The geometric approach to explore the *Bactrocera tau* complex (Diptera: Tephritidae) in Thailand

Sangvorn Kitthawee[1] and Jean-Pierre Dujardin [2,3]

[1] *Department of Biology, Faculty of Science, Mahidol University, Rama VI Rd., Bangkok 10400, Thailand, grskt@mahidol.ac.th, Tel: +662 201–5276, Fax: +662 354–7161* [2] *Center of Excellence for Vectors and Vector-borne Diseases (CVVD), Faculty of Sciences, Mahidol University (Thailand).* [3] *UMR IRD–CNRS 2724, Agropolis, 911, BP 64501, 34394, Montpellier (France).*

Abstract

1   Specimens of the genus *Bactrocera* were collected from several host plants in northern
2   and western Thailand. They were morphologically recognized as *Bactrocera tau* and were
3   subdivided according to host plant, geographic origin and time of collection, into eleven
4   samples.

5   A total of 264 males and 276 females were described at 12 landmarks of the right
6   wing. An exploratory analysis using kernel density estimates (KDE) was performed
7   on the multivariate morphometric space. Non-parametric classification highlighted the
8   existence of two non-overlapping clusters within both males and females. The clusters
9   were not congruent with geography. One cluster (cluster I) contained only one plant,
10 *Momordica cochinchinensis*, the other one (cluster II) contained five different plants
11 including *M. cochinchinensis*.

12 Further morphometric analyses on selected samples indicated that the influence of
13 plants on the shape of the wing could not explain satisfactorily the presence of two
14 clusters. Genetic techniques identified the presence of *B. tau* cryptic species C in *M.*
15 *cochinchinensis* from cluster I, and of *B. tau* cryptic species A in *Coccinia grandis* from
16 cluster II.

17 Our working hypothesis is that the two clusters identified by geometric morphomet-
18 rics were respectively species A and C.

19

Key words: *Bactrocera tau* complex, Thailand, morphometrics, host plants, species
identification.

1. Introduction

Fruit flies are the world's worst pests of fruits causing enormous economic loss every year (White and Elson-Harris, 1992; Aluja et al., 1996; Armstrong and Jang, 1997). Tephritid flies of the genus *Bactrocera* (Family Tephritidae) are of particular concern throughout much of Asia and Australia, where they constitute a significant threat to agricultural resources (Nagappan et al., 1971; Fletcher, 1987; Han et al., 1994; White, 1996; Kinnear et al., 1998; Kim et al., 1999).

In Thailand and other Southeast Asian (SEA) countries, the genus *Bactrocera* is known for being one of the major pests of tropical fruits and vegetables (Hardy, 1973; Drew and Romig, 1997). Most host plants belong to the family Cucurbitaceae, e.g. species of *Coccinia*, *Cucurbita*, *Cucumis*, *Luffa*, *Momordica*, *Trichosanthes*, etc. (Hardy, 1973; White and Elson-Harris, 1992). Various species are of great concern, such as *Bactrocera dorsalis* (Hendel), the oriental fruit fly, infesting a very wide range of fruits (Drew, 1989; Baimai et al., 2000); *Bactrocera* (Zeugodacus) *tau* (Walker), infesting a more restricted range of host plants, and the melon fly *Bactrocera cucurbitae* (Areekul, 1986; Yang et al., 1994). Compared with *B. cucurbitae*, *B. tau* is a more destructive species, especially in Taiwan and China (Yang et al., 1994; Chen, 2001).

Morphological variation within *B. tau* led some authors to suspect the presence of various species within the taxon (Hardy, 1973; White and Elson-Harris, 1992). Particularly, Drew and Romig (1997) suggested that *B. tau* is a large complex of sibling species in the SEA region. Cytogenetic study (Baimai et al., 2000), multilocus enzyme electrophoresis (MLEE) (Saelee et al., 2006) and DNA studies (Jamnongluk et al., 2003; Thanaphum and Thaenkham, 2003) confirmed that theory and recognized at least 7 species in the *B. tau* complex in Thailand. The members of the complex are cryptic species or morphologically very close, and have been labeled as species A, C, D, E, F, G and I (Baimai et al., 2000), with species A being *B. tau sensu stricto*.

Although the *B. tau* members were well classified by cytogenetics, MLEE and DNA techniques, their systematics still require an intensive development. Using the non-parametric kernel density estimates and the principal component analyses of shape, the present study explores the venation geometry of the wings as a character employable for pattern recognition. Examining the clusters as defined in our dataset by the kernel

2

density technique, we suggest that the geometric approach can help in the identification of cryptic taxa, and raises interesting questions about the possible effect of host plants and species competition on morphology.

2. MATERIALS AND METHODS

2.1. Insect samples

Oviposited eggs and larvae of the *B. tau* complex were collected from infested fruits of five host plant species in the family Cucurbitaceae: *Coccinia grandis* (CG), *Cucurbita moschata* (CMo), *Cucumis sativus* (CS), *Momordica cochinchinensis* (MC) and *Trichosanthes tricuspidata* (TT). They were obtained from three localities: Nan (NA) and Chiangmai (CM) in northern Thailand, and Kanchanaburi (KN) in western Thailand (Fig. 1). Fruits with ovipositional scars or marks of larval infestation were collected and kept in the laboratory with a code indicating location, host plant and time of collection (Table 1). The temperature of the laboratory was maintained at 27 ± 2 degrees Celsius, with $70\pm10\%$ relative humidity and a photoperiod of 12L:12D. Newly emerged adults were reared in transparent plastic cages ($12\times33\times18$cm). They were provided with 10% honey distilled water solution and sugar mixed with yeast hydrolyzate for at least two weeks to ensure all morphological characters developed well, especially the color and shape of abdominal bands typical of *B. tau* complex. The population density by fruit was not scored.

2.2. Specimen preparation and data collection

Left and right wings of the specimens were removed with forceps and mounted in Hoyer medium on glass microscopic slides. All slides were photographed by using a dissection stereo-microscope connected to a digital camera system with a 4x lens (40x). Twelve landmarks were digitized on the wings (Fig. 2) according to "type I" classification (venation intersections) (Bookstein, 1991). Only right wing was used unless damaged, in which case the left wing was used. To avoid possible optical distortion at the periphery of optical lens, each wing was located at the center of the visual field (Caro-Riaňo et al., 2008).

3

2.2.1. Repeatability

To reduce error at digitizing the landmarks, the same person collected the landmarks for the totality of the wings. The precision was estimated by comparing two sets of measurements on a subset of 42 individuals (21 males and 21 females). It was computed as the "repeatability" index (R) (Arnqvist and Mårtensson, 1998) of the first two principal components of shape ("relative warps", or RW, see 2.3.2), where R is provided by the ratio of the between-individual variance and the total variance.

2.3. Morphometric variables

2.3.1. Centroid size

We used the isometric estimator of size derived from coordinates data. This estimation of size is known as "centroid size" (CS). It is defined as the square root of the sum of the squared distances between the center of the configuration of landmarks and each separate landmark (Bookstein, 1991). The CS of different groups and sexes were compared by non-parametric analyzes based on permutations (1000 runs) allowing to compare both means and variances of size (Caro-Riaño et al., 2008).

2.3.2. Shape variables

Shape variables were obtained through the Generalized Procrustes Analysis (GPA) superimposition algorithm and subsequent projection of the Procrustes residuals into an euclidean space (Rohlf, 1999). Both non-uniform ("partial warps", strictly speaking) and uniform components (Rohlf, 1990) were used as shape variables[1]. These variables actually describe the differences in shape as deviations from an average configuration of landmarks, and their principal components (RW) are commonly used to illustrate the "morphological space", or "morphospace" (Bond et al., 2003). The shape variables were produced separately for males and females.

2.4. Kernel density estimates

Kernel density estimates (KDE) may be seen as a form of visual non-parametric spatial clustering. We used the bivariate density estimation as provided by the JMP®

---

[1]The uniform component describes stretching, compression or scission (global variation), and the non-uniform component corresponds to changes that occur at specific regions.

software (SAS Institute, 1995). It models a smooth surface that describes how dense the data points are at each point in that surface, and adds a set of contour lines showing the density. Local densities were estimated around each point of the morphospace defined by the two first RW. The resulting classification table was produced according to the modal clustering algorithm of the JMP® software (SAS Institute, 1995).

2.5. Allometry

The relative warps (RW) describing shape are obtained by removing the isometric change of size and are not necessarily free of allometric changes. The allometric residue of shape can be estimated in various ways, one of them is by performing a linear regression of the RW on centroid size. We chose the first relative warp because it was the component of shape that actually revealed data structuring (RW1, see Results). Thus, a linear regression analysis was performed between the first relative warp (RW1) as dependent variable and the centroid size as independent variable.

2.6. Morphometrics according to host plants

To examine shape variation according to host plants, we used only specimens having the same geographic origin (NA). These specimens allowed to study the effect of three different plants (CS, MC and TT), and were processed by taking into account, or not, the classification produced by the KDE. Results were illustrated by principal component analyses.

2.7. Genetic species identification

We used the mitochondrial DNA sequencing as described in Jamnongluk et al. (2003). The sequences of the COI gene were obtained from 12 specimens randomly selected from the clusters as classified by KDE: 7 specimens (2 males and 5 females) belonging to KN(CG)-26 (cluster II), and 5 specimens (2 males and 3 females) from KN(MC)-27 and NA(MC)-16 (both from cluster I). For each of these specimens, the percentage of identity was computed with the gene bank sequences of species A (AF400067.1) and species C (AF400069.1). According to Jamnongluk et al. (2003), these two species are genetically very distinct.

5

2.8. Software

Data collection, analyses and graphical output were performed using the various modules (COO, TET, MOG, COV and VAR) of specialized software developed by one of us (JPD) and freely available at http://www.mpl.ird.fr/morphometrics. COO was used to digitize landmarks on wings, TET to format the output for further analyses, MOG for performing the Generalized Procrustes Analysis, and then producing partial and relative warps, COV to perform regression analyses, and VAR to compare size among groups. The JMP® software (SAS Institute, 1995) was used to perform the bivariate density estimates and its graphical output.

3. RESULTS

3.1. Repeatability

The values obtained from the two first RW were 99% and 99% respectively, averaging to 99% for a total contribution of 60% to shape variation. For the set of eight first RW representing 91% of the total variation, the average repeatability was 91%. These values suggest the good quality of landmark recognition (Arnqvist and Mårtensson, 1998). On other insects, lower values of repeatability were obtained (Caro-Riaño et al., 2008).

3.2. Shape variation

The morphological shape space was illustrated by the contour plots based on kernel density estimates (KDE). Two large (non-overlapping) clusters were obtained within both males (Fig. 3) and females (Fig. 4). The clusters were not congruent with geography. In males as well as in females, one cluster contained insects from one plant only, *M. cochinchinensis* (MC), this cluster was labeled "I". Another cluster contained insects infesting five different plants including *M. cochinchinensis*, and was labeled "II" (Table 2). Five individuals were not included in the clusters (0.7% of males and 1.1% of females): they could correspond to undetected geometric anomalies making them behave as statistical outliers and were not considered further in this study.

3.2.1. Host plant effects

To study possible host effects on wing shape, we selected the specimens from the same geographic locality (NA, see Fig. 1). The specimens assigned to the cluster II assembled insects infesting three host plants: CS, MC and TT. Shape variation showed a consistent separation between the specimens infesting *M. cochinchinensis* and the ones from the other plants (Fig. 5, left). Performing the same analysis with additional specimens from NA belonging to cluster I, the resulting morphological space clearly showed two major groups corresponding to the two clusters, I and II (Fig. 5, right). Thus, in spite of one common host plant (MC), insects from the same geographic locality were still separated according to the clusters I and II.

7

3.3. Size variation

The centroid size variation was explored according to clusters defined by shape classification. The size of specimens from cluster I was significantly larger than the size of specimens from the other cluster, but the level of overlapping would not allow clear separation between clusters (Fig. 6).

Females were significantly larger on average than males ($P < 0.01$). As for shape, the comparison of mean size with respect to host plants was possible only for specimens assigned to the cluster II. Consistent differences were observed (Fig. 7).

The variance of size also presented significant variations. It was significantly different between clusters and sexes ($P < 0.05$ after Bonferroni correction), except between males and females of the cluster II. Between plants no significant difference in variance was found, except a larger variance in female insects collected from the CMo plant.

The only possible comparison of mean size according to geography exclusively was performed on the cluster I infesting the same plant (MC). The comparison was thus performed between NA, CM and KN: no significant difference was observed (details not shown).

3.4. Allometry

The regression of the RW1 on the centroid size was found to be statistically significant in males and in females ($P < 0.001$), with higher coefficient of determination (r^2) in males. For the total sample (Figs. 3 and 4), the coefficient of determination reached 43% in males and 21% in females. For the samples coming from NA, it was 27% in males (Fig. 5) and 22% in females.

3.5. Genetic species identification

From cluster II, both males showed 98% identity and females showed 97% to 98% identity with *B. tau* A. Conversely, the specimens randomly selected from cluster I showed high values of identity with species *B. tau* C (95% and 97% for the two males, 91%, 98% and 99% for the three females).

4. DISCUSSION

The multivariate analyses commonly used in morphometrics need the intervention of the user either previous to the analysis (discriminant anlysis) or after it (principal component analyses). Clustering techniques are able to identify groups without such intervention, they are based only on morphological similitude excluding any a priori taxonomic information.

The main result provided by our exploratory analysis was the existence in both sexes of two separate clusters within the total sample. These two clusters corresponded to distinct wing shapes, not to different host plants since they remained separate even in the same host plant. Although contributing to shape variation, size showed frequent overlapping between clusters, except when found in the same host plant. Because of the existence of two distinct shapes, and as confirmed by the genetic characterization of a few specimens, our data suggest that two probable taxa were present within the dataset.

4.1. Classification methods

Size showed variation seemingly according to host plants and sexes (Fig. 7), although we could not reject other factors such as the field conditions where fruits were collected, the larval density within fruits, and other uncontrolled parameters responsible for phenotypic plasticity. According to regression analyses, size significantly contributed to shape differences between clusters. However, the important overlapping of size observed between groups, especially between clusters (Fig. 6), indicates it would not have been able to identify the two main clusters observed. As an exception, size was not overlapping between clusters when found in the same fruit (Fig. 7, see NA(MC) and NA(MC) Cluster I), which could suggest character displacement (Grant, 1972), but we lack detailed information to fully confirm this hypothesis (Losos, 2000).

The geometry of the wing, even containing the influence of size variation, appeared as a more interesting candidate for classification. The morphological space described by the two first relative warps (RW1 and RW2, see Figs. 3 and 4) disclosed two separate clusters automatically recognized by the kernel density analysis, and completely determined on RW1. We used this technique of classification for its objectivity and intuitive simplicity,

9

₂₂₅ but other classification techniques exist which could be used together or alternatively

₂₂₆ (Baylac et al., 2003; Chengpeng et al., 2007).

4.2. The members of Bactrocera tau complex

₂₂₈   Our genetic classification recognized two species, *B. tau* C and A, in accordance with

₂₂₉ clusters I and II. This result led us to suggest the presence of these two species as the

₂₃₀ most parsimonious explanation for having two separate clusters in the morphospace of

₂₃₁ the wings. What was the probability to have another of the known seven species of the

₂₃₂ complex (Baimai et al., 2000) in the total sample ? A very first argument against the

₂₃₃ idea of more than two species in our total sample is that when submitted to a principal

₂₃₄ component analysis of superimposed shape coordinates, two well separated groups were

₂₃₅ apparent in each sex (Figs. 3 and 4). The second argument is that no genetically close

₂₃₆ species (hence likely morphologically close, also) to either A or C, were expected to be

₂₃₇ found in our sample. In the COI gene (Jamnongluk et al., 2003), allozyme (Saelee et al.,

₂₃₈ 2006) or heat shock protein classifications (Thanaphum and Thaenkham, 2003), there is

₂₃₉ no close species to A, but according to cytogenetics a close species is species E (Baimai

₂₄₀ et al., 2000). The latter was not likely to be present in our sample since it is specific

₂₄₁ to a host plant species we did not collect, *Strychnos thorellii* (Strychnaceae). According

₂₄₂ to COI gene and allozyme classifications, the species C constitutes a separate group

₂₄₃ together with species I (Jamnongluk et al., 2003; Saelee et al., 2006). Species I is also

₂₄₄ specific to another plant not collected by us, and its territory lies outside the geographic

₂₄₅ areas considered here. Finally, according to cytogenetics, another species, D, is close to

₂₄₆ C, but it is allopatric and a more southern species (Baimai et al., 2000) that has never

₂₄₇ been reported in the areas of our collections.

4.3. Host plant specificity

₂₄₉   As already observed by Baimai et al. (2000), the *M. cochinchinensis* cluster I, or

₂₅₀ putative species C, seems specific to that plant since it was collected from it in northern

₂₅₁ (NA, CM) and western (KN) Thailand. The other cluster (cluster II), which is the

₂₅₂ putative species A, was collected from five different host plants, indicating its more

₂₅₃ generalist behavior (Table 1). According to the host plant, shape and size showed some

₂₅₄ variation (Fig. 5, left; Fig. 7). This variation could be related to the host plants or

could be attributed to other uncontrolled conditions generating phenotypic plasticity. The possible influence, if any, of the host plant on insect morphology, could not produce any kind of morphological convergence between clusters developing in the same fruit. Putative species A is indeed able to develop in the same plant (MC) as the one specific to putative species C. Such situation did not make the two species more alike; on the contrary, the common infestation apparently resulted in non-overlapping sizes (Fig. 7, see NA(MC) and NA(MC) Cluster I), as well as completely separate shapes (Fig. 5, right). Although more difference in sympatry than in allopatry is suggestive of character displacement (Grant, 1972), we have no sufficient evidence allowing to clearly confirm that hypothesis (Losos, 2000). In our data, the possible influence of geographic isolation was not apparent.

5. Conclusion

Our results suggest that the geometric morphometric approach could represent a low-cost and fast technique to explore the *B. tau* complex. It appears as an informative tool to describe intraspecific variation related to host plants and/or other factors affecting morphology, as well as a promising candidate to reliably discriminate cryptic species A and C of *B. tau* in Thailand. The present study showed the existence of two different wing shapes within our sample of *B. tau* and provided arguments to relate them to species A and C, it did not test the methods specific to morphometric discrimination between them (Baylac et al., 2003).

11

6. Acknowledgments

References

Aluja, M., Hurtado, C., Liedo, P., Cabrera, M., Castillo, F., Guillen, J., Rios, E., 1996. Seasonal population fluctuations and ecological implications for management of *Anastrepha* fruit flies (Diptera: Tephritidae) in commercial mango orchards in Southern Mexico. Journal of Economical Entomology, 89(3): 654-667.

Areekul, S., 1986. Toxicity tests of some insecticides against fruit flies. II. *Dacus cucurbitae* Coquillett and *Dacus tau* (Walker). Kasetsart Journal of Natural Sciences, 20: 274-279.

Armstrong, J. W., Jang, E. B., 1997. An overview of present and future fruit fly research in Hawaii and U.S. mainland. In A.J. Allwood and R.A.I. Drew (Eds.). Management of Fruit Flies in the Pacific. Canberra: Australian Center for International Agricultural Research Proceeding Number, 76: 30-42.

Arnqvist, G., Mårtensson, T., 1998. Measurement error in geometric morphometrics: empirical strategies to assess and reduce its impact on measure of shape. Acta Zoologica Academiae Scientarum Hungaricae, 44(1-2): 73-96.

Baimai, V., Phinchongsakuldit, J., Sumrandee, C., Tigvattananont, S., 2000. Cytological evidence for a complex of species within the taxon *Bactrocera tau* (Diptera: Tephritidae) in Thailand. Biological Journal of the Linneaus Society, 69(3): 399-409.

Baylac, M., Villemant, C., Simbolotti, G., 2003. Combining geometric morphometrics with pattern recognition for the investigation of species complexes. Biological Journal of the Linnean Society, 80 (1): 89-98.

Bond, J., Beamer, D. A., Hedin, M. C., Sierwald, P., 2003. Gradual evolution of male genitalia in a sibling species complex of millipedes (Diplopoda: Spirobolida: Rhinocricidae: Anadenobolus). Invertebrate systematics, 17 (6): 711-717.

Bookstein, F. L., 1991. Morphometric tools for landmark data: Geometry and Biology. Cambridge University Press, Cambridge, 435 pp.

Caro-Riaño, H., Jaramillo, N., Dujardin, J. P., 2008. Growth changes in *Rhodnius pallescens* under simulated domestic and sylvatic conditions. Infection, Genetics and Evolution, 9(2):162-8.

Chen, S., 2001. The development and reproduction of pumpkin fly (*Bactrocera tau* (Walker)) (Diptera: Tryphritidae) and field occurrence investigation. Plant Protection Bulletin, Taichung, 43: 137-151.

Chengpeng, B., Saunders, M. C., McPheron, B. A., 2007. Wing pattern-based classification of the *Rhagoletis pomonella* species complex using genetic neural networks. International Journal of Computer Science and Applications, 4(3):1-14.

Drew, R. A. I., 1989. The tropical fruit flies (Diptera: Tephritidae: Dacinae) of the Australasian and Oceanian regions. Memoirs of the Queensland Museum, 26: 1-521.

Drew, R. A. I., Romig, M. C., 1997. Overview - Tephritidae in the Pacific and South-East Asia. In A.J. Allwood and R.A.I. Drew (Eds.). Management of Fruit Flies in the Pacific. Canberra: Australian Center for International Agricultural Research Proceeding Number, 76: 46-43.

Fletcher, B., 1987. The biology of dacine fruit flies. Annual Review of Entomology, 32: 115-44.

Grant, P. R., 1972. Convergent and divergent character displacement. Biological Journal of the Linnean Society, 4: 39-68.

13

Han, M. J., Lee, S. H., Ahn, S. B., Choi, J. Y., Choi, C. M., 1994. Distribution, damage and host plant of pumpkin fruit fly, *Paradacus depressa* (Shiraki). RDA (Rural Development Administration). Journal of Agricultural Science, 36: 346-350.

Hardy, D., 1973. The fruit flies (Tephritidae–Diptera) of Thailand and bordering countries. Pacific Insects Monograph, 31: 1–353.

Jamnongluk, W., Baimai, V., Kittayapong, P., 2003. Molecular phylogeny of tephritid fruit flies in the *Bactrocera tau* complex using the mitochondrial COI sequences. Genome, 46(1): 112-118.

Kim, T. H., Kim, J. S., Mun, J. H., 1999. Distribution and bionomics of *Bactrocera depressa* (Shriaki) in Chonbuck province. Korean Journal of Soil Zoology, 4: 26-32.

Kinnear, M., Bariana, H., Sved, J., Frommer, M., 1998. Polymorphic microsatellite marker for population analysis of a Tephritid pest species, *Bactrocera tryoni*. Molecular Ecology, 7: 1489-1495.

Losos, J. B., 2000. Ecological character displacement and the study of adaptation. PNAS, 97 (11): 5693-5695.

Nagappan, K., Kamalnathan, S., Santharaman, T., Ayyasan, M., 1971. Insecticidal trials for the control of the melon fly, *Dacus cucurbitae* Coq. infesting snake gourd, *Trichosanthes anguina*. Journal of Madras Agriculture, 58: 684-690.

Rohlf, F. J., 1990. Rotational fit (Procrustes) methods. In: Rohlf, F. J., Bookstein, F. L. (Eds.), Proceedings of the Michigan Morphometrics Workshop. Special Publiation Number 2. The University of Michigan Museum of Zoology. Ann Arbor, MI, p. 227-236. University of Michigan Museum, Ann Arbor, p. 380.

Rohlf, F. J., 1999. Shape statistics: Procrustes superimpositions and tangent spaces. Journal of Classification, 16: 197-223.

Saelee, A., Tigvattananont, S., Baimai, V., 2006. Allozyme electrophoretic evidence for a complex of species within the *Bactrocera tau* group (Diptera Tephritidae) in Thailand. Songklanakarin Journal of Science and Technology, 28(2): 249-257.

SAS Institute, I., 1995. JMP Statistics and Graphics Guide. Version 3.1. SAS Campus Drive. Cary, NC 27513. pp. 592.

Thanaphum, S., Thaenkham, U., 2003. Relationships of forms within the *Bactrocera tau* (Walker) (Diptera: Tephritidae) taxon based on heat shock protein 70 cognate sequences. Annals of the Entomological Society of America, 96(1): 44-53.

White, I., 1996. Fruit fly taxonomy: recent advances and new approaches. p. 253-258. In B.A. McPheron and G.J. Steck (Eds.). Fruit Fly Pests. A world assessment of their biology and management. St Lucie Press, Delray Beach, FL.

White, I., Elson-Harris, M., 1992. Fruit flies of economic significance; Their identification and bionomics. CAB International, Wallingford, UK.

Yang, P., Carey, J., Dowell, R., 1994. Tephritid fruit flies in China: Historical background and current status. Pan-Pacific Entomologist, 70: 159-167.

14

Host plant species	Locality	Code of the fruit	M	F
Coccinia grandis	Kanchanaburi	KN(CG)-26	16	16
Cucurbita moschata	Kanchanaburi	KN(CMo)-20	21	21
Cucurbita moschata	Kanchanaburi	KN(CMo)-30	21	20
Momordica cochinchinensis	Kanchanaburi	KN(MC)-27	21	35
Momordica cochinchinensis	Kanchanaburi	KN(MC)-31	20	14
Momordica cochinchinensis	Chiangmai	CM(MC)-1	67	82
Momordica cochinchinensis	Nan	NA(MC)-19	20	20
Momordica cochinchinensis	Nan	NA(MC)-16	21	21
Momordica cochinchinensis	Nan	NA(MC)-19/3	18	9
Trichosanthes tricuspidata	Nan	NA(TT)-38	19	17
Cucumis sativus	Nan	NA(CS)-32	20	21
			264	276

Table 1: Material of *Bactrocera tau* complex used in this study. M, males; F, females; KN, Kanchanaburi; CM, Chiangmai; NA, Nan; CG, *Coccinia grandis*; CMo, *Cucurbita moschata*; MC, *Momordica cochinchinensis*; TT, *Trichosanthes tricuspidata*; CS, *Cucumis sativus*. Numbers after plants abbreviations (-1, -16, -19, -19/3, -20, etc.) are codes referring to time of collection.

Code	Males N	Cluster I	Cluster II	Females N	Cluster I	Cluster II
KN(CG)-26	16	-	16	16	-	16
KN(CMo)-20	21	-	21	21	-	21
KN(CMo)-30	21	-	21	20	-	20
KN(MC)-27	21	20	-	35	35	-
KN(MC)-31	20	20	-	14	14	-
CM(MC)-1	67	45	22	82	40	41
NA(MC)-19	20	6	14	20	6	13
NA(MC)-16	21	21	-	21	21	-
NA(MC)-19/3	18	-	18	9	-	9
NA(TT)-38	19	-	18	17	-	16
NA(CS)-32	20	-	20	21	-	21
Total	262+2	112	150	273+3	116	157

Table 2: Morphometric classification of morphologically unidentified specimens into clusters I and II using bivariate density estimates. Only 2 males and 3 females were not included in the two clusters. KN, Kanchanaburi; CM, Chiangmai; NA, Nan; CG, *Coccinia grandis*; CMo, *Cucurbita moschata*; MC, *Momordica cochinchinensis*; TT, *Trichosanthes tricuspidata*; CS, *Cucumis sativus*. Numbers after plants abbreviations are codes referring to time of collection.

Figure 1: Geographic origin of the fruit flies: 186 flies from Nan (NA) and 149 flies from Chiangmai (CM) in northern Thailand; 205 flies from Kanchanaburi (KN) in western Thailand.
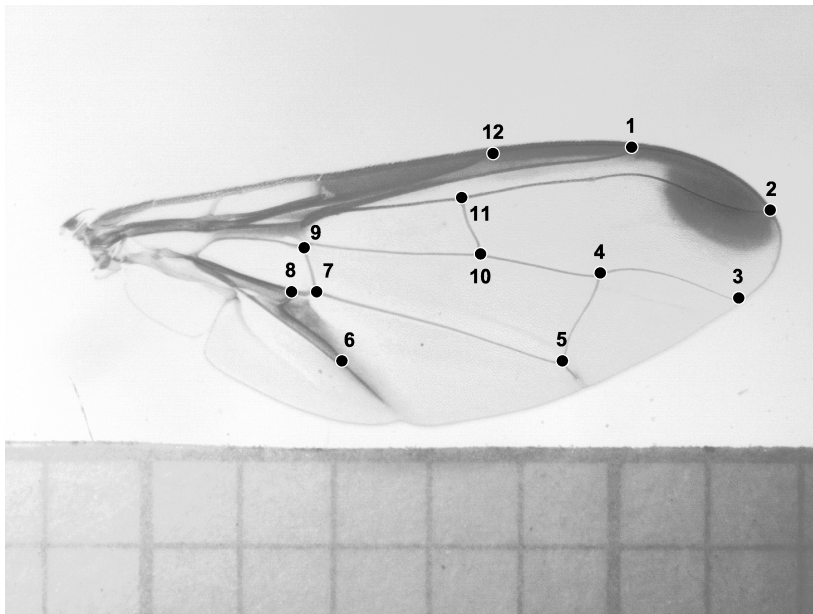
Figure 2: Fore wing of *Bactrocera tau* showing the 12 landmarks whose coordinates were used in morphometric analyses. Each landmark is the junction of two different veins, as required by Type I landmarks (Bookstein, 1991). Each picture contains a millimeter paper (see bottom) for true size scaling.
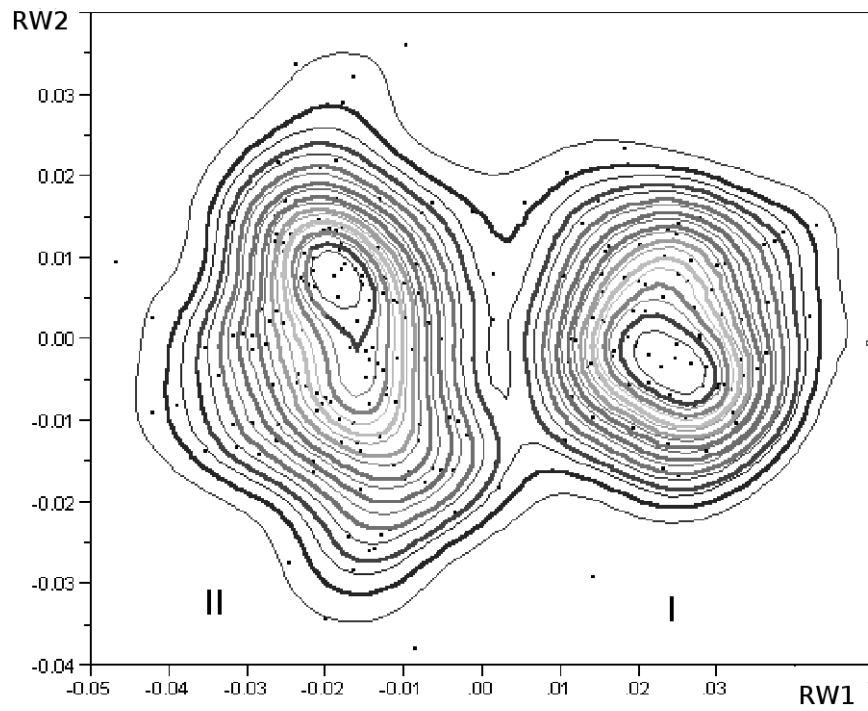
Figure 3: Morphospace of males. Bivariate density estimation models a smooth surface that describes how dense the data points are at each point in that surface, and adds a set of contour lines showing the density. The contour lines are quantile contours in 5% intervals with thicker lines at the 10% quantiles intervals. This means that about 5% of the points are below the lowest contour, 10% are below the next contour, and so forth. The highest contour has about 95% of the points below it. RW1 and RW2 stand for relative warps 1 and 2, i.e. first and second principal components of partial warps, respectively.
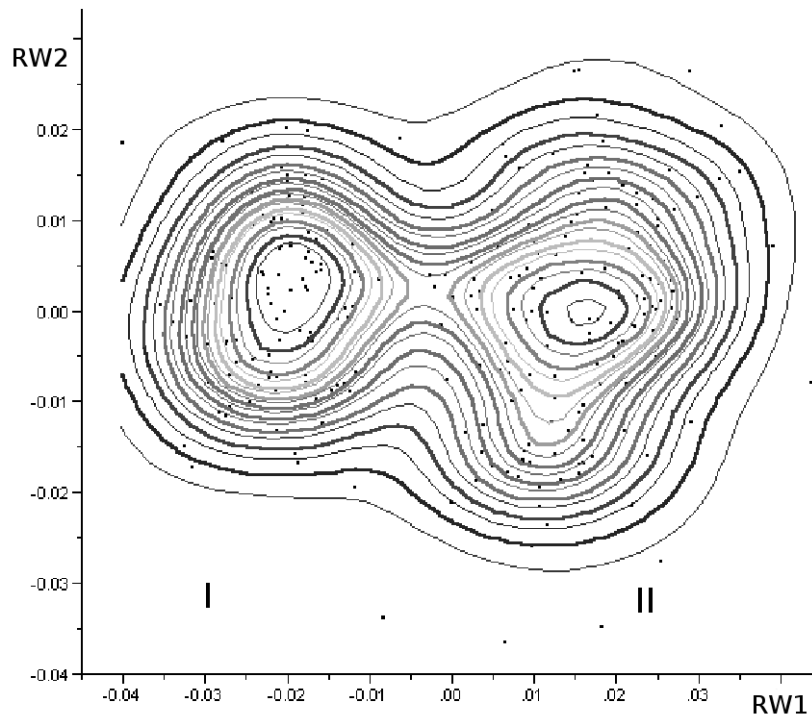
Figure 4: Morphospace of females. Bivariate density estimation models a smooth surface that describes how dense the data points are at each point in that surface, and adds a set of contour lines showing the density. The contour lines are quantile contours in 5% intervals with thicker lines at the 10% quantiles intervals. This means that about 5% of the points are below the lowest contour, 10% are below the next contour, and so forth. The highest contour has about 95% of the points below it. RW1 and RW2 stand for relative warps 1 and 2, i.e. first and second principal components of partial warps, respectively.
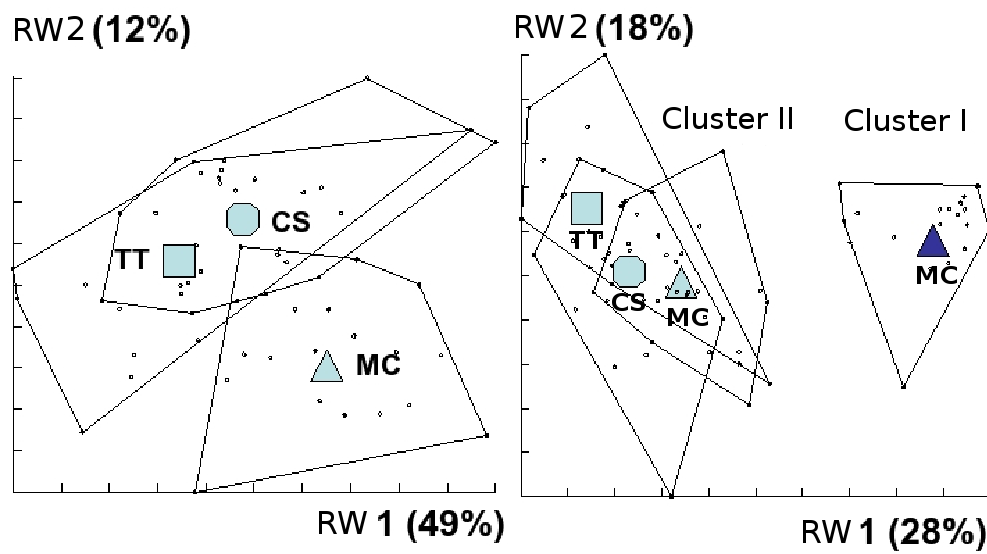
Figure 5: Morphological spaces of males from NA according to host plants: *C. sativus* (CS), *T. tricuspidata* (TT), and *M. cochinchinensis* (MC). Left: male specimens from cluster II, all of them coming from the same geographic origin (NA). Right: same composition as left part, plus cluster I male specimens coming from NA (cluster I exclusively infests MC). RW1 and RW2 stand for relative warps 1 and 2, i.e. first and second principal components of partial warps, respectively. The contribution of each RW to the total shape variation is given between parentheses. The same analyses on females produced similar results (female morphospace not shown)
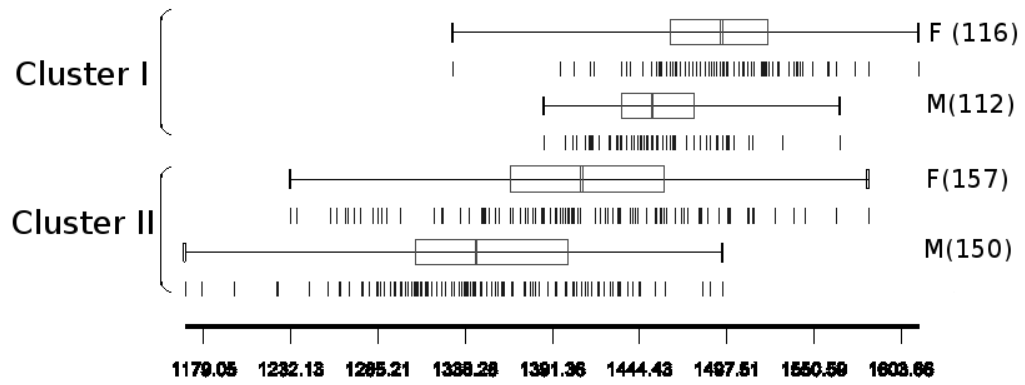
Figure 6: Centroid size variation among sexes and clusters (see Figs. 3 and 4) presented as quantile plots. Each box shows the median as a line across the middle and the quartiles (25th and 75th percentiles) as its ends. Units are pixels. M, males; F, females; between brackets: sample sizes.
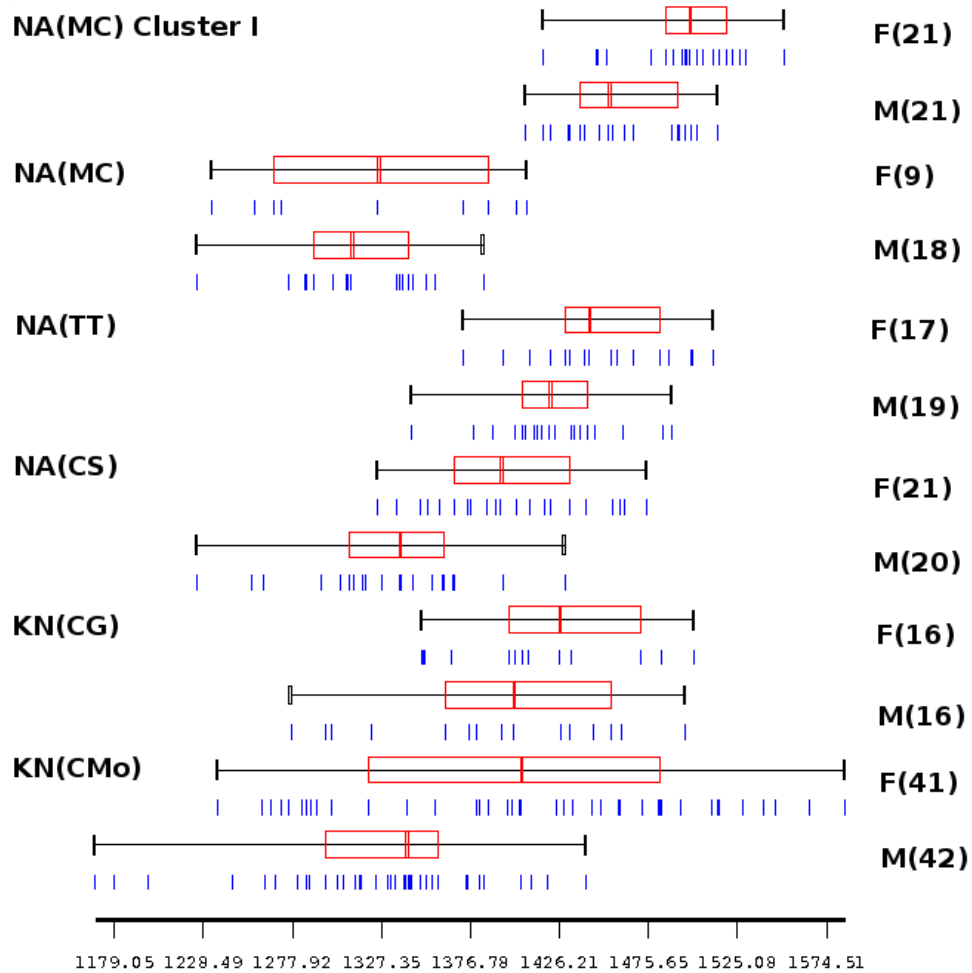
Figure 7: Centroid size variation among sexes and host plants. The variation is presented as quantile plots. Each box shows the median as a line across the middle and the quartiles (25th and 75th percentiles) as its ends. Except for the two first quantile plots named NA(MC) Cluster I (putative *B. tau* species C), all the quantile plots refer to specimens from cluster II (putative *B. tau* species A). Units are pixels. M, males; F, females; KN, Kanchanaburi; NA, Nan; CMo, *Cucurbita moschata*; CG, *Coccinia grandis*; MC, *Momordica cochinchinensis*; CS, *Cucumis sativus*; TT, *Trichosanthes tricuspidata*. Sample sizes are between brackets. The figure shows clearly that the size of the two clusters do not overlap in the same plant: see NA(MC) Cluster I versus NA(MC).

23