

Citation: Spitsyn VM, Kondakov AV, Bolotov NI, Thi Pham N, Gofarov MY, Bolotov IN (2018) DNA barcoding unravels contrasting evolutionary history of two widespread Asian tiger moth species during the Late Pleistocene. PLoS ONE 13(4): e0194200. https://doi.org/10.1371/journal.pone.0194200

Editor: Ben J Mans, Onderstepoort Veterinary Institute, SOUTH AFRICA

Received: October 6, 2017

Accepted: February 27, 2018

Published: April 4, 2018

Copyright: © 2018 Spitsyn et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: The field works and sampling were supported by the National Geographic Society (USA, grant no. 9821-15) (http://www. nationalgeographic.com). The molecular analyses were supported by the Federal Agency for Scientific Organizations (Russia, project no. 0409-2015-0143) (http://fano.gov.ru/en/). **RESEARCH ARTICLE**

DNA barcoding unravels contrasting evolutionary history of two widespread Asian tiger moth species during the Late Pleistocene

Vitaly M. Spitsyn^{1,2}*, Alexander V. Kondakov^{1,2}, Nikita I. Bolotov², Nhi Thi Pham^{3,4}, Mikhail Y. Gofarov^{1,2}, Ivan N. Bolotov^{1,2}

1 Lab for Molecular Ecology and Phylogenetics, Northern Arctic Federal University, Arkhangelsk, Russian Federation, 2 Institute of Biogeography and Genetic Resources, Federal Center for Integrated Arctic Research, Russian Academy of Sciences, Arkhangelsk, Russian Federation, 3 Institute of Ecology and Biological Resources, Vietnam Academy of Science and Technology, Hanoi, Vietnam, 4 Graduate University of Science and Technology, Vietnam Academy of Science and Technology, Hanoi, Vietnam

* vitalik91993@yandex.ru

Abstract

Populations of widespread pest insects in tropical areas are characterized by a complex evolutionary history, with overlapping natural and human-mediated dispersal events, sudden expansions, and bottlenecks. Here, we provide biogeographic reconstructions for two widespread pest species in the tiger moth genus Creatonotos (Lepidoptera: Erebidae: Arctiinae) based on the mitochondrial cytochrome c oxidase subunit I (COI) gene. The Asian Creatonotos transiens reveals shallow genetic divergence between distant populations that does not support its current intraspecific systematics with several local subspecies. In contrast, the more widespread Creatonotos gangis comprises at least three divergent subclades corresponding to certain geographic areas, i.e. Australia, Arabia + South Asia and Southeast Asia. With respect to our approximate Bayesian computation (ABC) model, the expansion of Creatonotos gangis into Australia is placed in the Late Pleistocene (~65-63 ka). This dating coincide with an approximate time of the earliest human migration into the continent (~65–54 ka) and the period of intervisibility between Timor and Australia (~65–62 ka). Our findings highlight that the drying Sunda and Sahul shelf areas likely support successful migrations of Asian taxa into Australia during the Pleistocene. The phylogeographic patterns discovered in this study can be used to improve the effectiveness of integrated pest control programs that is a task of substantial practical importance to a broad range of agricultural stakeholders.

Introduction

Although the history of faunal exchange between mainland Southeast Asia and Australia attracts an exceptional attention of scientists since Alfred Russel Wallace [1], the patterns of this process are not entirely known. Typically, faunas of these regions are distinct, with intermediate forms ranging across New Guinea, Wallacea, Sundaland and the Philippines



Competing interests: The authors have declared that no competing interests exist.

[2,3,4]. The molecular-based investigations revealed a higher level of endemism than previously recognized [3,4], e.g., 90% of Australian hawkmoths are endemic to the continent or to Australia, the Pacific Islands and the Papuan and Wallacean regions [4]. In contrast, several moth and butterfly species do not reveal deep genetic divergence between Australian and Southeast Asian populations, which indicates the possibility of long-distance dispersal events in these taxa during the Late Pleistocene [5,6,7,8]. The low sea levels through much of the Pleistocene support the exchange between the two continents via drying shelf areas creating Sundaland (mainland Southeast Asia with the Malay Peninsula and the Greater Sunda Archipelago) and Sahul (mainland Australia, Papua New Guinea and Tasmania) [9].

In the present study, we report the results of extensive phylogeographic analyses of *Creato*notos gangis (Linnaeus, 1763) and C. transiens (Walker, 1855), two widespread pest tiger moth species (Erebidae: Arctiinae: Arctiini). The first species, Creatonotos gangis, is widely distributed in tropical and subtropical areas of the Old World, including the Middle East, South Asia, China, mainland Southeast Asia, Indonesia, the Philippines, New Guinea and Australia [10,11,12,13]. With respect to the morphology-based taxonomy, it appears to be a complex of closely related taxa ranged from the Middle East to Northern Australia [14]. Several species and subspecies-level taxa were described in this species complex [14,15]. The second species, Creatonotos transiens, also has a broad range, situated within Southeast Asia [12]. Both Creatonotos gangis and C. transiens are typical generalist species, which may inhabit a variety of environments, including agricultural lands and even urban areas (e.g., Delhi) [16,17]. The extreme polyphagy of larvae could explain a broad ecological niche in both species [18,19,20,21]. Interestingly, Creatonotos gangis appears to be highly adapted to rice, because the highest percentage of larval survival to pupation and the shortest larval development period occurred on this plant [21], hence it is an abundant species on rice fields [22]. This tiger moth species is also considered a possible pest of economically important plants such as sweet potato, turmeric, tea, soybean and maize [23,24]. Additionally, Creatonotos gangis may cause an extensive defoliation of pomegranate trees [25]. Creatonotos transiens is also a pest of many agricultural and forest plants, including the dragon tree, Paulownia spp. [26].

These moths are characterized by the dual mating system (both males and females produce pheromone) and the gigantic scent-bearing male coremata (androconia) [27,28]. The male coremata are pneumatically eversible organs composed of two pairs of tubes, up to 37 mm long, each covered by approximately 3000 long scales, with a giant epidermal (trichogen) gland cell at the base of each scale [27]. 7-Hydroxy-6,7-dihydro-5H-pyrrolizine-1-carboxaldehyde is the major volatile component of the coremata [29]. The size of coremata and the pheromone biosynthesis are directly controlled by the quantity of hostplant-derived pyrrolizidine alkaloids ingested by the larvae [29,30,18]. Despite the fact that *Creatonotos* species are both economically and scientifically important, their phylogeo-graphic patterns are almost unknown, although several broad phylogenetic reconstructions contain sequences of certain *Creatonotos* species [31,32,33]. De Freina [34] provides a taxonomic revision of several African and Arabian *Creatonotos* taxa by means of a molecular approach.

Taking into account the above-mentioned broad ranges of our studied species, we assumed that they likely represent species complexes, each of which contains several cryptic species-level taxa. However, we found that the extensive distribution ranges of *Creatonotos gangis* and *C. transiens* do not reflect an ancient diversification pattern, and they could have been formed via multiple expansion events during the Late Pleistocene.

Materials and methods

Data collection, DNA extraction, PCR and sequencing

The barcode region of the fast evolving mitochondrial protein-coding cytochrome c oxidase subunit I (COI) gene was used in the present investigation. The sequence data set that combine our materials and published data comprises a total of 109 Creatonotos spp. sequences (S1 Table). For molecular analyses, we used 21 specimens of Creatonotos gangis and 20 specimens of C. transiens from the collection of the Russian Museum of Biodiversity Hotspots (RMBH), Federal Center for Integrated Arctic Research of the Russian Academy of Sciences, Arkhangelsk, Russian Federation. The total DNA was extracted from a single leg of each dry specimen using a standard phenol/chloroform procedure [35]. The standard primers C1-J-1718 and C1-N-2329R were used for the amplification of 612-bp-long barcode fragments of the COI gene [36]. The PCR mix contained approximately 200 ng of total cellular DNA, 10 pmol of each primer, 200 µmol of each dNTP, 2.5 µl of PCR buffer (with 10×2 mmol MgCl₂), 0.8 units Taq DNA polymerase (SibEnzyme Ltd., Novosibirsk, Russia), and H₂O added for a final volume of 25 µl. Thermocycling included one cycle at 95°C (4 min), followed by 36–38 cycles of 95°C (50 sec), 50°C (50 sec), and 72°C (50 sec) and a final extension at 72°C (5 min). Forward and reverse sequencing was performed on an automatic sequencer (ABI PRISM[®] 3730, Applied Biosystems) using an ABI PRISM[®] BigDye[™] Terminator v. 3.1 reagent kit. Forty-one new COI sequences were obtained in this study, which were deposited in NCBI GenBank. Additionally, the published COI sequences of Creatonotos spp. were supplied by NCBI GenBank and BOLD IDS. All sequences were checked manually using a sequence alignment editor (BioEdit version 7.2.5 [37]). The alignment of sequences, which contained no indels, was performed using the ClustalW algorithm implemented in MEGA6 [38].

Phylogenetic, phylogeographic and species delimitation analyses

For phylogenetic analyses, the sequence data set of *Creatonotos* spp. was collapsed into 61 unique *COI* haplotypes (648 bp in length) using an online FASTA sequence toolbox, FaBox v. 1.41 [39]. As an out-group, two additional *COI* haplotypes of *Arctia menetriesii* (Eversmann, 1846) and *A. tundrana* (Tshistjakov, 1990) were used (GenBank acc. nos. KM111174 and KM111175, respectively). The lacking sites were treated as missing data. The best models of sequence evolution as suggested the corrected Akaike Information Criterion (AICc) of MEGA6 [38] were as follows: 1st codon of *COI*: TN93, 2nd codon of *COI*: GTR, and 3rd codon of *COI*: TN93+G (G = 1.1). Phylogenetic relationships were reconstructed based on Bayesian inference implemented in MrBayes v. 3.2.6 [40]. The analyses were performed using the following parameters: nchains = 4, nruns = 2, samplefreq = 1000, temp = 0.1; 10% of the sampled trees were discarded as burn-in (pre-convergence part). Runs were conducted for 25 million generations. Convergence of the MCMC chains to the stationary distribution was checked visually based on the plotted posterior estimates using a MCMC trace analysis tool (Tracer v1.6 [41]). Calculations were performed at the San Diego Supercomputer Center through the CIPRES Science Gateway [42].

The phylogeographic analyses of *Creatonotos* spp. sequences were performed based on a median joining network approach using Network v. 5.0.0.1 software with default settings [43]. The data set of 99 sequences across the extensive area from the Arabian Peninsula to Northern Australia was used (S1 Table). To remove missing sites due to different length of available sequences, all sequences were cut in accordance with the minimal sequence length

leaving the data set of 456 bp long. Several shorter sequences were removed from the data set.

To test the hypothesis that *Creatonotos gangis* and *C. transiens* may represent complexes of closely related species-level taxa [14,15,34], we applied a species delimitation approach by using the Bayesian Poisson Tree Process (PTP) model [44]. Using the fifty-percent majority-rule Bayesian consensus tree (see above), two runs of 100,000 generations were executed with the first 10% discarded as a burn-in based on the convergence of log-likelihood values (S1 Fig). The two out-group taxa were retained for the analysis, because the number of in-group taxa was low that may biased the delimitation results. Additionally, we performed a standard barcoding gap analysis based on uncorrected *COI p*-distances [45,46] in order to supplement the PTP modeling results.

Population genetic analysis and approximate Bayesian computation

Population genetic diversity indexes (haplotype and nucleotide diversity), Tajima's *D*-test and Fu's *F*-test statistics and the mismatch distribution under a spatial expansion model were calculated using Arlequin v. 3.5.1.2 software to estimate the demographic histories of sampled populations [47,48]. A total sequence sample was subdivided in accordance with geographic areas and taxa: (i) *Creatonotos gangis*, Arabia and South Asia (N = 20); (ii) *C. gangis*, mainland Southeast Asia (N = 20); (iii) *C. gangis*, Eurasia (N = 40); (iv) *C. gangis*, Australia (N = 15); (vi) *C. gangis*, all available sequences (Eurasia + Lesser Sundas + Australia, N = 57); and (vii) *C. transiens*, Eurasia (N = 39). Nucleotide differences (uncorrected *p*-distance, %) between sequence groups with standard error values based on 1000 bootstrap replications were calculated in MEGA6 [38].

The time since population expansion (*t*, generations) was calculated using the equation:

t

$$=\tau/2\mu;$$
 (1)

where τ is a moment estimator that represents a unit of mutational time, inferred from the mode of mismatch distribution, and μ is a mutation rate (%) per site per generation [7,49,50,51]. The most appropriate external *COI* mutation rate of 1.77×10^{-8} substitutions/site/ year was applied, which is based on the mid-Aegean trench calibration using six genera of darkling beetles (Coleoptera: Tenebrionidae) [52].

Finally, the two integrative biogeographic scenarios concerning the possible origin of *Creatonotos gangis* populations were designed (S5 Table). Both scenarios assume that Arabian–South Asian and Australian populations independently derived from Southeast Asian population, but the scenario 1-CG suggests more ancient divergence times (S5 Table). These demographic scenarios were simulated for comparison using an approximate Bayesian computation (ABC) approach with DIYABC v. 2.1.0 software [53,54]. The primary sequence data set included three samples of *Creatonotos gangis* sequences: (i) Southeast Asia (N = 20); (ii) Arabia and South Asia (N = 20); and (iii) Australia (N = 15). Prior settings of the ABC analyses are presented in S5 Table. The HKY+G model was likely the best model of evolution for our data set in accordance with the AICc of MEGA6 [38]. A total of 2×10^6 simulated data sets were calculated. Pre-evaluations of model-prior combinations in ABC inference revealed that the prior settings were correctly assigned (S2 Fig; [53]). The model checking indicated that posterior combination under each scenario corresponded well with a target data set (S2 Fig), which confirmed the 'goodness-of-fit' of the models [53].

Results

Phylogenetic reconstruction, phylogeography and species delimitation

The Bayesian phylogeny of *Creatonotos* spp. based on the 61 *COI* haplotypes reveals the three species-level clades with high Bayesian posterior probabilities (BPP = 1.00) (Fig 1 and S1 Table). The haplotype of *Creatonotos omanirana* De Freina, 2007 is located within haplotypes of *C. gangis* from South Asia (India, Pakistan and Nepal) and the Lesser Sunda Archipelago (Flores). Among the taxa under discussion, *Creatonotos gangis* shows the strongest spatial structure of genetic diversity. There are three subclades of the *Creatonotos gangis* haplotypes with moderate support: Southeast Asia (BPP = 0.94), Australia (BPP = 0.91), and South Asia + the Arabian Peninsula + Lesser Sundas (BPP = 0.75). The Bayesian modeling did not resolve the position of several haplotypes from Myanmar, Thailand and Middle Andaman (gang11, gang14, gang15 and gang34) although the median-joining network analysis, which was calculated on the basis of the short length sequence data set, placed these haplotypes within the Southeast Asia haplogroup (see below). In contrast, *Creatonotos transiens* does not reveal such a spatial structure, with the only divergent haplotype from India (Western Ghats). All of the other haplotypes (Southeast Asia, South Asia, Taiwan and Borneo) are sampled into a single polytomic clade with moderate support (BPP = 0.75) (Fig 1).

The median-joining network reflects the same pattern, with three groups of closely related haplotypes (1–4 mutational steps) corresponding to the three species, which are separated by 20–28 nucleotide substitutions (Fig 2). The sequences of *Creatonotos transiens* show the 'star-like' network, which includes the two most frequently recorded haplotypes, which have broad range across Southeast Asia, with supplement of many local singletons. In contrast, the network of *Creatonotos gangis* is spatially structured, with several distant haplogroups ranged in Southeast Asia, Australia, and Arabia/South Asia (Fig 2). The haplotype of *Creatonotos omanirana* is located between haplotypes from Southeast Asia (Myanmar, Thailand, Middle Andaman and Flores) and Nepal. The highest Bayesian supported solution of the PTP model also support the three species-level MOTUs (Fig 1 and S1 Fig), but with moderate probabilities (0.45–0.72). The results of the barcoding gap analysis show a well-defined 3% species-level threshold (Fig 2). The mean genetic divergences between *Creatonotos gangis* populations are $\leq 1.78\%$, whereas the minimum interspecific distance of *Creatonotos* spp. is 6.41% (S2 Table).

Historical demography and divergence time estimates

With respect to the approach of Grant & Bowen [55], the Australian and Arabian-South Asian populations of *Creatonotos gangis* differ by large haplotype diversity (h ranges from 0.54 to 0.66) and small nucleotide diversity (π ranges from 0.2 to 0.4%) (S3 Table). All of the other samples of *Creatonotos gangis* and the sample of *C. transiens* reveal large haplotype diversity (h >0.85) and large nucleotide diversity (π >0.5%). The values of Fu's *FS* and Tajima's *D* tests indicate no deviation from mutation-drift equilibrium for the samples of *Creatonotos gangis*, whereas the *C. transiens* sample has significant negative values of both statistics revealing a possible historic demographic expansion (S3 Table).

The mismatch distribution analysis of the Southeast Asian and Arabian-South Asian populations of *Creatonotos gangis* resulted in a multimodal distribution with three peaks at 0, 5 and 8 bp (Southeast Asia) and at 0, 2 and 6 bp (Arabia and South Asia) (Fig 3). The general samples of *Creatonotos gangis* (Eurasia and all available sequences) also reveal multiple peaks, whereas the Australian population has a unimodal distribution with the maximum value at 0 bp. The sample of *Creatonotos transiens* from Eurasia shows a bimodal distribution with peaks at 1 and



Fig 1. Fifty-percent majority-rule consensus phylogenetic tree of *Creatonotos* spp. recovered from Bayesian inference analysis of an alignment comprising 61 *COI* haplotypes of *Creatonotos* spp. and two haplotypes of the out-group taxa (*Arctia menetriesii* and *A. tundrana*). The island localities are in bold. Black numbers near branches are Bayesian posterior probabilities (BPP). The red number near each primary clade is the probability of each species-level MOTU based on the highest Bayesian supported solution of the PTP model.

https://doi.org/10.1371/journal.pone.0194200.g001

PLOS ONE

6 bp. The Australian population of *Creatonotos gangis* and the sample of *C. transiens* from Eurasia reveal the lowest values of the parameter τ , which reflects the time since expansion, while all of the other samples of *C. gangis* return much larger moment estimator values (S3 Table). According to the 'mid-Aegean' mutation rate of 1.77×10^{-8} s/s/y (substitutions/site/year), the





Fig 2. Phylogeography of *Creatonotos* spp. (A) Median-joining network of *COI* sequences (see <u>S1 Table</u> for details). Photos (male specimens): *C. gangis* [Indonesia, Flores Island, voucher no. Sph0595] and *C. transiens* [Thailand, near Tham Lod Cave, voucher no. Sph0624] by Vitaly M. Spitsyn and *C. leucanioides* [Tanzania] by Roy Goff (with his permission; www.africanmoths.com). (B) Map of approximate collection localities of the specimens in accordance with the respective countries (color squares). Small squares indicate island sites. The map was created using ESRI ArcGIS 10 software (www.esri.com/arcgis); the base of the map was created with ESRI Data and Maps. (C) Frequency histogram of the barcoding gap analysis.

https://doi.org/10.1371/journal.pone.0194200.g002

mean time since expansion of the Arabian-South Asian population of *Creatonotos gangis* was 640 ka (thousands of years) and that of the Australian population was 67 ka (<u>S3 Table</u>). The time since expansion of *Creatonotos transiens* population was 196 ka.

The results of the ABC modeling reveal that the scenario 2-CG, which proposed more recent dispersal events into South Asia, Arabia and Australia, was assigned as the most likely scenario with a high posterior probability (Fig 4). The posterior predictive error rate, which was calculated over 1000 data sets using the direct approach, was 0.151. Type I and Type II errors for the choice of scenario 2-CG in accordance with the direct approach were 0.234 and 0.144, respectively. Modeling results, obtained under the scenario 2-CG, reveal the most recent dating of divergence times for the splits between populations of *Creatonotos gangis* (S4 Table). In particular, the Arabian-South Asian population most likely derived from the Southeast Asian population before the Late Pleistocene (mean age 143.0 ka, median age 109.0 ka, 95% CI 55.8–359.0 ka), while the Australian population most likely originated in the Late Pleistocene (mean age 62.6 ka, median age 65.1 ka, 95% CI 19.3–96.9 ka) (Fig 5). Our model predicts a relatively slow substitution rate in populations of *Creatonotos gangis* (mean rate 1.52×10^{-8} s/s/y, median rate 1.35×10^{-8} s/s/y, 95% CI $1.03 \times 10^{-8} - 2.58 \times 10^{-8}$ s/s/y).

PLOS ONE



Fig 3. Mismatch distributions of *Creatonotos* **spp. samples based on the mitochondrial** *COI* **gene.** Solid black lines indicate observed distribution, and solid red lines represent simulated distribution under a spatial expansion model. Dashed lines represent lower and upper confidence intervals (p = 0.01). (**A**) *C. gangis*, Eurasia (N = 40 sequences; Raggedness P = 0.342; Model (SSD) P = 0.317). (**B**) *C. gangis*, Arabia and South Asia (N = 20 sequences; Raggedness P = 0.264; Model (SSD) P = 0.273). (**C**) *C. gangis*, mainland Southeast Asia (N = 20 sequences; Raggedness P = 0.331). (**D**) *C. gangis*, Australia (N = 15 sequences; Raggedness P = 0.910; Model (SSD) P = 0.736). (**E**) *C. gangis*, the entire range (N = 57 sequences; Raggedness P = 0.715; Model (SSD) P = 0.557). (**F**) *C. transiens*, Eurasia (N = 39 sequences; Raggedness P = 0.671; Model (SSD) P = 0.596).

https://doi.org/10.1371/journal.pone.0194200.g003



Fig 4. ABC modeling of origin of the *Creatonotos gangis* **populations.** (A) Biogeographic scenarios that were tested under an ABC framework using *COI* gene sequences. Southeast Asian population: samples from Myanmar, Vietnam, Thailand, and South China, Arabian–South Asian population: samples from Oman, Pakistan, India, and Nepal, and Australian population. Effective population size: N_1 –Southeast Asian population; N_2 –Arabian–South Asian population; N_3 –Australian population; N_{2b} and N_{3b} –hypothetical founder population for Arabian–South Asian and Australian populations, respectively. Time intervals: t_{1a} and t_1 – time of the primary split between populations *Pop1* and *Pop2* under scenarios 1-CG (before the mid-Pleistocene) and 2-CG (since the mid-Pleistocene), respectively; t_{2a} and t_2 –time of split between Southeast Asian and Australian populations under scenarios 1-CG (before the Late Pleistocene) and 2-CG (during the Late Pleistocene), respectively; t_{2a} and t_2 –time of split between Southeast Asian and Australian populations under scenarios 1-CG (before the Late Pleistocene) and 2-CG (during the Late Pleistocene), respectively; t_{2b} and N_{3b} –hypothetic of low effective population size N_{2b} and N_{3b} since colonization of the Arabian–South Asian Region and Australia. Prior settings are presented in S5 Table. Evaluating the confidence in scenario choice using the direct (**B**) and linear regression (**C**) approaches.

https://doi.org/10.1371/journal.pone.0194200.g004

Discussion

Biogeography of *Creatonotos* spp. and possible Late Pleistocene expansions of Asian taxa into Sahul

In general, the results of our phylogenetic and population genetic analyses support the conclusion that both Creatonotos gangis and C. transiens are widespread polymorphic species. Creatonotos gangis comprises at least three divergent intraspecific subclades corresponding to certain geographic areas, i.e., Australia, Arabia + South Asia and Southeast Asia. This pattern may reflect ancient splitting events. Based on the approach of Grant & Bowen [55], large haplotype diversity together with small nucleotide diversity in Arabian-South Asian and Australian samples (h >0.5, π <0.5%) may indicate a population bottleneck since a founder event followed by rapid population growth and accumulation of mutations. According to this scenario, Southeast Asia appears to be the most probable ancestral area of *Creatonotos gangis*, from which this species has colonized other areas during the Late Pleistocene. The ABC modeling suggests that the Australian clade likely separated from the Southeast Asian populations at ~65-63 ka (median and mean estimations of the 2-CG scenario, respectively). This dating roughly coincided with that of the earliest human expansion crossing the sea-level-altered Indonesian Archipelago and arrive into Sahul at ~65-54 ka [56,57,58,59,60]. This human expansion was largely driven by orbital-scale global climate swings [58] that could have been accelerated a long-distance dispersal of Creatonotos gangis and other Asian insect taxa. The inter-island visibility modeling of Bergström et al. [56] provides an evidence for intervisibility between Timor and Australia at approximately 65-62 ka, which corresponds to the time of Creatonotos gangis expansion inferred from our ABC model. In contrast, specimens from the Lesser Sundas are not sister to the Australian lineage (Figs 1 and 2) but show affinities to the Arabian-South



Fig 5. Simplified summary of expansion routes inferred across populations of *Creatonotos gangis*, and examples of male specimens and habitat of Southeast Asian population. (A) Map of expansion routes. Color circles indicate an approximate range of each population: Southeast Asian population (blue), Arabian–South Asian population (green), and Australian population (orange). Red arrows indicate the directions of expansion routes. The numbers near arrows show the mean age of putative expansion events (in thousands of years) obtained from the ABC model. The map was created using ESRI ArcGIS 10 software (www.esri.com/arcgis); the base of the map was created with ESRI Data and Maps. (B) Male specimen, Maehongson, Thailand. (C) Male specimen, Kachin, Myanmar. (D) Paddy field, a typical habitat of the species, Thanh Hoa Province, Vietnam. (Photos: Vitaly M. Spitsyn).

https://doi.org/10.1371/journal.pone.0194200.g005

PLOS ONE

Asian subclade that may indicate a secondary, more recent expansion from the latter region into Wallacea. The Andaman Islands were likely recently colonized by the representatives of *Creatonotos gangis* from Southeast Asia (Figs 1 and 2).

These patterns differ from those discovered in many other insect taxa. For example, *Lampides boeticus* (Linnaeus, 1767), one of the most widely distributed butterfly species in the world, expands into Australia several times from populations of Sundaland and Wallacea [7]. Additionally, Australian populations of this butterfly have very low mtDNA divergence from corresponding Sundaic and Wallacean populations, which suggests an invasion since the Last Glacial Maximum. The same phylogeographic pattern with a small genetic divergence between Southeast Asian and Australian populations was observed in other widespread species, e.g., *Zizina otis* (Fabricius, 1787), *Danaus chrysippus* (Linnaeus, 1758), and *Asota caricae* (Fabricius, 1775) [5,6,8], and even in the migratory locust *Locusta migratoria* (Linnaeus, 1758) [61].

The other studied species, *Creatonotos transiens*, shows relatively shallow genetic divergence between populations across the entire distribution range, which suggests a possible sudden population expansion during the Late Pleistocene. However, in accordance with the approach of Grant & Bowen [55], large haplotype diversity together with large nucleotide diversity in the sample (h = 0.85, π = 0.6%) may indicate large stable population with long evolutionary history or secondary contact between differentiated lineages. Similar shallow genetic divergence was recorded in *Arctia plantaginis* (Linnaeus, 1758), another tiger moth species, within its range from Eurasia to North America [62]. The Polar tiger moth *Arctia tundrana* (Tshistjakov, 1990) shares closely related haplotypes across the Eurasian Arctic at the distance of ~5,000 km [63]. Rönkä et al. [64] show that many taxa within the widespread *Arctia caja* (Linnaeus, 1758) species complex form a single cluster with high support and very little genetic difference. Similar patterns were discovered in many moth and butterfly species from other families [5,6,8,65,66,67]. These findings highlight the role of long-distance dispersal events connecting distinct populations by gene flows, which leads to the lack of significant genetic divergences between individuals from geographically remote areas. We suggest that such species can exist as the huge metapopulations with broad ranges, which may cover extensive areas within one or even two continents, and hence these taxa do not reveal clear spatial differentiation on geographic races (subspecies).

Putative taxonomic implications

De Freina [15] described *Creatonotos omanirana* as a new species, which is closely related to *C. gangis*. However, our wide-scale phylogeny reveals that this taxon is rather a variation of *Creatonotos gangis*, because the available *COI* sequence of *C. omanirana* falls within the *C. gangis* clade and the morphological differences between these taxa are small [15,34] (Figs 1 and 2). Our phylogenetic reconstructions also suggest that the *Creatonotos gangis* clade may include several subspecies-level taxa, e.g., from Australia and from Lesser Sundas, but a broader sampling is needed to access the fully resolved phylogeny with well-supported subspecies-level clades. Additionally, *Creatonotos fasciatus* Candèze, 1927, a prospective sister species of *C. gangis* [14], has been lacking in our samples. This poorly known species is in need of future molecular investigations.

As for the intraspecific phylogeny of *Creatonotos transiens*, we found no clear support for the two subspecies that were considered valid by Dubatolov and Holloway [68]: *C. transiens koni* Miyake, 1909 from Taiwan and Japan, and *C. transiens vacillans* (Walker, 1855) from China, Indo-China, the Malay Peninsula, Sumatra and Borneo. Several other subspecies [68] are beyond the framework of the present study, and they are still waiting for molecular investigations. In contrast, our sample includes a divergent *Creatonotos transiens* sequence from the Western Ghats, in which a separate subspecies may be present, but it needs to be confirmed in the future.

In summary, our new molecular-based findings could be used as a supplement for future taxonomic revisions of *Creatonotos gangis* and *C. transiens* because both these species include several divergent intraspecific mtDNA lineages that may be considered separate evolutionary significant units.

Supporting information

S1 Fig. Bayesian species delineation using the Poisson Tree Process (PTP) model based on the distribution of nucleotide substitutions in the *COI* **Bayesian phylogeny of** *Creatonotos* **spp.** Terminal branches in blue indicate lineages that stand as separate Molecular Operational Taxonomic Units (MOTUs) and the clades in red are lumped into a single MOTU. Numbers near branches are support values of each MOTU based on the PTP model. (PDF)

S2 Fig. Test results of biogeographical scenarios 1-CG and 2-CG concerning origin of the *Creatonotos gangis* populations under an ABC framework using the *COI* gene sequences. (a) Pre-evaluation of prior combinations of scenarios. (b)-(c). Model checking to measure a mismatch between the parameters of posterior combination and observed data sets in scenarios 1- CG (b) and 2- CG (c). Scenarios are illustrated in Fig 4. A description of basic assumptions with prior settings for each scenario is presented in <u>S5 Table</u>. (PDF)

S1 Table. List of the mitochondrial *COI* sequences of *Creatonotos* spp. examined in the present study.

(PDF)

S2 Table. Genetic divergences (mean uncorrected *p*-distance \pm standard error estimations, %) between populations and taxa of *Creatonotos* spp. based on the mitochondrial *COI* gene fragment.

(PDF)

S3 Table. Molecular diversity indexes and population expansion test statistics of *Creatonotos* spp. samples based on the *COI* sequences. (PDF)

S4 Table. Model parameters estimated from posterior distribution of the scenario 2-CG concerning origin of *Creatonotos gangis* populations within the ABC framework. (PDF)

S5 Table. Prior assumptions and settings of the two biogeographical scenarios of the origin of *Creatonotos gangis* populations, which were tested under an ABC framework. (PDF)

Acknowledgments

We are very grateful to Dr. Vladimir V. Dubatolov and Dr. Oleg E. Kosterin for their helpful and insightful comments on the manuscript.

Author Contributions

Conceptualization: Nhi Thi Pham, Mikhail Y. Gofarov, Ivan N. Bolotov.

Data curation: Vitaly M. Spitsyn, Nikita I. Bolotov.

Formal analysis: Alexander V. Kondakov, Ivan N. Bolotov.

Methodology: Alexander V. Kondakov.

Writing - original draft: Ivan N. Bolotov.

Writing – review & editing: Vitaly M. Spitsyn, Nikita I. Bolotov, Nhi Thi Pham, Mikhail Y. Gofarov, Ivan N. Bolotov.

References

- 1. Wallace AR. On the zoological geography of the Malay Archipelago. Journal of the Proceedings of the Linnean Society: Zoology. 1860; 4: 172–184.
- Condamine FL, Toussaint EF, Clamens AL, Genson G, Sperling FA, Kergoat GJ. Deciphering the evolution of birdwing butterflies 150 years after Alfred Russel Wallace. Scientific Reports. 2015; 5. https://doi.org/10.1038/srep11860 PMID: 26133078

- Bolotov IN, Kondakov AV, Spitsyn VM, Gofarov MY, Kolosova YS () *Leptocneria vinarskii* sp. nov. (Lepidoptera: Erebidae: Lymantriinae), an overlooked Wallacean lineage of the Australian genus. Scientific Reports. 2017; 7: 12430. https://doi.org/10.1038/s41598-017-12797-3 PMID: 28963487
- Rougerie R, Kitching IJ, Haxaire J, Miller SE, Hausmann A, Hebert PD. Australian Sphingidae–DNA barcodes challenge current species boundaries and distributions. PLoS ONE. 2014; 9. e101108; https://doi.org/10.1371/journal.pone.0101108 PMID: 24987846
- Braby MF, Farias Quipildor GE, Vane-Wright RI, Lohman DJ. Morphological and molecular evidence supports recognition of *Danaus petilia* (Stoll, 1790) (Lepidoptera: Nymphalidae) as a species distinct from *D. chrysippus* (Linnaeus, 1758). Syst Biodivers. 2015; 13: 386–402.
- Craft KJ, Pauls SU, Darrow K, Miller SE, Hebert PDN, Helgen LE, et al. Population genetics of ecological communities with DNA barcodes: An example from New Guinea Lepidoptera. P Natl Acad Sci USA. 2010; 107 (11): 5041–5046. https://doi.org/10.1073/pnas.0913084107 PMID: 20202924
- Lohman DJ, Peggie D, Pierce NE, Meier R. Phylogeography and genetic diversity of a widespread Old World butterfly, *Lampides boeticus* (Lepidoptera: Lycaenidae). BMC Evol Biol. 2008; 8: 301. https://doi. org/10.1186/1471-2148-8-301 PMID: 18973689
- 8. Yago M, Hirai N, Kondo M, Tanikawa T, Ishii M, Wang M, et al. Molecular systematics and biogeography of the genus *Zizina* (Lepidoptera: Lycaenidae). Zootaxa. 2008; 1746: 15–38.
- 9. Lohman DJ, Bruyn M, Page T, Rintelen K, Hall R, Ng PKL, et al. Biogeography of the Indo-Australian archipelago. Annu Rev Ecol Evol. 2011; 42: 205–226.
- Bito D. An alien in an archipelago: Spathodea campanulata and the geographic variability of its moth (Lepidoptera) communities in the New Guinea and Bismarck Islands. J Biogeogr. 2007; 34: 769–778.
- Černý K. A review of the subfamily Arctiinae (Lepidoptera: Arctiidae) from the Philippines. Entomofauna. 2011; 32: 29–92.
- Dubatolov VV. Tiger moths of Eurasia (Lepidoptera, Arctiidae) (Nyctemerini by Rob de Vos & Vladimir V. Dubatolov). Neue Entomologische Nachrichten. 2010; 65: 1–106.
- 13. Van Eecke R. De Heterocera van Sumatra, IV. Zoologische Mededelingen, Leiden. 1927; 10: 90–156.
- Dubatolov VV. Creatonotos fasciatus (Candèze, 1927) (Lepidoptera, Arctiidae), a little known species from an Indochina sibling to C. gangis (Linnaeus, 1763), with a description of a new subspecies from Laos and Hong Kong. Euroasian Entomological Journal. 2010; 9: 727–730.
- 15. De Freina JJ. *Creatonotos omanirana* sp. n. aus dem Oman und dem Iran (Artiidae: Arctiinae). Nota Lepidopterologica. 2007; 30: 375–386 https://archive.org/details/notalepidopter3022007soci
- Shah SK, Mitra B. Moth (Insecta: Lepidoptera) fauna and their insect predators associated with the tea gardens and the surrounding natural ecosystem environs in Northern West Bengal, India. The Journal of Zoology Studies. 2015; 2: 1–5.
- Paul M, Das SK, Singh R, Shashank PR. Moth (Lepidoptera: Heterocera) fauna of Delhi with notes on their role as potential agricultural pests. Journal of Entomology and Zoology Studies. 2016; 4: 435–438.
- Allsopp PG. The biology of false wireworms and their adults (soil-inhabiting Tenebrionidae) (Coleoptera): a review. B Entomol Res. 1980; 70: 343–379.
- Boppré M, Schneider D. The biology of *Creatonotos* (Lepidoptera: Arctiidae) with special reference to the androconial system. Zool J Linn Soc-Lond. 1989; 96: 339–356.
- Wink M, Schneider D. Fate of plant-derived secondary metabolites in three moth species (Syntomis mogadorensis, Syntomeida epilais, and Creatonotos transiens). J Comp Physiol B. 1990; 160: 389– 400.
- Catindig JLA, Barrion AT, Litsinger JA. Developmental biology and host plant range of rice-feeding tiger moth *Creatonotus gangis* (L.). International Rice Research Notes. 1993; 18: 34–35.
- Catling D, Islam Z. Diversity and seasonal fluctuations of arthropod fauna in Bangladesh deepwater rice. Bangladesh Rice Journal. 2013; 17: 75–104.
- 23. Gurule SA, Nikam SM, Kharat AJ, Gangurde JH. Check-list of owlet and underwing moths (Lepidoptera: Noctuidae) from Nashik District, MS, India. Flora and Fauna. 2010; 16: 295–304.
- 24. Biswas O, Modak BK, Mazumder A, Mitra B. Moth (Lepidoptera: Heterocera) diversity of Sunderban Biosphere Reserve, India and their pest status to economically important plants. Journal of Entomology and Zoology Studies. 2016; 4: 13–19.
- Holland D, Hatib K, Bar-Ya'akov I. Pomegranate: botany, horticulture, breeding. Horticultural Reviews. 2009; 35: 127–191.
- Kumar M, Ahmad M. Quantitative study of food consumption, assimilation and growth in *Creatonotos transiens* Walker (Lepidoptera: Arctiidae) on *Paulownia fortunei*. Indian J Agr Res. 2006; 40: 42–46.

- Wunderer H, Hansen K, Bell TW, Schneider D, Meinwald J. Sex pheromones of two Asian moths (*Creatonotos transiens, C. gangis*; Lepidoptera—Arctiidae): behavior, morphology, chemistry and electro-physiology. Exp Biol. 1985; 46: 11–27.
- Wunderer H, De Kramer JJ. Dorsal ocelli and light-induced diurnal activity patterns in the arctiid moth Creatonotos transiens. J Insect Physiol. 1989; 35: 87–95.
- Schneider D, Boppré M, Zweig J, Horsley SB, Bell TW, Meinwald J, et al. Scent organ development in *Creatonotos* moths: regulation by pyrrolizidine alkaloids. Science. 1982; 215: 1264–1265. https://doi. org/10.1126/science.215.4537.1264 PMID: 17757544
- Boppré M. Insects pharmacophagously utilizing defensive plant chemicals (pyrrolizidine alkaloids). Naturwissenschaften. 1986; 73: 17–26.
- Wink M, Von Nickisch-Rosenegk E. Sequence data of mitochondrial 16S rDNA of Arctiidae and Nymphalidae: evidence for a convergent evolution of pyrrolizidine alkoloid and cardiac glycoside sequestration. J Chem Ecol. 1997; 23: 1549–1568.
- Zahiri R, Holloway JD, Kitching IJ, Lafontaine JD, Mutanen M, Wahlberg N. Molecular phylogenetics of Erebidae (Lepidoptera, Noctuoidea). Syst Entomol. 2012; 37: 102–124.
- Zaspel JM, Weller SJ, Wardwell CT, Zahiri R, Wahlberg N. Phylogeny and evolution of pharmacophagy in tiger moths (Lepidoptera: Erebidae: Arctiinae). PLoS ONE. 2014; 9. e101975; <u>https://doi.org/10. 1371/journal.pone.0101975</u> PMID: 25036028
- 34. De Freina JJ. Über das Artenspektrum der Gattung Creatonotos Hübner, 1816 auf der Arabischen Halbinsel mit einer Revision der afrotropischen Creatonotos leucanioides Holland, 1893—Artengruppe (Lepidoptera: Arctiidae, Spilosominae). Esperiana Buchreihe zur Entomologie. 2010; 15: 423–432.
- Sambrook J, Fritsch EF, Maniatis T. Molecular Cloning: A Laboratory Manual, (2nd ed.) 10.51–10.67 (Cold Spring Harbor: Cold Spring Harbor Laboratory Press, 1989).
- Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol Mar Biol Biotech. 1994; 3: 294– 299.
- Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acid S. 1999; 41: 95–98.
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol. 2013; 30: 2725–2729. <u>https://doi.org/10.1093/molbev/mst197</u> PMID: 24132122
- 39. Villesen P. FaBox: an online toolbox for FASTA sequences. Mol Ecol Notes. 2007; 7: 965–968.
- 40. Ronquist F, Teslenko M, Mark P, Ayres DL, Darling A, Höhna S, et al. MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. Syst Biol. 2012; 61: 539–542. https://doi.org/10.1093/sysbio/sys029 PMID: 22357727
- 41. Rambaut A, Suchard MA, Xie D, Drummond AJ. Tracer v1.6 http://beast.bio.ed.ac.uk/Tracer (2014).
- Miller M, Pfeiffer W, Schwartz T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. Gateway Computing Environments Workshop (GCE). 2010; 1–8.
- Bandelt HJ, Forster P, Röhl A. Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol. 1999; 16: 37–48. https://doi.org/10.1093/oxfordjournals.molbev.a026036 PMID: 10331250
- Zhang J, Kapli P, Pavlidis P, Stamatakis A. A general species delimitation method with applications to phylogenetic placements. Bioinformatics. 2013; 29: 2869–2876. https://doi.org/10.1093/bioinformatics/ btt499 PMID: 23990417
- Hebert PD, Cywinska A, Ball SL. Biological identifications through DNA barcodes. Proc R Soc B. 2003; 270: 313–321. https://doi.org/10.1098/rspb.2002.2218 PMID: 12614582
- Ashfaq M, Akhtar S, Khan AM, Adamowicz SJ, Hebert PD. DNA barcode analysis of butterfly species from Pakistan points towards regional endemism. Mol Ecol Resour. 2013; 13: 832–843. https://doi.org/ 10.1111/1755-0998.12131 PMID: 23789612
- 47. Excoffier L, Laval G, Schneider S. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evol Bioinform. 2005; 1: 47–50.
- Excoffier L, Lischer HEL. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour. 2010; 10: 564–567. <u>https://doi.org/10.1111/j. 1755-0998.2010.02847.x PMID: 21565059</u>
- 49. Epperson BK. Geographical Genetics. 356 (Princeton University Press, Oxford, 2003).
- Schenekar T, Weiss S. High rate of calculation errors in mismatch distribution analysis results in numerous false inferences of biological importance. Heredity. 2011; 107: 511. <u>https://doi.org/10.1038/hdy.</u> 2011.48 PMID: 21731052

- Fehér Z, Major Á, Krízsik V. Spatial pattern of intraspecific mitochondrial diversity in the Northern Carpathian endemic spring snail, *Bythinella pannonica* (Frauenfeld, 1865) (Gastropoda: Hydrobiidae). Org Divers Evol. 2013; 13: 569–581.
- Papadopoulou A, Anastasiou I, Vogler AP. Revisiting the insect mitochondrial molecular clock: the mid-Aegean trench calibration. Mol Biol Evol. 2010; 27: 1659–1672. https://doi.org/10.1093/molbev/ msq051 PMID: 20167609
- Cornuet JM, Ravigné V, Estoup A. Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1. 0). BMC Bioinformatics. 2010; 11: 401. https://doi.org/10.1186/1471-2105-11-401 PMID: 20667077
- 54. Cornuet JM, Pudlo P, Veyssier J, Dehne-Garcia A, Gautier M, Leblois R, et al. DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. Bioinformatics. 2014; 30: 1187–1189. https://doi.org/10.1093/bioinformatics/btt763 PMID: 24389659
- Grant WAS, Bowen BW. Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons for conservation. Journal of Heredity. 1998; 89(5): 415–426.
- 56. Bergström A, Nagle N, Chen Y, McCarthy S, Pollard MO, Ayub Q, et al. Deep roots for Aboriginal Australian Y chromosomes. Curr Biol. 2016; 26: 809–813. <u>https://doi.org/10.1016/j.cub.2016.01.028</u> PMID: 26923783
- Johnson CN, Alroy J, Beeton NJ, Bird MI, Brook BW, Cooper A, et al. What caused extinction of the Pleistocene megafauna of Sahul? Proc R Soc B. 2016; 283. 2015.2399; https://doi.org/10.1098/rspb. 2015.2399 PMID: 26865301
- Timmermann A, Friedrich T. Late Pleistocene climate drivers of early human migration. Nature. 2016; 538: 92–95. https://doi.org/10.1038/nature19365 PMID: 27654920
- Saltré F, Rodríguez-Rey M, Brook BW, Johnson CN, Turney CS, Alroy J, et al. Climate change not to blame for late Quaternary megafauna extinctions in Australia. Nat Comm. 2016; 7. https://doi.org/10. 1038/ncomms10511 PMID: 26821754
- Kealy S, Louys J, O'Connor S. Reconstructing Palaeogeography and Interisland Visibility in the Wallacean Archipelago During the Likely Period of Sahul Colonization, 65–45 000 Years Ago. Archaeol Prospect. 2017; 1570.
- Ma C, Yang P, Jiang F, Chapuis MP, Shali Y, Sword GA, et al. Mitochondrial genomes reveal the global phylogeography and dispersal routes of the migratory locust. Mol Ecol. 2012; 21: 4344–4358. https:// doi.org/10.1111/j.1365-294X.2012.05684.x PMID: 22738353
- Hegna RH, Galarza JA, Mappes J. Global phylogeography and geographical variation in warning coloration of the wood tiger moth (*Parasemia plantaginis*). J Biogeogr. 2015; 42: 1469–1481.
- Bolotov IN, Tatarinov AG, Filippov BY, Gofarov MY, Kondakov AV, Kulakova OI, et al. The distribution and biology of *Pararctia subnebulosa* (Dyar, 1899) (Lepidoptera: Erebidae: Arctiinae), the largest tiger moth species in the High Arctic. Polar Biol. 2015; 38: 905–911.
- Rönkä K, Mappes J, Kaila L, Wahlberg N. Putting *Parasemia* in its phylogenetic place: a molecular analysis of the subtribe Arctiina (Lepidoptera). Syst Entomol. 2016; 41: 844–853.
- Zhou C, Chen X, He R. COII phylogeography reveals surprising divergences within the cryptic butterfly Kallima inachus (Doyére, 1840) (Lepidoptera: Nymphalidae: Kallimini) in southeastern Asia. The Pan-Pacific Entomologist. 2012; 88: 381–398.
- Zhou LH, Wang XY, Xu GQ, Lei JJ. Mitochondrial DNA phylogeography of Spodoptera exigua across a broad geographic area in China. J Appl Entomol. 2016; 12371. https://doi.org/10.1111/jen.12371
- Vodă R, Dapporto L, Dincă V, Shreeve TG, Khaldi M, Barech G, et al. Historical and contemporary factors generate unique butterfly communities on islands. Scientific Reports. 2016; 6. <u>https://doi.org/10. 1038/srep28828 PMID: 27353723</u>
- Dubatolov VV, Holloway JD. A new species of the *Creatonotos transiens*-group (Lepidoptera: Arctiidae) from Sulawesi, Indonesia. Bonner zoologische Beiträge. 2006; 55: 113–121.