J. Microbiol. Biotechnol. (2015), 25(9), 1410–1416
http://dx.doi.org/10.4014/jmb.1503.03097

jmb

Wolbachia Sequence Typing in Butterflies Using Pyrosequencing

Sungmi Choi[1], Su-Kyoung Shin[1], Gilsang Jeong[2], and Hana Yi[1,3]*

[1]*BK21PLUS Program in Embodiment: Health-Society Interaction, Department of Public Health Sciences, Graduate School, Korea University, Seoul 136-701, Republic of Korea*
[2]*Department of Ecology and Evolution, National Institute of Ecology, Seocheon-gun, 325-813, Republic of Korea*
[3]*School of Biosystem and Biomedical Science, Korea University, Seoul 136-703, Republic of Korea*

Received: March 26, 2015
Revised: June 29, 2015
Accepted: July 1, 2015

First published online
July 2, 2015

**Corresponding author*
Phone: +82-2-3290-5644;
Fax: +82-303-940-2862;
E-mail: hanayi@korea.ac.kr

Supplementary data for this paper are available on-line only at http://jmb.or.kr.

pISSN 1017-7825, eISSN 1738-8872

Wolbachia is an obligate symbiotic bacteria that is ubiquitous in arthropods, with 25–70% of insect species estimated to be infected. *Wolbachia* species can interact with their insect hosts in a mutualistic or parasitic manner. Sequence types (ST) of *Wolbachia* are determined by multilocus sequence typing (MLST) of housekeeping genes. However, there are some limitations to MLST with respect to the generation of clone libraries and the Sanger sequencing method when a host is infected with multiple STs of *Wolbachia*. To assess the feasibility of massive parallel sequencing, also known as next-generation sequencing, we used pyrosequencing for sequence typing of *Wolbachia* in butterflies. We collected three species of butterflies (*Eurema hecabe, Eurema laeta*, and *Tongeia fischeri*) common to Korea and screened them for *Wolbachia* STs. We found that *T. fischeri* was infected with a single ST of *Wolbachia*, ST41. In contrast, *E. hecabe* and *E. laeta* were each infected with two STs of *Wolbachia*, ST41 and ST40. Our results clearly demonstrate that pyrosequencing-based MLST has a higher sensitivity than cloning and Sanger sequencing methods for the detection of minor alleles. Considering the high prevalence of infection with multiple *Wolbachia* STs, next-generation sequencing with improved analysis would assist with scaling up approaches to *Wolbachia* MLST.

Keywords: *Wolbachia*, *Eurema*, *Tongeia*, MLST, *wsp*, pyrosequencing

Introduction

Wolbachia is an obligate intracellular symbiotic bacterium. It belongs to the order Rickettsiales and is known to infect numerous species of arthropods and nematodes. The bacterium is known to manipulate the reproduction of its insect hosts *via* cytoplasmic incompatibility, parthenogenesis, male-killing, or feminization [8, 18, 30, 31]. *Wolbachia* species are classified into 11 (A–K) supergroups [22] based on multilocus sequence typing (MLST) and *Wolbachia* surface protein (WSP) typing [2]. The standard MLST system for *Wolbachia* determines a sequence type (ST) based on the combination of alleles for five conserved housekeeping genes (*ftsZ*, *gatB*, *coxA*, *hcpA*, and *fbpA*). The WSP typing system determines STs using four hypervariable regions (HVRs) of WSP: HVR1 (aa 52 to 84), HVR2 (85 to 134), HVR3 (135 to 185), and HVR4 (186 to 222) [1].

A host can be infected with a single strain of *Wolbachia*, or multiple strains, as has been observed in a wide range of insects [13, 20, 21]. Classical MLST can be used for the universal characterization of *Wolbachia*, but its application is currently limited to hosts infected with a single strain of *Wolbachia* [2]. To assess the diversity of *Wolbachia* in insects infected with multiple strains, sequencing of the *wsp* gene is generally conducted [32]. Amplification of ST-specific regions is also used, but this method has limited application because the design of specific primers is only possible after a strain's sequence is known. Improved methods for determining STs in insects infected with multiple strains would contribute to understanding *Wolbachia* diversity.

Butterflies are important plant pollinators; therefore, they are used as model systems in a variety of research fields [4]. *Wolbachia* has been detected in a wide range of butterfly species [24], with 24 *Wolbachia* STs recorded in butterflies (http://pubmlst.org/*Wolbachia*/) [25]. There are 268 known butterfly species in Korea, and five of these are yellow butterfly species (*Gonepteryx rhamni, Gonepteryx mahaguru, Colias erate, Eurema hecabe* and *Eurema laeta*) [26]. The yellow butterfly *E. hecabe* is widely distributed in East Asia and is known to be infected with *Wolbachia* [8, 15–17]. Japanese populations of *E. hecabe* are infected with two distinct strains of *Wolbachia*, *w*Hec1 and *w*Hec2 [17]. *Tongeia fischeri* is a small butterfly found in Eastern Europe and Northeastern Asia [12]; *T. fischeri* butterflies in west Siberia, Russia are all infected with a single *Wolbachia* ST300 and personal communication [11]. However, *Wolbachia* infection in butterflies has not been widely investigated in Korea. Considering the ecological role and abundance of butterflies, the infection status of *Wolbachia* in butterflies should be investigated. The accurate sequence typing result would provide the genetic framework for tracing the movement of *Wolbachia* within insect communities, the regional distribution of butterflies, and phenotypic effects on butterflies [2, 7].

Wolbachia strains are impossible to isolate from the host because they are obligate symbionts. Thus, the MLST for *Wolbachia* has depended on metagenomic cloning and sequencing methods. The public MLST database for *Wolbachia* is also based on cloning methods with *Wolbachia* strains found in diverse singly infected host species. Because the metagenome-based sequencing is the standard protocol for *Wolbachia* MLST, we thought next-generation sequencing (NGS)-based MLST would have merit as a massive screening technique for this particular taxon. Thus, we developed and applied pyrosequencing-based MLST and applied the method to three butterfly species (*E. hecabe, E. laeta*, and *T. fischeri*). Our results suggest that pyrosequencing-based MLST could be used for the large-scale screening of multiple *Wolbachia* STs.

Materials and Methods

Sample Collection and DNA Extraction

Three butterfly species (*E. hecabe, E. laeta,* and *T. fischeri*) were collected from Goheung-gun, Jeonnam, and Han River (Seoul, Korea), between October 2013 and September 2014. One adult butterfly for each species was collected and transported to our laboratory. The body of each butterfly was homogenized, and DNA was extracted from 100 µl of the homogenized sample using a QIAamp DNA Minikit (Qiagen, Germany).

DNA Amplification and Sanger Sequencing

Polymerase chain reaction (PCR) amplification of *Wolbachia wsp* and conserved housekeeping genes (*gatB*, *coxA*, *fbpA, ftsZ*, and *hcpA*) was performed using specific primers and protocols described by Baldo *et al*. [2]. GoTaq Colorless Master Mix (Promega, USA) and PCR Thermal Cycler Dice TP600 (Takara, Japan) were used for the PCR. PCR products were purified with the QIAquick PCR Purification Kit (Qiagen), and were ligated into the pTOP TA V2 vector using the TOPcloner TA core kit (Enzynomics, Korea). Plasmids were transformed into competent *Escherichia coli* cells; colonies containing recombinant plasmids (10 white colonies) were picked and grown overnight. Recombinant plasmid DNA was extracted from overnight *E. coli* cultures using a LaboPass Plasmid Miniprep Kit (Cosmo Genetech, Korea). The sizes of inserts were determined by restriction endonuclease digestion of the plasmid DNA using *Eco*RI (Takara, Japan). Plasmids with inserts of the correct size were subjected to Sanger sequencing by Macrogen Inc. (Korea).

Pyrosequencing

For pyrosequencing, *Wolbachia wsp* and the five housekeeping genes were amplified using barcoded fusion primers. The forward fusion primer was composed of 454-adaptor, key sequence, linker, and target gene-specific sequences (5′-CCTATCCCCTGTGTG CCTTGGCAGTC-TCAG-AC-target sequence-3′). The reverse fusion primer was composed of 454-adaptor, key sequence, barcode, linker, and target gene-specific sequences (5′-CCATCTCATCCCTGCGTGTCTCCGAC-TCAG-barcode-AC-target sequence-3′), with a unique barcode designed for each subject. The full list of oligonucleotide primer sequences used in the current study is summarized in Table S1. Each PCR was performed in a 50 µl volume and contained 1.25 µl of each primer, 25 µl of Taq DNA polymerase buffer, 21.5 µl of distilled water, and 1 µl of template DNA. Reactions were incubated at 94°C for 5 min, followed by 37 amplification cycles (94°C for 30 sec, the optimal annealing temperature for 45 sec, 72°C for 90 sec), and then a final elongation step at 70°C for 10 min, with the temperature then held at 4°C. Optimal PCR annealing temperatures were 53°C for *hcpA*, 54°C for *gatB* and *ftsZ*, and 55°C for *coxA*, *fbpA,* and *wsp*. To reduce amplification bias, three independent PCRs were performed and pooled. Amplicons were confirmed by agarose gel electrophoresis and purified using a QIAquick Gel Extraction kit (Qiagen). The DNA concentration was measured using a Picodrop (Bioneer), and equal quantities were mixed to create amplicon pools. Pyrosequencing was performed at ChunLab Inc. (Korea) using a Roche 454 GS Junior platform.

Sequence Analysis and Identification

All processing of pyrosequencing data was performed using Mothur ver. 1.29.2 [27]. The sequencing reads from the different samples were separated by their unique barcodes, and the sequences of the barcode, linker, and PCR primers were trimmed from both ends of the sequencing reads. For *gatB, coxA,* and *fbpA*

Table 1. Summary of generated pyrosequencing reads.

Characteristics	*Eurema hecabe*						*Eurema laeta*						*Tongeia fischeri*					
	gatB	*coxA*	*hcpA*	*ftsZ*	*fbpA*	*wsp*	*gatB*	*coxA*	*hcpA*	*ftsZ*	*fbpA*	*wsp*	*gatB*	*coxA*	*hcpA*	*ftsZ*	*fbpA*	*wsp*
Raw reads	1,108	59	504	1,564	938	2,299	2,361	3,519	740	1,857	567	2,631	2,109	3,329	643	860	205	845
Pre-filtered reads	949	36	376	1,232	793	994	1,994	2,279	545	1,406	482	1,380	2,104	3,245	621	844	200	480
No. of clusters[a]	209	12	210	306	276	342	344	518	240	327	192	440	33	132	58	36	19	52
Identified reads[b]	703	30	133	808	436	483	1,558	1,709	278	955	268	760	2,076	3,065	565	810	186	426

[a]A cluster cut-off value of a 2 nt difference.

[b]For *gatB*, *coxA*, *hcpA*, *ftsZ*, and *fbpA*, the top five largest clusters were chosen as the representative sequences and subjected to identification. For *wsp*, reads with an exact match to any hypervariable region (HVR1, 2, 3, or 4) were chosen for identification.

sequences, reads that were <350 bp were filtered out, and for *hcpA*, *ftsZ*, and *wsp*, reads that were <400 bp were filtered out using the screen following command: minlength = 350, maxhomop = 8, maxambig = 0. The pre-filtered reads were aligned using reference sequences obtained from GenBank and the *Wolbachia* MLST database (http://pubmlst.org/*Wolbachia*/). During clustering, sequencing errors were ignored by allowing for mismatch errors of up to 2 bp, given the substitution sequence error rate of pyrosequencing (0.5%) [6, 10, 14], using the pre.cluster command diffs = 2. Chimeras were identified and removed using UCHIME [5] (chimera.uchime). For the five housekeeping genes used in MLST, the top five largest clusters were chosen as representative sequences and subjected to identification. For the *wsp* gene, reads with an exact match to any HVR (HVR1, 2, 3, or 4) were chosen for identification. The allele numbers of representative sequences were identified according to the *Wolbachia* MLST database.

Phylogenetic Analysis

To align sequences and construct phylogenetic trees, MEGA5.2 [29] and MrBayes 3.2.3 [9] were used. The *wsp* gene sequences generated in this study were aligned with homologous sequences deposited in the *Wolbachia* MLST database. The consensus sequence comprised 495 bp for *wsp* gene fragments. A maximum-likelihood tree was inferred using the Jones-Yalor-Thorton substitution model and evaluated by 1,000 bootstrap replicates. A Bayesian tree was generated through Markov Chain Monte Carlo methods. The standard deviation was below 0.003 after 5,000,000 generations, and a 10% "burn in" of total samples was conducted.

Sequence Data Availability Statement

Raw data files of pyrosequencing are available in the NCBI Sequence Read Archive (Accession No. SRP 058686). Sequences obtained from Sanger sequencing were deposited in the GenBank database with the accession numbers KP763428–KP763455 and KR006333–KR006338.

Results

Sanger Sequencing-Based MLST

Wolbachia infection in the three butterfly species was confirmed by PCR amplification of the *wsp* gene (514–566 bp). Cloning and Sanger sequencing of the amplicons revealed that *T. fischeri* was infected with a single *Wolbachia* ST. In contrast, the yellow butterflies were infected with two *Wolbachia* STs, with two different *wsp* sequences seen among 38 clones. The five housekeeping genes were amplified from the same template DNAs, with at least 10 clones for each gene. The resultant *gatB*, *hcpA*, and *fbpA* sequences revealed that both yellow butterfly species were infected with at least two different *Wolbachia* STs by showing two STs for each gene. However, sequence variability was not consistently observed; for *coxA*, only one clone was successfully sequenced, thus resulting in a single allele type. For *ftsZ*, all clone sequences (10 clones for *E. hecabe*; 11 clones for *E. laeta*) contained a single allele (*ftsZ 36*).

Pyrosequencing Reads

The total number of pyrosequencing reads for each sample ranged from 59 to 3,519 (Table 1). After quality filtering, alignment, and the removal of chimeric sequences, the numbers of sequencing reads were reduced by 23%. The pre-filtered reads were grouped into 12–518 clusters, with a cluster cut-off value of a 2 nt difference. The size of the resulting clusters varied from 2–2,063 reads, with the top five largest clusters containing 59% of the pre-filtered reads. In most cases (genes and organisms) analyzed in this study, the increment curve of clusters reached a plateau before the fifth cluster (Fig. 1). In all cases, the allele identification was complete by the top five clusters, and no more allele was found even though the sixth and the rest of the clusters were identified fully. Thus, we considered the other clusters as error sequence-based reads, and the reads of the top five largest clusters were chosen as representative sequences and subjected to further identification.

Wolbachia STs Determined by MLST Allelic Profiles

Allele sequences obtained by Sanger sequencing and

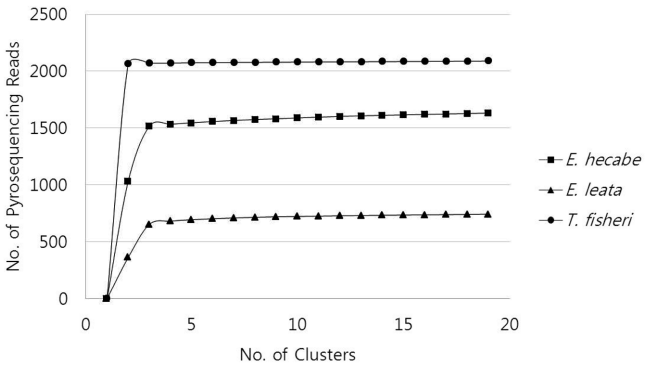

Fig. 1. Plot of the number of clusters as a function of the number of sequencing reads of *gatB*.
The curves grow rapidly at first, but then plateau as only the rarest clusters with a low number of sequencing reads remain.

pyrosequencing were identified using the *Wolbachia* MLST database. A summary of the allele IDs for the five housekeeping genes is provided in Table 2. A single ST in *T. fischeri* was evident from the Sanger sequencing and pyrosequencing results. According to our pyrosequencing results, the top five clusters of the tested genes were identified as an identical allele set and determined as ST41. In contrast, the two yellow butterfly species exhibited two alleles for *gatB* (39, 38), *coxA* (14, 38), *fbpA* (4, 42), *ftsZ* (36, 35), and *hcpA* (40, 29) (Table 2). Theoretically, this gene profile could result in 2^5 combinations, and consequently 32 different STs. However, according to the *Wolbachia* MLST database, only ST40 and ST41 were reported among the 32 theoretical allele profiles. *Eurema hecabe* and *E. laeta* were infected by the same *Wolbachia* population, and were equally infected with the mixed population of *Wolbachia* ST41 and ST40.

The allelic profiles generated from the pyrosequencing method revealed differences in their composition or proportion when compared with the Sanger sequencing ones. *CoxA* allele 14 and *ftsZ* allele 35 were not detected by Sanger sequencing, but were evident through pyrosequencing. The proportion of alleles in a population also differed, depending on the sequencing method employed. As an example, the proportion of *gatB* in *E. hecabe* was 53% according to pyrosequencing, but only 20% according to Sanger sequencing.

Table 2. Multilocus sequence typing allelic profiles of *Wolbachia* strains from three butterfly species.

Species	Allele					ST
	gatB	*coxA*	*hcpA*	*ftsZ*	*fbpA*	
Eurema hecabe	39	14[a]	40	36	4	41
	38	38	29	35[a]	42	40
Eurema laeta	39	14[a]	40	36[a]	4	41
	38	38	29	35[a]	42	40
Tongeia fischeri	39	14	40	36	4	41

[a]Alleles detected by pyrosequencing only.

Wolbachia Phylogeny Based on *wsp*

Based on our pyrosequencing results, 85% of clusters matched with at least one HVR. The proportion of reads containing HVR1 (0.7%) or HVR2 (8.9%) was lower than that for reads containing HVR3 (59.0%) or HVR4 (73.3%). Since sequencing of the *wsp* gene was started from HRV4 to HRV1, these differences were likely due to limited read lengths and higher error rates at the 3 end of the sequence. The length of *wsp* required for haplotype identification is 483 bp; however, the majority of sequences generated by pyrosequencing were shorter than this. For the *wsp* gene, reads that matched any HVR were chosen for identification. The *wsp* allele of *T. fischeri* was identified as *wsp*-10, which corresponded with the Sanger sequencing results. In the yellow butterfly species, both sequencing methods revealed two *wsp* alleles, *wsp*-64 and *wsp*-10, with unique HVR peptide profiles (10-8-10-8 and 35-35-38-44, respectively) (Table 3). The dominant *wsp* allele was *wsp*-10 with complete HVR peptide profiles; *wsp*-64 contained a unique partial HVR peptide profile.

A phylogenetic tree based on *wsp* sequences revealed that the *Wolbachia* strains detected in this study belonged to *Wolbachia* supergroup B (Fig. 2). The detected *wsp* alleles were positioned at different phylogenetic branches within supergroup B, and demonstrated a remarkable genetic distance. The *wsp*-64 sequences from the two yellow butterflies clustered with those for the orange butterfly (*Ariadne merione*), and the small brown plant hopper (*Laodephax striatellus*), and were distinctly separate from *wsp*-10. The two *wsp*-10 sequences from yellow butterflies clustered with that for Indian *E. hecabe*. The *wsp*-10

Table 3. *Wsp* allelic profiles of *Wolbachia* strains from yellow butterflies.

Species	Allele	HVR1	HVR2	HVR3	HVR4
Eurema hecabe	10	10	8	10	8
	64	35	35	38	44
Eurema laeta	10	10	8	10	8
	64	35	35	38	44
Tongeia fischeri	10	10	8	10	8

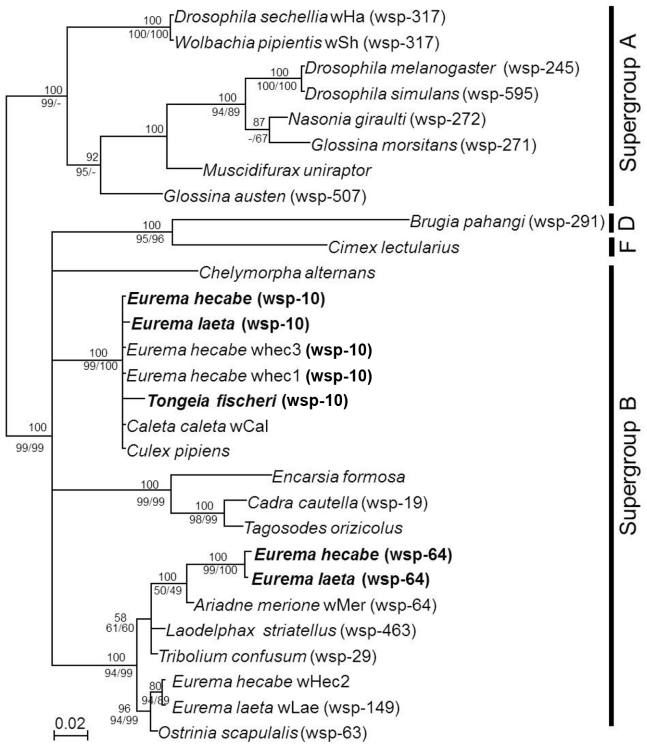

Fig. 2. Bayesian phylogenetic tree based on *wsp* loci (495 bp). Numbers in parentheses indicate *wsp* alleles. The numbers above the nodes are posterior probabilities from 5,000,000 generations. The numbers below the nodes are percentages of bootstrap support, from 1,000 resampled datasets of maximum-likelihood and/or neighbor-joining tree-inferring methods. The scale bar indicates a 0.02 nt change per position.

sequence of *T. fischeri* differed slightly from that in the yellow butterflies and belonged to a sister group.

Discussion

In the current study, we applied NGS technology to conduct MLST and investigate cases of *Wolbachia* infection involving multiple strains. Pyrosequencing read lengths fell within the size range of genes used in MLST (450–500 bp) [19]; therefore, a large number of trimmed pyrosequencing reads could be successfully assigned to allele number. According to the results of this study, pyrosequencing-based MLST is superior to Sanger-based methods with respect to sensitivity of detection. As an example, minor allele sequences, such as *coxA* allele 14 and *ftsZ* allele 35, were only detected by pyrosequencing in this study. However, there were several limitations with the pyrosequencing-based MLST. First, the time and cost effect of NGS-MLST is inferior to Sanger-based MLST. Second, the short sequence length of pyrosequencing reads hindered the complete identification of *wsp* hypervariable regions. Third, current NGS technologies, such as pyrosequencing, exhibit high error rates (0.49–2.8%) [7] compared with Sanger sequencing (0.0001–1%) [7]. The high error rates could result in over-calculation of STs, especially in cases where multiple strains are present. However, NGS technologies are rapidly evolving, and error rates are likely to decrease while read lengths increase. Improvements in experimental methods and bioinformatics analyses regarding next-generation sequencing-based MLST will allow us to conduct large-scale *Wolbachia* MLST in the near future.

The *coxA* of *E. hecabe* was hard to amplify. We tried various PCR conditions and various cloning vectors, but could get only one Sanger sequence. In line with Sanger sequencing results, only 59 reads of the *coxA* gene were obtained from pyrosequencing of *E. hecabe*. The *coxA* gene of *E. hecabe* probably had mutation in the primer region and produced a low yield of DNA amplification. This is a usual case that a part of MLST genes is not amplified or sequenced through the traditional *Wolbachia* MLST scheme [1, 3, 23]. Although the number of pyrosequencing reads were as small as 59 for *E. hecabe*, it was enough to determine STs in *E. hecabe*.

Wolbachia infection in *E. hecabe* and *E. laeta* was previously reported in India and Japan [24, 28]. Japanese *E. hecabe* and *E. laeta* are infected with both ST40 and ST41 [28], the same STs that we identified in both Korean yellow butterflies. In contrast, Indian *E. hecabe* (ST41) and *E. laeta* (ST149) are infected with a single *Wolbachia* ST [24]. This finding implies that Northeastern Asia might have homogeneous yellow butterfly populations. Further investigation of larger geographical regions is required to confirm this. Although two alleles were identified by MLST (ST41 and ST40) and *wsp* sequence typing (*wsp*-10 and *wsp*-64), it is difficult to conclude that only two strains of *Wolbachia* were present in the yellow butterflies. Certain *Wolbachia* strains are known to share the same *wsp* allele despite harboring different STs. For instance, *Wolbachia* strains sharing *wsp*-10 but harboring two distinct STs (ST41 and ST157) were reported in four *E. hecabe* butterflies from India [24]. The two Korean yellow butterflies we investigated were infected with at least two strains of *Wolbachia*.

We found that the three butterfly species we investigated were infected by a multitude of *Wolbachia* strains belonging to supergroup B. This is the first report of *Wolbachia* infection in butterflies from Korea, and would give basic knowledge for researchers on butterflies and insects.

However, the number of organisms and species we surveyed were limited. Further investigation of *Wolbachia* infection in a larger number of butterfly species across many geographical areas will be necessary to obtain an accurate indication of the distribution and diversity of *Wolbachia* species among Korea insect populations.

Acknowledgments

This work was supported by the Basic Science Research Programs through the National Research Foundation of Korea (NRF), funded by the Ministry of Science, ICT, and Future Planning (NRF-2013R1A1A3010041).

References

1. Atyame CM, Delsuc F, Pasteur N, Weill M, Duron O. 2011. Diversification of *Wolbachia* endosymbiont in the *Culex pipiens* mosquito. *Mol. Biol. Evol*. **28:** 2761-2772.
2. Baldo L, Dunning Hotopp JC, Jolley KA, Bordenstein SR, Biber SA, Choudhury RR, *et al*. 2006. Multilocus sequence typing system for the endosymbiont *Wolbachia pipientis*. *Appl. Environ. Microbiol*. **72:** 7098-7110.
3. Bing XL, Xia WQ, Gui JD, Yan GH, Wang XW, Liu SS. 2014. Diversity and evolution of the *Wolbachia* endosymbionts of *Bemisia* (Hemiptera: Aleyrodidae) whiteflies. *Ecol. Evol*. **4:** 2714-2737.
4. carol lb, ward bw, ehrlich pr. 2003. *butterlies: ecology and Evolution Taking Flight*. University of Chicago Press, IL.
5. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. 2011. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27:** 2194-2200.
6. Gilles A, Meglecz E, Pech N, Ferreira S, Malausa T, Martin JF. 2011. Accuracy and quality assessment of 454 GS-FLX titanium pyrosequencing. *BMC Genomics* **12:** 245.
7. Goodacre SL, Martin OY, Thomas CF, Hewitt GM. 2006. *Wolbachia* and other endosymbiont infections in spiders. *Mol. Ecol*. **15:** 517-527.
8. Hiroki M, Kato Y, Kamito T, Miura K. 2002. Feminization of genetic males by a symbiotic bacterium in a butterfly, *Eurema hecabe* (Lepidoptera: Pieridae). *Naturwissenschaften* **89:** 167-170.
9. Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17:** 754-755.
10. Huse SM, Huber JA, Morrison HG, Sogin ML, Welch DM. 2007. Accuracy and quality of massively parallel DNA pyrosequencing. *Genome Biol*. **8:** R143.
11. Ilinsky Y. 2012. Full information on strain Tfis_B. Available at http://pubmlst.org/perl/bigsdb/bigsdb.pl?db=pubmlst wolbachia isolates&page=info&id=474.
12. Jeratthitikul E, Hara T, Yago M, Itoh T, Wang M, Usami S, Hikida T. 2013. Phylogeography of Fischer's blue, *Tongeia fischeri*, in Japan: evidence for introgressive hybridization. *Mol. Phylogenet. Evol*. **66:** 316-326.
13. Kikuchi Y, Fukatsu T. 2003. Diversity of *Wolbachia* endosymbionts in heteropteran bugs. *Appl. Environ. Microbiol*. **69:** 6082-6090.
14. Kunin V, Engelbrektson A, Ochman H, Hugenholtz P. 2010. Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. *Environ. Microbiol.* **12:** 118-123.
15. Narita S, Kageyama D, Nomura M, Fukatsu T. 2007. Unexpected mechanism of symbiont-induced reversal of insect sex: feminizing *Wolbachia* continuously acts on the butterfly *Eurema hecabe* during larval development. *Appl. Environ. Microbiol*. **73:** 4332-4341.
16. Narita S, Nomura M, Kageyama D. 2007. Naturally occurring single and double infection with *Wolbachia* strains in the butterfly *Eurema hecabe*: transmission efficiencies and population density dynamics of each *Wolbachia* strain. *FEMS Microbiol. Ecol*. **61:** 235-245.
17. Narita S, Nomura M, Kato Y, Yata O, Kageyama D. 2007. Molecular phylogeography of two sibling species of *Eurema* butterflies. *Genetica* **131:** 241-253.
18. O'Neill SL, Giordano R, Colbert AM, Karr TL, Robertson HM. 1992. 16S rRNA phylogenetic analysis of the bacterial endosymbionts associated with cytoplasmic incompatibility in insects. *Proc. Natl. Acad. Sci. USA* **89:** 2699-2702.
19. Perez-Losada M, Cabezas P, Castro-Nallar E, Crandall KA. 2013. Pathogen typing in the genomics era: MLST and the future of molecular epidemiology. *Infect. Genet. Evol*. **16:** 38-53.
20. Raychoudhury R, Baldo L, Oliveira DC, Werren JH. 2009. Modes of acquisition of *Wolbachia*: horizontal transfer, hybrid introgression, and codivergence in the *Nasonia* species complex. *Evolution* **63:** 165-183.
21. Riegler M, Stauffer C. 2002. *Wolbachia* infections and superinfections in cytoplasmically incompatible populations of the European cherry fruit fly *Rhagoletis cerasi* (Diptera, Tephritidae). *Mol. Ecol*. **11:** 2425-2434.
22. Ros VI, Fleming VM, Feil EJ, Breeuwer JA. 2009. How diverse is the genus *Wolbachia*? Multiple-gene sequencing reveals a putatively new *Wolbachia* supergroup recovered from spider mites (Acari: Tetranychidae). *Appl. Environ. Microbiol*. **75:** 1036-1043.
23. Salunke BK, Salunkhe RC, Dhotre DP, Khandagale AB, Walujkar SA, Kirwale GS, et al. 2010. Diversity of *Wolbachia* in *Odontotermes* spp. (Termitidae) and *Coptotermes heimi* (Rhinotermitidae) using the multigene approach. *FEMS Microbiol. Lett*. **307:** 55-64.
24. Salunke BK, Salunkhe RC, Dhotre DP, Walujkar SA, Khandagale AB, Chaudhari R, *et al*. 2012. Determination of *Wolbachia* diversity in butterflies from Western Ghats, India, by a multigene approach. *Appl. Environ. Microbiol*. **78:** 4458-4467.

25. Salunkhe RC, Narkhede KP, Shouche YS. 2014. Distribution and evolutionary impact of *Wolbachia* on butterfly hosts. *Ind. J. Microbiol*. **54:** 249-254.
26. Lee SH, Kwon SG, Nam GP, Son JD, Lee JG, Park YK, *et al*. 2012. Studies on ecological environments and indoor-rearing conditions of the common grass yellow butterfly, *Eurema hecabe*. *J. Seric. Entomol. Sci*. **50:** 133-139.
27. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, *et al*. 2009. Introducing Mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol*. **75:** 7537-7541.
28. Tagami Y, Miura K. 2004. Distribution and prevalence of *Wolbachia* in Japanese populations of Lepidoptera. *Insect Mol. Biol*. **13:** 359-364.
29. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol*. **28:** 2731-2739.
30. Weeks AR, Breeuwer JA. 2001. *Wolbachia*-induced parthenogenesis in a genus of phytophagous mites. *Proc. Biol. Sci*. **268:** 2245-2251.
31. Werren JH, Baldo L, Clark ME. 2008. *Wolbachia*: master manipulators of invertebrate biology. *Nat. Rev. Microbiol*. **6:** 741-751.
32. Zhou W, Rousset F, O'Neil S. 1998. Phylogeny and PCR-based classification of *Wolbachia* strains using *wsp* gene sequences. *Proc. Biol. Sci*. **265:** 509-515.