Inconsistent phylogenetic topologies reveal genomic admixture in *Kallima* butterflies (Lepidoptera: Nymphalidae)

Shuting WANG [1,2], Wei ZHANG[1,2,3①]

1. *State Key Laboratory of Protein and Plant Gene Research, Peking University, Beijing 100871, China*

2. *School of Life Sciences, Peking University, Beijing 100871, China*

3. *Peking-Tsinghua Center for Life Sciences, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing 100871, China*

Abstract: *Kallima* butterflies are famous for their leaf-mimicking wing patterns. Yet the characterization of *Kallima* species is still under debate owing to their high phenotypic similarity. With the release of the *K. inachus* reference genome, phylogenetic studies based on genome-wide data have been carried out, thus improving the understanding of the evolutionary relationships of the genus *Kallima*. However, we noticed that there is some conflict between genome-based phylogenies and morphological classifications in butterflies. We further examined the cause of this conflict by conducting an in-depth study of the relationships among *Kallima* butterflies to test possible reticulate phylogenetic topologies. We constructed phylogenies based on various datasets (including SNPs in single-copy genes, coding sequences, neutral regions and all remaining sites across the genome) to compare the topologies, revealing the complex evolutionary history of *Kallima* butterflies. Our results suggest that the reticulate species topology may constitute a pervasive pattern present not only in species with adaptive radiations but also in gradually evolving species, with *Kallima* butterflies as an example.

Key words: *Kallima*; phylogenetic analysis; reticulate topology

系统发生拓扑冲突揭示枯叶蛱蝶属基因组交融（鳞翅目：蛱蝶科）

王姝婷 [1,2]，张蔚 [1,2,3①]

1. 北京大学蛋白与植物基因研究国家重点实验室，北京 海淀 100871；2. 北京大学生命科学学院，北京 海淀 100871；3. 北京大学前沿交叉学科研究院北大-清华生命科学联合中心，北京 海淀 100871

摘要：枯叶蛱蝶属 *Kallima* Doubleday, 1849 蝶类以模仿枯叶的翅图案而闻名，由于该属物种间表型高度相似，属内物种鉴定仍存在争议。随着枯叶蛱蝶（*Kallima inachus*）参考基因组的发布，基于全基因组数据的系统发育研究得以开展，从而增进了对于枯叶蛱蝶属演化关系的了解。然而，基于基因组的系统发生关系与基于形态的物种分类之间存在一些冲突。通过深入研究枯叶蛱蝶属蝶类之间的关系，检验可能的网状系统发育拓扑结构，本研究进一步探究了产生冲突的原因。基于不同的数据集（包括单拷贝基因、编码序列、中性区域和所有分析物种中均存在的单核苷酸多态性位点）构建的系统发生树呈现出不一致的拓扑结构，揭示了枯叶蛱蝶属复杂的演化历史。本研究结果表明，网状物种拓扑可能是一种普遍存在的模式，不仅存在于适应性辐射的物种中，也存在于渐进式演化的物种中，如枯叶蛱蝶属蝶类。

Accepted 5 September 2024. Published online 22 November 2024.

① Correspondence author, E-mail: weizhangvv@pku.edu.cn

关键词：枯叶蛱蝶属；系统发育分析；网状拓扑

Introduction

Kallima butterflies are famous for their leaf-mimicking wing patterns, with multiple species distributed in southeastern Asia, the Himalayas, and southern China (Küppers 2015). Due to their high phenotypic similarity, the characterization of *Kallima* species is still under debate (Shirôzu & Nakanishi 1984; Nakamura & Wakahara 2013). For example, *Junonia cymodoce* was once categorized as a species of *Kallima* (Shirôzu & Nakanishi 1984), and the *Kallima* butterflies on Hainan Island were once regarded as a subspecies of *Kallima inachus*, i.e., *K. i. alicia.* However, in recent years it has been revised as an independent species, i.e., *Kallima alicia* (Nakamura & Wakahara 2013; Nakamura 2014). Great efforts have been made to identify *Kallima* butterflies based on morphological features, such as wing pattern and size and the degree of apex and tail length (Küppers 2015; Nakamura & Wakahara 2013; Shirôzu & Nakanishi 1984). The most important properties for determining species are those that characterize the external genitals from all possible perspectives, such as the ring, valva, dorsum, saccus, etc. (Shirôzu & Nakanishi 1984; Nakamura & Wakahara 2013). Even though photographs and detailed descriptions make comparisons possible, some qualitative descriptions may confuse researchers when a holotype is not accessible.

Since distinguishing *Kallima* butterflies based on morphological characteristics alone is difficult, phylogenetic analyses based on DNA sequence information provide additional information on the identity of *Kallima* species (Zhou *et al*. 2012; Wang *et al*. 2022); e.g., a phylogenetic study of *K. inachus* identified relatively ancient haplotypes in Hainan and Tibetan populations based on *COII* evidence (Zhou *et al*. 2012). With the release of the reference genome of *K. inachus* (Yang *et al*. 2020), phylogenetic studies based on genome-wide data have been carried out, thus improving the understanding of the evolutionary relationships of the genus *Kallima*. For example, phylogenetic results based on genome-wide single nucleotide polymorphism (SNP) data revealed that the genus *Kallima* contains two major clades, the *inachus–alicia* clade and the *limborgii–paralekta–incognita* clade, with species and subspecies further determined within each clade (Wang *et al*. 2022). Notably, some of these genome-based analyses support previous morphological classifications, while others do not. For example, the current taxonomic status of *K. alicia* is consistent with the previous morphological classification (Nakamura 2014), whereas *Kallima incognita* used to be regarded as a subspecies of *Kallima limborgii* (Nakamura & Wakahara 2013). This conflict drew our attention since it may be caused by a lack of resolution due to morphological similarities or existing complex species relationships. Related to our concerns, recent studies have shown that even based on genome-wide data, for some species with reticulate relationships, a single bifurcating phylogeny, such as that of *Heliconius* butterflies, which have undergone adaptive radiation, has unresolved relationships with such reticulated species (Edelman *et al*. 2019).

Therefore, we conducted an in-depth study of the relationships among *Kallima* butterflies to test possible reticulate phylogenetic topologies. We constructed phylogenies based on various datasets (including SNPs in single-copy genes, coding sequences, neutral regions and all remaining sites across the genome) to compare the topologies, revealing the

complex evolutionary history of *Kallima* butterflies. We further characterized the major lineages leading to the complex topology using PhyloNet (Wen *et al*. 2018), ASTRAL (Zhang *et al*. 2018) and MP-EST (Liu *et al*. 2010). Our results suggest that reticulate species relationships may be common not only among species subject to radiative evolution but also among species subject to Darwinian gradualism, e.g., the genus *Kallima*.

Material and methods

Sample information

In this study, we analyzed whole-genome resequencing data from 15 samples from the genus *Kallima* and one sample from *Doleschallia bisaltide*. Two *Kallima* samples (VN01 and VN05) from Vietnam were obtained from the NCBI SRA database (PRJNA802711), while the rest were downloaded from the NCBI SRA database (PRJNA698415). The statistics and sampling information for these samples are shown in Table 1.

Table 1. Sample information and sequencing statistics

ID	Species	Sex	Location	Data (Gb)	Genotype calls (Qual>50)	Mean depth (Qual>50)	Alignment rate
dbiY	*Doleshcallia bisaltide*	NA	China: Yunnan	9.87	10,323,160	14.69	22.17%
FJ	*K. inachus chinensis*	male	China: Nanping	9.20	104,598,506	12.26	96.76%
MI-13	*K. inachus chinensis*	female	China: Ya'an	16.16	101,481,944	27.12	96.75%
Y79	*K. inachus chinensis*	Male	China: Leshan	14.53	100,664,164	24.37	95.91%
kinM	*K. inachus chinensis*	female	China: Yunnan, Kunming	14.81	105,906,790	18.82	96.62%
TW2	*K. inachus formosana*	male	China: Taiwan, Taizhong	13.11	102,212,717	17.61	96.22%
VN01	*K. inachus alboianchus*	male	Vietnam: Đồng Nai, Thác Mai	19.22	102,518,700	11.94	95.55%
VN05	*K. alicia kishii*	female	Vietnam: Lâm Đồng, Bảo Lộc	17.22	98,370,999	11.38	93.86%
GX	*K. alicia alicia*	male	China: Guangxi, Laibin	8.45	95,779,530	10.11	89.29%
Hn1	*K. alicia alicia*	male	China: Hainan, Wuzhishan	13.43	99,487,892	16.20	95.95%
mt4	*K. alicia shizuyai*	male	China: Tibet, Medog	11.67	91,792,358	14.79	89.53%
kliam62	*K. limborgii amplirufa*	female	Malaysia	11.69	91,792,358	14.79	92.79%
kpa57	*K. paralekta*	female	Indonesia: Java	9.22	90,277,857	13.27	91.51%
mt1	*K. incognita*	male	China: Tibet, Medog	12.05	96,368,615	14.09	91.97%
YN	*K. incognita*	male	China: Yunnan, Ruili	13.17	93,992,666	15.98	92.10%
LD	*K. knyvettii*	male	China: Tibet, Medog	13.23	83,977,909	15.30	86.10%

Data processing

We filtered low-quality data from the raw reads using Trimmomatic v0.38 (Bolger *et al*. 2014) and mapped the filtered reads to the *K. inachus* reference genome (Yang *et al*. 2020) using Bowtie2-2.3.4.3 (Langmead & Salzberg 2012) with the parameter "--very-sensitive-local". BAM files were generated from SAM files using SAMtools v1.9 (Li 2011) and further re-ordered, sorted, and duplicates marked using Picard tools v1.96 (http://broadinstitute.github.io/picard/). RealignerTargetCreator and IndelRealigner in GATK 3.7 (McKenna *et al*. 2010) were used to realign indels, and UnifiedGenotyper was used to call

SNPs with the following parameters: heterozygosity 0.05, stand_call_conf 50.0, stand_emit_conf 10.0, and dcov 250. For downstream analysis, we filtered VCF files using VCFtools 0.1.15 (Danecek *et al*. 2011) with the parameters "maf" 0.05 and "Qual" 50.

Phylogenetic analysis

VCF files were recoded according to these positions across the genome. BUSCO 3.0.2 (Waterhouse *et al*. 2018) was used for identifying single-copy genes in the *K. inachus* genome (Yang *et al*. 2020) with the lepidopteran database "lepidoptera_odb10". SNP sites within single-copy genes and CDS regions were extracted from the VCF file using the "--bed" flag of VCFtools 0.1.15 (Danecek *et al*. 2011). The neutral regions were identified according to Gronau *et al*. (2011). Filters were applied to the genome of *K. inachus* (Yang *et al*. 2020), including selection of scaffold sizes above 50 kb, masking of repetitive elements using Tandem Repeats Finder and RepeatMasker, exclusion of conserved noncoding elements and 100 bp neighboring regions by blasting against UCSC phastCons elements in the 27-way alignment for *Drosophila melanogaster*, exclusion of exons and 10 kb neighboring regions based on the genome annotation of *K. inachus*, exclusion of missing calls and calls with read depths greater than twice the average depth or less than half the average depth, and selection of 1 kb blocks at least 50 kb apart, which yielded 5,638 putative neutral loci in total (Wang *et al.* 2022). To filter all missing sites, we used the flag "--max-missing 1" with VCFtools 0.1.15 (Danecek *et al*. 2011). The newly recoded VCF files were converted to PHYLIP files using the vcf2phylip-master package (https://github.com/edgardomortiz/vcf2phylip). RAxML 8.2.12 (Stamatakis 2014) was used for maximum-likelihood phylogenetic tree construction with the GTRGAMMA model and 100 bootstrap replicates. Trees were visualized and annotated using ITOL v5 (Letunic & Bork 2021).

PhyloNet analysis

PhyloNet 3.8.2.9 (Wen *et al*. 2018) was used to explore the network of relationships in the genus *Kallima*. 150 single-copy genes were randomly selected, and the sequences of the genes were used as input for PhyloNet analysis, in MCMC_SEQ mode. The iteration chain length was 20,000,000, while the number of iteration burn-in was 2,000,000. Summarized results were generated using SummarizeMCMCResults in PhyloNet and further analyzed using Tracer 1.6 (http://beast.bio.ed.ac.uk/Tracer). The network was visualized using Dendroscope 3 (Huson & Scornavacca 2012).

ASTRAL and MP-EST analyses

ASTRAL 5.7.8 (Zhang *et al*. 2018) and MP-EST v3.0 (Liu *et al*. 2010) were used to construct phylogenetic trees based on the coalescence model. For each single-copy gene, a maximum-likelihood phylogenetic tree was constructed using RAxML 8.2.12 (Stamatakis 2014), while *D. bisaltide* was set as an outgroup in the rooted phylogeny. Gene trees were concatenated into one input file for downstream analyses. Output trees were viewed using Figtree (http://tree.bio.ed.ac.uk/software/Figtree/).

Results

We first annotated a total of 5,097 single-copy genes using BUSCO (Waterhouse *et al*.

2018), which refers to the conserved genes in Lepidoptera and contains approximately 212 Mb of SNPs; then, we constructed a maximum-likelihood phylogenetic tree based on these SNPs (Fig. 1A) and compared it with a previously reported species tree based on genome-wide SNPs (Wang *et al*. 2022), which resulted in a different topology. According to the previously reported species tree, *K. inachus* and *K. alicia* formed a clade, *Kallima paralekta* and *K. incognita* formed another clade and further clustered with *K. limborgii*, and *Kallima knyvettii* clustered with the latter clade (Wang *et al*. 2022). However, the tree based on single-copy gene SNPs showed a different topology: only the clade of *paralekta–incognita* was retained; *K. alicia* was clustered with *K. knyvettii*; and *K. inachus* formed a clade with *K. limborgii*, which in turn clustered with the *paralekta–incognita* clade. Within each clade, the relative positions of the population/subspecies samples within each species remained stable, which was consistent with previously reported results.

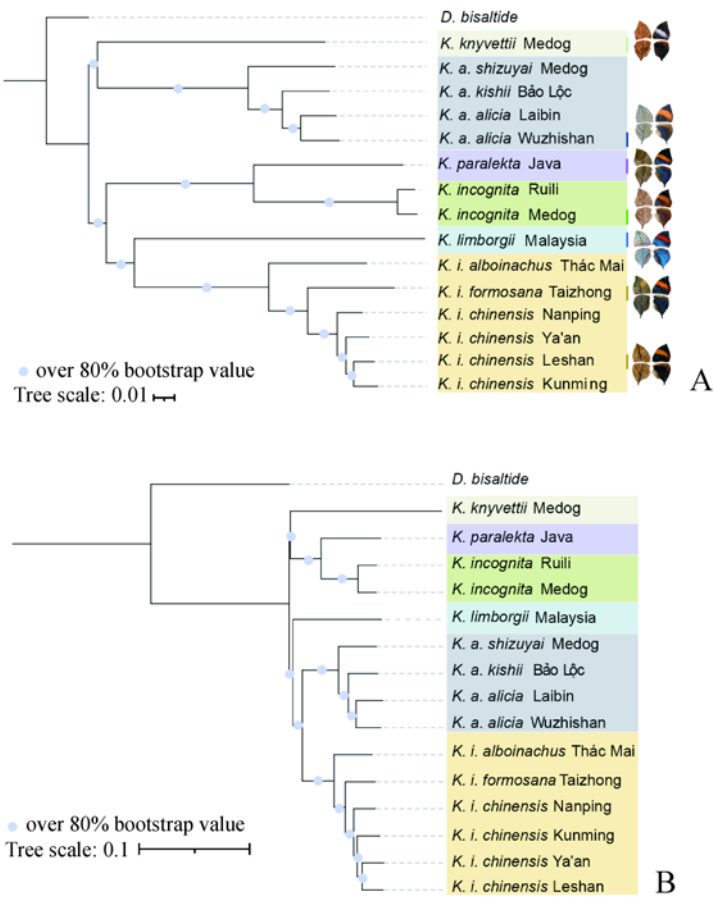

Figure 1. Maximum-likelihood phylogenetic trees constructed based on SNPs in the gene regions. A. Based on SNPs in single-copy genes; B. Based on SNPs in the coding regions.

Considering that single-copy genes include intron regions that may differ between species, we further constructed another phylogenetic tree based on approximately 0.25 Mb of SNPs in the coding sequences (CDSs) of all the genes across the genome (Fig. 1B). The

topology of this phylogeny was more similar to that of the previously reported species tree based on genome-wide SNPs (Wang *et al*. 2022), with the main difference being that *K. limborgii* was included in the *inachus–alicia* clade. This result supported our hypothesis that intronic regions may contribute more variation relative to the CDS and reflect the complex evolutionary history of the genus *Kallima*.

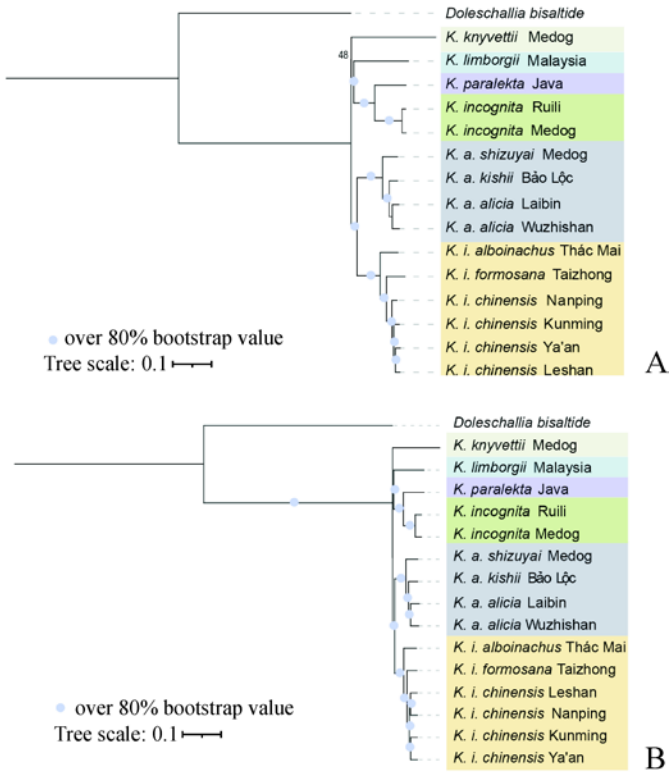

Figure 2. Maximum-likelihood phylogenetic trees constructed based on genome-wide SNPs. A. Based on SNPs in the neutral regions; B. Based on SNPs with missing data removed.

To further explore the evolutionary relationships of the genus *Kallima*, we considered identifying neutral regions (i.e., regions unaffected by selection and mainly distributed throughout the genome in intergenic regions (Gronau *et al*. 2011)) to continue the study because these regions may be conserved and show less variation due to strong selective pressure. The phylogenetic tree constructed based on neutral regions showed the same topology for two major clades (*inachus–alicia* and *paralekta–incognita–limborgii*) as the previously reported species tree based on genome-wide SNPs (Wang *et al*. 2022), differing only in the position of *K. knyvettii*, which appeared as the basal lineage of the genus *Kallima* in the tree based on neutral regions (Fig. 2A). This result suggested that *K. knyvettii* may contain more distinct variations from the other taxa, which was also supported by its alignment rate (86.1%), which was the lowest relative to those of the other *Kallima* species (Table 1).

We also considered the effect that the low alignment rate for the outgroup taxon *D. bisaltide* to the *K. inachus* reference genome was only 22.17% (Table 1), which may have led

to a large amount of missing data and biased the results. To address this issue, we excluded the effect of missing data by using a dataset of all the remaining SNPs (approximately 0.11 Mb) to construct a phylogeny, which also considered both coding sequences and noncoding sequences (Fig. 2B). The topology of this phylogeny was identical to that of the phylogenetic tree based on genome-wide SNP data (Wang *et al*. 2022), with two major clades of the *Kallima* genus, *knyvettii–limborgii–paralekta–incognita* and *alicia–inachus*.

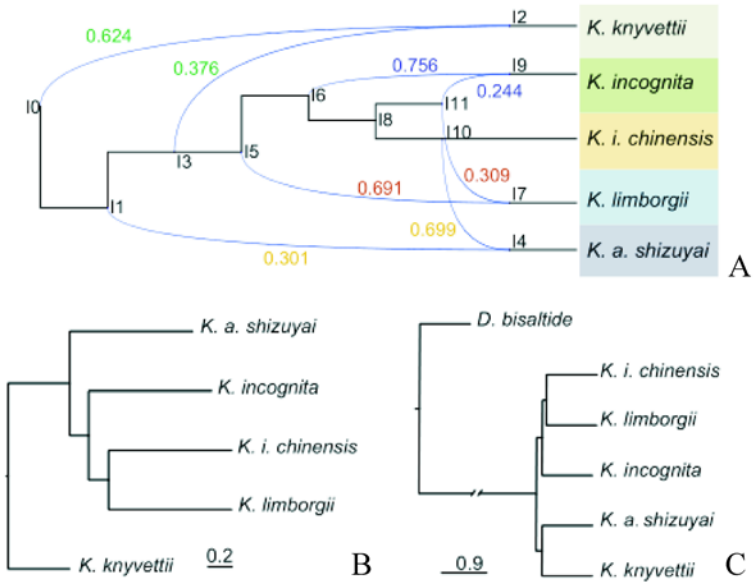

Figure 3. Phylogenetic estimates using multiple methods. A. The phylogenetic network was inferred using PhyloNet, which shows a reticulate evolutionary history of the genus *Kallima*. The blue lines indicate possible hybridization or incomplete lineage sorting, and the numbers on the side indicate inheritance probability; B. The species tree inferred using ASTRAL; C. The species tree inferred using MP-EST.

The above-mentioned results support the topology of the previously reported species tree but cannot rule out a potentially discordant evolutionary history in *Kallima* butterflies. Given that each gene may have a different evolutionary history, we therefore used randomly selected genes and integrated the information obtained from individual genes using PhyloNet (Wen *et al*. 2018). The results based on 150 randomly-selected single-copy genes showed reticulate evolutionary relationships among *Kallima* butterflies (Fig. 3A), with multiple ancestral lineages showing mixed genetic background. Since the possible reticulations largely involved ancestral lineages, we propose there might be a complex evolutionary history of *Kallima* butterflies in their early speciation process, which echoes our phylogenetic results based on a variety of datasets.

In addition to constructing phylogenetic trees using concatenated sequences, we also attempted to estimate phylogenetic trees based on the coalescence model using ASTRAL (Zhang *et al*. 2018) and MP-EST (Liu *et al*. 2010). MP-EST used rooted gene trees as input, whereas ASTRAL estimated phylogenies using unrooted trees. The result of ASTRAL analysis was therefore re-rooted using *K. knyvettii* as an outgroup (Fig. 3B). However, the proportion of input gene tree quartet trees satisfied by the species tree inferred using ASTRAL

was approximately 0.42, indicating strong discordance likely derived from incomplete lineage sorting. The topologies based on the coalescence model are consistent with each other (Figs 3B, 3C), showing major clades with similar topologies to that obtained from the maximum likelihood tree based on SNPs of the single-copy genes (Fig. 1A).

Discussion

In this study, we applied multiple genomic datasets to estimate the robustness of the phylogenetic topology of the *Kallima* genus, yielding results for the major clades in general agreement with the morphological taxonomy. With respect to the species *K. incognita*, which was revised to be an independent species in a previous study (Wang *et al*. 2022), we further demonstrated that *K. incognita* is more closely related to *K. paralekta* than to *K. limborgii* (Nakamura & Wakahara 2013). However, our sampling site was not the same as that used for the holotype of *K. l. incognita* (Nakamura & Wakahara 2013), and we were unable to obtain genomic DNA from that sample for additional validation. Thus, the relationships between the samples used for morphological and genomic analyses need to be further determined. We therefore suggest that genomic data may be necessary to compensate for the morphological data to determine the interspecific relationships among *Kallima* butterflies, as well as among other butterflies with similar wing patterns but lacking evidence of external genitals.

To gain deeper insight into the interspecific relationship, a phylogeny based on a single genomic dataset is informative but may not be sufficient because of the large amount of data and the characteristics of the datasets. For example, in this study, we constructed phylogenetic trees based on four genomic datasets that exhibited different topologies and were distinct from the previous species tree based on genome-wide SNP data (Wang *et al*. 2022). According to our results on *Kallima* butterflies, their interspecific relationship was well supported not only by the phylogenetic tree based on genome-wide SNP data, the fixation index, and PCA results but also by the unrooted tree based on coding sequences without the outgroup taxon (Wang *et al*. 2022). Our results suggested that multiple analyses can be integrated to reveal the real relationships among species.

However, despite determining the bifurcating species tree for the genus *Kallima*, we identified their underlying reticulated topologies, each of which may represent a unique evolutionary history for particular lineages. At the level of the genus *Kallima*, the results of phylogenic analyses based on multiple methods indicated that the intensity of genomic admixture is moderate, likely due to introgressive hybridization or incomplete lineage sorting, which requires further characterization through the incorporation of additional ingroup taxa. At the species level, comparisons between the neutral regions and SNP sites retained in all the species revealed variable phylogenetic positions of *K. knyvettii*, indicating its complex evolutionary history. These results raise new questions about the relationships and evolution of *Kallima* butterflies, which may be more complex than we previously thought; i.e., the reticulate species topology may be pervasive not only in species with adaptive radiations but also in gradually evolving species.

Acknowledgements

This project was supported by grants from the National Natural Science Foundation of China (32325009; 32170420), the Peking-Tsinghua Center for Life Sciences, and the State Key Laboratory of Protein and Plant Gene Research to WZ and grants from the China Postdoctoral Science Foundation (2023M730082; BX20230026) to SW.

Author contributions

WZ designed the study; SW performed the data analyses; and they together wrote, proofread and approved the manuscript.

References

Bolger AM, Lohse M & Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30: 2114–2120.

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G & Durbin R. 2011. The variant call format and VCFtools. *Bioinformatics*, 27: 2156–2158.

Edelman NB, Frandsen PB, Miyagi M, Clavijo B, Davey J, Dikow RB, García-Accinelli G, Van Belleghem SM, Patterson N, Neafsey DE, Challis R, Kumar S, Gilson RPM, Salazar C, Counterman BA, Papa R, Blaxter M, Reed RD, Dasmahapatra KK, Kronforst M, Joron M, Jiggins CD, McMillan WO, DiPalma F, Blumberg AJ, Wakeley J, Jaffe D & Mallet J. 2019. Genomic architecture and introgression shape a butterfly radiation. *Science*, 366: 594–599.

Gronau I, Hubisz MJ, Gulko B, Danko CG & Siepel A. 2011. Bayesian inference of ancient human demography from individual genome sequences. *Nature Genetics*, 43: 1031–1035.

Huson DH & Scornavacca C. 2012. Dendroscope 3: An interactive tool for rooted phylogenetic trees and networks. *Systematic Biology*, 61: 1061–1067.

Küppers PV. 2015. *Butterflies of the World*, *Part 44: Nymphalidae XXV: Kallima*. Verlag Goecke & Evers, Germany, 8 pp.

Langmead B & Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9: 357–359.

Letunic I & Bork P. 2021. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Research*, 49: W293–W296.

Li H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, 27: 2987–2993.

Liu L, Yu L & Edwards SV. 2010. A maximum pseudo-likelihood approach for estimating species trees under the coalescent model. *BMC Evolutionary Biology*, 10: 302.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M & DePristo MA. 2010. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, 20: 1297–1303.

Nakamura N. 2014. Distribution of *Kallima inachus* (Doyère 1840) and related species (Lepidoptera, Nymphalidae) in Indochina and adjacent regions with status alteration of *Kallima inachus alicia* Joicey & Talbot, 1921. *Butterflies* (*Teinopalpus*), 66: 22–39.

Nakamura N & Wakahara H. 2013. Notes on the butterflies of Laos (VIII): review of the genus *Kallima*

Doubleday, [1849] (Lepidoptera, Nymphalidae) with descriptions of two new subspecies. *Butterflies* (*Teinopalpus*), 63: 14–28.

Shirôzu T & Nakanishi A. 1984. A revision of the genus *Kallima* Doubleday (Lepidoptera, Nymphalidae): I. Generic Classification. *Lepidoptera Science*, 34: 97–110.

Stamatakis A. 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30: 1312–1313.

Wang S, Teng D, Li X, Yang P, Da W, Zhang YM, Zhang YB, Liu G, Zhang X, Wan W, Dong ZW, Wang DH, Huang S, Jiang ZS, Wang QY, Lohman DJ, Wu YJ, Zhang LL, Jia FH, Westerman E, Zhang L, Wang W & Zhang W. 2022. The evolution and diversification of oakleaf butterflies. *Cell*, 185: 3138–3152.e20.

Waterhouse RM, Seppey M, Simão FA, Manni M, Ioannidis P, Klioutchnikov G, Kriventseva EV & Zdobnov EM. 2018. BUSCO Applications from quality assessments to gene prediction and phylogenomics. *Molecular Biology Evolution*, 35: 543–548.

Wen D, Yu Y, Zhu J & Nakhleh L. 2018. Inferring phylogenetic networks using PhyloNet. *Systematic Biology*, 67: 735–740.

Yang J, Wan W, Xie M, Mao J, Dong Z, Lu S, He J, Xie F, Liu G, Dai X, Chang Z, Zhao RP, Zhang R, Wang ST, Zhang YM, Zhang W, Wang W & Li XY. 2020. Chromosome‐level reference genome assembly and gene editing of the dead-leaf butterfly *Kallima inachus*. *Molecular Ecology Resources*, 20: 1080–1092.

Zhang C, Rabiee M, Sayyari E & Mirarab S. 2018. ASTRAL-III: Polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics*, 19(S6): 153.

Zhou C, Chen X & He R. 2012. *COII* phylogeography reveals surprising divergencies within the cryptic butterfly *Kallima inachus* (Doyère 1840) (Lepidoptera: Nymphalidae: Kallimini) in southeastern Asia. *Pan-Pacific Entomologist*, 88: 381–398.