Pollen meta-barcoding reveals foraging preferences of honeybees (*Apis mellifera* L.) along space-time gradient in Japan

Grégoire Noël^{1¶}, Arnaud Mestrez^{1¶}, Philippe Lejeune², Frédéric Francis¹, Masayuki Miwa³, Koichi Uehara⁴, Ayako Nagase^{*4}

¹ Functional and Evolutionary Entomology – Gembloux Agro-Bio Tech (University of Liège).
Passage des Déportés 2, 5030 Gembloux (Belgique)

² TERRA Teaching and Research Centre (Forest is Life), Gembloux Agro-Bio Tech,
University of Liege, Passage des Déportés 2, 5030 Gembloux (Belgique)

³ Center for Environment, Health and Field Sciences, Chiba University, Kashiwa-no-ha 6-2-1,
Kashiwa, Chiba, Japan

⁴ Graduate school of Global and Transdisciplinary Studies, Chiba University, 1-33 Yayoi-cho,
Inage-ku, Chiba, Japan

* Corresponding author: anagase@chiba-u.jp

¶ These authors contributed equally to this work.

Short title: Foraging preferences of honeybees in different location and seasons in Japan

Abstract

1 The availability of pollen in urban-rural landscapes is an essential factor that influences the
 2 population dynamics of insect pollinators. The amount and diversity of pollen play a pivotal
 3 role in the foraging ecology of pollinators for their growth and health, but investigations on the
 4 spatio-temporal patterns of foraged plants remain rare, especially in cities as neo-ecosystems.
 5 Here, we explored the temporal foraging habits of a highly polylectic pollinator (*Apis mellifera*
 6 L.) in a study area, including different landscape classes from rural to urban areas. Mixed-pollen
 7 in each month and each location (N = 17) were analysed using DNA meta-barcoding to identify
 8 plants visited by honeybees. The results showed that the landscape class (rural, suburban and
 9 urban areas) explains spatial variations in the plant composition foraged by honeybees, but not
 10 in taxa richness. Furthermore, pollen diversity and plant composition showed a strong seasonal
 11 dependence. Furthermore, a higher plant richness and foraged woody taxa was found to occur
 12 in spring, which was mainly dominated by the genera *Prunus* and *Acer*. In summer and autumn,
 13 the genera *Trifolium* and *Plantago* of the herbaceous stratum were the most visited plants. The
 14 Fabaceae, Rosaceae, Brassicaceae, Plantaginaceae, and Onagraceae plant families were the
 15 most frequently observed in all combined samples. The present study contributes to a broader
 16 understanding of the ecology and floral preferences of honeybees, on which urban planning can
 17 rely to promote biodiversity in cities.

19

20 **Introduction**

21 Ongoing urbanisation is one of the main drivers of landscape degradation and pollinator
 22 biodiversity loss [1–5]. Indeed, floral resources are becoming scarcer under the pressure of
 23 urban fragmentation, and the increase in impervious surfaces is rendering nesting sites
 24 inaccessible to pollinators [6,7]. However, recent studies have revealed that cities can also act
 25 as a refuge for pollinators [8], particularly for bees [9]: (i) cities are less exposed to pesticides
 26 [5,10], (ii) urban management sustainably permits the maintenance of their floral resources
 27 [11,12], and (iii) urban areas are configured with a heterogeneity of green spaces, which would
 28 be favourable to the foraging preferences of bees [13]. Moreover, flowerbeds in the urban
 29 matrix are highly attractive and are a source of pollen and nectar for insect pollinators [14,15].
 30 Therefore, plant selection in greening projects must be pollinator-friendly. With an increasing
 31 popularity in beekeeping activities in cities, honeybees (*Apis mellifera* L.) contribute to urban
 32 plant pollination, generate profits of by-products, and provide environmental education [16,17].
 33 However, the massive introduction of urban honeybees has led to growing concerns about
 34 detrimental effects on wild pollinators through an increase in floral resource competition and
 35 the spillover of shared pathogen agents [18,19].

36 As a eusocial species, honeybees organise their floral resource collections (i.e. nectar and
 37 pollen) through a complex communication system within their colonies. According to plant
 38 phenology, honeybee scouts rapidly recruit their siblings to forage rich new patches of flowers
 39 using a characteristic waggle dance [20]. The foraging decision-making system of the colony
 40 can vary from day to day or within the same day following real-time nectar and pollen
 41 availability in the surroundings. Throughout its active seasons, the colony constantly maintains
 42 a balance between its biomass and energy management, combined with a diversity of floral

sources, to fully benefit from the abundance period, and makes reserves for incoming winters [21,22]. Pollen diversity provides substantial resources in terms of protein, lipid, vitamin, and mineral supplies [23]. Large amounts of pollen (15–30 kg) are collected annually, mainly for brood production during summer [21,24,25]. The quality and diversity of pollen are also essential for strong health and better immunity, as well as longevity and parasite tolerance of the bees and the colony [26–29].

The preservation of ecosystem functioning relies on the mutualistic networks of pollinators and plants. Several methods are used to assess these interactions: the observation of floral visits, digital tracking systems to capture floral visits, chemical signatures of pollen, pollen genetic sequencing, and pollen light microscopy [30]. Identifying a pollen species or genus by light microscopy from mixed pollen samples, also known as melissopalynology [31], is a time-consuming process, even for well-trained experts [32] that results in low taxonomic resolution, usually at the family or genus rank at best [33–36]. With the advent of high-throughput sequencing (HTS) techniques, DNA meta-barcoding approaches have become reliable methods to obtain faster taxonomic profiles with higher resolution of mixed pollen collected from bees or flowers [37–39]. To elucidate the plant taxonomic composition of mixed-pollen samples, the meta-barcoding process can be based on different genetic markers, such as the *rbcLa*, *matK*, *trnH-psbA*, *trnL*, and *ITS* regions (mainly *ITS2*), which require high inter-specific and low intraspecific variability [36,39–42]. These selected loci and the primer set used for amplification drive the range of taxonomic inferences [33]. The identification and quantification of plants visited by pollinators is an ecological issue for natural resource management and preservation in urban environments [36,43]. Indeed, the fragmentation of urban matrix usually leads to the creation of small, remote, and intensely maintained green spaces [44], which could influence the foraging distance of honeybee workers in terms of pollen richness and quantity [45]. The composition of agricultural or urban landscapes heavily impacts

the foraging preferences of honeybee workers [46,47]. As the most polylectic bee forager [48], honeybees can mitigate the degradation of floral resources by enlarging their foraging area [49,50]. Moreover, seasonal shifts greatly impact the pollen availability for honeybee colonies, according to the phenology of the floral resources [51,52]. In temperate climates, the foraged plant traits also vary according to the course of the seasons: spring is dominated by trees and shrubs, summer has more herbaceous species, and autumn is characterized by woody vines [52]. Nonetheless, the space-time effects combined with plant traits in the foraged plant community have yet to be studied extensively [53].

The present study aimed to determine the floral preferences of *A. mellifera* to promote pollinator-friendly plant biodiversity through integrated urban greening projects. To achieve this aim, the taxonomy of pollen foraged by honeybees was identified over the seasons along an urban-rural gradient from different locations in the Kanto region, Tokyo, and its surroundings in Japan. The research questions addressed were as follows: (i) How does the composition of the foraged flower community (and the foraged plant traits) vary according to an urban-rural gradient? (ii) How does the composition of foraged flower communities (and foraged plant traits) vary according to the course of the seasons? Lastly, the pollen availability of plant taxa composition foraged by honeybees was evaluated.

Material and methods

Study area and experimental set-up

We selected 17 apiary locations, with homogeneous climatic and altitudinal conditions (Table S1), distributed along the urbanisation gradient in the Kanto region of Japan (Fig 1). One hive per apiary was used for pollen sampling. From March to September 2019, pollen samples were collected using pollen traps at the entrance of the same hive (Fig S1). This experimental setup was used to collect pollen balls from the hind legs of honeybee foragers with a standardised

honeybee-size mesh and tray [54]. Then, the contents of the pollen traps were discharged into labelled 50-ml conical tubes and stored at -20°C . However, due to meteorological conditions and the personal schedules of the beekeepers, the sampling date, collection frequency, and operational time of the pollen trap varied from site to site. For subsequent analysis, the sampling dates were discretely pooled by month and analysed in R version 4.0.2 [55]. We used Spearman's rank-order correlation between the number of taxa per sample and the sampling length in hours to test whether the data could be treated independently of the sampling length. All samples were sent to the private company, Bioengineering Lab. Co., Ltd. (<https://www.gikenbio.com/>, consulted on 20/07/2020) for the processing of the meta-barcoding of the mixed pollen samples.

Fig 1. Selected hive locations along Tokyo bay (Japan). Each colour corresponds to the landscape type resulting from the cluster analysis of the study sites based on k-means approach. The map was drawn using *Openstreetmap France* from *mapview* in R [56].

Landscape analysis

Using remote sensing techniques, the landscape structure was investigated within a 6-km radius around each hive location (Table 1); this distance enclosed 95% of the forage area per colony [21]. With the help of Planet Labs Inc. (2020) [57], we used multi-spectral images (RGB, NIR) with 3-m pixel resolution. To fully exploit the potential of the data, the cloud cover condition was set to a maximum of 5%. Planet data are relevant for computing and mapping high-resolution terrestrial above-ground vegetation at the landscape scale [58]. For each planet image, the normalised difference vegetation index (NDVI) was computed using the red and near-infrared bands based on band rationing, which allowed for the delineation of the vegetation cover from other types of land cover [59]. Classes were created with the function *reclassified* from the *raster* package in R [60] by defining the NDVI threshold value to 0.2 to distinguish

the non-vegetation (NDVI: from -1 to 0.199) from the vegetation (NDVI: from 0.2 to 1) [61,62]. A majority filter, with a 6×6 filter kernel size, from the *whitebox* package in R [63], was applied to smooth the result and aggregate regions of high uncertainty. Landscape classifications at the site level were performed using demographic data (i.e. number of inhabitants per admin units) and landscape metrics from the *lconnect* [64] and *landscape metrics* [65] packages in R. To classify our sites along an urban-rural gradient [44,66], we retained selected data: number of inhabitants per km^2 (dpop), the integral index of connectivity (IIC) [67], the effective mesh size (MESH) [68], Shannon's evenness index (SHEI) [69], the vegetation cover proportion (veg cover), vegetation patch density (Pd) [70], and the median vegetation class NDVI (NDVI median) (Table 1).

Table 1. Mean and standard deviation of variables among the landscape classes. The units of landscape variables are given in square brackets. The standard deviation is provided after the mean.

Cluster	dpop [in./ km^2]	IIC [-]	MESH [ha]	NDVI median [-]	SHEI [-]	veg cover [%]	Pd [nbr/ km^2]
<i>rural</i>	398 (291)	0,3533 (0.0366)	0.49 (0.05)	2414.8 (904.9)	0,93 (0.05)	0.67 (0.05)	0.000044 (0.00004)
<i>suburba</i>	6475 (4839)	0.0572 (0.0342)	0.48 (0.05)	186.3 (122.9)	0.93 (0.1)	0.56 (0.08)	0.000117 (0.00003)
<i>urban</i>	10741 (2104)	0.0023 (0.0021)	0.32 (0.04)		0.51 (0.12)	0.27 (0.08)	0.000276 (0.00003)
<i>urban</i>	16532	0.0009	0.32		0.38	0.19	0.000437
<i>centre</i>	(1499)	(0.0006)	(0.03)	1.4 (1.1)	(0.14)	(0.07)	(0.00010)

We conducted principal component analysis (PCA) of the landscape dataset to visualise the differences among our study sites. The unsupervised k-means clustering method was applied to delineate the landscape category along the urban-rural gradient into k groups. Before initiating the analysis, the data were standardised using the *scale* function in R to make variables comparable. As a result, the clustering algorithm was independent of any variable unit. The number of k groups required to be defined as the first step was determined using the elbow method [71]. The k-means partitioning analysis was performed using the *k-means* function with 25 random sets [72] and the *factoextra* package in R for PCA graphical representations [73].

Molecular techniques

DNA extraction

First, pollen samples (0.5 g) were lyophilized using a lyophilizer freeze dryer VD-250R (TAITEC, Koshigaya, Saitama, Japan). After being ground at 1500 rpm for 2 min using a ShakeMaster NEO homogeniser (bms, Shinjyuku, Tokyo, Japan), DNA was extracted using the protocol of MPure Bacterial DNA Extraction Kit (MP Biomedicals, Irvine, CA, USA). DNA purification of the samples was performed using the MPure-12 Automated Nucleic Acid Purification System (MP Biomedicals, Irvine, CA, USA). Quality control of DNA extracts was conducted using Synergy H1 (BioTek, Winooski, VT, USA) and QuantiFluor dsDNA System (Promega, Madison, WI, USA).

Library preparation and DNA sequencing

Libraries were produced using a 2-step tailed polymerase chain reaction (PCR) method. The first PCR amplification was conducted using internal transcribed spacer (ITS1) primers designed by Masamura [74], coupled with MiSeq-specific adapters and Illumina index sequences. The second PCR amplification was conducted using index primers. PCR reactions

were carried out in a reaction volume of 10 μ L containing 1.0 μ L of 10 $\times$ Ex Buffer, 0.8 μ L of nucleoside triphosphate dNTPs (each at 2.5 mM), 0.5 μ L for both forward and reverse primer at a concentration of 10 μ M, 2.0 μ L of DNA template normalized at 0.5 ng/ μ L, 0.1 μ L of DNA polymerase ExTaq at 5 U/ μ L (TaKaRa, Otsu, Shiga, Japan) and 5.1 μ L of double-distilled water. The PCR profile was as follows: 2 min of denaturation at 95°C, followed by 10 cycles with 30 s of denaturation at 95°C, 30 s of annealing at 57°C, 30 s of elongation at 72°C, and a final elongation at 72°C for 5 min. The PCR products were purified using AMPure XP (Beckman Coulter, Brea, CA, USA). Library concentrations were determined using a Synergy H1 microplate reader (BioTek, Winooski, VT, USA) and a QuantiFluor dsDNA System (Promega). Library quality was evaluated using a fragment analyser (Advanced Analytical Technologies, Ankeny, IA, USA) with a dsDNA 915 Reagent Kit (Agilent, Santa Clara, CA, USA). The generated library was sequenced using MiSeq Illumina technology (Illumina, San Diego, CA, USA) through a 2 $\times$ 300 paired-end run.

Data processing

We used “FASTX Barcode Splitter” from Fastx toolkit, a short-reads pre-processing tool, to extract only the Miseq reads with sequences readings that matched the primers used exactly [75]. Next, the reads were denoised and filtered using Sickle software [76] with an overlap quality value of 20. Trimmed reads with fewer than 150 bases were discarded. The remaining reads were merged using FLASH (version 1.2.11) paired-end merge script [77] using the following conditions: fragment length after merging of 420 bases, read fragment length of 280 bases, and minimum overlap length of 10 bases. The paired readings were aligned with the reference sequences using the open-source bioinformatics pipeline Qiime 2.0 [78]. A Qiime workflow script was used for OTU creation and taxonomic assignment. A naïve Bayesian classifier was used for taxonomic assignment with the output sequence [79,80]. Following the

taxonomy classification step, taxa-abundance data and operational taxonomic unit (OTU) data were applied to the R environment [55]. First, the assignment of all OTUs below the identity threshold of 97% was discarded [46,81]. Next, the number of reads was sorted by genus and sample (i.e. site and collection date), and was then expressed as the ratio between the read count and the sum of read number per sample for each genus. Genera accounting for less than 0.05% of the total number of readings for a single sample were excluded to prevent false positives. Moreover, a sample was removed if it accounted for less than 1000 reads to limit inferences from insufficient sequencing depth [82].

Taxonomic analysis

Given the poor correlation between the actual pollen proportion per foraged plant taxa and the sequencing read proportions from one locus [36,41], the analyses undertaken were based exclusively on the incidence (i.e. presence/absence binary arrays) dataset. This process prevents the misinterpretation of DNA metabarcoding data [83]. The richness of the pollen samples (i.e. the number of distinct taxa of foraged plants) was analysed as a function of the months and landscape classes. Differences were assessed using analysis of variance (ANOVA) on richness (response), across independent landscape classes (between-subject factors), and along time (within-subjects factor). Prior to the test, the data were mathematically root-squared transformed to respect the normality (Shapiro–Wilk test) assumption for each combination of factor levels. Then, the assumptions of homogeneity of variance (Levene’s test), homogeneity of covariances (Box’s M-test), and sphericity (Mauchly’s test) of the landscape classes were verified at each level of the time variable. Finally, Tukey’s post-hoc test was applied to analyse the pairwise differences. For multivariate analysis, the pollinated plant taxonomic composition of the samples was studied across sites, sampling periods, and landscape classes using the Jaccard dissimilarity metric from the *vegan* package in R [84]. This asymmetric distance

coefficient addresses the problem of double zero, which is essential when studying data on community composition along a gradient. Differences in plant composition between sampling periods and landscape classes were investigated by permutation-based multivariate analysis ($n = 999$) of variance using the *adonis* function [85]. If the PERMANOVA results were significant, a post-hoc multilevel pairwise analysis with Bonferroni correction was performed using the *pairwiseAdonis* package in R [86]. The dissimilarities in the structures of pollinated plant communities were displayed using non-metric multidimensional scaling (NMDS) with 999 permutations. All graphics were generated using the *ggplot2* package in R [87].

Indicator species and trait-based analysis

Similarity percentage (SIMPER) analysis was applied to identify how the taxonomic composition differed from the environmental conditions (landscape type) and changes (season). This step allowed us to identify the sampled taxa that contributed significantly to the dissimilarities among the months or landscapes. Finally, to analyse the foraging preferences of honeybees concerning plant traits, each identified taxa in the samples was characterised by its nature, including herbaceous (no woody stems above ground) or woody taxon (tree, shrub, liana), and its native status, including native, alien, or cultivar taxa. The plant trait database was built using information from Ylist [88] and ©Species2000 [89] for the Japanese plant dataset. To determine if the proportion between the different traits varied with the seasons and landscape types, the G.-test of independence for contingency table was performed using the *RVAideMemoire* package in R [90]. The G-test is based on the log likelihood ratio and tests whether the relative proportions of one categorical variable (i.e. plant nature or nativity status) are independent of the second categorical variable (i.e. season or landscape). Next, post-hoc pairwise comparisons were conducted between pairs of proportions using the Bonferroni correction of the p-values [91].

Results

Landscape classification

The method of differentiating vegetation from impervious surfaces using NDVI provided convincing results after crosschecking, even in the complex environment of an urban matrix. The two first dimensions of the landscape PCA from landscape variables of our sampling sites described a high percentage of the variance (axis 1 = 81.6% and axis 2 = 11.0%; Fig 2). According to the elbow method (Fig S2) of the k-means partitioning, we classified our sampling sites into four landscape classes according to a rural-urban gradient: rural ($n = 3$), suburban ($n = 3$), urban ($n = 5$), and urban centre ($n = 6$). The urbanised locations were driven by a much higher demographic density compared to the other landscapes (Fig S3). Moreover, following the decrease in the proportion of vegetation along the rural-urban gradient, it can be assumed that the higher patch density in the cities was induced by the presence of many smaller plots, such as private garden patches. In contrast, the rural sites demonstrated a higher connectivity between the patches. Finally, suburban landscapes occupied an intermediary position between rural and metropolitan areas.

Fig 2. Landscape cluster analysis of selected locations based on k-means approach. The axes represent the first two principal components of the PCA analysis. The dot shapes and colours represent the resulted landscape classes: circle/red = rural landscape; square/blue = suburban landscape; triangle/green = urban landscape; reversed triangle/purple = urban centre landscape.

Taxonomic analysis

Illumina sequencing generated a total of 8,179,602 paired-end raw reads for the 143 pollen samples from the 17 sites throughout the 7 months of pollen sampling. After assembling and

filtering, 6,799,314 reads (83.2%) were obtained for analysis, with a mean count of $47,548 \pm 27,464$ (SD) reads per sample. By taxonomic assignment of the meta-barcoding dataset, we identified 307 plant flower taxa from 74 families and 187 genera. Prior to the analysis, the richness was not correlated with the sampling length, showing a very weak relationship (r_s [143] = -0.17 , $p < 0.05$), allowing us to consider the statistical independence of all our pollen samples.

Plant richness ranged between 3 and 42 pollinated plant taxa per sample, with an average of 12 ($SD = 6.2$). Mixed ANOVA was not able to detect a significant interaction between landscape type and collection time on the squared root taxa richness ($F = 0.95$; $p = 0.52$). The follow-up tests determined that the landscape type implied no statistically significant differences ($F = 2.00$; $p = 0.12$), but the pollen richness (squared root plant taxa richness) varied significantly across the months ($F = 5.22$; $p < 0.001$). The early spring (March and April) showed significantly higher pollen diversity (Tukey's test: $p < 0.05$) than the late growth season (July, August, and September) (Fig 3).

Fig 3. Boxplots of squared root taxa richness by sampling period. Letters and boxplot colours indicate significant differences according to Tukey's test ($p < 0.05$).

NMDS displayed high variability in the composition of foraged pollen across the months and seasons (Fig 4). The greatest discontinuity separated spring (March, April, and May) from autumn (September). Concerning the floral composition foraged by the honeybees, May and August served as transition months to subsequent seasons. The permutation tests revealed that the month period ($F = 6.87$; $R^2 = 0.23$; $p < 0.001$), site ($F = 1.27$; $R^2 = 0.1$; $p < 0.01$), and landscape class ($F = 2.01$; $R^2 = 0.03$; $p < 0.001$) were significant explanatory variables of pollen composition in the samples; however, the sampling period was attributed a larger proportion of the variance. From pairwise comparisons (i.e. letters from Fig 4), the urbanised sites hosted

similar plant communities. Moreover, the structure of the plant communities varied significantly over the months until late summer and early autumn (i.e. August and September), when the floral composition harboured similar foraged plant communities.

Fig 4. Non-metric multidimensional scaling (NMDS) of plant communities from the pollen incidence data. Dot shapes correspond to the landscape classes of pollen samples. Dot colours correspond to the sampling months, and the seasons are displayed by 80% prediction confidence ellipses. Letters indicate significant differences ($p < 0.05$) according to the pairwise post-hoc comparisons, with Bonferroni correction, of the foraged plant communities among the landscape gradient and the sampling period.

Indicator species and trait-based analysis

The plant traits foraged by the honeybees varied significantly according to the nature ($G = 99.0$, $p < 0.001$) and native status ($G = 69.1$, $p < 0.001$) over the study months (Fig 5B), while only the plant strata showed significant differences ($G = 10.7$, $p < 0.05$) according to the landscape classes (Fig 5A).

Effect of landscape

Significantly more herbaceous plant taxa structured the rural landscape (Fig 5A), even if plant composition traits were independent of landscape type. A total of 35 plant taxa were shared among all the landscapes over the sampling months, while 151 plant taxa were exclusively related to specific landscapes, corresponding to 27%, 22%, 24%, and 26% of the total plant taxa for rural, suburban, urban, and urban centre landscapes, respectively (Fig 5C). Plant families, such as Fabaceae, Rosaceae, Brassicaceae, Asteraceae, Plantaginaceae, and Onagraceae, were the most frequent taxonomic families encountered in all samples. However, their proportions varied according to the landscape (Fig S4). In urban and urban centre areas,

leguminous plants prevailed in rural and suburban landscapes. However, the suburban landscape showed a higher frequency for the Brassicaceae, Ranunculaceae, and Rosaceae plant families. Surprisingly, anemophilous plants of the Poaceae family were more frequently foraged in the urbanised landscape than in the countryside (Fig S4).

Effect of sampling period

The proportion of foraged woody taxa decreased significantly over the seasons ($G = 87.5$, $p < 0.001$), with a peak of 46% in April and a low of 10% in September. Over the sampling months, honeybees foraged mainly on alien plant taxa (Fig 5B). Cultivar taxa were more foraged in spring than in the other two seasons ($G = 32.9$, $p < 0.001$). The most visited plant genera in March and April included *Prunus*, *Helleborus*, *Brassica*, and *Acer* taxa (Fig 5D). In contrast, four of the 11 most frequent genera emerged in late spring in May. Between April and June, a noticeable phenological turnover in the pollen composition (Fig 5D) was observed, with May serving as a transition bridge. This has already been highlighted by the discontinuities in the NMDS ordination (Fig 4). Following this shift, the genus *Trifolium* spp. was highly dominant in the June and July samples. In addition, the herbaceous genera *Plantago* and *Oenothera* spp. were also found in large proportions in combination with the woody genera *Mallotus* and *Hydrangea* spp. In August, the taxa from the genus *Oenothera* were the most represented with *Trifolium*, despite a reduction in its occurrence. A shift in pollen composition trends, with a reduction in highly proportional occurrence genera, was observed in August and September. In other words, plants detected in August and September were more distributed between the genera. Only the genus *Allium* showed an increase from August to September. Finally, *Trifolium* spp., *Rosa* spp., and *Allium* spp. were the only genera that were observed throughout the study period.

Fig 5. Proportional occurrences of the different plant traits, Venn diagram, and heatmap according to the landscape classes and the sampling months. (A) Proportional occurrences of the different plant traits (plant nature and nativity status) along the different landscape classes. Letters on top of the bar cluster homogeneous landscapes according to the significant results of post-hoc pairwise comparison with Bonferroni correction ($p < 0.05$). (B) Proportional occurrence of the different plant traits (plant nature and nativity status) across the sampling period. Letters on top of the bar cluster denote homogeneous sampling periods according to the significant results of post-hoc pairwise comparison with Bonferroni correction ($p < 0.05$). (C) Venn diagram indicating the overlap of plant taxa according to the different landscape classes. "N" corresponds to the number of samples per landscape. (D) Heatmap representing 11 most frequent genera (N = 62 taxa; 47% of the meta-barcoding dataset) ordered in descending order by their prevalence in all samples according to their monthly proportional occurrence. The proportional occurrence corresponds to the ratio of the number of observations for a genus per month to the number of samples for the specific month. The 11 genera were chosen according to the condition that their observation frequency is superior to 10% of the total occurrence of at least one season.

Discussion

This study revealed interesting patterns of honeybee foraging habits along the urban-rural gradient throughout the 2019 active season among 17 different sites. In the present study, we used an unconventional approach, namely k-means clustering from landscape metrics, to categorise our sampling sites into four landscape classes. These landscape variables allowed the estimation of the effects of some ecological processes at the landscape level (i.e. foraging and plant dispersal) in assessing the diversity, connectivity, and aggregation of the patches [92,93]. Despite the convincing results of the grouping method, several reservations are worth

mentioning. First, the selected foraging radius of 6 km accentuated the spatial autocorrelation issues on the landscape variables by increasing the foraging area overlaps among the sites [94]. This was not tested in the present study because we considered that there was no spatial influence on foraging preferences between each sampling colony due to the high variations in their foraging behaviour [95,96], even at the local scale for colonies of the same apiary [97]. Indeed, the foraging behaviour of honeybee colonies is mainly driven by: (i) the temporal colony needs as a function of its strength; (ii) the high density of pollen and nectar resources available near the colony; (iii) the capacity to make hourly colonial decisions for the most profitable flower patches; and (iv) the ability to tightly modulate its pollen reserves to buffer the colony from seasonal pollen breaks (e.g. long rainy periods, rarefaction of floral resources at the end of summer, etc.) [21]. Second, the 3-image resolution from satellites may result in some limitations, especially in complex landscapes, such as urban matrices. Therefore, this spatial resolution issue could be counterbalanced with the use of specific cameras, such as red-green-blue (RGB) or multispectral cameras, mounted on unmanned aerial vehicles to characterise floral identification and surfaces [98,99]. Nonetheless, this research domain is still in its infancy, particularly in the data processing of imagery classification by deep learning [100]. However, despite these spatial limitations, the approach led to satisfactory classification, which paves the way for further investigations.

The use of DNA meta-barcoding allowed for the identification of the pollen composition of the samples. This method yields to the identification of 307 taxa, which is higher than previous studies owing to a greater sampling effort among sites and seasons [46,47,52]. This reflects a great diversity of plant resources and highlights foraging patterns, regardless of the landscape. Through our molecular analysis, we gauged potential biases that may arise from PCR amplification, sequencing errors, or arbitrary choices in the bioinformatics pipeline [101,102]. Moreover, recent studies aimed to compare pollen quantification from complex samples using

a metagenomics approach against light microscopy. They recommended using a multi-locus meta-barcoding strategy in combining plastid loci, such as the *matK* or *trnL* markers with ITS2, to improve the correlation between the number of reads and classic palynology identification [103,104]. However, in the present study, we used one pair of ITS1 primers [74], which led to the consideration of all the generated datasets on their incidences due to the poor correlations between the number of sequence reads and visually identified pollen grains [36,41,105]. Furthermore, the robustness of the correlation depended primarily on the pollen counting methodology, DNA marker selection, plant taxonomy dataset, and number of PCR cycles [52,104]. Despite the need for further studies and improvements, this clearly provides a strong foundation for the study plant-pollinator interactions and the floral preferences of pollinating insects.

The landscape type did not influence forage plant richness, as reported in previous studies [49,51]. However, considering the spatial variation of taxa composition among the sites of our area of interest (i.e. beta diversity) [106], our results showed a spatial structure of foraged plant communities in countryside, suburban, and urban environments (i.e. by merging the urban and urban centre areas) (Fig 4). It is likely that honeybee colonies modify their foraging preferences due to the high prevalence of unattractive ornamental flowers in urban landscapes [14]. The urban matrix also offers smaller spread patches and less dense floral resources, which contribute to the foraging change of workers [47,97]. This shift in the prospected flora is also coupled to fulfil the nutritional demand with a diverse and complementary floral diet [107,108]. Therefore, it shows the importance of taking beta diversity and not only the local richness into account to understand the community structure of foraged plants throughout space and time scales to protect local biodiversity and support conservation planning [53,109]. Despite the landscape structure of the foraged plant community, the trait-based analysis revealed no significant preference for honeybee colonies, except for herbaceous plants in rural landscapes (Fig 5A).

In all the samples, 35 plant taxa were shared throughout all the landscapes studied, corresponding to 45% of all read counts. The top three plant families were Fabaceae, Rosaceae, and Brassicaceae in the samples (Fig S4). The presence of grasses (Poaceae) may seem surprising among the most frequent families in the samples (Fig S4), given that these plants are considered unsuitable for *A. mellifera* resource needs [110]. In view of their dominance, it seems unlikely that anemophilic pollen is the result of contamination by pollen blown from flowers onto the body of the bees, such as the rice paddy field in Japan [111]. Other recent studies have shown that pollinators (i.e. bees and syrphids), particularly *Apis* bees, interact with wind-pollinated plant species for their nutrient or nesting requirements [112]. This means that these grass species could be implemented in rural and urban management recommendations.

In contrast to the minor variations in landscape diversity, we observed a strong seasonal effect on plant richness (Fig 3), the foraged plant community, and traits (Fig 4 and 5B). We observed higher plant richness and foraged woody taxa in spring than in the other two seasons (Fig 3 and 5B), as reported previously [82,113,114]. The genera *Prunus* spp. and *Acer* spp. (Fig 5D) dominate the foraged woody stratum during this season, as these taxa constitute an important source of pollen for the early season [82,103,114,115]. Indeed, bee breads with high proportions of both genera were positively correlated with a higher protein content [107]. Particularly, the complexity and high range of foraged plants is known to be beneficial to the “nutritional value” of these bee bread stocks, and thus to honeybee immunity [28]. After the spring period, the proportion of foraged herbaceous strata gradually substituted woody taxa to reach approximately 90% of herbaceous foraged taxa in September, in agreement with previous studies [82,114]. This growing herbaceous stratum is mainly dominated by *Trifolium* spp. and *Plantago* spp. which, because of their long flowering period or floral consistency, explains the lower richness of taxa foraged in summer and autumn [114,116,117]. Moreover, clover species (i.e. *Trifolium* spp.) are highly ubiquitous in grasslands, such as meadows for rural areas or

parks and gardens for urban areas [114,118] and may contribute to the concentration effect of amino acid content [107]. Concerning the temporal dynamics of the native traits (Fig 5B), the observations point to a decrease in the frequency of cultivars in favour of exotic taxa, which contradicts previous studies [119,120]. The highly anthropized and fragmented environments of the Tokyo region and its surroundings could explain the predominance of non-native floral species. However, this statement should be mitigated and requires further investigation, such as a complete plant inventory of the study site. Finally, we observed a transitional change from August by a rupture of the dominant flower prevalence, which may correspond to the seasonal dearth of floral resources (Fig 4 and 5B) [114,121]. From the perspective of the honeybees, they mitigate this effect by increasing their foraging range, requiring extra effort for sometimes worthless rewards [121]. Therefore, in an urban greening context, it would be relevant to ensure the diversity and abundance of floral resources near apiaries by targeting seasonal gaps in pollen collection [11]. Finally, the study of other co-variables, such as brood monitoring or estimating the pollen collection/reserve of each colony, could be used to compare the conditions of each sampled colony and improve our understanding of the foraging patterns of the colonies [122].

The availability of floral resources needs to be sufficient to host both domesticated honeybees and local wild pollinators. The percentage of impervious surfaces plays a major role in the pollination of biodiversity [123]. Therefore, populations of managed honeybees must be regulated to ensure wild pollinator populations are not adversely affected [18,124]. In contrast, green areas must be managed and well distributed to meet the demands of the pollinator community [125]. In summary, well-managed cities could play an active role in the preservation of insect pollinators, and thus provide hotspots for pollination services [126]. However, to do so, decision-makers will need to focus on regulating the introduction of honeybees (selection of native bee species, colony density, and control of pathogens and parasites) and on the

440 availability of resources (proportion of impermeable surface area, melliferous plant species,
441 and landscape diversity).

442 **Acknowledgments**

The sequencing data have been deposited with links to BioProject accession number PRJNA74415 in the NCBI BioProject database (<https://www.ncbi.nlm.nih.gov/bioproject/>).

This work was supported by JSPS KAKENHI (grant number 18KK0121) and the Yamada Research Grant. We thank Dr. Junko Kawai at Chiba University for his useful advice on DNA meta barcoding. The pollen was provided by following beekeepers in Japan: Ryuta Aikyo, Sayuri Aoki, Shinta Hegizono, Honami Kawabe, Tamaki Maruhashi, Haruko Nagae, Emi Nobunaga, Nobuyuki Okada, Takumi Sakou, Yuichi Shibasaki, Yosuke Toyomasu, Misa Uchino, and Tomoyuki Yuzurihara.

References

1. Concepción ED, Moretti M, Altermatt F, Nobis MP, Obrist MK. Impacts of urbanisation on biodiversity: The role of species mobility, degree of specialisation and spatial scale. *Oikos*. 2015;124: 1571–1582. doi:10.1111/oik.02166
2. Sánchez-Bayo F, Wyckhuys KAG. Worldwide decline of the entomofauna: A review of its drivers. *Biol Conserv*. 2019;232: 8–27. doi:10.1016/j.biocon.2019.01.020
3. Potts SG, Biesmeijer JC, Kremen C, Neumann P, Schweiger O, Kunin WE. Global pollinator declines: Trends, impacts and drivers. *Trends Ecol Evol*. 2010;25: 345–353. doi:10.1016/j.tree.2010.01.007
4. Ahrné K, Bengtsson J, Elmqvist T. Bumble Bees (*Bombus* spp) along a Gradient of Increasing Urbanization. *PLoS One*. 2009;4: : e5574. doi:doi:10.1371/journal.pone.0005574
5. Fortel L, Henry M, Guilbaud L, Guirao AL, Kuhlmann M, Mouret H, et al. Decreasing abundance, increasing diversity and changing structure of the wild bee community (Hymenoptera: Anthophila) along an urbanization gradient. *PLoS One*. 2014;9: : e104679. doi:10.1371/journal.pone.0104679
6. Harrison T, Winfree R. Urban drivers of plant-pollinator interactions. *Funct Ecol*. 2015;29: 879–888. doi:10.1111/1365-2435.12486

7. Wenzel A, Grass I, Belavadi V V, Tschardt T. How urbanization is driving pollinator diversity and pollination – A systematic review. *Biol Conserv.* 2020;241: 108321. doi:<https://doi.org/10.1016/j.biocon.2019.108321>
8. Hall DM, Camilo GR, Tonietto RK, Smith DH, Ollerton J, Ahn K, et al. The city as a refuge for insect pollinators. *Conserv Biol.* 2016. doi:10.1111/cobi.12840
9. Theodorou P, Radzevičiūtė R, Lentendu G, Kahnt B, Husemann M, Bleidorn C, et al. Urban areas as hotspots for bees and pollination but not a panacea for all insects. *Nat Commun.* 2020;11: 1–13. doi:10.1038/s41467-020-14496-6
10. Muratet A, Fontaine B. Contrasting impacts of pesticides on butterflies and bumblebees in private gardens in France. *Biol Conserv.* 2015;182: 148–154. doi:10.1016/j.biocon.2014.11.045
11. Baldock KCR. Opportunities and threats for pollinator conservation in global towns and cities. *Curr Opin Insect Sci.* 2020;38: 63–71. doi:10.1016/j.cois.2020.01.006
12. Pardee GL, Philpott SM. Native plants are the bee’s knees: local and landscape predictors of bee richness and abundance in backyard gardens. *Urban Ecosyst.* 2014;17: 641–659. doi:10.1007/s11252-014-0349-0
13. Baldock KCR, Goddard MA, Hicks DM, Kunin E, Mitschunas N, Osgathorpe LM, et al. Where is the UK ’s pollinator biodiversity ? The importance of urban areas for flower- visiting insects. *Proc R Soc Biol Sci.* 2015;282: 20142849. doi:10.1098
14. Garbuzov M, Samuelson EEW, Ratnieks FLW. Survey of insect visitation of ornamental flowers in Southover Grange garden, Lewes, UK. *Insect Sci.* 2015;22: 700–705. doi:10.1111/1744-7917.12162
15. Garbuzov M, Ratnieks FLW. Quantifying variation among garden plants in attractiveness to bees and other flower-visiting insects. Thompson K, editor. *Funct Ecol.* 2014;28: 364–374. doi:10.1111/1365-2435.12178
16. Tanaka K, Nozaki A, Nakadai H, Shiwa Y, Shimizu-Kadota M. Using pollen DNA metabarcoding to profile nectar sources of urban beekeeping in Kōtō-ku, Tokyo. *BMC Res Notes.* 2020;13: 1–6. doi:10.1186/s13104-020-05361-2
17. Cho Y, Lee D. ‘Love honey, hate honey bees’: reviving biophilia of elementary school students through environmental education program. *Environ Educ Res.* 2018;24: 445–460. doi:10.1080/13504622.2017.1279277
18. Geslin B, Gauzens B, Baude M, Dajoz I, Fontaine C, Henry M, et al. Massively Introduced Managed Species and Their Consequences for Plant–Pollinator Interactions. *Adv Ecol Res.* 2017;57: 1–53. doi:10.1016/j.talanta.2011.11.033

19. Ropars L, Dajoz I, Fontaine C, Muratet A, Geslin B. Wild pollinator activity negatively related to honey bee colony densities in urban context. Blenau W, editor. PLoS One. 2019;14: e0222316–e0222316. doi:10.1371/journal.pone.0222316
20. von Frisch K. Tanzsprache und Orientierung der Bienen. Tanzsprache und Orientierung der Bienen. Berlin, Heidelberg: Springer Berlin Heidelberg; 1965. doi:10.1007/978-3-642-94916-6
21. Seeley TD. The wisdom of the hive. illustrate. Harvard University Press; 1995.
22. Alaux C, Allier F, Decourtye A, Odoux J-F, Tamic T, Chabirand M, et al. A ‘Landscape physiology’ approach for assessing bee health highlights the benefits of floral landscape enrichment and semi-natural habitats. Sci Rep. 2017;7: 40568. doi:10.1038/srep40568
23. Haydak MH. Honey Bee Nutrition. Annu Rev Entomol. 1970;15: 143–156. doi:10.1146/annurev.en.15.010170.001043
24. Avni D, Hendriksma HP, Dag A, Uni Z, Shafir S. Nutritional aspects of honey bee-collected pollen and constraints on colony development in the eastern Mediterranean. J Insect Physiol. 2014;69: 65–73. doi:10.1016/j.jinsphys.2014.07.001
25. Avni D, Dag A, Shafir S. Pollen sources for honeybees in Israel: Source, periods of shortage, and influence on population growth. Isr J Plant Sci. 2009;57: 263–275. doi:10.1560/IJPS.57.3.263
26. Wilde J, Bratkowski J, Siuda M. Pollen collection by 3 subspecies of honeybee *Apis mellifera* L. Acta Biol Univ Daugavp. 2003;3: 101–106.
27. Di Pasquale G, Salignon M, Le Conte Y, Belzunces LP, Decourtye A, Kretzschmar A, et al. Influence of pollen nutrition on honey bee health: Do pollen quality and diversity matter? Zeil J, editor. PLoS One. 2013;8: e72016–e72016. doi:10.1371/journal.pone.0072016
28. Alaux C, Ducloz F, Crauser D, Le Conte Y. Diet effects on honeybee immunocompetence. Biol Lett. 2010;6: 562–565. doi:10.1098/rsbl.2009.0986
29. Wang H, Zhang SW, Zeng ZJ, Yan WY. Nutrition affects longevity and gene expression in honey bee (*Apis mellifera*) workers. Apidologie. 2014;45: 618–625. doi:10.1007/s13592-014-0276-3
30. Cornman RS, Otto CR V, Iwanowicz D, Pettis JS. Taxonomic characterization of honey bee (*Apis mellifera*) pollen foraging based on non-overlapping paired-end sequencing of nuclear ribosomal loci. Jan E, editor. PLoS One. 2015;10: e0145365–e0145365. doi:10.1371/journal.pone.0145365

31. Ohe W von der, Persano Oddo L, Piana ML, Morlot M, Martin P. Harmonized methods of melissopalynology. *Apidologie*. 2004;35: 18–25. doi:10.1051/apido
32. De França Alves R, De Assis Ribeiro DosSantos F. Plant sources for bee pollen load production in Sergipe, northeast Brazil. *Palynology*. 2014;38: 90–100. doi:10.1080/01916122.2013.846280
33. Bell KL, De Vere N, Keller A, Richardson RT, Gous A, Burgess KS, et al. Pollen DNA barcoding: Current applications and future prospects. Van der Bank M, editor. *Genome*. 2016;59: 629–640. doi:10.1139/gen-2015-0200
34. Kaškonienė V, Venskutonis PR. Floral Markers in Honey of Various Botanical and Geographic Origins: A Review. *Compr Rev Food Sci Food Saf*. 2010;9: 620–634. doi:10.1111/j.1541-4337.2010.00130.x
35. Rahl M. Microscopic identification and purity determination of pollen grains. *Methods Mol Med*. 2008;138: 263–269. doi:10.1007/978-1-59745-366-0_22
36. Richardson RT, Lin C-H, Sponsler DB, Quijia JO, Goodell K, Johnson RM. Application of ITS2 Metabarcoding to Determine the Provenance of Pollen Collected by Honey Bees in an Agroecosystem. *Appl Plant Sci*. 2015;3: 1–6. doi:10.3732/apps.1400066
37. Bell KL, Fowler J, Burgess KS, Dobbs EK, Gruenewald D, Lawley B, et al. Applying Pollen DNA Metabarcoding to the Study of Plant–Pollinator Interactions. *Appl Plant Sci*. 2017;5: 1600124. doi:10.3732/apps.1600124
38. Pornon A, Escaravage N, Burrus M, Holota H, Khimoun A, Mariette J, et al. Using metabarcoding to reveal and quantify plant-pollinator interactions. *Sci Rep*. 2016;6: 27282. doi:10.1038/srep27282
39. Sickel W, Ankenbrand MJ, Grimmer G, Holzschuh A, Härtel S, Lanzen J. Increased efficiency in identifying mixed pollen samples by meta - barcoding with a dual - indexing approach. *BMC Ecol*. 2015;15: 1–9. doi:10.1186/s12898-015-0051-y
40. Hawkins J, De Vere N, Griffith A, Ford CR, Allainguillaume J, Hegarty MJ, et al. Using DNA metabarcoding to identify the floral composition of honey: A new tool for investigating honey bee foraging preferences. Labra M, editor. *PLoS One*. 2015;10: e0134735–e0134735. doi:10.1371/journal.pone.0134735
41. Kraaijeveld K, de Weger LA, Ventayol García M, Buermans H, Frank J, Hiemstra PS, et al. Efficient and sensitive identification and quantification of airborne pollen using next-generation DNA sequencing. *Mol Ecol Resour*. 2014;15: 8–16. doi:10.1111/1755-0998.12288

42. Galimberti A, De Mattia F, Bruni I, Scaccabarozzi D, Sandionigi A, Barbuto M, et al. A DNA Barcoding Approach to Characterize Pollen Collected by Honeybees. *PLoS One*. 2014;9: e109363. Available: <https://doi.org/10.1371/journal.pone.0109363>
43. Hennig EI, Ghazoul J. Plant-pollinator interactions within the urban environment. *Perspect Plant Ecol Evol Syst*. 2011;13: 137–150. doi:10.1016/j.ppees.2011.03.003
44. Bastin L, Thomas CD. The distribution of plant species in urban vegetation fragments. *Landsc Ecol*. 1999;14: 493–507.
45. Garbuzov M, Schürch R, Ratnieks FLW. Eating locally: dance decoding demonstrates that urban honey bees in Brighton, UK, forage mainly in the surrounding urban area. *Urban Ecosyst*. 2015;18: 411–418. doi:10.1007/s11252-014-0403-y
46. Danner N, Keller A, Ha S, Steffan-dewenter I. Honey bee foraging ecology : Season but not landscape diversity shapes the amount and diversity of collected pollen. *PLoS One*. 2017;12: 9–12. doi:10.1371/journal.pone.0183716
47. Lucek K, Galli A, Gurten S, Hohmann N, Maccagni A, Patsiou T, et al. Metabarcoding of honey to assess differences in plant-pollinator interactions between urban and non-urban sites. *Apidologie*. 2019;50: 317–329. doi:10.1007/s13592-019-00646-3
48. Butz Huryn VM. Ecological Impacts of Introduced Honey Bees. *Q Rev Biol*. 1997;72: 275–297. doi:10.1086/419860
49. Steffan-Dewenter I, Kuhn A. Honeybee foraging in differentially structured landscapes. *Proc R Soc B Biol Sci*. 2003;270: 569–575. doi:10.1098/rspb.2002.2292
50. Danner N, Molitor AM, Schiele S, Härtel S, Steffan-Dewenter I. Season and landscape composition affect pollen foraging distances and habitat use of Honey bees. *Ecol Appl*. 2016;26: 1920–1929. doi:10.1890/15-1840.1
51. Danner N, Keller A, Härtel S, Steffan-Dewenter I. Honey bee foraging ecology: Season but not landscape diversity shapes the amount and diversity of collected pollen. *PLoS One*. 2017;12: 9–12. doi:10.1371/journal.pone.0183716
52. Sponsler DB, Shump D, Richardson RT, Grozinger CM. Characterizing the floral resources of a North American metropolis using a honey bee foraging assay. *Ecosphere*. 2020;11. doi:10.1002/ecs2.3102
53. Richardson RT, Eaton TD, Lin CH, Cherry G, Johnson RM, Sponsler DB. Application of plant metabarcoding to identify diverse honeybee pollen forage along an urban–agricultural gradient. *Mol Ecol*. 2021;30: 310–323. doi:10.1111/mec.15704
54. Mahmood R, Asad S, Sarwar G, Ahmad W, Qadir ZA, Balouch A, et al. Pollen

collection and pollen foraging behaviour of honeybees (*apis mellifera*) during different time intervals from *brassica campestris* L. *Pak J Zool.* 2017;49: 733–736. doi:10.17582/journal.pjz/2017.49.2.sc4

55. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria; 2020. Available: <https://www.r-project.org/>
56. Appelhans T, Detsch F, Reudenbach C, Woellauer S. mapview: Interactive Viewing of Spatial Data in R. 2019. Available: <https://cran.r-project.org/package=mapview>
57. Team P. Planet Application Program Interface: In Space for Life on Earth. 2020.
58. Miller GJ, Morris JT, Wang C. Estimating aboveground biomass and its spatial distribution in coastal wetlands utilizing planet multispectral imagery. *Remote Sens.* 2019;11: 2020. doi:10.3390/rs11172020
59. Xue J, Su B. Significant remote sensing vegetation indices: A review of developments and applications. *J Sensors.* 2017;2017: 1–17. doi:10.1155/2017/1353691
60. Hijmans RJ, van Etten J, Sumner M, Cheng J, Baston D, Bevan A, et al. raster: Geographic Data Analysis and Modeling. 2020.
61. Taufik A, Syed Ahmad SS, Azmi EF. Classification of landsat 8 satellite data using unsupervised methods. *J Telecommun Electron Comput Eng.* 2016;8: 37–40.
62. Hashim H, Abd Latif Z, Adnan NA. Urban vegetation classification with NDVI threshold value method with very high resolution (VHR) Pleiades imagery. *Int Arch Photogramm Remote Sens Spat Inf Sci - ISPRS Arch.* 2019;42: 237–240. doi:10.5194/isprs-archives-XLII-4-W16-237-2019
63. Lindsay JB. Whitebox GAT: A case study in geomorphometric analysis. *Comput Geosci.* 2016;95: 75–84. doi:10.1016/j.cageo.2016.07.003
64. Mestre F, Silva B. lconnect: Simple Tools to Compute Landscape Connectivity Metrics. 2019. Available: <https://github.com/FMestre1/lconnect/issues>
65. Hesselbarth MHK, Sciaini M, With KA, Wiegand K, Nowosad J. Landscapemetrics: an open-source R tool to calculate landscape metrics. *Ecography (Cop).* 2019;42: 1648–1657.
66. Hadley AS, Betts MG. The effects of landscape fragmentation on pollination dynamics: absence of evidence not evidence of absence. *Biol Rev.* 2012;87: 526–544. doi:10.1111/j.1469-185X.2011.00205.x

67. Saura S, Pascual-Hortal L. A new habitat availability index to integrate connectivity in landscape conservation planning: Comparison with existing indices and application to a case study. *Landsc Urban Plan.* 2007;83: 91–103. doi:10.1016/j.landurbplan.2007.03.005
68. Spanowicz AG, Jaeger JAG. Measuring landscape connectivity: On the importance of within-patch connectivity. *Landsc Ecol.* 2019;34: 2261–2278. doi:10.1007/s10980-019-00881-0
69. Shannon CE. A Mathematical Theory of Communication. *Bell Syst Tech J.* 1948;27: 623–656. doi:<https://doi.org/10.1002/j.1538-7305.1948.tb00917.x>
70. Threlfall CG, Walker K, Williams NSG, Hahs AK, Mata L, Stork N, et al. The conservation value of urban green space habitats for Australian native bee communities. *Biol Conserv.* 2015;187: 240–248. doi:10.1016/j.biocon.2015.05.003
71. Kodinariya TM, Makwana PR. Review on determining number of Cluster in K-Means Clustering. *Int J Adv Res Comput Sci Manag Stud.* 2013;1: 90–95.
72. Strickland JS. Predictive Analytics using R. Lulu Inc; 2014.
73. Kassambara A, Fabian M. factoextra: Extract and Visualize the Results of Multivariate Data Analyses. 2020. Available: <https://cran.r-project.org/package=factoextra>
74. MASAMURA N, KIKUCHI R, NAGATOMI Y. Developments of an Identification Method for Foreign Substances of Plant Origin Using ITS 1 Region. *BUNSEKI KAGAKU.* 2014;63: 245–253. doi:10.2116/bunsekikagaku.63.245
75. Hannon GJ. FASTX-Toolkit. In: 2010 [Internet]. Available: http://hannonlab.cshl.edu/fastx_toolkit.
76. Joshi N, Fass J. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files. 2011. Available: <https://github.com/najoshi/sickle>
77. Magoč T, Salzberg SL. FLASH: Fast length adjustment of short reads to improve genome assemblies. *Bioinformatics.* 2011;27: 2957–2963. doi:10.1093/bioinformatics/btr507
78. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol.* 2019;37: 852–857. doi:10.1038/s41587-019-0209-9
79. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Appl Environ Microbiol.* 2007;73: 5261 LP – 5267. doi:10.1128/AEM.00062-07

80. Keller A, Danner N, Grimmer G, Ankenbrand M, von der Ohe K, von der Ohe W, et al. Evaluating multiplexed next-generation sequencing as a method in palynology for mixed pollen samples. *Plant Biol (Stuttg)*. 2015;17: 558–566. doi:10.1111/plb.12251
81. Smart MD, Cornman RS, Iwanowicz DD, Pettis JS, Spivak MS, Otto CR V. Pollinator Ecology and Management A Comparison of Honey Bee-Collected Pollen From Working Agricultural Lands Using Light Microscopy and ITS Metabarcoding. 2017; 1–12. doi:10.1093/ee/nvw159
82. Sponsler DB, Shump D, Richardson RT, Grozinger CM. Characterizing the floral resources of a North American metropolis using a honey bee foraging assay. *Ecosphere*. 2020;11. doi:10.1002/ecs2.3102
83. Lamb PD, Hunter E, Pinnegar JK, Creer S, Davies RG, Taylor MI. How quantitative is metabarcoding: A meta-analytical approach. *Mol Ecol*. 2019;28: 420–430. doi:10.1111/mec.14920
84. Oksanen J, Blanchet Michael FGF, Kindt R, Legendre P, McGlinn D, Minchin PR, et al. *vegan: Community Ecology Package*. 2019. Available: <https://cran.r-project.org/package=vegan>
85. Anderson MJ. A new method for non-parametric multivariate analysis of variance. *Austral Ecol*. 2001;26: 32–46. doi:10.1111/j.1442-9993.2001.01070.pp.x
86. Martinez Arbizu P. pairwiseAdonis: Pairwise multilevel comparison using adonis. 2020.
87. Wickham H. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York; 2016. Available: <https://ggplot2.tidyverse.org>
88. Yonekura K, Kajita T. Japanese botanical name-scientific name index YList. 2007.
89. Roskov Y, Ower G, Orrell T, Nicolson D, Bailly N, Kirk PM, et al. Species 2000 & ITIS Catalogue of Life, 2019 Annual Checklist. In: Species 2000: Naturalis, Leiden, the Netherlands [Internet]. 2019. Available: <http://www.catalogueoflife.org/annual-checklist/2019>
90. Hervé M. RVAideMemoire: Testing and Plotting Procedures for Biostatistics. 2020. Available: <https://cran.r-project.org/package=RVAideMemoire>
91. MacDonald PL, Gardner RC. Type I error rate comparisons of post hoc procedures for $I \times J$ chi-square tables. *Educ Psychol Meas*. 2000;60: 735–754. doi:10.1177/00131640021970871
92. Baguette M, Blanchet S, Legrand D, Stevens VM, Turlure C. Individual dispersal,

- landscape connectivity and ecological networks. *Biol Rev.* 2012;88: 310–326. doi:10.1111/brv.12000
93. Doherty TS, Driscoll DA. Coupling movement and landscape ecology for animal conservation in production landscapes. *Proc R Soc B Biol Sci.* 2018;285: 20172272. doi:10.1098/rspb.2017.2272
94. Plant RE. *Spatial data analysis in ecology and agriculture using R*. Boca Raton: CRC Press; 2012.
95. Visscher PK, Seeley TD. Foraging strategy of honeybee colonies in a temperate deciduous forest. *Ecology.* 1982;63: 1790–1801. doi:10.2307/1940121
96. Oldroyd BP, Rinderer TE, Buco SM. Intra-Colonial Foraging Specialism by Honey Bees (*Apis mellifera*) (Hymenoptera: Apidae). *Behav Ecol Sociobiol.* 1992;30: 291–295. Available: <http://www.jstor.org/stable/4600690>
97. De Vere N, Jones LE, Gilmore T, Moscrop J, Lowe A, Smith D, et al. Using DNA metabarcoding to investigate honey bee foraging reveals limited flower use despite high floral availability. *Sci Rep.* 2017;7: 1–10. doi:10.1038/srep42838
98. Librán-Embid F, Klaus F, Tschardt T, Grass I. Unmanned aerial vehicles for biodiversity-friendly agricultural landscapes - A systematic review. *Sci Total Environ.* 2020;732: 139204. doi:10.1016/j.scitotenv.2020.139204
99. Chen B, Jin Y, Brown P. An enhanced bloom index for quantifying floral phenology using multi-scale remote sensing observations. *ISPRS J Photogramm Remote Sens.* 2019;156: 108–120. doi:10.1016/j.isprsjprs.2019.08.006
100. Pritt M, Chern G. *Satellite Image Classification with Deep Learning*. 2020 [cited 26 Feb 2021]. Available: <http://arxiv.org/abs/2010.06497>
101. Zhou J, Wu L, Deng Y, Zhi X, Jiang Y-H, Tu Q, et al. Reproducibility and quantitation of amplicon sequencing-based detection. *ISME J.* 2011;5: 1303–1313. doi:10.1038/ismej.2011.11
102. Lee CK, Herbold CW, Polson SW, Wommack KE, Williamson SJ, McDonald IR, et al. Groundtruthing Next-Gen Sequencing for Microbial Ecology–Biases and Errors in Community Structure Estimates from PCR Amplicon Pyrosequencing. *PLoS One.* 2012;7: e44224. Available: <https://doi.org/10.1371/journal.pone.0044224>
103. Richardson RT, Lin C, Quijia JO, Riusech NS, Goodell K, Johnson RM. Rank-based characterization of pollen assemblages collected by honey bees using a multi-locus metabarcoding approach. *Appl Plant Sci.* 2015;3: 1500043. doi:10.3732/apps.1500043

104. Baksay S, Pornon A, Burrus M, Mariette J, Andalo C, Escaravage N. Experimental quantification of pollen with DNA metabarcoding using ITS1 and trnL. *Sci Rep.* 2020;10: 4202. doi:10.1038/s41598-020-61198-6
105. Keller A, Danner N, Grimmer G, Ankenbrand M, von der Ohe K, von der Ohe W, et al. Evaluating multiplexed next-generation sequencing as a method in palynology for mixed pollen samples. Mock H-P, editor. *Plant Biol.* 2014;17: 558–566. doi:10.1111/plb.12251
106. Legendre P. Interpreting the replacement and richness difference components of beta diversity. *Glob Ecol Biogeogr.* 2014;23: 1324–1334. doi:https://doi.org/10.1111/geb.12207
107. Donkersley P, Rhodes G, Pickup RW, Jones KC, Power EF, Wright GA, et al. Nutritional composition of honey bee food stores vary with floral composition. *Oecologia.* 2017;185: 749–761. doi:10.1007/s00442-017-3968-3
108. Hendriksma HP, Shafir S. Honey bee foragers balance colony nutritional deficiencies. *Behav Ecol Sociobiol.* 2016;70: 509–517. doi:10.1007/s00265-016-2067-5
109. Socolar JB, Gilroy JJ, Kunin WE, Edwards DP. How Should Beta-Diversity Inform Biodiversity Conservation? *Trends Ecol Evol.* 2016;31: 67–80. doi:10.1016/j.tree.2015.11.005
110. Decourtye A, Mader E, Desneux N. Landscape enhancement of floral resources for honey bees in agro-ecosystems. *Apidologie.* 2010;41: 264–277. doi:10.1051/apido/2010024
111. Kimura K, Yoshiyama M, Saiou K, Nirasawa K, Ishizaka M. Examination of mass honey bee death at the entrance to hives in a paddy rice production district in Japan: The influence of insecticides sprayed on nearby rice fields. *J Apic Res.* 2014;53: 599–606. doi:10.3896/IBRA.1.53.5.12
112. Saunders ME. Insect pollinators collect pollen from wind-pollinated plants: implications for pollination ecology and sustainable agriculture. Packer L, editor. *Insect Conserv Divers.* 2018;11: 13–31. doi:10.1111/icad.12243
113. Lau P, Bryant V, Ellis JD, Huang ZY, Sullivan J, Schmehl DR, et al. Seasonal variation of pollen collected by honey bees (*Apis mellifera*) in developed areas across four regions in the United States. Blenau W, editor. *PLoS One.* 2019;14: e0217294–e0217294. doi:10.1371/journal.pone.0217294
114. Brodschneider R, Gratzner K, Kalcher-Sommersguter E, Heigl H, Auer W, Moosbeckhofer R, et al. A citizen science supported study on seasonal diversity and monoflorality of pollen collected by honey bees in Austria. *Sci Rep.* 2019;9: 16633. doi:10.1038/s41598-019-53016-5

115. Requier F, Odoux JF, Tamic T, Moreau N, Henry M, Decourtye A, et al. Honey bee diet in intensive farmland habitats reveals an unexpectedly high flower richness and a major role of weeds. *Ecol Appl*. 2015;25: 881–890. doi:10.1890/14-1011.1
116. Liolios V, Tananaki C, Dimou M, Kanelis D, Goras G, Karazafiris E, et al. Clasificación del polen de las plantas melíferas en función de su aportación de proteínas para las abejas de la miel. *J Apic Res*. 2015;54: 582–592. doi:10.1080/00218839.2016.1173353
117. Donaldson-Matasci MC, Dornhaus A. How habitat affects the benefits of communication in collectively foraging honey bees. *Behav Ecol Sociobiol*. 2012;66: 583–592. doi:10.1007/s00265-011-1306-z
118. Critchley CNR, Fowbert JA, Wright B. Dynamics of species-rich upland hay meadows over 15 years and their relation with agricultural management practices. *Appl Veg Sci*. 2007;10: 307–314. doi:https://doi.org/10.1111/j.1654-109X.2007.tb00429.x
119. Williams NM, Cariveau D, Winfree R, Kremen C. Bees in disturbed habitats use, but do not prefer, alien plants. *Basic Appl Ecol*. 2011;12: 332–341. doi:10.1016/j.baae.2010.11.008
120. Urbanowicz C, Muñoz PA, McArt SH. Honey bees and wild pollinators differ in their preference for and use of introduced floral resources. *Ecol Evol*. 2020;10: 6741–6751. doi:10.1002/ece3.6417
121. Park B, Nieh JC. Seasonal trends in honey bee pollen foraging revealed through DNA barcoding of bee-collected pollen. *Insectes Soc*. 2017;64: 425–437. doi:10.1007/s00040-017-0565-8
122. Delaplane KS, Steen J Van Der, Guzman-nova E. Standard methods for estimating strength parameters of *Apis mellifera* colonies. *J Apic Res*. 2013;52: 1–12. doi:10.3896/IBRA.1.52.1.03
123. Geslin B, Le Féon V, Folschweiller M, Flacher F, Carmignac D, Motard E, et al. The proportion of impervious surfaces at the landscape scale structures wild bee assemblages in a densely populated region. *Ecol Evol*. 2016;6: 6599–6615. doi:10.1002/ece3.2374
124. Mallinger RE, Gaines-day HR, Gratton C. Do managed bees have negative effects on wild bees ? A systematic review of the literature. *PLoS One*. 2017;12: 1–32.
125. Blackmore LM, Goulson D. Evaluating the effectiveness of wildflower seed mixes for boosting floral diversity and bumblebee and hoverfly abundance in urban areas. Stewart A, Bezemer M, editors. *Insect Conserv Divers*. 2014;7: 480–484. doi:10.1111/icad.12071

126. Baldock KCR, Goddard MA, Hicks DM, Kunin WE, Mitschunas N, Morse H, et al. A systems approach reveals urban pollinator hotspots and conservation opportunities. *Nat Ecol Evol.* 2019;3: 363–373. doi:10.1038/s41559-018-0769-y

443 **Supplementary information**

444 **Tables**

445 **Table S1.** Climatic data of the study sites

Location	Prefectu	Elevation [m]	Mean Temperature [°C]	Precipitati
Enokisawa	Chiba	14	15.6	1428
Kuwata	Chiba	23	15.6	1428
Yachiyo	Chiba	31	14.9	1394
Ichihara	Chiba	36	15.5	1550
Nerima	Tokyo	38	15.1	1448
Shiba	Tokyo	6	15.4	1442
Togo	Tokyo	29	15.4	1442
Colombin	Tokyo	24	15.4	1442
Shinjyuku	Tokyo	32	15.4	1442
Toyosu	Tokyo	6	15.4	1442
Yamatecho	Kanagaw	23	15.6	1554
Ishikawach	Kanagaw	29	14.7	1488
Gumyoji	Kanagaw	13	15.6	1554

Honmoku	Kanagaw	12	15.6	1554
Nishichiba	Chiba	17	15.3	1435
Kashiwano	Chiba	19	14.7	1358
Inohana	Chiba	17	15.3	1435

446

447 Supplementary Figure legends

448 **Fig S1. Pollen traps.** (A) Pollen trap set-up at the entrance of the selected hive in the Nishi-
449 Chiba campus of the Chiba University (25 March 2020). (B) Close-up of pollen trap with pollen
450 balls collected in the trail in the Nishi-Chiba campus of the Chiba University (17 June 2020).

451 **Fig S2. Elbow clustering plot.** Elbow method plotting the total within sum of square explained
452 in function of the number of k clusters. The elbow of the curve suggests the number of groups
453 to retain for k-means clustering analysis.

454 **Fig S3. Principal component analysis (PCA) biplot of individual sites and landscape**
455 **variables.** The landscape variables (in blue) are Pd: Patch density [patches/km²], dpop:
456 demographic density (number of inhabitant/km²), MESH_veg: effective mesh size of
457 vegetation [-], IIC: Integral index of connectivity [-], veg_cover: vegetation cover (%),
458 NDVI_median: median of the NDVI of the cells superior to 0.2 [-], and SHEI: Shannon's
459 evenness index [-].

460 **Fig S4. Bar plot showing the nine most frequent families observed in the samples in**
461 **contrast with the landscape classes.** Colour proportions correspond to the landscape classes.

Fig 1.

Fig 2.

Fig 3.

bioRxiv preprint doi: <https://doi.org/10.1101/2021.08.05.455320>; this version posted August 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Fig 4.

Commented [A1]: Please note that this figure was not visible in the file that you sent.

Fig 5.

bioRxiv preprint doi: <https://doi.org/10.1101/2021.08.05.455320>; this version posted August 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Fig S1.

Optimal number of clusters

bioRxiv preprint doi: <https://doi.org/10.1101/2021.08.05.455320>; this version posted August 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Fig S2.

Fig S3.

Fig S4.