ORIGINAL ARTICLE

A distinct new subspecies of the white-rumped shama *Copsychus malabaricus* at imminent risk of extinction

Meng Yue Wu[1] · Frank E. Rheindt[1]

Abstract

We here describe a new subspecies of the white-rumped shama *Copsychus malabaricus* from the western islands of the Thai Malay Peninsula. The new subspecies is mitogenomically distinct from other members of the complex. Morphologically, its tail is longer and the black breast in males is much less extensive than in males of all other taxa. The discovery of a new shama in a relatively well-explored part of Southeast Asia brings to light a gap in our taxonomic understanding of tropical Asian birds, which continue to be understudied. The new Langkawi shama is the subject of extensive specialized poaching efforts, and its survival may be at risk, calling for immediate conservation action.

Keywords Morphometrics · Novel subspecies · Integrative taxonomy · Museum collections

Zusammenfassung

Eine markante neue Unterart der Weißbürzel-Schama *Copsychus malabaricus*, die unmittelbar vom Aussterben bedroht ist

Wir beschreiben hier eine neue Unterart der Weißbürzel-Schama *Copsychus malabaricus* von der Inselkette westlich der Thai-Malayischen Halbinsel. Die neue Unterart unterscheidet sich grundlegend von anderen Mitgliedern des Artkomplexes in ihrem mitochondrialen Genom. Morphologisch ist ihr Schwanz länger, und die schwarze Brust des Männchens ist viel weniger ausgedehnt als bei den Männchen aller anderen Taxa. Die Entdeckung einer neuen Schama in einem relativ gründlich erschlossenen Gebiet Südostasiens entblößt eine Lücke in unserem taxonomischen Verständnis der Vögel des tropischen Asiens, die nach wie vor wenig erforscht sind. Die neue Langkawi-Schama ist das Ziel umfangreicher, spezialisierter Wilderei und ist daher direkt vom Aussterben bedroht. Sofortige Schutzmaßnahmen sind dringend erforderlich.

Introduction

The white-rumped shama *Copsychus malabaricus* complex is an important songbird radiation that spans widely across tropical Asia (Roberts et al. 2020; Eaton et al. 2021). White-rumped shamas are known for their melodious song, making them one of the most highly sought-after targets in the Asian songbird trade (Nash 1993; Eaton et al. 2015). This popularity has led to the recent endangerment and even extinction-in-the-wild of a number of taxa in the complex (Ng et al. 2017; Symes et al. 2018; Rheindt et al. 2019, 2020). Many cases of endangerment and extinction, however, have gone under the conservation radar because the most widely accepted taxonomic treatment considers many component taxa to be subspecies of one widespread species ranging from the Indian subcontinent in the west to Borneo and Java in the east: the white-rumped shama *C. malabaricus*.

In the present contribution, we formally describe a new taxon of white-rumped shama (Figs. 1, 2) from the Langkawi archipelago and adjacent islands on the west coast of the northern Thai Malay Peninsula (Fig. 3), henceforth referred to by its vernacular name Langkawi shama. The new taxon is morphologically distinct, falls outside the main clade of white-rumped shamas in terms of mitochondrial DNA (mtDNA), and has long been overlooked by taxonomists,

1 Department of Biological Sciences, National University of Singapore, 16 Science Drive 4, Singapore, Republic of Singapore

Fig. 1 Photographs of male shama specimens in comparison, taken at the Lee Kong Chian Natural History Museum (Singapore). Specimens are shown in ventral, dorsal and lateral orientation and are ordered from *left* to *right*: holotype of the Langkawi shama from Langkawi Island (Malaysia); Langkawi shama from Tarutao Island (Thailand); *tricolor* from Perak (peninsular Malaysia); *melanurus* from Sipora Island in the Mentawai group (Indonesia)

although it is well known among bird breeders, traders and keepers in northernmost Malaysia. We carried out biometric measurements and sequenced the mitogenome from historic degraded DNA of a specimen collected in 1916 from Langkawi Island, stored in the Lee Kong Chian Natural History Museum, Singapore, which constitutes the holotype of our description.

The classification of the white-rumped shama complex has been relatively stable with only minor incongruences over the past century (Peters 1964; Clements et al. 2019; Gill et al. 2021; HBW 2019), although novel treatments over the last 5 years have begun to afford species status to some of the most conspicuous members of the complex (e.g., Lim et al. 2017; Eaton et al. 2021). The new taxon has an archipelagic distribution that is geographically adjacent to the white-rumped shama subspecies *macrourus* from most of mainland Southeast Asia north of the Isthmus of Kra and subspecies *tricolor* from the rainforest belt of Sumatra, West Java and the Thai Malay Peninsula south of the Isthmus (Fig. 3a; Roberts et al. 2020). Ca. 500 km to the west of the archipelago inhabited by the new taxon are the Andaman Islands, which host a distinct shama that has been afforded species rank in most classifications over the last 50 years: the Andaman shama *C. albiventris*. Our comparisons of the

new taxon concentrate on these three adjacent taxa, but also encompass all other known taxa of white-rumped shama.

Methods and materials

Mitogenome sequencing and analysis

We sequenced the DNA contained within 5 toepads from historic museum specimens and 9 fresh blood samples, representing populations from Langkawi Island, *tricolor* and *macrourus* from the nearby Thai Malay Peninsula, Andaman Island *albiventris*, Bornean *stricklandii*, Indian *malabaricus*, and the Barusan Island populations of *hypolizus* and *melanurus*. Toepad samples were provided by museums, whereas most of the fresh blood samples were collected by us in the field under proper permit (Fig. 3, Table 1 in ESM).

DNA was extracted using the DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany) and libraries were generated using NEBNext Ultra DNA Library Prep Kits for Illumina (New England Biolabs, Ipswich, Massachusetts). Fresh blood samples were extracted and prepped following the manufacturer's protocol with an additional RNase treatment applied in extraction. Meanwhile toepad samples

Fig. 2 Photographs of female shama specimens in comparison, taken at the Lee Kong Chian Natural History Museum (Singapore). Specimens are shown in ventral, dorsal and lateral orientation and are ordered from *left* to *right*: Langkawi shama from Tarutao Island (Thailand); *macrourus* from peninsular Thailand; *melanurus* from Sipora Island in the Mentawai group (Indonesia)

were extracted and prepped under sterile conditions inside a dedicated ancient DNA facility with slight modifications to the manufacturer's protocol as previously described (Chattopadhyay et al. 2019). Multiple extraction and library negatives were included to detect contamination. Libraries were sequenced on an Illumina Hiseq 4000 platform to produce 150-bp paired-end reads.

We removed adaptor sequences with cutadapt (Martin 2011) and eliminated reads that potentially constitute exogenous contamination using FastQScreen (Wingett and Andrews 2018) by removing reads that had mapped to a human genome (GenBank Assembly Accession: GCA_000001405.28), bacterial genomes and an *Aerodramus* swiftlet genome (unpublished data) because swiftlets were sequenced on the same Illumina lane. Mitogenomes were harvested by mapping the reads

to the complete mitogenome of an Oriental Magpie Robin *C. saularis* (GenBank Accession: NC_030603) using BWA-MEM (Li, 2013). Initial quality filtering was carried out as previously described (Wu et al. 2020). We employed ANGSD to call consensus fasta sequences for each sample using doFasta 3 (Korneliussen et al. 2014). The fasta sequences were aligned to a Seychelles Magpie Robin *C. sechellarum* partial mitogenome (GenBank Accession: MN356447.1) using Muscle (Edgar 2004). We trimmed 147 bp from the end of each sequence to remove a poor-quality stretch and obtained sequences 15,433 bp in length. We ran two types of phylogenetic analyses with MEGAX, neighbor-joining and maximum likelihood under the GTR+Inverse+Gamma model as identified by a model selection test implemented within the program (Kumar et al. 2018). We also annotated the partial mitogenome
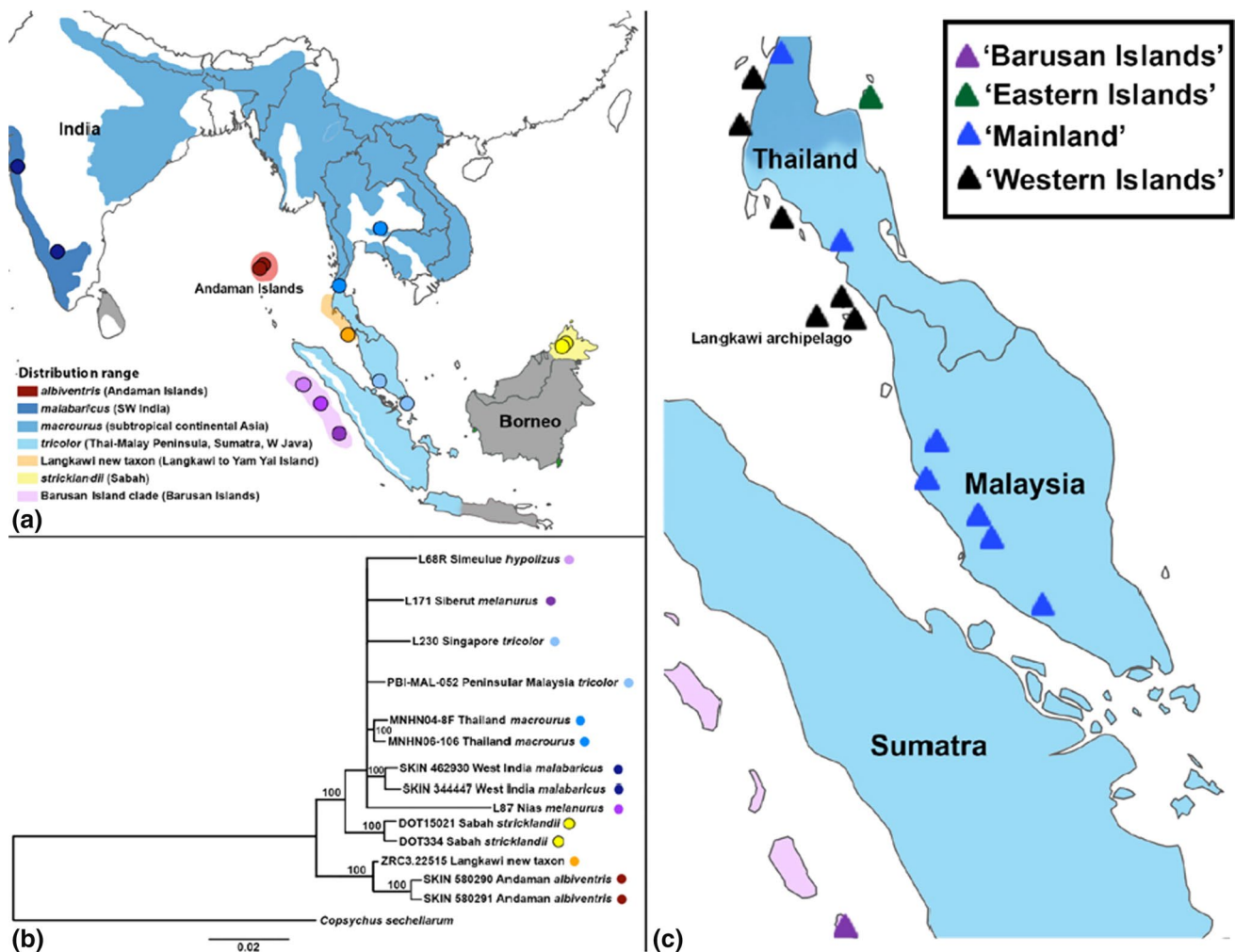
Fig. 3 a Collection localities of DNA samples ($n = 14$) and distribution ranges of taxa in the white-rumped shama *Copsychus malabaricus* complex. Distribution ranges of taxa not used in mitogenomic analysis given in grey; various smaller-island forms that are distant to the new taxon are omitted from the map. Barusan Island clade includes *melanurus*, *opisthochrus* and *hypolizus*. b Maximum likelihood topology for mitogenomic alignment (15,433 bp). Bootstrap values depicted are supported by bootstrap = 100 in both maximum likelihood and neighbor-joining analyses. Branch support values below 100 not provided. *Coloured dots* behind sample names refer to collection localities in (a). c Sample localities of museum specimens used in morphometric analysis and plumage inspection. Specimens with imprecise collection localities are not mapped. Colour coding for specimens divided into four populations: (1) 'Barusan Islands' refers to specimens collected on the island of Sipora in the Barusan group ($n = 2$); (2) 'Eastern Islands' refers to specimens collected on the island of Ko Samui east of the Thai Malay Peninsula ($n = 2$); 'Mainland' refers to specimens collected on the Thai Malay Peninsula ($n = 18$); and (4) 'Western Islands' refers to specimens collected on islands west of the Thai Malay Peninsula ($n = 15$), comprising the range of the new taxon

using MITOS (Bernt et al. 2013) to extract cytochrome c oxidase subunit I (COI) and cytochrome b (cyt-*b*) sequences. We then calculated uncorrected *p* distances in MEGAX separately for the mitogenome, COI and cyt-*b*.

Morphometric measurements and plumage inspection

We measured 37 specimens of white-rumped shama deposited in the Raffles Collection of the Lee Kong Chian Natural History Museum, Singapore, from localities across the Thai Malay Peninsula that mostly fall within the range of *tricolor* and the intergradation zone between *tricolor* and *macrourus* from the Isthmus of Kra (Fig. 3c, Table 2 in ESM). We also measured two specimens from the island of Sipora in the Mentawai archipelago off western Sumatra, also known as the Barusan Islands (Fig. 3c, Table 2 in ESM). Specimens were measured for flattened wing length (from the carpal joint to the tip), upper mandible length (from the bill tip to where the mandible meets the skull), tail length (from the longest rectrix tip to the uropygial gland) and tarsus length (from tibiotarsus). All specimens were inspected

for plumage coloration, with the following representative specimens being used in direct comparison to the holotype (with voucher numbers in brackets): *suavis* (ZRC3.22503, ZRC3.22498, ZRC3.22507), *barbouri* (ZRC3.22599), *stricklandii* (ZRC3.22488), *omissus* (ZRC3.22482), and *javanus* (ZRC3.4175). Biometric analyses were divided by sex and the different populations were compared. We carried out both a Wilcoxon test with Bonferroni correction and Kruskal–Wallis test in R to identify any traits that might be significantly different among populations (R Core Team 2021). Localities of the specimens were also noted down.

Results

We here describe the new taxon of the *Copsychus malabaricus* complex from the Langkawi archipelago and islands to the north at the subspecies level:

Copsychus malabaricus ngae, subspecies nova
   English name: Langkawi Shama
   Malay name: Murai Batu Langkawi
   Zoobank identifier: urn:lsid:zoobank.org:act:0329091A-C8BC-4671-BBCD-2E3E59A5EA0A

Holotype

ZRC3.22515 (Fig. 1); adult male collected on 9 Dec 1916 on Dayang Bunting Island south of Langkawi Island, Kedah, Malaysia, deposited in the Raffles Collection of the Lee Kong Chian Natural History Museum, Singapore.

Description of holotype

Colour definitions follow the Munsell Colour Wheel, with colour identification codes given within brackets (http://www.andrewwerth.com/color/). Head, nape and mantle black (N1). Throat black (N1), extending down to the upper breast but not as far as nearby peninsular populations. Lower breast is rufous (2.5YR 4/10) and gradually lightens towards the vent. Discrete dividing line between black on upper and rufous on lower breast. Rump and thigh feathers white (N9). Wings are black (N1) turning dark brownish (7.5YR 2/2) on major remiges, which is likely due to post-mortem fading. Tail has eight white (N9) inner rectrices distinctly shorter than the 2 + 1 dark-brown outermost rectrices. One dark outer rectrix seems to be lost on the holotype. The outermost rectrices' brown colouration resembles that of the remiges (7.5YR 2/2) and is probably black in life with post-mortem fading based on comparison with photos of live individuals. The two outermost dark brown rectrices are roughly twice as long as the white inner rectrices, while the second outermost dark rectrix is intermediate in length. Tarsus on the

specimen is dark orange (5YR 5/10), probably with considerable post-mortem discolouration. Mandible grey (N3), likely black in life with post-mortem fading. Wing 94 mm, mandible 22.7 mm, tarsus 23.7 mm and tail 186 mm.

Diagnosis

Individuals from the island group west of the Thai Malay Peninsula, roughly from the Langkawi Island group northwards almost to the Thai–Burmese border, exhibit pronounced morphological differences from adjacent peninsular populations to the east, from the range of *tricolor*, *macrourus* and the intergradation zone between *tricolor* and *macrourus*, warranting recognition as a discrete subspecies *ngae*.

Biometrically, male individuals of *ngae* are comparable to males of adjacent populations attributable to *tricolor, macrourus* and intergrades between *tricolor* and *macrourus* in all parameters except tail length (Figs. 1–2, 4; Table 1). The male tail length of the new subspecies is significantly longer compared to these peninsular mainland populations (Fig. 4; Table 1). Tail length differences are so pronounced that they are instantly discernible even on cursory inspection of specimen series, with extremely limited overlap (Fig. 4). Only the tail length of a nameless insular population east of the peninsula (from Ko Samui) comes close to that of *ngae* (Fig. 4; Table 1); given our low sample size of Ko Samui individuals ($n = 2$), tail length differences between Ko Samui and *ngae* did not reach statistical significance level, but the tail lengths of both Ko Samui specimens were at the lower end of the range of *ngae*'s tail lengths, with 67% of *ngae* specimens exceeding the tails of both Ko Samui samples.

Even in comparison with all other subspecies of white-rumped shama, including those that are geographically more distant, *ngae* generally exhibits the longest tail lengths in both sexes (Table 2; Rasmussen and Anderton 2012; Kloss 1921; Hartert 1902; Vorderman 1893; Salvadori 1886; Sclater 1861). Only *malabaricus* and *leggei* from the Indian subcontinent come close, approaching the lower bounds of *ngae*'s tail length variation (Table 2; Rasmussen and Anderton 2012).

In male plumage, *ngae* is separated from all 14 other taxa of white-rumped shama by a much lesser extent of black on the breast, which does not reach as far down onto the lower breast (Fig. 1). Importantly, even the two relatively long-tailed insular male individuals from Ko Samui to the east of the Peninsula show black anterior underparts extending far down onto the lower breast akin to mainland birds, setting them apart from the lesser extent of black in *ngae*, and arguing against their inclusion as a disjunct insular population of *ngae* (Fig. 1).

Additionally, *ngae* differs from *tricolor* and *macrourus* on the mainland and on Ko Samui in its white rather than rufous thigh feathers (although some geographically
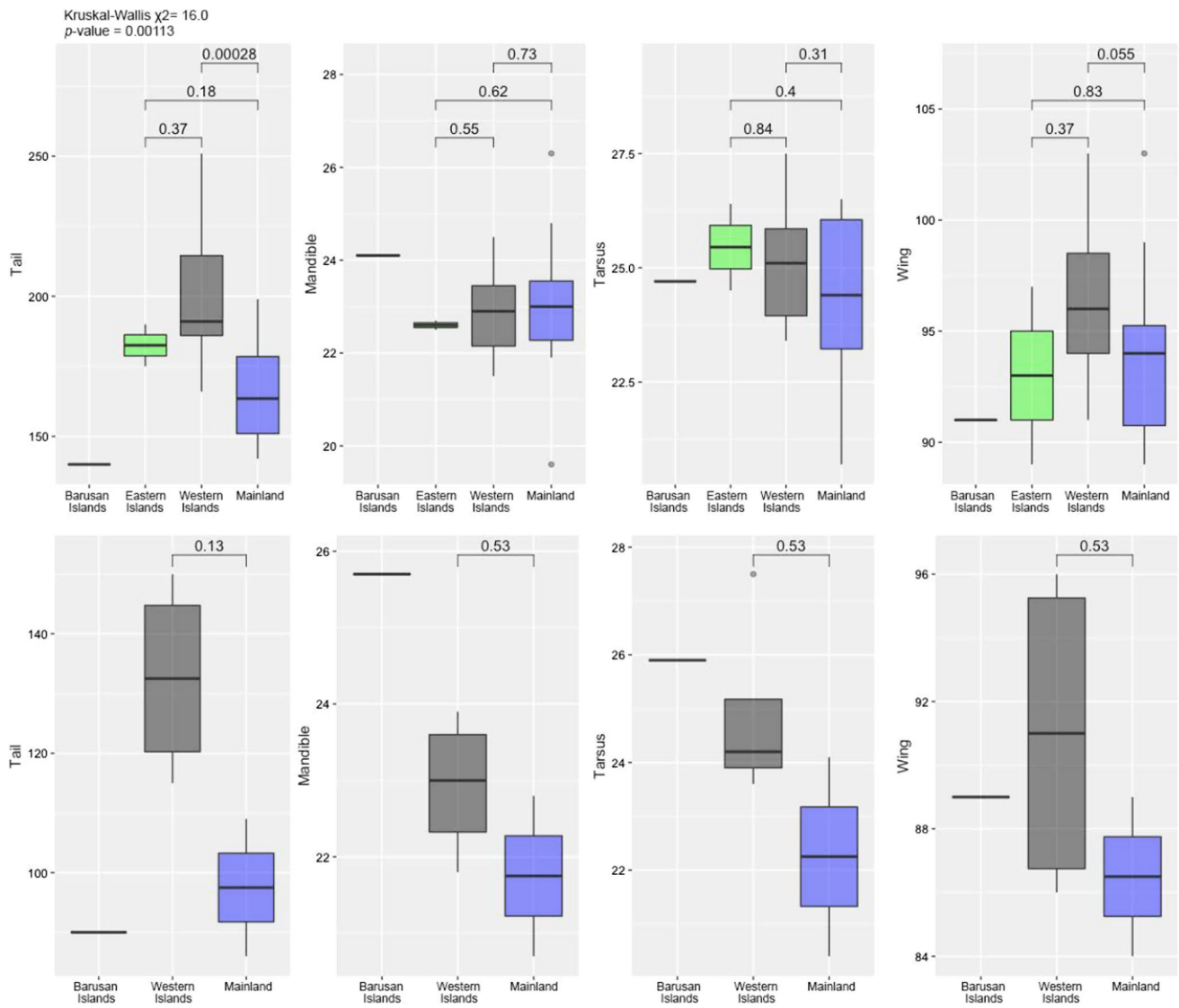
Fig. 4 Box plots of morphometric measurements for males (top row) and females (bottom row). The *p* values are shown above the plots for each pairwise Wilcoxon test comparison (except comparisons involving the Barusan Islands, for which sample size was too low).

Kruskal–Wallis test results are only shown for significant *p* values. Population labels (e.g., 'Eastern Islands', 'Western Islands' etc.) follow those in Fig. 3c, with 'Western Islands' referring to the new taxon

distant individuals of *macrourus* from north of the Isthmus of Kra [e.g., ZRC 3.22438 from Pak Jong] also possess white thigh feathers). Further afield, *suavis* and *barbouri* are also separated from *ngae* by their rufous rather than white thighs. The two distinct Bornean taxa *stricklandii* and *barbouri* are set apart from *ngae* by their bright white (not black) crowns (Lim et al. 2017; Eaton et al. 2021). Moreover, *barbouri, mirabilis, nigricauda, opisthochrus, hypolizus* and *melanurus* differ in their complete or near-complete lack of white colour on the rectrices (Figs. 1, 2; Rheindt et al. 2019; Eaton et al. 2021). In addition, the two eastern and central Javan taxa *omissus* and *javanus*

are characterised by much paler orange on the belly, often divided from an extensive black breast by a white band.

The Andaman shama *C. albiventris* is usually considered a distinct monotypic species (e.g., Roberts et al. 2020; Gill et al. 2021), but in our mtDNA analyses *ngae* unexpectedly emerged closer to the Andaman shama than to traditional white-rumped shamas (see below). In morphological terms, however, Andaman shamas differ distinctly from *ngae*, as well as from all other taxa considered herein, in their extensively white (not rufous or orange) lower breast and belly.

Table 1 Ranges of measurements of regional populations in the *Copsychus malabaricus* complex taken from Lee Kong Chian Natural History Museum specimens, followed by means and standard deviation in brackets

Population	n	Sex	Tail	Wing	Upper mandible	Tarsus
Eastern Islands	2	M	175–190 (183 ± 10.6)	89–97 (93 ± 5.66)	22.5–22.7 (22.6 ± 0.141)	24.5–26.4 (25.45 ± 1.34)
Western Islands	11	M	166–251 (202 ± 25.3)	91–103 (96.4 ± 3.33)	21.5–24.5 (22.9 ± 0.938)	23.4–27.5 (25.1 ± 1.37)
Barusan Islands	1	M	140	91	24.1	24.7
Thai-Malay Peninsula	16	M	142–199 (166 ± 16.1)	89–103 (93.9 ± 3.83)	19.6–26.3 (23.0 ± 1.46)	20.7–26.5 (24.3 ± 1.96)
Western Islands	4	F	115–150 (133 ± 16.7)	86–96 (91 ± 5.23)	21.8–23.9 (22.9 ± 0.954)	23.6–27.5 (24.9 ± 1.78)
Barusan Islands	1	F	90	89	25.7	25.9
Thai-Malay Peninsula	3	F	86–109 (97.5 ± 16.3)	84–89 (86.5 ± 3.54)	20.7–22.8 (21.75 ± 1.48)	20.4–24.1 (22.25 ± 2.62)

Population labels (e.g., 'Eastern Islands', 'Western Islands' etc.) follow those in Fig. 3c, with 'Western Islands' referring to the new taxon. Measurements given in millimetres

M male, *F* female

Table 2 Ranges of tail measurements of taxa within the *Copsychus malabaricus* complex taken from the literature

Taxon	Sex	Tail	Source
malabaricus	M	177–215	Rasmussen and Anderton (2012)
	F	115–128	
macrourus	M	128–165	Rasmussen and Anderton (2012)
	F	103–112	
leggei	M	159–172	Rasmussen and Anderton (2012)
	F	115–119	
tricolor	M	145–175	Kloss (1921)
	F	105–126	
omissus	M	130–144	Kloss (1921)
	F	91–122	Hartert (1902)
javanus	M	138–142	Kloss (1921)
	F	102	
nigricauda	M	160	Vorderman (1893)
melanurus	M	150	Salvadori (1886)
suavis	M	114	Sclater (1861)
stricklandii	M	127	Motley and Dillwyn (1855)
barbouri	M	142–150	Bang and Peters (1927)
	F	112–115	

Measurements given in millimetres

M male, *F* female

Etymology

The new taxon is named in honour of Ng Chin Wei, who first made our research team aware of this distinct population by connecting us to a Malaysian birdkeeper in Penang named Mel Ivan Blake, who recounted stories of a highly-prized long-tailed variety of shama from Langkawi which is sold at inflated price levels in the markets of Penang and Kedah states, and who showed us video footage of this new shama illustrating its remarkable distinctness. The epithet *ngae* is in the genitive singular and is inalienable.

Phylogenetic analysis

Mitogenomic comparison of samples and taxa from a large geographic range confirmed the genetic distinctness of the Langkawi Shama (Fig. 3). Uncorrected mitochondrial pairwise divergence between the Langkawi Shama and all other members of the white-rumped shama complex sequenced exceeded 2.8% for the mitogenome (3.18% for COI, 3.68% for cyt-*b*), which is roughly consistent with the threshold that has been postulated in the literature as evidence of species level recognition of closely related bird species (Hebert et al. 2004; Kerr et al. 2007). In terms of mtDNA, *ngae* groups much closer with the Andaman shama with maximum branch support (Fig. 3).

Morphometric analysis

Tail measurements of the specimens on islands along the western coast of the Thai Malay Peninsula, comprising the postulated range of *ngae*, exhibited nearly non-overlapping ranges with all other white-rumped shama populations, and significantly so with the sample set from the mainland of the Thai Malay Peninsula (Fig. 4). The male wing measurements of *ngae* are also considerably longer than those from the peninsular mainland populations at a nearly significant level (Fig. 4).

Given the small sample size for female specimens ($n = 7$ in total; see Table 1 for breakdown), we refrained from statistical testing of morphometric differences in females.

Discussion

Taxonomic synthesis

We demonstrate that populations of white-rumped shamas from offshore islands to the west of the Thai Malay Peninsula exhibit pronounced morphological differences from populations on the nearby mainland. In terms of mtDNA, these insular populations are not even embedded within members of the white-rumped shama *C. malabaricus* as traditionally circumscribed, but they instead emerge as sister to the Andaman shama *C. albiventris*. We describe this population unit under a novel taxonomic epithet, *ngae*, but there is a legitimate question whether our placement of *ngae* as a subspecies of the white-rumped shama *C. malabaricus* is justified, or whether it should perhaps be elevated to species level or subsumed under the Andaman Shama *C. albiventris* as suggested by mtDNA (Fig. 3b).

In our assignment of taxonomic rank, we adopted an integrative approach that is primarily guided by the multidimensional Biological Species Concept (Mayr 1942, 1996, 2000; Bock 1986; Mayr and Ashlock 1991). Such integrative taxonomy uses multiple lines of evidence from different traits that may contribute to speciation, and has been widely used and found effective in delineating avian species (Helbig et al. 2002; Padial et al. 2010; Gill 2014; Sangster 2014, 2018; Ng et al. 2016; Gwee et al. 2017; Hosner et al. 2018).

Morphologically, the most distinguishing character between the Langkawi Shama and nearby peninsular populations is the tail length, sometimes growing almost twice as long in the former compared to the latter populations (Figs. 1, 2, 4). This phenotype leads to differences in courtship posture, especially in male individuals in full song, as attested to by breeders who highly value the Langkawi Shama on account of these peculiarities (M.I. Blake, pers. comm.). Noticeable differences in the extent of black on the breast additionally contribute to rendering *ngae* one of the most distinct taxa in the complex, although perhaps not quite as distinct as the two Bornean taxa with white rather than black crowns (*stricklandii* and *barbouri*).

Another strong line of evidence for *ngae*'s distinctness lies in its mitogenome. Our mitogenomic tree encompassed almost all recognised taxa from Southeast Asia in addition to the nominate *malabaricus* from India as well as the Andaman Shama *C. albiventris* from the Andaman Islands. The Langkawi shama emerged as most closely related to the Andaman shama on the basis of mtDNA (Fig. 3b), even though the latter lives across the Andaman Sea ~500–1000 km west of the island chain inhabited by *ngae* (depending on which island is taken as a reference). If mitogenomic evidences were taken at face value,

ngae would have to be incorporated as a subspecies of *C. albiventris*, or even upgraded to monotypic species status. However, phylogenetic reconstructions based solely on mtDNA are often subject to artifacts generated by genetic introgression (Rheindt and Edwards 2011), sometimes involving unsampled extinct lineages that can make the species tree impossible to reconstruct (Zhang et al. 2019). Unfortunately, our attempts to recover genomic DNA from the holotype's toepad beyond the mitogenome failed, even though we were more successful with the historic specimens of other shama taxa (data not shown), precluding an analysis that would entail genome-wide markers from the nucleus. At a minimum, we interpret the mitogenomic signature of the new taxon *ngae* to mean that it has undergone an unusual phylogenetic trajectory, and has not been in extensive recent gene flow with mainland populations that would have facilitated an influx of continental mtDNA into *ngae*.

In summary, our assignment of subspecies status to *ngae* despite considerable morphological and mitogenomic differences is a conservative approach. The difference in tail length may well play a role in sexual selection resulting in some degree of reproductive isolation contributing to the distinct mitogenome of *ngae* and suggesting possible species level differentiation. We also cannot completely rule out the possibility of *ngae* being more closely related to *C. albiventris* as reflected in mtDNA. There is a possibility that *ngae* could be the founding population of *C. albiventris* or share a common ancestor, possibly a dispersive super-tramp population that colonised both island chains. The change in plumage colouration in *C. albiventris* to a white belly may have been the result of mutations at few loci, leading to founder event speciation. Future analyses encompassing bio-acoustic data and DNA of all white-rumped shama taxa, hopefully including genome-wide markers, may well prove that *ngae* deserves to be recognized at the species level or reclassified under *C. albiventris*. In the meantime, it is hoped that its current assignment will not stand in the way of urgent conservation measures that may bring this terminally endangered taxon back from the brink of extinction.

Distribution range

On the basis of the diagnostic morphological traits described herein, the distribution range of the Langkawi shama encompasses—at a minimum—the islands on the western coast of the Thai Malay Peninsula from the Langkawi group near the Thai–Malaysian border in the south almost to the Thai–Burmese border in the north (Fig. 3c). Specimens from the islands of Tarutao, Phi Phi Don, Ra, Yam Yai (now part of Laem Son National Park) and other islands of the Langkawi archipelago exhibited the same morphological traits as the holotype.

The northern limit of the new taxon remains obscure. North of the Thai–Burmese border lies the Mergui archipelago, a dense agglomeration of ~800 small islands off the coast of Tanintharyi (formerly Tenasserim). Unfortunately, we had no access to specimen material from these islands, and it is doubtful whether such museum material exists in any considerable quantities. For the time being, Yam Yai Island—one of the northernmost islands of Thailand's west coast—should be considered the northern boundary of *ngae* until data from Myanmar become available (Fig. 3c).

There is a noteworthy specimen from Tanjong Andu, Dinging (=present day Manjung in the surroundings of the city of Sitiawan, Perak, Malaysia) (ZRC3.22527) with an exceptional tail length of 199 cm, consistent with the range of variation of the Langkawi Shama (Table 2 in ESM). We have no further information on the collection locality of this specimen, but conservatively attributed it to the mainland population. The word 'tanjong' is the Malay term for 'cape' or 'land spit', certainly placing this locality on the coast. It is possible, and indeed likely, that the specimen was collected on Pangkor Island, which is part of Manjung District. If so, the range of the Langkawi shama may extend south of Langkawi Island down the Malay Peninsula to islands off Perak. This hypothesis is somewhat contradicted by a lack of evidence of Langkawi shamas on the well-studied island of Penang, which is geographically in between Langkawi and Manjung District. Alternatively, specimen ZRC3.22527 may have been mislabeled, or may be an example of a true outlier of continental birds in terms of tail length. Unfortunately, its breast colouration is somewhat intermediate in the extent of black between *ngae* and *tricolor*, adding little resolution to this open question.

Conservation status

There is a distinct lack of information about the Langkawi shama's survival in the wild. White-rumped shamas are highly traded in the region (Nash 1993; Shepherd and Cassey 2017), perhaps slightly more extensively in Malaysia than in Thailand, and the long tail of the Langkawi shama renders it even more desirable and susceptible to poaching, and has created a veritable market hype. At the time of writing, Penang market prices for Langkawi shamas with a moderate tail length (~175–180 mm) are approximately US$ 2000, whereas individuals with tails around 200 mm can command prices of up to US$ 4000, and males along the upper spectrum of tail lengths (>230 cm) sell for a minimum of US$ 5000 (M.I. Blake, pers. comm.).

Very little modern ornithological exploration has been carried out on the islands that form the Langkawi shama's range. Based on verbal accounts, the bird can no longer be found on the main island of Langkawi itself (M.I. Blake, pers. comm.), where habitat destruction in the face of growing tourism and other development would have acted in concert with extreme poaching pressure to exterminate the population (Samat 2010). Most wild-caught birds in the Malaysian trade are said to be sourced from tiny offshore islands around Langkawi, such as our holotype specimen from Pulau Dayang Bunting, or may have been smuggled across from Thailand. The population status in Thailand is unknown, but potentially lower levels of poaching pressure coupled with the presence of many more islands raise hopes that Langkawi shamas may still occur in healthy numbers here. For instance, Tarutao and several other islands have been designated as national parks in the past few decades.

Future surveys, especially in Thailand, coupled with an assessment of any shama populations on the Mergui archipelago in Myanmar, are urgently required to assess the threat status of this newly described bird. At worst, the bulk of the taxon's population may be found in illegal captivity, which would not bode well for its survival, as Langkawi shamas are exceedingly rare in the market and most captive individuals would be cross-paired with members of other shama taxa to produce hybrid offspring. A survey of its population size is needed now to implement long-lasting conservation measures to preserve this new taxon in perpetuity.

References

Bang O, Peters JL (1927) Birds from Maratua Island, off the east coast of Borneo. Occas Papers Boston Soc Nat Hist 5:235–242

Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsch G, Pütz J, Middendorf M, Stadler PF (2013) MITOS: improved de novo metazoan mitochondrial genome annotation. Mol Phylogenet Evol 69:313–319

Bock WJ (1986) Species concepts, speciation and macroevolution. In: Iwatsuki K, Raven PH, Bock WJ (eds) Modern aspects of species. University of Tokyo Press, Tokyo, pp 31–57

Chattopadhyay B, Garg KM, Mendenhall IH, Rheindt FE (2019) Historic DNA reveals Anthropocene threat to a tropical urban fruit bat. Curr Biol 29:R1299–R1300

Clements JF, Schulenberg TS, Iliff MJ, Billerman SM, Fredericks TA, Sullivan BL, Wood CL (2019) The eBird/Clements Checklist of Birds of the World: v2019. https://www.birds.cornell.edu/clementschecklist/download/. Accessed 19 September 2020

Eaton JA, Shepherd CR, Rheindt FE, Harris JBC, van Balen S, Wilcove DS, Collar NJ (2015) Trade-driven extinctions and near-extinctions of avian taxa in Sundaic Indonesia. Forktail 31:1–12

Eaton JA, van Balen B, Brickle NW, Rheindt FE (2021) Birds of the Indonesian Archipelago. Greater Sundas and Wallacea, 2nd edn. Lynx Edicions, Barcelona

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797

Gill FB (2014) Species taxonomy of birds: which null hypothesis? Auk 131:150–161. https://doi.org/10.1642/AUK-13-206.1

Gill F, Donsker D, Rasmussen P (2021) IOC World Bird List (v11.1). http://www.worldbirdnames.org/new/ioc-lists/master-list-2/ Accessed 19 September 2020. Doi: https://doi.org/10.14344/IOC.ML.11.1

Gwee CY, Christidis L, Eaton JA, Norman JA, Trainor CR, Verbelen P, Rheindt FE (2017) Bioacoustic and multi-locus DNA data of *Ninox* owls support high incidence of extinction and recolonisation on small, low-lying islands across Wallacea. Mol Phylogenet Evol 109:246–258

Hartert E (1902) On the birds from Pahang Eastern Malay Peninsula. Novitates Zoologicae 9:572

Handbook of the Birds of the World & BirdLife International (2019) Handbook of the Birds of the World and BirdLife International digital checklist of the birds of the world. Version 4. http://datazone.birdlife.org/userfiles/file/Species/Taxonomy/HBW-BirdLife_Checklist_v4_Dec19.zip. Accessed 19 September 2020

Hebert PDN, Stoeckle MY, Zemlak TS, Francis CM (2004) Identification of birds through DNA barcodes. PLoS Biol 2:e312

Helbig AJ, Knox AG, Parkin DT, Sangster G, Collinson M (2002) Guidelines for assigning species rank. Ibis 144:518–525

Hosner PA, Campillo LC, Andersen MJ, Sánchez-González LA, Oliveros CH, Urriza RC, Moyle RG (2018) An integrative species delimitation approach reveals fine-scale endemism and substantial unrecognized avian diversity in the Philippine Archipelago. Conserv Genet 19:1153–1168

Kerr KCR, Stoeckle MY, Dove CJ, Weigt LA, Francis CM, Hebert PDN (2007) Comprehensive DNA barcode coverage of North American birds. Mol Ecol Notes 7:535–543

Kloss CB (1921) New and Known Oriental birds. J Federated Malay States Mus 10:207–213

Korneliussen TS, Albrechtsen A, Nielsen R (2014) ANGSD: analysis of next generation sequencing data. BMC Bioinformatics 15:356

Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35:1547

Li H (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv preprint arXiv:1303.3997.

Lim HC, Gawin DF, Shakya SB, Harvey MG, Rahman MA, Sheldon FH (2017) Sundaland's east–west rain forest population structure: variable manifestations in four polytypic bird species examined using RAD-Seq and plumage analyses. J Biogeogr 44:2259–2271

Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. Embnet J 17:10–12

Mayr E (1942) Systematics and the origin of species, from the viewpoint of a zoologist. Harvard University Press, Cambridge

Mayr E (1996) What is a species and what is not? Philos Sci 63:262–277

Mayr E (2000) The multidimensional biological species concept. In: Wheeler QD, Meier R (eds) Species concepts and phylogenetic theory: a debate. Columbia University Press, New York

Mayr E, Ashlock P (1991) Principles of systematic zoology, 2nd edn. McGraw-Hill, New York

Motley J, Dillwyn LL (1855) Contributions to the natural history of Labuan, and the adjacent coasts of Borneo. J. Van Voorst, London

Nash SV (1993) Sold for a song: the trade in Southeast Asian non-CITES birds. Cambridge, UK, TRAFFIC International

Ng EYX, Eaton JA, Verbelen P, Hutchinson RO, Rheindt FE (2016) Using bioacoustic data to test species limits in an Indo-Pacific island radiation of *Macropygia* cuckoo doves. Biol J Lin Soc 118:786–812

Ng EYX, Garg KM, Low GW, Chattopadhyay B, Oh RRY, Lee JGH, Rheindt FE (2017) Conservation genomics identifies impact of trade in a threatened songbird. Biol Cons 214:101–108

Padial JM, Miralles A, De la Riva I, Vences M (2010) The integrative future of taxonomy. Front Zool 7:16

Peters JL (1964) Check-list of the birds of the world, vol 10. Harvard University Press, Cambridge, pp 69–72

R Core Team (2021) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Rasmussen PC, Anderton JC (2012) Birds of South Asia. The Ripley Guide, vol 2. Lynx Edicions, Barcelona

Rheindt FE, Edwards SV (2011) Genetic introgression: an integral but neglected component of speciation in birds. Auk 128:620–632

Rheindt FE, Baveja P, Ferasyi TR, Nurza A, Rosa TS, Haminuddin Ramadhan R, Gwee CY (2019) The extinction-in-progress in the wild of the Barusan Shama *Copsychus [malabaricus] melanurus*. Forktail 35:30–37

Rheindt FE, Gwee CY, Baveja P, Ferasyi TR, Nurza A, Rosa TS, Haminuddin (2020) A taxonomic and conservation re-appraisal of all the birds on the island of Nias. Raffles Bull Zool 68:496–528

Roberts GE, Male TD, Conant S (2020) White-rumped Shama (Copsychus malabaricus), version 1.0. In: Billerman SM (ed) Birds of the world. Ithaca, NY, USA, Cornell Lab of Ornithology

Salvadori T (1886) Catalogo delle collezioni Ornitologiche fatte presso Siboga in Sumatra, e nell'Isola Nias. Ann Mus Civ Stor Nat Genova 24:549

Samat N (2010) Assessing land use land cover changes in Langkawi island: towards sustainable urban living. Malays J Environ Manag 11:48–57

Sangster G (2014) The application of species criteria in avian taxonomy and its implications for the debate over species concepts. Biol Rev 89:199–214

Sangster G (2018) Integrative taxonomy of birds: the nature and delimitation of species. In: Tietze DT (ed) Bird species: how they arise, modify and vanish. Fascinating life sciences series. Cham, Springer

Sclater PL (1861) On a new species of the genus *Copsychus* from Borneo. Proc Zool Soc Lond 1861:185–187

Shepherd CR, Cassey P (2017) Songbird trade crisis in Southeast Asia leads to the formation of IUCN SSC Asian Songbird Trade Specialist Group. J Indonesian Nat Hist 5:3–5

Symes WS, Edwards DP, Miettinen J, Rheindt FE, Carrasco LR (2018) Combined impacts of deforestation and wildlife trade on tropical biodiversity are severely underestimated. Nat Commun 9:4052

Vorderman AG (1893) Bijdrage tot de kennis der vogels van den Kangean-archipel. In: Koninklijke Natuurkundige Vereeniging in Nederlandsch Indië, Natuurkundig tijdschrift voor Nederlandsch Indië, vol. 53. Batavia: Lange.

Wingett SW, Andrews S (2018) FastQ Screen: a tool for multi-genome mapping and quality control. F1000Research 7.

Wu MY, Low GW, Forcina G, van Grouw H, Lee BPYH, Oh RRY, Rheindt FE (2020) Historic and modern genomes unveil a

domestic introgression gradient in a wild red junglefowl population. Evol Appl 13:2300–2315. https://doi.org/10.1111/eva.13023

Zhang D, Tang L, Cheng Y, Hao Y, Xiong Y, Song G, Qu Y, Rheindt FE, Alström P, Jia C, Lei F (2019) 'Ghost introgression' as a cause of deep mitochondrial divergence in a bird species complex. Mol Biol Evol 36:2375–2386