Journal of Biogeography, **30**, 193–209

Plant collecting spread and densities: their potential impact on biogeographical studies in Thailand

J. A. N. Parnell[1]*, D. A. Simpson[2], J. Moat[2], D. W. Kirkup[2], P. Chantaranothai[3], P. C. Boyce[2], P. Bygrave[2], S. Dransfield[2], M. H. P. Jebb[4], J. Macklin[1], C. Meade[5], D. J. Middleton[6], A. M. Muasya[7], A. Prajaksood[1,3], C. A. Pendry[8], R. Pooma[9], S. Suddee[9] and P. Wilkin[2] [1]*Herbarium, School of Botany, Trinity College Dublin, Dublin, Ireland,* [2]*Royal Botanic Gardens Kew, Richmond, Surrey, UK,* [3]*Department of Biology, Faculty of Science, Khon Kaen University, Khon Kaen, Thailand,* [4]*National Botanic Gardens, Glasnevin, Dublin, Ireland,* [5]*Institute of Bioengineering and Agroecology, National University of Ireland, Maynooth, Co. Kildare, Ireland,* [6]*Harvard University Herbaria, Divinity Avenue, Cambridge, MA, USA,* [7]*National Museums of Kenya, Nairobi, Kenya,* [8]*Royal Botanic Gardens Edinburgh, Inverleith Row, Edinburgh, UK and* [9]*Royal Forest Herbarium, Royal Forest Department, Bangkok, Thailand*

Abstract

Aims To produce representative aggregate maps of plant collection locations in Thailand and discuss their impact on biogeographical studies in Thailand and the surrounding region.

Location Thailand.

Methods A representative data set comprising 6593 plant specimen records for Thailand has been assembled. The data set contains ± all known collections for fifteen representative plant families and further records for another 104. All records are localized to Changwat (province), 6441 to at least quarter degree square.

Results Analysis shows that the spread of collecting activity in Thailand is markedly uneven; 20% of collections come from a single Changwat (Chiang Mai) and 53% of Changwat have fifty or fewer collections. The distribution of collections by Changwat and by quarter degree square is erratic with most squares and Changwat having few collections, both in proportionate and absolute terms. Some of the most densely forested Changwats and squares appear undercollected. Distribution maps for common, easily recognized tree species in the genus *Syzygium* show distributional gaps.

Conclusions Thailand is defined as an undercollected country. Even within the few well-collected quarter degree squares the spread of collecting is still poor; almost all collections being localized to one of three mountain ranges or their foothills. There are many gaps in collecting activity which make impossible a straightforward interpretation of biogeographical pattern. It is argued that targeted collecting activity is needed, that assembly of this type of data set is therefore essential and that our data set and its interpretation is a model for all countries in the region.

Keywords

Thailand, biogeography, biodiversity, collection density, conservation, plants, GIS, flora of Thailand, collection, Southeast Asia.

*Correspondence: Herbarium, School of Botany, Trinity College Dublin, Dublin 2, Ireland. E-mail: jparnell@tcd.ie

INTRODUCTION

Thailand is bordered on the west by Myanmar (Burma), on the east by Lao PDR and Cambodia, and in the south by peninsular Malaysia and the Andaman and China seas. It covers an area of *c.* 515,000 km^2 and is therefore only slightly smaller than the largest country in the EU (France) and is 20% larger than the third largest American state (California). Thailand has a diverse array of habitats but was originally more or less completely covered in forest. The forests were and are of two broad types: evergreen (including Malayan mixed Dipterocarp, wet seasonal evergreen, dry evergreen, montane, limestone, peat swamp and mangrove) and deciduous (mixed and dry Dipterocarp) (Santisuk *et al.*, 1991). Clearance for agriculture and other uses has reduced forest cover, to perhaps as low as 20% (Santisuk *et al.*, 1991), much of which may be degraded. Thailand has a wide altitudinal range (the highest point, Doi Inthanon is 2600 m) and is the geographical centre of a biogeographical realm stretching from eastern India to Vietnam and south China (Santisuk *et al.*, 1991). The serious deterioration of Thailand's natural and environmental resource has been identified by the Royal Thai Government as a constraint on continued development and is obviously of importance in terms of maintenance of biodiversity (Santisuk *et al.*, 1991).

The plant diversity of Thailand is being documented by a major long-term project which was initiated in 1957–58 (Larsen & Warncke, 1966; Smitinand & Larsen, 1966; Larsen, 1979, 1988). It has resulted in accounts of *c.* 30–40% of the angiosperm flora in Thailand, with estimates of the size of the flora varying from *c.* 10,000–12,500 species (Santisuk *et al.*, 1991; Parnell, 2000). Despite fairly rapid progress the project will not be complete for many years (estimates ranging from another 100 years (Santisuk *et al.*, 1991) to, a perhaps overoptimistic, 30 years (Parnell, 2000).

There are a number of problems constraining biogeographical studies of the flora of Thailand and Southeast Asia. Of particular relevance to this paper are issues relating to plant collection rates and densities.

A recently published estimate for collecting density in Thailand is *c.* 0.5 specimens per km^2 (Parnell, 2000), a value at least twice that estimated for much of Malesia (0.25 specimens per km^2) but half that of peninsular Malaysia with one specimen per km^2 (Johns, 1995). In addition, Campbell (1988) suggested a minimum minimal botanical collection density of one specimen per km^2 for specimens housed within the borders of a country, regardless of those deposited abroad. Holmgren *et al.* (1990) indicate that there are only 157,000 specimens held in Thailand; immediately neighbouring countries have fewer. Therefore, even allowing for a 20% increase in plant holdings since 1990, Thailand and all immediately surrounding countries are still well below the minimum. Toledo & Sosa (1993) showed that, based on Campbell's criterion, 40% of counties in Latin America and the Caribbean have adequate collections. Therefore, it appears that Thailand and all immediately surrounding countries are in a poorer state in terms of collection density than many neotropical countries. Such deficiencies are particularly worrying when indices of current collecting activity are calculated.

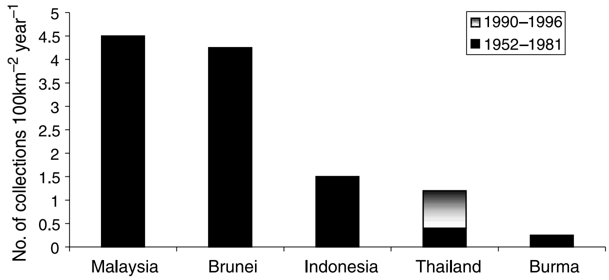

Figure 1 Number of collections per 100 km^2 per year for various Southeast Asian countries for the period 1952–81; data from Prance & Campbell (1988). Data for the period 1990–96 are available only for Thailand.

Data from Prance & Campbell (1988) show that the rate of collecting in Thailand was by far the lowest of the countries sampled in Southeast Asia and Malesia (at 0.4 specimens per 100 km^2 per year) for the period 1952–81 (Fig. 1). Prance & Campbell (1988) also show that in most of Southeast Asia (including Thailand) and Malesia, that there was a significant decline in the collecting rate between 1974 and 1981 compared with 1952–81. Indeed, Prance & Campbell (1988) note that the collecting rate for Thailand between 1974 and 1981 effectively fell to zero, a position previously indicated by Larsen (1979). However, that situation has since been reversed with collecting activity between 1990 and 1998 rising to 1.2 collections per 100 km^2 per year (Parnell, 2000), a rate comparable to that for Indonesia over the period 1952–81 (Fig. 1).

For many countries, especially in the tropics, there are no detailed, reliable data to indicate the spread or density of collecting within that country. However, two notable exceptions are recent analyses of large collection-based data bases assembled for Amazonia and Guyana (Kress *et al.*, 1998; Funk *et al.*, 1999; Funk & Richardson, 2002). Amongst other things, these authors have used their assembled data sets to identify areas which are undercollected, of high species diversity and those with high endemism. They also link their analyses to conservation studies. This paper presents the results of analyses we have undertaken on plant collecting spread and density in Thailand, using a large data set derived from herbarium collections representing over 100 plant families. From these we are able to identify areas of both high and low collecting activity, providing a model applicable to study of the flora of Southeast Asia in general terms. We discuss the implications of current spread and density of collecting on associated biogeographical studies. We exemplify this in more detail with the distribution patterns of two species of *Syzygium* (Myrtaceae) in Thailand. Myrtaceae are one of the larger families of flowering plants in Thailand with twelve genera and *c.* 115 species (Chantaranothai, 1989; Chantaranothai & Parnell, 1994). *Syzygium* is the largest genus in that family with eighty-four species in Thailand (Parnell & Chantaranothai, 2002). All *Syzygium* species are trees or shrubs and are relatively conspicuous; many are also common and widespread. Therefore, a study of their distribution patterns is likely to generate models applicable to other taxa.

METHODS

The data set comprises specimen records accumulated by a large number of workers on the Thai flora. In general, these workers have been responsible for accounts of particular families for the Flora of Thailand project. Other data in the data base largely result from recent collecting expeditions in Thailand with identifications made by acknowledged experts (Simpson *et al.*, 1995; Parnell *et al.*, unpubl. data). Therefore, the species identifications in the data base data set were produced by systematic specialists actively working on the families concerned. Thus, these data represent specimens from a broad range of habitat type, have been obtained from many herbaria, represent collecting activity over a *c.* 100-year time period and are correctly identified. They are not based on the work of any single institution or individual. Neither are the families at the limits of their distributions being confined to one or other side of the Isthmus of Kra (van Steenis, 1950); therefore, there are unlikely to be artefacts associated with this data set which would give false indications of collecting density due to the fact that the plant material concerned is at or near a well-known biogeographical limit. The data base includes all known collections, or the vast majority of them, for the following plant families: Annonaceae (*Uvaria* group), Apocynaceae, Araceae, Araliaceae, Cyperaceae, Dioscoreaceae, Eriocaulaceae, Lamiaceae, Lecythidaceae, Myrtaceae, Orobanchaceae, Poaceae (Bambusoideae), Plantaginaceae, Polygalaceae and Santalaceae. In addition, a number of records for Euphorbiaceae are included as well as scattered records for 104 other families. Nearly all the records have been localized to varying degrees of resolution: 6593 records to one of the seventy-six Changwat (Provinces) (Table 1, Fig. 2) and 6441 to at least quarter degree (Fig. 3). In summary, the data base is broadly based and are very likely to be representative of plant collecting activity in Thailand as a whole.

Data were assembled into an Excel spreadsheet and transferred to Access prior to mapping and data analysis. For map construction the collection locality latitudes and longitudes were first transformed from the recorded degree and minute format into decimal degrees for display within the Geographical Information Systems (GIS). The data base containing the reformatted co-ordinates was then imported into ESRI Arcview GIS as a table and this table was used to create an event theme (also known as a coverage or a layer) which plotted the individual collections as separate points. The following analyses were undertaken.

Collection density per Changwat

Up-to-date coverage for the Changwat was created within the GIS by manipulation of existing coverage for Thailand (ESRI ArcWorld 1 : 3 m). Four new Changwat, established in 1996 and not included in the existing coverage, were produced by splitting some of the the older Changwat (e.g. Ubon Ratchathani which was split in the northern third to form Ubon Ratchathani and Amnat Charoen). This data set was joined by Changwat name to the table of collections. The collection density was calculated by dividing the number of collections by the area of each Changwat. Data were displayed as blue to red dichromatics, using standard deviation (SD) from the mean (where blue = negative SD suggesting undercollection and red = positive SD, suggesting many collections).

Collection density per quarter degree square

The analyses of the number of collections by grid square were performed by making spatial joints between the collection (point) theme and the appropriate polygon theme (grid). The resulting tables, summarizing the number of points per polygon, were then joined to the polygon theme to enable graphical display of the collecting densities. The percentage area of land was calculated for each grid cell by combining the country boundary with the grid cells. The collection density was calculated for each grid by dividing the collections by the percentage land area. The results were again displayed as blue to red dichromatics, using the SD from mean (where blue = negative SD, suggesting undercollection and red = positive SD, suggesting many collections).

Changwat, vegetation and percentage of forest cover in each Changwat

Vegetation, as defined by the Global Land Cover Characteristics (GLCC) data set (http://edcdaac.usgs.gov/glcc/glcc.html), was mapped using the simple International Geosphere Biosphere Programme classification (IGBP) (Belward *et al.*, 1999). This map shows vegetation cover in Thailand as of the mid-1990s. The map was then reclassified into forest and non-forest (forest was defined as: evergreen needleleaf forest, evergreen broadleaf forest, deciduous needleleaf forest, deciduous broadleaf forest and mixed forest). This coverage was overlaid with the Changwat coverage to give the percentage of each Changwat that is forested.

Collection density by forest area

A map was developed in a similar manner to collection density per Changwat, except that collection density was calculated by dividing the number of collections within each Changwat per 100 km^2 by the forested area of each Changwat.

Number of collections per vegetation class

Sampling was carried out of collection localities in relation to major vegetation classes defined in the GLCC vegetation

Table 1 Area and collection data for each Changwat

Changwat No.	Changwat name	Changwat area (km^2)	No. of collections per Changwat	Collections per Changwat (%)	No. of collections per 100 km^2 per Changwat	Forest cover per Changwat (%)	No. of collections within forest per 100 km^2 per Changwat
1	Mae Hong Son	13,396	105	1.59	0.78	94	0.83
2	Chiang Mai	22,712	1312	19.90	5.78	75	7.74
3	Chiang Rai	11,413	66	1.00	0.58	49	1.17
4	Phayao	6113	14	0.21	0.23	35	0.65
5	Nan	12,222	106	1.61	0.87	78	1.11
6	Lamphun	4079	33	0.50	0.81	39	2.05
7	Lampang	12,559	65	0.99	0.52	50	1.04
8	Phrae	6619	31	0.47	0.47	64	0.73
9	Uttaradit	7830	71	1.08	0.91	43	2.11
10	Tak	16,007	116	1.76	0.72	69	1.05
11	Sukothai	7098	11	0.17	0.15	22	0.70
12	Phitsanulok	10,703	150	2.28	1.40	21	6.83
13	Kamphaeng Phet	8771	27	0.41	0.31	14	2.14
14	Phichit	4259	1	0.02	0.02	0	0.00
15	Nakhon Sawan	9786	33	0.50	0.34	4	9.26
16	Phetchabun	12,466	77	1.17	0.62	15	4.16
17	Loei	10,450	386	5.85	3.69	15	24.93
18	Nong Bua Lum Phu	4445	16	0.24	0.36	4	8.48
19	Udon Thani	10,715	43	0.65	0.40	1	51.94
20	Nong Khai	7424	92	1.40	1.24	3	37.23
21	Sakon Nakhon	9589	142	2.15	1.48	6	24.08
22	Nakhon Phanom	4892	30	0.46	0.61	5	11.54
23	Mukdahan	4928	61	0.93	1.24	16	7.55
24	Kalasin	6877	14	0.21	0.20	2	9.84
25	Maha Sarakham	5749	24	0.36	0.42	0	0.00
26	Khon Kaen	10,732	105	1.59	0.98	1	80.32
27	Chaiyaphum	13,234	143	2.17	1.08	19	5.67
28	Nakhon Ratchasima	20,808	151	2.29	0.73	9	7.76
29	Buri Ram	10,212	15	0.23	0.15	4	4.18
30	Surin	8647	26	0.39	0.30	3	10.39
31	Roi Et	7756	18	0.27	0.23	1	30.26
32	Yasothon	4102	2	0.03	0.05	2	2.80
33	Amnat Charoen	5045	10	0.15	0.20	7	3.00
34	Si Sa Ket	8985	30	0.46	0.33	7	5.01
35	Ubon Ratchathani	13,735	58	0.88	0.42	7	5.92
36	Uthai Thani	6575	14	0.21	0.21	35	0.61
37	Kanchanaburi	19,418	241	3.66	1.24	40	3.11
38	Ratchaburi	5279	36	0.55	0.68	7	9.38
39	Phetchaburi	6194	32	0.49	0.52	29	1.80
40	Prachuap Khiri Khan	6271	76	1.15	1.21	46	2.62
41	Chai Nat	2497	5	0.08	0.20	0	42.04
42	Sing Buri	848	0	0.00	0.00	0	0.00
43	Lop Buri	6508	11	0.17	0.17	0	92.12
44	Suphan Buri	5426	1	0.02	0.02	11	0.16
45	Ang Thong	979	5	0.08	0.51	0	0.00
46	Ayutthaya	2498	4	0.06	0.16	12	1.39
47	Saraburi	3033	69	1.05	2.28	5	47.97
48	Nakhon Nayok	2198	2	0.03	0.09	30	0.31
49	Nakhon Pathom	2113	3	0.05	0.14	20	0.70
50	Pathum Thani	1485	21	0.32	1.41	2	58.45
51	Nonthaburi	599	5	0.08	0.83	15	5.72
52	Krung Thep	1565	84	1.27	5.37	1	697.85
53	Samut Prakan	938	4	0.06	0.43	0	0.00
54	Samut Songkhram	413	0	0.00	0.00	21	0.00
55	Samut Sakhon	868	0	0.00	0.00	0	0.00
56	Sa Kaeo	7458	32	0.49	0.43	14	3.13
57	Prachin Buri	4570	52	0.79	1.14	25	4.64

Table 1 *continued*

Changwat No.	Changwat name	Changwat area (km^2)	No. of collections per Changwat	Collections per Changwat (%)	No. of collections per 100 km^2 per Changwat	Forest cover per Changwat (%)	No. of collections within forest per 100 km^2 per Changwat
58	Chachoengsao	5588	3	0.05	0.05	8	0.64
59	Chon Buri	4412	151	2.29	3.42	6	56.94
60	Rayong	3690	37	0.56	1.00	34	2.92
61	Chanthaburi	6200	181	2.75	2.92	43	6.84
62	Trat	2915	181	2.75	6.21	38	16.37
63	Chumphon	5656	103	1.56	1.82	59	3.11
64	Ranong	3404	139	2.11	4.08	58	6.98
65	Surat Thani	13,129	215	3.26	1.64	71	2.32
66	Phangnga	4094	131	1.99	3.20	76	4.19
67	Phuket	566	64	0.97	11.31	70	16.26
68	Krabi	4505	55	0.83	1.22	77	1.58
69	Nakhon Si Thammarat	10,578	177	2.68	1.67	39	4.26
70	Phatthalung	3602	33	0.50	0.92	30	3.09
71	Trang	4775	173	2.62	3.62	58	6.28
72	Satun	2669	166	2.52	6.22	51	12.27
73	Songkhla	8147	180	2.73	2.21	43	5.18
74	Pattani	1986	52	0.79	2.62	38	6.87
75	Yala	4535	67	1.02	1.48	67	2.21
76	Narathiwat	4389	204	3.09	4.65	78	5.99
		516,930	6593				

map. A histogram was produced which showed (1) actual collections for a vegetation class based on the collecting locality data and (2) expected collections for a given vegetation class. The latter was calculated by normalizing from the pixel count of the vegetation map to the number of collections, using the formula: number of pixels per vegetation class × (total collections/total number of pixels).

Number of collections relative to distance from populated places and roads

Data for both roads and populated places in Thailand were obtained from a digital chart of the world (DCW) (ESRI, 1993). Distance image maps of Thailand were produced from these data using ESRI ArcView Spatial Analysis (ESRI, 1996). From the maps, distance classes were identified with the distances quoted in kilometres. Collection localities were also overlaid onto the maps. Histograms were produced using the collection data and the distance classes, one each for the roads and populated places. Each histogram showed (1) the actual number of collections for a distance class and (2) the expected number of collections for that distance class. The latter was calculated by using a similar formula to that for vegetation class.

Maps for exemplar species

Finally, we selected *Syzygium cinereum* (Kurz) P. Chantaranothai & J.Parn and *S. cumini* (L.) Skeels as exemplar species for which we mapped individual distributions using the methods described above. Both are relatively unspecialized species which occur in Thailand in a variety of forest types up to 1000 m altitude.

RESULTS

Collection density per Changwat

Table 1 and Fig. 2 show the spread of plant collecting activity across Thailand. Three Changwat have no collections for any of the families surveyed and a further eleven Changwat have ten or fewer collections (Table 1). Fifty-three per cent of the Changwat have fewer than 50 collections each and 72% less than 100 collections. Moreover, the distribution of collections by Changwat is erratic (Table 1) as 19.9% come from Chiang Mai (Changwat No. 2), 5.9% from Loei (Changwat No. 17) and 3.7% from Kanchanaburi (Changwat No. 37). Only one other Changwat [Narathiwat (Changwat No. 76)] accounts for at least 3% of the total collections in the data base. In terms of collections per unit area the most densely collected Changwat are, in descending order:

Phuket (Changwat No. 67) @ 11.3 collections per 100 km^2
Trat (Changwat No. 62) @ 6.2 collections per 100 km^2
Satun (Changwat No. 72) @ 6.2 collections per 100 km^2
Chiang Mai (Changwat No. 2) @ 5.8 collections per 100 km^2
Krung Thep (Changwat No. 52) @ 5.4 collections per 100 km^2
Ranong (Changwat No. 64) @ 4.1 collections per 100 km^2
Loei (Changwat No. 17) @ 3.7 collections per 100 km^2 and
Narathiwat (Changwat No. 76) @ 3.1 collections per 100 km^2.

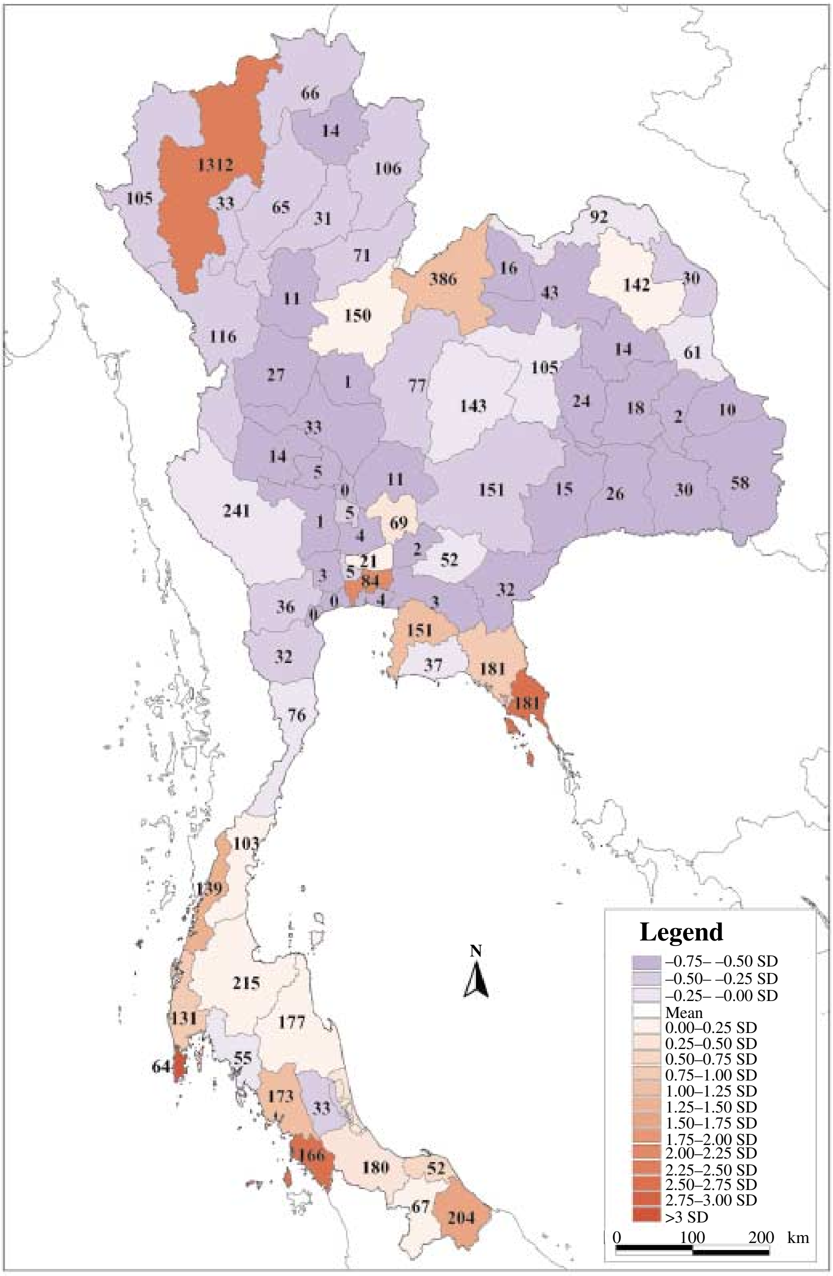

Figure 2 Map of Thailand divided into Changwat, showing number of collections (bold figures) and collection density (blue to red dichromatic shading) per Changwat. Mean = 1.394 collections per 100 km^2, with 1 SD = 1.888 (i.e. +1 SD = 3.282 collections per 100 km^2).

Changwat area is strongly positively correlated with the number of collections (Pearson product moment correlation; $r = 0.531$; d.f. = 74; $P = 0.001$) for all georeferenced collections and therefore, not unexpectedly, the larger the Changwat the larger the number of collections that have been made therein. However, there are notable exceptions as the second largest Changwat in terms of area (Nakon Ratchasima, Changwat No. 28) is fourteenth in terms of the number of collections made per Changwat and therefore thirtyninth in terms of collection density per 100 km^2 (Table 1).

Collection density per quarter degree square

Figure 3 shows both the spread of collecting activity and collecting density by quarter degree squares. The different shading indicates different collecting densities; these are calculated as standard deviations above or below the theoretical mean number collections for each quarter degree square. Forty-seven (19.2%) of the 245 squares have SDs above the mean, of which eight (3.3%) are two or more above. Of the remaining squares, 181 (73.8%) have collecting densities below the mean, although only five (2%)

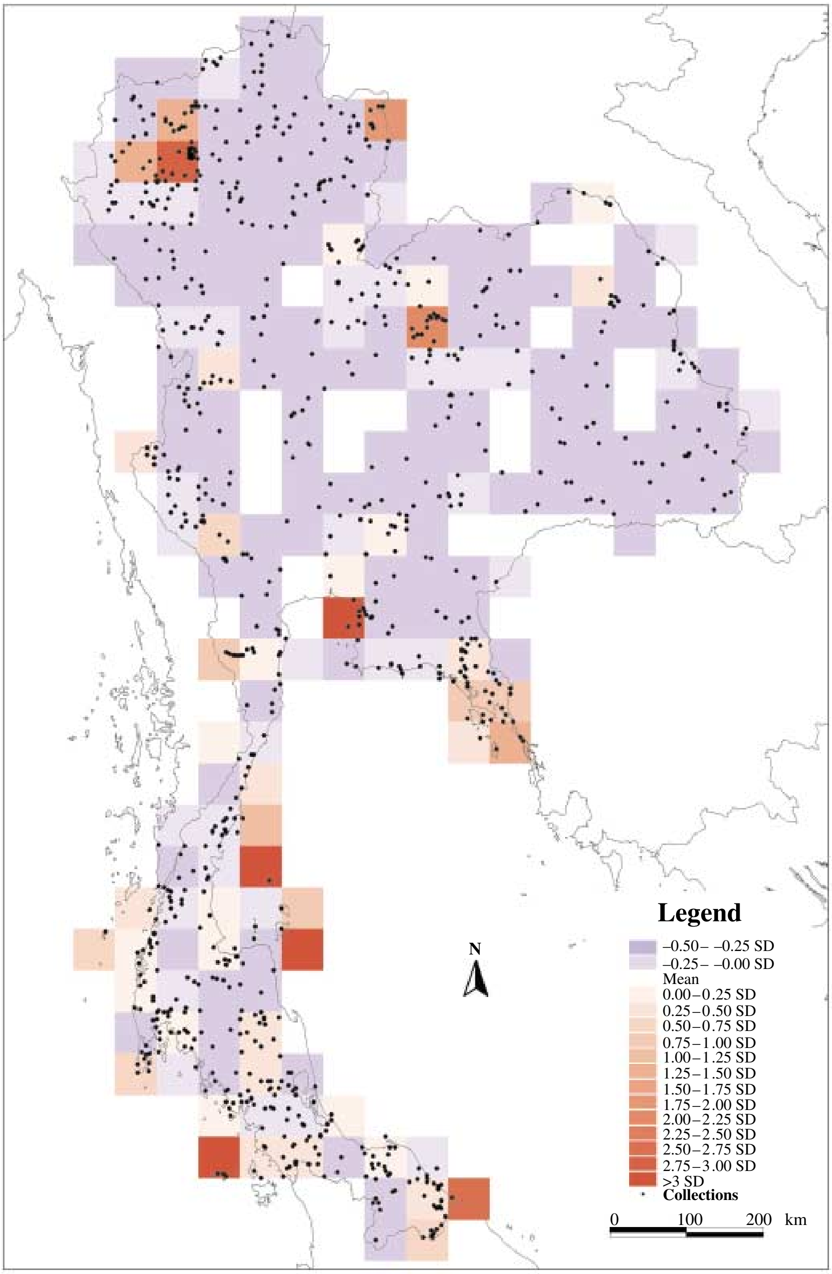

Figure 3 Map of Thailand showing collection density (blue to red dichromatic shading) per quarter degree square. Mean collection density = 72.24 collections per 100 km^2 with 1 SD = z162.577. Collection localities also shown as small dots.

have no collections recorded from them at all. The five southernmost highly collected squares are located either on boundaries or include islands (Fig. 3).

Changwat, vegetation and percentage of forest cover in each Changwat

Figures 4 and 5 show that forest cover, as defined in this paper, is largely confined to north-western, north-eastern and peninsular Thailand with Mae Hong Song (Changwat No. 1) possessing the most forest. To some extent the partitioning of forest by Changwat, shown in Fig. 5, masks the fact that significant areas of forest occur across Changwat boundaries (Fig. 4) – for example, along the southern boundary area of Nakhon Ratchasima (Changwat No. 28) and that small, but significant forest areas occur elsewhere – for example, in Loei (Changwat No. 17) or Chaiyaphum (Changwat No. 27).

Collection density by forest area

Figure 6 shows the collection density by forest area. Comparison of Fig. 6 with Fig. 5 indicates that, in general, Changwat with much forest have low collection densities whilst those with little forest have higher densities. For

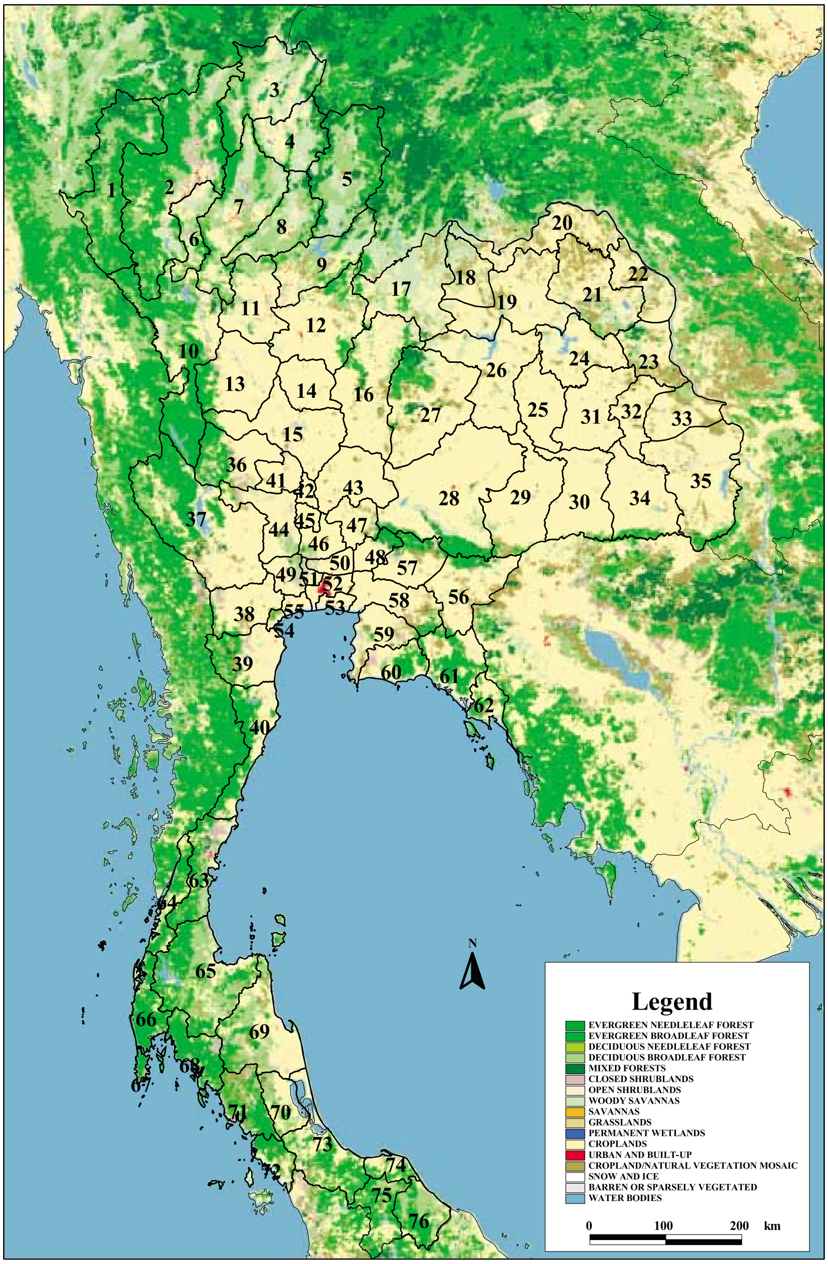

Figure 4 Map of Thailand showing vegetation type and Changwat boundaries. Changwat numbers indicated in bold.

example, Krung Thep (Changwat No. 52) has eighty-four collections with a 1% forest cover which is almost 700 collections per 100 km^2 of forest. However, Mae Hong Son (Changwat No. 1), has 94% forest cover, but only 105 collections and a collection density per 100 km^2 of forest of < 1.

Number of collections per vegetation class

The histogram of actual and expected collection numbers per vegetation class is shown in Fig. 7. The actual number of collections is higher than expected in four classes (evergreen broadleaf forest, deciduous broadleaf forest, closed shrubland and urban/built up areas), while lower in the remaining classes. The highest number of collections is found in the evergreen forest and croplands. Differences between actual and expected numbers were particularly marked in evergreen broadleaf forest, where there was a 95% excess of actual collection numbers over the expected.

Number of collections relative to distance from populated places and roads

Figure 8 shows the distance image map for populated places. Areas of the greatest distance from populated places are predominantly in montane regions in parts of the west, north and centre, as well as near the Cambodian border in the east

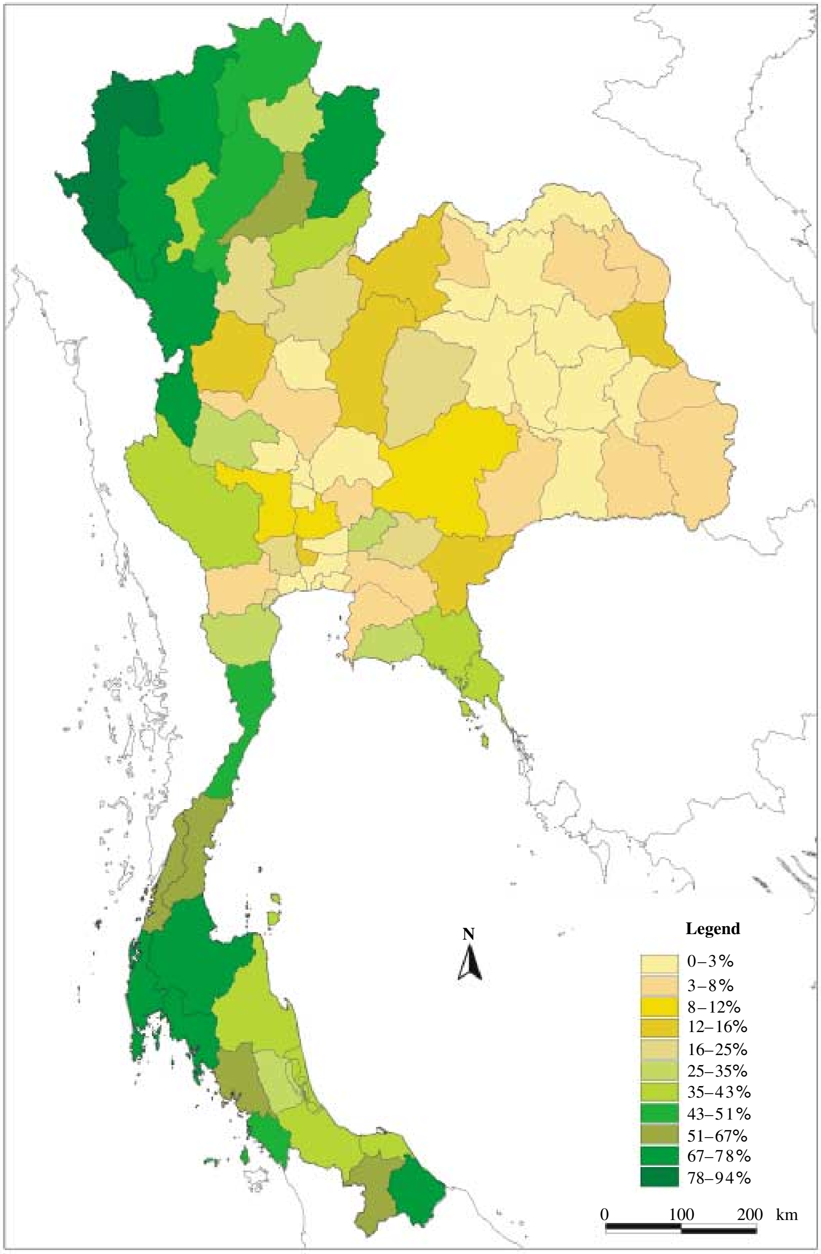

Figure 5 Map of Thailand showing the percentage of forest cover in each Changwat based on Global Land Cover Characteristics and the simple International Geosphere Biosphere Programme classification data.

and southeast. Figure 9 shows the histogram for actual and expected numbers of collections in each distance class for populated places. A general pattern of decreasing collection numbers with increasing distance from populated places is seen, although the expected numbers increase in the 4–8 km size class before showing a decrease in subsequent classes.

Figure 10 shows the distance image map for roads. Again, areas of greatest distance from roads are in the montane regions. Figure 11 shows the histogram for actual and expected numbers of collections in each distance class for roads. The greatest numbers are in the 0–4 km class with rapidly reducing numbers in subsequent classes.

Maps for exemplar species

Figure 12 shows the distribution of the two *Syzygium* species. Both show widespread distributions, but with large gaps, especially in central Thailand. This figure also shows that some apparent gaps in the distribution of these species have been eliminated by records found post-1989.

DISCUSSION

Our results demonstrate that collecting activity is unevenly spread over Thailand, and confirm that the country is

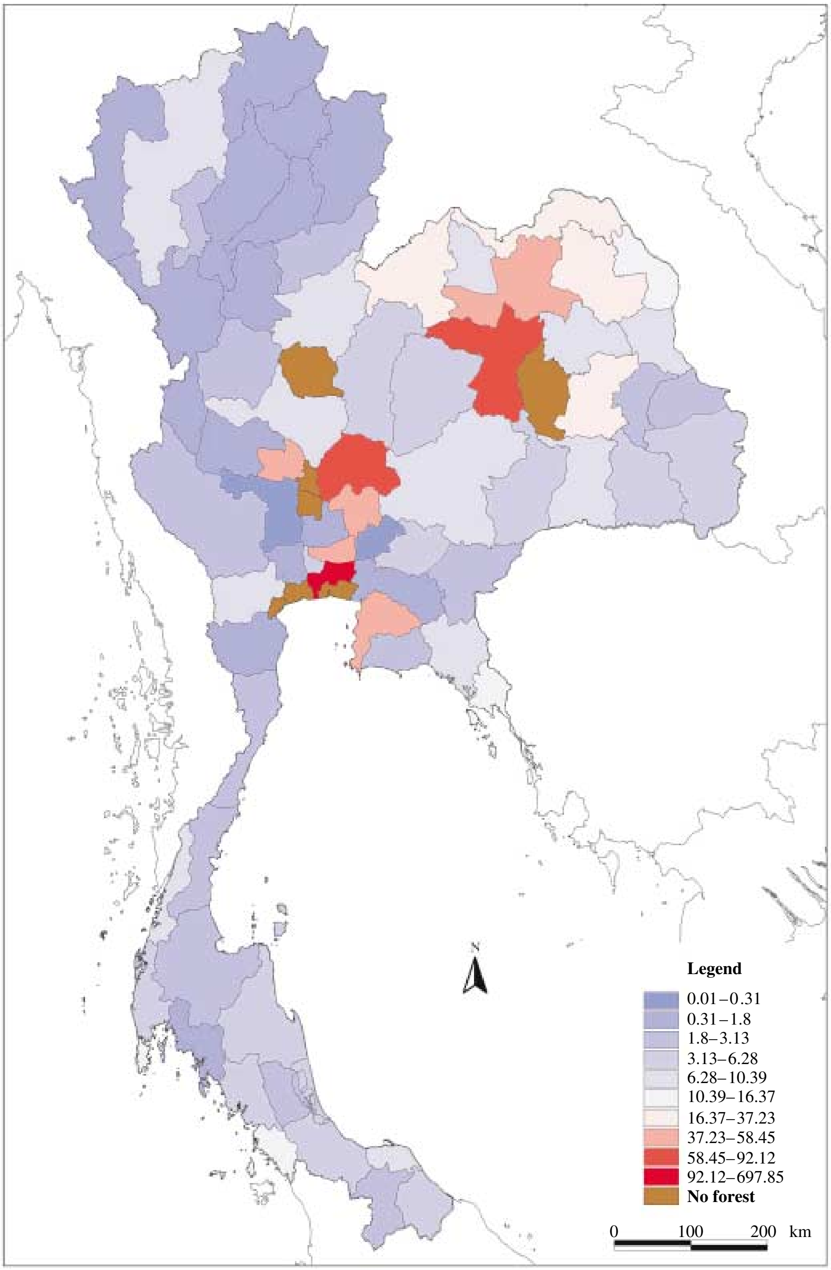

Figure 6 Map of Thailand divided into Changwat, showing collection density (blue to red dichromatic shading, non-forested areas brown) per 100 km^2 of forest per Changwat. Mean = 19.85 collections per 100 km^2. Density is not classified by SD as this gave no differentiation for the negative areas.

generally undercollected, much of it poorly so. The mean number of collections at 1.3 per 100 km^2 is very low, compared with parts of the world that are well known floristically, for example the British Isles, where collections in equivalent-sized areas and families may number thousands. Only 3.3% of quarter degree squares in Thailand with positive standard deviations above 2 could be regarded as well-collected. With 73.4% of the quarter degree squares below the mean and 2% uncollected, the definition of Thailand as an undercollected country is justified.

The reasons for such an uneven spread and low density of collecting are varied. First, it is clear from Fig. 3 that five of the eight quarter degree squares which have collection densities more than 2 SDs above the mean are centred on islands or include country boundaries. The problem with islands is that the resolution of the collection data is to the nearest minute, so that for a small island most locations will be at one locality, so pushing all collections to one grid cell with a small land area. Nevertheless, the collection densities calculated in this paper are accurate. In a sense our results confirm what many field biologists know – that islands are popular to collect on. Secondly, relatively few collectors have worked in Thailand, even in recent years when interest in the Thai flora has increased through the Flora of Thailand

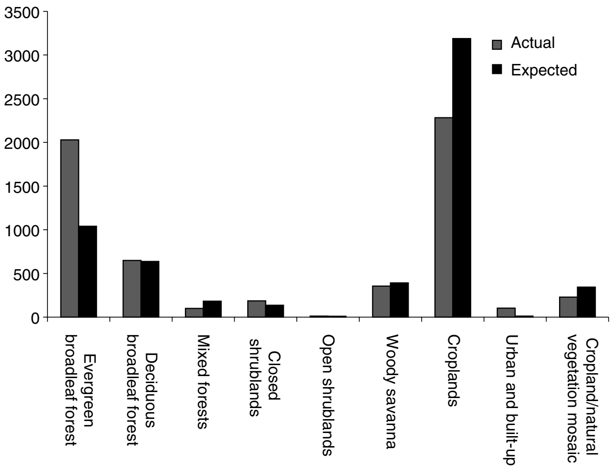

Figure 7 Histogram of the actual and expected number of collections per vegetation class. The *y*-axis indicates number of collections.

project. Many recent workers have concentrated their activities in the national parks, leaving other areas uninvestigated; such uninvestigated areas appear less attractive because they are heavily cultivated with little pristine natural vegetation. Thirdly, other areas have not or have rarely been worked, either because they are remote with poor communications making travel difficult, or because travel has been prohibited for political reasons (e.g. along parts of the Lao and Cambodian borders).

If islands are excluded, the Changwat and quarter degree squares with the highest number of collections coincide with three national parks protecting mountains which have been the focus of collecting activity over many years, namely Doi Suthep, Doi Inthanon [both in Chiang Mai Changwat (Changwat No. 2)] and Phu Kradung [in Loei Changwat (Changwat No. 17)]. Figure 3, which pinpoints individual collecting localities, shows that even in the quarter degree squares covering these Changwat the spread of collecting is poor; almost all the collections in these squares are from these mountains or the surrounding foothills. A similar situation applies to some of the other quarter degree squares with lower positive SDs which also cover national parks. Figure 3 also emphasizes that there are large parts of Thailand, especially near the western and northern border and in the latitudinal centre of the country, where few collections appear to have been made; these include some national parks.

Our data clearly show that mountainous areas are more likely to appear as hot-spots of diversity, partly because they are much more intensively and repeatedly sampled. Similar patterns, reflecting high species numbers on mountains, occur throughout Malesia (e.g. Van Steenis, 1963; Van Steenis & Van Balgooy, 1966; Van Balgooy, 1975, 1984, 1993; Beaman & Beaman, 1990), notably, for example, in Borneo and New Guinea. In the past, such high diversity and its patterning on mountains has been attributed largely to a number of factors including altitude, physical area, geographical position, soil type or the effects of man (e.g. Grubb, 1977; Edwards *et al.*, 1990; Cox & Moore, 1993; Lovett, 1999). Although there is considerable debate about the patterning of diversity on mountains [many authors suggest that the middle altitudes are the richest (e.g. Janzen, 1996) whilst others suggest that diversity decreases with increasing altitude (Stevens, 1992)], it is commonly accepted that mountains in the tropics are centres of diversity for at least terrestrial plants and endemic birds (Hawksworth & Kalin-Arroyo, 1995). Our data show that, in part, in Thailand, and possibly elsewhere, mountainous regions appear as hot-spots of diversity simply because they have been heavily collected relative to the surrounding areas – i.e. lowland regions have been undersampled and their diversity is likely to be underestimated relative to neighbouring mountainous regions. The situation is further complicated by the obvious fact that in Thailand most of the lowlands have been cleared for agriculture making it even more difficult to estimate their natural diversity (Santisuk *et al.*, 1991). However, we do not argue that such undersampling invalidates the case for mountains being centres of diversity; rather that it may overemphasize it.

The analysis of collecting activity in relation to remaining forest area has demonstrated that the Changwat, with much remaining forest cover, are often relatively poorly collected. For the purposes of this section of our analysis we have had to assume that all collections from a Changwat are forest-based; this could be an overestimation as many collections come from outside forested areas, although the latter may have been forested when the collections were made (see below). Therefore, our analysis does not necessarily indicate true rates of collection within a vegetation type, i.e. within forest. Nevertheless, the analysis does indicate that high numbers of collections are not necessarily linked to the presence of forest. As the forests of Thailand are the greatest store of biodiversity in the country (Santisuk *et al.*, 1991), our analysis appears to indicate that far more sampling is required in forested areas.

Such inconsistencies are highlighted by the distribution patterns of the two *Syzygium* species, where some distributional gaps have been eliminated, but where many gaps are still present for two taxa that are easily collectable, common and conspicuous. Either these remaining gaps are real (and therefore indicative of phenomena of biological and/or biogeographical interest) or artefacts of low collecting densities. If the latter applies then any consideration of distribution pattern or inferences made from such patterns are unsound.

Our analysis of the number of collections occurring in a given vegetation class partly reflects what may be expected, i.e. that high numbers of collections are from broadleaf evergreen forest, because it is the most botanically diverse of the classes and is also the type of vegetation which collectors tend to visit most often. The high number for croplands is of interest because this class should be less diverse and therefore less frequented and undercollected. This may be due to several causes, including the collection of large numbers of easily obtained weedy species in these areas, the occurrence

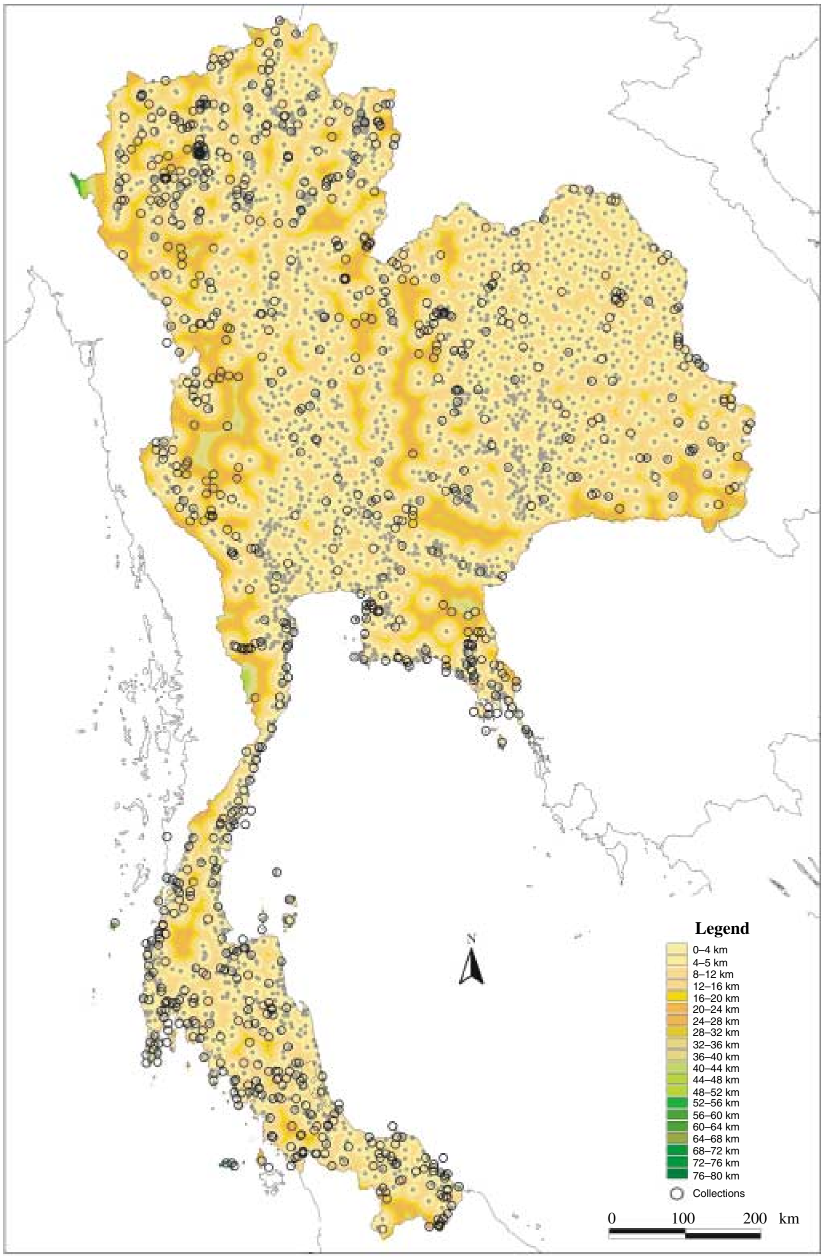

Figure 8 Distance image map of Thailand for populated places based on digital chart of the world (DCW) (ESRI, 1993) data.

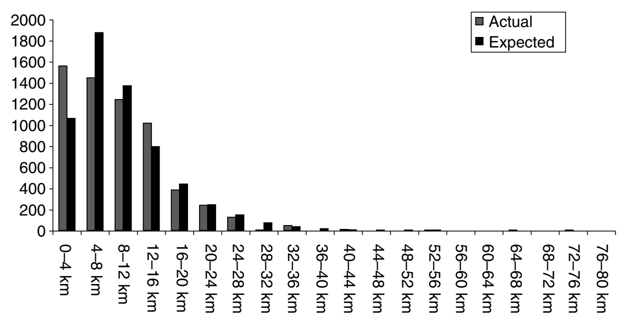

Figure 9 Histogram of the actual and expected number of collections in relation to distance from populated places. The *y*-axis indicates number of collections.

of very small patches of forest below the pixel size and not recorded as forest in the analysis (e.g. in stream valleys), or deforestation. All three causes may come into play, but deforestation must be significant especially when it is considered that the collections date back to the beginning of the twentieth century when forest was much more widespread. Many of these earlier collections could have been from what were then forested areas but which are now croplands.

Our analysis of collections relative to distance from populated places gives a result that suggests most collecting activity takes place within 8 km or less of a populated place. While this may partly reflect the true picture of

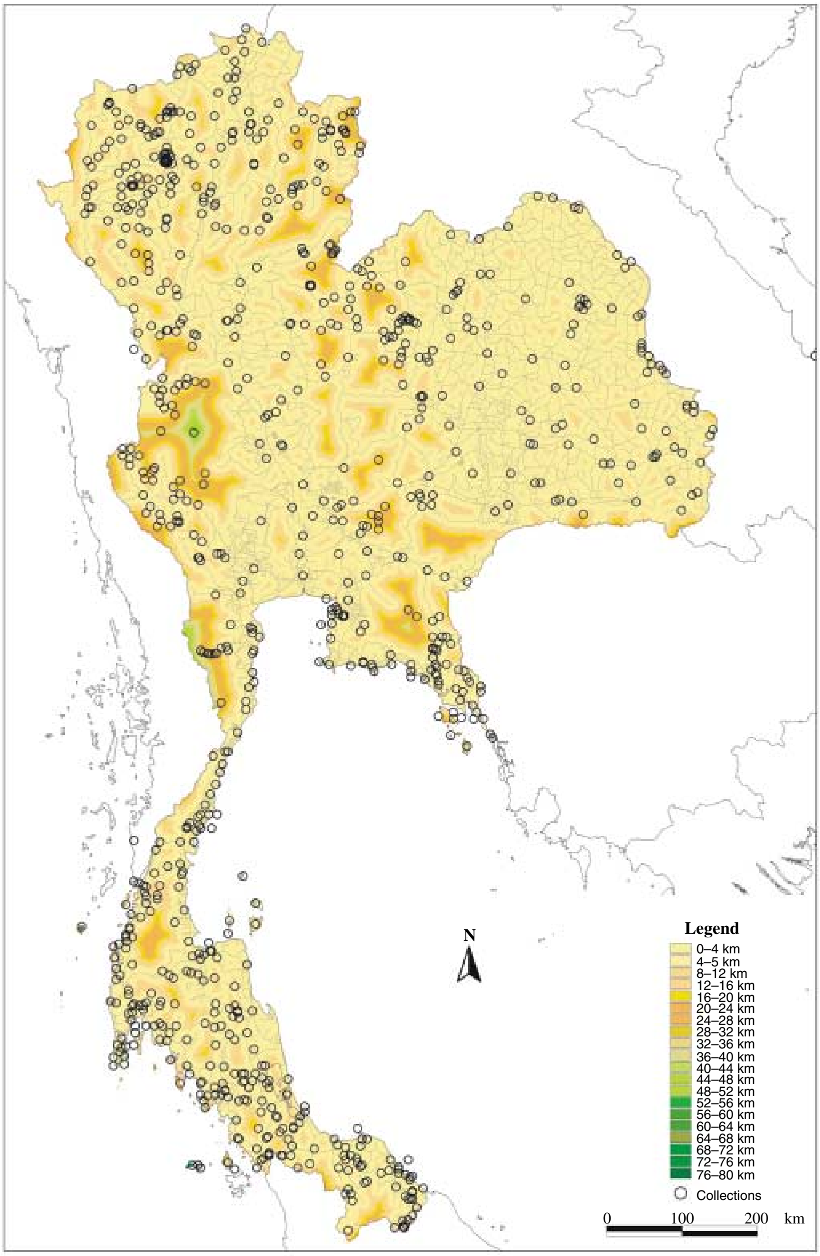

Figure 10 Distance image map of Thailand for roads based on digital chart of the world (DCW) (ESRI, 1993) data.

collecting activity, the method of georeferencing a specimen could also be important. Specimens without a GPS reference have to be georeferenced from gazetteers, based on locality information given on the label. This will often be the name of the nearest populated place, although that place may be several kilometres away from the collecting site. This suggests that populated places have a 'sucking effect' on the georeferencing of specimens. Initial studies (not presented) indicate that the 'sucking effect' extends to a radius of 2 km around a populated place, i.e. if a collection is within 2 km of a populated place then it will tend to be given the grid reference of that place. However, the greater the distance from a populated place that the specimen was collected the more likely it is that the collector would be 'forced' to provide a true locality, resulting in a more precise georeference.

A similar situation may arise when geoferencing in relation to roads. While it is well known that collections can be biased towards roads, there is also the likelihood that the collection was made some distance from the road but was georeferenced to it. Even if a collection was made by the side of the road there may be inaccuracies in the precise location along the road. Initial studies suggest that there may be a 1–2-km discrepancy in georeferencing specimens

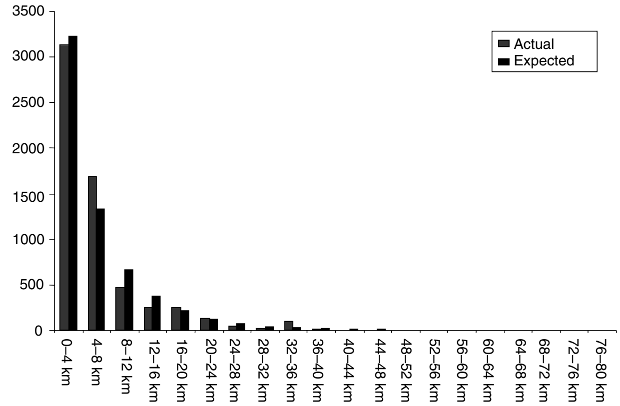

Figure 11 Histogram of the actual and expected number of collections in relation to distance from populated places. The *y*-axis indicates number of collections.

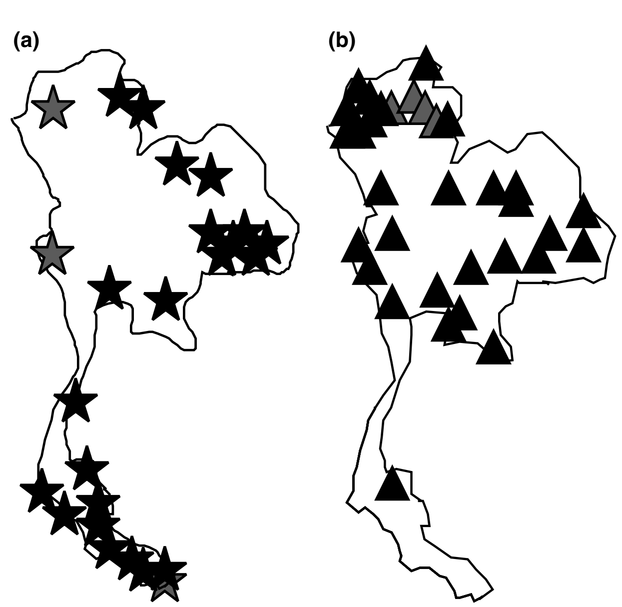

Figure 12 Distributions of (a) *Syzygium cinereum* (Kurz) P. Chantaranothai & J. Parn. and (b) *S. cumini* (L.) Skeels in Thailand. Grey symbols indicate material found post-1989.

in relation to roads. An additional problem when analysing these data is that the road data available for Thailand is out of date, and does not show up some of the roads or tracks that are known to exist. So caution needs to be exercised here.

Therefore, we suggest that at least some collection data may be biased towards populated places and roads rather than the actual collecting localities. Given the age of some of the Thai collections, for which locality data refer solely to the nearest populated place or road, this may be significant and requires further investigation, especially as this situation will apply elsewhere.

As far as we are aware ours is the only analysis of the type for Southeast Asia, although there are a number of currently active floristic projects (e.g. Roos, 1996) in the region and similar types of analysis have been undertaken in the neotropics. Therefore, it is difficult to assess the potential impact that our work could have on countries in Southeast Asia other than Thailand. Nevertheless, our involvement in and knowledge of the progress and state of floristic study of the immediately surrounding countries to Thailand suggests that, although there are differences (for example, all other immediately surrounding countries in the region contain more forest but, in general, have a poorer conservation infrastructure and only Vietnam has a similarly large population to Thailand; Hamilton & Davis, 1998), many of the pressures causing loss of biodiversity factors are consistent over the region. Therefore, our analysis is more likely to be widely applicable. This view is reinforced by a comparison with similar analyses for similar data assembled for Latin America and the Caribbean (Toledo & Sosa, 1993), Guyana (Funk *et al.*, 1999; Funk & Richardson, 2002) and Amazonia (Kress *et al.*, 1998) amongst others. One of the principal outcomes of these authors' work is suggested targeted areas for collecting and centres of biodiversity with concomitant high conservation value.

Funk & Richardson (2002) argue that there are three limitations to the use of collection data for conservation decision-making, viz. that the data are: (1) geographically biased, favouring more easily accessed areas; (2) taxonomically incomplete, including only easy-to-study species, which gives weight to a few taxa; and (3) temporally biased, based on one survey and not usually carried out during the wet season. In relation to these points we have shown that our data set does not necessarily favour more accessible areas; indeed these may be undersampled. Secondly, we argue that many of the groups we have included are taxonomically difficult, some notoriously so (e.g. Cyperaceae, Myrtaceae and Santalaceae). Finally, our data set is not temporally biased, with collecting activity spread over the entire year. Therefore, it does not necessarily suffer the limitations outlined above.

We believe the principal results of our analyses are transferable and the implications of our findings, when related to biogeographical studies, are significant. Given the data that have emerged, we suggest that detailed localized analyses of biogeographical patterns and identification of centres of diversity in Thailand and in many countries in the region are based on incomplete knowledge and are, somewhat, premature. This is especially important as it may impact on conservation planning.

In general, it is accepted that, amongst other attributes, a network of conservation sites (reserves) need to be complementary in nature, that is, where each reserve is selected so as to maximally add to the biodiversity already extant in the network. Vane-Wright *et al.* (1994) applied this idea to Thailand, using data for Hawk moths and Tiger beetles and suggested a prioritized set of reserves. A preliminary analysis of our data set indicates that the suggested reserves fall within areas which are relatively well-collected and that few fall outside. However, as Funk *et al.* (1999) indicate in respect to Guyana, the efficacy of conservation decisions must be measured against the utilized data's quality and reliability. As she states, the really important question is, 'How well do the data represent the biodiversity' of a country. Funk *et al.* (1999) indicate that two important criteria are the coverage and completeness of the data. In terms of coverage we are aware that one of the most widespread and ecologically important tree families in Thailand is as yet not in our data base (the Dipterocarpaceae). In terms of completeness this paper has shown that there are gaps in terms of collecting. Our results therefore parallel those of Funk *et al.* (1999) in Guyana and Kress *et al.* (1998) who found significant areas with few or no collections. For these reasons, and although we believe that our data set will be useful in terms of conservation prioritization and biodiversity assessment, we prefer to await the acquisition of additional data before this is undertaken. We do not wish it to appear that we totally agree with Bullock (as quoted in Lawton *et al.*, 1994) that 'the distribution of a taxon is that of its students and the diversity of a site is a matter of serendipity'. Rather, we wish to balance Bullock's view against the facts that it is unlikely, in the extreme, that tropical regions of the world will, for many decades, approach an asymptote for collecting of new taxa (indeed that is what makes them exciting places to work in) and that forest (biodiversity) is being lost very rapidly.

We believe that the current data set is adequate for a preliminary analysis of biogeographical patterning and its relationship to various parameters – notably topography, climate, conservation, agriculture and sea level changes – and this paper presents some of these findings. In the latter case it is of note that sea level has both risen (Chappel *et al.*, 1996) and fallen (Hanebuth *et al.*, 2000) in the recent past; therefore submerging and revealing land and that sea level rise has never received a significant historical biogeographical treatment in SE Asia (Parnell *et al.*, unpubl. data).

For Thailand we now have a tool which should enable future plant collecting to be targeted on undercollected areas, which in turn will lead to more accurate modelling of biogeographical patterns in the flora. Moreover, given current data basing activities in various herbaria, it should be possible to perform similar analyses for many other tropical countries. Such analyses should be rapidly undertaken so as to allow collecting activity to be targeted and carried out in an economical and efficient way. In turn this will validate hypotheses relating to biogeographical pattern. Without such analyses, biogeographical inferences will, in many cases, remain speculative.

ACKNOWLEDGMENTS

We are grateful to the many Thai colleagues who have helped us in carrying out field work in Thailand, especially Dr Kongkanda Chayamarit and staff at the Royal Forest Herbarium, Bangkok. We also thank the curators of herbaria who have loaned specimens and/or provided facilities during study visits and the referee for helpful suggestions on an early draft of the paper.

REFERENCES

Beaman, J.H. & Beaman, R.S. (1990) Diversity and distribution patterns in the flora of Mount Kinabalu. *The plant diversity of Malesia* (eds P. Baas, K. Kalkman and R. Geesink), pp. 147–160. Kluwer Academic Publishers, Dordecht.

Belward, A.S., Estes, J.E. & Kline, K.D. (1999) The IGBP-DIS 1-km land-cover data set DISCover: a project overview. *Photogrammetric Engineering and Remote Sensing*, **65**, 1013–1020.

Campbell, D.G. (1988) The importance of floristic inventory in the tropics. *Floristic inventory of tropical countries* (eds D.G. Campbell and H.D. Hammond), pp. 6–30. New York Botanical Garden, New York.

Chantaranothai, P. (1989) *The taxonomy of Eugenia* L. *sensu lato (Myrtaceae) in Thailand*. Unpublished PhD Thesis. Trinity College, Dublin.

Chantaranothai, P. & Parnell, J. (1994) A revision of *Acmena, Cleistocalyx, Eugenia s.s.* and *Syzygium* (Myrtaceae) in Thailand. *Thai Forest Bulletin*, **21**, 1–123.

Chappell, J., Omura, A., Esat, T., McCulloch, M. Pandolfi, J., Ota, Y. & Pillans, B. (1996) Reconciliation of late Quaternary sea levels derived from coral terraces at Huon Peninsula with deep sea oxygen isotope records. *Earth and Planetary Science Letters*, **141**, 227–236.

Cox, C.B. & Moore, P.D. (1993) *Biogeography. An ecological and evolutionary approach*, 326 pp. Blackwell Scientific Publications, Oxford.

Edwards, I.D., Payton, R.W., Proctor, J. & Riswan, S. (1990) Altitudinal zonation of the rain forests in the Manusela National Park, Seram, Maluku, Indonesia. *The plant diversity of Malesia* (eds P. Baas, K. Kalkman and R. Geesink), pp. 161–175. Kluwer Academic Publishers, Dordecht.

ESRI (1993) *Digital chart of the world (DCW)*. Environmental Systems Research Institute, Redlands, CA.

ESRI (1996) *ArcView spatial analysis*. Environmental Systems Research Institute, Redlands, CA.

Funk, V.A. & Richardson, K.S. (2002) Systematic data in biodiversity studies: use it or lose it. *Systematic Biology*, **51**, 303–316.

Funk, V.A., Zermoglio, M.F. & Nasir, N. (1999) Testing the use of specimen collection data and GIS in biodiversity exploration and conservation decision making in Guyana. *Biodiversity and Conservation*, **8**, 727–751.

Grubb, P.J. (1977) Control of forest growth and distribution on wet tropical mountains. *Annual Review of Ecology and Systematics*, **8**, 83–107.

Hamilton, A.C. & Davis, D. (1998) *Regional overview: China and East Asia. Centres of plant diversity, Volume 2, Asia, Australasia and the Pacific* (eds S.D. Davis, V.H. Heywood and A.C. Hamilton), pp. 145–174. WWF & IUCN, Cambridge.

Hanebuth, T.J.J., Stattegger, K. & Grootes, P.M. (2000) Rapid flooding of the Sunda Shelf: a late-glacial sea-level record. *Science*, **288**, 1033–1035.

Hawksworth, D.L. & Kalin-Arroyo, M.T. (1995) Magnitude and distribution of biodiversity. *Global biodiversity assessment* (eds V. Heywood and R.T. Watson), pp. 107–192. Cambridge University Press for the United Nations Environment Programme, Cambridge.

Holmgren, P.K., Holmgren, N.H. & Barnett, L.C. (1990) *Index Herbariorum. Part 1. The herbaria of the World*, 693 pp. International Association for Plant Taxonomy, New York.

Janzen, D.H. (1996) Why are there so many species of insects? *Proceedings of the International Congress of Entomologists*, 84–94.

Johns, R.J. (1995) Malesia – an introduction. *Curtis's Botanical Magazine*, **12**, 52–62.

Kress, W.J., Heyer, W.R., Acevedo, P., Coddington, J., Cole, D., Erwin, T.L., Meggers, B.J., Pogue, M., Thornington, R.W., Vari, R.P., Weitzman, M.J. & Weittzman, S.H. (1998) Amazonian biodiversity: assessing conservation priorities with taxonomic data. *Biodiversity and Conservation*, **7**, 1577–1587.

Larsen, K. (1979) Exploration of the Flora of Thailand. *Tropical botany* (eds K. Larsen and L.B. Holm-Nielsen), pp. 125–333. Academic Press, London.

Larsen, K. (1988) Botany in Aarhus 1963–1988. The first 25 years of the Botanical Institute, University of Aarhus. *Reports from the Botanical Institute, University of Aarhus*, **17**, 1–92.

Larsen, K. & Warncke, E. (1966) Report on the Flora of Thailand project – I. Expedition. *The Natural History Bulletin of the Siam Society*, **21**, 251–262 + 10 plates.

Lawton, J.H., Prendergast, J.R & Eversham, B.C. (1994) The numbers and spatial distribution of species: analyses of British data. *Systematics and conservation evaluation* (eds P.L. Forey, C.J. Humphries and R.I. Vane-Wright), pp. 177–195. Clarendon Press, Oxford.

Lovett, J.C. (1999) Tanzanian forest tree plot diversity and elevation. *Journal of Tropical Ecology*, **15**, 689–694.

Parnell, J. (2000) The conservation of biodiversity: aspects of Ireland's role in the study of tropical plant diversity with particular reference to the study of the flora of Thailand and *Syzygium*. *Biodiversity: the Irish dimension* (ed. B. Rushton), pp. 205–216. Royal Irish Academy Special Publication, Dublin.

Parnell, J. & Chantaranothai, P. (2002) Myrtaceae. *Flora of Thailand*, **Vol. 7** (eds T. Santisuk, K. Larsen, I. Nielsen, K. Chayamarit, I. Hedge, J. Dransfieldd, B. Hansen, C. Phengkhlai, J. Vidal, K. Iwatsuki, J.Parnell and M. Roos), pp. 778–914. The Forest Herbarium, Royal Forest Department, Bangkok, Thailand.

Prance, G.T. & Campbell, D.G. (1988) The present state of tropical floristics. *Taxon*, **37**, 519–548.

Roos, M.C. (1996) Charting tropical plant diversity: Europe's contribution and potential. *Systematics Agenda 2000 the challenge for Europe* (eds S. Blackmore and D. Cutler), pp. 55–88. Linnean Society Occasional Publication Number 1. Samara Publishing for the Linnean Society of London, London.

Santisuk, T, Smitinand, T., Hoamuangkaew, W., Ashton, P., Sohmer, S.H. and Vincent, J.R. (1991) *Plants for our future. Botanical research and conservation needs in Thailand*, 48 pp. Royal Forest Department Thailand, Bangkok, Thailand.

Simpson, D., Parnell, J., Chantaranothai, P. & Middleton, D. (1995) The Royal Botanic Gardens Kew, Khon Kaen University and Trinity College Dublin expeditions to Thailand 1990 and 1993. *Thai Forest Bulletin*, **23**, 50–61.

Smitinand, T. & Larsen, K. (1966) I. The flora of Thailand project. Report on the work in 1966. *The Natural History Bulletin of the Siam Society*, **21**, 341.

Stevens, G.C. (1992) The elevational gradient in altitudinal range: an extension of Raportt's latitudinal rule to altitude. *American Naturalist*, **140**, 893–911.

Toledo, V.M.. & Sosa, V. (1993) Floristics in Latin America and the Caribbean: an evaluation of the numbers of plant collections and botanists. *Taxon*, **42**, 355–364.

Van Balgooy, M.M.J. (1975) *Pacific Plant Areas. 3*, 386 pp. Rijksherbarium, Leiden.

Van Balgooy, M.M.J. (1984) *Pacific Plant Areas. 4*, 270 pp. Rijksherbarium, Leiden.

Van Balgooy, M.M.J. (1993) *Pacific Plant Areas. 5*, 251 pp. Rijksherbarium, Leiden.

Van Steenis, C.G.G.J. (1950) The delimitation of Malaysia and its main plant geographic regions. *Flora Malesiana*, Series 1, Vol. 1, lxxvi–cvi. Nordhoff-Kolff N.V., Djakarta.

Van Steenis, C.G.G.J. (1963) *Pacific Plant Areas. 1*, 297 pp. National Institute of Science and Technology, Manila.

Van Steenis, C.G.G.J. & Van Balgooy, M.M.J. (1966) *Pacific Plant Areas. 2. Blumea*, **5**, Suppl., 312 pp.

Vane-Wright, R.I., Smith, C.R. and Kitching, I.J. (1994) Systematic assessment of taxic diversity by summation. *Systematics and Conservation Evaluation* (eds P.L. Forey, C.J. Humphries and R.I. Vane-Wright), pp. 309–326. Clarendon Press, Leiden.

BIOSKETCHES

J. Parnell is a higher plant taxonomist, a member of the editorial board of the Flora of Thailand and Head of the Department of Botany, Trinity College, Dublin, who has been involved in revisions of a number of families for the Flora of Thailand project.

D. Simpson is a higher plant taxonomist, a member of the editorial board of the Flora of Thailand, Head of Commelinid monocots in Kew, UK, and has revised the Cyperaceae for the Flora of Thailand project.

J. Moat is a Geographical Information Systems (GIS) specialist and Head of the GIS Unit at Kew.

D. Kirkup is a higher plant taxonomist at Kew and computer specialist interested in interactive keys, Loranthaceae of Africa, pollination and floral anatomy.

P. Chantaranothai is a higher plant taxonomist who has revised a number of families for the Flora of Thailand project and an Associate Professor in Biology at Khon Kaen University.

P. Boyce is a higher plant taxonomist specializing in Araceae.

P. Bygrave is a higher plant taxonomist specializing in Annonaceae.

S. Dransfield is a higher plant taxonomist specializing in Bambusoid Poaceae of Southeast Asia. She is an Honorary Research Fellow at Kew.

M. Jebb is a higher plant taxonomist and Curator of the Herbarium at Glasnevin, who has revised a number of families for Flora Malesiana and the Flora of Thailand.

J. Macklin is a specialist in the Santalaceae of Southeast Asia who has revised the family for the Flora of Thailand.

C. Meade is a post-doctoral researcher at the University of Maynooth and specialist on the *Uvaria* group of Annonaceae.

D. Middleton is a tropical plant systematist specializing in Apocynaceae of Asia and has produced a number of generic revisions and a family revision for the Flora of Thailand.

A. Muasya is a higher plant taxonomist specialist in Cyperaceae petaloid monocotyledons and Lobeliaceae.

C. Pendry is a specialist in Polygalaceae who has revised the family for a number of Southeast Asian floras.

A. Prajaksood is a higher plant taxonomist specializing in Eriocaulaceae of Thailand and lecturer in Biology at Khon Kaen University.

R. Pooma is a specialist in Dipterocarpaceae and the Flora of Thailand.

S. Suddee is a specialist in Labiatae and the Flora of Thailand.

P. Wilkin is a specialist in Dioscoreaceae in the Old World. He is Head of Alismatid and Lilioid Monocots in the Herbarium at Kew.