2

The Mosquito Fauna: From Metric Disparity to Species Diversity

Jean-Pierre Dujardin[1], P. Thongsripong[2] and Amy B. Henry[3]
[1]*UMR 5090 MIVEGEC, Avenue Agropolis, IRD, Montpellier*
[2]*Department of Tropical Medicine, Medical Microbiology, and Pharmacology, University of Hawaii at Manoa, Honolulu, Hawaii*
[3]*Department of Microbiology, University of Hawaii at Manoa, Honolulu, Hawaii*
[1]*France*
[2,3]*USA*

1. Introduction

Biodiversity, encompassing diversity of genes, species, and ecosystems, is fundamental to biology (Gaston & Spicer, 2004), yet tools to monitor it are insufficient. Biodiversity can be estimated by using the number of species (species richness) in a community, and/or by this number together with the proportion of species (species evenness), and/or by other more indirect estimators among which is morphological variation.

Studies on morphological and biological diversity have highlighted the complexity of the diversity structure, that is, the relationship between morphological and taxonomic diversity (Foote, 1992). Most studies acknowledge a certain level of dissociation between morphological diversity and species richness, suggesting that taxonomic and morphological diversity patterns are distinct ones (Foote, 1993; Moyne & Neige, 2007; Roy et al., 2001; Roy & Foote, 1997; Vasil'ev et al., 2010). Occasionally, the use of metric diversity as a proxy for species richness has been suggested (Dolan et al., 2006).

However, appropriate morphological disparity metrics and sets of morphological characters are not clearly defined (Navarro, 2003; Roy & Foote, 1997; Wills et al., 1994). Although most studies have used one or two measures of disparity to quantify and characterize the occupation of morphospace, multiple measures might be necessary to fully detect changes in patterns of morphospace occupation (Ciampaglio et al., 2001).

Moreover, organisms present an indefinitely large number of potentially quantifiable traits, and in practice only a small number of features can be studied. Therefore, we cannot strictly measure morphological diversity, but diversity with respect to some set of traits. The common practice is to seek broad coverage of morphology, using as many characters as possible. The present study is disputing this common practice, showing that the opposite strategy could be more informative: to seek for elementary coverage of morphology.

Our hypothesis is that when grouping many traits morphological variation becomes too complex to validly reflect a single factor like species richness. Organismal morphology is the result of many biological causes that are not only evolutionary factors but also environmental

and historical. Instead of considering a conglomerate of traits to " cover " the organismal morphology, we suggest to decompose the global morphology into more elementary units and test their variation against species richness.

A set of anatomical landmarks in the wing represents a suitable tool to explore that hypothesis since it can be decomposed into subsets of different landmarks. In our approach, the complete set of anatomical landmarks of the wing would represent the broadest morphological coverage, and its decomposition into smaller configurations of landmarks would provide more elementary units of morphology. Do different subsets (combinations) of landmarks reflect biodiversity in the same way as the total set? Landmark-based geometric morphometrics, which is applied here, provides a convenient way for measuring morphological variety, requiring only the recognition of homologous landmarks in all individuals under comparison. This condition applies very well to mosquito wings, because their venation pattern is almost identical among different species and higher taxa, including different tribes.

2. Materials & methods

2.1 The insects

We used a total sample of 480 individuals (one wing per individual). They were tentatively identified using available morphological keys (Rattanarithikul et al., 2005; Rattanarithikul, Harrison, Panthusiri, Peyton & Coleman, 2006; Rattanarithikul, Harrison, Harbach, Panthusiri & Coleman, 2006.; Rattanarithikul et al., 2010). A total of 10 genera and 43 species of Culicidae (mosquitoes) was found, with unidentified species pooled into one putative "species" (Table 1). This collection was a representative set of higher taxa within Culicidae: it contained indeed the two subfamilies (Anophelinae and Culicinae) and, within the Culicinae, 6 tribes out of 11.

2.2 Mosquito collection

The mosquito collection was done during the rainy season of 2008 (June-August) along a forest-agro-urban landscape gradient within Nakhon Nayok province, central Thailand. Six habitat types: forest, fragmented forest, rice field, rural, suburban, and urban were identified and characterized across the landscape gradient (Table 1). For each habitat type, four sites were picked as representative of similar habitat range. Forest sites were situated along the border of the pristine Khao Yai National Park. Fragmented forest sites were on the edge of disturbed vegetation patch not far from the National Park, where human settlements were sparse and traditional small-scale agricultures were practiced. Rice field sites were further away from the National park and situated in the lowland closer to the main river hence the big-scale and irrigated rice agricultures were possible. The rural, suburban, and urban sites distributed based on the distance from the centre of town. The rural sites were more than 7 km from town; the suburban were within 5 km from town; and the urban sites were in the center of the town.

Four types of adult mosquito trap: BG sentinel, Mosquito Magnet, CDC UV light traps, and CDC backpack aspirator, were used in order to maximize the variety of mosquito samples collected. In each site, mosquitoes were collected for 24 hours using 8 BG traps, 2 Mosquito Magnet traps, and 3 of 3 to 10 minute-long aspirations for the day trapping, and 8 UV light traps, 8 BG traps, and 2 Mosquito Magnet traps for the night trapping. A total of over

80,000 mosquitoes were collected. For the morphometric study, only a subset of the female mosquitoes (Table 1) were examined.

Genus	Species	*n*	F	FF	R	RF	SU	U
Aedes	*aegypti*	*12*	.	.	1	3	5	3
	albopictus	*7*	.	4	2	1	.	.
	lynetopennis	*14*	.	4	10	.	.	.
	mediolineatus	*12*	.	6	6	.	.	.
	vexans	*25*	.	7	16	.	2	.
	unknown	*20*	1	5	7	2	5	.
Aedomyia	*catasticta*	*1*	.	1	.	.	.	.
Anopheles	*baezai*	*1*	.	.	.	.	1	.
	barbirostris	*10*	2	3	.	1	3	1
	kochi	*6*	.	6	.	.	.	.
	minimus	*2*	.	2	.	.	.	.
	peditaeniatus	*5*	.	.	1	2	2	.
	phillippines	*5*	.	3	2	.	.	.
	tessellatus	*9*	.	4	.	.	5	.
	vagus	*37*	.	21	12	.	1	3
	unknown	*7*	.	7	.	.	.	.
Armigeres	*dentatus*	*1*	1	.	.	.	.	.
	magnus	*2*	2	.	.	.	.	.
	malayi	*1*	.	1	.	.	.	.
	subalbatus	*10*	2	4	1	.	1	2
	unknown	*2*	.	2	.	.	.	.
Coquillettidia	*crassipes*	*8*	.	.	.	8	.	.
	unknown	*5*	.	5	.	.	.	.
Culex	*bitaeniorhynchus*	*22*	5	4	1	4	7	1
	brevipalpis	*7*	.	2	2	2	.	1
	fuscocephala	*6*	.	5	.	.	.	1
	gelidus	*18*	.	4	1	5	4	4
	mocthogenes	*1*	.	1	.	.	.	.
	nigropunctatus	*12*	10	.	.	.	1	1
	quinquefasciatus	*11*	.	1	.	.	4	6
	sinensis	*28*	3	11	1	7	6	.
	tritaeniorhynchus	*1*	.	1	.	.	.	.
	vishnui	*53*	1	29	8	1	11	3
	unknown	*24*	2	14	1	3	2	2
Ficalbia	*minima*	*3*	.		3		.	.
Heizmania	unknown	*7*	7		.		.	.
Mansonia	*annulifera*	*5*	.	3	.		.	2
	indiana	*6*	.		.	1	5	.
	uniformis	*11*	.	7	.	1	3	.
	chamberlainai	*5*	.		3	1	1	.
	hybrida	*7*	2	2	2	1	.	.
	luzonensis	*7*	2	1	1	.	.	3
	metallica	*2*	.	.	.	.	2	.
	unknown	*1*	.	1	.		.	.

Genus	Species	*n*	F	FF	R	RF	SU	U
Uranotaenia	*campestris*	2	.	2	.	.	.	.
	lateralis	1	.	.	.	.	1	.
	lutescens	4	.	.	.	1	3	.
	micans	7	.	1	.	3	3	.
	nivipleuraura	3	2	.	.	.	1	.
	subnormalis	2	.	.	.	.	2	.
	unknown	22	17	2	1	1		1

Table 1. List of genera and species of Culicidae identified on morphological ground, with their repartition according to the habitat. F, forest; FF, fragmented forest; RF, rice field; R, rural; SU, semi-urban and U, urban. For statistical tests, unknown species have been pooled into one single taxon (92 specimens). *n*, number of specimens submitted to morphometric analyses.

2.3 Shape of the wing

The shape of the mosquito wings was described by 13 landmarks (see Fig. 1).

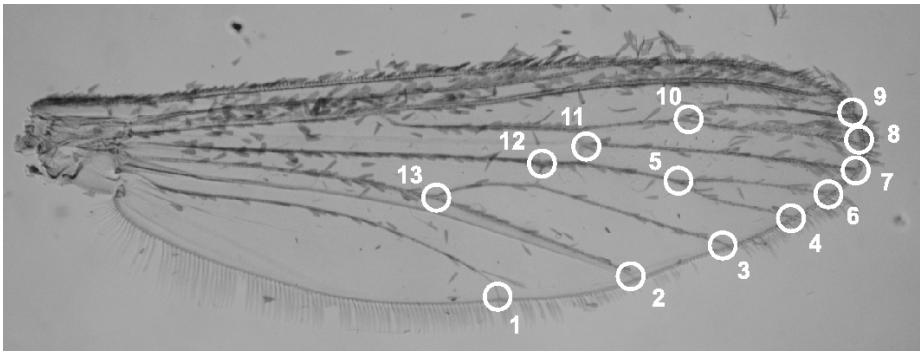


Fig. 1. Mosquito wing. Landmarks are labelled according to the order of digitization.

The decomposition of shape used a total of 254 different LM configurations out of the 7814 possible ones, i.e. the totality (13) of landmarks (LM), the 13 combinations of 12 LM, and 240 combinations of decreasing numbers of LM. For each subset of LM, going from 11 LM to 4 LM, 30 different combinations were tested (see Table 2).

The above testing of 254 LM configurations was replicated on ten different species sequences (see Table 3).

2.4 Species sequence

A sequence refers here to a succession of species assemblages with increasing richness. The mosquitoes were randomly sampled with replacement into 22 assemblages of increasing species richness. These 22 assemblages constituted a sequence where each unit represented a community with different species richness, starting from 2 species, 4 species, 6 species, and so on until 44 species. A partial representation of a sequence is shown Table 3; ten such sequences were constituted. The range of categorical units (44 species) was not modified.

	(4)	(5)	0	(10)	(11)
1	8-9-11-13	1-4-6-8-11	0	1-2-5-7-8-10-11-12-13	1-3-4-5-7-8-10-11-12-13
2	4-5-7-13	3-5-9-10-11	0	1-3-4-5-6-7-9-10-11-12	1-2-3-4-5-6-9-10-11-12-13
3	2-6-7-13	5-9-10-11-12	0	1-2-3-4-5-7-8-10-11-12	1-2-3-4-7-8-9-10-11-12-13
4	2-3-8-13	4-6-7-10-13	0	2-3-5-6-7-8-9-10-12-13	2-3-4-5-7-8-9-10-11-12-13
5	1-2-3-10	1-4-6-9-13	0	3-4-5-6-7-8-9-10-11-13	1-2-3-6-7-8-9-10-11-12-13
6	4-5-12-13	1-2-3-4-5	0	3-4-5-6-8-9-10-11-12-13	1-2-4-6-7-8-9-10-11-12-13
7	5-8-10-13	2-3-7-9-11	0	1-2-4-6-7-8-9-10-11-12	2-3-4-5-6-7-8-10-11-12-13
8	2-9-12-13	2-3-5-7-12	0	1-2-6-7-8-9-10-11-12-13	1-2-3-4-5-6-7-9-11-12-13
9	1-3-8-12	1-3-7-8-10	0	1-2-3-5-6-7-9-11-12-13	2-3-4-6-7-8-9-10-11-12-13
10	1-4-5-9	1-5-7-12-13	0	1-2-4-5-6-9-10-11-12-13	1-2-3-4-5-6-7-9-10-11-12
11	6-8-10-13	2-4-7-8-10	0	2-3-4-6-7-8-9-10-11-13	1-2-3-5-6-7-8-9-10-11-12
12	4-7-8-9	3-4-8-11-12	0	1-2-3-4-5-6-7-8-12-13	1-3-4-5-6-7-8-9-10-11-12
13	1-3-6-9	1-5-8-9-10	0	1-2-3-5-7-8-10-11-12-13	1-2-3-4-5-7-8-9-10-11-13
14	1-5-8-10	3-5-9-11-12	0	2-3-4-5-6-7-8-9-10-13	1-2-4-5-6-7-8-9-10-12-13
15	1-4-8-9	3-8-9-10-11	0	1-2-4-5-6-7-8-10-12-13	1-2-3-4-5-6-7-8-9-10-12
0	0	.	0	0	0
30	3-5-8-12	2-5-8-9-13	0	1-2-3-4-5-7-9-11-12-13	1-2-3-4-5-6-7-8-9-12-13

Table 2. A partial representation of the sets of landmarks used to examine a sequence of increasing taxonomic richness (see Table 3). Landmarks were numbered from 1 to 13 (Fig. 1). The first row refers to the number of landmarks (between brackets). The first column enumerates the 30 combinations of landmarks randomly generated for each number of landmarks. The second column partially shows the 30 combinations to represent 4 landmarks. To save space, only some combinations for 4, 5, 10 and 11 landmarks are represented here. To these 240 (30*8) combinations, we added the 13 combinations of 12 landmarks, and the total number of landmarks (13).

2.5 Species assemblage

In the building of sample sets of increasing taxonomic richness (a sequence), species were not " added " to previous ones, they were resampled at each step from the total available. This can be seen in the sequence shown Table 3. The program was not simulating a temporal variation where species progressively accumulate in a given environment, but a spatial sampling of taxa where groups represent communities of various taxonomic richness.

2.6 Sampling individuals

To measure metric disparity (MD), a total of 50 specimens by assemblage was typically used. However, for assemblages of low SR (2 to 8 species), the sample size could be less than 50 (28.6 +- 9.6). On average, the sample size was 45.8 +- 8.7. The abundance of each species within an assemblage was approximately the same. For instance, an assemblage of 10 species contained approximately 5 specimens per species, while an assemblage of 24 species could contain for instance 20 species with 2 individuals and 4 species containing 3 individuals. This situation of high evenness is generally not the one found in natural conditions.

2.7 Metric disparity

To estimate morphological diversity, we considered only the geometric, landmark-based approach. Morphological diversity was estimated by the metric disparity (MD) index. For

(*SR*)		(*2*)		(*4*)		(*6*)		(*8*)	(*etc.*)	(*40*)	(*44*)
1	13	21									
etc.	.	.									
30	3	14									
1	15	18	23	34							
etc.	.	.	.	.							
30	3	12	14	38							
1	7	17	22	35	37	41					
etc.	.	.	.	.	.	.					
30	3	5	6	10	26	38					
1	11	12	16	22	24	35	38	42			
etc.	.	.	.	.	.	.	.	.			
30	3	6	7	13	18	20	26	37			
1	6	9	18	19	22	28	34	38	etc.		
etc.	.	.	.	.	.	.	.	.	etc.		
30	3	6	9	13	15	24	27	38	etc.		
etc.	etc.	etc.	etc.	etc.	etc.	etc.	etc.	etc.	etc.		
1	1	2	5	6	8	11	12	13	etc.	41	
etc.	.	.	.	.	.	.	.	.	etc.	.	
30	1	3	4	6	7	8	10	11	etc.	44	
1-30	1	2	3	4	5	6	7	8	etc.	42	43 44

Table 3. A partial representation of a sequence of increasing taxonomic richness (numbers between brackets) as used by our simulation programme. Each number represents a mosquito species. There were 44 different taxa. The first set of rows shows two assemblages of two randomly selected species (species "13", species "21" and species "3", species "14") among the 30 species pairs randomly generated. The second set of rows shows two such assemblages of 4 species among the 30 random combinations of 4 species, and so on till reaching an assemblage of 44 species. This makes a total of 22 assemblages of increasing species richness, each one sampled 30 times among species. The last assemblage, containing the totality of the species, was sampled 30 times among individuals (*1-30*). Ten such sequences were generated, and each one was explored by using 254 different landmarks configurations (Table 2).

each of N wings, after Procrustes superposition using the Generalized Procrustes Algorithm (GPA) (Rohlf, 1990) and partial warps (PW) computation as in Rohlf (1996), the sum of squared PW was obtained. This sum was divided by the degrees of freedom (N-1) to compute MD (Zelditch et al., 2004). To estimate MD for an assemblage with, for instance, 2 species, 30 random pairs of species were used (see Table 3) and the average MD value was considered.

2.8 Habitat heterogeneity

The sample composition did not allow valid statistics using the habitat as a categorical unit (instead of the species). To evaluate the importance of the environment on the metric disparity (MD), we performed a simple two-way ANOVA with taxa and habitat as effects. The variable used was the individual sum of squared PW because this sum is the term directly used

to compute MD (see above). Five ANOVA were performed, one using the totality of LM to compute the PW, and four using selected combinations of LM. The latter were chosen according to their relationship to species diversity: they were two configurations of LM that produced MD highly correlated to the species richness (SR), and two others not related to SR (see Table 7).

2.9 Biodiversity

To estimate the biodiversity, we used the species richness index (SR, or the number of categorical units) which is the total number of species found in the community (an assemblage). Since evenness was maintained as high as possible, biodiversity indexes like the Shannon index (Shannon & Weaver, 1949) or the Simpson one (Simpson, 1949) were not examined.

2.10 Diversity structure

Diversity structure (DS) is described here as the relationship between MD and SR (Foote, 1992). It was estimated by the determination coefficients (squared linear correlation coefficients) (see Table 5), and illustrated graphically (Fig. 2). An average estimate of MD was used during all correlations. For each specific combination of LM, at each level of SR, MD was an average value derived from the 30 different assemblages (Table 3). For one specific number of LM, for example 4, this average was performed taking into account also 30 combinations of 4 LM (Table 2).

2.11 Software

A special TclTk (`http://www.tcl.tk/software/tcltk/`) script was written where Procrustes superposition (GPA), partial warps (PW) computations, as well as metric diversity (MD) estimations, made use of procedures extracted from the CLIC package (`http://www.mpl.ird.fr/morphometrics/clic/index.html`). Table 3 was computed using STATA (**?**). Figure 2 used the GNUMERIC spreadsheet (`http://projects.gnome.org/gnumeric/`).

3. Results

Globally, the effects of SR on various LM configurations of the same wing confirmed our initial hypothesis: different aspects of shape did not vary the same way in response to the same factor. The relationship between species richness (SR) and metric disparity (MD) was estimated by the determination coefficient (see DS, for "diversity structure" in Tables 4, 5 and 6). This coefficient does not inform about the sign of the correlation between SR and MD, which was not necessary since these correlation coefficients were all positive. The species richness contributed to the metric disparity according to the number of landmarks used to compute MD (Tables 4 and 5), and also according to the identity of landmarks involved (Table 6). We subdivide our results according to these two aspects of shape composition: the number (see paragraph 3.1) and the identity (see paragraph 3.2) of LM. We then show the ANOVA output for some remarquable configurations of LM (high DS, low DS): it estimates the possible role of habitat heterogeneity on the metric disparity (see paragraph 3.3).

3.1 Species richness and the number of landmarks

We observed that the configurations involving a low number of landmarks varied more frequently in very close accordance with species richness. Actually, for the same species arrangements, some configurations of the wing involving a very low number of landmarks (LNLM configurations) had either very high (column "DS > 75", see Table 4) or very low (column "DS < 50", see Table 4) prediction power on species diversity, while more complex anatomical configurations of the same wing showed a more stable but less predictive relationship with SR.

LM_nb	total	sampled	DS $\leq$ 55	55 < DS < 75	DS $\geq$ 75
4	715	(4%)	26%	45%	29%
5	1287	(2%)	21%	71%	8%
6	1716	(2%)	17%	71%	12%
7	1716	(2%)	12%	82%	9%
8	1287	(2%)	16%	77%	7%
9	715	(4%)	6%	87%	7%
10	286	(10%)	10%	83%	6%
11	78	(38%)	5%	86%	9%
12	13	(100%)	0%	98%	2%
13	1	(100%)	0%	100%	0%

Table 4. For each number of landmarks (LM_nb), the column "total" gives the total number of possible configurations among the possible landmark positions on the wing (a total of 13), the column "sampled" indicates the percentage of such configurations that have been studied; except for 12 and 13 LM, we always examined 30 random configurations for each number of landmark. Thus, 4% for instance is 30 out of 715. DS, or the "diversity structure" was estimated here by the determination coefficient (expressed in percentages) between metric disparity and species richness. The three last columns refers to the frequency at which a given DS was observed. It can be seen that a determination coefficient lower than 55% or higher than 75% between species richness and metric disparity was observed more frequently when using a low number of LM

Both the best and worst DS scores between SR and MD were obtained with configurations made from a low number of landmarks (LNLM). The range of scores progressively decreased with the addition of more LM (see column SD of Table 5). With more numerous LM however, the best predictive values did not reach such high levels as with fewer LM. Thus, more LM meant a more stable assessment of diversity structure (DS, or the relationship between MD and species richness), with no occurrence of very high values (Table 5).

3.2 Species richness and the identity of landmarks

Not only were we able to disclose different diversity structures according to the number of landmarks, but also according to specific configurations of landmarks. A partial output is presented Table 6. The highest determination coefficient (94%) observed was obtained with a specific combination of 4 LM (see 2-3-8-13, Table 4); other combinations involving the same number of LM gave much lower predictability (see 1-3-8-12, Table 4). It could be as low as 27% (see the columns MAX and MIN of Table 5).

NLM	DS	SD	MIN	MAX	STABILITY
4	64.7%	13.8	27	94	5.8
5	63.4%	7.7	35	85	6.0
6	65.4%	8.3	40	89	5.7
7	65.3%	5.8	41	86	6.4
8	64.1%	6.5	40	83	6.6
9	66.4%	3.6	49	83	6.1
10	64.5%	3.9	47	79	6.8
11	65.8%	3.4	48	80	6.3
12	67.5%	2.5	59	71	2.2
13	64.0%	0.6	57	62	3.5

Table 5. Determination coefficients (column DS) as percentages representing the contribution of species richness (SR) to metric disparity (MD) according to the number of landmarks (NLM) used to compute MD. These coefficients allow comparisons of the diversity structures (DS), i.e. the relationship between MD and SR according to NLM. Columns DS lists the determination coefficients and SD their standard deviation; MIN is the minimum value of DS and MAX its maximum. The table, except last column, represents an average DS derived from an average sequence of groups having increased SR; it was computed from the ten replicated sequences used in the study (see Table 3). The last column (STABILITY) is the standard deviation of the DS mean scores obtained from one sequence to another; it indicates how stable were the DS for the same LM configurations across 10 different random sequences of species. For each number of landmarks (each row), DS and SD values were averaged over 30 combinations of different landmarks (except for 12 LM which had only 13 possible configurations and of course for 13LM, see Table 2) and over ten replicated sequences, each one providing an average estimate from 30 random assemblages of species (Table 3)

For the same specimens and the same species arrangements, some shape components (i.e. landmarks configurations) of the wing varied in accordance with the number of species and others did not. Furthermore, a LM configuration highly predictive of SR, like the set of landmarks 2, 3, 8, 13 or the landmarks 2, 6, 7, 13, remained predictive regardless of the species sequences. The same observation applied for non-predictive sets of landmarks, like for instance the set of landmarks 1, 4, 5, 6 or 1, 3, 6, 9. This stability was verified across the ten replicates (see last column of Table 5).

3.3 The habitat heterogeneity

The highly predictive LNLM configurations, like 2-3-8-13 and 2-6-7-13 used in the ANOVA, were both affected by species richness only ($P < 0.0001$), not by the habitat heterogeneity ($P > 0.0500$), while the TNLM and two poorly predictive LNLM (1-4-5-6 and 1-3-6-9) were affected by both species and habitat (Table 7).

4. Discussion

To explore the diversity structure (DS), i.e. the relationship between metric and biological diversity, our model tested the effect of species richness (SR) on the metric disparity (MD) computed from 254 possible combinations of landmarks. Our data showed that the DS

	(4)	MIN	MAX	(11)	MIN	MAX
1	8-9-11-13	71	93	1-3-4-5-7-8-10-11-12-13	56	75
2	4-5-7-13	44	62	1-2-3-4-5-6-9-10-11-12-13	55	67
3	2-6-7-13	71	90	1-2-3-4-7-8-9-10-11-12-13	58	70
4	2-3-8-13	77	94	2-3-4-5-7-8-9-10-11-12-13	58	76
5	1-2-3-10	44	72	1-2-3-6-7-8-9-10-11-12-13	57	76
6	4-5-12-13	54	74	1-2-4-6-7-8-9-10-11-12-13	57	76
7	5-8-10-13	52	72	2-3-4-5-6-7-8-10-11-12-13	58	75
8	2-9-12-13	62	80	1-2-3-4-5-6-7-9-11-12-13	56	70
9	1-3-8-12	37	48	2-3-4-6-7-8-9-10-11-12-13	61	77
10	1-4-5-9	39	62	1-2-3-4-5-6-7-9-10-11-12	49	63
11	6-8-10-13	51	71	1-2-3-5-6-7-8-9-10-11-12	49	68
12	4-7-8-9	74	89	1-3-4-5-6-7-8-9-10-11-12	50	68
13	1-3-6-9	64	78	1-2-3-4-5-7-8-9-10-11-13	56	77
14	1-5-8-10	47	65	1-2-4-5-6-7-8-9-10-12-13	52	73
15	1-4-8-9	65	79	1-2-3-4-5-6-7-8-9-10-12	48	67
.	.	.	.	.	.	.
30	3-5-8-12	42	63	1-2-3-4-5-6-7-8-9-12-13	54	76

Table 6. Minimum (MIN) and maximum (MAX) determination coefficients as percentages representing the contribution of species richness to metric disparity according to the configurations of landmarks used. The 8-9-11-13 formula means the configuration of four landmarks using landmarks 8, 9, 11 and 13 as represented in Fig. 1. As in Table 2 the LM configurations are classified according to the number of LM (number between brackets). To save sapece, we present only a partial output of the data for 4LM and 11LM. For each configuration of landmarks (each row), values were obtained from ten replicated sequences, each one providing an average estimate from 30 random assemblages of species (Table 3). It can be seen that the contribution of species richness to metric disparity (MD) can be very high when MD is computed from a low number of LM (column 4LM), which does not seem to be the case for configurations involving more landmarks (column 11LM). Table 5 provides the average DS over all the tested configurations for each number of LM (not only 4 and 11).

depended on the aspects of shape that were considered (the number of LM, the configuration of LM).

4.1 Our model

The ability to detect morphological trends and occupation patterns within morphospace might depend on using the appropriate measure(s) of disparity. Since there is no clear indication about which index is best, some authors used a variety of measurements (Ciampaglio et al., 2001; Navarro, 2003). Our model presented results regarding only the MD index: it is the most commonly applied measurement of morphological disparity in landmark-based geometric studies (Zelditch et al., 2004).

In trying to include 50 individuals per group within our model, we were unable to consider the effect of sampling variation. However, metric disparity, as measured here, is relatively insensitive to sample size variation and has been shown to be a stable estimate when using a sample of 50 individuals (Navarro, 2003).

Source	Partial SS	df	MS	F	Prob > F
Model (13 LM)	0.036945	48	0.000770	14.710	0.000000
species richness	0.030978	43	0.000720	13.770	0.000000
habitat heterogeneity	0.005266	5	0.001053	20.130	0.000000
Residual	0.022545	431	0.000052		
Total	0.059491	479	0.000124		

Source	Partial SS	df	MS	F	Prob > F
Model (2-3-8-13)	0.001168	48	0.000770	10.240	0.000000
species richness	0.001126	43	0.000024	11.020	0.000000
habitat heterogeneity	0.000025	5	0.000005	2.140	0.059800
Residual	0.001025	431	0.000002		
Total	0.002193	479	0.000005		

Table 7. Two-way ANOVA showing the contribution of species richness and habitat heterogeneity to the individual sum of squared PW: this sum per individual is the value on which metric disparity is directly estimated (as an average). The ANOVA was performed on values computed from the total number of landmarks (TNLM configuration, top) or from the following configuration of four landmarks: 2-3-8-13 (bottom). For the latter (and also for LNLM highly predictive configurations like 2-6-7-13, not shown here), only species richness contributed to metric variation. Both species richness and habitat heterogeneity contributed to the variation of the total set of landmarks (as well as of LNLM poorly predictive configurations like 1-4-6-9 or 1-3-5-6, not shown here)

The model did not randomize the total number of species, which was always fixed to 44 taxa, nor did it randomize the sequence of species richness (2, 4, 6,..., 44), two parameters that are likely to affect the diversity structure of some landmark configurations.

There is however no *a priori* reason to think that these shortcomings would reduce the interest to use smaller configurations of shape to simplify the interpretation of the diversity structure.

More critical may be that our model did not take into account variation in species evenness. In each assemblage of species for a given sequence, the model was designed to get an approximately equal number of specimens within each species. In the natural conditions, there is generally no such evenness, and, because the MD index is an average value, this might affect the correlation between MD and SR. Rare taxa are often assumed to exhibit unusual morphologies because of specialized life habits and could thus contribute disproportionately to the disparity of an assemblage (Deline, 2009). But wings of mosquitoes, because they are generally very similar, are unlikely to have such an effect. Their general similarity may cause the opposite effect to take place: rare taxa may fill the mophospace among the common species and would likely lower MD (Deline, 2009).

Finally, in our simulation the successive assemblages contained an increasing number of species randomly selected from a total pool of mosquitoes, they were not obtained by " adding " species to previously selected ones. Because of this specific design, our simulation is not relevant to temporal follow-up of metric diversity, however it is specific to the spatial comparison of biodiversity.

4.2 Species diversity: The number of landmarks

Although correlations between MD and SR could be weak or not significant, they were positive. A positive correlation has an intuitive explanation: more diversity means more forms. Weak and not significant correlation between SR and MD obtained from some combinations of LM could be the consequence of a few species immediately occupying the extremes of the morphospace. In that situation, by filling in the morphospace between disparate taxa, increasing taxonomic diversity was not able to have a significant effect on MD estimates (Roy & Foote, 1997).

The intensity of the correlation between SR and MD depended on the number of LM. Unexpectedly, the low number of LM (LNLM) configurations gave the best predictive power on SR (Fig. 2, left side). This is counter-intuitive since it is the current belief that more shape would have more taxonomic contents and hence would be more useful to distinguish species. We verified that by using one of the most predictive configuration of 4 LM (2-3-8-13) only 30% of individuals could be correctly assigned to their species, while more than 75% could be correctly attributed with 13 LM (the TNLM configuration). How could it be that a better discriminating configuration of LM could give a lower taxonomic prediction than a poorly discriminating one? The answer could lie in the observation that a LNLM configuration also could have no relationship at all with SR, and that the TNLM configuration is gathering both the predictive and non-predictive subsets of LM. Thus, the mixing of better predictive with poorly predictive components of shape into a larger set of LM was likely to mix opposite trends and blur the relationship between MD and SR (Fig. 2, right side). In other words, when a larger number of LM is used as a proxy for SR, a larger number of influences is also allowed which is not limited to SR. This explanation raises the question whether in our material there were other influences affecting the variation of the LM configurations that were not related to SR.

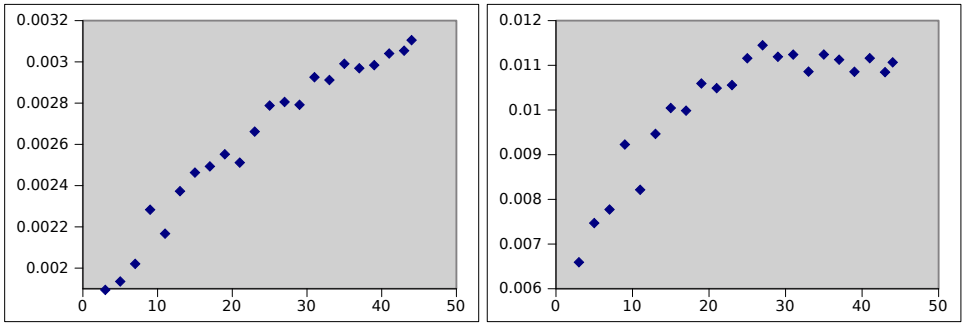


Fig. 2. Two graphs representing the diversity structure (DS), i.e. the relationship between species richness (SR, on the horizontal axis) and metric diversity (MD, on vertical axis). The left graph shows the DS appearing when using a set of four landmarks (namely the landmarks 2, 3, 8 and 13; see Fig. 1). The right graph shows the DS using the total number (13) of LM. When using the totality of LM (right graph) there is a first list of increasing MD values up to a SR of 30 species. This first list is common to the two graphs. In the right one however, after 30 species, there is a " plateau " breaking the global correlation between SR and DM

4.3 Species diversity: The identity of landmarks

For a given number of landmarks, the same configurations produced high (or low) correlation with SR regardless of the taxonomic composition of the species sequence. For instance, the 2-3-8-13 configuration always produced high correlation with SR regardless of the sequence of species considered, and the 1-3-6-9 configuration always produced one of the worst correlations. Thus, there seemed to be specific configurations of LM that could reflect the number of species, and others that could not, regardless of the species sequence used. This observation raises two questions:

- Are there any LM configurations that could be used in mosquitoes as a proxy for species richness estimations?
- What are the sources (other than randomness) that generated variation for the LM configurations that were not influenced by the species richness?

The first question is about the effect of taxonomic diversity on wing shape variation: can the latter be used as a proxy for indirect species richness estimation?

Before starting to answer that question, one could ask what the interests are in making an indirect estimation of species richness. The answer is about time. Metric disparity extracted from wing venation is much faster to obtain than metrics requiring a taxonomic identification of each collected individual. Mosquitoes are often distinguished on the basis of labile characters which may be lost on damaged specimens, making species diagnostic difficult in a group where species are numerous (3500 species) and taxonomists are unfortunately not. An indirect but fast comparative estimation of SR would be welcome.

Our results showed that some landmarks configurations were highly predictive for SR. If these landmark configurations were known in advance, their variation could be used indeed as a proxy for SR. In our study, their performances seemed stable regardless of species arrangement in different groups (Table 5, last column). However, whether the high scoring of these LM configurations would be maintained in other studies on other mosquito species is still to be investigated.

The second question refers to other possible meanings of MD in our data set. Other than SR, what could be the cause of shape variation? It probably has many causes, certainly among which is species evenness (Deline, 2009), but also functional and ecological attributes (Roy & Foote, 1997), environmental conditions (Vasil'ev et al., 2010), founder effects (David, 1999; Whitlock & Fowler, 1999), endemicity (Magniez-Jannin et al., 2000) or reproduction mode (Baltanas et al., 2002). The only factor we could discuss with our data was the habitat, an environmental parameter to which insect metric properties are known to be sensible (Benítez et al., 2011; Tantowijoyo & Hoffmann, 2011). Insects were collected in the forest, in rural and urban areas. It was difficult to address the question of habitat influence for a given species since the within species sample sizes did not allow valid statistics inference (see Table 1). Thus, the habitat as explored here was not free from the possible effects due to different species compositions. Our ANOVA analyzes could however provide some indications. When both species richness and habitat heterogeneity were significantly contributing to the MD, a good relationship between MD and SR was not observed. A broad morphological covering (TNLM), or some unpredictive elementary units of morphology (LNLM configurations like 1-4-6-9 or 1-3-5-6), were under the influence of both species and habitat heterogeneities (Table

7). On the contrary, the highly predictive LNLM configurations (like 2-3-8-13 or 2-6-7-13) were apparently under the influence of SR only (Table 7).

5. Perpectives

Almost two decades ago, Foote (1993) claimed that " *discordances between morphological and taxonomic diversity demand to be interpreted biologically, not explained away as artifact of taxonomic practice* ". Morphological disparity is obviously under the influence of factors other than the mere species number, and we showed here for instance the likely influence of the mosquito habitat. However, this does not preclude the possibility to use shape variation as an indicator of species richness. We considered this possibility through decomposing wing shape into elementary components and comparing their respective relationships with taxonomic richness.

We suggest that the use of elementary units of shape (LNLM) could allow one to focus on a single factor, like species richness, and that in this regard the use of many characters (TNLM) has the inconvenience of mixing various effects, making a clear interpretation difficult. In a recent study showing good parallelism between SR and metric diversity, a single character was used (Dolan et al., 2006).

If the objective of the morphometric analysis was to accurately reflect one factor, then the use of LNLM is recommended, although not any LNLM configuration. The remaining question is: which LNLM configuration to use?

The answer to such a question certainly implies to explore the relationship of the LNLM configurations with known factors other than possible species richness, like the habitats, the localities of collection, etc. Any combination of landmarks which would vary under the influence of such parameters would be less likely to reliably reflect species richness alone. However, a more definitive answer cannot be provided through the use of a model which did not reflect natural conditions closely enough. As explained above, an investigation is still needed to evaluate the interference of species evenness on the relationship between MD and SR (Deline, 2009), an issue not contemplated in our model.

6. Conclusion

At this stage, our data confirmed that metric properties of a given community contain hidden but accurate information about species richness. Our model suggests that this information is likely to be found through the examination of some elementary shape configurations rather than of a global multivariate projection of many morphological traits.

7. Acknowledgment

We thank R. Rattanarithikul for her kind help in taxonomic identification, as well as P. Kittayapong and C. Apiwhatnasorn (Mahidol University, Bangkok, Thailand) who encouraged this work. This study has been supported by the TRF/BIOTEC Special Program for Biodiversity Research and Training Grant BRT R_352052 and the Commission on Higher Education, RMU5080060 (Thailand), as well as by the IRD grants number HC3165-3R165-GABI-ENT2 and HC3165-3R165-NV00-THA1.

8. References

Baltanas, A., Alcorlo, P. & Danielopol, D. L. (2002). Morphological disparity in populations with and without sexual reproduction: a case study in *Eucypris virens* (Crustacea : Ostracoda), *Biological Journal of the Linnean Society 75(1): 9–19*.

Benítez, H. A., Briones, R. & Jerez, V. (2011). Intra and inter-population morphological variation of shape and size of the chilean magnificent beetle, *Ceroglossus chilensis* in the Baker River basin, Chilean Patagonia, *Journal of Insect Science, 11(94): 1–9*.

Ciampaglio, C. N., Kemp, M. & McShea, D. W. (2001). Detecting changes in morphospace occupation patterns in the fossil record: characterization and analysis of measures of disparity, *Paleobiology* 27(4): 695–715.

David, P. (1999). A quantitative model of the relationship between phenotypic variance and heterozygosity at marker loci under partial selfing, *Genetics, 153: 1463–1474*.

Deline, B. (2009). The effects of rarity and abundance distributions on measurements of local morphological disparity, *Paleobiology* 35(2): 175–189.

Dolan, J. R., Jacquet, S. & Torréton, J.-P. (2006). Comparing taxonomic and morphological biodiversity of Tintinnids (planktonic ciliates) of New Caledonia, *Limnology and Oceanography, 51 (2)* pp. 950–958.

Foote, M. (1992). Rarefaction Analysis Of Morphological And Taxonomic Diversity, *Paleobiology* 18(1): 1–16.

Foote, M. (1993). Discordance and Concordance Between Morphological and Taxonomic Diversity, *Paleobiology* 19(2): 185–204.

Gaston, K. J. & Spicer, J. I. (2004). Biodiversity: An introduction, *Oxford: Blackwell Publishing*.

Magniez-Jannin, F., David, B. & Dommergues, J. L. (2000). Analysing disparity by applying combined morphological and molecular approaches to French and Japanese carabid beetles, *Biological Journal of the Linnean Society 71(2): 343–358*.

Moyne, S. & Neige, P. (2007). The space-time relationship of taxonomic diversity and morphological disparity in the Middle Jurassic Ammonite radiation, *Palaeogeogr. Palaeoclimatol. Palaeoecol., 248: 82–95*.

Navarro, N. (2003). MDA: a MATLAB-based program for morphospace-disparity analysis, *Computers & Geosciences, 29: 655–664*.

Rattanarithikul, R., Harrison, B. A., Panthusiri, P. & Coleman, R. E. (2005). Illustrated keys to the mosquitoes of Thailand I. Background; geographic distribution; lists of genera, subgenera, and species; and a key to the genera., *Southeast Asian J Trop Med Public Health, 36(Suppl 1): 1–80*.

Rattanarithikul, R., Harrison, B. A., Panthusiri, P., Peyton, E. L. & Coleman, R. E. (2006). Illustrated keys to the mosquitoes of Thailand III. Genera *Aedeomyia*, *Ficalbia*, *Mimomyia*, *Hodgesia*, *Coquillettidia*, *Mansonia*, and *Uranotaenia*., *Southeast Asian J Trop Med Public Health, 37(Suppl 1): 1–85*.

Rattanarithikul, R., Harrison, B., Harbach, R., Panthusiri, P. & Coleman, R. (2006.). Illustrated keys to the mosquitoes of Thailand IV. *Anopheles*., *Southeast Asian J Trop Med Public Health 37(Suppl.2): 1–128*.

Rattanarithikul, R., Harbach, R. E., Harrison, B. A., Panthurisi, P., Coleman, R. E. & Richardson, J. (2010). Illustrated keys to the mosquitoes of Thailand. VI. Tribe Aedini, *Southeast Asian J Trop Med Public health, Vol. 41 (Suppl.1): pp. 225*.

Rohlf, F. J. (1990). Rotational fit (Procrustes) methods, *in* F. Rohlf & F. Bookstein (eds), *Proceedings of the Michigan Morphometrics Workshop. Special Publiation Number 2.*

The University of Michigan Museum of Zoology. Ann Arbor, MI, pp380, University of Michigan Museums, Ann Arbor, pp. 227–236.

Rohlf, F. J. (1996). Morphometric spaces, shape components and the effects of linear transformations, *in* L. F. Marcus, M. Corti, A. Loy, G. Naylor & D. Slice (eds), *Advances in Morphometrics. Proceedings of the 1993 NATO-ASI on Morphometrics*, New York: Plenum Publ. NATO ASI, ser. A, Life Sciences, pp. 117–129.

Roy, K., Balch, D. P. & Hellberg, M. E. (2001). Spatial patterns of morphological diversity across the Indo-Pacific: analyses using strombid gastropods, *Proc. R. Soc. Lond., 268: 2503–2508*.

Roy, K. & Foote, M. (1997). Morphological approaches to measuring biodiversity, *Trends In Ecology & Evolution* 12(7): 277–281.

Shannon, C. E. & Weaver, W. (1949). The mathematical theory of communication, *University of Illinois Press, Urbana*.

Simpson, E. H. (1949). Measurement of diversity, *Nature, 163: 688*.

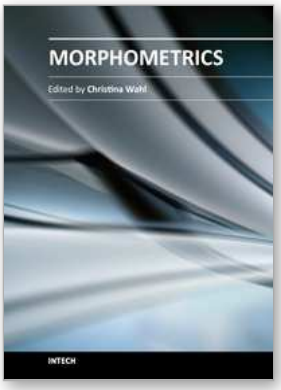
Tantowijoyo, W. & Hoffmann, A. A. (2011). Variation in morphological characters of two invasive leafminers, *Liriomyza huidobrensis* and *L. sativae*, across a tropical elevation gradient, *Journal of Insect Science, 11(69): 1–16*.

Vasil'ev, A. G., Vasil'eva, I., Gorodilova, Y. V. & Chibiryak, M. V. (2010). Morphological disparity in populations with and without sexual reproduction: a case study in *Eucypris virens* (Crustacea : Ostracoda), *Russian Journal of Ecology, 41(2): 153–158*.

Whitlock, M. C. & Fowler, K. (1999). The distribution of phenotypic variance with inbreeding, *Evolution, 53(4): 1143–1156*.

Wills, M. A., Briggs, D. E. G. & Fortey, R. A. (1994). Disparity As An Evolutionary Index - A Comparison Of Cambrian And Recent Arthropods, *Paleobiology* 20(2): 93–130.

Zelditch, M. L., Swiderski, D. L., Sheets, H. D. & Fink, W. L. (2004). *Geometric morphometrics for biologists: a primer*, Elsevier, Academic Press. New-York.



Morphometrics

Edited by Prof. Christina Wahl

ISBN 978-953-51-0172-7
Hard cover, 108 pages
Publisher InTech
Published online 02, March, 2012
Published in print edition March, 2012

It is human nature to measure things, and this holds true for science as well as everyday life. The five papers in this book demonstrate the usefulness of a morphometric approach to a variety of subjects in natural history, including systematics, phenotypic plasticity in response to environmental variation, and ontogenetic adaptation. As our understanding of genetic control mechanisms and epigenetics has matured over the last several decades, it has become clear that morphometric assessment continues to be important to our overall understanding of natural variability in growth and form. The tremendous growth of our knowledge base during the last century has necessitated that we find new ways to measure and track greater detail as well as greater numbers of parameters among populations and individuals.

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821