Honey Pollen: Using Melissopalynology to Understand Foraging Preferences of Bees in Tropical South India

Raja Ponnuchamy [1,2], Vincent Bonhomme [1¤a¤b], Srinivasan Prasad [1], Lipi Das [3], Prakash Patel [3], Cédric Gaucherel [1], Arunachalam Pragasam

and provide methodological recommendations for further melissopalynological experimental designs.

Materials and Methods

Ethics statement

The study was conducted in privately owned areas where prior permission was obtained: site 1 to which two of the authors (LD and PP) are affiliated, and site 2 managed by Bernard Declercq and Deepika Kundaji, Auroville. The study did not include any human subjects or vertebrate animals, and honey was collected with minimal disturbance to the honeybees.

Study site

The study was conducted in four locations in two sites that are 6 km apart. The first three locations were in site 1, spread over 160 ha at 40–50 m a.s.l. (11°57′18.3″ N, 79°45′57.2″ E). Location 4 was in site 2 (11°59′20.4″ N, 79°47′4.2″ E and 50 m a.s.l.), spread over 10 ha (Figure 1).

In the Puducherry region, rainfall is distributed over a part of the summer and a part of the (tropical) winter, i.e., from July to January. Much of the rain occurs during the northeast monsoon between October and December (records from Puducherry weather station). Mean temperature in December was 25°C for the years from 1911 to 1961 [24]. In the last three decades (1980–2009), the pattern remained similar.

Bee species, beehive locations, and surrounding vegetation

As part of an ongoing eco-restoration project, particular effort was made to rear the Asiatic honeybee *Apis cerana* Fabricius, one of the native honeybees of India [25]. The bees were reared in wooden beehives placed in four locations. The locations were chosen to reflect the complex mosaic of vegetation (Figure 1; [26]). One beehive per location was considered in this study.

The four locations are Garden and Tropical Dry Evergreen Forest (GT), Coconut Grove (CG), Agricultural Field (AF), and Scrub Jungle (SJ). GT contains ornamental plants, several introduced drought-tolerant *Acacia* spp., Casuarina junghuTerove

variance explained by the successive PC axes were calculated using the Eigen values of the variance/covariance matrix. LDA was performed using the package MASS [41] to test whether the discriminant linear combinations of pollen taxa were able to discriminate between groups. The standardized discriminant coefficients were used to compare the relative importance of the pollen taxa in order to discriminate within/between groups of samples. Three LD axes were obtained for Month and Location (4 levels each) and two LD axes were obtained for the factor Year (3 levels each).

The LD axes that brought the maximum between-groups differences were considered for each of the factors. LDA was first performed on the complete dataset of 80 taxa. We then tested if LDA can also be used in our case as a tool to classify additional pollen spectra based on this discriminant function, using the leave-one-out cross-validation implemented in the MASS package [41]. The idea behind this validation is simple: every pollen spectrum is successively removed from the dataset and an LDA is performed and then used to classify the pollen spectrum removed. For the final analyses, we retained only the 51 taxa that occurred at least 3 times in our dataset of 42 samples.

Results

Honey pollen content

The pollen analyses of the 42 honey samples collected during the three years yielded 80 pollen taxa/types: 72 dicotyledonous and 8 monocotyledonous, encompassing 41 botanical families spread into seven life forms namely, trees, shrubs, epiphytes, herbs, climbers, grasses, epiphytes and sedges (Table S1). On average, 6983 ± 3097 pollen grains (average ± 1 SD are given throughout the text) were counted (min = 283, max = 12356, median = 7591), and 18 ± 7 pollen types (min = 5, max = 30, median = 18.5) were identified per sample.

Three families accounted for more than half the total number of pollen grains: Arecaceae (29%), Anacardiaceae (14%), and Mimosaceae (11%). Arecaceae was present in all, but one, samples. The three most abundant life forms, expressed as percentages of the total pollen count, were trees (39%), shrubs (24%), and herbs (20%). Forty-four out of 80 taxa were recorded at $\geq$3% of the total (Figure 2). The graph at the extreme right in this figure shows the total number of pollen counted in each sample. Twenty-six honey samples were found to be uni-floral (one pollen taxon occurring at $\geq$45% of the total pollen count per sample), represented by 14 predominant pollen taxa (Figure 3). As expected, pollen spectra and predominant pollen taxa varied in space and time, a few examples being *Lannea*, *Dodonaea*, *Phoenix*, *Borassus*, and *Cocos* (Figure 2).

Statistical analyses

We performed MANOVAs on the distance matrices calculated using the two similarity indices (Table 1 & Table 2). All the tested factors were highly significant and hence the pollen spectra were not highly replicable.

The average Bray-Curtis (BC) index was 0.3 and the average binary Bray-Curtis (bBC) index was 0.4, graphically shown as heatmaps (Figure 4). Taken individually, each factor was highly significant ($P < 10^{-4}$) and retained in the final model. Only the two interactions Year $\times$ Location ($P < 10^{-3}$) and Year $\times$ Month

Figure 3. Photomicrographs showing the 14 predominant pollen types *i.e.*, uni-floral origin (a single pollen type represented >45% of total observed pollen types in a sample) in the 42 honey samples arranged ascending order of family: A–B *Lannea*; C–D *Mollugo*; E–F *Borassus*; G–H *Cocos*; I–J *Phoenix*; K–L *Delonix/Peltophorum*; M–N *Evolvulus*; O–P Compositae-echinate; Q–R *Securinega*; S–T *Acacia*; U–V *Mimosa pudica*; W–X Rhamnaceae; Y–Z *Atalantia*; A1–B1 *Dodonaea*.
doi:10.1371/journal.pone.0101618.g003

($P<10^{-4}$) were significant in the bBC model; only Year × Month ($P<10^{-2}$) was retained in the final BC model (Tables 1a & 1b). The heatmaps showed accordingly no well-defined patterns: when samples were arranged location-wise, there were no higher similarity values along the matrix diagonal (Figure 4), as expected. The only ''structure'' was that the least similar sample pairs (blue and shades of blue) correspond to samples in different locations in a given year and month (Figure 4, BC model). Similar heatmaps were obtained when arranged month-wise and year-wise.

Only a few highly abundant taxa (*Dodonaea*, *Lannea*, *Phoenix*, and *Acacia*) appeared distinctive in the plots on the first two PC axes. The first two PC axes captured 38% (PC1 = 21%; PC2 = 17%) of the total variance (Figure 5D). The graphical relative distribution of the pollen taxa showed two trends on the plot using PC1 and PC2 as principal axes (Figure 5A–D). Moreover, *Dodonaea* and *Lannea* were pronounced on the two directions along two PC axes. First, a clear Month effect was identified (Figure 4A). Then, in the factors Year (Figure 5B) and Location (Figure 5C), such grouping was not as pronounced. In addition, most multi-floral honey samples overlapped.

In the final LDA, retaining only 51 taxa, samples were well classified with reference to all three factors (Figure 6). The discriminant coefficient values were high for 15 taxa with reference to one or more factors (Figure 6). *Madhuca* was identified as the most discriminant taxa for Month and Year and, overall, the taxa with the highest discriminant value. Other taxa with high discriminant values for all factors were *Cassia*, *Grewia*, Commelinaceae, and Malvaceae. Most of the taxa with high discriminant power were ''low abundance'' taxa. The discriminant coefficient values were low (<0.01) for the remaining 35 taxa (Figure 6). Several of the set of taxa with coefficient values closer to zero corresponded to ''high abundance'' taxa (*e.g.*, *Dodonaea*, *Lannea*, *Phoenix*, *Acacia*, and *Cocos*), of which many were pronounced along the first two PC axes. Results of leave-one-out cross validation also indicated that only one sample was misclassified for each of the tested factors: Month (''CGn08''), Year (''GTn09''), and Location (''SJa09''). Two out of 3 misclassified samples were designated as multi-floral origin. The analysis helped delineate the pollen taxa of importance *vis-à-vis* the considered factors, found in both high and low proportions in the honey samples.

Discussion

Our results showed that pollen spectra were equally comparable between Locations and also between Months and Years; the importance of this result, is that it helped to demonstrate the complexity of ecological/environmental phenomena involved in the process of foraging by bees in a heterogeneous and complex landscape. This shows that single, random samples of honey are not likely to provide reliable replicates of the pollen spectra. Furthermore, samples and taxa groups were well delineated based on the three factors considered; the importance of this result is that we now have a tool to classify additional pollen spectra, even when there is a low overall replicability.

The honey pollen content reflected the vegetation characterized by Tropical Dry Evergreen Forest (TDEF) species typical of Coromandel Coast [42–44]; markers were both predominant (*e.g.*, *Lannea*, *Dodonaea*, and *Mollugo*) and less-represented (*e.g.*, Melastomataceae/Combretaceae, *Drypetes* and *Glycosmis*). Taxa such as *Lannea*, *Dodonaea*, *Cocos* and *Borassus* have been reported to be foraged by the same bee species in other parts of south India [8,45]. Some taxa such as Poaceae and Cyperaceae, generally reported in very low proportions, were frequently found in our samples, sometimes in considerable proportions. The present study, one of the first comprehensive melissopalynological contributions to the TDEF in the context of plant-pollinator interaction, is unique in that it documents, over time and space, differences in the pollen contents of honey, even within the confined landscape mosaic. At larger spatial sampling scales, this method is likely to throw up even more differences. As floral availability in the immediate vicinity of the beehive was ensured at all sites during most seasons, we assumed that the bees did not go into neighbouring sites to forage.

Bees are considered to be predominant pollen vectors in tropical forests [46–49], yet studies limited to bees in Southeast Asia are rare [48], thus there is an urgent need for forest bee community level data [50] for natural regeneration and restoration.

We found 80 pollen taxa from 42 samples. In a compilation of honey collected across a few hundred kilometers of Andhra Pradesh in south India, 104 taxa were reported from 164 samples [8]. In general, melissopalynological studies used random sampling because the main concern was determining the broad geographic

Table 1. MANOVA table for binary Bray-Curtis' distance.

MANOVA on binary Bray-Curtis' distances						
Response variable: Bray-Curtis' distances	df	SSQ	MSQ	F	R^2	Pr(>F)
Year	2	1.3910	0.6955	4.3182	0.0906	<0.0001
Location	3	1.1849	0.3950	2.4522	0.0772	<0.0001
Month	3	4.5184	1.5061	9.3511	0.2944	<0.0001
Year × Location	6	1.5822	0.2637	1.6373	0.1031	0.0021
Year × Month	6	3.2889	0.5482	3.4033	0.2143	<0.0001
Residuals	21	3.3823	0.1610		0.2204	
Total	41	15.3477			1	

doi:10.1371/journal.pone.0101618.t001

Table 2. MANOVA table for Bray-Curtis' distances. df, SSQ, MSQ stand for degrees of freedom, sum of squares, mean squares.

MANOVA on Bray-Curtis' distances						
Response variable: Binary Bray-Curtis' distances	df	SSQ	MSQ	F	R^2	Pr(>F)

Figure 5. Multivariate analyses (PCA) showing the structure of pollen spectra in reduced dimensionality of absolute pollen frequency. A = Month-wise, B = Year-wise, C = Location-wise, and D = Percentage of Eigen value and overall variance distribution.
doi:10.1371/journal.pone.0101618.g005

origins of honey, which did not require long term monitoring. Thanks to our methodology of achieving high pollen counts, we report here, in a smaller geographic space, a comparable number of taxa, allowing us to exploit other statistical analyses.

We found a higher degree of similarity with qualitative index (bBC) and no structure with quantitative index (BC), probably due to the complex web of factors influencing the pollen content of honey. Apart from phenology and such other ''external'' ecological/environmental factors, factors related to bee behaviour such as the individual ability of some bees to remove more or less pollen from the nectar they collect [23] also come into play in this complexity. Similarity indices and ordination analyses were used to classify honey samples based on their spatial location over two provinces in Spain [20]. They found similar trends, the focus being botanical and again broad geographic origins. However, the temporal variations were not taken into account in that study. To our knowledge, in India only one study [51] has calculated qualitative indices on 6 samples collected in the same season; quantitative indices have not been reported for India.

Recent studies have effectively used similarity indices as well as ordination techniques such as PCA and LDA [21,52–53] to classify the honey samples in terms of their botanical and geographic origins. Because of the temporal dimension in our study adding MANOVA was useful. Combining multivariate techniques (MANOVA, PCA, and LDA) with the pollen counts helped us to classify the samples in terms of factorial influences and understand the dynamics of bee foraging preferences.

Ordination analyses helped delineate pollen taxa such as *Dodonaea*, *Lannea*, *Phoenix* and *Acacia* collected by the bees in large proportions. Species corresponding to these taxa have a distinct flowering season in contrast with *Cocos*, which remained available through most seasons. Even during the peak flowering of these taxa, the bees continued to visit and gather pollen from *Cocos*. This finding supports the results of others [8,45] regarding the bees' preference of *Cocos* pollen.

Lannea and *Dodonaea* were consistently recorded during summer and winter seasons, respectively, but their abundances varied inter-annually due to variations in rainfall. Seasonal variations in rainfall and soil water availability drive flowering periodicity [54–56]. In a tropical context, the physical condition of the sites, the neighboring vegetation, and the effect of animals influence the variation of individual phenology [57]. These may have reflected spatial differences in the pollen assemblages within the vegetation mosaic of our study area, which however seemed less pronounced than the other (temporal) factors like phenology. Thus, this method can be used for tracing floral phenology across years (scarce versus good flowering) and its effect on bees as the major pollinator in the community.

Discriminant analysis helped to highlight taxa found in small proportions in the honey, such as *Madhuca*, *Cassia*, *Grewia*, Commelinaceae and Malvaceae. Our direct observations [Pon-

Figure 6. Multivariate analyses (LDA) showing the group membership of honey with reference to spatio-temporal factor and discriminant coefficient value of individual pollen type. For Month and Location, the third LD axis did not bring major between-groups differences; therefore only the first two axes have been retained.
doi:10.1371/journal.pone.0101618.g006

nuchamy (2014), PhD Thesis, Pondicherry University, India] corroborate that they frequently visit the species corresponding to these taxa. Though frequently in low abundances, the consistent presence of these taxa in the honey suggests that, rather than serving as the bees' reward, they may actually be getting the benefit of pollination (by the bees). Some other studies [21–22,53] also used discriminant analyses to highlight low abundance taxa. Though not in the purview of the present study, this highlights the need to quantify the pollen present in the nectar [58], for the tropical plant taxa, documented here as potential bee-plants (Table S1).

Though our analyses highlighted only a few outliers or misclassified samples, it seems likely that this may be a result of source scarcity or bees' preference. Some additional taxa with low discriminant power such as *Atalantia*, *Phoenix* and *Ziziphus* are among the broad geographic indicators of the Coromandel Coastal environment.

Part of the reason for the apparent wide variation in the pollen data results may be directly attributed to sampling a combination of both pollen cells and honey cells. We accept that such sampling could account for much of the differences found in the samples collected during different seasons as well as years. In other words, our results reflect not only the primary nectar sources of the bees but also the primary pollen collecting plant sources. Studies in other geographic areas [9,46] have shown that the primary nectar and primary pollen foraging plants are usually different. It is true that sometimes in a hive a minor amount of pollen might be collected from a primary nectar source, as the bees eliminate much, but not all, of the pollen they collected with the nectar; the reverse is sometimes also true when the bee has no time to remove large quantities of pollen from the nectar, but rarely are the same

32. Moore PD, Webb JA, Collinson ME (1991) Pollen analysis. 2nd edition, Blackwell Scientific Publications, 216pp & 271 plates.
33. Huang T-C. (1972) Pollen Flora of Taiwan. National Taiwan University Botany Department Press, 297pp & 177 plates.
34. Vasanthy G (1976) Pollen of the South Indian Hills. Institut Francais de Pondichery, 74pp & 34 plates.
35. Nayar TS (1990) Pollen flora of Maharashtra State, India. Today & Tomorrow's printers & publishers, New Delhi, 157pp & 67 plates.
36. Tissot C, Chikhi H, Nayar TS (1994) Pollen of wet evergreen forests of the Western Ghats, India. Publications du departement d'écologie, Institut Francais de Pondichéry, 55pp & 78plates.
37. Louveaux J, Maurizio A, Vorwohl G (1978) Methods of melissopalynology. Bee World 59: 139–157.
38. Grim EC (1987) TILIA and TILIA * Graph. Illinois state Museum, Springfield, IL, USA (http://wwwncdcnoaagov/paleo/tiliahtml).
39. Oksanen FJ, Blanchet G, Kindt R, Legendre P, Minchin PR, et al. (2011) vegan: Community Ecology Package. R package version 2.0-2. http://CRAN.R-project.org/package = vegan.
40. Dray S, Dufour AB (2007) The ade4 package: implementing the duality diagram for ecologists. Journal of Statistical Software 22: 1–20.
41. Venables WN, Ripley BD (2002) Modern Applied Statistics with S. Fourth