metabolic patchwork in the nested symbiosis of mealybugs. *Curr Biol.* 21:1366-1372.
- Mira A, Moran N. 2002. Estimating population size and transmission bottlenecks in maternally transmitted endosymbiotic bacteria. *Microb Ecol.* 44:137-143.
- Moran NA. 2003. Tracing the evolution of gene loss in obligate bacterial symbionts. *Curr Opin Microbiol.* 6:512-518.
- Moran NA, McCutcheon JP, Nakabachi A. 2008. Genomics and evolution of heritable bacterial symbionts. *Annu Rev Genet.* 42:165-190.
- Moran NA, Munson MA, Baumann P, Ishikawa H. 1993. A molecular clock in endosymbiotic bacteria is calibrated using the insect hosts. *Proc R Soc Lond B Biol Sci.* 253:167-171.
- Moya A, Peretó J, Gil R, Latorre A. 2008. Learning how to live together: genomic insights into prokaryote-animal symbioses. *Nat Genet.* 9:218-229.
- Nakamura Y, et al. 2009. Prevalence of *Cardinium* bacteria in planthoppers and spider mites and taxonomic revision of "*Candidatus Cardinium hertigii*" based on detection of a new *Cardinium* group from biting midges. *Appl Environ Microb.* 75:6757-6763.
- Nakamura Y, Yukuhiro F, Matsumura M, Noda H. 2012. Cytoplasmic incompatibility involving *Cardinium* and *Wolbachia* in the white-backed planthopper *Sogatella furcifera* (Hemiptera: Delphacidae). *Appl Entomol Zool.* 47:273-283.
- Noda H. 1977. Histological and histochemical observation of intracellular yeastlike symbionts in the fat body of the smaller brown planthopper, *Laodelphax striatellus* (Homoptera: Delphacidae). *Appl Entomol Zool.* 12:134-141.
- Noda H, Kawahara N. 1995. Electrophoretic karyotype of intracellular yeast-like symbionts in rice planthoppers and anobiid beetles. *J Invertebr Pathol.* 65:118-124.
- Noda H, Koizumi Y. 2003. Sterol biosynthesis by symbionts: cytochrome P450 sterol C-22 desaturase genes from yeastlike symbionts of rice planthoppers and anobiid beetles. *Insect Biochem Mol Biol.* 33:649-658.
- Noda H, Omura T. 1992. Purification of yeast-like symbionts of planthoppers. *J Invertebr Pathol.* 59:104-105.
- Ogata H, et al. 2001. Mechanisms of evolution in *Rickettsia conorii* and *R. prowazekii*. *Science* 293:2093-2098.
- Petersen TN, Brunak S, von Heijne G, Nielsen H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods.* 8:785-786.
- Quinn NL, et al. 2008. Assessing the feasibility of GS FLX pyrosequencing for sequencing the Atlantic salmon genome. *BMC Genomics* 9:404.
- Rambaut A, Drummond A. 2009. Tracer v1.5: an MCMC trace analysis tool. Available from: <http://beast.bio.ed.ac.uk/> (cited December 1, 2009).
- Rannala B, Yang Z. 2007. Inferring speciation times under an episodic molecular clock. *Syst Biol.* 56:453-466.
- Rasko DA, Myers GS, Ravel J. 2005. Visualization of comparative genomic analyses by BLAST score ratio. *BMC Bioinformatics* 6:2.
- Rose MD, Winston FM, Heiter P. 1990. *Methods in yeast genetics: a laboratory course manual.* New York: Cold Spring Harbor Laboratory Press.
- Ruan J, et al. 2008. TreeFam: 2008 update. *Nucleic Acids Res.* 36:D735-D740.
- Sabree ZL, Kambhampati S, Moran NA. 2009. Nitrogen recycling and nutritional provisioning by *Blattabacterium*, the cockroach endosymbiont. *Proc Natl Acad Sci U S A.* 106:19521-19526.
- Sasaki T, Kawamura M, Ishikawa H. 1996. Nitrogen recycling in the brown planthopper, *Nilaparvata lugens*: involvement of yeast-like endosymbionts in uric acid metabolism. *J Insect Physiol.* 42:125-129.
- Shigenobu S, Watanabe H, Hattori M, Sakaki Y, Ishikawa H. 2000. Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera sp.* APS. *Nature* 407:81-86.
- Silva FJ, Latorre A, Moya A. 2001. Genome size reduction through multiple events of gene disintegration in *Buchnera* APS. *Trends Genet.* 17:615-618.
- Suh SO, Noda H, Blackwell M. 2001. Insect symbiosis: derivation of yeast-like endosymbionts within an entomopathogenic filamentous lineage. *Mol Biol Evol.* 18:995-1000.
- Thao ML, et al. 2000. Cospeciation of psyllids and their primary prokaryotic endosymbionts. *Appl Environ Microbiol.* 66:2898-2905.
- Urban JM, Cryan JR. 2012. Two ancient bacterial endosymbionts have coevolved with the planthoppers (Insecta: Hemiptera: Fulgoroidea). *BMC Evol Biol.* 12:87.
- Vashishtha A, Sharama K, Lakhnypaul S. 2011. Co-existence, phylogeny and putative role of *Wolbachia* and yeast-like symbiont (YLS) in *Kerria lacca* (Kerr). *Curr Microbiol.* 63:206-212.
- Vogel KJ, Moran NA. 2013. Functional and evolutionary analysis of the genome of an obligate fungal symbiont. *Genome Biol Evol.* 5:891-904.
- Wernegreen JJ. 2002. Genome evolution in bacterial endosymbionts of insects. *Nat Rev Genet.* 3:850-861.
- Wetzel JM, et al. 1992. Diversity in steroidogenesis of symbiotic microorganisms from planthoppers. *J Chem Ecol.* 18:2083-2094.
- Wu D, et al. 2006. Metabolic complementarity and genomics of the dual bacterial symbiosis of sharpshooters. *PLoS Biol.* 4:e188.
- Xiao G, et al. 2012. Genomic perspectives on the evolution of fungal entomopathogenicity in *Beauveria bassiana*. *Sci Rep.* 2:483.
- Xue J, et al. 2014. Genomes of the rice pest brown planthopper and its endosymbionts reveal complex complementary contributions for host adaptation. *Genome Biol.* 15:521.
- Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol.* 24:1586-1591.
- Yang Z, Rannala B. 2006. Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. *Mol Biol Evol.* 23:212-226.
- Zhao C, Waalwijk C, de Wit PJ, van der Lee T, Tang D. 2011. EBR1, a novel Zn2Cys6 transcription factor, affects virulence and apical dominance of hyphal tip in *Fusarium graminearum*. *Mol Plant Microbe Interact.* 24:1407-1418.

Associate editor: Daniel Sloan