For: Systematic Entomology

One and the same: integrative taxonomic evidence that *Bactrocera invadens* (Diptera: Tephritidae) is the same species as the Oriental Fruit Fly *Bactrocera dorsalis*

Schutze, M.K.[1,2], Mahmood, K.[3], Pavasovic, A.[1], Bo, W.[4,5], Newman, J.[1], Clarke, A.R.[1,2], Krosch, M.N.[6], & Cameron, S.L.[1]

[1]Queensland University of Technology, Brisbane, Queensland, Australia

[2]Plant Biosecurity Cooperative Research Centre, Canberra, Australian Capital Territory, Australia

[3]Pakistan Museum of Natural History, Garden Avenue, Islamabad, Pakistan

[4]Beneficial Insects Institute, Fujian Agricultural and Forestry University, Fuzhou, Fujian Province, China

[5]FAO/IAEA Division of Nuclear Techniques in Food and Agriculture, Vienna and Seibersdorf, Austria

[6] Centre for Water in the Minerals Industry, University of Queensland, St Lucia, Queensland, Australia

Short title: Integrative taxonomy of *B. invadens* & *B. dorsalis*

Abstract

The invasive fruit fly *Bactrocera invadens* Drew, Tsuruta & White, and the Oriental fruit fly *Bactrocera dorsalis* (Hendel) are highly destructive horticultural pests of global significance. *Bactrocera invadens* originates from the Indian subcontinent and has recently invaded all of sub-Saharan Africa, while *B. dorsalis* principally occurs from the Indian subcontinent toward southern China and Southeast Asia. High morphological and genetic similarity has cast doubt over whether *B. invadens* is a distinct species from *B. dorsalis*. Addressing this issue within an integrative taxonomic framework, we sampled from across the geographic distribution of both taxa and: (i) analysed morphological variation, including those characters considered diagnostic (scutum colour, length of aedeagus, width of post-sutural lateral vittae, wing size, and wing shape); (ii) sequenced four loci (ITS1, ITS2, *cox1*, and *nad4*) for phylogenetic inference; and (iii) generated a *cox1* haplotype network to examine population structure. Molecular analyses included the closely related species, *Bactrocera kandiensis* Drew & Hancock. Scutum colour varies from red-brown to fully black for individuals from Africa and the Indian subcontinent. All individuals east of the Indian subcontinent are black except for a few red-brown individuals from China. The post-sutural lateral vittae width of *B. invadens* is narrower than *B. dorsalis* from eastern Asia, but the variation is clinal with sub-continent *B. dorsalis* populations intermediate in size. Aedeagus length, wing shape, and wing size cannot discriminate the two taxa. Phylogenetic analyses failed to resolve *B. invadens* from *B. dorsalis*, but did resolve *B. kandiensis*. *Bactrocera dorsalis* and *B. invadens* shared *cox1* haplotypes, yet the haplotype network pattern does not reflect current taxonomy or patterns in thoracic colour. Some individuals of *B. dorsalis/B. invadens* possessed haplotypes more closely related to *B. kandiensis* than to conspecifics, suggestive of mitochondrial introgression between these species. The combined evidence fails to support the delimitation of *B. dorsalis* and *B. invadens* as separate biological species. Consequently, existing biological data for *B. dorsalis* may be applied to the invasive population in Africa. Our recommendation, in line with other recent publications, is that *B. invadens* be synonymised with *B. dorsalis*.

Background

Fruit flies of the sub-family Dacinae (Diptera: Tephritidae) include some of the world's most important horticultural pests (White & Elson-Harris, 1992). Within Dacinae, species of the genus *Bactrocera* Macquart (Drew & Hancock, 2000) have diversified prolifically in the Southeast Asian and Pacific regions over the last 40 million years (Drew & Hancock, 2000; Krosch *et al.*, 2012). To differentiate diversity in this species-rich genus, it has been divided taxonomically into 22 subgenera and over 20 species complexes (informal species groups within subgenera) (Drew, 1989). The best known is the Oriental Fruit Fly, *Bactrocera* (*Bactrocera*) *dorsalis* (Hendel) complex, because it includes the most widely distributed and damaging pest species in the genus (Drew, 1989; Clarke *et al.*, 2005).

The *B. dorsalis* species complex (hereafter the '*dorsalis* complex') contains over 100 taxa that share a defined set of morphological characters, principally a mostly black scutum and abdominal terga III-V with a medial longitudinal dark band and variable dark patterns on the lateral margins (Drew & Hancock, 1994; Drew & Romig, 2013). While most members of the complex are readily identifiable and of little to no economic importance, the recently described *B. invadens* Drew, Tsuruta & White is morphologically very similar to *B. dorsalis*, and with similar economic pest status. This species was first detected in Africa in 2003 and has since become a destructive and highly invasive member of the complex, attacking over 40 fruit species and recorded from more than 30 African countries (Lux *et al.*, 2003; Goergen *et al.*, 2011; Khamis *et al.*, 2012).

When first reported in Kenya, *B. invadens* was considered an "unusually variable" invasive population of *B. dorsalis* (Lux *et al.*, 2003: p. 358). These flies were initially identified as *B. dorsalis* because they were collected in methyl eugenol baited traps (no other African tephritid was known to respond to methyl eugenol) and they possessed morphological characters consistent with *B. dorsalis* (Lux *et al.*, 2003). The African fly was considered a new species and named *B. invadens* following examination of specimens of the same *B. dorsalis*-like species from Sri Lanka, the purported native range (Drew *et al.*, 2005; Drew *et al.*, 2007). According to the formal taxonomic description by Drew

et al. (2005) and a recent major revision of tropical fruit flies by Drew & Romig (2013), *B. invadens* is distinguished from *B. dorsalis* by: 1) a mostly dark orange-brown scutum with a dark fuscous to black lanceolate pattern, 2) a longer aedeagus, 3) a scutum with narrower postsutural vittae, 4) a dark transverse band on the abdominal tergite III which broadly reaches tergite IV, and 5) a dark antereolateral marking on abdominal tergite V extended mesally. The abdominal characters are not referred to in Drew & Romig (2013), with scutum colour, aedeagus length, and postsutural lateral vittae the only diagnostic features provided. These morphological characters are, however, sufficiently variable to render some individuals of *B. invadens* virtually inseparable from *B. dorsalis* (Drew *et al.*, 2005). The question therefore remains: how reliable are diagnostic characters of *B. invadens* in distinguishing it from *B. dorsalis*? And if not, is *B. invadens* a distinct species?

Due to its economic impact, most studies on *B. invadens* have an applied focus on host use, seasonality and invasion dynamics (Mwatawala *et al.*, 2006; Ekesi *et al.*, 2007; Rwomushana *et al.*, 2008; Khamis *et al.*, 2009; De Meyer *et al.*, 2010), temporal occurrence and comparative demographic parameters (Vayssières *et al.*, 2005; Salum *et al.*, 2014), interactions with other fruit fly species and their parasitoids (Mohamed *et al.*, 2008; Ekesi *et al.*, 2009; Rwomushana *et al.*, 2009; Van Mele *et al.*, 2009), and the development of market access protocols (Grout *et al.*, 2011; Hallman *et al.*, 2011). These considerable research efforts are based on the assumption that *B. invadens* is a biologically distinct species from *B. dorsalis*, a fundamental issue which is receiving increased attention. If *B. invadens* and *B. dorsalis* are the same species, the considerable existing regulatory arrangements and literature on *B. dorsalis* may be applied to the invasive population in Africa.

Of those studies investigating the biological relationship between *B. invadens* and *B. dorsalis*, results show that: 1) aedeagi of *B. invadens* from Sri Lanka are significantly longer than those of *B. dorsalis* from Taiwan (Drew *et al.*, 2008), 2) male pheromone constituents following methyl eugenol feeding between *B. dorsalis* and *B. invadens* are identical (Tan *et al.*, 2011), 3) there exists extremely low wing-morphometric differences between these two species together with the lowest estimate of evolutionary divergence between *B. dorsalis* and *B. invadens* following mitochondrial DNA analysis among multiple taxa (including *Bactrocera kandiensis* Drew & Hancock, another *dorsalis*-complex

95 species) (Khamis *et al.*, 2012), 4) molecular analyses across a range of tephritid taxa have found no
96 significant genetic differentiation between *B. invadens* and *B. dorsalis* (Frey *et al.*, 2013; Leblanc *et*
97 *al.* 2013; San Jose *et al.* 2013), and 5) *B. invadens* and *B. dorsalis* are fully sexually compatible as
98 demonstrated by random mating and viable offspring to the second hybrid generation (Bo *et al.*,
99 2014). Despite mounting evidence supporting their conspecificity, a recent major revision of South-
100 east Asian fruit flies maintains *B. invadens* as a valid species which is no longer considered a member
101 of the *dorsalis* complex (Drew & Romig, 2013).

102 Given that morphological characters based on a limited amount of material collected from Africa
103 (Kenya, Benin, Cameroon, and Uganda) and Sri Lanka (Asia) were the only features used to
104 originally separate *B. invadens* from *B. dorsalis* (Drew *et al.*, 2005), we re-examined purportedly
105 diagnostic characters of both taxa from across a much wider geographic range to determine if
106 variation was indeed discontinuous and supportive of two biologically distinct species or in fact
107 continuous and supportive of a single morphologically variable species. We therefore focussed on the
108 characters previously used to differentiate the two putative species most recently by Drew & Romig
109 (2013) (*i.e.*, scutum colour, width of postsutural lateral vittae, and aedeagus length). Additionally, we
110 applied geometric morphometric wing shape analysis due to its demonstrated capacity in resolving
111 fine-scale variation between real and putative cryptic insect taxa, as well as intraspecific population
112 structure (Aytekin *et al.*, 2007; Schutze *et al.*, 2012b; Krosch *et al.*, 2013).

113 We generated genetic datasets in addition to morphological and morphometric analyses, because
114 morphologically identical populations may consist of multiple cryptic biological species (Bickford *et*
115 *al.*, 2007). Using two nuclear and two mitochondrial loci, which have proven discriminatory power
116 for cryptic taxa within the *dorsalis* complex (Boykin *et al.*, 2014), we carried out Bayesian and
117 Maximum Likelihood phylogenetic analyses to test if samples of *B. invadens* and *B. dorsalis* form
118 distinct and well supported clades as predicted for two species, or if individuals of *B. invadens* emerge
119 unresolved within a larger *B. dorsalis* clade as predicted for one species. Moreover, we used one
120 mtDNA locus (*cox1*) to construct a minimum spanning haplotype network; this form of analysis is
121 well suited to inferring intraspecific relationships (Bandelt *et al.*, 1999). Wherever possible, we used

the same individuals as for morphological analysis to strengthen conclusions drawn within an integrative taxonomic framework (Schlick-Steiner *et al.*, 2010; Yeates *et al.*, 2011).

Samples of *Bactrocera papayae* Drew & Hancock from our previous work on the *dorsalis* complex from the Indo/Malay Archipelago (Schutze *et al.*, 2012b; Krosch *et al.*, 2013) were included in analyses due to considerable evidence that *B. papayae* is the same biological species as *B. dorsalis* (Perkins *et al.*, 1990; Medina *et al.*, 1998; Tan, 2003; Mahmood, 2004); hence inclusion of this material in a geographically wide-ranging study involving *B. dorsalis* is justified. We interpret our results in the context of the unified species concept (*sensu* de Queiroz, 2007), for which no single species character (*e.g.*, mate compatibility, genetic divergence, or morphological difference) is relied upon for their delimitation; rather, multiple lines of data are independently analysed to evaluate evidence for, or against, separately evolving metapopulation lineages. We discuss our findings within the context of the taxonomic history of *B. dorsalis*, particularly the relationship between these taxa and *Dacus ferrugineus* Fabricius, a species described in the late 18th Century and a junior synonym of *B. dorsalis.*

Materials and methods

Specimens

Twenty individuals from each of 13 locations were examined for morphological variation (n = 260). Of these, 200 were newly acquired for this study (Table 1) and combined with 60 specimens from Thailand, Taiwan, and Malaysia that were part of previous studies (Schutze *et al.*, 2012b; Krosch *et al.*, 2013). Molecular analyses included 94 newly acquired specimens (Table 1) which were combined with 312 individuals from Taiwan, Thailand, Philippines, Indonesia, and Malaysia which formed the study of Boykin *et al.* (2014). Males were used for all analyses because they are readily caught using lure-traps to which females do not respond. All voucher specimens are at the Queensland University of Technology, Brisbane, Australia.

African samples were collected from Benin, the Democratic Republic of Congo, Mozambique, Sudan, and Kenya. Specimens from all locations except Kenya were collected from the wild between 2005

148 and 2011 using methyl eugenol traps (Table 1) and supplied via Marc De Meyer of the Royal
149 Museum for Central Africa, Tervuren, Belgium. Kenyan samples were sourced from a third-
150 generation laboratory colony maintained at the UN/FAO-International Atomic Energy Agency
151 (IAEA) Insect Pest Control Laboratory (IPCL) (Seibersdorf, Austria), initiated in March 2012.

152 Indian subcontinent samples were collected from Pakistan, Nepal, India, and Sri Lanka. Specimens
153 from all locations except Dahanu (India) were methyl eugenol trap-collected from the wild between
154 1992 and 2012 (Table 1). Samples from Dahanu were sourced from a first-generation laboratory
155 colony maintained at the IAEA-IPCL, having been reared from wild-infested *Musa* sp. (banana)
156 collected in November 2012. Twenty individuals from each location were screened for each
157 morphological analysis whereas ten individuals per location were included for molecular analysis
158 (although not all individuals amplified for all loci). With respect to Indian samples, only individuals
159 from Dahanu were used in molecular analyses because we were unable to amplify DNA from
160 Bangalore material due to its age (1992; pinned loan material from the British Museum of Natural
161 History BMNH, London, U.K.). Bangalore specimens were used for wing-shape analysis as the wings
162 of specimens from Dahanu were too badly damaged. Remaining morphological analyses (aedeagus
163 and lateral vittae morphometrics and scutum colour variation) were conducted on Dahanu material.

164 East Asian samples were from China, Thailand, Peninsular Malaysia, Taiwan, Indonesia and the
165 Philippines. Specimens from Taiwan, Thailand, and Malaysia were used in comparative
166 morphological analyses; specimens from all locations were used in molecular analyses. Further,
167 specimens from Malaysia, Thailand, Indonesia, and the Philippines included individuals traditionally
168 classified as *B. papayae* and *Bactrocera philippinensis* Drew & Hancock. However, as *B.*
169 *philippinensis* has been synonymised with *B. papayae* (Drew & Romig, 2013) and there is now
170 considerable evidence that these two species are synonymous with *B. dorsalis*, we deemed it
171 appropriate to include them here as part of the wider study. All East Asian specimens were collected
172 from the wild into methyl eugenol traps between 2009 and 2012.

We included other species from both within and outside the *B. dorsalis* complex as part of our molecular analysis. Those in the complex included *Bactrocera carambolae* Drew & Hancock (*n* = 61), *B. kandiensis* (*n* = 9), *Bactrocera opiliae* (Drew & Hardy) (*n* = 19), *Bactrocera cacuminata* (Hering) (*n* = 19) and *Bactrocera occipitalis* (Bezzi) (*n* = 22); those from outside the complex were *Bactrocera musae* (Tryon) (*n* = 20) and *Bactrocera tryoni* (Froggatt) (*n* = 9). All sequences used for molecular analysis, except *B. kandiensis*, had been acquired in the earlier study of Boykin *et al.* (2014) with collection data reported therein. We analysed eight specimens of *B. kandiensis* which were collected into methyl eugenol traps in Sri Lanka (caught May 2007). Note that *B. kandiensis* is only included in the molecular analysis as too few samples were obtained for morphological analysis.

Morphology and morphometrics

Four morphological features were examined as part of this study: scutum colour variation, post-sutural lateral vittae, aedeagus length, and wing size and shape. While abdominal colour pattern is listed as differing between *B. dorsalis* and *B. invadens*, we did not include it as part of our study as it was too variable. Only specimens identified as *B. dorsalis* or *B. invadens* were included for analysis (*i.e.*, no outgroups), of which we examined twenty individuals for each morphological feature from all new locations (except Myanmar and Yunnan, China) in addition to three locations included in a previous examination of *B. dorsalis*: Taiwan, Thailand, and Malaysia. We excluded specimens from Myanmar and Yunnan because all individuals died as teneral adults and were unsuitable for morphological analysis.

Scutum: Scutum colour is a continuous variable and defining variants is largely arbitrary. However, in an attempt to document variation across the geographic range, we divided scutum colour into one of three types based on figure 4 in Drew *et al.* (2005): pale, intermediate, or dark. Pale forms were entirely pale-brown or with negligible black colouration (< 10% of the scutum with black markings; see the first two images of fig. 4 in Drew *et al.*, 2005); intermediate forms possessed a weak to strong black lanceolate pattern on an otherwise pale-brown scutum (see images 3–6 in fig. 4 of Drew *et al.*, 2005); and dark forms had entirely, or almost entirely, black scutums whereby the lanceolate pattern

was no longer discernable (> 80% of the scutum is black, see images 7-8 in fig. 4 Drew *et al.*, 2005). This character was not subjected to statistical analysis due to the subjective nature of categorising scutum colour. Instead, the three colour forms are simply presented graphically as frequency charts.

Post-sutural lateral vittae: Measurements of post-sutural vittae width were made at the widest point of the vittae using an eye-piece micrometer mounted into a Leica MZ6 stereo-microscope. Analysis of variance with *post hoc* Tukey test was used to assess for significant differences among sample sites.

Aedeagus: Abdomens were removed and immersed in 10% KOH solution overnight to soften the integument prior to dissection. Each aedeagus was excised from remaining genitalic structures, fully straightened out on a microscope slide and measured as for vittae. Aedeagus length was measured from the base of the aedeagus to the start of (and excluding) the distiphallus, following Krosch *et al.* (2013). Analysis of variance with *post hoc* Tukey test was used to assess for significant differences among sample sites.

Wing shape: One wing from each fly was removed for slide mounting, image capture, and analysis. Usually the right wing was dissected; if damaged, the left was used (~4% of specimens across the total dataset). Wings were mounted in Canada balsam and air-dried prior to image capture using an AnMo Dino-Eye microscope eye-piece camera (model # AM423B) mounted into a Leica MZ6 stereo-microscope. Fifteen wing landmarks were selected followed Schutze *et al.* (2012a) and using the computer program tpsDIG2 v.2.16 (Rohlf, 2010).

Landmark coordinate data were imported into the computer program MorphoJ v.1.04a (Klingenberg, 2011) for shape analysis. Data were subjected to Procrustes superimposition to remove all but shape variation (Rohlf, 1999). Multivariate regression of the dependant wing-shape variable against centroid size (independent variable; see below) was conducted to assess the effect of wing size on wing shape (*i.e.*, allometry) (Drake & Klingenberg, 2008; Schutze *et al.*, 2012a). The statistical significance of this regression was tested by permutation tests (10,000 replicates) against the null hypothesis of independence (MorphoJ v.1.04a). Subsequent analyses were undertaken in MorphoJ v.1.04a using the

residual components as determined from the regression of shape on centroid size to correct for allometric effect.

The size of each wing (centroid size) was calculated in MorphoJ v.1.04a. Centroid size is an isometric estimator of size calculated as the square root of the summed distances of each landmark from the centre of the landmark configuration (see fig. 1.10 in Zeldich *et al.*, 2004). Analysis of variance with *post hoc* Tukey test was used to assess for significant differences among sample sites.

Canonical variate analysis (CVA) on wing shape data was undertaken on 13 *a priori* groups based on collection location. Significant differences were determined via permutation tests (1000 permutation rounds) for Mahalanobis distances among groups ($\alpha = 0.05$; Bonferroni corrected). We also tested for isolation by distance (IBD) (Wright, 1943), whereby we conducted a subset CVA using only individuals from the native range of *B. dorsalis* and *B. invadens* (*i.e.*, all Asian and Indian subcontinent samples to the exclusion of invasive African samples) upon which we undertook regression analysis (SPSS v.21) on pair-wise geographic distance (km) *vs.* Mahalanobis distances calculated from CVA. We did this because *B. dorsalis* has demonstrated a strong IBD effect with respect to wing shape variation within a biogeographical context in Southeast Asia (Schutze *et al.*, 2012b). African samples were excluded from this analysis because they are a recent invasive population (detected in 2003) and hence geographic distance would be artificially inflated with respect to any differences in wing shape.

Molecular analysis

DNA extraction and PCR: Total genomic DNA was extracted from 8–14 individuals from each of the sampled locations using the Bioline Isolate II extraction kit with minor modifications. The modifications involved a pre-crushing step where three legs from each individual were placed in lysis buffer and crushed using either a micro-pestle or three mm ball bearings using a Qiagen mixer mill. Four gene fragments were amplified for the molecular component of this study, which consisted of two nuclear (ITS 1 and 2) and two mitochondrial loci (*cox1* and *nad4*). Primer sequences for ITS1 and ITS2 are as per Boykin *et al.* (2014), *cox1* as per Folmer *et al.* (1994) and *nad4* are Teph_ND4F2 (5'-

WCC WAA RGC TCA TGT WGA AGC TCC-3') and Teph_ND4R2 (5'-WCC CCC TCT AAA TGA ATA AAY WCC-3'). Note, the same *nad4* region was amplified in the present study as in Boykin *et al.* (2014); however, the primers reported in Boykin *et al.* (2014) are incorrect. Each PCR reaction contained 2.5 µL of template DNA, 1 x MyTAQ PCR buffer (Bioline), 0.5 Units of MyTAQ polymerase and 2.5 mM MgCl2, in a total reaction volume of 25 µL. PCR cycling conditions consisted of an initial denaturation step for three minutes at 94°C, followed by 25–30 cycles of 94°C for 30 seconds, 47–52°C for 30 seconds and 72°C for 30 seconds, and a final extension at 72°C for five minutes. PCR products for each gene fragment were purified using a Bioline Isolate PCR purification kit. Cycle sequencing of purified PCR products were conducted using ABI Big Dye® Terminator v.3.1 chemistry. Following a standard isopropanol precipitation clean-up, fragments were sequenced on an ABI 3500 genetic analyser. Trace files were corrected and contigs formed using Sequencher ver. 5.0 (GeneCodes Corporation, 2004).

Phylogenetic Methods: Individual sequences for each of the four loci were aligned in MEGA 5.2.2 (Tamura *et al.*, 2011). Alignments of *cox1* and *nad4* were trivial as no indels were found in this study; ITS1 and ITS2 were aligned by eye (alignments available from the authors upon request). Individual alignments were concatenated in MEGA and partitioned by codon position for protein-coding genes or loci for ribosomal RNA genes. Evolutionary models were inferred for each partition using ModelTest ver 3.6 (Posada & Crandall, 1998)) (ITS1: HKY+G; ITS2: GTR+I+G; *cox1*-1st: GTR-I; *cox1*-2nd: F81; *cox1*-3rd: GTR; *nad4*-1st: GTR+G; *nad4*-2nd: GTR; *nad4*-3rd: HKY). Phylogenetic analyses were run using likelihood and Bayesian inference methods. Likelihood analyses used the RAxML Blackbox webserver (http://phylobench.vital-it.ch/raxml-bb/index.php) (Stamatakis *et al.*, 2008), using separate partitions, a gamma model of rate heterogenetic, estimated proportions of invariant sites, and branch-lengths optimised on a per locus basis. Bayesian analyses used MrBayes ver 3.2 (Ronquist *et al.*, 2012) using unlinked partitions, two independent runs each with three hot and one cold chains, for 10 million generations. Convergence between runs was monitored within MrBayes (standard deviation of split frequencies <0.001) and in Tracer v1.5.4 (Rambaut & Drummond, 2010). Two parallel datasets were analysed, one composed of all specimens for which at

least two of the four loci had been sequenced (Dataset #1, 406 specimens) and one in which all specimens had been sequenced for all four loci (Dataset #2, 293 specimens). Our previous analyses of *B. dorsalis* group flies (Boykin *et al.*, 2014) has shown that phylogenetic analyses is robust to such missing data for this gene set.

A haplotype network using *cox1* data was constructed using the median-joining method followed by maximum parsimony post-processing in Network Version 4.6.1.1 (Bandelt *et al.*, 1999). This allows evolutionary relationships among individuals to be inferred under a statistical framework that does not force bifurcation and was thus compared with relationships resolved using phylogenetic methods.

Results

Morphology and morphometrics

Scutum colour variation: All three colour variants were observed for flies collected from sites across Africa and the Indian subcontinent, albeit with varying relative proportions (Fig. 1). For instance, most individuals from Benin, India, and Nepal had black scutums, whereas most Sri Lankan and Congolese flies had intermediate scutums (black lanceolate pattern). Pale forms were the least frequently observed form, except for the Kenyan sample; however, note that these flies were taken from a laboratory colony.

All east-Asian flies (with one exception) possessed predominantly black scutums with no pale or intermediate forms present. The one exception was the sample from Wuhan for which two flies (out of 20 screened) had a pale scutum. While not examined here, all specimens of *B. dorsalis* collected from further along the Indo-Malay Archipelago and into the Philippines previously examined by the authors possessed a fully black scutum.

Post-sutural lateral vittae: Post-sutural lateral vittae of African *B. invadens* specimens ranged from 0.13 – 0.21 mm, *B. dorsalis* and *B. invadens* from the Indian subcontinent ranged from 0.13 – 0.22 mm, and *B. dorsalis* from eastern Asia ranged from 0.15 – 0.23 mm. Post-sutural lateral vittae width varied significantly across sample sites ($F_{12,247} = 10.76$, $P < 0.001$; Fig. 2A) with no significant

differences among flies from Africa, Pakistan, India, and Sri Lanka; flies from these locations had the narrowest vittae, ranging from an average width of 0.16 mm (Benin) to 0.17 mm (Pakistan). Nepalese flies had significantly wider lateral vittae (mean width = 0.18 mm) than flies from some African locations (Benin, D.R. Congo, and Kenya) and Sri Lanka; however, they did not differ from Mozambican, Sudanese, Indian, or Pakistani samples. The widest vittae belonged to individuals from all east-Asian locations (Thailand, Taiwan, Malaysia, and China; ranging from an average of 0.19 – 0.20 mm for Thai and Chinese flies, respectively) and males from these sites differed significantly from those from every other location in Africa and the Indian subcontinent (except for Nepal and Pakistan).

Aedeagus: While similar to the vittae analysis, in that there was a significant difference among populations for aedeagus length ($F_{12,247} = 15.45$, $P < 0.001$), there was no west-to-east trend from Africa to eastern Asia for aedeagus length (Fig. 2B). Aedeagi from African *B. invadens* specimens ranged from 2.41 – 2.97 mm, *B. dorsalis and B. invadens* from the Indian subcontinent ranged from 2.38 – 2.91 mm, and *B. dorsalis* from eastern Asia ranged from 2.35 – 3.00 mm. Significant aedeagus length variation was observed within Africa; for example, Congolese males had significantly shorter aedeagi (average of 2.64 mm) compared to Mozambican or Kenyan colony flies (2.71 mm and 2.76 mm, respectively). Furthermore, there was significant variation in aedeagus length among samples from eastern Asia: Malaysian aedeagi were significantly longer (above 2.8 mm long) compared to those from other locations in the region (all under 2.8 mm long). Only males from the Indian subcontinent possessed aedeagi of statistically similar lengths among all locations from within the region; there were varying levels of overlap in aedeagus length among populations from this region and those from Africa and eastern Asia.

Wing shape: Wing centroid size significantly varied among sampled populations ($F_{12,247} = 5.013$, $P < 0.001$) and there was a significant allometric effect (4.09%; $P < 0.0001$). While there were differences in wing size among locations, there was no longitudinal trend from Africa to eastern Asia as observed for vittae (Fig. 2C). Similar to the aedeagus analysis, we found significant variation among samples

from Africa and eastern Asia, respectively, whereas all samples from the Indian subcontinent did not significantly vary from each other.

Congolese wings were the smallest of all African locations (average centroid size of 6.06) and they were significantly different from Kenyan and Mozambican flies that possessed the largest wings (average centroid sizes of 6.49 and 6.61, respectively). Sudanese and Beninese flies had wings that were not significantly different from any other African location or from each other (average centroid sizes of 6.39 and 6.35, respectively). There was significant variation in the east-Asian samples, with the smallest wings belonging to Malaysian flies (average centroid size of 5.96; the smallest wings of the entire dataset) which were significantly different from Taiwanese and Chinese flies (average centroid sizes of 6.40 and 6.46, respectively). Contrary to African and east-Asian samples, all wings sampled from the Indian subcontinent were not significantly different from each other with respect to size; ranging from an average centroid size of 6.34 (Nepalese wings) to 6.48 (Indian wings).

Canonical variate analysis following correction for allometric effect (due to the significant result reported above) produced 12 canonical variates of which the first two explained 63% of the variation (Fig. 3). Group Mahalanobis distances were significantly different for all comparisons except among the following locations: i) Sudan, Benin, and the Democratic Republic of the Congo and ii) Nepal and Pakistan. All African groups were closest neighbours except for Kenya, which separated in multidimensional space from other African samples relative to both Pakistan and Nepal (*i.e.*, Pakistani and Nepalese wings were more similar in shape to Sudanese, Congolese, Beninese, and Mozambican wings than Kenya was to any of the other African locations). The remaining Indian subcontinent groups (India and Sri Lanka) were relatively different from both African samples and those from the northern Indian subcontinent (*i.e.*, Pakistan and Nepal). Further, despite their relative geographic proximity, wings from Indian flies were considerably different to Sri Lankan wings. Malaysian, Taiwanese and Thai wings were more similar in shape to wings from Pakistan and Nepal than those from India or Sri Lanka. Chinese wings were highly similar in shape to those from the southern Indian subcontinent, particularly Sri Lanka (Mahalanobis distance between China and Sri Lanka = 2.51).

Canonical variate analysis on Asian samples (*i.e.*, excluding Africa) was conducted on eight *a priori* defined groups: Pakistan, Nepal, India, Sri Lanka, Thailand, Wuhan, Taiwan, and Malaysia, resulting in seven canonical variates of which the first two explained 74% of the variation. There was no significant association between Mahalanobis distance and geographic distance (r^2 = 0.001; P = 0.850; Fig. 4).

Molecular analysis

Phylogenetics: New sequences were generated for up to four loci per specimen and combined with sequences from a previous study (Boykin *et al.*, 2014) for phylogenetic analyses (GenBank accession numbers JX099580-JX099755; KC446030-KC447278; KM 453245- KM453574; see Table S1). Bayesian and Maximum likelihood analyses for both Dataset #1 (two out of the four loci analysed, 406 individuals) and Dataset #2 (all four loci, 293 individuals) yielded similar phylogenetic topologies; albeit with varying levels of nodal support with highest values generally obtained for the Bayesian analysis using Dataset #1 (Fig. 5). All outgroups were well resolved, including *B. carambolae* recovered as sister to the larger *B. dorsalis* clade. The ingroup clade contained previously sequenced data for *B. dorsalis, B. papayae* and *B. philippinensis* from Southeast Asia (collectively termed *'B. dorsalis* s.l.') (Boykin *et al.*, 2014), in addition to *de novo* data from individuals obtained from the expanded range of east Asia (*B. dorsalis*), the Indian subcontinent (*B. dorsalis, B. invadens*, and *B. kandiensis*) and Africa (*B. invadens*). Specimens of *B. dorsalis* additional to the study of Boykin *et al.* (2014) were from Myanmar, China (Wuhan and Yunnan), India, Nepal, and Pakistan. All newly included *B. dorsalis* and African specimens of *B. invadens* fell within the broader *B. dorsalis* clade with no evidence of statistically supported subclades. All African *B. invadens* specimens were either completely unresolved within the broader *B. dorsalis* clade or emerged as two weakly supported clades that included individuals from across all African countries sampled and representing the range of scutum colour variation (see Suppl. Figs 1 and 2). The same was true for Sri Lankan *B. invadens* specimens which were either fully unresolved within the *B. dorsalis* clade or formed small, poorly supported groups nested within the larger clade. Scutum colour (*i.e.*, red-brown, black, or intermediate) did not align with the phylogeny in any consistent pattern (Fig. 5 inset).

Bactrocera kandiensis is the only subclade within *B. dorsalis* s.l. with significant nodal support (Fig. 5). This was driven by a unique indel pattern present in the ITS1 locus which, while diagnostic, is shorter than the indel that which occurs in the same locus in *B. carambolae* (Boykin *et al.*, 2014). Further, some specimens of *B. dorsalis* possessed *cox1* haplotypes more closely related to *B. kandiensis* than other *B. dorsalis* haplotypes; these included five specimens from Sri Lanka (Bd1561-1564 and Bd1566), two from Myanmar (Bd1580 and Bd1582) and seven from India (Bd1691-1697). As these specimens did not possess the *'B. kandiensis* ITS1 indel', they resolved with the remainder of *B. dorsalis* in multi-locus analyses.

The median-joining network largely conformed to that presented in Schutze *et al.* (2012b) that comprised east Asian *B. dorsalis* s.l. haplotypes, with new sequences from this paper placed throughout (Fig. 6). The central starburst-like pattern remained, with numerous singletons radiating from a common, widespread haplotype. Individuals of *B. kandiensis* formed a divergent and diverse cluster and were connected to the network by a very long branch, demonstrating that the position of this taxon in the multilocus phylogeny is not driven solely by ITS1 indel patterns. Within the *B. kandiensis* cluster were several haplotypes from *B. dorsalis* flies from Sri Lanka, India, and Myanmar; although, no haplotypes were shared between *B. dorsalis* and *B. kandiensis* flies. There was no apparent separation of haplotypes from flies identified morphologically as *B. invadens* or *B. dorsalis*; haplotypes of these two taxa generally were scattered throughout the network. Indeed, four *cox1* haplotypes were shared between the two taxa; one sampled from Nepal and Thailand populations, one from Sudan and Thailand, one from Taiwan, Malaysia, Nepal, and Pakistan and the common widespread haplotype sampled from all locations except Benin, Mozambique, Kenya, and Sudan (Suppl. Fig. 3). Likewise, there was no clear geographical pattern in the network, although there were few haplotypes shared among individuals from broadly different regions. No clustering of haplotypes was apparent for individuals with different scutum colours (Suppl. Fig. 2). Five haplotypes were shared by individuals that differed at this trait (often the same haplotype was shared among sites and/or among *B. dorsalis* and *B. invadens* flies).

Discussion

The interpretation of our combined results strongly suggests that *B. dorsalis* and *B. invadens* are one biological species. Genetic data, both at the multi-locus and haplotype levels, fail to show evidence of distinct clades or unique haplotypes, respectively, consistent with the presence of two species as reported in other studies (Khamis *et al.*, 2012; Frey *et al.*, 2013; San Jose *et al.*, 2013). Examination of morphology shows a great deal of variation among populations, some with apparent geographic structuring that relates to current taxonomy (especially scutum colour), but with other traits showing no such structure; in all cases the patterns in morphology do not align with any apparent genetic variation. To place our results within the broader context of *B. dorsalis* taxonomy and species delimitation over the centuries, a brief summary of the confusing taxonomic history of *B. dorsalis* is provided.

Taxonomic history of Bactrocera dorsalis

The species now known as *B. dorsalis* was first described by Fabricius in the late 18th Century as a rust-red coloured fly from 'India Orientali' under the name *Musca ferruginea* (Fabricius, 1794). Note that whilst India Orientali can be interpreted as the East Indies (Pont, 1995), the type specimen described by Fabricius is considered to be from East India (Drew & Romig, 2013). Further, treatments of other Fabrician collections, *e.g.*, hymenopterans (van der Vecht, 1961), state that India Orientali usually refers to India, rather than other parts of Southeast Asia (*e.g.*, Indonesia). Fabricius (1805) subsequently transferred this species to the genus *Dacus* Fabricius, a combination that persisted until the 20th Century. In the early part of the 1900s, however, the morphological variability of *D. ferrugineus* was noted by Froggatt (1910), specifically the scutum colour which ranged from black to rust-red. The black scutum variety was soon described by Hendel (1912) as a new species, *D. dorsalis* Hendel, following examination of specimens from Formosa (= Taiwan) and with a black scutum the chief discriminatory character separating it from *D. ferrugineus* (which possessed a red-brown scutum).

Nevertheless, the view of the 'Formosan type' as a distinct species was not universally accepted, with studies over the next 50 years regarding *D. dorsalis* as either a species in its own right (Perkins, 1938)

or simply a dark variety of *D. ferrugineus* (Bezzi, 1916; Miyake, 1919; Shiraki, 1933; Munro, 1939). Critically, Hendel himself accepted that *D. dorsalis* represented a black variety of *D. ferrugineus* and, after examining more specimens, explicitly stated that specimens from Taiwan corresponded with the Fabrician description (Hendel, 1915). Munro (1939) found north-west Indian specimens to exhibit a full range of thoracic colour forms (from pale, through intermediates, to dark) with no additional structural characters present to further distinguish any of these forms from each other, leading him to conclude they all belonged to the same species.

In the late 1960s, Hawaiian taxonomist D.E. Hardy undertook a revision of what was, by then, commonly known as the 'Oriental Fruit Fly' (Hardy, 1969). The key outcomes of this revision were the following: i) the species name *ferrugineus* was invalid as it was preoccupied by another fly described by Scopoli (1763); ii) the only valid name available for the Fabrician species was Hendel's *D. dorsalis*; iii) that *D. ferrugineus* must therefore become a junior synonym of *D. dorsalis*; iv) that individuals of this species with a red-brown scutum colour were teneral adults yet to develop their final black-scutum colouration; v) *D. dorsalis* was characterised as possessing only a black (or mostly black) scutum; and finally, vi) a number of closely-related species existed which were to be placed in the newly formed, 16 member, *D. dorsalis* species complex. It was at this critical point that red-brown scutum forms were subsumed in subsequent treatments of *D. dorsalis*, including in major revisions towards the end of the 20th Century by which time *D. dorsalis* had been reassigned to genus *Bactrocera* and the complex expanded to more than 70 species with a black, or mostly black, scutum one of their defining characters (Drew, 1989; Drew & Hancock, 1994; Drew & Romig, 2013). Importantly, Hardy's (1969) revision referred to previous work which detailed the range of colour forms, such as the Indian study by Munro (1939); however, Hardy specifically noted that Munro examined a limited sample range of 39 specimens, and that Hardy himself never observed such variability in the many thousands of specimens he examined from India and Pakistan (Hardy, 1969).

A morphologically variable fly closely allied to *B. dorsalis* was reported from Africa in 2003 (Lux *et al.*, 2003). Although the newly detected species was considered highly variable, it showed morphological characters that were consistent with *B. dorsalis* (Lux *et al.*, 2003). Our examination of

specimens from Sri Lanka in comparison to African specimens has led to the conclusion that the invasion likely originated from the Indian subcontinent (*n.b.,* the same region from which *D. ferrugineus* was likely first described by Fabricius) and that these morphologically variable flies were a new species which was described by Drew *et al.* (2005) as *B. invadens*. An important side note for taxonomic consideration is that the type locality of *B. invadens* is Kenya (*i.e*., invasive range), not Asia (*i.e.*, native range and type locality).

Morphological variation

Clearly there was confusion during the last century over the identity of *B. dorsalis* in relation to *D. ferrugineus*, particularly with respect to scutum colour variation. Earlier studies, such as that of Munro (1939), described specimens from the Indian subcontinent as exhibiting a range of thoracic colour forms, yet towards East Asia the darker, mostly black form, predominated. Our assessment of scutum colour variation of newly acquired specimens reflects this pattern, with a range of colour forms across the Indian subcontinent (Fig. 1) and entirely black forms occurring eastward into the rest of Asia, with the exception of a small number of individuals from China. All flies in our study were collected from traps placed in the wild or sourced from colony material and were fully mature specimens, thereby conflicting with Hardy's (1969) view that mature red-brown specimens of *B. dorsalis* do not exist. Indeed, the presence of a range of thoracic colour forms is similarly documented, either directly or through inference, in contemporary studies of material from the Indian subcontinent. Drew *et al.* (2007) recorded *B. invadens* in Bhutan, but considered as doubtful by Drew & Romig (2013); and a recent illustrated key on Indian fruit flies states that while *B. invadens* does not occur in India, specimens keying out as *B. dorsalis* may more closely match descriptions given for *B. invadens* (David & Ramani, 2011). Further, a detailed survey of *B. dorsalis* from Bangladesh clearly demonstrated individuals to possess the range of thoracic and abdominal colour forms typical of African and Sri Lankan *B. invadens* (Leblanc *et al.*, 2013). Given the evidence at hand, it is difficult to accept Hardy's assertion that such colour variation does not exist in *B. dorsalis*.

Other presumably diagnostic morphological characters measured here, namely aedeagus length and width of post-sutural lateral vittae, do not conform to variation expected under a two-species hypothesis. *Bactrocera invadens* is reported to possess longer aedeagi and narrower post-sutural lateral vittae than *B. dorsalis* (Drew *et al.*, 2008; Drew *et al.*, 2005). Our results demonstrate neither of these characters possesses diagnostic value as they show either no pattern at all (*i.e.*, aedeagus length, Fig. 2B) or they are continuously variable across a geographic cline (*i.e.*, vittae width, Fig. 2A). Such results may, therefore, be indicative population level variation rather than species level variation, similar to the latitudinal variation in aedeagus length documented for *B. dorsalis* in Southeast Asia (Krosch *et al.*, 2013).

Geometric morphometric shape analysis is a more sensitive tool for assessing morphological variation than simple linear measurements (Krosch *et al.*, 2013; Schutze *et al.*, 2012b). This study extends wing shape analysis from East Asia into the Indian subcontinent and Africa, revealing potentially insightful patterns of variation and points of origin. For example, while wing shape is highly similar among all African populations of *B. invadens* (as expected for a relatively newly established invasive population), there is greater difference in wing shape among populations of *B. dorsalis* throughout the native range of the Indian subcontinent and East Asia (Fig. 3). Furthermore, the wing shape of African flies is most similar to those from the northern range of the Indian subcontinent, namely Nepal and Pakistan (Fig. 3), while those from further south (*i.e.*, India and Sri Lanka) have wings that are relatively different in shape from African and northern Indian subcontinent populations (Fig. 3). Wing shape can, under some circumstances, demonstrate a highly significant isolation by distance signature, as found in the study of *B. dorsalis* s.l. in Southeast Asia (Schutze *et al.*, 2012b). In that study, wing shape was superior to population genetic data at resolving isolation by distance signatures given a specific biogeographic hypothesis, demonstrating that as geographic distance between populations increased so did relative differences in wing shape. We therefore consider wing shape analysis to be a valuable tool for inferring the origin of an invasive species such as *B. invadens*, and that the African invasion may have come from the northern Indian subcontinent rather than Sri Lanka, as previously

thought. This hypothesis could be tested further by a targeted population genetic study using markers of contemporary gene flow (*e.g.*, microsatellites or RAD-tag).

A 'one-species' hypothesis is supported by molecular data

While the significant morphological variation could be interpreted as the forms present in Africa and the Indian subcontinent representing different species (*i.e.*, *B. invadens* in the west, *B. dorsalis* in the east), it is not supported by the molecular evidence. These data fail to resolve specimens from Africa and the Indian subcontinent as distinct from the broader *B. dorsalis* clade, a clade incorporating material from across the geographic range of *B. dorsalis* and *B. invadens* (Fig. 5). Moreover, mapping scutum colour onto individuals from the *dorsalis* clade reveals little to no pattern in the distribution of colour forms (Fig. 5 inset). Previous molecular studies have found similar results. Neighbour-joining analysis of the *cox1* barcoding gene, for instance, resolved specimens from a Hawaiian lab colony of *B. dorsalis* as a group within a larger clade of *B. invadens* from Africa and Sri Lanka, while splitting *B. invadens* into two clades, one with specimens from Africa and Sri Lanka (along with *B. dorsalis*), the other grouping Sri Lankan *B. invadens* with Sri Lankan *B. kandiensis* (Khamis *et al.*, 2012). While Khamis *et al.* (2012) did not conclude that *B. dorsalis* and *B. invadens* were the same species, a more recent multi-locus phylogenetic study of a number of *Bactrocera* species found the following: i) *B. invadens* to be polyphyletic within the *B. dorsalis* s.l. clade, ii) that *B. invadens* was genetically indistinguishable from many of the pest species within the group and therefore, iii) did not support the placement of *B. invadens* as an independent species outside of *B. dorsalis* s.l. (San Jose *et al.*, 2013). This conclusion was also reached by Frey *et al.* (2013) who explicitly stated that *B. invadens* should be synonymised with *B. dorsalis* as a result of their *cox1* barcode study. These earlier molecular studies, while extensive in their own right, have incorporated a relatively limited sample range for either *B. dorsalis* or *B. invadens*, yet they nevertheless demonstrate that *B. dorsalis* and *B. invadens* are most likely a single species. This conclusion is reinforced in our phylogenetic study, which represents the most extensive molecular analysis of *B. invadens* and *B. dorsalis* from across much of their native and invasive geographic ranges. We also found haplotypes to be shared between individuals identified as either *B. dorsalis* or *B. invadens* in our analysis of the *cox1* haplotype

network (Fig. 6). Furthermore, the most common and widespread haplotype includes individuals that exhibit the full range of scutum colour forms from Thailand, Malaysia, India, Nepal, China, Pakistan, Sri Lanka, and the Democratic Republic of the Congo (Supplemental Figs. 2 and 3).

The curious case of B. kandiensis – *evidence of introgression?*

Our analysis of the *cox1* gene revealed an unexpected association between *B. kandiensis*, *B. dorsalis*, and *B. invadens*, in that some Sri Lankan, Indian, and Burmese specimens that were morphologically identified as either *B. dorsalis* or *B. invadens* possessed *cox1* sequences more closely related to *B. kandiensis* haplotypes than conspecifics (Fig. 6). This was not reflected by nuclear data which, to the contrary, revealed an indel that consistently separated all *B. kandiensis* specimens from *B. dorsalis*/*invadens*. Our results reflect the barcode study of Khamis *et al.* (2012), who found a large proportion of Sri Lankan *B. invadens* specimens to be more closely related to *B. kandiensis* than to African *B. invadens* or Hawaiian *B. dorsalis*. Importantly, however, Khamis *et al.* (2012) examined a single gene, *cox1* and on only Sri Lankan specimens rather than those from other locations across the Indian subcontinent.

The presence of *B. kandiensis* haplotypes among *B. invadens* or *B. dorsalis* individuals raises the possibility of introgression: the permanent incorporation of genes from one population into another via hybridization (Dowling & Secor, 1997). Introgression in dacine fruit flies has been proposed for other taxa, such as the Australian species, *B. tryoni*, for which horizontal introgression of genetic material from its sister species, *Bactrocera neohumeralis* (Hardy), was proposed as a potential adaptive mechanism allowing the expansion of *B. tryoni* into new climate regions (Lewontin & Birch, 1966). Under the current scenario, the presence of *B. kandiensis* mtDNA haplotypes in the genome of *B. dorsalis*, but not the reverse (*i.e.*, *B. dorsalis* haplotypes in *B. kandiensis*), implies that such hybridisation, if it has occurred, was unidirectional. Moreover, this pattern must have resulted from sex-biased couplings whereby *B. dorsalis* males mate with *B. kandiensis* females to produce offspring with *B. dorsalis* morphology but with *B. kandiensis* mtDNA. Preliminary field cage mating tests examining compatibility between *B. kandiensis* and *B. dorsalis* revealed no evidence of assortative

mating between these two species (unpublished data), reinforcing the potential for them to interbreed in sympatry. *Bactrocera kandiensis* is recorded only from Sri Lanka; however, it is suspected to also exist in southern India (Kapoor, 2005) and may occur sympatrically with *B. dorsalis* potentially as far east as Myanmar. This remains a hypothesis in need of further testing, with the incorporation of a wider sample range and continued research into mating compatibility between these taxa. Furthermore, we advocate additional studies including other species of the *dorsalis* complex, such as *Bactrocera caryeae* (Kapoor). This economically important species is considered allopatric to *B. kandiensis* as it is recorded from India but not Sri Lanka, yet they emerge as sister taxa in phylogenetic analyses (Krosch *et al.* 2012) and are distinguished based on abdominal colour pattern (Drew & Romig, 2013).

Conclusions

Our integrative molecular and morphological study of *B. invadens* and *B. dorsalis* from across a wide geographic distribution supports the hypothesis that they represent a single biological species. These data, in accordance with mounting evidence from other studies (Bo *et al.*, 2014; Frey *et al.*, 2013; Khamis *et al.*, 2012; San Jose, 1999; Tan *et al.*, 2011), highlight the need for formal synonymy between *B. dorsalis* and *B. invadens* and a subsequent revision of the current description of the Oriental fruit fly to encompass a wider range of morphological colour variants, particularly with respect to scutum colour. Further, given the taxonomic history of this species, we argue that the fly described as *B. invadens* is likely conspecific with that described by Fabricius (1794) as *M. ferrugineus*. The type specimen of *ferrugineus* designated by Fabricius and held by the Natural History Museum of Denmark is almost completely destroyed and so this proposition cannot be directly tested. However, what remains of this specimen (a thorax and partial abdomen) bears a strong resemblance to present-day *B. invadens* (Fig. 7), and a later-collected individual from Sri Lanka (coll. 1899 and labelled by Hendel as *Chaetodacus ferrugineus* F.; specimen in the Natural History Museum, Vienna, Austria) has been confirmed as *B. invadens* (Drew & Romig, 2013). Synonymising these species will have a profound impact on quarantine and trade access, especially for sub-Saharan Africa which has been devastated by the rapid and destructive expansion of this species across the

592 continent. However, we stress that although *B. invadens* is the same biological species as *B. dorsalis*,
593 the potential for biological differences among populations remains, especially considering the broad
594 geographic distributions and environmental conditions where these species are found. Finally, our
595 data revealing the potential of hybridising introgression between *B. dorsalis/invadens* and *B.*
596 *kandiensis* further exemplifies the need for further research into the mechanisms of speciation and the
597 evolution of the *B. dorsalis* species complex.

Acknowledgements

We sincerely thank Marc De Meyer, S. Vijaysegaran, Sundar Tiwari, Disna Gunawardana, Ramesh Hire, and the Insect Pest Control Laboratory of the UN-FAO International Atomic Energy Agency (IAEA) for providing specimens used in this project. We also thank Thomas Pape and Ole Karsholt of the Natural History Museum of Denmark for providing an image of the type specimen of *D. ferrugineus*; and the British Museum of Natural History for material loaned from their collections. This research was undertaken with the support of the UN-FAO IAEA Coordinated Research Project 'Resolution of cryptic species complexes of tephritid pests to overcome constraints to SIT application and international trade'. MKS was supported by the Australian Government CRC National Plant Biosecurity and Plant Biosecurity CRC. SLC was supported by the Australian Research Council, Future Fellowships scheme (FT120100746).

609 **References**

610 Aytekin, A.M., Terzo, M., Rasmont, P., *et al.* (2007) Landmark based geometric morphometric
611 analysis of wing shape in *Sibiricobombus* Vogt (Hymenoptera: Apidae: Bombus Latreille).
612 *Annales de la Société Entomologique de France*, **43**, 95–102.

613 Bandelt, H-J., Forster, P. & Röhl, A. (1999) Median-joining networks for inferring intraspecific
614 phylogenies. *Molecular Biology and Evolution*, **16**, 37–48.

615 Bezzi, M. (1916) On the fruit flies of the genus *Dacus* (s.l.) occurring in India, Burma, and Ceylon.
616 *Bulletin of Entomological Research*, **7**, 99–121.

617 Bickford, D., Lohman, D.J., Sodhi, N.S., *et al.* (2007) Cryptic species as a window on diversity and
618 conservation. *Trends in Ecology and Evolution*, **22**, 148–155.

619 Bo, W., Ahmad, S., Dammalage, T., *et al.* (2014) Mating compatibility between *Bactrocera invadens*
620 and *Bactrocera dorsalis* (Diptera: Tephritidae). *Journal of Economic* Entomology, **107**, 623–
621 629.

622 Boykin, L.M., Schutze, M.K., Krosch, M.N., *et al.* (2014) Multi-gene phylogenetic analysis of south-
623 east Asian pest members of the *Bactrocera dorsalis* species complex (Diptera: Tephritidae)
624 does not support current taxonomy. *Journal of Applied Entomology*, **138**, 235–253.

625 Clarke, A.R., Armstrong, K.F., Carmichael, A.E., *et al.* (2005) Invasive phytophagous pests arising
626 through a recent tropical evolutionary raditation : the *Bactrocera dorsalis* complex of fruit
627 flies. *Annual Review of Entomology*, **50**, 293–319.

628 Gene Codes Corporation (2004) Sequencher. 5.0 ed.: Gene Codes Corporation, Inc, Madison,
629 Wisconsin.

630 David, K.J. & Ramani, S. (2011) An illustrated key to fruit flies (Diptera: Tephritidae) from
631 Peninsular India and the Andaman and Nicobar Islands. *Zootaxa*, **3021**, 1–31.

632 De Meyer, M., Robertson, M.P., Mansell, M.W., *et al.* (2010) Ecological niche and potential
633 geographic distribution of the invasive fruit fly *Bactrocera invadens* (Diptera, Tephritidae).
634 *Bulletin of Entomological Research*, **100**, 35–48.

635 Dowling, T.E. & Secor, C.L. (1997) The role of hybridization and introgression in the diversification
636 of animals. *Annual review of Ecology and Systematics*, **28**, 593–619.

637 Drake, A.G. & Klingenberg, C.P. (2008) The pace of morphological change: Historical
638 transformation of skull shape in St. Bernard dogs. *Proceedings of the Royal Society of*
639 *London, B Biological Sciences*, **275**, 71–76.
640 Drew, R.A.I. (1989) The tropical fruit flies (Diptera: Tephritidae: Dacinae) of the Australasian and
641 Oceanian regions. *Memoirs of the Queensland Museum*, **26**, 1–521.
642 Drew, R.A.I. & Hancock, D.L. (1994) The *Bactrocera dorsalis* complex of fruit flies (Diptera:
643 Tephitidae: Dacinae) in Asia. *Bulletin of Entomological Research Supplement Series*, **Suppl.**
644 **No. 2**, i–iii + 1–68.
645 Drew, R.A.I. & Hancock, D.L. (2000) Phylogeny of the Tribe Dacini (Dacinae) based on
646 morphological, distributional, and biological data. In: Aluja M, Norrbom, A.L (ed) *Fruit Flies*
647 *(Tephritidae): Phylogeny and Evolution of Behavior*. Boca Raton: CRC Press, 491–504.
648 Drew, R.A.I. & Romig, M.C. (2013) Tropical fruit flies of South-east Asia. CAB International,
649 Wallingford, UK.
650 Drew, R.A.I., Raghu, S. & Halcoop, P. (2008) Bridging the morphological and biological species
651 concepts: studies on the *Bactrocera dorsalis* (Hendel) complex (Diptera: Tephritidae:
652 Dacinae) in South-east Asia. *Biological Journal of the Linnean Society*, **93**, 217–226.
653 Drew, R.A.I., Romig, M.C. & Dorji, C. (2007) Records of dacine fruit flies and new species of *Dacus*
654 (Diptera: Tephritidae) in Bhutan. *Raffles Bulletin of Zoology*, **55**, 1–21.
655 Drew, R.A.I., Tsuruta, K. & White, I.M. (2005) A new species of pest fruit fly (Diptera: Tephritidae:
656 Dacinae) from Sri Lanka and Africa. *African Entomology*, **13**, 149–154.
657 Ekesi, S., Billah, M.K., Nderitu, P.W., *et al.* (2009) Evidence for competitive displacement of
658 *Ceratitis cosyra* by the Invasive Fruit Fly *Bactrocera invadens* (Diptera: Tephritidae) on
659 mango and mechanisms contributing to the displacement. *Journal of Economic Entomology*,
660 **102**, 981–991.
661 Ekesi, S., Nderitu, P.W. & Rwomushana, I. (2007) Field infestation, life history and demographic
662 parameters of the fruit fly *Bactrocera invadens* (Diptera: Tephritidae) in Africa. *Bulletin of*
663 *Entomological Research*, **96**, 379–386.
664 Fabricius, J.C. (1794) *Entomologia systemica emendata et aucta*: Hafniae.

665 Fabricius, J.C. (1805) *Systema antliatorum secundum ordines, genera, species*: Brunsvigae.
666 Folmer, O., Black, M., Hoeh, W., *et al.* (1994) DNA primers for amplification of mitochondrial
667 cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine*
668 *Biology and Biotechnology*, **3**, 294–299.
669 Frey, J.E., Guillén, L., Frey, B., *et al.* (2013) Developing diagnostic SNP panels for the identification
670 of true fruit flies (Diptera: Tephritidae) within the limits of COI-based species delimitation.
671 *BMC Evolutionary Biology*, **13**.
672 Froggatt, W.W. (1910) *Fruit flies (Family Trypetidae)*: Department of Agriculture, New South Wales.
673 Grout, T.G., Daneel, J.H., Mohamed, S.A., *et al.* (2011) Cold susceptibility and disinfestation of
674 *Bactrocera invadens* (Diptera: Tephritidae) in oranges. *Journal of Economic Entomology*,
675 **104**, 1180–1188.
676 Goergen, G., *et al.* (2011) *Bactrocera invadens* (Diptera: Tephritidae), a new invasive fruit fly pest
677 for the Afrotropical region: host plant range and distribution in West and Central Africa.
678 *Environmental Entomology*, **40**, 844–854.
679 Hallman, G.J., Myers, S.W., Jessup, A.J., *et al.* (2011) Comparison of in vitro heat and cold tolerances
680 of the new invasive species *Bactrocera invadens* (Diptera: Tephritidae) with three known
681 tephritids. *Journal of Economic Entomology*, **104**, 21–25.
682 Hardy, D.E. (1969) Taxonomy and distribution of the Oriental fruit fly and related species
683 (Tephritidae-Diptera). *Proceedings of the Hawaiian Entomological Society*, **20**, 395–428.
684 Hendel, F. (1912) H. Sauter's Formosa-Ausbeute. Genus *Dacus* (Dipt.). *Suplementa Entomologica*, **1**,
685 13–24.
686 Hendel, F. (1915) H. Sauter's Formosa-Ausbeute. Tephritinae. *Annales Musei Nationalis Hungarici*,
687 **13**, 424–467.
688 Kapoor, V.C. (2005) Taxonomy and biology of economically important fruit flies of India. *Israel*
689 *Journal of Entomology*, **35–36**, 459–475.
690 Khamis, F.M., Karam, N., Ekesi, S., *et al.* (2009) Uncovering the tracks of a recent and rapid
691 invasion: the case of the fruit fly pest *Bactrocera invadens* (Diptera: Tephritidae) in Africa.
692 *Molecular Ecology*, **18**, 4798–4810.

693 Khamis, F.M., Masiga, D.K., Mohamed, S.A., *et al.* (2012) Taxonomic identity of the invasive fruit
694 fly pest, *Bactrocera invadens*: concordance in morphometry and DNA barcoding. *PLoS ONE*
695 **7**, e44862.
696 Klingenberg, C.P. (2011) MorphoJ: an integrated software package for geometric morphometrics.
697 *Molecular Ecology Resources*, **11**, 353–357.
698 Krosch, M.N., Schutze, M.K., Armstrong, K.F., *et al.* (2013) Piecing together an integrative
699 taxonomic puzzle: microsatellite, wing shape and aedeagus length analyses of *Bactrocera*
700 *dorsalis* s.l. (Diptera: Tephritidae) find no evidence of multiple lineages in a proposed contact
701 zone along the Thai/Malay Peninsula. *Systematic Entomology*, **38**, 2–13.
702 Krosch, M.N., Schutze, M.K., Armstrong, K.F., *et al.* (2012) A molecular phylogeny for the Tribe
703 Dacini (Diptera: Tephritidae): systematic and biogeographic implications. *Molecular*
704 *Phylogenetics and Evolution*, **64**, 513–523.
705 Leblanc, L., Hossain, M.A., Khan, S.A., *et al.* (2013) A preliminary survey of the fruit flies (Diptera:
706 Tephritidae: Dacine) of Bangladesh. *Proceedings of the Hawaiian Entomological Society*, **45**,
707 51–58.
708 Lewontin, R.C. & Birch, L.C. (1966) Hybridization as a source of variation for adaptation to new
709 environments. *Evolution*, **20**, 315–336.
710 Lux, S.A., Copeland, R.S., White, I.M., *et al.* (2003) A new invasive fruit fly species from the
711 *Bactrocera dorsalis* (Hendel) group detected in East Africa. *Insect Science and its*
712 *Application*, **23**, 355–361.
713 Mahmood, K. (2004) Identification of pest species in Oriental fruit fly, *Bactrocera dorsalis* (Hendel)
714 (Diptera: Tephritidae) species complex. *Pakistan Journal of Zoology*, **36**, 219–230.
715 Medina, F.I.S., Carillo, P.A.V., Gregorio, J.S., *et al.* (1998) The mating compatibility between
716 *Bactrocera philippinensis* and *Bactrocera dorsalis*. In: Tan K.H. (ed) *Abstracts, 5th*
717 *International Symposium on Fruit Flies of Economic Importance, 1-5 June.* Penang,
718 Malaysia, **155**.
719 Miyake, T. (1919) Studies on the fruit flies of Japan. *Bulletin of the Imperial Central Agricultural*
720 *Experiment Station in Japan*, **2**, 85–164.

721 Mohamed, S.A., Ekesi, S. & Hanna, R. (2008) Evaluation of the impact of *Diachasmimorpha*
722 *longicaudata* on *Bactrocera invadens* and five African fruit fly species. *Journal of Applied*
723 *Entomology*, **132**, 789–797.
724 Munro, H.K. (1939) The fruit fly, *Dacus ferrugineus* Fabr., and its variety *dorsalis* Hendel in
725 northwest India. *Indian Journal of Entomology*, **1**, 101–105.
726 Mwatawala, M.W., De Meyer, M., Makundi, R.H., *et al.* (2006) Seasonality and host utilization of the
727 invasive fruit fly, *Bactrocera invadens* (Dipt., Tephritidae) in central Tanzania. *Journal of*
728 *Applied Entomology*, **130**, 530–537.
729 Perkins, F.A. (1938) Studies in oriental and Australian Trypaneidae. Part 2. Adraminae and Dacinae
730 from India, Ceylon, Malaya, Sumatra, Java, Borneo, Philippine Islands, and Formosa.
731 *Proceedings of the Royal Society of Queensland*, **49**, 120–144.
732 Perkins, M.V., Fletcher, M.T., Kitching, W., *et al.* (1990) Chemical studies of rectal gland secretions
733 of some species of *Bactrocera dorsalis* complex of fruit flies (Diptera: Tephritidae). *Journal*
734 *of Chemical Ecology*, **16**, 2475–2487.
735 Pont, A.C. (1995) The dipterist C.R.W. Wiedemann (1770-1840). His life, work and collections.
736 *Steenstrupia*, **21**, 125–154.
737 Posada, D. & Crandall, K.A. (1998) MODELTEST: Testing the model of DNA substitution.
738 *Bioinformatics*, **14**, 817–818.
739 Rambaut, A. & Drummond, A.J. (2010) *Tracer v1.5.4*. Available at: http://beast.bio.ed.ac.uk/Tracer.
740 Rohlf, F.J. (1999) Shape statistics: Procrustes superimpositions and tangent spaces. *Journal of*
741 *Classification*, **16**, 197–223.
742 Rohlf, F.J. (2010) tpsDIG2, digitize landmarks and outlines. 2.16 ed. Department of Ecology and
743 Evolution, State University of New York at Stony Brook.
744 Ronquist, F., Teslenko, M., van der Mark, P., *et al.* (2012) MrBayes 3.2: Efficient Bayesian
745 phylogenetic inference and model choice across a large model space. *Systematic Biology*, **61**,
746 539–542.
747 Rwomushana, I., Ekesi, S., Callistus, K.P.O.O., *et al.* (2009) Mechanisms contributing to the
748 competitive success of the invasive fruit fly *Bactrocera invadens* over the indigenous mango

749 fruit fly, *Ceratitis cosyra*: the role of temperature and resource pre-emption. *Entomologia*
750 *Experimentalis et Applicata*, **133**, 27–37.
751 Rwomushana, I., Ekesi, S., Gordon, I., *et al.* (2008) Host plants and host plant preference studies for
752 *Bactrocera invadens* (Diptera: Tephritidae) in Kenya, a new invasive fruit fly species in
753 Africa. *Annals of the Entomological Society of America*, **101**, 331–340.
754 San Jose, M., Leblanc, L., Geib, S.M., *et al.* (2013) An evaluation of the species status of *Bactrocera*
755 *invadens* and the systematics of the *Bactrocera dorsalis* (Diptera: Tephritidae) complex.
756 *Annals of the Entomological Society of America*, **106**, 684–694.
757 Salum, J.K., Mwatawala, M.W., Kusolwa, P.M. & De Meyer, M. (2014) Demographic parameters of
758 the two main fruit fly (Diptera: Tephritidae) species attacking mango in Central Tanzania.
759 *Journal of Applied Entomology*, **138**, 441–448.
760 Schlick-Steiner, B.C., Steiner, F.M., Seifert, B., *et al.* (2010) Integrative taxonomy: a multisource
761 approach to exploring biodiversity. *Annual Review of Entomology*, **55**, 421–438.
762 Schutze, M.K., Jessup, A. & Clarke, A.R. (2012a) Wing shape as a potential discriminator of
763 morphologically similar pest taxa within the *Bactrocera dorsalis* species complex (Diptera:
764 Tephritidae). *Bulletin of Entomological Research*, **102**, 103–111.
765 Schutze, M.K., Krosch, M.N., Armstrong, K.F., *et al.* (2012b) Population structure of *Bactrocera*
766 *dorsalis s.s.*, *B. papayae* and *B. philippinensis* (Diptera: Tephritidae) in southeast Asia:
767 evidence for a single species hypothesis using mitochondrial DNA and wingshape data. *BMC*
768 *Evolutionary Biology*, **12**, DOI: 10.1186/1471-2148-1112-1130.
769 Shiraki, T. (1933) A systematic study of Trypetidae in the Japanese empire. *Memoirs of the Faculty of*
770 *Science and Agriculture Taihoku Imperial University*, **8**, pp. 509.
771 Stamatakis, A., Hoover, P. & Rougemont, J. (2008) A rapid bootstrap algorithm for the RAxML
772 webservers. *Systematic Biology*, **75**, 758–771.
773 Tamura, K., Peterson, D., Peterson, N., *et al.* (2011) MEGA5: molecular evolutionary genetics
774 analysis using maximum likelihood, evolutionary distance, and maximum parsimony
775 methods. *Molecular Biology and Evolution*, **28**, 2731–2739.

776 Tan, K.H. (2003) Interbreeding and DNA analysis of sibling species within the *Bactrocera dorsalis*
777 complex. *Recent trends on sterile insect technique and area-wide integrated pest*
778 *management - economic feasibility, control projects, farmer organization and Bactrocera*
779 *dorsalis complex control study*: pp. 113–122.
780 Tan, K.H., Tokushima, I., Ono, H., *et al.* (2011) Comparison of phenylpropanoid volatiles in male
781 rectal pheromone gland after methyl eugenol consumption, and molecular phylogenetic
782 relationship of four global pest fruit fly species: *Bactrocera invadens, B. dorsalis, B. correcta*
783 and *B. zonata*. *Chemoecology*, **21**, 25–33.
784 van der Vecht, J. (1961) *Hymenoptera Sphecoidea Fabriciana*. EJ Brill.
785 Van Mele, P., Vayssières, J-F., Adandonon, A., *et al.* (2009) Ant cues affect the oviposition behaviour
786 of fruit flies (Diptera: Tephritidae) in Africa. *Physiological Entomology*, **34**, 256–261.
787 Vayssières, J.-F., Goergen, G., Lokossou, O., *et al.* (2005) A new *Bactrocera* species in Benin among
788 mango fruit fly (Diptera: Tephritidae) species. *Fruits*, **60**, 371–377.
789 White, I.M. & Elson-Harris, M.M. (1992) *Fruit flies of economic significance: their identification and*
790 *bionomics*, Wallingford UK: CAB International.
791 Wright, S. (1943) Isolation by distance. *Genetics*, **28**, 114–138.
792 Yeates, D.K., Seago, A.E., Nelson, L.A., *et al.* (2011) Integrative taxonomy, or iterative taxonomy?
793 *Systematic Entomology*, **36**, 209–217.
794 Zeldich, M.L., Swiderski, D.L. Sheets, H.D, & Fink, W.L. (2004) *Geometric Morphometrics for*
795 *Biologists: A Primer*. Elsevier Academic Press, New York, U.S.A.

Table 1. Collection data for newly acquired specimens of *Bactrocera dorsalis* and *Bactrocera invadens* used for morphological and molecular analyses in the current study.

					Molecular				Morphological			
Country	**Location**	**Latitude**	**Longitude**	**Date**	**COI**	**ND4-3**	**ITS1**	**ITS2**	**Aedeagus**	**Lateral vittae**	**Scutum**	**Wing shape**
Benin	Mts. Kouffé	8.73	2.07	25 Sept 2005	5	5	5	2	20	20	20	20
D.R. Congo	Kisangani	0.52	25.2	March-April 2011	4	4	5	4	20	20	20	20
Kenya	IAEA Colony	n/a	n/a	Initiated March 2012**	5	5	4	4	20	20	20	20
Mozambique	Cuamba	-14.8	36.53	Nov 2007 - Jan 2008	4	4	5	4	20	20	20	20
Sudan	Singa	13.16	33.96	Aug-Oct 2009	4	4	5	5	20	20	20	20
Pakistan	Islamabad	33.72	73.05	25 July 2012	9	8	10	7	20	20	20	20
Nepal	Dhankuta	27.00	87.33	5 Sept-20 Oct 2012	10	9	10	10	20	20	20	20
India	Bangalore	13.05	77.60	May 1992	0	0	0	0	0	0	0	20
India	Dahanu	19.76	72.97	Nov 2012	9	9	9	7	20	20	20	0
Sri Lanka	Central Province	7.27	80.64	May 2007	9	10	10	9	20	20	20	20
China	Wuhan	30.58	114.30	1 Nov 2012	5	5	4	5	20	20	20	20
China	Yunnan	n/a	n/a	Nov 2012	5	5	3	5	0	0	0	0
Myanmar	n/a	n/a	n/a	Nov 2012	10	10	9	8	0	0	0	0

Figure captions

Figure 1. Geographic distribution of scutum phenotypes for *Bactrocera dorsalis* and *Bactrocera invadens* from 1) Benin, 2) Democratic Republic of Congo, 3) Kenya, 4) Mozambique, 5) Sudan, 6) Pakistan, 7) Nepal, 8) India, 9) Sri Lanka, 10) Thailand, 11) Taiwan, 12) Malaysia, and 13) China. Twenty specimens of either species were examined per location with relative proportion of pale, intermediate, or fully black scutums shown.

Figure 2. Morphometric results for (A) post-sutural lateral vittae width, (B) aedeagus length, and (C) wing centroid size for *Bactrocera dorsalis* and *Bactrocera invadens* collected from Africa and Asia. $N = 20$ for each location with error bars 1 standard error about the mean. Locations sharing the same letter are not statistically different ($\alpha = 0.05$) following ANOVA with *post hoc* Tukey test.

Figure 3. Plot of first two variates following canonical variate analysis of geometric morphometric wing shape data for *Bactrocera dorsalis* and *Bactrocera invadens* collected from Africa, Indian subcontinent, and Eastern Asia. Twenty wings were analysed per location, with respective regions shaded in each of the three plots.

Figure 4. Regression of Mahalanobis distance (between pairwise geographic localities generated from CVA of wing shape data) against geographic distance (km) for *Bactrocera dorsalis* and *Bactrocera invadens* collected from the Indian subcontinent and eastern Asia.

Figure 5. Phylogenetic tree generated from Bayesian and Maximum Likelihood analysis for *Bactrocera dorsalis*, *Bactrocera invadens*, and outgroups. *East Asian specimens include *Bactrocera papayae* and *Bactrocera philippinensis*. Nodal supports presented for each analytical approach and for both 2/4 and 4/4 loci analyses. Ingroup specimens from the Indian subcontinent, Africa, and eastern Asia are highlighted in light grey, dark grey, and black, respectively. Specimen identities have been removed for clarity (provided in supplementary figure 1). Inset figure shows scutum colour pattern mapped onto *Bactrocera dorsalis/invadens* clade.

Figure 6. Median joining haplotype network generated from *cox1* sequence data of *Bactrocera dorsalis, Bactrocera invadens*, and *Bactrocera kandiensis* collected from Africa, the Indian subcontinent, and eastern Asia.

Figure 7. Holotype of *Dacus ferrugineus* Fabricius located in the Natural History Museum of Denmark. While nearly entirely destroyed, the taxonomically informative 'red-brown' colour of the thorax is still evident. Photo credit: Verner Michelsen.

Supplementary Figure S1. Phylogenetic tree generated from Bayesian and Maximum Likelihood analysis for *Bactrocera dorsalis, Bactrocera invadens,* and outgroups. Nodal supports presented for each analytical approach and for both 2/4 and 4/4 loci analyses. Ingroup specimens from the Indian subcontinent, Africa, and eastern Asia are highlighted in light grey, dark grey, and black, respectively.

Supplementary Figure S2. Median joining haplotype network generated from *cox1* sequence data of *Bactrocera dorsalis* and *Bactrocera invadens* collected from Africa, the Indian subcontinent, and eastern Asia. Different colours represent different scutum phenotypes from pale brown, to intermediate, to fully black.

Supplementary Figure S3. Median joining haplotype network generated from *cox1* sequence data of *Bactrocera dorsalis* and *Bactrocera invadens* collected from Africa, the Indian subcontinent, and eastern Asia. Different colours represent different collection countries.