We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

4,300
Open access books available

117,000
International authors and editors

130M
Downloads

154
Countries delivered to

TOP 1%
Our authors are among the most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com
Chapter 3

Molecular Phylogeny and Taxonomy of Lepidoptera with Special Reference to Influence of Wolbachia Infection in the Genus Polytremis

Weibin Jiang

Abstract

This chapter provides a case of genus Polytremis Mabille, 1904 (Lepidoptera: Hesperiidae), to explain the molecular phylogeny and taxonomy of Lepidoptera and the influence of Wolbachia infection. Earlier studies of Lepidoptera were focused mainly on the morphological classification, population distribution, and identification of new species. As the supplementary to morphological research, analysis of DNA has been widely used in the phylogenetic studies of Lepidoptera. The study provides a conservative estimate that the Wolbachia infection rate in Polytremis nascens Leech (1893) is 31%, and no significant difference in the prevalence is found between the sexes. The Wolbachia infection mainly prevails in populations of P. nascens in southern China, which influence the diversity of mtDNA in P. nascens by a Wolbachia-induced sweep. The Wolbachia infection rate in Polytremis fukia Evans (1940) is 47% and shows a weak association existed between mitochondrial DNA haplotypes and wFuk1 infection status.

Keywords: Lepidoptera, microsatellite, mitochondrial genome, molecular phylogeny, taxonomy, Wolbachia

1. Introduction to molecular phylogeny and taxonomy of Lepidoptera

Butterflies and moths (Lepidoptera) have long served as a model system for ecological and evolutionary studies on the basis of the high degree of diversity and complexity, which constitute one of the most diverse insect orders with more than 157,000 described species. Earlier studies of Lepidoptera were focused mainly on the morphological classification, population distribution, and identification of new species. As the supplementary to morphological research, analysis of
DNA has been widely used in the phylogenetic and taxonomic studies of Lepidoptera. The case of genus *Polytremis* will be discussed as follows.

2. The phylogeny of the butterfly genus *Polytremis*

The family Hesperiidae includes more than 4000 species, commonly known as “skippers,” of which Hesperini is the largest subfamily. *Polytremis* Mabille (1904) is a genus of subfamily Hesperini, which has 18 members and is restricted geographically to the continental part of the southeastern Palaearctic and northern oriental regions. They have a thick body and relatively small wings. These wings are commonly dark brown or yellowish brown [1]. The main external features are characterized by the unspined mesotibia, the absence of a cell spot on the underside of each hind wing, and the serial, linear, and semi-hyaline spots. Male genitalia are distinguished by the elongated harpe, swollen tegumen, and bifid uncus [2].

2.1. The method of constructing and analyzing phylogenetic tree

The specimens from 15 of the estimated 18 species in the genus *Polytremis* were collected, from different localities. The DNA was isolated from leg tissue. The mitochondrial cytochrome c oxidase I (COI) gene, recommended as the universal and standard barcoding marker for species identification [3], was amplified approximately 490 bp. For nuclear DNA, three expansion segments of 18S rDNA and 28S rDNA were chosen, the slowly evolving genes used normally in higher classification studies [4]. The haplotype sequence matrix was used for all subsequent phylogenetic analyses. MEGA v4.0 was used to calculate the intra- and interspecific genetic divergences based on the K2P model [5]. Phylogenetic trees were constructed by the maximum-likelihood (ML) methods using PAUP 4.0b10 [6]. Relationships among the mitochondrial COI and concatenated sequence (mitochondrial COI + nuclear rDNA) haplotypes were represented as a haplotype network obtained by the software DnaSP4.90 [7] and Network4.5 using the median-joining method [8].

2.2. Genetic divergence, phylogenetics, and network of genus *Polytremis*

Figure 1A reveals five main clades and shows the ML tree based on the data set of COI. Clade I contained eight species: *Polytremis suprema* and *Polytremis gigantean*, *Polytremis caerulescens*, *Polytremis kiraizana* Sonan, 1938, and *Polytremis matsuii* Sugiyama, 1999 are first clustered with a strong support value, followed by clustering of *Polytremis gotama communis* and *Polytremis nascens* Leech, 1893. Then *Polytremis jigongi* Zhu, 2012 is revealed [2]. Two haplotypes of *P. jigongi* are clearly separated from other species form and strongly supported lineages. Clade II contains only *Polytremis theca*. It was reported to include three subspecies. They show a greater intraspecific genetic distance than some interspecific genetic distances in the genus *Polytremis*. Furthermore, the sister group relationship between *Polytremis zina* Evans, 1932 and *Polytremis pellucida* Murray, 1875 in Clade III is confirmed. Clade IV contains only one species, *Polytremis mencia* Moore, 1877. Clade V contains three species: *Polytremis discreta* Elwes & Edwards, 1897, *Polytremis lubricans* Herrich-Schäffer, 1869, and *Polytremis eltola* Hewitson, 1869, which are distributed sympatrically in the oriental region throughout Malaya and India.
The level of DNA sequence divergence reflected the taxonomic hierarchy of the original species. The lowest intraspecific COI genetic distance was observed between *P. suprema* and *P. gigantea* (K2P distance 1.7%). Except for *P. theca*, the intraspecific distances were shorter. The COI data confirmed the sister group relationship between *P. suprema* and *P. gigantea*, which form a monophyletic group together with *P. matsui*, *P. caerulescens*, and *P. kiraizana*. All of their interspecific distances were smaller than 3% (K2P distance). We can infer from morphological traits. These five species shared many morphological traits including ear-like uncus with a pair of processes, the absence of a cornuti, and thin coecum penis. *P. theca* was the only species for which the intraspecific genetic distance was greater than some interspecific
genetic distances based on COI in the genus Polytremis. However, the distance was much less than the average interspecific genetic distances of the genus Polytremis (K2P distance 7.9%) [9]. *P. theca* was widely distributed in the south of the Qinling Mountains in China, except in the Hainan Province and the southern tropical regions of Yunnan Province [10]. Three subspecies of *P. theca* were reported based on morphological features of the wings. Our specimens included two of them, namely, *Polytremis theca theca* and *Polytremis theca fukia*. The COI tree revealed these two distinct haplotype lineages without intermediates (K2P distance 4.2%). Additionally, the subspecies were separated by nuclear rDNA sequence, and the K2P distance was 0.3%, suggesting the possible existence of a sibling species paired with allopatric distribution.

The average interspecific rDNA genetic distance (K2P distance 1.0%) was far less than that of COI (K2P distance 7.9%). Except for *P. caerulescens* and *P. gotama*, other species in *Polytremis* could be distinguished with rDNA. These two species could be separated in the COI, and K2P distance was 1.9% but showed no variation in the rDNA. Based on mitochondrial and nuclear markers, the differences observed between results may contribute to recent separation, introgressive hybridization, or incomplete lineage sorting [11]. Because *Polytremis* is a fairly old genus and the splits of COI of the two lineages are also quite old, it seems that incomplete lineage sorting may not be an appealing explanation for the discordance. Additionally, *P. gotama* has been described as an independent species by COI data and morphological features [12]. A specimen of *P. gotama* and three specimens of *P. caerulescens* (from two populations) revealed two distinct haplotype lines without intermediates (K2P distances 4.9%). The observation indicates that they were two species based on the molecular and morphological level. Nevertheless, more specimens of the two species from different population need to be collected and analyzed in the future to see if this pattern is recovered consistently and further confirm the relationship of them. Instead, some arguments favor the assumption of recent separation [9]. As far as their morphological characteristics were concerned, *P. caerulescens* was considered to be closely related to *P. gotama* on the basis of the structural similarity of the cell spot on the upper side of the hind wing and the male genitalia, which were not found in the other *Polytremis* species [13]. Additionally, only these two species were observed and captured at altitudes higher than 2000 m. *P. gotama* was a little smaller than *P. caerulescens*. They both varied in other morphological traits which clearly support the existence of two closely related but distinct species, including male stigma on the upper side of the forewing and the ground color of the wings. rDNA showed no sequence variation, whereas K2P distances of the COI fragments reached 4.9%, suggesting a possible recent separation of *P. gotama* and *P. caerulescens*.

In Lepidoptera, thresholds have been proposed as 3% for COI [3]. The intraspecific and interspecific genetic divergences did not fall into separate intervals, and an obvious “barcode gap” did not occur in COI in our study of *Polytremis* [9]. It was entailed by two factors. Firstly, all intraspecific distances were less than 3% except for *P. theca*. However, we inferred *P. t. theca* and *P. t. fukia* could be a sibling species pair according the morphological and molecular data in the study. Secondly, the interspecific genetic distances among five sister species (*P. gigantean*, *P. suprema*, *P. caerulescens*, *P. kiraizana*, and *P. matsuii*) were less than 3%, which caused the interspecific and intraspecific genetic divergences to overlap from 1.7 to 3%. Regardless, for COI, the overlaps between intraspecific and interspecific variations would not affect
identification in a thoroughly sampled evaluation [14]. The *Polytremis* species could be clustered with a well support and distinguished by tree-based methods, suggesting that the COI sequence could be used to correctly identify almost all species in the genus as a DNA barcode (Figure 1A). Additionally, the markers of the nuclear rDNA sequences used in our studies, three expansion segments of 18S and 28S rDNA, have been proposed as a reasonable alternative to mitochondrial COI. These markers could avoid problems of mitochondrial markers such as introgression and pseudogenes and identify or delimit species or species-like units as they are not inherited maternally [9].

2.3. Conclusion: combined morphological and molecular analysis

A total of 20 morphological characters yield a two-cluster solution with hierarchical cluster analysis. The first cluster includes 12 species of *Polytremis*, and the second includes the remaining 3 species and outgroups. All supported clades from the combined data matrix are also appearing in the molecular data matrix. ML tree based on COI constructed in this study showed that individuals belonging to the same species formed a monophyletic cluster. Meanwhile, there was considerable congruence in topology of the interspecies level for both mtDNA COI and concatenated sequences of ML trees indicating certain clades were well differentiated phylogenetically (Figure 1) [9]. The strong support for the monophyly of *Polytremis* was found in the analyses of the concatenated alignments and COI [9]. In genus *Polytremis*, the results obtained by hierarchical cluster analysis showed traditional classification was basically consistent with molecular phylogeny. However, because the morphological characters and character states were commonly homologous in *Polytremis*, the morphological analysis resulted in only limited resolution based on just 20 morphological characters. On contrary, molecular classification provided a lower artificial and more precise taxonomic rank [15]. Thus, the combination of the morphological and molecular matrix was better resolved for understanding of the phylogeny in the genus.

3. Taxonomic status of two sibling species of *Polytremis* (Lepidoptera: Hesperiidae)

The skipper *P. theca* is widely distributed in south China, except Taiwan, Hainan, and the southern tropical regions of Yunnan Province. Three subspecies have been recorded: *P. t. theca* [9] (west Sichuan and south Shaanxi Province), *P. t. fukia* [16] (Zhejiang to west Sichuan Province), and *Polytremis theca macrotheca* Huang, 2002 (Northwest Yunnan Province). In a preliminary study of molecular phylogeny of the genus *Polytremis* Mabille, 1904 using mitochondrial cytochrome c oxidase I (COI), we found the inter-subspecific distance between *P. t. theca* and *P. t. fukia* ranged up to 4.2%, which is higher than some interspecific genetic distances in *Polytremis*. Additionally, *P. theca* is also the only species whose intraspecific distance is more than 3%; thresholds of species identification have been proposed in *Lepidoptera* for COI, in genus *Polytremis* [9]. Thus, we suspected the possible existence of a sibling species paired or the cryptic diversity in the species.
3.1. Genetic divergences and haplotype networks

All 46 samples yielded high quality of DNA. A total of 19 haplotypes were identified in all 46 samples, and the haplotype network was constructed and presented in Figure 2. There was no shared haplotype among the four taxa. Haplotypes of the same taxon differed from each other by no more than five mutation distance. The five mutation distances existed between the haplotype Ptt I and Ptt III of P. t. theca. The potential ancestral haplotype of P. t. fukia, defined by its central position in the network, was designated as Ptf I, which was found in three samples from Tianmushan, one from Jinggangshan, and one from Anjiangping. Ptf II was the most common haplotype in P. t. fukia and shared with 10 samples. Haplotype Ptf XIII was identified in two samples from Maoershan and one sample from Anjiangping. The remaining haplotypes of P. t. fukia occurred in only one individual (Figure 2A).

The data set of nuclear wingless contains 390 nucleotide positions without gaps or stop codons, of which 18 positions are variable and 9 are parsimony informative. In total, 10 haplotypes were found in all samples, in which 2 haplotypes were found in P. t. theca, 4 in P. t. fukia, 3 in P. nascens, and 1 in P. mencia. Haplotypes of the same taxon differed from each other by no more than two mutation distances, in particular, the 2 haplotypes in 13 samples of P. t. theca differed by only one-mutation distance. Five nucleotide substitutions were observed between the potential ancestral haplotypes of P. t. theca (Ptt I) and that of P. t. fukia (Ptf I) (Figure 2B).

Overall, P. t. theca had a lower diversity than P. t. fukia according to the result of analysis of both mitochondrial COI and nuclear wingless. The haplotype diversity (Hd) and nucleotide diversity (π) for P. t. theca and P. t. fukia are given in Table 1. Additionally, they differed from each other by 5.07 ± 0.49% (4.3–5.9% divergence) for the COI sequences and by 1.70 ± 0.27% (1.3–2.1% divergence) for the wingless sequences.

3.2. Population structure and phylogenetic analysis

The analysis of molecular variance (AMOVA) for the COI sequences of P. t. theca and P. t. fukia revealed that 88.53% of the genetic variation was among populations and 11.47% was within populations (Table 2). The average ΦST value is 0.896 (p < 0.01), suggesting significant genetic variation among the populations. Pair-wise estimates of FST (0.885) and gene flow (Nm = 0.065) between P. t. theca and P. t. fukia suggest that the subspecies in this species are highly differentiated [16].

Mitochondrial haplotypes sampled from P. theca form well-supported clades that closely correspond with subspecific boundaries delimited primarily on the basis of wing color and pattern (Figure 3A). The haplotype clades associated with both subspecies are deeply genetically divergent, differing from each other by 5.07 ± 0.49% (4.3–5.9% divergence). This degree of divergence suggests that evolutionary separation of both subspecies occurred about 0.81 highest probability density (HPD = 0.53–1.28) Mya, likely sometime during the Pleistocene based on a molecular clock calibration of 3.54% pair-wise divergence per million years for
Figure 2. Network profile of (A) COI and (B) wingless gene haplotypes based on the nucleotide sequences of *P. t. theca*, *P. t. fukia*, *P. nascens*, and *P. mencia*. 
a homologous mtDNA fragment in other insect species [17]. It is noteworthy for the nuclear wingless sequences that P. t. theca and P. mencia are considered distinct species with a genetic divergence of 0.65 ± 0.15%, while the P. t. fukia is currently considered a subspecies of the P. t. theca despite 1.70 ± 0.27% sequence divergence. The phylogenetic of wingless gene indicates that the P. theca is paraphyly with three species sisters to the clade of P. t. theca (Figure 3B). While 100% of P. t. fukia constitutes one separate clade, the clade consisting of P. t. theca also includes P. pellucid, P. zina, and P. mencia. The clade consisting of P. t. theca is not monophyletic, but complex. This suggests that P. t. theca and P. t. fukia differ from each other, as evident from the COI tree where they form two separate clades (Figure 3A). Concordance between strongly differentiated mtDNA, nuclear haplotype clades, and phenotypic variation supports the hypothesis that both subspecies of P. theca deserve recognition at the species level under the general lineage concept of species.

### 3.3. Demographic inference and estimation of divergence times

Demographic history changes were analyzed for P. t. theca and P. t. fukia populations through neutrality tests and mismatch distribution. The neutrality tests reveal that the mitochondrial COI appear to be not evolving neutrally as Fu’s F values in P. t. fukia group are negative significantly (Table 1). The Tajima’s D and Fu’s F values were nonsignificantly positive in

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>df</th>
<th>Sum of squares</th>
<th>Variance components</th>
<th>Percentage variation</th>
<th>Φ statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among populations</td>
<td>1</td>
<td>169.650</td>
<td>9.98266 Va</td>
<td>88.33</td>
<td>–</td>
</tr>
<tr>
<td>Within populations</td>
<td>11</td>
<td>45.269</td>
<td>1.29341 Vb</td>
<td>11.47</td>
<td>0.896 ( p &lt; 0.01 )</td>
</tr>
<tr>
<td>Total</td>
<td>12</td>
<td>214.919</td>
<td>11.27607</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fixation index</td>
<td></td>
<td>0.8853</td>
<td>0.8853</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Analysis of molecular variance (AMOVA) for the COI sequences of P. t. theca and P. t. fukia.

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>df</th>
<th>Sum of squares</th>
<th>Variance components</th>
<th>Percentage variation</th>
<th>Φ statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among populations</td>
<td>1</td>
<td>169.650</td>
<td>9.98266 Va</td>
<td>88.33</td>
<td>–</td>
</tr>
<tr>
<td>Within populations</td>
<td>11</td>
<td>45.269</td>
<td>1.29341 Vb</td>
<td>11.47</td>
<td>0.896 ( p &lt; 0.01 )</td>
</tr>
<tr>
<td>Total</td>
<td>12</td>
<td>214.919</td>
<td>11.27607</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fixation index</td>
<td></td>
<td>0.8853</td>
<td>0.8853</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Genetic diversity and neutrality tests calculated for P. t. theca and P. t. fukia.

Ns, number of samples; Nh, number of haplotypes; Hd, haplotype diversity; Nv, number of variable sites; \( \pi \), nucleotide diversity; SD, standard deviation; D, Tajima’s D statistic; F, Fu’s F statistic.

Significant difference.

Ns, number of samples; Nh, number of haplotypes; Hd, haplotype diversity; Nv, number of variable sites; \( \pi \), nucleotide diversity; SD, standard deviation; D, Tajima’s D statistic; F, Fu’s F statistic.

Significant difference.

### 3.3. Demographic inference and estimation of divergence times

Demographic history changes were analyzed for P. t. theca and P. t. fukia populations through neutrality tests and mismatch distribution. The neutrality tests reveal that the mitochondrial COI appear to be not evolving neutrally as Fu’s F values in P. t. fukia group are negative significantly (Table 1). The Tajima’s D and Fu’s F values were nonsignificantly positive in
P. t. theca group and all P. theca sample group. The mismatch analysis yielded a unimodal distribution of pair-wise differences for P. t. fukia compared to the multimodal distribution of P. t. theca samples and the pooled samples. According to Rogers and Harpending [18], the observed curves with unimodal represent population expansion and the observed curves with many peaks or resemblance to expected curves mean equilibrium population, which further elucidates the demographic history of P. theca. The results suggest population expansion in P. t. fukia and population equilibrium in P. t. theca. We still confirm the result of the population size change in haplotype network (Figure 2). Statistical parsimony network reflects genealogical relationships of the mtDNA haplotypes, that is, single mutation steps separate adjacent haplotypes in the network and older haplotypes are placed at internal branching points, whereas younger ones occur toward the tip positions [19]. The haplotype network of P. t. fukia displays a star-like pattern (Figure 2). Haplotype I, the second most common and geographically widespread in central-west of China, is in the star’s center, and derivatives are connected to it by short branches [16]. Based on coalescence theory, the star-like topologies for this cluster strongly suggest the effect of a population expansion [20]. Divergence time analysis with an uncorrelated lognormal relaxed clock run in Bayesian MCMC analysis of molecular sequences (BEAST) produced a tree with a topology similar to ML tree (Figure 4).

Figure 3. The maximum-likelihood tree for (A) mitochondrial COI and (B) nuclear wingless haplotypes of Polytremis.
P. t. theca diverged from P. t. fukia around 0.81 (HPD = 0.53–1.28) million years ago (Mya) during the Pleistocene (node a in Figure 4). P. theca diverged from other congeners included in the analysis about 1.36 (HPD = 1.02–1.53) Mya during the Pleistocene eras (node b in Figure 4).

In our study, a higher $F_{ST}$ value indicated a lower level of gene flow ($Nm$) and higher genetic differentiation among populations. The results of two-level AMOVA show that significant genetic variation exists among the examined populations. These results provide a second line of support to a conclusion that the P. t. fukia is a different species [16].

### 3.4. Conclusion

There is a small region of overlap in west Sichuan province in the distribution of P. t. theca and P. t. fukia, but otherwise they are not sympatric. P. t. theca inhabits the higher elevations of west Sichuan and south Shaanxi Province [21]. P. t. fukia occurs in the whole Southern China except Taiwan, Hainan, and south Yunnan Province, from Zhejiang to west Sichuan Province [22]. According to the description on morphological variation between P. t. theca and P. t. fukia [23], we found a different morphological feature existing in female genitalia except for the different colors and spot numbers in some part of wings (Table 3). The lateral edge of lamella antevaginalis of female P. t. fukia is more rounded than that of P. t. theca. We suspected the taxonomic status of the subspecies from their geographic separation and the morphological variation. Our investigations and analyses revealed significant molecular and biogeographical differences between P. t. theca and P. t. fukia. We propose that P. t. fukia should be treated as a distinct species called Polytremis fukia [16] under the phylogenetic species concept. In fact, it has been
recently found that other species previously considered subspecies based on morphology are in fact sibling species that passed unnoticed until the advent of molecular techniques [24]. Results from our study strengthen information about the *Polytremis* species complex and help in developing appropriate integrated pest management strategies for these insect pests.

4. The application of mitochondrial genome and microsatellite DNA in *Polytremis*

With the development of the research, the single molecular fragment cannot meet the research requirements. New molecular markers need to be explored. Recently, the mitochondrial genome has become one of the important molecular markers to explore different categories of *Lepidoptera*. Additionally, the ideal molecular marker, microsatellite DNA, has become very prevalent in molecular ecological studies including inferring population genetic structure, exploring taxonomic status, and studying reproductive ecology.

4.1. Complete mitochondrial DNA genome of *P. nascens* and *Polytremis jigongji*

*Lepidoptera* is the second largest insect order in the world and contains more than 160,000 species. However, the information currently existing on lepidopteran mitogenomes is limited. To date, only 236 complete or nearly complete mitogenome sequences have been determined to belong to six superfamilies. The phylogenetic inference based on the variations at such short gene sequences is not always robust and may lack sufficient resolution compared to the phylogenies based on longer mitogenome sequences [25]. Additionally, mitogenomes are also applied to the studies on comparative and evolutionary genomics [26], molecular evolution [27], phylogeography [28], etc. Thus, further insight into *Lepidoptera* phylogeny and evolution awaits more related species sequences to be determined [29].

The complete mtDNA genome of *P. nascens* was a circular molecule of 15,392 bp in length, including the standard 13 protein-coding genes (PCGs), 2 ribosomal RNA (rRNA) genes, 22
transfer RNA (tRNA) genes, and 1 noncoding region. The overall-based composition was 39.7% A, 40.7% T, 7.7% G, and 11.9% C, with a slight A+T bias of 80.4%. Thirteen PCGs and two rRNA were first identified using an open reading frame (ORF) finder, specifying the invertebrate mitochondrial genetic code. Then, they were calibrated by sequence similarity using published lepidopteran mitogenome sequences. All PCGs use standard ATN (ATT, ATG, or ATA) as the start codon except COX1 that uses CGA. Eight PCGs (ND2, ATP8, ATP6, COIII, ND3, ND4, ND6, and ND1) employ the typical stop codon TAA, while the remaining five PCGs terminate with a single T [22].

The complete mtDNA genome of *P. jigongi* was a circular molecule of 15,353 bp in length, including the standard 13 protein-coding genes (PCGs), 2 ribosomal RNA (rRNA) genes, 22 transfer RNA (tRNA) genes, and 1 noncoding region. Its organization and arrangement are identical to other skippers [30, 31]. The overall-based composition was 39.8% A, 41.1% T, 7.6% G, and 11.5% C, with a slight A+T bias of 80.9%. Thirteen PCGs and two rRNA were identified using an open reading frame (ORF) finder and calibrated by sequence similarity using published lepidopteran mitogenome sequences. All PCGs use standard ATN (ATT, ATG, or ATA) as the start codon except COX1 that uses CGA. Eight PCGs (ND2, ATP8, ATP6, COIII, ND3, ND4, ND6, and ND1) employ the typical stop codon TAA, while the remaining five PCGs terminate with TA or T [22]. Alignment of amino acid sequences of each of individual 13 PCGs was performed through Clustal X [32], and the phylogenetic analysis was carried out using neighbor-joining (NJ) method with MEGA version 5.0 program [33] (Figure 5).

**Figure 5.** Phylogenetic tree of the lepidopterans based on 13 PCG nucleotide sequences of the mitogenome using NJ analysis.
4.2. Isolation and characterization of microsatellite loci in *P. nascens* and *P. fukia*

Microsatellites are highly polymorphic and codominant molecular markers based on simple repeated and frequent sequences common in all-living organisms, which have proven to be a powerful tool available in population genetic and evolutionary studies [34]. The ideal molecular marker has become very prevalent in studies of insects over the last 10 years [35]. Variability of the 12 polymorphic microsatellites was surveyed in 53 individuals of *P. nascens*. We found 53 different multilocus genotypes, and the number of alleles per locus ranged from 3 to 12. Observed (HO) and expected heterozygosity (HE) values ranged from 0.33 to 0.71 and from 0.61 to 0.90. Only one locus (Polynyas 13) \((p < 0.01)\) that deviated from HWE showed significant heterozygote deficit in the populations of Lianglu, Shengtangshan, and Maoershan [33]. Analysis performed with Micro-Checker showed that the deviation was attributed to the homozygote excess with null allele frequency of 0.2041, 0.2914, and 0.3533 in the three populations, respectively [33]. No linkage disequilibrium was detected for any pair of loci \((p > 0.01)\) in any populations following Bonferroni correction [33]. Comparisons of pair-wise F\(_{ST}\) among six populations show the genetic difference based on 12 microsatellite loci (Table 4). The population of Baishanzu showed the largest pair-wise F\(_{ST}\) values, corresponding to relatively geographic isolation from the other five populations [33].

The variability of the 11 polymorphic microsatellites was surveyed in 21 individuals of *P. fukia*. We found 21 different multilocus genotypes, and the number of alleles per locus ranged from 5 to 10. Observed heterozygosity (Ho) values ranged from 0.48 to 0.65 and expected heterozygosity (He) values ranged from 0.69 to 0.86. Polymorphic information content (PIC) ranged from 0.59 to 0.88 per locus, and all markers were highly informative (PIC > 0.5) [33]. All loci were in Hardy-Weinberg equilibrium, consistent with inbreeding and/or Wahlund effects. No linkage disequilibrium was detected in any pair of loci [33]. Micro-Checker software found no evidence of scoring error due to stuttering or large allele dropout. We also tested the selected primers for amplification on eight other *Polytremis* species: *P. caerulescens* Mabille, *P. jigongi* Zhu, *P. theca* Evans, *P. zina* Evans, *P. mencia* Moore, *P. lubricans* Herrich-Schäffer, *P. eltola* Hewitson, and *P. discrete* Elwes & Edwards. Of the 11 markers, we found that cross-amplification was successful for 10 loci in at least one conge-

<table>
<thead>
<tr>
<th></th>
<th>Baishanzu</th>
<th>Hailuogou</th>
<th>Lianglu</th>
<th>Shengtangshan</th>
<th>Maoershan</th>
<th>Kuankuoshui</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baishanzu</td>
<td>0.044</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hailuogou</td>
<td>0.050</td>
<td>0.003</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lianglu</td>
<td>0.029</td>
<td>0.025</td>
<td>0.022</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shengtangshan</td>
<td>0.023</td>
<td>0.021</td>
<td>0.020</td>
<td>0.013</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maoershan</td>
<td>0.026</td>
<td>0.012</td>
<td>0.014</td>
<td>0.022</td>
<td>0.017</td>
<td></td>
</tr>
<tr>
<td>Kuankuoshui</td>
<td>0.014</td>
<td>0.012</td>
<td>0.014</td>
<td>0.022</td>
<td>0.017</td>
<td></td>
</tr>
</tbody>
</table>

Table 4. Comparisons of multilocus pair-wise F\(_{ST}\) values \((p < 0.05)\) among 6 regional populations of *P. nascens* based on 12 microsatellite loci.
neric species [33]. Two primer loci amplified satisfactorily among all congeneric species, indicating that these loci may be useful for population genetics, including phylogeography, species cohesion and delimitation, and barriers to gene flow for other congeneric and related species [33].

5. Wolbachia infection and influence in Polytremsis species

Wolbachia may be the most widespread endosymbiont in terrestrial ecosystems. The maternally inherited endosymbiotic bacteria infect perhaps two-thirds of present-day insect species [36, 37]. The transmission of Wolbachia is primarily vertical and secondarily horizontal [38]. The bacteria manipulate the reproduction of their host to ensure their vertical transmission by cytoplasmic incompatibility, feminization, male killing, and parthenogenesis [39]. Wolbachia can potentially influence mitochondrial variation of their hosts. The linkage disequilibrium is expected to occur between them since they are co-transmitted maternally. The rapidly spread of Wolbachia in the host populations can result in the hitch-hiking effect of mitochondrial DNA [31]. One particular mitochondrial haplotype can reduce mtDNA polymorphism in the infected population and sweep through a population [19]. Wolbachia may also drive introgression of mtDNA following hybridization events between sibling species: the introduction and spread of Wolbachia in a novel species result in spread of the mtDNA from the neighboring species [40].

Wolbachia infections have been reported in various Lepidoptera families such as Papilionidae, Lycaenidae, Pieridae, Nymphalidae, Hesperiidae, Pyralidae, Noctuidae, and Lasiocampidae. A few butterfly species harboring the bacteria have been thoroughly studied. We have reported the molecular phylogeny of the genus in a prior study [9]. Meanwhile, we have preliminarily screened for the presence of Wolbachia in Polytremsis and found at least three species (P. nascens Leech, 1893, P. theca Evans, 1937 and P. pellucid Murray, 1875) are infected with Wolbachia.

5.1. Wolbachia infection status and genetic structure in natural populations of P. nascens

We surveyed Wolbachia infection rates in 14 regional populations of P. nascens. Twenty-one specimens (31%) infected with Wolbachia in seven regional populations were detected. However, all specimens are free from infection with Wolbachia in the other seven regional populations. The infection rates between the female and male butterflies show no significant difference ($\chi^2 = 0.65, p > 0.05$). The further rearing experiment would be made to reveal whether there was a sex-ratio distortion in P. nascens induced by the strain of Wolbachia. A positive relationship of Wolbachia infections and latitudinal distribution has also been found. The lower or no Wolbachia infection rates in central-west of China may indicate that incidence is apparently lower in regions experiencing longer dry seasons and higher average daily temperatures. The higher Wolbachia prevalence occurs in more southerly moist and temperate populations (Figure 6). This has been observed in the beetle Chelymorma alternans and in ants of the genus Solenopsis [41, 42].
Many explanations have been proposed that deviation from neutral evolution in a species and the absence of diversity in mtDNA can be associated with either a genome-wide bottleneck effect or a selective sweep on mtDNA [43]. In our study, the uninfected butterflies show higher mtDNA polymorphism than butterflies infected with Wolbachia (Table 5). The perfect concordance of Wolbachia infection status and mtDNA polymorphism suggests that the mitochondrial genetic structure of the host insects may be strongly affected by the Wolbachia infection. Decreased mtDNA polymorphism as a consequence of Wolbachia infection has also been reported in several other insects [38, 40]. We also found that the mtDNA genes of the butterflies infected with Wolbachia deviated significantly from neutral evolution according to both D (−2.3303, \( p < 0.05 \)) and F values (−3.7068, \( p < 0.05 \)), while this was not so for uninfected ones [31]. These results suggest that the populations of *P. nascens* have

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Number of haplotypes</th>
<th>Haplotype diversity</th>
<th>Number of variable sites (S)</th>
<th>( \pi )</th>
<th>SD (( \pi ))</th>
<th>D</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>67</td>
<td>12</td>
<td>0.797</td>
<td>119</td>
<td>0.0197</td>
<td>0.0012</td>
<td>1.4795</td>
<td>2.0907*</td>
</tr>
<tr>
<td>Infection</td>
<td>21</td>
<td>2</td>
<td>0.095</td>
<td>12</td>
<td>0.0006</td>
<td>0.0005</td>
<td>−2.3303*</td>
<td>−3.7068*</td>
</tr>
<tr>
<td>Free from infection</td>
<td>46</td>
<td>10</td>
<td>0.750</td>
<td>77</td>
<td>0.0122</td>
<td>0.0015</td>
<td>0.8835</td>
<td>1.9320</td>
</tr>
</tbody>
</table>

*\( N \), number of samples; \( \pi \), nucleotide diversity; SD, standard deviation; D, Tajima’s D statistic; F, Fu’s F statistic.  
*\( \ast \) Significant difference.*

Table 5. Mt haplotype and nucleotide diversity estimates from infected and uninfected samples.

Figure 6. Distribution of mtDNA haplotypes among populations of *P. nascens* collected in China.
recently been subjected to a Wolbachia-induced sweep, making the mtDNA undergo purifying selection. Additionally, we analyzed the population size change of P. nascens infected with Wolbachia and all P. nascens, respectively, by the software DnaSP4.90 [7] and got no evidence for population expansion in Wolbachia-infected group [18]. However, we got unimodal curves representing population expansion in P. nascens. A selective sweep would erase variability in the population even under population expansion, potentially eliminating any evidence of past demographic processes. Moreover, recent studies also revealed that there are some other endosymbionts known to manipulate host reproduction like Wolbachia, for example, Arsenophonus, Cardinium, and Rickettsia [44]. A wide range of insect species can host more than one endosymbiont [42]. The infection status of the secondary endosymbionts in P. nascens needs to be further studied.

Although there seems to be a strong association existed between mtDNA haplotypes and Wolbachia infection status, the association between mtDNA haplotypes, Wolbachia infection, and geographical distribution is weak [31]. In our preliminary experiment, we found that two sympatrically distributed sister species of P. nascens (P. theca and P. pellucid) are infected with Wolbachia. We could not eliminate the possibility of the multiple introgression events and hybridizations between the species pairs [40]. Thus, we excluded the infected group (Clade I) in the analysis of biogeographical implications of P. nascens. It is notable that if we only take the uninfected group into account, the conclusion can be drawn that P. nascens probably has two genetically diverse and geographically localized clades in China based on the mtDNA haplotype phylogeny and networks (Figure 6) [31]. They are the central-west of China clade (H1–H3) and the Eastern and Southern China clade (H4–H10). These two clades are isolated mainly by the Yangtze River, with the exception of a specimen in HH [31]. Our results were similar to those obtained from the striped stem borer, Chilo suppressalis [45], and the melon fly, Bactrocera cucurbitae [46], which suggest that the Yangtze River Range has acted as a substantial barrier to gene flow. This contrasts with studies of the beet armyworm, Spodoptera exigua, and the migratory locust, Locusta migratoria, and the cotton bollworm Helicoverpa armigera, which showed little or no evidence that the Yangtze River Range limits gene flow [47–49].

The network analysis reflects genealogical relationships of the mtDNA haplotypes [31]. The single mutation steps separate adjacent haplotypes in the network and older haplotypes are placed at internal branching points, whereas younger ones occur toward the tip positions [50]. The haplotype network of P. nascens displays a star-like pattern (Figure 6). Haplotype 1 (H1) is in the star’s center, and derivatives are connected to it by short branches. The haplotype is the most common and geographically widespread in central-west of China. The star-like topologies for this cluster strongly suggest the effect of a population expansion based on coalescence theory [20]. We still confirm the result of the population size change of P. nascens with the software DnaSP4.90 and got unimodal curves representing population expansion [31]. The most common H1 had strong support as the ancestral haplotype due to its representation in a significant proportion of individuals in all populations and its basal location in the network [31]. A reticulation connecting multiple haplotypes from the Eastern and Southern China clade was formed in the network [31].
5.2. A prevalence survey of Wolbachia in *P. fukia*

Of the butterflies examined by diagnostic PCR for 16S rRNA, 47% (15/32) were *Wolbachia* positive. The infection rates in female and male are 69% (11/16) and 25% (4/16). For characterization of *Wolbachia* strains, we amplified a segment of the *Wolbachia* cell cycle gene *ftsZ*. BLAST searches of the *ftsZ* sequences yielded a significant sequence similarity between the *Wolbachia* strains infecting *P. fukia* and the *Wolbachia* supergroup A typically found in insects. Phylogenetic analyses suggested that the bacteria are subdivided into two strains with a genetic divergence of approximately 2.8%. These strains will be referred to as *wFuk1* and *wFuk2* in the following.

**Figure 7A** shows the ML tree based on the data set of the concatenated sequences and supports the monophyly of *P. fukia*. On the phylogeny, the sequences are split into three clades supported by high bootstrap values. Clade I exclusively consists of the *P. fukia* from two geographical populations in Southwest side of their distribution area (Maoershan and Anjiangpin). These butterflies are either free from *Wolbachia* infection or infected with *wFuk2*. Clade II exclusively consists of the butterflies from four geographical populations (West Tianmushan, Qingliangfeng, Wuyishan, and Jinggangshan), which are either free from *Wolbachia* infection or infected with *wFuk2*. These two clades are consistent with the distribution of geographical population. Eight females constitute Clade III are invariably infected with *wFuk1*. In our data set comprising 32 individuals (16♀ and 16♂), we did not detect any male *P. fukia* infected.

![Figure 7](http://dx.doi.org/10.5772/intechopen.69958)

**Figure 7.** (A) Maximum-likelihood phylogeny on the basis of the concatenated mitochondrial sequences and (B) network on the basis of the concatenated mitochondrial sequences constructed with software Network4.5.
with wFuk1. All eight individuals infected with wFuk1 are female. Males belong to Clade I or Clade II and, if infected, the strains wFuk2. All mtDNA of P. fukia were used for network construction with the software Network4.5 using the median-joining method (Figure 7B). Twenty-three haplotypes are clustered into three groups in accordance with three clades in ML tree (see Figure 7A). Two haplotypes were shared by the individuals infected with Wolbachia and free from Wolbachia. The Wolbachia-induced sweep has been shown in P. nascaes [51]. However, such Wolbachia effects on mtDNA variation in P. fukia examined in this study were not as conspicuous as in the above studies. Still, mtDNA variation in the P. fukia is likely to be weakly associated with the presence of wFuk1, although the sample size was not large enough to draw a firm conclusion [37]. The nucleotide diversity in wFuk1-infected butterflies is smaller than that of uninfected butterflies (Table 6) and their mt concatenated sequences form a clade solely (Figure 7), presumably reflecting a Wolbachia sweep. This tendency was consistently seen in all the mitochondrial regions examined. A larger sample will reveal more in detail the association of Wolbachia infection status with variation of mtDNA in P. fukia.

### Acknowledgements

This study was financially supported by grants from National Natural Science Foundation of China (no. 31401997) and Young Teacher Training Scheme for Colleges and Universities in Shanghai (no. ZZssd15068).

### Author details

Weibin Jiang

Address all correspondence to: jiangwb@shnu.edu.cn

College of Life and Environmental Sciences, Shanghai Normal University, Shanghai, People’s Republic of China

<table>
<thead>
<tr>
<th>N</th>
<th>Number of haplotypes</th>
<th>Hd</th>
<th>SD (Hd)</th>
<th>Number of variable sites (S)</th>
<th>π</th>
<th>SD (π)</th>
<th>D</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>All sequences</td>
<td>32</td>
<td>23</td>
<td>0.980</td>
<td>0.012</td>
<td>50</td>
<td>0.008</td>
<td>0.001</td>
<td>0.568</td>
</tr>
<tr>
<td>Infection with wFuk1</td>
<td>8</td>
<td>4</td>
<td>0.821</td>
<td>0.101</td>
<td>5</td>
<td>0.001</td>
<td>0.000</td>
<td>0.840</td>
</tr>
<tr>
<td>Infection with wFuk2</td>
<td>7</td>
<td>7</td>
<td>1.000</td>
<td>0.076</td>
<td>32</td>
<td>0.008</td>
<td>0.001</td>
<td>0.077</td>
</tr>
<tr>
<td>Free from infection</td>
<td>17</td>
<td>14</td>
<td>0.978</td>
<td>0.027</td>
<td>36</td>
<td>0.007</td>
<td>0.002</td>
<td>0.522</td>
</tr>
</tbody>
</table>

N, number of samples; Hd, haplotype diversity; π, nucleotide diversity; SD, standard deviation; D, Tajima’s D statistic; F, Fu’s F statistic.

Table 6. Mt haplotype and nucleotide diversity estimates from infected and uninfected samples.
References


[10] Huang H. A list of butterflies collected from Nujiang and Dulongjiang, China with descriptions of new species, new subspecies and revisional notes. Neue Entomologische Nachrichte. 2003;55:3-114


Shoemaker DD, Dyer KA, Ahrens M, McAbee K, Jaenike J. Decreased diversity but increased substitution rate in host mtDNA as a consequence of *Wolbachia* endosymbiont infection. Genetics. 2004;168:2049-2058


