

# SCIENTIFIC REPORTS



OPEN

## Taxonomic status and molecular phylogeography of two sibling species of *Polytremis* (Lepidoptera: Hesperidae)

Weibin Jiang<sup>1</sup>, Haiyan He<sup>1</sup>, Yingdong Li<sup>1</sup>, Mengyi Ren<sup>1</sup>, Yazhong Ma<sup>1</sup>, Lingli Zheng<sup>1</sup>, Jianqing Zhu<sup>2</sup> & Weidong Yu<sup>1</sup>

Received: 17 July 2015

Accepted: 08 January 2016

Published: 10 February 2016

The skipper *Polytremis theca* species complex is widely distributed in the south of the Qinling Mountains in China. A recent study of the *Polytremis* genus suggested that this species might encompass two differentiated lineages. We tested this hypothesis, by carrying out a phylogenetic study of this agricultural pest based on nationwide sampling and the evaluation of mitochondrial and nuclear DNA markers. We show that this species is actually an amalgamation of two sibling taxa (*P. t. theca* and *P. t. fukia*), which displayed levels of genetic divergence as great as those generally found between sister species in the *Polytremis* genus, suggesting that they actually correspond to two distinct species. The Divergence time estimates suggest that an active period of speciation within *Polytremis* occurred within the Pleistocene eras. Based on its distinct phylogenetic placement and geographical isolation, we suggest that the subspecies should be elevated to full species status under the phylogenetic species concept, which has significant management implications.

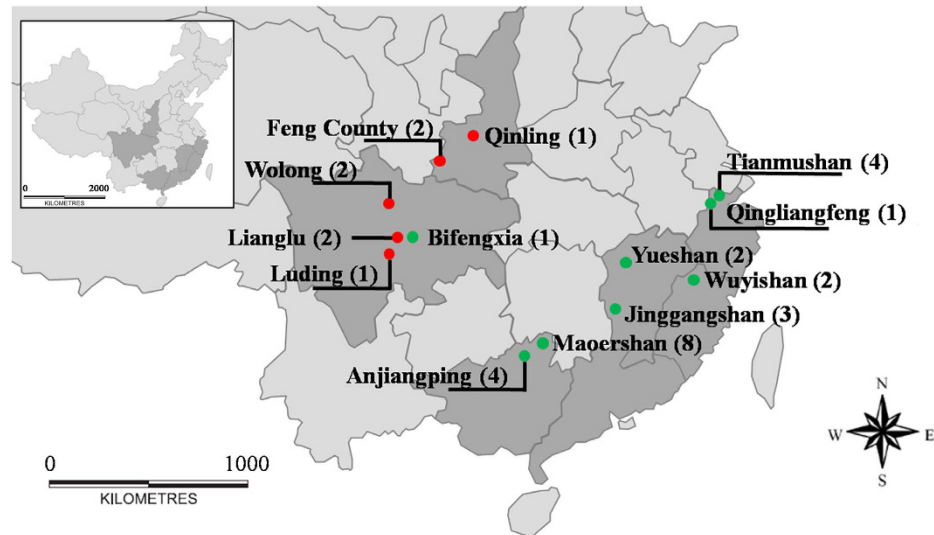
Evans<sup>1</sup> described a lot of skipper species from Europe, Asia and Australia. He often described varieties, now considered subspecies under the current Code of Zoological Nomenclature<sup>2</sup>, when specimens exhibited morphology similar to known taxa. It appears to follow the classical typological species concept where small differences represented regional variation in the species. Close inspection of such intraspecific variations has sometimes revealed that the species' actually contain several closely related multiple species, so-called 'sibling species' or 'cryptic species', that are difficult to distinguish morphologically from each other, at least without skilful eyes<sup>3–5</sup>. With the development of molecular techniques and the collection of greater number of specimens, some of his subspecies were argued, revised and elevated to species rank. Chiba & Eliot<sup>6</sup>, Devyatkin & Monastyrskii<sup>7</sup>, Guo *et al.*<sup>8</sup> reevaluated *Parnara guttatus batta* Evans, 1949 by morphological features, molecular data and biogeographical implications and treated it as an independent species named *Parnara batta* Evans, 1949. Chiba *et al.*<sup>9</sup> and Hsu<sup>10</sup> regarded *Polytremis menica kirazana* Evans, 1949 as valid species-level named *Polytremis kiraizana* Sonan, 1938 by further morphological and phylogenetic analysis.

The skipper *Polytremis theca* species complex is widely distributed in the south of the Qinling Mountains in China, except Hainan Province and the southern tropical regions of Yunnan Province<sup>1,11,12</sup>. Evans<sup>1</sup> described *P. t. theca* Evans, 1937 from Sichuan Province. He then described the *P. t. fukia* Evans, 1940 from Fujian Province that was lighter in coloration than the typical *P. theca* and existed 1 or 2 white spots in space Cu<sub>2</sub> on the upperside of forewing. *P. t. theca* and *P. t. fukia* are nearly separated geographically.

If genetic differentiation has occurred, the taxonomic status of *P. theca* species complex can be reevaluation by molecular techniques. 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 inter-specific genetic distances in *Polytremis*<sup>13</sup>. Additionally, *P. theca* is also the only species whose intra-specific distance is more than 3%, a thresholds of species identification have been proposed in Lepidoptera for COI, in genus *Polytremis*<sup>14</sup>. Thus, we suspected the possible existence of a sibling species paired or the cryptic diversity in the species. However, this assertion was based on

<sup>1</sup>Shanghai Normal University, College of Life and Environmental Sciences, Shanghai, People's Republic of China.

<sup>2</sup>Shanghai Zoological Park, Shanghai, People's Republic of China. Correspondence and requests for materials should be addressed to W.Y. (email: ywd@shnu.edu.cn) or J.Z. (email: zzzjjq@gmail.com)



**Figure 1.** Distribution of the specimens of *P. t. theca* collected in China; red circle, *P. t. theca*, green circle, *P. t. fukia*. We modified the map from Jiang *et al.*<sup>13</sup>

a very small number of samples and required confirmation in a larger sample. To answer the question, we conducted molecular phylogenetic analyses of *P. theca* species complex collected from 13 geographic localities in China using a region of mitochondrial genes (COI) and a region of nuclear genes (*wingless*).

## Materials and methods

**Samples collection.** We collected specimens of *P. t. theca* (8) from five local regions and *P. t. fukia* (25) from eight different localities in China from 2008 to 2014 (Fig. 1; Table 1), and preliminarily identified them based on traditional wing and genitalia morphology characters recommended by Evans<sup>1</sup>, Eliot<sup>15</sup> and Chou<sup>16</sup>. Our sampling of 33 specimens covered the major range of the *P. theca* (Fig. 1). The outgroups included eight specimens of *P. nascens* Leech, 1893, five specimens of *P. mencia* Moore, 1877 (Table 1). All specimens were caught in the field, preserved by dehydration in small envelopes and dried with silica desiccant for further processing.

**DNA extraction and sequencing.** The DNA was isolated from legs of adult butterfly using a QIAamp DNA Mini kit (Qiagen, Hilden, Germany) essentially following the manufacturer's instructions but with some modification. Briefly, after adding proteinase K and buffer AL (QIAGEN®), the mixed homogeneous solution was incubated at 70 °C for 2 h. Subsequently, 200 µL of 100% ethanol was added and the mixture transferred to a QIAamp spin column. The mixture in the spin column was subjected to 3 cycles of centrifugation at full speed (14,000g) for 1 min and the filtrate was returned to the spin column to increase the amount of DNA obtained. Partial sequences of mitochondrial gene COI (487 bp) was amplified and sequenced with primers HCO2198: 5'-TAA ACT TCA GGG TGA CCA AAA AAT CA-3' and LCO1490: 5'-GGT CAA CAA ATC ATA AAG ATA TTG G-3'<sup>17</sup>. The nuclear DNA fragments of *wingless* (389) were amplified and sequenced with primers WG1z: 5'-GAR TGY AAR TGY CAY GG-3' and WG2a: 5'-ACT ICG CAR CAC CAR TGG AAT GTR CA-3'<sup>18</sup>. The PCR for all amplicons were conducted in 20 µL final volume reactions containing 1 µL template DNA, 0.4 µM each of the amplification primers, 200 µM each dNTP, 1.5 mM MgCl<sub>2</sub>, 50 mM KCl, 10 mM Tris-HCl (pH 8.3) and 2 U Taq polymerase (Takara, Otsu, Shiga, Japan). The thermal profile used was as follows. COI: 95 °C denaturation for 5 min followed by 35 cycles of 94 °C for 45 s, 42 °C for 1 min, 90 °C for 90 s and a final elongation step at 72 °C for 10 min. *Wingless*: 95 °C denaturation for 5 min followed by 30 cycles of 94 °C for 60 s, annealing at 55 °C for 60 s and extension at 72 °C for 2 min, with a final extension of 72 °C for 10 min.

Extraction blanks were run in all reactions to control for contamination during the extraction and PCR processes. The amplification products were subjected to electrophoresis in a 2% (w/v) agarose gel in TAE buffer (0.04 M Tris-acetate, 0.001 M EDTA) with a DL1000 ladder size marker (Takara, Otsu, Shiga, Japan) to determine whether the amplification reactions were successful. In addition, after electrophoresis in agarose gels, the amplification products were extracted using the Wizard SV Gel and PCR Clean-up System (Promega, Madison, WI, USA) for sequencing. Finally, all the haplotypes obtained were deposited in GenBank (Table 1).

**Statistical analyses.** The sequence data of the mitochondrial COI and nuclear *wingless* were aligned with published homologous sequence from genus *Polytremis* (e. g., *P. eltola* Hewitson, 1869, *P. discreta* Elwes & Edwards, 1897, etc.) as well as *Borbo cinnara* Wallace, 1866 and *Pseudoborbo bevani* Moore, 1878 (presented in Table 1) respectively, translated to amino acid sequences to check for nuclear mitochondrial pseudogenes (numts) and pruned to remove redundant sequences with Bioedit v.7.0<sup>19</sup>. The haplotype sequence matrix was used for all subsequent phylogenetic analyses (Table 1). Phylogenetic trees were constructed by the ML (maximum-likelihood) methods with PhyML<sup>20</sup>. Modeltest 3.7<sup>21</sup> was used to select the optimal nucleotide substitution models following the Akaike Information Criterion (AIC). In the ML analysis, a heuristic search was

Locality	Specimen ID	Sex	Date	Altitude	Accession number		
					Haolotype (COI/wingless)	COI	wingless
<i>Polytrems theca theca</i>							
Lianglu, Tianquan County, Sichuan Province (N 29°55'; E 102°23')	Ptt_LL 1	♀	6-Jul-11	not recorded	Ptt I/Ptt I	KC684410'	KR911938
	Ptt_LL 2	♂	11-Aug-11	not recorded	Ptt I/Ptt II	KC684410'	KR911939
Wolong, Wenchuan County, Sichuan Province (N 31°29'; E 103°35')	Ptt_WL 1	♂	6-Aug-11	not recorded	Ptt II/Ptt II	KC684411'	KR911939
	Ptt_WL 2	♂	6-Aug-11	not recorded	Ptt II/Ptt II	KC684411'	KR911939
Luding, Sichuan Province (N 29°55'; E 102°15')	Ptt_LD 1	♂	19-Jul-11	not recorded	Ptt II / Ptt II	KC684411'	KR911939
Feng County, Shanxi Province (N 33°54'; E 106°31')	Ptt_FC 1	♀	19-Jul-11	not recorded	Ptt III/Ptt I	KR911924	KR911938
	Ptt_FC 2	♂	19-Jul-11	not recorded	Ptt III/Ptt I	KR911924	KR911938
Qinling, Shanxi Province 34°14'N 103°54'E	Ptt_QL 1	♂	16-Jul-10	not recorded	Ptt III/Ptt I	KR911924	KR911938
<i>Polytrems theca fukia</i>							
Qinliangfeng, Lin'an County, Zhejiang Province (N 30°59'; E 118°52')	Ptf_QLF 1	♂	11-Apr-13	800 m	Ptf II/Ptf I	KC684406'	KR911935
Tianmushan, Lin'an County, Zhejiang Province (N 30°19'; E 119°25')	Ptf_TMS 1	♀	8-May-09	not recorded	Ptf I/Ptf I	KC684409'	KR911935
	Ptf_TMS 2	♂	20-Sep-08	not recorded	Ptf I/Ptf I	KC684409'	KR911935
	Ptf_TMS 3	♂	17-Aug-09	not recorded	Ptf I/Ptf I	KC684409'	KR911935
	Ptf_TMS 4	♂	20-Aug-13	300 m	Ptf II/Ptf I	KC684406'	KR911935
Wuyishan, Fujian Province (N 27°44'; E 118°01')	Ptf_WYS 1	♂	19-Sep-08	not recorded	Ptf III/Ptf II	KC684408'	KR911936
	Ptf_WYS 2	♂	19-Sep-08	not recorded	Ptf III/Ptf II	KC684408'	KR911936
Jingangshan, Ji'an county, Jiangxi Province (N 26°45'; E 114°17')	Ptf_JGS 1	♀	19-Oct-10	not recorded	Ptf I/Ptf I	KC684409'	KR911935
	Ptf_JGS 2	♀	19-May-10	not recorded	Ptf II/Ptf I	KC684406'	KR911935
	Ptf_JGS 3	♀	20-May-10	not recorded	Ptf II/Ptf I	KC684406'	KR911935
Yueshan, Fengxin County, Jiangxi Province (N 28°79'; E 115°18')	Ptf_YS 1	♀	13-Jul-13	400–600 m	Ptf II/Ptf I	KC684406'	KR911935
	Ptf_YS 2	♀	13-Jul-13	400–600 m	Ptf IV/Ptf I	KR911919	KR911935
Maoershan, Xin'an County, Guangxi Province (N 25°49'; E 110°00')	Ptf_MES 1	♀	8-Jul-11	400 m	Ptf V/Ptf II	KC684407'	KR911936
	Ptf_MES 2	♂	7-Jul-11	400 m	Ptf II/Ptf III	KC684406'	KR911937
	Ptf_MES 3	♂	7-Jul-11	400 m	Ptf VI/Ptf II	KR911920	KR911936
	Ptf_MES 4	♀	7-Jul-11	400 m	Ptf VII / Ptf III	KR911921	KR911937
	Ptf_MES 5	♂	8-Jul-11	400 m	Ptf II/Ptf III	KC684406'	KR911937
	Ptf_MES 6	♂	8-Jul-11	400 m	Ptf II/Ptf III	KC684406'	KR911937
	Ptf_MES 7	♂	8-Jul-11	400 m	Ptf VIII/Ptf I	KR911922	KR911935
	Ptf_MES 8	♂	9-Jul-11	2000 m	Ptf VIII/Ptf I	KR911922	KR911935
Anjiangpin, Lingui County, Guangxi Province (N 25°14'; E 110°12')	Ptf_AJP 1	♂	13-Jul-11	1200 m	Ptf I/Ptf III	KC684409'	KR911937
	Ptf_AJP 2	♂	13-Jul-11	1200 m	Ptf II/Ptf II	KC684406'	KR911936
	Ptf_AJP 3	♂	13-Jul-11	1200 m	Ptf IX/Ptf III	KR911923	KR911937
	Ptf_AJP 4	♂	14-Jul-11	1350 m	Ptf VIII/Ptf II	KR911922	KR911936
Bifengxia, Yaan County, Sichuan Province (N 30°03'; E 102°62')	Ptf_BFX 1	♀	2-Aug-14	not recorded	Ptf II/Ptf I	KC684409'	KR911935
Outgroups							
<i>Polytrems nascens</i>							
Zhanghe, Langao County, Shaanxi Province (N 32°32'; E 108°14')	Pn_ZH 1	♂	30-Jul-12	1600 m	Pn I/Pn I	KJ574014'	KR911931
	Pn_ZH 2	♂	30-Jul-12	1600m	Pn II/Pn I	KC684397'	KR911931
Hailuogou, Tianquan County, Sichuan Province (N 29°34'; E 102°04')	Pn_HLG 1	♂	27-Jul-06	2300 m	Pn I/Pn II	KJ574014'	KR911932
	Pn_HLG 2	♂	27-Jul-06	2300 m	Pn II/Pn II	KC684397'	KR911932
Lianglu, Tianquan County, Sichuan Province (N 29°55'; E 102°23')	Pn_LL 1	♂	3-Sep-10	1450 m	Pn I/Pn II	KJ574014'	KR911932
Fenghuangshan, Hanyin County, Shaanxi Province (N 32°90'; E 108°50')	Pn_FHS 1	♀	3-Aug-12	1400 m	Pn II/Pn I	KC684397'	KR911931
	Pn_FHS 2	♀	3-Aug-12	1600 m	Pn III/Pn I	KJ574015'	KR911931
Houhe, Wufeng County, HuBei Province (N 30°22'; E 110°68')	Pn_HH 1	♂	9-Jul-13	1200 m	Pn II/Pn III	KC684397'	KR911933
Continued							

Locality	Specimen ID	Sex	Date	Altitude	Accession number		
					Haolotype (COI/wingless)	COI	wingless
<i>Polytrems mencia</i>							
Tianmushan, Lin'an County, Zhejiang Province (N 30°19'; E 119°25')	Pm_TMS 1	♂	5-Jun-11	not recorded	Pm I/Pm I	KC684415*	KR911940
	Pm_TMS 2	♂	15-Jun-10	not recorded	Pm I/Pm I	KC684415*	KR911940
	Pm_TMS 3	♂	30-May-10	not recorded	Pm II/Pm I	KC684414*	KR911940
	Pm_TMS 4	♀	20-Sep-08	not recorded	Pm III/Pm I	KC684416*	KR911940
Ningbo, Zhejiang Province (N 29°52'; E 121°32')	Pm_NB 1	♂	23-Aug-11	not recorded	Pm IV/Pm I	KC684417*	KR911940
<i>Polytrems pellucid</i>							
Tianmushan, Lin'an County, Zhejiang Province (N 30°19'; E 119°25')	121112009	♀	19-Sep-08	not recorded	<i>P. pellucid</i>	KC684393*	KR911941
<i>Polytrems zina</i>							
Tianmushan, Lin'an County, Zhejiang Province (N 30°19'; E 119°25')	121116021	♂	18-Aug-09	not recorded	<i>P. zina</i>	KC684395*	KR911942
<i>Polytrems discreta</i>							
Baoxing County, Sichuan Province (N 30°22'; E 102°48')	121116024	♂	2-Jul-08	not recorded	<i>P. discreta</i>	KC684392*	KR911943
<i>Polytrems lubricans</i>							
Baoxing County, Sichuan Province (N 30°22'; E 102°48')	121116027	♂	2-Jul-08	not recorded	<i>P. lubricans</i>	KC684391*	KR911944
<i>Polytrems eltola</i>							
Jingxiu, Guangxi Province (N 24°07'; E 110°11')	121116030	♂	28-Jul-11	not recorded	<i>P. eltola</i>	KC684389*	KR911945
<i>Polytrems gigantean</i>							
Qingchengshan, Sichuan Province (N 30°53'; E 103°34')	121119033	♀	28-Aug-11	not recorded	<i>P. gigantean</i>	KC684403*	KR911929
<i>Polytrems matsuii</i>							
Hongya, Sichuan Province (N 29°54'; E 103°22')	121119036	♂	4-Jun-11	not recorded	<i>P. matsuii</i>	KC684400*	KR911927
<i>Polytrems caerulescens</i>							
Lianglu, Tianquan County, Sichuan Province (N 29°55'; E 102°23')	121119039	♂	21-Jul-11	not recorded	<i>P. caerulescens</i>	KC684399*	KR911925
<i>Polytrems jigongi</i>							
Tianmushan, Lin'an County, Zhejiang Province (N 30°20'; E 119°23')	121119042	♂	11-Jul-09	not recorded	<i>P. jigongi</i>	KC684404*	KR911934
<i>Polytrems kiraizana</i>							
Qilai, Taiwan Province (N 24°01'; E 121°22')	121124051	♂	11-Jul-91	not recorded	<i>P. kiraizana</i>	KC684401*	KR911926
<i>Polytrems suprema</i>							
Jingxiu, Guangxi Province (N 24°07'; E 110°11')	121124052	♂	31-Jul-11	not recorded	<i>P. suprema</i>	KC684402*	KR911928
<i>Polytrems gotama</i>							
Lijiang, Yunnan Province (N 26°51'; E 100°13')	130101067	♂	28-Jul-06	not recorded	<i>P. gotama</i>	KC684396*	KR911930
<i>Borbo cinnara</i>							
Taidong, Taiwan Province (N 22°59'; E 120°59')	121124069	♂	17-Sep-07	not recorded	<i>B. cinnara</i>	KC684418*	KR911947
<i>Pseudoborbo bevani</i>							
Nantou, Taiwan Province (N 23°55'; E 120°41')	121124067	♀	14-Sep-03	not recorded	<i>P. bevani</i>	KC684419*	KR911946

**Table 1.** List of specimen information used for this study. \* indicates the sequences retrieved from GenBank.

conducted. The starting tree for branch-swapping was from stepwise addition. Nodal support of the ML tree was estimated by 1000 bootstraps. Network profile of the haplotypes identified in *P. t. theca*, *P. t. fukia*, *P. nascens* and *P. mencia* was constructed with Network4.5 using the median-joining method<sup>22</sup>. The haplotype diversity (Hd) and nucleotide diversity ( $\pi$ ) for *P. t. theca* and *P. t. fukia* were estimated by DnaSP4.90<sup>23</sup>.

For the COI data set, pairwise  $F_{ST}$  was also calculated with Arlequin v3.0<sup>24</sup>, which accurately reflects patterns of genetic variation. Correspondingly, gene flow was estimated in Arlequin v3.0<sup>24</sup>. Two level hierarchical analyses

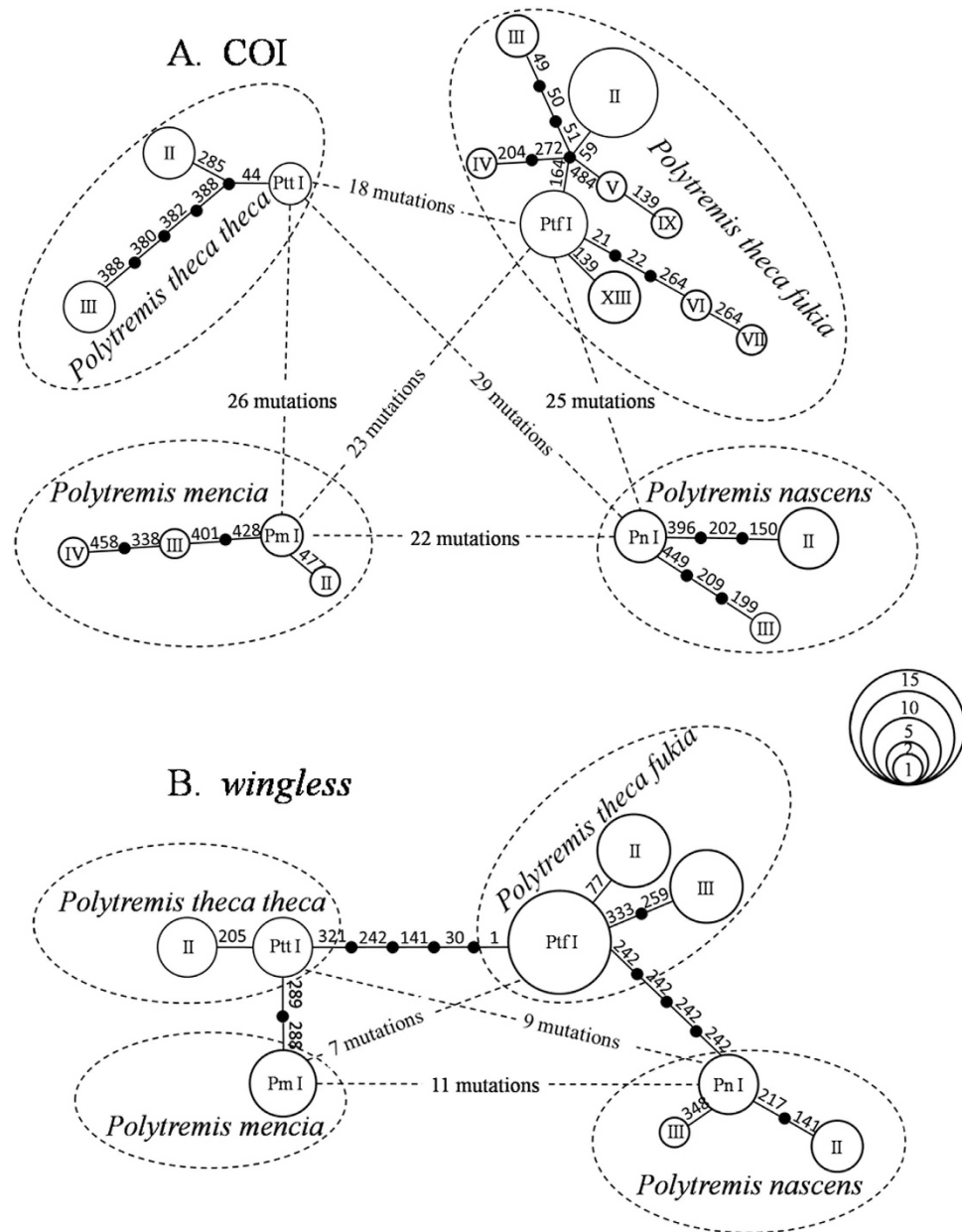
A. COI		11111111111111111111222222222333333333333333333444444444444	
		122444455556677801123344555689900026667823345678888901112234578	
		8124569013989479139519690284519246914792548970840289613692819874	
Ptf I	ACCTCATATATTTCATTTAATTCTTCTCTCTGTTAACTCGTTTTGAACCTTCTCGACATTTTGC		5
Ptf II	.....C.....T.....		9
Ptf III	.....CGC.....C.....		2
Ptf IV	.....C.....C.....A.....		1
Ptf V	.....C.....G.....		1
Ptf VI	.GG.....C.....T.....A.....		1
Ptf VII	.GG.....C.....T.....A.....C.....		1
Ptf VIII	.....C.....G.....T.....		3
Ptf IX	.....C.....G.....G.....		1
Ptt I	...CTT...C...GC...GCC...ATTC.....G...TT...ATT...CA...		2
Ptt II	...TT...C...GC...GCC...ATTC.....C...G...TT...ATT...CA...		3
Ptt III	...TT...C...GC...GCC...AATTC.....G...TTCAT...TATT...CA...		3
Pn I	...TT...CT...AT...CCT...CCT...ATTCA...C...T...TA...C...A...TT...TCT...C...A...		3
Pn II	...TT...CT...AT...CCT...CC...ATTCA...C...T...TA...C...A...TT...T...T...C...A...		4
Pn III	...TT...CT...AT...CCT...CC...ATTCA...CTT...TA...C...A...TT...T...T...CA...A...		1
Pm I	TT...T...CAT...T...C...ATT...T...T...TA...C...TT...T...T...TA...TT...A...		2
Pm II	TT...T...CAT...T...C...ATT...T...T...TA...C...TT...T...T...TA...TT...A...		1
Pm III	TT...T...CAT...T...C...ATT...T...T...TA...C...TT...T...T...A...TTC...A...		1
Pm IV	TT...T...CAT...T...C...ATT...T...T...TA...A...TT...T...T...A...TTC...CA...		1
B. <i>wingless</i>		122222222223333	
		2374014445688234	
		12071572459289138	
Ptf I	CTGGGGCCAAATCCCC		12
Ptf II	...A.....		6
Ptf III	.....G...T...		6
Ptt I	G...A...A...T...T...		4
Ptt II	G...A...AA...T...T...		4
Pn I	.C.....GG...C....		4
Pn II	.C...A...T...GG...C....		3
Pn III	.C.....GG...C...T...		1
Pm I	G...A...A...T...TT...		5

**Figure 2.** Sequence nucleotide substitutions of (A) the 19 haplotypes identified in 45 COI gene sequences and (B) the 9 haplotypes identified in 45 *wingless* gene sequences of *P. t. theca* (Ptt), *P. t. fukia* (Ptf), *P. nascens* (Pn) and *P. mencia* (Pm). The numbers on top of sequences indicates the nucleotide position which is relative to beginning of the fragment investigated in this study. The last column shows the number of individuals shared each haplotype. Dots (·) denote the identity with the reference sequence (Ptf I).

of molecular variance (AMOVA) were conducted to evaluate possible population genetic structure of *P. t. theca* and *P. t. fukia* using Arlequin v3.0 with 1,000 permutations. We calculated Tajima's  $D^{25}$  and Fu's  $F$  statistic<sup>26</sup> and ran 10,000 coalescent simulations for each statistic to create 95% confidence intervals investigate the historical population demographics and testing whether the sequences conformed to the expectations of neutrality. Pairwise mismatch distribution analyses were performed for all *P. theca* specimens and two subspecies specimens separately to find the evidence of past demographic expansions using DnaSP4.90<sup>23</sup>. The times to the most recent common ancestor of the major lineages and the whole population were estimated using relax-clock molecular dating estimation implemented in the BEAST 1.5.2<sup>27</sup>. Analyses using the HKY model of nucleotide substitution with gamma distributed rate variation among sites were performed. The Yule speciation method was assumed and the nucleotide substitution rate of 3.54% per million years that has generally been calibrated for COI in insect<sup>28</sup> was used. Chains were run for 50 million generations, with the first 20% discarded as burn-in. The results were summarized through TRACER 1.5<sup>29</sup>.

## Result

**Genetic Divergences and haplotype networks.** All 46 samples yielded high-quality of DNA and were successfully sequenced for the mitochondrial COI and nuclear *wingless* (accession numbers KR911919-KR911947, Table 1, Fig. 2). The data set of COI alignment contains 487 nucleotide positions, of which 64 positions are variable and 47 are parsimony informative. The mean base composition of the fragment shows a strong bias of A + T (T 39.4%, C 18.2%, A 28.7% and G 13.7%), as found commonly in insect mitochondrial genomes<sup>30</sup>. Nineteen haplotypes were identified in all 46 samples (three in *P. t. theca*, nine in *P. t. fukia*, three in *P. nascens* and four in *P. mencia*) and the haplotype network was constructed and presented in Fig. 3. 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 distance 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 West Tianmu Mountains, one from Jinggang Mountains, and one from Lingui. Ptf II was the most common haplotype in *P. t. fukia* and shared with ten samples. Haplotype Ptf III was found in two samples from Wuyishan. Haplotype Ptf XIII was identified in two sampled from Maoershan and one sampled from Anjiangpin. The remaining haplotypes of *P. t. fukia* occurred in only one individual. At least 18 nucleotide substitutions were observed between the potential ancestral haplotypes of *P. t. theca* (Ptt I) and that



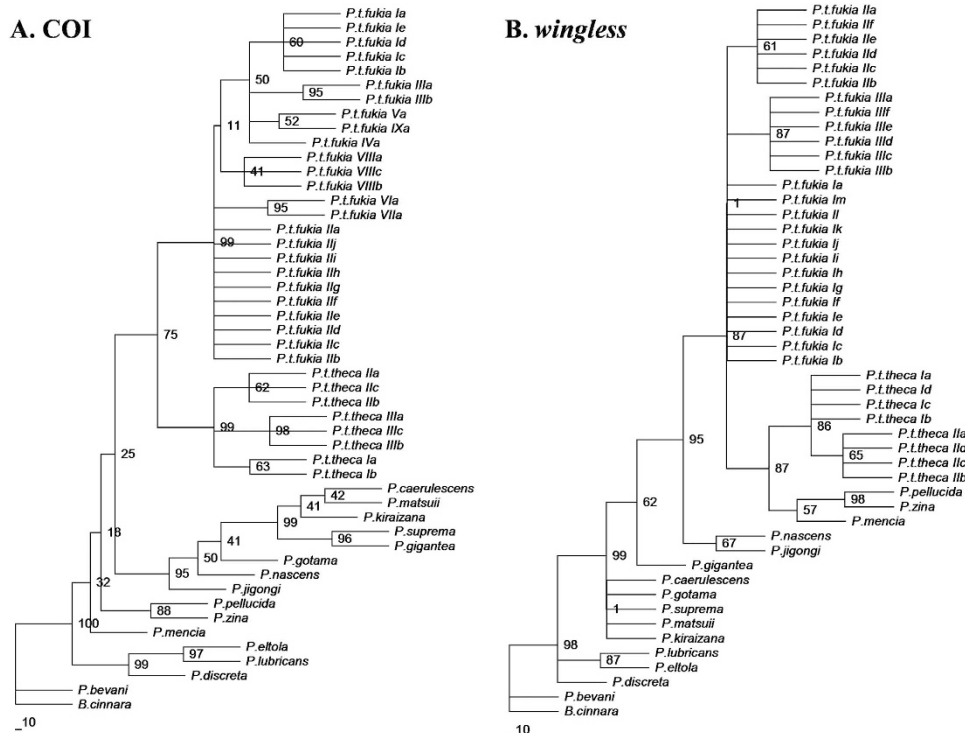
**Figure 3.** Network profile of (A) COI and (B) *wingless* gene haplotypes based on the nucleotide sequences of *P. t. theca*, *P. t. fukia*, *P. nascens* and *P. mencia*. Each haplotype is represented by a circle. The area of the circles is proportional to the number of individuals that shared the haplotype. The circles labelled 1, 2, 5, 10 and 15 on the right hand side represent the number of individuals. The number of unique and shared haplotypes identified in different locations is shown in Table 1. The order of mutations (which were scored relative to haplotype Ptt I) on a branch is arbitrary.

of *P. t. fukia* (Ptf I). In addition, the haplotype (Pn I) of *P. nascens* differed from Ptt I and Ptf I by 29 and 25 nucleotide substitutions while the haplotype (Pm I) of *P. mencia* differed from Ptt I and Ptf I by 26 and 23 nucleotide substitutions, respectively (Figs 2 and 3).

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, ten haplotypes were found in all samples, in which two haplotypes were found in *P. t. theca*, four in *P. t. fukia*, three in *P. nascens* and one in *P. mencia*. All the haplotypes were used for network construction with the software Network 4.5 using the median-joining method (Fig. 3). Haplotypes of the same taxon differed from each other by no more than two mutation distance, in particular, the two 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). In addition, the haplotype (Pn I) of *P. nascens* differed from Ptt I and Ptf I by nine and five nucleotide substitutions while the haplotype (Pm I) of *P. mencia* differed from Ptt I and Ptf I by two and seven nucleotide substitutions, respectively (Figs 2 and 3).

	Ns	Nh	Hd	Nv	$\pi$	SD ( $\pi$ )	D	F
All <i>P. theca</i> samples (COI)	33	12	0.875	38	0.0207	0.0019	0.281	2.368
<i>P. t. theca</i> samples (COI)	8	3	0.750	6	0.0064	0.0048	1.598	2.631
<i>P. t. fukia</i> samples (COI)	25	9	0.803	13	0.0050	0.0070	-1.014	-1.886*
All <i>P. theca</i> samples (wingless)	33	5	0.773	9	0.0078	0.0057	1.122	1.509
<i>P. t. theca</i> samples (wingless)	8	2	0.571	1	0.0015	0.0010	1.444	1.100
<i>P. t. fukia</i> samples (wingless)	25	3	0.640	3	0.0030	0.0020	1.080	1.159

**Table 2. 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.

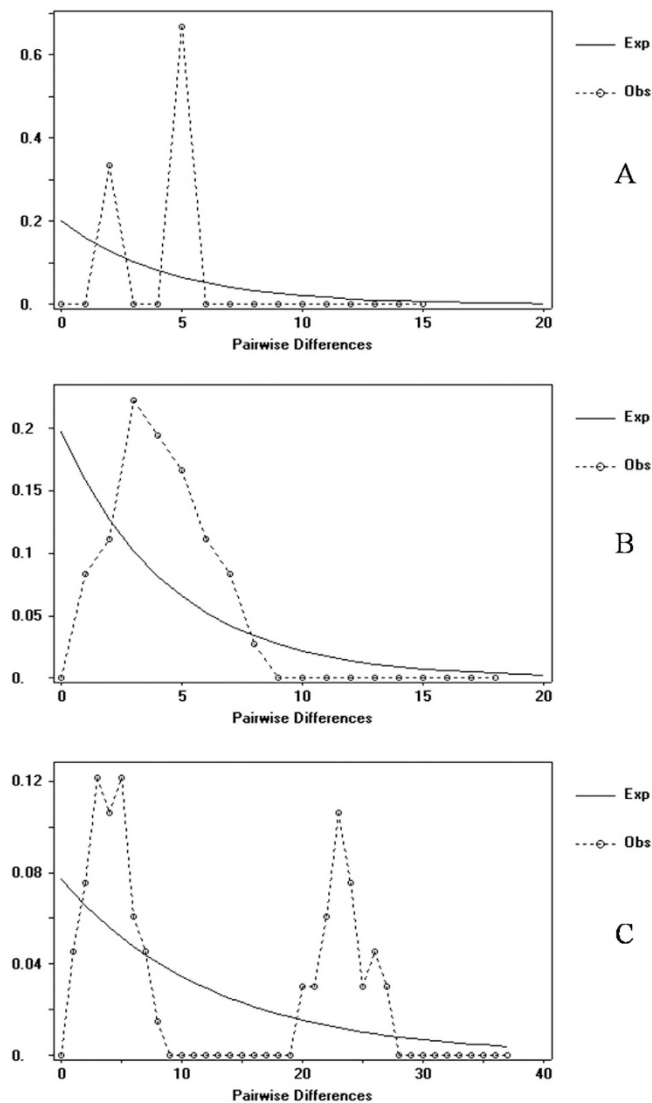


**Figure 4. The maximum-likelihood tree for mitochondrial COI and nuclear *wingless* haplotypes of *Polytremis*.** (A) The submodel GTR + I + G was selected for the COI gene sequence data; (B) The submodel GTR + G was selected for the *wingless* sequence data.

Source of variation	df	Sum of squares	Variance components	Percentage variation	$\Phi$ Statistic
Among populations	1	169.650	9.98266 Va	88.53	—
Within populations	11	45.269	1.29341 Vb	11.47	0.896 ( $p < 0.01$ )
Total	12	214.919	11.27607		
Fixation index	0.8853				

**Table 3. Analysis of molecular variance (AMOVA) for the COI sequences of *P. t. theca* and *P. t. fukia*.**

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 ( $\pi$ ) for *P. t. theca* and *P. t. fukia* are given in Table 2. Additionally, they differed from each other by  $5.07 \pm 0.49\%$  (4.3–5.9% divergence) for the COI sequences and by  $1.70 \pm 0.27\%$  (1.3–2.1% divergence) for the *wingless* sequences.



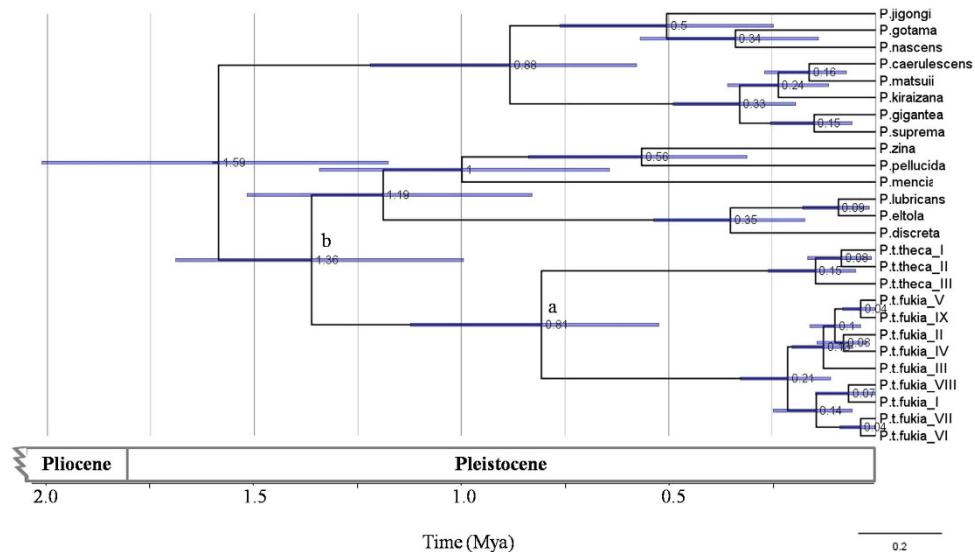
**Figure 5.** Mismatch distributions of pairwise nucleotide differences for (A) *P. t. theca*, (B) *P. t. fukia* and (C) *P. t. theca* and *P. t. fukia*. X axis: Pairwise Differences. Y axis: Frequency. The circles show the observed distribution of pairwise difference. The solid lines represent the expected equilibrium distributions.

**Phylogenetic and population structure analysis.** Phylogenetic analyses were performed using ML method. For the COI gene sequence data, the submodel GTR + I + G was selected. For the *wingless* sequence data, the submodel GTR + G was selected. Figure 4A shows the ML tree based on the data set of COI. On the phylogeny, the haplotypes of *P. theca* were split into two discrete clades. One clade with three haplotypes consisted exclusively of the *P. t. theca* (strongly supported by the bootstrap probability of 99%), whereas the other clade with nine haplotypes exclusively consisted of the *P. t. fukia* (strongly supported with bootstrap probability of 99%). Figure 4B shows the ML tree based on the data set of *wingless*. Three haplotypes, constituted a monophyletic group, exclusively containing *P. t. fukia* butterflies whereas two haplotypes of the *P. t. theca*, *P. pellucida* Murray, 1875, *P. zina* Evans, 1932 and *P. mencia* Moore, 1877 constituted a strongly supported clade (89%). On the other hand, two *P. t. theca* haplotypes formed a strongly supported clade (86%). Mitochondrial COI and nuclear *wingless* strongly supports the distinction between *P. t. theca* and *P. t. fukia*.

The 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 3). The average  $\Phi_{ST}$  value as 0.896 ( $p < 0.01$ ), suggesting significant genetic variation among the populations. Pairwise estimates of  $F_{ST}$  (0.885) and gene flow ( $Nm = 0.065$ ) between *P. t. theca* and *P. t. fukia* suggests that the subspecies in this species are highly differentiated.

**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 2). The Tajima's  $D$  and Fu's  $F$  values were non-significantly positive in





**Figure 6.** Bayesian Inference (BI) tree of mtDNA datasets for *Polytrema* using uncorrelated lognormal relaxed clock in BEAST v1.5.2, shows estimates of time since the most recent common ancestor for each major node and main mitochondrial clades. Posterior probabilities of nodes are shown below node branch when higher than 0.95. The violet bars indicate 95% highest posterior density interval (HPD) of the node ages.

	wing					genital
	Color of cilia of wings	Color of Underside ground	Number of spots in space Cu2 of the forewing	Color of scales scattered in costa and subapical area of forewing	Color of scales scattered in discal area and dorsum of hindwing	Ductus bursae
<i>P. t. theca</i>	brown	yellowish brown	0	greenish ochreous	greenish ochreous	thin
<i>P. t. fukia</i>	greyish white	greenish ochreous	1 or 2	greyish white	greyish white	thick

**Table 4.** Different morphological features of genitals and wings between *P. t. theca* and *P. t. fukia*.

*P. t. theca* group and all *P. theca* samples group. The mismatch analysis yielded a unimodal distribution of pairwise differences for *P. t. fukia* (Fig. 5B) compared to the multimodal distribution of *P. t. theca* samples (Fig. 5A) and the pooled samples (Fig. 5C). According to Rogers and Harpending<sup>31</sup>, the observed curves with unimodal representing population expansion and the observed curves with many peaks or resemblance to expected curves mean equilibrium population, which further elucidating the demographic history of *P. theca*. The results suggest that population expansion in *P. t. fukia* and population equilibrium in *P. t. theca*. Divergence time analysis with an uncorrelated lognormal relaxed clock run in BEAST produced a tree with a topology similar to ML tree (Fig. 6). *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 Fig. 6). *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 Fig. 6).

## Discussion

There is a small region of overlap in central Sichuan province in the distribution of *P. t. theca* and *P. t. fukia*, but otherwise they are not sympatric (Fig. 1). *P. t. theca* inhabits the higher elevations of west-central Sichuan Province and the Qinling Mountains<sup>16</sup>. *P. t. fukia* occurs in the whole southeastern of China, from northern Zhejiang Province, Jiangxi Province to east-central Yunnan Province<sup>9,32,33</sup>. According to the description on morphological variation between *P. t. theca* and *P. t. fukia*<sup>12</sup>, we found a different morphological feature existing in female genitalia except for the different color and spot number in some part of wings (Table 4). The ductus bursae of female *P. t. theca* is thinner than that of female *P. t. fukia*. We suspected the taxonomic status of the subspecies from their geographic separation and the morphological variation.

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 (Fig. 4A). The haplotype clades associated with both subspecies are deeply genetically divergent, differing from each other by  $5.07 \pm 0.49\%$  (4.3–5.9% divergence). This degree of divergence suggests that evolutionary separation of both subspecies occurred about 0.81 (HPD = 0.53–1.28) Mya, likely sometime during the Pleistocene based on a molecular clock calibration of 3.54% pairwise divergence per million years for a homologous mtDNA fragment in other insect species<sup>28</sup>. (fossil data and geographical calibration event are not available for this group). It is noteworthy for the nuclear *wingless* sequences that *P. t. theca* and *P. mencia* are considered distinct species with a genetic divergence of  $0.65 \pm 0.15\%$  while the *P. t. fukia* is currently considered a subspecies of the *P. t. theca* despite  $1.70 \pm 0.27\%$

sequence divergence. The phylogenetic of *wingless* gene indicates that the *P. theca* is paraphyly with three species sister to the clade of *P. t. theca* (Fig. 4B). 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 (Fig. 4A). Concordance between strongly differentiated mtDNA, nuclear haplotype clades and phenotypic variation supports the hypothesis that both subspecies of *P. theca* deserves recognition at the species-level under the general lineage concept of species<sup>4,34</sup>.

Two neutrality tests were chosen for the demographic history analysis, and significantly negative values of neutrality statistics can be indicative of background selection, but are also consistent with either population subdivision or expansion. Fu's  $F_s$  is significantly negative in *P. t. fukia* group reveal that the mtDNA of *P. t. fukia* appear to be not evolving neutrally. Additionally, we analysed the population size change of *P. t. theca*, *P. t. fukia* and all *P. theca*, respectively by the software DnaSP4.90<sup>23</sup> and got no evidence for population expansion in *P. t. theca* group<sup>31</sup> (Fig. 5A). However, we got unimodal curves representing population expansion in *P. t. fukia*<sup>31</sup> (Fig. 5B). The results suggest that population expansion in *P. t. fukia*. We still confirm the result of the population size change in haplotype network (Fig. 3). 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<sup>35</sup>. The haplotypes network of *P. t. fukia* displays a star-like pattern (Fig. 3). Haplotype I, the second most common and geographically widespread in central-west of China, is in the star's centre and derivatives are connected to it by short branches. Based on coalescence theory, the star-like topologies for this cluster strongly suggest the effect of a population expansion<sup>36</sup>. 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.

In conclusion, 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* Evans 1940 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<sup>37–40</sup>. Results from our study strengthen information about the *Polytremis* species complex and help in developing appropriate integrated pest management strategies for these insect pests.

## References

- Evans, W. H. In *A catalogue of the Hesperidae from Europe, Asia and Australia in the British Museum (Natural History)* (Trustees of the British Museum, London, 1949).
- ICZN International Commission on Zoological Nomenclature. In *International Code of Zoological Nomenclature* (4th ed, London, 1999).
- Futuyma, D. In *Evolutionary Biology* (3rd edn, Sinauer Associates, Massachusetts, 1997).
- Coyne, J. A. & Orr, H. A. In *Speciation* (Sinauer Associates, Sunderland, 2004).
- Hebert, P. D. N. *et al.* Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptus fulgerator*. *Proc. Natl. Acad. Sci. USA* **101**, 14812–14817 (2004).
- Chiba, H. & Eliot, J. N. A revision of the genus *Parnara* Moore (Lepidoptera, Hesperidae) with special reference to the Asian species. *Tyô to Ga*. **42**, 179–194 (1991).
- Devyatkin, A. L. & Monastyrskii, A. L. Hesperidae of Vietnam 12- a further contribution to the Hesperidae fauna of North and Central Vietnam. *Atalanta* **33**, 137–155 (2002).
- Guo, D., Fan, X. L. & Wang, M. Taxonomic Status of *Parnara batta* Evans (Lepidoptera: Hesperidae) Inferred from Mitochondrial COI Gene Sequences. *J. South China Agr. Univ.* **31**, 43–46 (2010).
- Chiba, H. *et al.* Jinhaku Sonan's skipper type collection deposited at Taiwan Agricultural Research Institute (Lepidoptera: Hesperidae). *Zootaxa* **2202**, 48–58 (2009).
- Hsu, Y. F. Notes on the Little Knows Females of Two Hesperiid Species. *Chinese J. Entomol.* **9**, 77–80 (1989).
- Huang, H. A list of butterflies collected from Nujiang and Dulongjiang, China with descriptions of new species, new subspecies and revisional notes. *Neue Entomologische Nachrichten* **55**, 3–114 (2003).
- Zhu, J. Q. In *A Taxonomic Study on the tribe Baorini Doherty, 1886 (Lepidoptera: Hesperidae: Hesperinae) of China* (Shanghai Normal University Press, Shanghai, 2012).
- Jiang, W. B. *et al.* Molecular phylogeny of the butterfly genus *Polytremis* (Hesperidae, Hesperinae, Baorini) in China. *PLoS ONE* **8**, e84098 (2013).
- Hebert, P. D. N., Ratnasingham, S. & Waard, J. R. Barcoding animal life: cytochrome c oxidase subunit I divergences among closely related species. *P. Roy. Soc. Lond. B Bio.* **270**, S96–S99 (2003).
- Eliot, J. N., Corbet, A. S. & Pendlebury, H. M. In *The butterflies of the Malay Peninsula* (4th edn, United Selangor Press, Kuala Lumpur, 1992).
- Chou, I. In *Monographia Rhopalocerorum Sinensium* (Henan Scientific and Technological Press, Zhengzhou, 1994).
- Rand, D. B., Heath, A., Suderman, T. & Pierce, N. E. Phylogeny and life history evolution of the genus *Chrysoritis* within the Aphnaeini (Lepidoptera: Lycaenidae), inferred from mitochondrial cytochrome oxidase I Sequences. *Mol. Phylogenet. Evol.* **17**, 85–96 (2000).
- Regier, J. C., Cook, C. P., Mitter, C. & Hussey, A. A phylogenetic study of the 'bombycoid complex' (Lepidoptera) using five protein-coding nuclear genes, with comments on the problem of macrolepidopteran phylogeny. *Syst. Entomol.* **33**, 175–189 (2013).
- Hall, T. A. BioEdit: a user-friendly biological sequence alignment editor an analysis program for Windows 95/98/NT. *Nucleic Acids symp.* **41**, 95–98 (1999).
- Guindon, S. & Gascuel, O. PhyML—A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52**, 696–704 (2003).
- Posada, D. & Crandall, K. A. MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**, 817–818 (1998).
- Bandelt, H. J., Foster, P. & Röhl, A. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* **16**, 37–48 (1999).
- Rozas, J., Sánchez-Delbarrio, J. C., Messeguer, X. & Rozas, R. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**, 2496–2497 (2003).

24. Excoffier, L., Laval, G. & Schneider, S. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evol. Bioinform.* **1**, 47–50 (2005).
25. Tajima, F. Statistical-method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–595 (1989).
26. Fu, Y. X. & Li, W. H. Statistical tests of neutrality of mutations. *Genetics* **133**, 693–709 (1993).
27. Drummond, A. J. & Rambaut, A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7**, 214 (2007).
28. Papadopoulou, A., Anastasiou, I. & Vogler, A. P. Revisiting the insect mitochondrial molecular clock: the mid-Aegean trench calibration. *Mol. Biol. Evol.* **27**, 1659–1672 (2010).
29. Rambaut, A. & Drummond, A. Tracer1.5. The University of Edinburgh, Edinburgh, Scotland, United Kingdom. URL <http://beastbioedacuk/> (2007).
30. Simon, C., Frati, F., Beckenbach, A., Crespi, B. & Liu, H. Evolution, Weighting, and Phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Ann. Entomol. Soc. Am.* **87**, 651–701 (1994).
31. Rogers, A. R. & Harpending, H. Population growth makes waves in the distribution of pairwise genetic differences. *Mol. Biol. Evol.* **9**, 552–569 (1992).
32. Xiong, H. L., Lin, C. & Chen, L. Danaidae and Hesperidae Resources in Maolan Nature Reserve and Their Faunal analysis. *Guizhou Agr. Sci.* **39**, 24–27 (2011).
33. Xu, G. Q. & Duan, H. S. Study on Resources of Butterfly in Hubei Province. *J. Jiangnan Univ.* **38**, 85–94 (2010).
34. de Queiroz, K. *In The general lineage concept of species and the defining properties of the species category* (MIT Press, Cambridge, 1999).
35. Narita, S., Nomura, M., Kato, Y. & Fukatsu, T. Genetic structure of sibling butterfly species affected by *Wolbachia* infection sweep: evolutionary and biogeographical implications. *Mol. Ecol.* **15**, 1095–1108 (2006).
36. Slatkin, M. & Hudson, R. R., Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics* **129**, 555–562 (1991).
37. Andres, C., Franke, F., Bleidorn, C., Bernhard, D. & Schlegel, M. Phylogenetic analysis of the *Lacerta agilis* subspecies complex. *Syst. Biodivers.* **12**, 43–54 (2014).
38. Evans, III E. L., Martinez-Gomez, J. E. & Sehgal, R. N. M. Phylogenetic relationships and taxonomic status of the endemic Socorro Warbler (*Setophaga pitiayumi graysoni*). *J. Ornithol.* **156**, 363–370 (2015).
39. Mulcahy, D. G. Phylogeography and species boundaries of the western North American Nightsnake (*Hypsiglena torquata*): Revisiting the subspecies concept. *Mol. Phylogenet. Evol.* **46**, 1095–1115 (2008).
40. Piffaretti, J. *et al.* Molecular phylogeny reveals the existence of two sibling species in the aphid pest *Brachycaudus helichrysi* (Hemiptera: Aphididae). *Zool. Scr.* **41**, 266–280 (2012).

## Acknowledgements

We thank for Huang Hao (Qingdao, China), Chen Zhibin (Shanghai Zoological Park, Shanghai, China), Tang Liang, Hu Jiayao, Yin Ziwei, Li Lizhen and their colleagues (Shanghai Normal University, Shanghai, China) for providing valuable comments and specimens throughout the project. This study was financially supported by grants from National Natural Science Foundation of China (No. 31401997), Innovation Program of Shanghai Municipal Education Commission–China (No. 14YZ068) and Innovation Program for University Student in Shanghai, China (No. A-9103-15-035088 and A-9103-15-035089).

## Author Contributions

W.J. and J.Z. initiated the project. W.J. performed bioinformatics analyses. W.J. and W.Y. designed experiments. H.H., Y.L., M.R., Y.M. and L.Z. carried out experiments. W.J. and J.Z. analysed data and interpreted results. W.J. wrote the manuscript.

## Additional Information

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Jiang, W. *et al.* Taxonomic status and molecular phylogeography of two sibling species of *Polytremis* (Lepidoptera: Hesperidae). *Sci. Rep.* **6**, 20820; doi: 10.1038/srep20820 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>