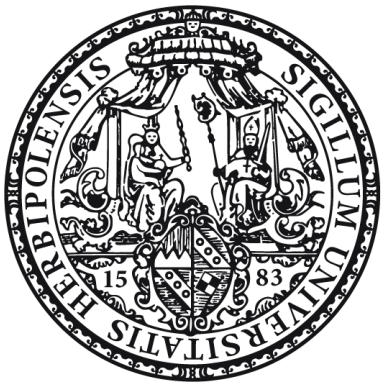
HONEY BEE FORAGING IN AGRICULTURAL LANDSCAPES



Dissertation zur Erlangung des
naturwissenschaftlichen Doktorgrades
der Julius-Maximilians-Universität Würzburg

vorgelegt von

Nadja Danner

geboren in Erlangen

Würzburg 2016

Eingereicht am: .....

Mitglieder der Promotionskommission:

Vorsitzender: .....

Gutachter : Prof. Dr. Ingolf-Steffan-Dewenter .....

Gutachter: Prof. Dr. Ricarda Scheiner .....

Tag des Promotionskolloquiums: .....

Doktorurkunde ausgehändigt am: .....

Abstract

Dissertation

Honey bee foraging in agricultural landscapes

by Nadja Danner

1. Today honey bee colonies face a wide range of challenges in modern agricultural landscapes which entails the need for a comprehensive investigation of honey bees in a landscape context and the assessment of environmental risks. Within this dissertation the pollen foraging of honey bee colonies is studied in different agricultural landscapes to gain insight into the use of pollen resources and the influence of landscape structure across the season. General suggestions for landscape management to support honey bees and other pollinators are derived.

2. Decoding of waggle dances and a subsequent spatial foraging analysis are used as methods in Chapters 4 and 5 to study honey bee colonies in agricultural landscapes. The recently developed metabarcoding of mixed pollen samples was applied for the first time in honey bee foraging ecology and allowed for a detailed analysis of pollen, that was trapped from honey bees in front hive entrances (Chapter 6).

3. Pollen identification through molecular sequencing and DNA barcoding has been proposed as an alternative approach to light microscopy, which still is a tedious and error-prone task. In this study we assessed mixed pollen probes through next-generation sequencing and developed a bioinformatic workflow to analyse these high-throughput data with a newly created reference database. To evaluate the feasibility, we compared results from classical identification based on light microscopy from the same samples with our sequencing results. Abundance estimations from sequencing data were significantly correlated with counted abundances through light microscopy. Next-generation sequencing thus presents a useful and efficient workflow to identify pollen at the genus and species level without requiring specialized palynological expert knowledge.

4. During maize flowering, four observation hives were placed in and rotated between 11 landscapes covering a gradient in maize acreage. A higher foraging frequency on maize fields compared to other landuse types showed that maize is an intensively used pollen resource for honey bee colonies. Mean foraging distances were significantly shorter for maize pollen than for other pollen origins, indicating that effort is put into collecting a diverse pollen diet. The percentage of maize pollen foragers did not increase with

maize acreage in the landscape and was not reduced by grassland area as an alternative pollen resource. Our findings allow estimating the distance-related exposure risk of honey bee colonies to pollen from surrounding maize fields treated with systemic insecticides.

5. It is unknown how an increasing area of mass-flowering crops like oilseed rape (OSR) or a decrease of semi-natural habitats (SNH) change the temporal and spatial availability of pollen resources for honey bee colonies, and thus foraging distances and frequency in different habitat types. Sixteen observation hives were placed in and rotated between 16 agricultural landscapes with independent gradients of OSR and SNH area within 2 km to analyze foraging distances and frequencies. SNH and OSR reduced foraging distance at different spatial scales and depending on season, with possible benefits for the performance of honey bee colonies. Frequency of pollen foragers per habitat type was equally high for SNH, grassland and OSR fields, but lower for other crops and forest. In landscapes with a small proportion of SNH a significantly higher density of pollen foragers on SNH was observed, indicating the limitation of pollen resources in simple agricultural landscapes and the importance of SNH.

6. Quantity and diversity of collected pollen can influence the growth and health of honey bee colonies, but little is known about the influence of landscape structure on pollen diet. In a field experiment we rotated 16 honey bee colonies across 16 agricultural landscapes (see also Chapter 5), used traps to get samples of collected pollen and observed the intra-colonial dance communication to gain information about foraging distances. Neither the amount of collected pollen nor pollen diversity were related to landscape diversity. The revealed increase of foraging distances with decreasing landscape diversity suggests that honey bees compensate for a lower landscape diversity by increasing their pollen foraging range in order to maintain pollen amount and diversity.

7. Our results show the importance of diverse pollen resources for honey bee colonies in agricultural landscapes. Beside the risk of exposure to pesticides honey bees face the risk of nutritional deficiency with implications for their health. By modifying landscape composition and therefore availability of resources we are able to contribute to the wellbeing of honey bees. Agri-environmental schemes aiming to support pollinators should focus on possible spatial and temporal gaps in pollen availability and diversity in agricultural landscapes.

Zusammenfassung

Dissertation

Sammelverhalten von Honigbienen in der Agrarlandschaft

von Nadja Danner

1. Honigbienen stehen heutzutage vor einer Vielzahl von Herausforderungen in der modernen Agrarlandschaft, was umfassende Untersuchungen von Honigbienen im Landschafskontext erforderlich macht. Im Rahmen dieser Arbeit wurde das Pollensammeln von Honigbienenvölkern in verschiedenen Agrarlandschaften studiert, um Einblick in die Nutzung von Pollenressourcen und auf den Einfluss der Landschaftsstruktur zu gewinnen.

2. Die Dekodierung von Schwänzeltänzen und eine anschließende räumliche Analyse des Sammelverhaltens werden als Methoden in den Kapiteln 4 und 5 eingesetzt, um Bienenvölker in Agrarlandschaften zu untersuchen. Das kürzlich entwickelte Metabarcoding von gemischten Pollenproben wurde zum ersten Mal in der Honigbienenökologie angewandt und ermöglichte eine detaillierte Analyse von Pollenproben, die per Pollenfallen vor den Stockeingängen gesammelt wurden (Kapitel 6).

3. Pollenbestimmung durch molekulare Sequenzierung und DNA Barcoding wurde als Alternative zur Lichtmikroskopie vorgeschlagen, die immer noch sehr mühsam und fehlerbehaftet ist. In dieser Studie bestimmten wir gemischte Pollenproben durch Next-Generation-Sequenzierung und entwickelten einen bioinformatischen Arbeitsablauf um diese Hochdurchsatz-Daten mit einer neu kreierten Referenzdatanbank zu analysieren. Um die Durchführbarkeit zu evaluieren verglichen wir Ergebnisse aus der klassischen Identifizierung via Lichtmikroskopie derselben Proben mit unseren Sequenzier-Ergebnissen. Häufigkeitsschätzungen auf Basis der Sequenzierdaten waren signifikant mit den gezählten Häufigkeiten via Lichtmikroskopie korreliert. Next-Generation-Sequenzierung stellt daher einen nützlichen und effizienten Arbeitsablauf dar, um Pollen auf dem Gattungs- und Artniveau zu bestimmen ohne spezielles palynologisches Expertenwissen zu benötigen.

4. Während der Maisblüte wurden vier Beobachtungsstöcke in 11 Landschaften mit einem Maisflächengradienten platziert und zwischen diesen rotiert. Maisfelder wurden intensiver genutzt als Flächen anderer Landnutzungstypen. Die mittleren Sammeldistanzen waren signifikant niedriger für Maispollen als Pollen anderer Herkunft, was darauf hinweist, dass Aufwand in das Sammeln einer diversen Pollendiät gesetzt wird. Der Anteil an Maispollensammlerinnen stieg nicht mit der Maisanbaufläche in der Landschaft

und wurde nicht durch Grünlandfläche als alternative Pollenressource reduziert. Unsere Ergebnisse ermöglichen die Schätzung des entfernungsbezogenen Expositionsrisikos von Honigbienenvölker auf Pollen aus den umliegenden Maisfeldern, die mit systemischen Insektiziden behandelt werden.

5. Es ist nicht bekannt, wie eine Zunahme von Massentrachten wie Raps (OSR) oder eine Abnahme von halbnatürlichen Habitaten (SNH) die zeitliche und räumliche Verfügbarkeit von Pollenressourcen für die Honigbienen, und damit Sammeldistanzen und -frequenzen in verschiedenen Lebensraumtypen verändert. Sechzehn Beobachtungsstöcke wurden in 16 Agrarlandschaften mit unabhängigen Gradienten an OSR- und SNH-Fläche innerhalb von 2 km platziert und regelmäßig rotiert, um Sammeldistanzen und -frequenzen zu analysieren. SNH und OSR reduzierten die Sammeldistanzen auf verschiedenen räumlichen Skalen und je nach Saison, mit möglichen Vorteilen für die Leistungsfähigkeit von Bienenvölkern. Die Häufigkeit der Pollensammler pro Habitattyp war gleich hoch für SNH, Grünland und OSR, aber niedriger für andere Kulturen und Wald. In Landschaften mit einem kleinen Anteil von SNH wurde eine deutlich höhere Dichte von Pollensammlerinnen auf SNH beobachtet, was auf die Begrenzung der Pollenressourcen in einfachen Agrarlandschaften und die Bedeutung von SNH hinweist.

6. Menge und Diversität des gesammelten Pollens können das Wachstum und die Gesundheit von Honigbienenvölkern beeinflussen, aber es ist wenig über den Einfluss der Landschaftsstruktur auf die Pollendiät bekannt. In einem Feldexperiment rotierten wir 16 Honigbienenkolonien über 16 Agrarlandschaften (siehe auch Kapitel 5), nutzten Pollenfallen um Proben des gesammelten Pollens zu nehmen und beobachteten die intrakoloniale Tanzkommunikation, um Informationen über die Sammeldistanzen zu erhalten. Weder Pollenmenge noch -diversität waren von der Landschaftsdiversität abhängig. Der offenbarte Anstieg von Sammeldistanzen mit abnehmender Landschaftsdiversität legt nahe, dass Honigbienen durch die Erweiterung des Pollensammelbereichs eine niedrigere Landschaftsdiversität kompensieren, um Pollenmenge und -diversität zu erhalten.

7. Unsere Ergebnisse zeigen die Bedeutung eines diversen Pollenangebots für Bienenvölker in der Agrarlandschaft. Neben dem Risiko einer Exposition gegenüber Pestiziden, stehen Bienenvölker vor der Gefahr von Mangelernährung mit Auswirkungen auf ihre Gesundheit. Durch eine Änderung der Landschaftzusammensetzung und damit der Verfügbarkeit von Ressourcen können wir zum Wohlergehen der Honigbienen beitragen. Agrarumweltmaßnahmen mit dem Ziel Bestäuber zu unterstützen, sollten sich

auf mögliche räumliche und zeitliche Lücken in der Pollenverfügbarkeit und Vielfalt in der Agrarlandschaft konzentrieren.

Contents

1 General Introduction

Supplying plenty of honey beside pollen, propolis, royal jelly and beeswax, honey bees (*Apis mellifera* L.) have been of interest for people ever since. Native to Eurasia and Africa, they are today domesticated on all continents except Antarctica. With their social way of life (Winston, 1987) and the unique communication through dances within the colony (von Frisch, 1946), honey bees certainly are among the most fascinating insects and have drawn the attention of many researchers. One of the significant findings in ecology was their importance as pollinators for agricultural production (Free, 1970), where a whole pollination industry is based on, but also for wild plants (Rollin et al., 2013). Today honey bee colonies face a wide range of challenges in modern agricultural landscapes which entails the need for a comprehensive investigation of honey bees in a landscape context and the assessment of environmental risks (Härtel and Steffan-Dewenter, 2014).

1.1 Honey bees in the modern agricultural landscape

The ongoing intensification of agriculture (see Tscharntke et al., 2005) leads to landscape homogeneity, mass-flowering events of monocultures and increased exposure to pesticides and genetically modified crops while semi-natural habitats (Chapter 5) are decreasing. For example, the demand for biofuels as renewable energy is resulting in a constantly increasing proportion of maize acreage (Meissle et al., 2009, ; Chapter 4). Similarly, the area of oilseed rape has been increasing worldwide (Food and Agriculture Organization of the United Nations, 2016, ; Chapter 5). As a consequence, spatial and temporal shortages of food resources arise which might influence health and sustainability of honey bee colonies, but also other pollinators, in a negative way (Requier et al., 2015; Pasquale et al., 2016). Foraging distances of honey bees – determined via decoding waggle dances - can be an indicator for the resource availability in a landscape, i.e. bees have to take larger distances to meet the colony's demand for pollen when pollen resources are low. Only few studies deal with the decoding of dances to address ecological questions (Härtel and Steffan-Dewenter, 2014) although it is a unique possibility to gain insight into honey bee foraging on a landscape scale. Similarly, very few studies deal with the performance of honey bee colonies in response to landscape variables (Sponsler and Johnson, 2015). Within

FIGURE 1.1: Honey bee (*Apis mellifera* L.) collecting pollen on chamomile.

this dissertation the foraging of honey bee colonies - specifically for pollen (Figure 1.1) - is studied in differently structured agricultural landscapes to gain insight into the use of pollen resources and the influence of landscape structure. General suggestions for landscape management to support honey bees and other pollinators are derived.

Landscape structure can be described by the composition and the configuration of landscape elements, each of which in turn may be represented by different indices (McGarigal and Marks, 1994). Within our studies we focussed on landscape composition, i.e. the term landscape structure always refers to some measure of landscape composition. We assume that landscape composition is a suitable indicator for available pollen resources, e.g. a higher proportion of floriferous semi-natural habitats in a landscape should correlate with a higher availability of pollen.

1.2 The importance of pollen

Pollen is of specific interest since it is (1) the primary source of protein, lipids and other nutrients including essential amino acids (2) potentially harmful in the case of genetic modification, (3) a significant pathway for pesticides into the honey bee colony. Pollen is stored in only small amounts within the colony which bears the threat of underdevelopment due to shortages of pollen resources in the agricultural landscape. Pollen diet affects a variety of processes and life-history traits (see Dolezal et al., 2016) like immune response (Alaux et al., 2010) and pathogen resistance (Di Pasquale et al., 2013), and influences individual bees physiologically, e.g. their lipid content (Toth and Robinson, 2005).

1.3 Environmental risks

The wide use of pesticides and genetically modified crops in agricultural landscapes worldwide entails the assessment of environmental risks about the impacts on non-target organisms. Feeding on pollen and nectar exposes honey bee larvea and adults directly to the environment (Babendreier et al., 2004; Hendriksma et al., 2013; Krupke et al., 2012). Especially mass-flowering crops like maize and oilseed rape are potential exposure pathways to chemicals applied in agriculture. Neonicotinoid pesticides like Imidacloprid distribute within the plant and can be detected in maize pollen (Bonmatin et al., 2003). They are highly toxic to honey bees (Krupke et al., 2012) and thought to contribute to the observed declines of honey bee colonies (vanEngelsdorp et al., 2008). Negative effects of oilseed rape treated with a combination of neonicotinoids on pollinators have also been shown (Rundlöf et al., 2015). Sub-lethal effects on reproduction and behaviour as well as synergistic interactions with parasites have been reported (see references in Blacquière et al., 2012).

Maize and oilseed rape are among the most widely grown genetically modified crops worldwide (FAO). Within Europe currently GM maize MON810 is allowed to be grown, while no genetically modified plants are grown anymore within Germany. In contrast to pesticides, no harm of honey bees through pollen of genetically modified crops has been proven so far (Duan et al., 2008; Malone and Burgess, 2009; Hendriksma et al., 2012; Hendriksma et al., 2013). However, dose-dependent pleiotropic effects may be possible (Steijven, Steffan-Dewenter, and Härtel, 2015).

Data about foraging distances or pollen input of specific crops can be used to estimate exposure risks in agricultural environments. Results should be taken into account when designing laboratory ERA studies, e.g. to determine realistic dosages for survival experiments.

1.4 Research questions

In Chapter 3 the method of multiplexed next-generation sequencing was evaluated for mixed pollen samples (applied in Chapter 6). In Chapter 4 the importance of maize as a pollen source with its implications for honey bee health is investigated in a field experiment on landscape scale. It was revealed (1) whether maize is frequently used, (2) how foraging distances differ between maize and other pollen, (3) how landscape composition and (4) alternative pollen sources influence maize pollen foraging. In Chapter 5 we focussed on mass-flowering oilseed rape and semi-natural habitats as pollen sources. We studied (1) whether landscape composition influences foraging distances for pollen, (2) whether there are seasonal differences in this influence, (3) the importance of OSR compared to

SNH as a pollen source and (4) the influence of landscape composition on foraging frequencies in different habitat types. In Chapter 6 we analyzed pollen samples collected by honey bee colonies to answer (1) whether the amount of collected pollen depends on landscape diversity or season, (2) whether landscape diversity or season influence richness and diversity of the pollen diet, (3) whether honey bees compensate for a lower landscape diversity by increasing their foraging range to maintain pollen amount and diversity and (4) which are the most abundant pollen taxa in the pollen diet over time.

2 Methods in honey bee foraging ecology

Decoding of waggle dances and a subsequent spatial foraging analysis are unique methods that reveal the interactions between honey bee colonies and the landscape (Chapters 4 and 5). The recently developed metabarcoding of mixed pollen samples (Chapter 3) was applied for the first time in honey bee foraging ecology (Chapter 6) and allowed for a detailed pollen analysis.

2.1 Decoding dances

Decoding the well-known waggle dance is a unique tool to gain insight into honey bee foraging in the (agricultural) landscape (Seeley, 1995). The most rewarding food resources are propagated by the dancing forager bees to recruit nest mates for further exploitation. In glass-sided observation hives (Figure 2.1) they can be easily observed. While the angle of the straight waggle run relative to the vertical encodes the direction (it corresponds to the angle between the horizontal lines hive - advertised resource and hive - sun azimuth), the duration of a dance unit encodes the distance between hive and resource (von Frisch, 1946). The measured dance unit might be the straight run only or a whole dance circuit. The ladder was used within our studies (Chapter 4 and 5), since they are based on in field observations of waggle dances and taking the longer measure increased accuracy. However, (Couvillon et al., 2012) suggest that the duration of the straight waggle run better represents the distance. They also applied an alternative approach: filming of honey bee colonies in observation hives and a subsequent analysis of performed dances on the screen. Filming allows for decoding of simultaneously performed dances and repeated measurements to increase accuracy, but needs some extra time capacity in addition to field work. An advantage of in field observation on the other hand, as applied in our studies, is the longer presence on site and therefore higher probability to detect anomalies and react instantly, e.g. to swarming behavior or robbery. The detailed procedure of decoding is described in the respective Methods sections in Chapters 4, 5 and 6.

FIGURE 2.1: Observation hive installed on a table in the field (left), with opened door and two brood frames (middle) and during observation under a light proof tent (right).

2.2 Spatial foraging analysis

Distance and direction of food resources result in discrete points in the landscape that represent important foraging locations of honey bee colonies. We combined them with digital land-use maps for a comprehensive analysis of honey bee foraging (Chapters 4 and 5). The measure of foraging frequency was developed, which is the number of foraging locations per hectare of the respective land use type they intersect with in a landscape and per hour observation time. It was applied to compare the foraging intensities on different land use types and assess their values as pollen resources in the agricultural landscape.

2.3 Pollen analysis

Honey bee related pollen can be sampled from honey bee guts, from bee bread within the hive or from collected pollen loads by using pollen traps in front of hive entrances. Within our research pollen loads were usually sampled by using traps we developed specifically for observation hives (Figure 2.2). They catch a representative amount of pollen loads from the hind legs of returning foragers as any other conventional pollen trap. In a more specific case we wanted to identify the pollen of dancing bees without disturbing them (Chapter 4). We assumed that observed dancing bees carry the same pollen type as any other returning forager within the observation period if pollen colors are identical. Honey bees returning with pollen loads in front of the hive entrance were catched with little plastic tubes, allowing for softly squeezing bees against a mesh on the other end of the tube and pick the pollen from their hind legs for microscopical identification. The pollen color was identified in the field and then compared to the pollen carried by dancing bees. This method has limitations since determining colors behind the glass of the observation hives under poor lighting conditions is difficult. However, we were able to distinguish maize pollen, the focus species, since it has a really bright yellow color and rough texture. To give an overview,

FIGURE 2.2: Pollen trap with removable 5mm hole grid, lid and drawer for collecting the pollen sample. Developed for observation hives of honey bee colonies.

pollen loads can be distinguished (1) morphological which allows for a rough analysis of diversity that does not rely on expert knowledge but on a good ability to distinguish colors and texture; (2) through light microscopy, which delivers a more precise picture of collected pollen but highly relies on palynological expert knowledge and fairly reaches species level, and (3) by metabarcoding of mixed pollen samples. The ladder represents an efficient workflow to identify pollen at the species and genus level (Chapter 3) and was successfully applied on ecological data in Chapter 6.

3 Evaluating multiplexed next-generation sequencing as a method in palynology

This chapter is published as: A. Keller, N. Danner, G. Grimmer, M. Ankenbrand, K. von der Ohe, W. von der Ohe, S. Rost, S. Härtel and I. Steffan-Dewenter (2015). Evaluating multiplexed next-generation sequencing as a method in palynology for mixed pollen samples. Plant Biology 17:558-566.

Key-words: DNA barcoding; high throughput; internal transcribed spacer 2; ITS2; meta-barcoding; molecular ecology; phylotyping; pollination, plant–pollinator interactions.

RESEARCH PAPER

Evaluating multiplexed next-generation sequencing as a method in palynology for mixed pollen samples

A. Keller[1,2], N. Danner[1], G. Grimmer[1,2], M. Ankenbrand[3], K. von der Ohe[4], W. von der Ohe[4], S. Rost[5], S. Härtel[1] & I. Steffan-Dewenter[1]

1 Department of Animal Ecology and Tropical Biology, University of Würzburg, Biocenter, Würzburg, Germany
2 DNA Analytics Core Facility, University of Würzburg, Biocenter, Würzburg, Germany
3 Department of Bioinformatics, University of Würzburg, Biocenter, Würzburg, Germany
4 LAVES Institut für Bienenkunde, Celle, Germany
5 Department of Human Genetics, University of Würzburg, Biocenter, Würzburg, Germany

Keywords
DNA barcoding; high throughput; internal transcribed spacer 2; ITS2; meta-barcoding; molecular ecology; phylotyping; pollination, plant–pollinator interactions.

Correspondence
A. Keller, Department of Animal Ecology and Tropical Biology, Biocenter, Am Hubland, 97074 Würzburg, Germany.
E-mail: a.keller@biozentrum.uni-wuerzburg.de

Editor
H.-P. Mock

Received: 14 April 2014; Accepted: 25 August 2014

doi:10.1111/plb.12251

ABSTRACT

The identification of pollen plays an important role in ecology, palaeo-climatology, honey quality control and other areas. Currently, expert knowledge and reference collections are essential to identify pollen origin through light microscopy. Pollen identification through molecular sequencing and DNA barcoding has been proposed as an alternative approach, but the assessment of mixed pollen samples originating from multiple plant species is still a tedious and error-prone task. Next-generation sequencing has been proposed to avoid this hindrance. In this study we assessed mixed pollen probes through next-generation sequencing of amplicons from the highly variable, species-specific internal transcribed spacer 2 region of nuclear ribosomal DNA. Further, we developed a bioinformatic workflow to analyse these high-throughput data with a newly created reference database. To evaluate the feasibility, we compared results from classical identification based on light microscopy from the same samples with our sequencing results. We assessed in total 16 mixed pollen samples, 14 originated from honeybee colonies and two from solitary bee nests. The sequencing technique resulted in higher taxon richness (deeper assignments and more identified taxa) compared to light microscopy. Abundance estimations from sequencing data were significantly correlated with counted abundances through light microscopy. Simulation analyses of taxon specificity and sensitivity indicate that 96% of taxa present in the database are correctly identifiable at the genus level and 70% at the species level. Next-generation sequencing thus presents a useful and efficient workflow to identify pollen at the genus and species level without requiring specialised palynological expert knowledge.

INTRODUCTION

Palynology, the scientific study of pollen and identification of its origin, plays an important role in studying mechanisms of plant–pollinator interactions (Wilcock & Neiland 2002), resource use of flower-visiting animals (Wcislo & Cane 1996; Kleijn & Raemakers 2008) and climate-related variation of plant communities through time (Tzedakis 1993; Sugita 1994; Marchant *et al.* 2001). Pollen grains often display a species-specific morphology, with diverse structure and sculpture. However, it remains difficult to delineate between closely related species when using light microscopy (Mullins & Emberlin 1997). As a result, many pollen types are simply grouped at genus or family level (Davies & Fall 2001) and data analyses on pollen diversity are strongly limited (Bagella *et al.* 2013). DNA barcoding, *i.e.* to identify and classify organisms according to a nucleotide sequence, has often and successfully been applied to all major groups of organisms, including plants and their pollen (Hebert *et al.* 2003; Zhou *et al.* 2007; Chen *et al.* 2010). Accordingly, molecular tools to analyse pollens have also substantially increased in their application and show great potential, especially with difficult and fossil taxa as well as taxa having low taxonomic information (Bennett & Parducci 2006; Zhou *et al.* 2007; Wilson *et al.* 2010).

Barcoding is further a promising new approach in ecology to directly determine the diversity of organisms in environmental samples (Sheffield *et al.* 2009; Valentini *et al.* 2009), *i.e.* samples that represent a mixture of species, *e.g.* faeces, soil or pollen collections, for which identification with classical methods is difficult or incomplete (Wilson *et al.* 2010). To analyse mixed sets of pollens originating from different plant organisms with DNA barcoding, however, is still a tedious and error-prone task, requiring manual separation of pollens to taxa, each to be amplified and sequenced individually. Studies evaluating applicability of high-throughput techniques for pollen materials are currently lacking (Wilson *et al.* 2010; Taylor & Harris 2012) or are restricted to specific investigations using quantitative real-time polymerase chain reaction (qrtPCR), where prior information about present organisms is required (Agodi *et al.* 2006; Schnell *et al.* 2010). Palynology would therefore benefit from species-level determination from mixed samples, larger counts, higher processing speed, improved objectivity and automation

to be attractive for large-scale studies (Stillman & Flenley 1996). Molecular methods based on high-throughput DNA sequencing could provide the required features to extend and improve classical pollen determination. Valentini *et al.* (2010) proposed next-generation sequencing (NGS) as a suitable method for this task. We agree with this idea, and thus in this study we evaluated performance and reliability of the new sequencing and bioinformatic strategies by directly comparing them with data obtained from light microscopy.

Specifically, we address the following challenges that emerge in DNA barcoding with mixed pollen samples. (i) A laboratory routine has to be defined that can be applied to all major plant clades, requiring universality of amplification priming regions and adequate length to be suitable for next-generation sequencing while holding enough sequence variation to differentiate between species. This routine includes DNA extraction, amplification, sample multiplexing, library preparation, sequencing with high-throughput devices and raw data cleanup. Also, (ii) a mapping algorithm must be developed that adequately maps the obtained sequences in their full length to reference samples, preferably in a hierarchical progression with confidence values for each level of taxonomy. Further, this algorithm has to perform sufficiently well to be able to process high-throughput data on a standard desktop computer and produce results in a reasonable time. (iii) A comprehensive reference database is required to derive the desired taxonomic annotations.

Several genetic marker regions have been proposed for DNA barcoding in plants that match the above requirements, foremost presence and feasibility for amplification in all investigated taxa, as well as low intraspecific but high interspecific variability to succeed in being species-specific (Hebert *et al.* 2003; Zhou *et al.* 2007; Chen *et al.* 2010; Hollingsworth *et al.* 2011). In this study, we use the internal transcribed spacer 2 (ITS2) region, which has been shown to be suitable as a barcode for plants (92.7% successful identifications in 6600 samples (Chen *et al.* 2010; Buchheim *et al.* 2011). Also, the enclosed genetic regions (5.8 S and 28 S) are highly conserved throughout the eukaryotes. Thus a universal primer for the analysis of probes consisting of multiple organisms is applicable, with a low risk of excluding taxa from the amplification (White *et al.* 1990; Keller *et al.* 2009; Chen *et al.* 2010). A further reason for choosing this marker is that a comprehensive ITS2 database already exists (Koetschan *et al.* 2010), enabling preparation of reference sequences suitable for our needs.

We approached the targeted tasks by combining and adapting existing molecular and bioinformatic tools to develop new functionalities for DNA barcoding of pollen samples that consist of multiple taxa. We then evaluated the performance and quality of the molecular and bioinformatic workflow by comparing our results with data from classical light microscopy identification of pollen samples. Further, we tested the applicability for samples with low pollen content and performed computer-based simulations to validate whether the bioinformatic classification pipeline is trustworthy.

MATERIAL AND METHODS

Pollen collection

The honeybee pollen samples were collected in 12 different landscapes in the region around Bayreuth, Germany. The distance between landscapes was at least 3 km, leading to diverse pollen inputs, depending on the surrounding floral resources. In the centre of each landscape we established a honeybee colony (*Apis mellifera carnica* L.) with a pollen trap in front of the hive entrance. Returning foragers had to pass through a 5-mm grid, removing the pollen load from their hind legs. From 21 July 2009 to 12 August 2009, every 1–3 days accumulated pollen loads were removed from the traps and stored as individual samples at −18 °C until the end of the sampling period. Pollen samples were dried at 30 °C for 1 week. Further, to assess variability in resource use of honeybees at one location, samples from three colonies located at the same study site were separately analysed (in the following designated as samples 12a, 12b and 12c). From each of the 14 samples (one per colony), 20% of the collected pollen was randomly taken and mixed for further analyses.

We performed next generation sequencing (NGS) as well as microscopy assessment of the samples. The samples were split into independent aliquots for these separate, blinded analyses. NGS was performed with samples by AK, GG and MA, whereas samples were classified through classical light microscopy by ND with expert guidance from KvO, without knowledge of the other group's results.

Two further pollen samples were obtained from solitary bee nests (*Osmia bicornis* L.) in October 2012 by swabbing the cell walls with cotton buds (Keller *et al.* 2013). In contrast to the relatively pure pollen samples obtained from honeybees, this experiment reflects samples strongly contaminated with nest building material (soil) and faeces, which is challenging to analyse with traditional methods. Solitary bee samples were thus only processed with NGS.

Classical pollen identification

Pollen samples were first analysed using light microscopy in the LAVES Institut für Bienenkunde, Celle, Germany. For microscopic pollen determination, 10 mg pollen loads of each sample were homogenised in 50 ml demineralised water with a magnetic stirrer for 1 h. An aliquot of 15 µl of the solution and 30 µl demineralised water were transferred to a slide, distributed equally over an area of the size of a cover glass and embedded in glycerine:gelatin after complete dehydration, following the method of Behm *et al.* (1996). From each sample, 500 randomly selected pollen grains were determined to genus level, and where possible to species level. Very rarely occurring pollen types were not determined (Behm *et al.* 1996).

Molecular pollen identification

Second pollen identification was done using DNA barcoding of the ITS2 region. The main working steps described below were: DNA extraction, amplification, sequencing, bioinformatic clean-up and taxonomic classification.

DNA extraction, amplification and sequencing

For each sample, 2 g pollens were added to 4 ml bidest H_2O and homogenised with an electronic pestle within a plastic tube. Of this emulsion, 200 µl (~ 50 mg pollens) were taken for the following extraction. We ground the aliquot with a TissueLyser LT (Qiagen, Hilden, Germany) and extracted DNA using the Machery-Nagel (Düren, Germany) NucleoSpin Food Kit;

we followed the special supplementary guidelines for pollen samples provided by the manufacturer. For PCR amplification we used the primers S2F and ITS4R originally designed by Chen *et al.* (2010) and White *et al.* (1990) to span a mean region of approximately 350 bp; this covers the complete ITS2 region. We adapted these primers to match 454 sequencing purposes and multiplexing by adding the 454 specific Adapters A and B, the linker key, and a variable multiplex identifier (MID). Thus the forward 'fusion' primer was 5′-CGT ATC GCC TCC CTC GCG CCA TCA GAT GCG ATA CTT GGT GTG AAT -3′ and the reverse 'fusion' primer was 5′-CTA TGC GCC TTG CCA GCC CGC TCA GXX XXX XXX XXT CCT CCG CTT ATT GAT ATG C-3′, where the X region designates a variable multiplex identifier (MID). In total, 16 MIDs were taken from the official Roche technical bulletin (454 Sequencing Technical Bulletin No. 005-2009, April 2009) to be able to process all our samples with one sequencing chip.

The PCR reaction mixes consisted of 0.25 μl of each forward and reverse primer (each 30 μM molar), 3 μl template DNA and 25 μl Phusion High-Fidelity DNA polymerase PCR 2x MasterMix (Thermo Scientific, Waltham, MA, USA). Bidest H_2O was added to a reaction volume of 50 μl. Samples were initially denatured at 94 °C for 4 min, then amplified using 25 cycles of 95 °C for 40 s, 49 °C for 40 s and 72 °C for 40 s. A final extension (72 °C) of 5 min was added at the end of the programme to ensure complete amplification. All samples were amplified in ten separate aliquots to reduce random effects on the community during PCR amplification (Fierer *et al.* 2008). PCR amplicons of these ten replicates were combined, gel-electrophoresed, trimmed for amplicon length and cleaned with the HiYield PCR Clean-up Kit (Real Biotech Corp., Banqiao City, Taiwan) according to the manufacturer's description. Cleaned samples were quantified using a Qubit II Flurometer (Invitrogen/Life Technologies, Carlsbad, CA, USA) and the dsDNA High-Sensitivity Assay Kit (also Invitrogen/Life Technologies) as described in the vendor's protocol. We used the BioAnalyzer 2200 (Agilent, Santa Clara, CA, USA) with High Sensitivity DNA Chips (also Agilent) for verification of fragment length distributions. Pyrosequencing and library preparation was performed according to guidelines for the GS junior (Roche, Basel, Switzerland). Sequencing was performed in-house with a GS junior device at the Department of Human Genetics (University of Würzburg, Germany) with original Roche GS junior titanium chemistry.

Bioinformatic clean-up

Data was demultiplexed into the different samples using the MID adapter sequences and the QIIME software (Caporaso *et al.* 2010; Kuczynski *et al.* 2011). During this step, only sequences spanning both priming regions were further used, *i.e.* only completely sequenced amplicons. Primers, adapters and MIDs were trimmed. Chimeric checking and quality filtering was also performed during this step. We restricted data to high-quality reads with a phred score ≥27 (Kunin *et al.* 2010), and no reads with ambiguous characters were included in the following downstream analyses.

Hierarchical classification

Taxonomic assignments were performed with the RDP (Ribosomal Database Project) classifier (Wang *et al.* 2007) and an ITS2-specific, novel reference set created and evaluated as described below. Further, we applied a bootstrap cut-off at 85% as classification threshold with respect to the maximum f-measure in the training database evaluation (see below).

Method comparison statistics

Most of the analyses were performed at a generic level, as both methods yielded some taxa only assignable to this level. With a generic analysis, all identified taxa were directly comparable. With these data, we compared taxon richness and identified species overlaps and differences obtained from the two methods. Rarefaction curves for each plot were generated with R (R Development Core Team 2010) in the NGS data to evaluate species richness in relation to sequencing depth. Abundance was assessed relatively as percentage of total number of reads and percentage of 500 pollen grains (Behm *et al.* 1996) for NGS and light microscopy, respectively. We used overall and per plot abundance of these relative accounts to compare between the two methodologies with Pearson's product moment correlation using R (R Development Core Team 2010).

Molecular reference database training

Taxonomic classifications with DNA barcodes are currently mostly done *via* phylogenetic analyses (Buchheim *et al.* 2011), pair-wise alignments with specific reference sequences (Chen *et al.* 2010) or BLAST searches (Basic Local Alignment Search Tool; Altschul *et al.* 1990) in GenBank (Benson *et al.* 2010) or other nucleotide databases. The first methods require that prior knowledge of taxonomy is present to select suitable taxa for inclusion into the recalculated phylogenetic tree or alignment. This is not feasible for mixed pollen collections, where the included taxa are unknown prior to assessment or stem from very different taxonomic groups. BLAST searches have to be performed very carefully, as hits may include local alignments, and identity calculations may thus be based only on parts of the query and reference sequences. Further, the raw output of a BLAST search is often obscured as many hits are not taxonomically annotated or flagged as 'environmental samples'. A novel approach to tackle these drawbacks has been proposed with a Bayesian classification algorithm (Wang *et al.* 2007). This provides hierarchical taxonomic assignments of DNA sequences and is well accepted in the scientific community, as especially high throughput analyses profit from the efficiency and accuracy of the algorithm (Caporaso *et al.* 2010). Currently, the only publicly available training sets are limited to bacterial 16 S (Wang *et al.* 2007) and fungal large ribosomal subunit (Liu *et al.* 2012).

In this study, a new ITS2 training set was designed for plants. We used the ITS2-Database as an original database that is restricted to structure-validated sequences (Koetschan *et al.* 2010). All ITS2 sequences matching the taxonomic group Viridiplantae and with a sequence length between 200 bp and 400 bp were downloaded, resulting in 73,853 sequences (accessed 3 March 2013). The taxonomy for each sequence was assigned using the GI (GenBank Identifier) and the corresponding NCBI taxonomy (Federhen 2012) with Perl scripting and reformatted to be usable with the python script 'assign taxonomy.py' of the QIIME (Caporaso *et al.* 2010) package. Additionally, RDP required formats of these pre-processed files were generated. Training was performed with the RDP

classifier version 2.2 (Wang *et al.* 2007) as implemented in QIIME. Before training the final set, we evaluated the performance by varying several parameters of the underlying data to maximise effectiveness and allow quality estimations of the assignments as described below.

Pre-clustering evaluation
Because of intraspecific variation (Song *et al.* 2012) and sequencing errors in the underlying data (Kunin *et al.* 2010), pre-clustering of reference sequences prior to training may prove useful to increase reliability of the results (Lan *et al.* 2012). Thus, from the full dataset we generated 11 separate training sets differing in the pre-clustering threshold of sequences before the actual training. Clusters of sequences were generated at identity levels of 90%, 91% ... 100%, and only the most abundant sequence of each cluster was picked. This also generated an even distribution of taxonomic units in the sets. To assess the assignment quality and depth, each sequence was reclassified to the training set. Then, starting from the root of the taxonomy of each sequence, every taxonomic level of the assignment was compared to the correct taxonomy. If the bootstrap of an assignment was <0.8, the level (and all sub-levels) was considered unassignable. If there was a mismatch between assigned taxonomy and expected taxonomy, the number of remaining sub-levels (plus one) was called erroneous level. The number of assigned levels before the first mismatch or unassignable level was called correct level.

Cut-off and assignment quality evaluation
To estimate assignment qualities, the test and training data must be distinct sets. Further, we wanted to evaluate the effectiveness in identifying 'new species' that do not have representatives in the training data (Lan *et al.* 2012). The complete ITS2 reference dataset was thus, for testing purposes, artificially split into three sets representing 'training data', 'test data A' with references and 'test data B' without references. This was achieved using the following procedure: species with multiple sequences were separated into 'test data A' (one sequence) and 'training data' (remaining sequences). Species with only a single deposited sequence were assigned to category 'test data B'. For this evaluation purpose, the algorithm was trained only with the set 'training data' (36,418 sequences). According to the measures for the RDP classifier evaluation performed by Lan *et al.* (2012) for the original 16S dataset, we estimated the number of 'true positive' (TP) and 'false negative' (FN) assignments by classifying sequences of 'test data A' (10,635 sequences), where references were present in the 'training data'. Only correct assignments were considered as TP, whereas wrong assignments (to a different species) were added to the list of FNs. Similarly, we classified sequences of 'test data B' (26,800 sequences) to determine the number of 'true negative' (TN) and 'false positive'(FP) hits. With these, we calculated sensitivity $SN = \frac{TP}{TP + TN}$ to identify existing taxa and specificity $SP = \frac{TN}{TN + FP}$ to leave sequences without references unclassified. Using these split datasets, we were able to estimate SN at species and genus level, whereas SP was only assessable at the species level. We optimised our assignment bootstrap value for classification by maximising the f-measure as the harmonic mean of sensitivity and specificity at species level $= \frac{2 \times SN \times SP}{SN + SP}$.

RESULTS

Pollen high-throughput sequencing and classification

In total, our study produced 14,924 raw sequences for pollen samples passing Roche's quality filtering of the 454 junior sequencing device. Of these, 9310 ITS2 sequences matched our extended quality standards. The remainder was dismissed as too short (<200 bp), with low quality score (<27), excess homopolymers (>5 bp), chimeric or mismatches in primer regions (Caporaso *et al.* 2010; Kunin *et al.* 2010). After removal of adapters and primers, mean sequence length was 348.3 bp ($\pm$ 28 bp SD), spanning the complete ITS2 region. Individual samples comprised 219–1179 reads, with mean read length of 330.5 $\pm$ 3.8 bp to 363.9 $\pm$ 68.2 bp. Beside plant sequences, we also found several fungal sequences, belonging to *Issatchenkia occidentalis*, *Cochliobolus sativus*, *Phoma* sp. and *Lewia infectoria*, which regularly inhabit or infect plant tissues.

Honeybee pollen samples

For the samples collected by honeybees, 98.9% of all reads were assignable to genus level, with a bootstrap confidence higher or equal than 0.85. At the species level, we were able to classify 61.6% of our reads using the same bootstrap cut-off. Reducing the filter's required sequence length to 150 bp did not produce any new classifiable plant taxa. Taxon richness was not correlated with the number of reads within a sample (Pearson's correlation, $r = -0.099$, $df = 12$, $t = -0.3453$, $P > 0.05$). Rarefaction showed that we reached a plateau regarding genera richness in all samples (Fig. 1A). These observations suggest that the sequencing depth was adequate to assess the underlying taxon richness.

We identified a total of 29 different genera of 16 families when we combined the results from molecular sequencing and microscopy (Table 1). Further, 24 taxa were also identifiable at species level. With NGS we found 13 genera that were not identified through microscopy, whereas four genera (*Heracleum, Carduus, Phacelia, Convolvulus*) that were identified by light microscopy were missing in the NGS results, despite having references in the database. One genus (*Vitis*) had no conclusive reference sequence in the database and was thus also not identifiable with the NGS method.

From phenology of the pollens and presence at plots, we assume that a misidentification of very similar pollens occurred with light microscopy, which was revealed by NGS: *Tanacetum* and *Scorzoneroides* were both manually misclassified as *Taraxacum*. We observed higher intra-generic taxon richness for *Trifolium*, *Hypochaeris* and *Chamerion* through NGS, yet less in *Centaurea* (Fig. 1B). Improvement of the taxonomic assignment was found in four genera, where species levels were obtainable only through NGS. However, *Helianthus* was only classified at genus level, whereas microscopy was able to identify it as *H. annuus*.

Based on NGS data, taxon richness within the samples ranged from four to 12 taxa that were at least classifiable at genus level (Fig. 1B). Correspondingly, diversity ranged from four to 12 taxa for the microscopy assessment. Pollen diversity collected using the three colonies from site 12 was 12, ten and ten taxa, respectively. The compositional profile was similar for the dominant pollen taxa in all three samples, but still showed considerable variation (Fig. 1B).

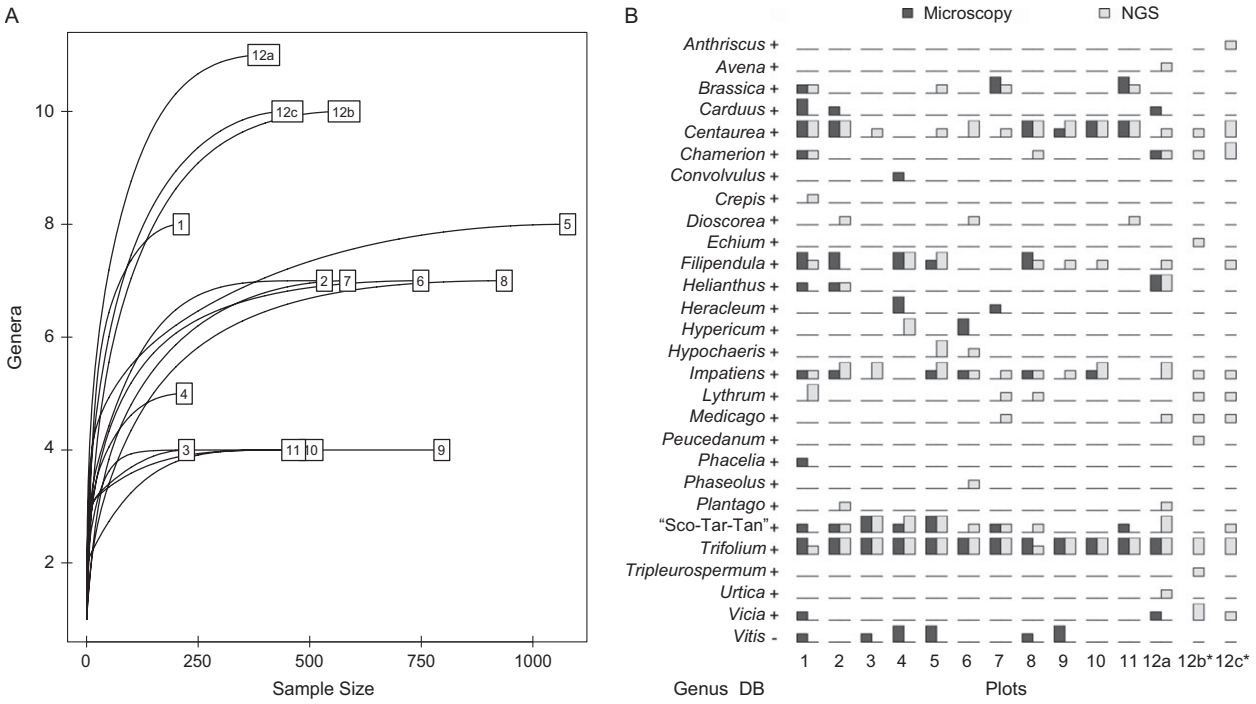


Fig. 1. A: Rarefaction of genera richness obtained for each honeybee sample with respect to sequencing depth. B: Plot-based comparison of pollen identification through optical microscopy and NGS. Taxonomic assignments are illustrated at the genus level. Positive identification of a taxonomic unit within a sample is indicated in the community matrix as dark grey for microscopy and light grey for NGS. Relative abundance estimations are indicated by size at two levels, *i.e.* ≥5% (fully filled box) and <5% (half-filled box) of total abundance within a sample. Genera misidentified in optical microscopy were combined for direct comparison and are indicated by quotation marks in abbreviated form (Tar = *Taraxacum*, Sco = *Scorzoneroides*, Tan = *Tanacetum*). Availability in the reference database is indicated in the column DB. *For sample 12, three samples were taken from the same study site but different colonies. All three samples were analysed using NGS to evaluate repeatability, yet optical microscopy was only performed for 12a.

Table 1. Plant families with their number of genera and number of species assessed by next generation sequencing (NGS) and optical microscopy.

	NGS		microscopy	
family	#genera	#species	#genera	#species
Apiaceae	2	2	1	1
Asteraceae	7	11	4	6
Balsamicaceae	1	1	1	1
Boraginaceae	1	2	1	1
Convolvulaceae	0	0	1	1
Brassicaceae	1	1	1	1
Dioscoreaceae	1	1	0	0
Fabaceae	4	10	2	4
Hypericaceae	1	2	1	1
Lythraceae	1	1	0	0
Onagraceae	1	3	1	1
Plantaginaceae	1	3	0	0
Poaceae	1	1	0	0
Rosaceae	1	1	1	1
Urticaceae	1	1	0	0
Vitaceae	0	0	1	1
total	24	40	15	19

Over all samples we found a strong correlation of abundance estimations between the two identification methods (Pearson's correlation, r = 0.86, t = 8.71, df = 26, $P < 0.001$; Fig. 2). This relationship is also reflected on a per plot basis, yet with a lower correlation coefficient (Pearson's correlation, r = 0.66, t = 17.36, df = 390, $P < 0.001$). These results indicate that the abundance estimates of taxa within plots show relatively high similarity between the two methods.

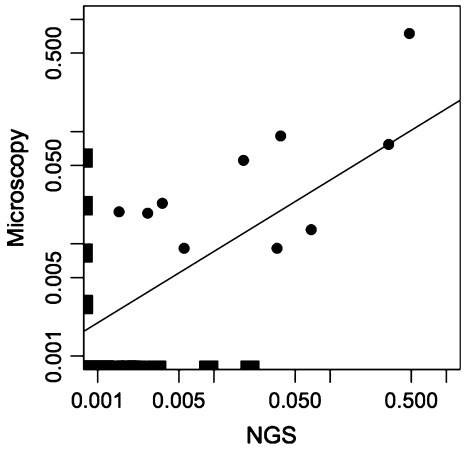


Fig. 2. Overall log-scaled relative abundance comparison of genera between the two classification strategies. Rectangles at the axes represent genera only found with one of the two sampling techniques. Pearson's correlation r = 0.86, t = 8.71, df = 26, $P < 0.001$

Pollens in solitary bee nests

Pollen samples from both solitary bee nests were successfully processed, with 100% of reads identifiable at genus level despite high contamination of the samples with nesting material and faeces. Both samples harboured *Brassica* sp. and *Dioscorea* sp. pollen, the latter most likely *Dioscorea* (*Tamus*) *communis* as the only representative of the Dioscoreaceae present in the sampling region.

Molecular reference database training

Pre-clustering of data prior to training of the RDP classifier did not improve the overall performance of classifications (Fig. 3). This was the case both for depth of the assignment as well as the mean number of incorrectly assigned levels, which, respectively, increase and decrease with higher pre-clustering thresholds. We thus used a cut-off at 100% sequence identity, which equals unique sequences, for the final training set. With that, of the 73,853 tested database sequences, 55,028 were positively identifiable at species and a further 10,518 at genus level. Surprisingly, 6104 sequences were assignable only to phylum level; they likely represent contamination in the reference database.

Regarding determination of the optimal cut-off threshold, specificity and sensitivity of the novel/known classifications are shown with their dependency of the bootstrap (Fig. 4). The best classification by means of f-measures is achieved with a bootstrap cut-off of 0.85. Both specificity and sensitivity at this threshold for species level were approximately 70%. At genus level, sensitivity to correctly identify a genus increased to 96%. We thus recommend this threshold when using the RDP classifier with the generated training data.

Currently, all sequences in the reference dataset accumulate to 37,435 different plant species and 6162 genera according to NCBI taxonomy (Federhen 2012). The complete reference dataset is available for download and public use at http://www.dna-analytics.biozentrum.uni-wuerzburg.de.

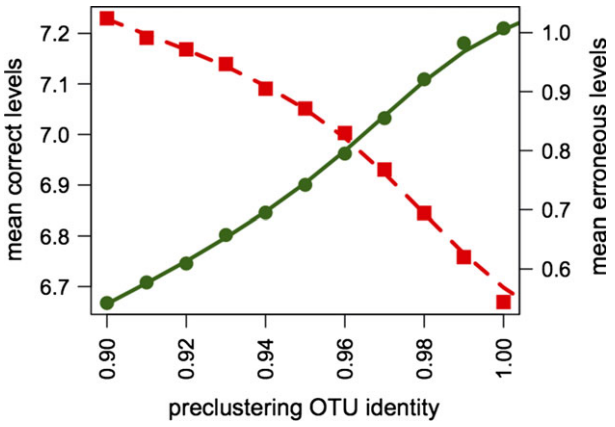


Fig. 3. Pre-clustering evaluation: Starting from the root of the taxonomy of each sequence, every taxonomic level of the assignment was compared to its correct lineage. The overall mean of correct assignments according to the different pre-clustering levels is presented as dots in the figure (left scale). Similarly, each sequence was tested for erroneous levels of classification with means displayed as squares and the scale on the right side.

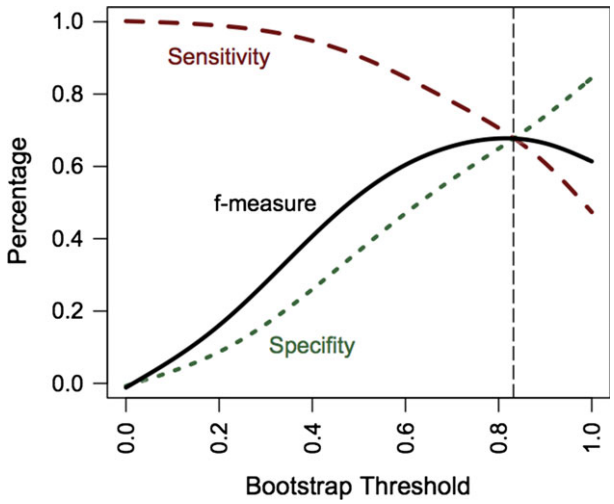


Fig. 4. Dependence of sensitivity and specificity by the bootstrap threshold. Sensitivity to identify at species level is illustrated with a single-dashed line. Specificity is displayed as a dotted line. The harmonic mean of both species-level measures is displayed as a solid black curve, maximised at approximately 0.85 as the suggested optimal classification threshold.

DISCUSSION

The demand for methods to identify pollen samples at a high-throughput level is increasing for many applications in ecology and paleo-climatology (Bennett & Parducci 2006; Zhou *et al.* 2007; Sheffield *et al.* 2009; Valentini *et al.* 2009; Wilson *et al.* 2010; Taylor & Harris 2012). DNA barcoding is a frequently and successfully applied method, yet pollens of mixed samples originating from more than one source are currently not assessable through standard methods. Valentini *et al.* (2010) proposed that NGS may counter this deficiency, *i.e.* to investigate such mixed samples by identifying all included plant organisms together, without manual separation. The goals of this study were thus to develop, and moreover evaluate, a molecular laboratory procedure and bioinformatic analysis for such a task. The complete workflow was applied to pollen samples from two different studies (in total 16 samples). The resulting gene sequences allowed us to successfully identify taxon richness and abundance of the underlying samples. The resulting taxonomic resolution is similar or better than results from classical light microscopy. Details of the performance of each individual step of the workflow and the resulting methodological and biological relevance are discussed below.

High-throughput pollen sequencing

In general, our laboratory workflow was suitable for processing mixed pollen probes through NGS. However, quality filtering according to our rigorous restrictions reduced the obtained sequences from approximately 15,000 sequences to 10,000. Most of them were removed due to failure to include both primer regions and/or multiplex identifier due to low quality scores towards the end of sequences or short read lengths (Caporaso *et al.* 2010). The former indicates that a large proportion of reads was not fully sequenced with sufficient quality,

whereas the latter shows that the primers also amplified shorter fragments than the intended plant ITS2 region. Not fully sequenced reads are a technical issue that is regularly improved by increasing read length and quality through new generations of sequencing devices and chemistry (Metzker 2009). Improvements can also be expected by applying paired-end strategies, as quality near the ends will increase, or using technologies with general lower sequencing error rates. Shorter, fully sequenced sequences are project-specific problems, but are also expected: as a drawback of universal primers, they will also amplify fungal ITS2 (White *et al.* 1990) ranging from approximately 100 to 250 bp, and even other eukaryotic protists with far shorter ITS2 regions (Keller *et al.* 2009). Further, the existence of non-functional pseudo-genes is known (Harpke & Peterson 2008). Thus studies investigating plant ITS2 sequences should account for a sufficient overhead of estimated sequences per sample during project design related to sequencing technology and potential contamination from unwanted organisms (Parameswaran *et al.* 2007). The remaining high-quality reads showed a high proportion of classifiable sequences (~99%), whereas reduction of the minimum sequence length had no impact on plant species diversity. Both observations suggest that the filters are adequate to concentrate on the data of interest, *i.e.* plant sequences.

Classification pipeline

To be able to use the RDP classifier (Wang *et al.* 2007) for taxonomic assignments with plants and with the ITS2 marker, we re-trained the algorithm with structurally verified sequences obtained from the ITS2 database (Koetschan *et al.* 2010). The underlying dataset incorporates more than 70,000 different plant sequences and represents a cross-section throughout the Viridiplantae. Sequences originate from all biogeographic regions of the world since the primary database is GenBank (Benson *et al.* 2010). Currently, all sequences in the reference dataset represent 37,435 different plant species and 6162 genera according to NCBI taxonomy (Federhen 2012). Exemplarily for the data analysed in this study, the dataset covers 79% of all vascular plant genera and 54% of species known to exist within the Federal state of Bavaria, Germany, where our samples were obtained (comprehensive plant database http://www.bayernflora.de, accessed 6 November 2013; Staatliche Naturwissenschaftliche Sammlungen Bayern 2013). As 99% of reads were classifiable to genus level and only one genus (*Vitis*) of the assessed 29 genera in total was missing in the reference database, most of the abundant and bee relevant plant genera seem to be included. Further, the classifier's dataset is updateable to match the constantly increasing number of sequences deposited in GenBank and the ITS2 database in the future (Wang *et al.* 2007).

In the computational evaluation of database and classifier for an ITS2 dataset, we obtained values comparable to those of existing datasets published for bacteria (Wang *et al.* 2007) and fungi (Liu *et al.* 2012). Taxonomic classifications performed best regarding sensitivity, *i.e.* to identify taxa existing in the database, and specificity, *i.e.* to restrain from classifying organisms without references, at a bootstrap threshold level of approximately 0.85 (Lan *et al.* 2012). Species- and genus-level sensitivity to correctly identify sequences with this bootstrap were 70% and 96%, respectively. This is similar to the classifier's preferred level used to classify microbial organisms (0.80; Wang *et al.* 2007; Lan *et al.* 2012). From a technical perspective, it is thus valid to also apply the classification algorithm for ITS2 sequences of plants.

Comparison of assessment methods

Using NGS, we were clearly able to improve palynology diversity assessments in comparison with traditional optical microscopy. This appears in novel taxa that were identified, as well as improvement of classification of taxa and better possibilities to distinguish species within a genus. Further, some misidentifications of pollen through microscopy were revealed that were caused by very similar morphological appearance of closely related species. Also, molecular assessments were successful for solitary bee nest samples, where swabs included pollens as well as contaminating material. Sequencing assessments were repeatable, identifying similar diversity in samples obtained from different bee colonies placed within the same landscape.

However, using the high-throughput approach we also encountered limitations, which are partly related to the data used for training of the classifier. Regarding the Vitaceae, the ITS2 database is currently lacking acceptable reference sequences. We validated the only existing sequence, which was very short (~200 bp) and derived from a whole genome shotgun sequencing study (assembled sequence from short length reads, GenBank ID: AM462492.2; Velasco *et al.* 2007). Due to intra-genomic variation of the ITS2 (Song *et al.* 2012), we assume the assembly yielded a consensus, stacked ITS2 sequence, not usable for barcoding purposes or that a non-ITS2 region was falsely identified as such by the ITS2 database annotation algorithm (Keller *et al.* 2009). We therefore dismissed the sequence as missing within the reference database. In general, taxa missing or with inadequate sequences in the underlying database are not identifiable. As shown exemplarily for the geographic region of Bavaria, 22% of known plant genera are missing, and thus the current coverage is far from complete (Staatliche Naturwissenschaftliche Sammlungen Bayern 2013). Also, valid sequences with wrong taxonomic annotations may lead to mis-training of the classification model regarding the respective taxa (Bridge *et al.* 2003). This is highlighted by a proportion of sequences re-classified in the evaluation to a different phylum, suggesting wrong taxonomic annotation of GenBank database sequences. To address limitations of the underlying database (missing or misclassified sequences) in a given research question, we suggest that applied studies should also consider reviewing one cross-section pool of all samples in parallel through optical means to verify the overall richness of taxa relevant for the study. This will also maintain comparability between studies applying traditional and molecular approaches. Despite these database-specific drawbacks, the classifier produced taxonomic assignments that are congruent with light microscopy, and thus corroborating the positive technical evaluation of the pipeline above with a direct comparison of biological data.

Abundance estimations for both methods showed a strong correlation, suggesting that abundance estimates based on high-throughput sequencing regarding high or low sequence frequency of taxa within the sample are valid. In our study, we took care to reduce amplification biases through PCR with ten aliquots of each sample simultaneously (typical in microbiota

studies: three, Fierer *et al.* 2008) and a low number of amplification cycles (Suzuki & Giovannoni 1996). Nevertheless, abundances retained from PCR-amplified DNA samples must be regarded critically, as amplification biases through priming preference of specific taxonomic groups, random effects and the exponential nature of the amplification process cannot be excluded (Suzuki & Giovannoni 1996; Spooner 2009). Abundances are thus likely better interpreted as categorical (*e.g.* high abundance, low abundance) than with linear association. With the advent of increased sequencing throughput and third-generation single molecule sequencers without need for amplification (Metzker 2009; Roberts *et al.* 2013), improved abundance estimations from sequencing are likely in the near future.

Cost per sample was almost equal for both applied methods when considering time and consumables. As the trend of sequencing technologies moves rapidly toward higher throughput and resulting multiplexing possibilities (Metzker 2009; Kozich *et al.* 2013), we expect price efficiency per sample with NGS to outpace optical assessments in the near future.

Fields of application

Various applications arise for the proposed method. These include studies of pollen material of various origin, including plants themselves, pollinators, soil samples and wind collections. The results of such assessments are of great importance in identifying the diversity and specialisation of plant–pollinator interaction networks (Bosch *et al.* 2009) and also in supporting agricultural and ecological management decisions (*e.g.* Girard *et al.* 2012; Odoux *et al.* 2012). Further, paleo-ecological and climate change-associated studies investigating fossil pollens may also profit (Bennett & Parducci 2006).

Special attention is currently required in quality control of honeybee products, including the geographic origin, correct labelling of different varieties based on the used floral resources and detection of contamination from genetically modified (GM) crops (Picard-Nizou *et al.* 1995; Hemmer 1997). As pollen is naturally incorporated into honey and protocols to isolate pollens are common usage (Sowunmi 1976), high-throughput sequencing and classification may make a large contribution to this endeavour by facilitating the analytical process and inclusion of references from plant taxa throughout the world (Sowunmi 1976; Ruoff *et al.* 2007).

Furthermore, the methodology may be equivalently applied to other questions not only related to pollens. Other target samples are naturally occurring communities of plants, (*e.g.* green algae) or artificially mixed probes of plant tissue fragments (Schlumbaum *et al.* 2008). As the primers used in this study also efficiently amplify fungal ITS2 sequences, ancillary information is automatically gained about this group, including pathogens as *Ascosphaera* spp. that may be present in collected pollen samples and vectored through harvesting flights of worker bees (Gilliam 1990; White *et al.* 1990).

CONCLUSIONS

Expert knowledge is essential to adequately identify pollens through traditional light microscopy, while taxonomic expertise is also often restricted to specific plant groups or geographic regions. Further, mixed samples of pollens from several plant origins present a problem in current palynology. With this study we evaluated NGS to approach pollen assessments through molecular techniques including their bioinformatic analysis. The analytical pipeline is designed for high-throughput data, but also adaptable to single sequences. It is a useful technique, broadening the assessment capabilities from expert labs to all work groups with access to standard molecular laboratory equipment. Further, our results show that this assessment method improves the standard technique with regard to taxonomic depth, overall diversity and rectifying misidentifications.

ACKNOWLEDGEMENTS

We greatly appreciate the help of staff from the Department of Human Genetics (University of Würzburg, Germany), especially C. Müller-Reible and T. Haaf for placing the 454 device at our disposal for this study. We also thank F. Förster (Department of Bioinformatics, University of Würzburg, Germany) for providing a useful taxonomy Perl module for generation of the training sets, and two anonymous reviewers for their valuable comments.

DATA ACCESSIBILITY

Sequences have been deposited at the ENA:SRA (https://www.ebi.ac.uk/ena) and are accessible under study accession number PRJEB5016. The used training set alongside installation and application notes, is available for download at http://www.dna-analytics.biozentrum.uni-wuerzburg.de.

REFERENCES

Agodi A., Barchitta M., Grillo A., Sciacca S. (2006) Detection of genetically modified DNA sequences in milk from the Italian market. *International Journal of Hygiene and Environmental Health*, **209**, 81–88.

Altschul S.F., Gish W., Miller W., Myers E.W., Lipman D.J. (1990) Basic local alignment search tool. *Journal of Molecular Biology*, **215**, 403–410.

Bagella S., Satta A., Floris I., Caria M.C., Rossetti I., Podani J. (2013) Effects of plant community composition and flowering phenology on honeybee foraging in Mediterranean sylvo-pastoral systems. *Applied Vegetation Science*, **16**, 689–697.

Behm F., von der Ohe K., Henrich W. (1996) Zuverlässigkeit der Pollenanalyse von Honig: bestimmung der Pollenhäufigkeit. *Deutsche Lebensmittel-Rundschau*, **92**, 183–188.

Bennett K.D., Parducci L. (2006) DNA from pollen: principles and potential. *The Holocene*, **16**, 1031–1034.

Benson D.A., Karsch-Mizrachi I., Lipman D.J., Ostell J., Sayers E.W. (2010) GenBank. *Nucleic Acids Research*, **38** (Suppl. 1), D46–D51.

Bosch J., Martín González A.M., Rodrigo A., Navarro D. (2009) Plant–pollinator networks: adding the pollinator's perspective. *Ecology Letters*, **12**, 409–419.

Bridge P.D., Roberts P.J., Spooner B.M., Panchal G. (2003) On the unreliability of published DNA sequences. *New Phytologist*, **160**, 43–48.

Buchheim M., Keller A., Koetschan C., Forster F., Merget B., Wolf M. (2011) Internal Transcribed Spacer 2 (nu ITS2 rRNA) sequence–structure phylogenetics: towards an automated reconstruction of the green algal tree of life. *PLoS One*, **6**.

Caporaso J.G., Kuczynski J., Stombaugh J., Bittinger K., Bushman F.D., Costello E.K., Fierer N., Pena A.G., Goodrich J.K., Gordon J.I. (2010) QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, **7**, 335–336.

Chen S., Yao H., Han J., Liu C., Song J., Shi L., Zhu Y., Ma X., Gao T., Pang X. (2010) Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. *PLoS ONE*, **5**, e8613.

Davies C.P., Fall P.L. (2001) Modern pollen precipitation from an elevational transect in central Jordan and its relationship to vegetation. *Journal of Biogeography*, **28**, 1195–1210.

Federhen S. (2012) The NCBI taxonomy database. *Nucleic Acids Research*, **40**, D136–D143.

Fierer N., Hamady M., Lauber C.L., Knight R. (2008) The influence of sex, handedness, and washing on the diversity of hand surface bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 17994–17999.

Gilliam M. (1990) *Chalkbrood disease of honey bees, Apis mellifera, caused by the fungus, Ascosphaera apis: a review of past and current research*. Vth International Colloquium on Invertebrate Pathology and Microbial Control, Adelaide, Australia, pp 398–402.

Girard M., Chagnon M., Fournier V. (2012) Pollen diversity collected by honey bees in the vicinity of *Vaccinium* spp. crops and its importance for colony development This article is part of a Special Issue entitled 'Pollination biology research in Canada: perspectives on a mutualism at different scales'. *Botany*, **90**, 545–555.

Harpke D., Peterson A. (2008) 5.8 S motifs for the identification of pseudogenic ITS regions. *Botany*, **86**, 300–305.

Hebert P.D., Cywinska A., Ball S.L. (2003) Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **270**, 313–321.

Hemmer W. (1997) *Foods derived from genetically modified organisms and detection methods*. Agency for Biosafety Research and Assessment of Technology Impacts of the Swiss Priority Programme Biotechnology of the Swiss National Science Foundation, http://www.bats.ch/bats/publikationen/1997-2_gmo/gmo_food.pdf.

Hollingsworth P.M., Graham S.W., Little D.P. (2011) Choosing and using a plant DNA barcode. *PLoS ONE*, **6**, e19254.

Keller A., Schleicher T., Schultz J., Müller T., Dandekar T., Wolf M. (2009) 5.8S–28S rRNA interaction and HMM-based ITS2 annotation. *Gene*, **430**, 50–57.

Keller A., Grimmer G., Steffan-Dewenter I. (2013) Diverse microbiota identified in whole intact nest chambers of the red mason bee *Osmia bicornis* (Linnaeus 1758). *PLoS One*, **7829**, 6.

Kleijn D., Raemakers I. (2008) A retrospective analysis of pollen host plant use by stable and declining bumble bee species. *Ecology*, **89**, 1811–1823.

Koetschan C., Forster F., Keller A., Schleicher T., Ruderisch B., Schwarz R., Muller T., Wolf M., Schultz J.. (2010) The ITS2 Database III – sequences and structures for phylogeny. *Nucleic Acids Research*, **38**, D275–D279.

Kozich J.J., Westcott S.L., Baxter N.T., Highlander S.K., Schloss P.D. (2013) Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Applied and Environmental Microbiology*, **79**, 5112–5120.

Kuczynski J., Stombaugh J., Walters W.A., González A., Caporaso J.G., Knight R. (2011) Using QIIME to analyze 16S rRNA gene sequences from microbial communities. *Current Protocols in Bioinformatics*, **10**, 17.

Kunin V., Engelbrektson A., Ochman H., Hugenholtz P. (2010) Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. *Environmental Microbiology*, **12**, 118–123.

Lan Y., Wang Q., Cole J.R., Rosen G.L. (2012) Using the RDP classifier to predict taxonomic novelty and reduce the search space for finding novel organisms. *PLoS One*, **7**, e32491.

Liu K.-L., Porras-Alfaro A., Kuske C.R., Eichorst S.A., Xie G. (2012) Accurate, rapid taxonomic classification of fungal large-subunit rRNA genes. *Applied and Environmental Microbiology*, **78**, 1523–1533.

Marchant R., Behling H., Berrio J.C., Cleef A., Duivenvoorden J., Hooghiemstra H., Kuhry P., Melief B., Geel B.V., Van der Hammen T. (2001) Mid- to Late-Holocene pollen-based biome reconstructions for Colombia. *Quaternary Science Reviews*, **20**, 1289–1308.

Metzker M.L. (2009) Sequencing technologies—the next generation. *Nature Reviews Genetics*, **11**, 31–46.

Mullins J., Emberlin J. (1997) Sampling pollens. *Journal of Aerosol Science*, **28**, 365–370.

Odoux J.-F., Feuillet D., Aupinel P., Loublier Y., Tasei J.-N., Mateescu C. (2012) Territorial biodiversity and consequences on physico-chemical characteristics of pollen collected by honey bee colonies. *Apidologie*, **43**, 561–575.

Parameswaran P., Jalili R., Tao L., Shokralla S., Gharizadeh B., Ronaghi M., Fire A.Z. (2007) A pyrosequencing-tailored nucleotide barcode design unveils opportunities for large-scale sample multiplexing. *Nucleic Acids Research*, **35**, e130.

Picard-Nizou A., Pham-Delegue M., Kerguelen V., Douault P., Marilleau R., Olsen L., Grison R., Toppan A., Masson C. (1995) Foraging behaviour of honey bees (*Apis mellifera* L.) on transgenic oilseed rape (*Brassica napus* L. var. *oleifera*). *Transgenic Research*, **4**, 270–276.

R Development Core Team (2010) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria (01/19).

Roberts R.J., Carneiro M.O., Schatz M.C. (2013) The advantages of SMRT sequencing. *Genome Biology*, **14**, 405.

Ruoff K., Luginbühl W., Kilchenmann V., Bosset J.O., von der Ohe K., von der Ohe W., Amadò R. (2007) Authentication of the botanical origin of honey using profiles of classical measurands and discriminant analysis. *Apidologie*, **38**, 438–452.

Schlumbaum A., Tensen M., Jaenicke-Després V. (2008) Ancient plant DNA in archaeobotany. *Vegetation History and Archaeobotany*, **17**, 233–244.

Schnell I.B., Fraser M., Willerslev E., Gilbert M.T.P. (2010) Characterisation of insect and plant origins using DNA extracted from small volumes of bee honey. *Arthropod-Plant Interactions*, **4**, 107–116.

Sheffield C.S., Hebert P.D., Kevan P.G., Packer L. (2009) DNA barcoding a regional bee (Hymenoptera: Apoidea) fauna and its potential for ecological studies. *Molecular Ecology Resources*, **9** (Suppl. 1), 196–207.

Song J., Shi L., Li D., Sun Y., Niu Y., Chen Z., Luo H., Pang X., Sun Z., Liu C. (2012) Extensive pyrosequencing reveals frequent intra-genomic variations of internal transcribed spacer regions of nuclear ribosomal DNA. *PLoS One*, **7**, e43971.

Sowunmi M. (1976) The potential value of honey in palaeopalynology and archaeology. *Review of Palaeobotany and Palynology*, **21**, 171–185.

Spooner D.M. (2009) DNA barcoding will frequently fail in complicated groups: an example in wild potatoes. *American Journal of Botany*, **96**, 1177–1189.

Staatliche Naturwissenschaftliche Sammlungen Bayern. (2013) Botanischer Informationsknoten Bayern, Germany.

Stillman E., Flenley J.R. (1996) The needs and prospects for automation in palynology. *Quaternary Science Reviews*, **15**, 1–5.

Sugita S. (1994) Pollen representation of vegetation in Quaternary sediments: theory and method in patchy vegetation. *Journal of Ecology*, **82**, 881–897.

Suzuki M.T., Giovannoni S.J. (1996) Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. *Applied and Environmental Microbiology*, **62**, 625–630.

Taylor H.R., Harris W.E. (2012) An emergent science on the brink of irrelevance: a review of the past 8 years of DNA barcoding. *Molecular Ecology Resources*, **12**, 377–388.

Tzedakis P. (1993) Long-term tree populations in northwest Greece through multiple Quaternary climatic cycles. *Nature*, **364**, 437–440.

Valentini A., Pompanon F., Taberlet P. (2009) DNA barcoding for ecologists. *Trends in Ecology & Evolution*, **24**, 110–117.

Valentini A., Miquel C., Taberlet P. (2010) DNA barcoding for honey biodiversity. *Diversity*, **2**, 610–617.

Velasco R., Zharkikh A., Troggio M., Cartwright D.A., Cestaro A., Pruss D., Pindo M., FitzGerald L.M., Vezzulli S., Reid J. (2007) A high quality draft consensus sequence of the genome of a heterozygous grapevine variety. *PLoS One*, **2**, e1326.

Wang Q., Garrity G.M., Tiedje J.M., Cole J.R. (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology*, **73**, 5261–5267.

Wcislo W.T., Cane J.H. (1996) Floral resource utilization by solitary bees (Hymenoptera: Apoidea) and exploitation of their stored foods by natural enemies. *Annual Review in Entomology*, **41**, 257–286.

White T., Bruns T., Lee S., Taylor J. (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis M., Gelfand D., Shinsky J., White T. (Eds), *PCR-protocols: a Guide to Methods and Applications*. Academic Press, San Diego, pp 315–322.

Wilcock C., Neiland R. (2002) Pollination failure in plants: why it happens and when it matters. *Trends in Plant Science*, **7**, 270–277.

Wilson E.E., Sidhu C.S., LeVan K.E., Holway D.A. (2010) Pollen foraging behaviour of solitary Hawaiian bees revealed through molecular pollen analysis. *Molecular Ecology*, **19**, 4823–4829.

Zhou L.J., Pei K.Q., Zhou B., Ma K.P. (2007) A molecular approach to species identification of Chenopodiaceae pollen grains in surface soil. *American Journal of Botany*, **94**, 477–481.

4 Maize pollen foraging by honey bees in relation to crop area and landscape context

This chapter is published as: Nadja Danner, Stephan Härtel, Ingolf Steffan-Dewenter (2014). Maize pollen foraging by honey bees in relation to crop area and landscape context. Basic and Applied Ecology 15:677-684.

Key-words: *Apis mellifera*, *Zea mays*, resource use, foraging distances, bee decline, exposure risk.

Summary

The increasing demand for insect pollinated crops and high recent losses of honey bee colonies raise concerns about food security. Systemic insecticides are recognized as one of the drivers of worldwide honey and wild bee declines. Particularly honey bees in agricultural environments are exposed to pesticides when they collect crop pollen and nectar. However, landscape scale studies which analyze pollen use and foraging distances of honey bees on mass-flowering crops like maize to evaluate potential exposure risks are currently lacking. In an experimental approach on a landscape scale we took advantage of intra-colonial dance communication to gather information about the location of utilized pollen resources. During maize flowering, four observation hives were placed in and rotated between 11 different landscapes which covered a gradient from low to high maize acreage. A higher frequency of dances for foraging locations on maize fields compared to other land use types shows that maize is an intensively used pollen resource for honey bee colonies. Mean foraging distances were significantly shorter for maize pollen than for other pollen origins. The percentage of maize pollen foragers did not increase with maize acreage in the landscape. The proportion of grassland area providing alternative pollen sources did not reduce the percentage of maize pollen foragers. Our findings allow estimating the distance-related exposure risk of honey bee colonies to pollen from surrounding maize fields treated with systemic insecticides. Similarly, the results can be used to estimate the exposure to transgenic maize pollen, which is relevant for honey production in European countries. Provision of alternative pollen resources within agri-environmental schemes could potentially reduce exposure risk to pesticide

contaminated crop pollen.

Introduction

The honey bee (*Apis mellifera* L.) is a globally distributed pollinator and plays an important role in maintaining the ecosystem service of pollination in agricultural landscapes (Klein et al., 2007; Potts et al., 2010). Pollination and the survival of pollinators are at risk since bees face a number of potential threats from agricultural intensification, such as exposure to pesticides (Henry et al., 2012b; Pettis et al., 2012; Stokstad, 2013), pesticide/bee-pathogen interactions (Blacquière et al., 2012; Pettis et al., 2013), lower immunocompetence due to monofloral pollen diets (Alaux et al., 2010) or temporal gaps in pollen availability due to monocultures, habitat loss and fragmentation (Potts et al., 2010). On the other hand, the proportion of agricultural production that depends on pollinators shows a strong increase in the last two decades (Aizen et al., 2009).

The few studies that have used the spatial information provided by bee dances to analyze the foraging ecology of this major pollinator indicate that landscape structure, resource availability and season impact foraging distances (Steffan-Dewenter and Kuhn, 2003). However, the way how crops and landscape composition shape the spatial distribution of used floral resources is still unknown, despite its general relevance for maintaining pollination services and protecting honey bees from potentially negative impacts of intensive agriculture, including pesticide applications and their interaction with pathogens (Härtel and Steffan-Dewenter, 2014).

In central Europe the demand for biofuels as renewable energy is resulting in a constantly increasing proportion of maize acreage (Meissle et al., 2009). On a global scale, maize cultivation is increasingly dominated by genetically modified (GM) varieties and characterised by intensive pesticide application including systemic neonicotinoids. Being a wind-pollinated mass flowering crop, maize provides huge amounts of pollen within a flowering period of 2-5 weeks on a landscape scale. It was reported to be a pollen source for honey bees (Keller, Fluri, and Imdorf, 2005; Odoux et al., 2012) despite its visually unattractive flowers compared to most insect pollinated plants.

Feeding on pollen and nectar exposes larvae and adult bees directly to the environment (Babendreier et al., 2004; Hendriksma et al., 2013; Krupke et al., 2012). Mass-flowering crops are a potential exposure pathway to chemicals applied in agriculture. Particularly neonicotinoid pesticides which are used for crop seed dressing and distribute within the plant can be detected in maize pollen (Bonmatin et al., 2003). Neonicotinoids are highly toxic for honey bees (Krupke et

al., 2012) and synergistic interactions with spreading honey bee diseases are reported (Blacquière et al., 2012). In agricultural landscapes the exposure to neonicotinoids is thought to contribute to the observed declines of honey bee colonies (vanEngelsdorp et al., 2008). In Europe seed coating with three commonly used neonicotinoids was banned in 2013 for two years, but in many other countries of the world the application is still agricultural practice. The potential harm of honey bees through crop pollen contaminated with pesticides is proven while for GM pollen no detrimental effects have been found (Duan et al., 2008; Malone and Burgess, 2009). However, there is another reason why the exposure to GM pollen is a relevant issue at least within Europe. Pollen is an inherent compound of honey and bee keepers are highly interested to produce GM pollen free products. In view of the observed detrimental effects of neonicotinoids on pollinators (Henry et al., 2012b) and the expansion of GM crops, it is astonishing that the resource use of crop pollen including maize has never systematically been analyzed to estimate exposure risks in agricultural environments.

In this study we address pollen foraging of honey bee colonies in selected landscapes that cover a gradient from low to high percentage of maize acreage. In order to investigate the importance of maize as a pollen source on a landscape scale with its implications on bee health we test the following hypotheses: (1) Maize is a frequently used pollen resource for honey bee colonies; (2) foraging distances differ between maize pollen and other pollen species; (3) the proportion of maize pollen foragers increases with the amount of maize acreage in the landscape; (4) in landscapes with alternative pollen sources the proportion of maize pollen foragers is reduced.

Materials and Methods

Study region and sites

The study was performed in summer 2009 around Bayreuth in northern Bavaria, Germany. The study region is characterized by a mix of intensively managed cropland, extensive grasslands and differently sized forest fragments (Table 4.1). Cultivation of barley and maize is accounting for most of the cropland area (24% and 16% respectively) within the region. We selected 11 circular landscapes with a radius of 1500 m (Steffan-Dewenter and Kuhn, 2003), the maximum possible maize acreage gradient within the study region and independent gradients of grassland and crop area (Table 1). In each landscape maize crop fields were located at distances from below 100 m up to 1500 m from the centre. The minimum distance between landscape centres was 3000 m.

TABLE 4.1: Land use characteristics for 11 landscapes in a 1500 m radius around observation hives. Means ± standard errors and ranges (%).

Parameter	Mean ± SE	Range
Crop area total	43.7 ± 3.5	31.8-64.5
Maize area	11.1 ± 2.2	3.0-22.4
Grass land area	24.1 ± 1.3	17.8-31.0
Forest area	19.9 ± 3.2	5.3-36.0

Land cover data

Within a radius of 1500 m the cover of maize area was calculated on the basis of mapping maize fields on site in spring 2009 (landscape examples in Figure 4.1). The cover of land use types grassland, other crop land (excluding maize) and forest was calculated on the basis of digital land use maps (Bayerisches Landesvermessungsamt 2009). Forest area was excluded from linear regression analysis since it was correlated with maize and crop area (Pearson's R = -0.8 and -0.6 respectively).

Observation hives

To observe honey bee waggle dances we used four glass-sided observation hives with two Zander brood frames each (comb area 3056 cm^2). Colonies were built three weeks before starting with observations using young, mated queens (*Apis mellifera carnica*) obtained from the LAVES Institute for Apidology in Celle, Germany. All queens hatched in 2009, derived from a single mother and were mated at the same queen-mating station to assure minimal genetic differences between colonies. Brood frames with similar amounts of brood cells, honey and pollen stocks (assessed via Liebefeld method) were transferred to observation hives along with queens and approximately 4000 workers per hive. Returning foragers were restricted to perform their dances on one side of the comb, which always was the side for observation of dances. Therefore a diagonal wooden block in the passage at the bottom of the observation hive was installed, closing the gaps between the brood frame and the hive.

Experimental design

The four observation hives were placed in the centre of four out of 11 landscapes. Hives were moved to four other landscape centres each night after termination of flight activity on a day with suitable weather conditions for foraging. The pseudo-random rotation of the observation hives between landscapes was based on the following criteria: during the study period every landscape was occupied by at least three out of the four observation hives to account for potential colony

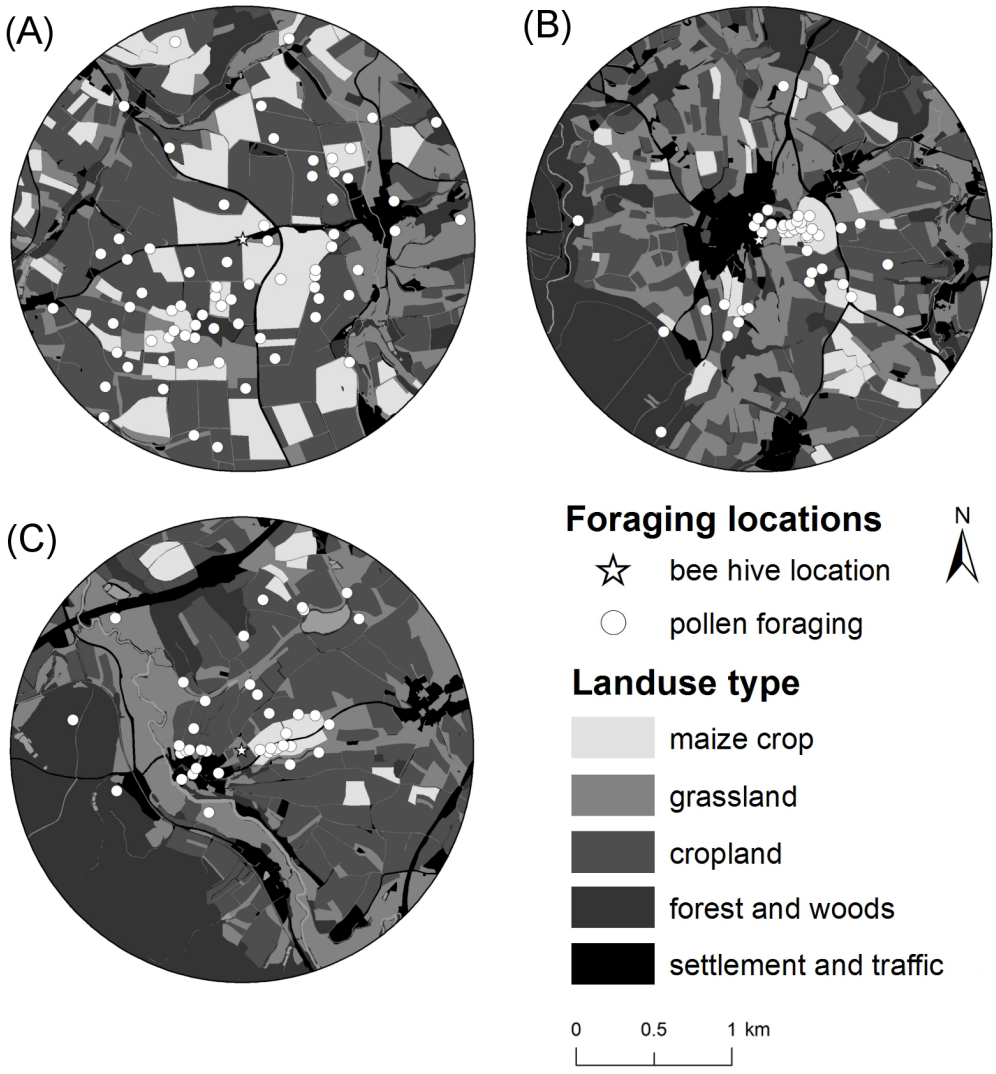


FIGURE 4.1: Examples of landscapes with high (A), medium (B) and low (C) proportion of maize crop area in the foraging range of 1.5 km. All dots indicate pollen foraging locations of a honey bee colony in the centre of each landscape accumulated during the study time.

effects. Landscapes with low, middle and high proportions of maize were always occupied at the same time.

Observation and decoding of bee dances

Honey bee waggle dances were observed from 22 July until 7 August. Except for days with unsuitable weather conditions, observation directly followed rotation of colonies at night. One observation period lasted normally 90 min (at least 30 min, depending on dance activity) and observation was terminated if there was no dancing performed for 15 min. Over the whole study period, observation periods in each landscape were distributed over day time as equally as possible to avoid bias due to species-specific time of the day for pollen supply. All colonies were observed at least once on each observation day. We considered a series of circuits as a single dance, each circuit consisting of a straight waggle run and the return run. For each dance we recorded the duration of a series of circuits, the corresponding number of circuits, the average angle of the waggle runs relative to the vertical, the time of day and the color of pollen carried by the dancing bee. Only dances of pollen-carrying foragers with a minimum of five consecutive circuits were decoded. The angle of sun position relative to the north added to the recorded waggle run angle provides the direction of the pollen foraging site indicated in each dance. Flight distance (y) was calculated via the mean duration of a single dance circuit (x) according to a third-order polynomial fit ($y = 92.137 - 346.659 * x + 228.454 * x^2 - 10.963 * x^3$) based on data presented by von Frisch (1965). The determined locations of pollen foraging sites were plotted into a land use map using the software ArcGIS 10 (ESRI, 2011).

We assumed that observed dancing bees carry the same pollen type as any other returning forager during one observation period if the pollen colors were identical. For each definable pollen color of pollen loads of dancing bees, a returning forager carrying the same pollen color was captured in front of the hive entrance. The pollen was carefully detached from their hind legs and stored for microscopic identification to allow for differentiation between dances for maize pollen and other pollen species (grouped together and not further specified). Maize dances determined in this way were used to analyze pollen foraging distances and proportions of maize pollen foragers. This method was applied to include all observed recruitments which to a small extent naturally went beyond our chosen 1500 m radius for land use mapping.

In order to analyze the relative use of different land use types within 1500 m, foraging locations below 1500 m were assigned to the land use type they directly intersected with (maize crop, other crops, grassland) in GIS maps. Variation in dance frequency for different land use types was calculated as number of dances per hour observation time and per hectare of the corresponding land use area in the landscape. In contrast to the method of pollen analysis, here we are restricted

to the mapped radius but obtain more specific spatial information concerning the underlying land use.

Statistical analysis

Statistical analysis was performed using the open source statistic software R (R Development Core Team, 2011). Linear mixed models were applied to analyze foraging distances (R package lme4; Bates, Maechler, and Bolker, 2011). Site and colony were implemented as crossed random factors, maize area and pollen type (maize pollen / other pollen) as explanatory variables. Foraging distances were square-root-transformed and maize foraging distances log-transformed in order to achieve normal distribution and homogeneity of variance of the models residues. Chi^2 and P-values were calculated by Likelihood ratio tests between the full model and the model without an explanatory variable to check for its significance. A general linear model (family = binomial) was applied to analyze percentages of dances for maize pollen. P-values were calculated by Chi^2-tests between models with and without an explanatory variable to select significant explanatory variables (Crawley, 2007). A Likelihood ratio based R^2 was calculated (R package MuMIn; Barton, 2012). An ANOVA with post hoc comparisons after Tukey and "BH" correction (Benjamini and Yekutieli, 2001) was performed to check for differences between log transformed dance frequency per hour and hectare between land use types.

Results

Foraging distances

During maize flowering a total of 614 bee dances for pollen sources in 11 landscapes with varying maize acreage were recorded and decoded. 125 dances (19%) advertised for maize pollen (determined by color / pollen analysis) and 489 for other pollen species m (see supplementary Table 4.S1). The mean pollen foraging distance was 820 m +/- 582 m with a range from 14 – 4439. 90% of all foraging locations were within 1538 m around the hive, proving that our chosen radius of 1500 m for land-use mapping was adequate. The mean foraging distances of bees that collected maize pollen were significantly lower (589 m $\pm$ 41 m, range 27 – 3040 m) than for other pollen origins (879 m $\pm$ 27 m, range 14 – 4439 m; lmer: Chi^2 = 20.25, $P < 0.001$; Figure 4.2). The maize area in a landscape radius of 1500 m did not influence maize pollen foraging distances (lmer: Chi^2 = 0.42, P = 0.52). 5% of all maize pollen foraging locations were beyond 1456 m.

Maize pollen foraging and landscape context

We found a significantly higher dance frequency for the land use type maize compared to forests, grasslands and other crops (ANOVA, F_4 = 5.9, P = 0.003; Figure

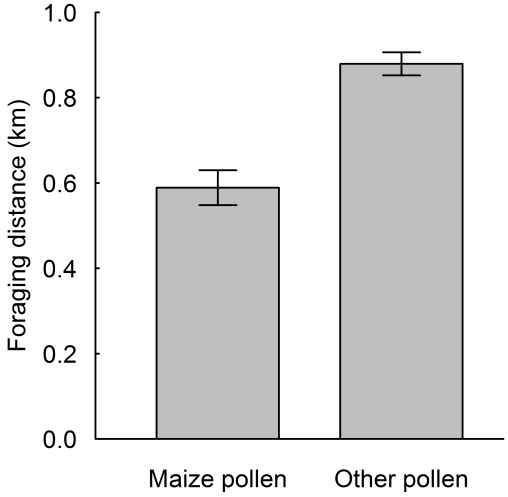


FIGURE 4.2: Mean foraging distances for maize pollen and other pollen ± SE in 11 landscapes during maize flowering (linear mixed model: $F_{1,614}$= 20.04, $P < 0.001$).

4.3). In contrast to our hypothesis, the proportion of maize pollen foragers did not increase with the area of maize fields in the surrounding landscape, instead it significantly decreased (Figure 4.4A). To test whether a flower-rich habitat type modulated the preferences of pollen foragers in our colonies we also included the area of grassland as an alternative pollen source in the model. However, we found no significant influence of grassland area on the proportion of maize pollen foragers (Figure 4.4B).

Discussion

Foraging distances

Our method for assigning a single dance to a certain pollen type made it possible to distinguish between dances for maize pollen and dances for other pollen species even for distances above 1500 m. This is the first study providing foraging distances for a specific pollen type, in our case maize pollen. Maize crop fields in each landscape were located at distances from below 100 m up to 1500 m from the hive, thus a sufficient supply of maize pollen was available at all distances within the main foraging range. The longer distances for other pollen than maize indicate that instead of a pure maize diet honey bees prefer to collect diverse pollen types, although this leads to higher costs in terms of energy. Landscape structure influences foraging distances (Steffan-Dewenter and Kuhn, 2003), and thus it can be expected that changing proportions of different land-use types affect foraging distances. The lack of an influence of maize area on pollen foraging distances suggests that the pollen demand for small honey bee colonies is already met by the pollen yield of only a few maize fields next to the hive. A similar observation of

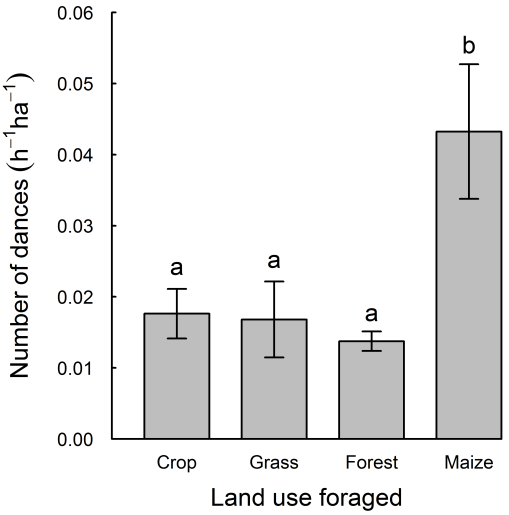


FIGURE 4.3: Habitat preferences based on dance frequency measured as number of dances per hour observation time and hectare of the corresponding land-use type (different letters indicate significant differences between groups: Tukey post hoc test and "BH" correction $P < 0.05$; (Benjamini and Yekutieli, 2001).

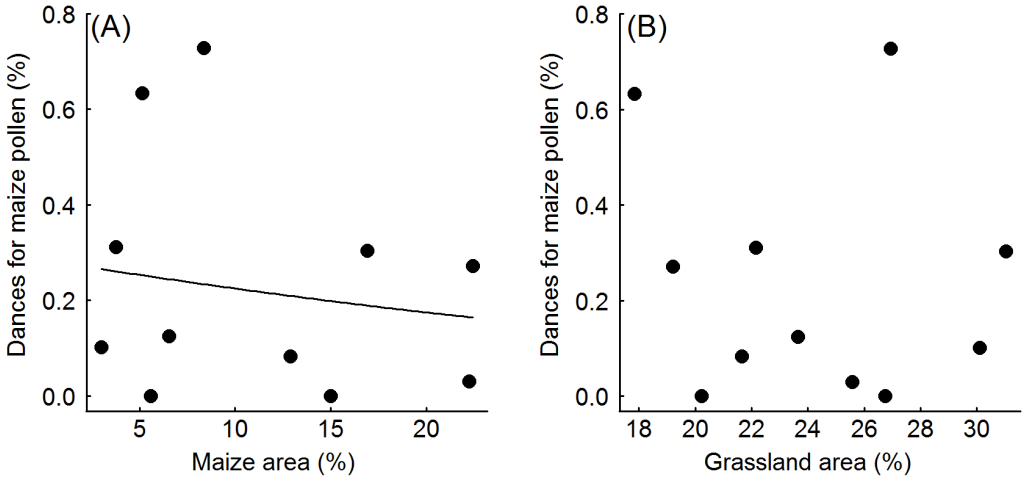


FIGURE 4.4: Percentage of maize pollen foragers per landscape (N = 11) in relation to maize area (A) and grassland area (B). Chi^2-tests: df = 1;$P < 0.001$; R^2 = 0.35 and df = 1; P = 0.14; R^2 = 0.08, respectively.

intense use of a small acreage of maize was made in an agricultural landscape in France (Odoux et al., 2012). This kind of foraging patterns is supported by visual analysis of the 11 GIS maps (examples in Figure 4.1). They show that only up to four maize fields, mostly in direct vicinity of the hive, out of 12 to 59 maize fields in the entire radius/landscape, intersect with foraging locations for maize pollen. This finding is in accordance with the significantly lower foraging distances for maize pollen as discussed above.

Maize pollen foraging and landscape context

We show that maize is a highly relevant pollen source for honey bees. Through analysing pollen input data of 114 studies it is already known that maize is among the most frequently detected pollen species in honey bee colonies (Keller, Fluri, and Imdorf, 2005). However, this meta-analysis gives no information about maize pollen quantity, landscape structures and foraging distances. Odoux et al. (2012) report the relevance of maize pollen as well, studying only a single landscape. With over 50% of the pollen harvest it was the dominant species during 5 weeks of maize pollen shedding despite its restricted cultivation area (4%) within a radius of 2500 m. We show that maize, being a mass-flowering crop, is intensively foraged during flowering while pollen resources on other land use types are probably less dense and foraging locations are therefore more "diluted". Our experimental design allowed us to independently test for effects of maize acreage and other land use types. In contrast to our hypothesis foraging for maize pollen measured as proportion of dances for maize pollen did not increase with maize area in a landscape. This result demonstrates that it is not possible to reduce the proportion of maize pollen foragers by a reduction of maize area within the foraging range. Most probably a few maize fields close to the hive already meet the demand for maize pollen. However, the underlying mechanisms resulting in the observed decrease can hardly be explained within the scope of our study and a detailed discussion would require additional experiments. Remarkably, there was a flowering field (about one hectare mixture of different clover species, thistle species and sunflower among other honey bee pollinated species) sown in the context of agri-environmental measures next to the honey bee hive location in the landscape with the second highest proportion of maize area (22.2%). Here we observed an extremely low proportion of maize foragers (0.03%) despite the presence of maize fields likewise next to the hive. Excluding this unusual data point from the analysis results in a non-significant decrease of the proportion of maize pollen foragers. Furthermore, foraging for maize pollen was not influenced by grassland area. In contrast to our hypothesis, grassland as alternative pollen source did not modulate maize pollen foraging. However, our result has to be interpreted with caution since we could not control for the actual resource availability on grassland. Mowing of grassland in late summer might have reduced

available resources. Generally, we assume that highly rewarding alternative floral resources could alter foraging patterns and thereby reduce the percentage of maize pollen foragers. The extremely low proportion of maize foragers in the landscape with the flowering field (see above) indicates that flower-rich habitats could prevent honey bees from maize pollen foraging.

Potential exposure risks

Our data are suitable to estimate the potential exposure risk to maize fields treated with systemic insecticides. It has been shown that neonicotinoids are present in maize pollen of plants grown from treated seeds (Krupke et al., 2012). In laboratory studies a chronic toxicity of the neonicotinoid Imidacloprid was detected with the chronic lethal dose being 4000 times lower than the acute lethal dose (Suchail, Guez, and Belzunces, 2001). Further sub-lethal effects on reproduction and behaviour as well as a synergistic interaction with the gut parasite Nosema ceranae have been reported (see references in Blacquière et al., 2012). Given that maize flowering takes approximately 2-5 weeks and maize fields are intensively foraged during that time, our data demonstrate the potential exposure risk of bee brood and nurse bees to toxic neonicotinoids during a critical colony stage when long-living winter bees are developing in temperate regions. The potential hazard of maize fields cannot be controlled by maize area within the main foraging range, as shown by the lack of a positive correlation with the percentages of maize dances. Thus, a reduction in exposure to maize pollen can only be reached by a large distance of the honey bee colony to the next maize field, together with provision of highly rewarding alternative floral resources. Our data show that 95% of maize foraging flights were performed within 1500 m around the hive. We suggest that the exposure risk to maize pollen could be decreased significantly by keeping a distance of more than 1500 m to the next maize field. However, maize foraging distances might increase under other circumstances, e.g. in landscapes with no alternative floral resources at shorter distances. We assume that honey bees would not forage on maize fields out of their main foraging range as long as other resources are available. Therefore, the establishment of attractive flower-rich habitats might be a successful measure to actively prevent honey bees from foraging on maize fields. In order to assess the actual exposure to crop pollen under field conditions, pesticide loads in maize pollen or other crops and quantitative input data in relation to the respective crop area have to be collected. Currently only very few data are available on the actual exposure of honey bee larvae or nurse bees to maize pollen grains (Babendreier et al., 2004; Hendriksma et al., 2013). Further studies are needed to determine the spatial distribution of pollen and nectar resource use to estimate how real world combinations and loads of pesticides enter and affect honey bee colonies (Pettis et al., 2013).

Conclusions

In conclusion, maize is an important pollen source for honey bees during late summer. The data on foraging distances can be used to estimate colony exposure risks to pesticides and will also help to declare minimum distances between bee hives and GM maize fields to avoid GM pollen contamination of honey. Following our results we suggest that a distance of 1500 m could reduce exposure risk to maize pollen by 95% in comparable landscapes, which would need to be approved by further studies.

Supplement

Table 4.S1: Descriptive statistics of honey bee pollen foraging distances [m] at 11 landscapes during flowering time of maize. foraging distances (m)

Parameter	total	maize pollen	other pollen
N	614	125	489
Min	14	27	14
Max	4439	3040	4439
Median	701	432	781
Mean	820	589	879
SE	23	41	27
95 %	1788	1456	1884

5 Season and landscape composition affect pollen foraging distances and habitat use of honey bees

This chapter is published as: Nadja Danner, Anna Maria Molitor, Susanne Schiele, Stephan Härtel, Ingolf Steffan-Dewenter (2016). Season and landscape composition affect pollen foraging distances and habitat use of honey bees. Ecological Applications 26:1920-1929.

Key-words: *Apis mellifera* L., semi-natural habitat, resource competition, textit-Brassica napus, oilseed rape, pollination, land-use intensification.

Summary

Honey bees show a large variation in foraging distances and use a broad range of plant species as pollen resources, even in regions with intensive agriculture. However, it is unknown how increasing areas of mass-flowering crops like oilseed rape (OSR) or a decrease of semi-natural habitats (SNH) change the temporal and spatial availability of pollen resources for honey bee colonies, and thus foraging distances and frequency in different habitat types. We studied pollen foraging of honey bee colonies in 16 agricultural landscapes with independent gradients of OSR and SNH area within 2 km and used waggle dances and digital geographic maps with major land cover types to reveal the distance and visited habitat type on a landscape level. Mean pollen foraging distance of 1347 decoded bee dances was 1015 m ($\pm$ 26 m). In spring, increasing area of flowering oilseed rape within 2 km reduced mean pollen foraging distances from 1324 m to only 435 m. In summer, increasing cover of SNH areas close to the colonies (within 200 m radius) reduced mean pollen foraging distances from 846 to 469 m. Frequency of pollen foragers per habitat type, measured as the number of dances per hour and hectare, was equally high for SNH, grassland and OSR fields, but lower for other crops and forests. In landscapes with a small proportion of SNH a significantly higher density of pollen foragers on SNH was observed, indicating that pollen resources in such simple agricultural landscapes are more limited. Overall, we

conclude that semi-natural habitats and mass-flowering crops can reduce foraging distances of honey bee colonies at different scales and seasons with possible benefits for the performance of honey bee colonies. Further, mixed agricultural landscapes with a high proportion of SNH reduce foraging densities of honey bees in semi-natural habitats and thus possible competition for pollen resources.

Introduction

The honey bee, *Apis mellifera* L., has been managed for centuries in agricultural landscapes and provides substantial pollination for many crops and wild plant species (Klein et al., 2007). Today pollinators are globally under decline which is associated with agricultural intensification (Potts et al., 2010). The resulting loss of semi-natural habitats providing nesting sites is considered a major threat for wild pollinators (Brown and Paxton, 2009). Other effects like lack of floral resources or increased pesticide use concern both wild pollinators and managed honey bees. Potential drivers of honey bee colony losses in particular can be grouped into pests and pathogens, environmental stressors and lack of genetic diversity and vitality (Neumann and Carreck, 2010; Potts et al., 2010). Environmental stressors, e.g. malnutrition or exposure to agrochemicals, are linked to agricultural intensification and reduced landscape heterogeneity. Increasing areas of mass-flowering crops like oilseed rape (OSR) and a decrease of semi-natural habitats (SNH) as multifloral resource change the temporal and spatial availability, quality and diversity of food resources for honey bee colonies in fragmented agricultural landscapes. However, the consequences for pollen foraging of honey bees in these fragmented landscapes, the influence of landscape composition and the relevance of specific habitat types are mainly unknown.

Pollen is the only protein source for honey bees and essential for colony growth and development (Haydak, 1970). Its quality and diversity has been identified as an important factor for honey bee health (Alaux et al., 2010; Di Pasquale et al., 2013). In contrast to nectar, it is stored in only small amounts within the hive (Seeley, 1995). To secure pollen supply during breeding, a continuous availability of resources, presumably ascertained by semi-natural habitats or weeds in arable landscapes is required (Requier et al., 2015).

In intensive farming systems crops can contribute significantly to pollen harvest of honey bee colonies (Odoux et al., 2012). The influence of mass-flowering crops like oilseed rape on pollen foraging of honey bees is not well known, though. Honey bees are thought to be tightly associated with mass flowering crops like OSR (Rollin et al., 2013). They collect huge amounts of OSR nectar and are important pollinators for winter oilseed rape (Stanley, Gunning, and Stout, 2013). However, the probability of pollen collection on OSR seems to be low (Odoux et al., 2012; Woodcock et al., 2013; Garbuzov et al., 2015; Requier et al., 2015), but

was never systematically analyzed while taking account of the influence of landscape composition. In general honey bees are attracted by mass-flowering crops but little is known about the use of SNH relative to OSR (Rollin et al., 2013).

Foraging is an energy consuming effort for the honey bee and it can be assumed that shorter flight distances are of advantage for colony health and development. Foraging distances vary depending on landscape structure, resource availability and season (Steffan-Dewenter and Kuhn, 2003), but few studies analyzed pollen foraging distances in particular (Waddington et al., 1994; Steffan-Dewenter and Kuhn, 2003; Danner, Härtel, and Steffan-Dewenter, 2014; Couvillon et al., 2015). Foraging distances might be reduced and therefore colony fitness increased in landscapes with a high amount of floral resources, provided by semi-natural habitats or mass-flowering crops (Requier et al., 2015). Further pollen supply in the landscape depends on season and landscape structure (Steffan-Dewenter and Kuhn, 2003). Intensively managed landscapes with mass-flowering crops and little or no SNH often provide only pollen for short time periods, but see Requier et al. (2015). Larger foraging distances as a result of resource scarcity might arise to meet the pollen demand of honey bee colonies. The development of honeybee colonies follows seasonal stages showing a strong growth in spring with a high demand for pollen which turns regressive in summer when the swarming period ends (Seeley, 1995). With the seasonal changes in pollen supply and demand there might be temporal differences in the influence of landscape composition on honey bee habitat preferences and foraging distances.

In this study we took the perspective of honey bee colonies by decoding waggle dances to test (1) the influence of the amount of OSR and SNH in a landscape on pollen foraging distances and (2) the frequency of pollen foraging on OSR compared with SNH and other land-use types. Successful foragers communicate about the location of most rewarding food resources to their nest mates using the waggle dance (von Frisch, 1965; Seeley, 1995). Distance and direction are encoded in the duration of the waggle run and in the angle of the run relative to the vertical, respectively. There is variation in these signals (Schürch et al., 2013), but decoding a large number of dances and combining them with digitalized geographical maps is a unique tool to study the landscape ecology of honey bees (Couvillon, Schürch, and Ratnieks, 2014b; Danner, Härtel, and Steffan-Dewenter, 2014; Garbuzov et al., 2015; Härtel and Steffan-Dewenter, 2014; Steffan-Dewenter and Kuhn, 2003).

In this study we aimed to quantify pollen foraging of honeybee colonies in relation to landscape context. We selected 16 landscapes with independent gradients in the percent cover of OSR fields and SNH to answer the following questions: (1) Do pollen foraging distances depend on the amount of OSR and SNH within the main flight range? (2) Are there seasonal differences in the influence of landscape composition on pollen foraging distances? (3) Do honeybee pollen foragers prefer mass-flowering OSR compared to mixed floral resources in SNH? (4)

TABLE 5.1: Land-use characteristics for 16 landscapes in a 2000 m radius around observation hives. Means ± standard errors and ranges (%) are shown. Crop total includes oilseed rape.

Habitat type	Mean ± SE	Range
Oilseed rape	5.6 ± 1.0	0.2-12.6
Seminatural habitat	4.6 ± 1.2	0.2-14.3
Crop total	66.2 ± 4.0	37.8-93.3
Forest	16.7 ± 3.0	0.1-33.6
Settlement	7.1 ± 0.8	2.3-13.8

How does landscape composition influence pollen foraging frequency of honey bees in different habitat types?

Materials and Methods

Study region and sites

The study region is situated in a 40 km radius around Würzburg, Germany. The landscape is dominated by agriculture with cultivation of wheat and barley (together about 50% of agricultural land). Cultivation of mass flowering oilseed rape (OSR) accounts for about 8% of agricultural land. Intensive wine-growing is established on sun-exposed hills next to the river Main. Seminatural habitats (SNH) are present at varying extent and typically represented by flower rich calcareous grassland, extensive meadows and hedges. We selected 16 circular landscapes with 2 km radius and a minimum distance between landscape centres (observation hive position) of 4 km (Figure 5.1). Landscape selection aimed to maximize an OSR area gradient and an independent SNH area gradient over all landscapes (Pearson's product-moment correlation; R = -0.16, N = 16, P = 0.55). OSR and SNH area gradients reached from 0 to 13 percent and 0 to 14 percent of total landscape area, respectively (Table 5.1). The distance of OSR/SNH from the honey bee colony could also influence foraging on the respective food source. We assured that the distance to the nearest OSR field/SNH from the landscape centre was correlated with OSR and SNH area within 2 km, respectively (Pearson's product-moment correlation; OSR: R = -0.59, N = 16, P = 0.016; SNH: R = -0.50, N = 16, P = 0.048). Distance to the nearest field ranged from 52 to 1752 m for OSR and 0 to 1126 m for SNH. The calculation of areas and landscape selection was based on individual mapping in combination with digital land-use maps (provided by Bayerische Landesvermessungsverwaltung) and processing in ArcGIS 10 (ESRI, 2011).

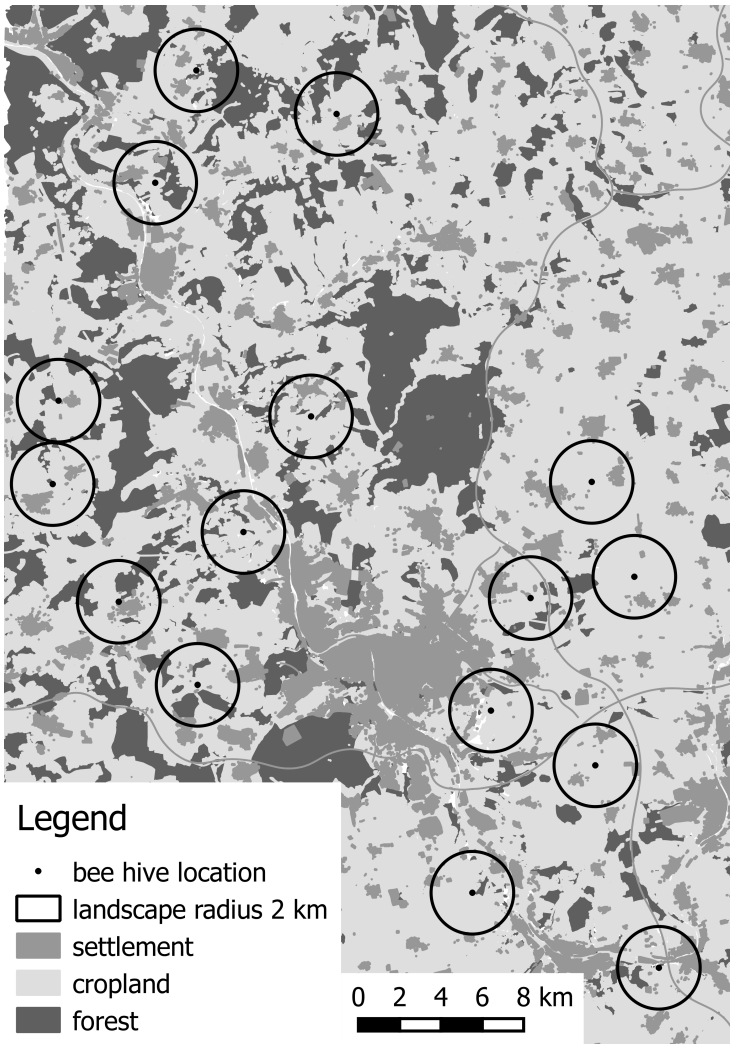


FIGURE 5.1: Geographical distribution of 16 study sites (landscapes, 2 km radius each) around Würzburg, Germany.

Landscape variables

OSR fields were individually mapped within the study region in winter/ early spring 2012 and digitalised into vector data. SNH vector data were obtained from the Bavarian mapping of biotopes provided by the Bavarian State Office for the Environment. For our purpose SNH included calcareous grasslands, species rich extensive meadows, natural hedges, initial shrubs, grassland fallows and orchards. OSR and SNH area were each calculated on the 5 spatial scales 200 m, 500 m, 900 m, 1400 m, and 2000 m radius around the landscape centers. The choice of maximum radius was based on a previous study by Danner et al. (2014) where over 96% of dances indicated foraging locations within 2000 m. In our case almost 90% of foraging locations were within the mapped landscape radius of 2000 m proving that it is a suitable scale for landscape analysis of honey bee foraging. The correlation between variables on all scales (within and between scales) was always $R < 0.5$ (Pearson's product-moment correlation). For the analysis of foraging frequency on different land use types vector data for other cropland (excluding OSR), grassland, forest and settlement within the study region was obtained from Bayerisches Landesvermessungsamt.

Observation hives and colonies

Sixteen glass-sided observation hives, each with two Zander brood frames were used for observations. 16 colonies were built using artificial swarms with young, mated queens (*Apis mellifera carnica*). All queens derived from a single mother and were mated at the same queen-mating station (Gehlberg, Germany) to assure minimal genetic differences between colonies. Brood frames with similar amounts of brood cells, honey and pollen stocks were transferred to each observation hive along with a queen and approximately 4000 worker bees. Returning foragers were restricted to perform their dances on one side of the comb by a diagonal wooden block in the passage at the bottom of the observation hive. A tent of lightproof cloth prevented light induced confusion of bees during observations (Danner, Härtel, and Steffan-Dewenter, 2014).

Sampling design

One observation hive was placed during spring and summer 2012 in the center of each of the 16 landscapes (Figure 5.1). In total we performed seven observation rounds, divided into a block of four rounds during OSR bloom (season spring, 18/04/12 to 24/05/12), and a second block of three rounds after bloom of OSR (season summer, 16/07/12 to 20/08/12). SNH provided flower resources in both seasons and allowed us to analyze its effect in interaction with a mass-flowering crop (OSR) in spring and in the absence of it in summer, respectively. During one observation round lasting several days, pollen waggle dances of each colony were observed for 90 minutes (based on previous studies; Steffan-Dewenter and Kuhn,

2003; Danner, Härtel, and Steffan-Dewenter, 2014). After finishing one round, all hives were moved over night five landscape centers further following a fixed order that was determined by the optimal connection through paved roads. This rotation scheme ensured that bees would not find their way back to the previous landscape since distances were around 10 km minimum, and allowed data collection from seven independent colonies in each landscape (Steffan-Dewenter and Kuhn, 2003).

Observation and decoding of bee dances

Over the whole study period, observation units per landscape were relatively equally distributed over the day. For each waggle dance we monitored the duration of a series of circuits, the correspondent number of circuits, the average angle of the waggle runs relative to the vertical and the time of day. We decoded only dances of pollen carrying foragers with a minimum of five consecutive circuits (Couvillon et al., 2012). The sun azimuth relative to the north was added to the waggle run angle to determine the direction of the pollen foraging site indicated by each dance. Flight distance (y) was calculated via the mean duration of a single dance circuit (x) according to a third-order polynomial fit ($y = 92.137 - 346.659 * x + 228.454 * x^2 - 10.963 * x^3$) based on data presented by von Frisch (1965), previously used by Steffan-Dewenter and Kuhn (2003), Waddington et al. (1994), Beekman et al. (2004), and Beekman and Ratnieks (2000). The locations of pollen foraging sites were plotted into land-use maps using the software ArcGIS 10 (ESRI 2011). Dance frequency for different land-use types was calculated as number of dances per hour observation time and per hectare of the corresponding land-use area (Danner, Härtel, and Steffan-Dewenter, 2014). Therefore foraging locations within 2000 m around the hives were assigned to the six land-use types they directly intersected with in ArcGIS maps (OSR, other cropland, SNH, grassland, forest, settlement). A new approach, developed by Schürch et al. (2013), is based on spatial probability distributions of the resource locations advertised by dances and was recently applied for determining foraging preferences in the landscape (Couvillon, Schürch, and Ratnieks, 2014a). In our study region landscapes were characterized by small patches of focal habitats like OSR or SNH (in average 1.7 and 0.4 hectare, respectively) in contrast to a relatively large mean patch area of 130 ha in Couvillon, Schürch, and Ratnieks (2014a). Therefore, we based the analysis on original foraging locations, also matching with OSR and SNH patches, which helped to avoid a possible bias of estimated foraging locations.

Statistical analysis

Statistical analysis was performed using the open source statistic software R (R Development Core Team, 2015). Separate linear mixed models (R package lme4;

Bates, Maechler, and Bolker, 2015) were applied per season to analyze foraging distances in the spring season with OSR and SNH area as continuous explanatory variables and in the summer season with SNH area only (questions 1 and 2). To determine the most relevant spatial scale for each variable (OSR and SNH area) in spring, we performed single models for each scale and variable first and then combined both variables at their most predictive scale in one model to test their significance (Graf et al., 2005). In detail one model per scale was performed with OSR area as explanatory variable on the five scales 200 m, 500 m, 900 m, 1400 m and 2000 m radius, resulting in five models. The scale of the model with the lowest AIC value was chosen to enter the final analysis. This procedure was repeated with SNH as explanatory variable. All single scale models contained landscape and colony as random factors. We did not include random slopes for OSR and SNH at this point since they would be scale-dependent and hinder the comparison of fixed effects (the different scales) only. In the final analysis the random structures were tested with the full models considering colony and landscape as random factors and OSR and SNH area on their respective scale as possible random slopes for each factor. The random structures resulting in the lowest AIC values of the models were chosen: colony as random intercept with random slopes for OSR and SNH, and landscape as random intercept. Further model simplification was done by stepwise reduction of fixed effects and Likelihood ratio tests. Foraging distance data were log-transformed in order to achieve normal distribution and homogeneity of variance of the model residues. Interaction between the predictors was tested. Chi^2 and P values (significance level 0.05) were calculated by Likelihood ratio tests. The difference of the slope from zero was tested post hoc. The whole procedure of scale determination and modelling was repeated for season summer considering only SNH as explanatory variable. The final random structure contained colony as random intercept with random slope for SNH, and landscape as random intercept.

Linear mixed models were performed to analyze log transformed dance frequency in each season (questions 3 and 4). Explanatory variables were land use type, SNH area, OSR area and the one-way interaction with SNH and OSR area in spring. Land use type, SNH area and the interaction between both entered the model for summer season. Post hoc comparisons after Tukey and "BH" corrections (Benjamini and Yekutieli, 2001) were applied. Both models contained landscape and colony as random factors. In all cases model diagnostic plots were checked for validity of the full model.

TABLE 5.2: Seasonal effects of OSR area, SNH area, and land use type on foraging distances and frequencies. Results from Likelihood ratio tests between mixed effects models with and without the respective explanatory variable or interaction (stepwise backward testing) are shown.

Explanatory variable, by response and season	**Chi** ^{2}	**df**	P
Foraging distance			
Spring			
OSR area	4.04	1	0.044
SNH area	0.16	1	0.69
Summer			
SNH area	4.86	1	0.027
Foraging frequency			
Spring			
Land use type	106	5	<0.001
Land use type × OSR area	13.61	5	0.018
Land use type × SNH area	19.61	5	0.001
Summer			
Land use type	41.27	4	<0.001
Land use type × SNH area	15.22	4	0.004

Results

Pollen foraging distances

Pollen foraging distances of 1347 observed and decoded bee dances ranged between 35-9510 m with a mean of 1015 m (± 26 s.e.m.). 90% of all dances were within 2193 m and 75% were within 1355 m around the hives (Figure 5.2). The influence of OSR and SNH area on pollen foraging distances was analyzed on the five scales 200 m, 500 m, 900 m, 1400 m and 2000 m. The most predictive scales (chosen via comparing AIC values of single scale models, see Supplement Table 1) of explanatory variables were 200 m for SNH and 2000 m for OSR in spring and 200 m for SNH in summer. Differences of models between scales were in general small ($\Delta AIC \leq 2$), indicating the absence of a strong scale effect. In spring (n = 940 dances) the OSR area within 2000 m influenced overall pollen foraging distances most significantly (Figure 5.3A; for statistics see Table 5.2), whereas the area of SNH had no effect. Distances decreased with increasing OSR area from 1324 m to 435 m (post hoc test for slope: P = 0.045). There was no interaction between SNH and OSR. In summer (N = 407 dances) SNH area within 200 m around the hive reduced mean pollen foraging distances from 846 m to 469 m (Figure 5.3B; post hoc test for slope: P = 0.024).

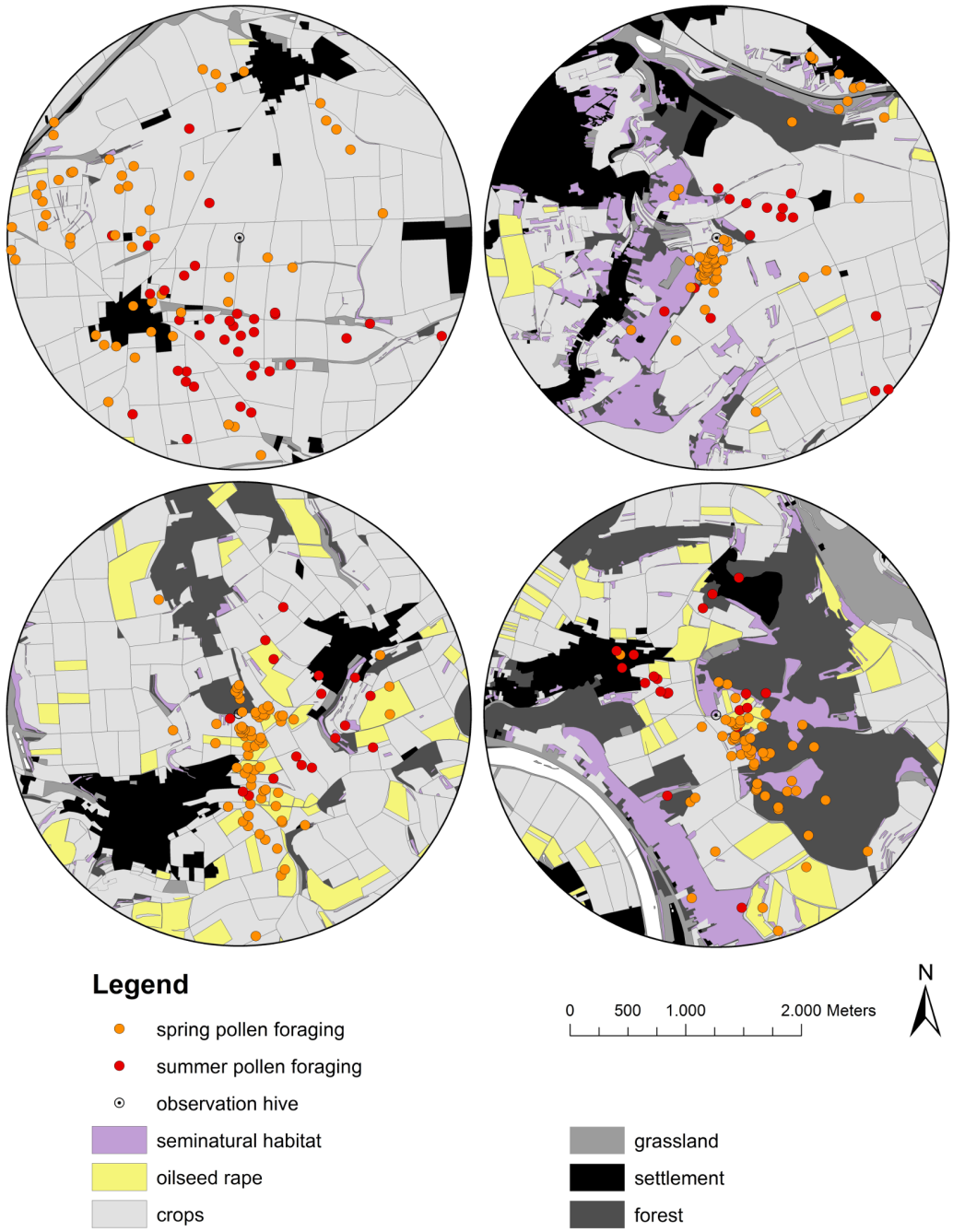


FIGURE 5.2: Pollen foraging locations in four exemplary landscapes out of 16, representing the independent gradients of oilseed rape (Brassica napus) and semi-natural habitat area. Foraging locations derive from decoded waggle dances that were observed in observation hive colonies placed in the landscape centers.

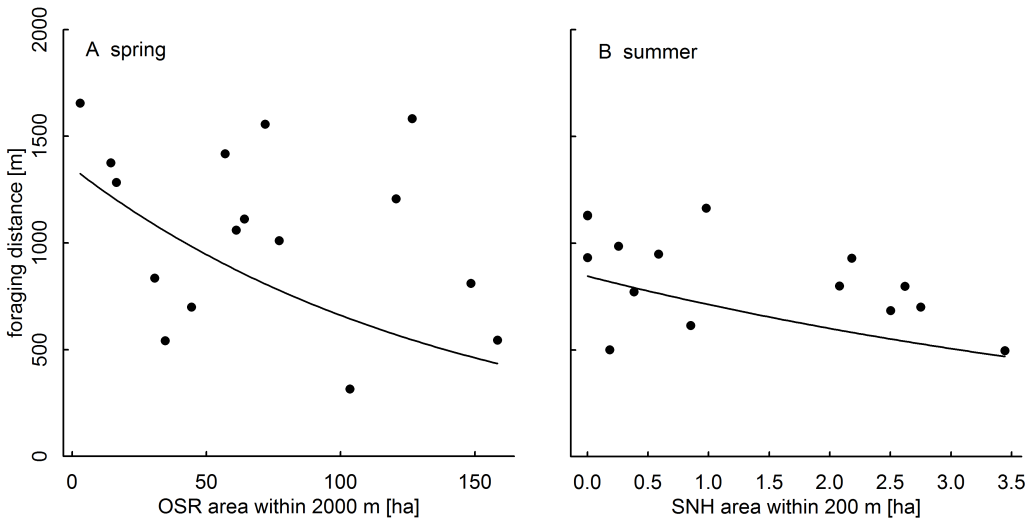


FIGURE 5.3: Pollen foraging distances in relation to OSR and SNH area in 16 landscapes in spring (A) and summer (B). Displayed dots show distance means per landscape while fitted lines derive from analysis of original data (N = 940 dances and N = 407 dances in spring and summer, respectively), back-transformed for plotting.

Pollen foraging frequency in different habitat types

We assessed the value of different habitat types as pollen resources for honey bee colonies by pollen foraging frequency measured as number of dances per hour and hectare (Danner et al., 2014). There were significant differences between land use types within seasons (Figure 5.4; for statistics see Table 2). In spring OSR was as frequently foraged as other pollen sources like SNH and grassland. In summer the number of dances per observation unit decreased in general. Settlement was as frequently foraged as SNH and grassland. Other crops and forest played a subordinate role in both seasons regarding number of dances per hour and hectare. We found a significant interaction between habitat type and SNH and OSR area within 2000 m (Figure 5.4). Post hoc multiple comparisons of slopes revealed that foraging frequency on SNH significantly increased with decreasing SNH area in both seasons (spring: $P < 0.001$; summer: $P = 0.002$). For example, decreasing SNH area from 100 ha to 10 ha would increase dance frequency in spring by 2.5 times and in summer by almost 4 times. Further, with increasing OSR area in spring, foraging frequency on forest increased ($P = 0.027$).

Discussion

Our results show that a mass-flowering crop (OSR) and semi-natural habitats (SNH) influence pollen foraging distances on different scales and depending on season. Pollen foragers did not prefer mass-flowering oilseed rape compared to

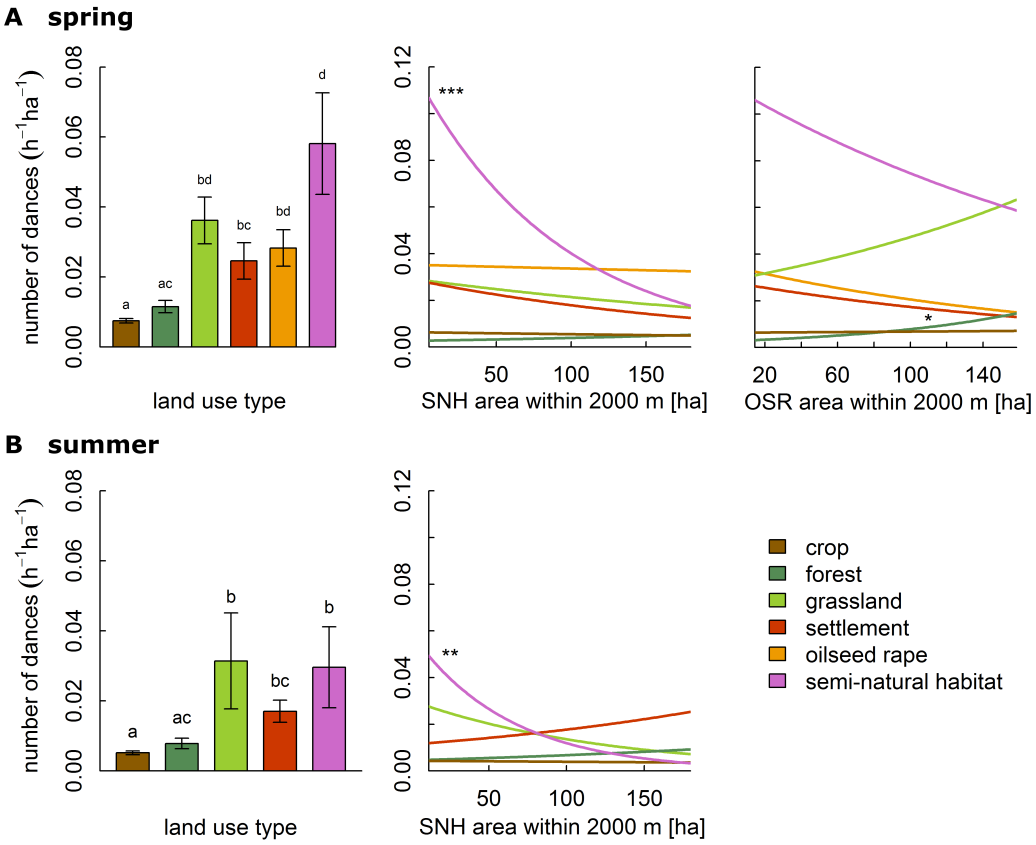


FIGURE 5.4: Effects of semi-natural habitats (SNH) on pollen foraging for different land use types in spring (A) and summer (B). Bar plots show mean dance frequencies ± standard error of the mean (SEM) measured as number of dances per hour and hectare. Different lowercase letters above the bars indicate significant differences ($P < 0.05$). Other plots show the effect of SNH and oilseed rape (OSR) area within 2000 m on dance frequency. Crops in spring excluded OSR, in summer included former OSR fields. Asterisks indicate slopes significantly different from zero (* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; see Table S2 for post hoc tests for slopes). Model estimates were back transformed for plotting fitted lines.

mixed floral resources in semi-natural habitats, whose importance as pollen resource was influenced by landscape composition. Foraging frequency on SNH increased in spring and in summer with decreasing SNH area in the landscape. Our findings have general implications for the understanding of honey bee foraging in agricultural landscapes and for agri-environmental management to foster pollen resources for bees. While our results underpin the importance of semi-natural habitats as pollen sources for honeybees, they also indicate a strong increase of foraging frequency in simple agricultural landscapes, possibly resulting in resource competition with wild bees.

Pollen foraging distances and landscape composition

Mean pollen foraging distances in this study (1 km) are in the range of previous reported distances of 0.7 km (suburban landscape; Waddington et al., 1994), 0.8 km (agricultural landscape; Danner, Härtel, and Steffan-Dewenter, 2014), 1.1 km (suburban/agricultural landscape; Couvillon et al., 2015) and 1.6 km (agricultural landscape; Steffan-Dewenter and Kuhn, 2003). Foraging distances of honey bees are variable and depend on temporarily occurring patches of food resources, landscape structure and dominant habitat types (Steffan-Dewenter and Kuhn, 2003; Visscher and Seeley, 1982). In intensively used landscapes mass-flowering crops like OSR in particular might reduce foraging distances since they massively enhance the availability of resources – at least for a short time period, but empirical data are mainly lacking so far. Our study demonstrates that in spring OSR area within 2000 m around the hives reduced overall pollen foraging distances (Figure 5.4A). Lower distances in landscapes with high OSR area confirm that OSR has the potential to reduce pollen foraging effort in terms of flight distances. However, our study and others indicate that oilseed rape is not as preferentially used as expected by its dominance (Odoux et al., 2012; Woodcock et al., 2013; Garbuzov et al., 2015; Requier et al., 2015).

SNH area was a relevant factor for pollen collection on smaller scales than 2000 m. Our results show that SNH areas close to the hive (within 200 m) reduced overall pollen foraging distances and consequently foraging effort in summer. The effect of SNH might differ depending on specific plant communities. However, our study provides more general information about the significance of SNH in agricultural landscapes. Comparing single scale models during the scale selection and modelling procedure (see statistics section) reveals small differences between models ($\Delta AIC < 2$; Supplement Table 5.S1), suggesting that the influence of SNH does not strongly depend on its scale. It is therefore most relevant on small scales like 200 m but probably still important on larger scales as well. Similarly, in a study with African honey bees foraging ranges increased during a shortage of resources (Schneider and McNally, 1993). A larger foraging range of a colony is assumed to raise the probability of discovering more suitable food source patches (Visscher and Seeley, 1982). Shorter flight distances in turn should

point out that the available resources like OSR and SNH are suitable for honey bee colonies. Steffan-Dewenter and Kuhn (2003) found no significant differences of overall mean foraging distances between simple (assumed to supply resources less continuous) and complex landscapes. However, pollen foraging distances were significantly higher in simple landscapes after OSR bloom. Assuming that a simple landscape would most likely correspond to one with low SNH area in our study, this is in accordance with our results.

OSR raises the pollen offer in the landscape for a short time period in spring and might reduce energy effort of honey bee colonies. In contrast a landscape with SNH provides pollen continuously with positive influence on energy effort. Requier et al. (2015) suggest that honey bees use a wide variety of pollen resources, even during OSR flowering, in order to ensure colony health. According to our observed effect of SNH area on foraging distance this is of relevance especially in summer when resources in our studied landscapes tend to be scarce compared to spring. In summer, profitable resources (comparable to OSR in spring and besides SNH) would be sunflower and maize that were barely grown in these landscapes. The missing effect of SNH area on foraging distances in spring might be due to the higher demand for pollen in the stage of colony growth. Apparently even high amounts of SNH close to the hive do not meet the complete demand for pollen (quantity or quality/diversity) in spring, resulting in high foraging distances. A poor availability of pollen resources might have negative consequences for colony development and health (Keller, Fluri, and Imdorf, 2005).

Very few published studies quantitatively measure colony success in response to landscape variables (Sponsler and Johnson, 2015). Besides foraging distances as a measure of colony success, honey production or colony size can be related to landscape variables, e.g. colony size was positively correlated with forest land cover in an intensively managed agricultural landscape in France (Odoux et al., 2014). A study performed in Denmark reveals that colonies situated in agricultural landscapes were significantly less productive than colonies situated in urban areas (Lecocq et al., 2015), pointing out the poor suitability of certain agricultural landscapes for honey bee colonies. To overcome obvious shortages in pollen resources after periods of mass flowering crops and in landscapes with low SNH, a suitable landscape management is needed. Accounting for about one third of foraging flights (Fewell and Winston, 1992) makes pollen foraging a relevant part of energy effort that can be shaped by landscape composition. We emphasize the importance of semi-natural habitats in the neighborhood of honey bee colonies. Further, the establishment of flower strips, organic farming and diversifying flowering crops are possible agri-environmental schemes for a pollinator supporting landscape management (see Decourtye, Mader, and Desneux (2010) for a review). However, to answer questions like how much resources in the agricultural landscape are actually needed for optimal colony development depending on season, further studies are needed.

Pollen foraging frequeny and landscape composition

To gain insight into the intensity of resource use on different land use types in different seasons we analyzed dance frequencies, i.e. the number of dances (resource locations advertised by dances) per hour and hectare (Danner, Härtel, and Steffan-Dewenter, 2014). Overall, we observed a lower dance frequency in summer compared to spring. Since we controlled for colony size throughout the experiment by removing bees when necessary, there was no natural colony growth that could have influenced dance frequency. The colony regulates its pollen foraging intensity in accordance with changing colony needs which derive from the amount of stored pollen and the demand of adult bees and larvae (Camazine, 1993). A lower dance frequency in summer might therefore reflect a decreasing demand for pollen after the intensive larval rearing in spring.

Comparing dance frequencies for different land-use types revealed that grassland, settlements (gardens) and SNH were frequently foraged for pollen regardless of the season. OSR is a rewarding pollen resource in spring, which is underlined by its influence on pollen foraging distances, but OSR was not as preferentially visited as could be assumed for a mass-flowering crop. Oilseed rape is known to be foraged intensively for nectar (Nedić et al., 2013) while there is only little known for pollen. In a study in an agricultural landscape in western France OSR pollen never represented more than 29% pollen weight of a weekly pollen load sample while other single weed species represented up to 98% (Odoux et al., 2012). Although it is not known whether pollen input from a certain plant species correlates with its related dance frequency, this result indicates that OSR pollen is not a dominating species in the pollen harvest of honey bees. In accordance (Woodcock et al., 2013) and (Garbuzov et al., 2015) report a low probability of pollen foraging on OSR, based on observed visitation rates in OSR fields and dance decoding, respectively. While the study of Garbuzov et al. (2015) was performed in a single landscape, our study allows the independent comparison of OSR and SNH landscape gradients. Similarly, a preference of other pollen than oilseed rape in intensive farmlands was indicated by analyzing pollen samples from traps in front of the hive (Requier et al., 2015). This is also consistent with general low proportions of OSR pollen in nests of solitary bees (Holzschuh et al., 2013). The frequent utilization of SNH and grassland in spring and summer suggests that honey bees try to recruit nest mates for collecting a diverse pollen diet throughout the year, instead of only concentrating on mass-flowering crops. However, the influence of mass flowering crops on pollen foraging pattern appears to be crop and season specific. In a previous study analyzing pollen foraging in landscapes during bloom of maize we detect a different pattern with a significantly higher pollen foraging frequency on mass-flowering maize fields compared to all alternative habitat types including grassland and settlement (Danner, Härtel, and Steffan-Dewenter, 2014). In temperate regions, maize is in bloom during late summer when alternative resources are scarce compared to spring which

might also be a cause for the observed difference.

As a further new aspect in our study we incorporated the effect of landscape context on foraging frequencies. Foraging on SNH was concentrated in both seasons when SNH area within the landscape decreased, while this effect was absent for other land use types as well as for the influence of OSR area in spring. Interestingly, OSR area had a positive but relatively small effect on foraging on forest. Since foraging frequency on forest was generally very low compared to other land use types, this effect could be biologically less important. However, Requier et al. (2015) report that woody plant species (which are only partly forest species) are important pollen resources during OSR bloom. In order to clarify a possible relationship between OSR and foraging in forest habitats individual studies should be designed. Nevertheless, our results showed the importance of SNH as a pollen resource for honey bees especially when its proportion in the landscape is rather limited (Steffan-Dewenter et al., 2002). Furthermore, as wild bees are known to prefer SNH (Rollin et al., 2013), the interspecific competition for pollen resources between honey bees and wild bees might be more severe in such landscapes (Holzschuh et al., 2013; Härtel and Steffan-Dewenter, 2014), an aspect not considered in the few studies on honey bee – wild bee competition (Steffan-Dewenter and Tscharntke, 2000; Hudewenz and Klein, 2013). Integrating SNH in landscape management would benefit honey bees but also other pollinators like wild bees (Rollin et al., 2013) or butterflies (Ockinger and Smith, 2007). Rollin et al. (2013) suggest to support different bee groups with distinct management strategies, e.g. promoting mass-flowering crops to benefit honey bees and semi-natural herbaceous habitats to benefit wild bees. We emphasize the cross-group effectiveness and importance of SNH, particularly regarding the continuous provision of diverse pollen resources. In conclusion, our study provides clear evidence that landscape composition has an influence on foraging distances and consequently energy effort of honey bee colonies. SNH in the vicinity to honey bee colonies reduced foraging distances in summer and was one of the most important pollen sources with even increasing importance in landscapes with low SNH cover. While OSR is an important nectar source it was not an outstanding pollen forage which suggests that diverse pollen sources within the foraging range are very important for honey bee colonies. Maintaining semi-natural habitats and promoting agri-environmental schemes to provide diverse and continuously available pollen resources will improve the wellbeing of honey bee colonies, the coexistence with wild bees and the provisioning of crop pollination services in agricultural landscapes.

Supplement

Table 5.S1: AIC values of single scale models with landscape and colony as random factors. In bold the lowest value for each explanatory variable OSR and

SNH. Variables entered final analysis on the scales 200 m (SNH) and 2000 m (OSR).

Scale [m]	**AIC (OSR)**	**AIC (SNH)**
200	2353.70	**2352.91**
500	2352.02	2353.10
900	2353.04	2353.06
1400	2352.16	2353.56
2000	**2351.41**	2353.69

Table 5.S2: Multiple slope comparisons for two linear mixed models analyzing dance frequency in spring and summer. Spring: explanatory variables land use (six levels crop, forest, grass, settle, osr, snh), OSR area, SNH area, interaction land use with OSR area, interaction land use with SNH area. Summer: explanatory variables land use (five levels crop, forest, grass, settle, snh), SNH area and the interaction. Random factors colony and landscape.

	Spring				**Summer**			
Linear Hypothesis	**Estimate**	**Std. Error**	**z value**	**Pr(>|z|)**	**Estimate**	**Std. Error**	**z value**	**Pr(>|z|)**
osrslope crop - 0 == 0	-0.0006223	0.0021378	-0.291	0.8411				
osrslope forest - 0 == 0	0.0106504	0.0037513	2.839	0.0271 *				
osrslope grass - 0 == 0	0.0050028	0.0070697	0.708	0.5750				
osrslope settle - 0 == 0	-0.0048626	0.0038267	-1.271	0.4077				
osrslope osr - 0 == 0	-0.0053787	0.0032879	-1.636	0.3056				
osrslope snh - 0 == 0	-0.0040502	0.0037931	-1.068	0.4090				
snhslope crop - 0 == 0	-0.0017222	0.0016850	-1.022	0.4090	-0.0009148	0.0019633	-0.466	0.64124
snhslope forest - 0 == 0	0.0036596	0.0027309	1.340	0.4077	0.0040450	0.0035316	1.145	0.42009
snhslope grass - 0 == 0	-0.0029497	0.0026933	-1.095	0.4090	-0.0077986	0.0033813	-2.306	0.05272 .
snhslope settle - 0 == 0	-0.0045290	0.0025630	-1.767	0.3056	0.0045987	0.0058007	0.793	0.53488
snhslope osr - 0 == 0	-0.0004134	0.0025186	-0.164	0.8696				
snhslope snh - 0 == 0	-0.0102964	0.0022000	-4.680	3.44e-05 ***	-0.0152008	0.0042504	-3.576	0.00174 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Adjusted p values reported -- BH method)

6 Applying DNA metabarcoding in honey bee foraging ecology – Season but not landscape structure shapes the diversity of collected pollen

This chapter is in preparation for submission as: Nadja Danner, Alexander Keller, Stephan Härtel, Ingolf Steffan-Dewenter (in prep.). Applying DNA metabarcoding in honey bee foraging ecology – Season but not landscape structure shapes the diversity of collected pollen.

Key-words: *Apis mellifera* L., next generation sequencing, landscape diversity, foraging distance, *Brassica napus*, oilseed rape

Summary

The availability of pollen in agricultural landscapes is essential for the successful growth and reproduction of honey bee colonies (*Apis mellifera* L.). The quantity and diversity of collected pollen can influence the growth and health of honey bee colonies, but little is known about the influence of landscape structure on pollen diet. In a field experiment we rotated 16 honey bee colonies across 16 agricultural landscapes, used traps to get samples of collected pollen and observed the intra-colonial dance communication to gain information about foraging distances. DNA metabarcoding was applied to analyze mixed pollen samples. Neither the amount of collected pollen nor pollen diversity were related to landscape diversity. However, we found a strong seasonal variation in the amount and diversity of collected pollen in all sites independent of landscape diversity. The revealed increase of foraging distances with decreasing landscape diversity suggests that honey bees compensated for lower landscape diversity by increasing their pollen foraging range on order to maintain pollen amount and diversity. Our results underline the importance of a diverse pollen diet for honey bee

colonies. Agri-environmental schemes aiming to support pollinators should focus on possible spatial and temporal gaps in pollen availability and diversity in agricultural landscapes.

Introduction

Pollen is the main protein source for honey bees and also provides lipids, vitamins and minerals (Haydak, 1970). The availability of pollen is essential for the growth, development and reproduction of honey bee colonies. A honey bee colony in temperate zones collects about 15–30 kg of pollen per year, almost all of which is consumed, with a reserve of about 1 kg kept in the colony at any time (Pernal and Currie, 2001; Winston, 1987). Its survival is positively influenced by the pollen amount collected in the course of the season (Smart et al., 2016). Therefore, pollen resources have to be available continuously and sufficiently to assure pollen supply for the colony. Besides quantity (Smart et al., 2016), the diversity of collected pollen is important for honey bee health (Alaux et al., 2010; Di Pasquale et al., 2013). Honey bee colonies forage on an area of more than 10 km^2 and several studies report on the diversity of collected pollen (Alves and Santos, 2014; Baum et al., 2004; Requier et al., 2015; Koeppler, Vorwohl, and Koeniger, 2007). But the way how land use change and agricultural intensification influence the amount and diversity of collected pollen remains largely unexplained. Earlier studies indicate that foraging distances of honey bee colonies increase in structurally simple landscapes with a high proportion of arable land and depend on the availability of crop-based pollen resources (Steffan-Dewenter and Kuhn, 2003; Danner, Härtel, and Steffan-Dewenter, 2014). Honey bees expanded their foraging range for pollen when there was less semi-natural habitat with a diversity of resources available (Danner et al., 2016). Landscape composition (beside landscape configuration another quality of landscape structure) can not only be characterized by the area of specific habitat types like SNH but also by landscape diversity, which evaluates richness and evenness aspects of the landscape, i.e. the number and area of different habitat types (MacGarigal and Marks, 1994; Nagendra, 2002). We assume that landscape diversity (based on the Shannon index) provides a good indication of the potential diversity of available resources. It might influence foraging distances as well as the amount and diversity of pollen diet, which has rarely been addressed (Piroux et al., 2014). Less diverse landscapes with a lower resource availability – at least outside mass-flowering crops - and diversity might lead to a lower input and diversity of pollen. However, since honey bees are able to discriminate between different pollen diets (Hendriksma and Shafir, 2016) and rely on a wide variety of resources (Odoux et al., 2012; Requier et al., 2015) they might compensate for a lower landscape diversity – in case resources are insufficient - by increasing their foraging range. Pollen availability is not only a function of landscape diversity but also subject to seasonal

fluctuations. For example the presence of mass-flowering resources in our study is mostly reflected by season. Oilseed rape is flowering in April and May while potential mass-flowering crops in summer like sunflower and maize are almost lacking in our study sites. Together with seasonal shifting demands by the colony, season could therefore influence the pollen diet also directly.

A major restriction for research on the foraging ecology of honey bees is the time-consuming identification of pollen grains via light microscopy and the limited taxonomic resolution, mostly to the family level (Keller et al., 2015). Next-generation sequencing presents a useful and efficient workflow to identify pollen at the genus and species level without requiring specialised palynological expert knowledge (Keller et al., 2015). Sequencing data can also be used for abundance estimation. A subsequent analysis of species composition and diversity (Shannon index) gives insight into the pollen resource use of pollinators and its variability. In this study we analyzed pollen foraging of honey bee colonies in differently structured agricultural landscapes across the season to answer the following questions: (1) Does the amount of collected pollen (dry weight) depend on landscape diversity or season? (2) Is the richness and diversity of pollen influenced by landscape diversity or season? (3) Do honey bees compensate for a lower landscape diversity by increasing their pollen foraging range to maintain pollen amount and diversity? (4) Which are the most abundant pollen taxa and how do they vary over time?

Materials and Methods

Study region and experimental design

This study was conducted around Würzburg, Germany, where the landscape is dominated by agriculture including cultivation of mass-flowering crops like oilseed rape. Semi-natural habitats like flower rich calcareous grasslands, extensive meadows and hedges are present at varying extent. We used 16 observation hives, each with a honey bee colony of approximately 4000 worker bees on two brood frames, for observation of waggle dances and collection of pollen samples. We selected 16 landscapes with 2 km radius each and with an overall gradient in landscape diversity. Each observation hive was set up in the center of one landscape. Seven observation rounds were performed from 18 April to 20 August 2012 (Figure 6.1), each comprising of 90 minutes waggle dance observation at each colony (lasting several days), a full day of collecting pollen loads simultaneously from all colonies, and a final rotation of all hives to be set up in another landscape for the next observation round (except for the last round). The rotation of the hives followed a fixed scheme that was logistically determined and that

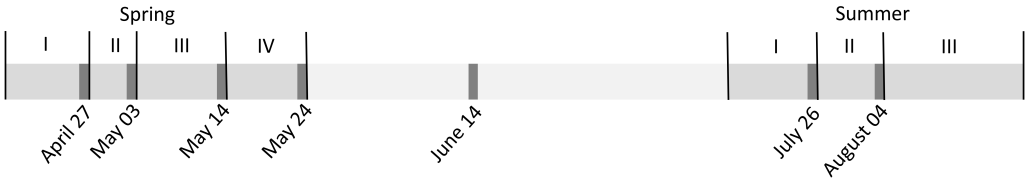


FIGURE 6.1: Timeline of the study period from 18/04/12 to 20/08/12. Seven days of pollen sampling in 16 landscapes are shown in darkgrey, seven periods of waggle dance observation in middle grey and no data collection in light grey. Spring I - Summer III refer to the seven observation rounds as described in subsection Experimental design.

assured a minimum distance of movement of 10 km for each colony, which prevented them from flying back to the previous landscape. The rotation allowed for data collection from seven independent colonies in each landscape.

Observation and decoding of waggle dances

Waggle dances of bees returning with pollen loads were observed to calculate foraging distances for pollen as described in Danner et al. (2016). The duration of a series of circuits and the correspondent number of circuits was recorded for each bee carrying pollen and dancing at least five consecutive circuits (Couvillon, 2012). Foraging distance (y) was calculated via the mean duration of a single dance circuit (x) according to a polynomial fit ($y = 92.137 - 346.659 * x + 228.454 * x^2 - 10.963 * x^3$) based on data by von Frisch (1965). Only dances of bees carrying pollen and with five consecutive circuits were decoded (Couvillon, 2012).

Landscape diversity

We measured landscape diversity via the Shannon index, which is widely used and recommended for landscape management in an ecological framework (Nagendra, 2002). It is defined as

$$SHDI = -\sum_{i=1}^{N} P_i \cdot \ln P_i$$

where N is the number of habitat types and P_i is the proportional abundance of habitat type i. Landscape diversity is independent of specific habitat types and therefore better comparable between studies. The calculation of landscape diversity for all 16 landscapes was based on the area of seven main land use types within 2 km (Table 6.1; Danner, Härtel, and Steffan-Dewenter, 2014; Danner et al., 2016). Vector data for cropland, grassland, other agricultural land, hedges, forest and settlement was obtained from Bayerisches Landesvermessungsamt. Further, vector data for semi-natural habitat was obtained from the Bavarian mapping of biotopes (Bayerisches Landsamt für Umwelt).

TABLE 6.1: Land-use characteristics for 16 landscapes in a 2000 m radius around observation hives. Means ± standard errors and ranges (%) are shown.

Habitat type	**Mean ± SE**	**Range**
Crop total	66.2 ± 4.0	37.8-93.3
Forest	16.7 ± 3.0	0.1-33.6
Settlement	7.1 ± 0.8	2.3-13.8
Seminatural habitat	4.6 ± 1.2	0.2-14.3
Other agricultural land	3.2 ± 1.3	0.0-18.1
Grassland	4 ± 0.7	0.3-10.6
Hedges	2.1 ± 0.7	0.0-9.5

Pollen sampling and analysis

Pollen sampling at all colonies was performed on six dates within the observation rounds and one additional date in June. Pollen loads were collected via pollen traps with 5mm hole grids in front of the hive entrances and frozen at -20 ° C. For further analysis pollen samples were dried at 40° Celsius and 30% relative humidity for 48 hours. The dry weight of all samples was determined. Samples were homogenized before taking sub-samples. DNA from $0.003g$ pollen grains was isolated as described by (Keller et al., 2015) using the Macherey-Nagel Food Kit (Düren, Germany) and the supplementary protocol of the kit dedicated to pollen samples.

The amplification PCR for the ITS2 marker was performed in three separate $10\mu L$ reactions in order to avoid PCR bias (Fierer et al., 2008). Primers were ITS-S2F (Chen et al., 2010) and ITS4R (White et al., 1990), but modified for sample multiplexing according to (Sickel et al., 2015).

Each reaction contained $5\mu L$ 2x Phusion Master Mix (New England Biolabs, Ipswich, MA, USA), $0.33\mu M$ each of the forward and reverse primers, $3.34\mu L$ PCR grade water and $1\mu L$ DNA template. PCR conditions were as follows: initial denaturation at 95° Celsius for $4min$, 37 cycles of denaturation at 95° Celsius for $40s$, annealing at 49° Celsius for $40s$ and elongation at 72° Celsius for $40s$; followed by a final extension step at 72° Celsius for $5min$. Triplicate reactions of each sample were combined after PCR and DNA amounts between samples normalized using the SequalPrep Normalization Plate Kit (Invitrogen GmbH, Darmstadt, Germany). Pooled multiplexed samples were quality controlled using a Bioanalyzer High Sensitivity DNA Chip (Agilent Technologies, Santa Clara, CA, USA) and quantified with the dsDNA High Sensitivity Assay (Life Technologies GmbH, Darmstadt, Germany).

Sequencing was performed on the Illumina MiSeq using 2x150 cycles v2 chemistry (Illumina Inc., San Diego, CA, USA). Raw sequence data was deposited at the European Nucleotide Archive (ENA, http://www.ebi.ac.uk/ena) with the

project accession number PRJEB15870.

Raw reads were joined using QIIME v1.8.0 (Caporaso et al., 2010) and filtered with USEARCH v8.0.1477 (Edgar, 2010) to remove low quality data (<Q20, <150 bp, ambiguous base-pairs). Reads were classified first by a global query with USEARCH with a threshold of >97% sequence identity. For this, only reference sequences of plants occurring in Bavaria according to http://www.bayernflora.de (accessed on 2015/01/24) were subset from the complete ITS2-database, covering 90.4% of genera known in this area. Sequences that did not match such references were secondly classified to the highest possible taxonomic group, but maximally genus, using UTAX as implemented in USEARCH v8.0.1477 (Edgar, 2010), the UTAX-reference database of (Sickel et al., 2015) and the scripts for parsing deposited in https://www.github.com/iimog/meta-barcoding-dual-indexing.

Data was imported into R (R Development Core Team, 2015) using the phyloseq package (McMurdie and Holmes, 2013). We filtered Chlorophyta from the dataset and transformed raw read numbers into relative amounts. Species below a minimum relative abundance of 1% per sample were removed. Six of 102 samples were omitted due to an insufficient number of reads, determined via rarefaction curves of all samples. We calculated the richness per sample as the number of different plant species pollen originated from. As a measure for diversity we calculated the Shannon index per sample based on the number of sequencing reads per plant taxon as an estimate for abundance. Pollen samples were comparable between landscapes and seasonal dates due to the standardized colony sizes throughout the experiment.

Statistical analysis

We analyzed dry weight, richness and diversity of pollen samples, each in response to season (date), landscape diversity and their interaction using the open source statistic software R (R Development Core Team, 2015). Response variables were square root transformed if necessary. Linear mixed models were applied using the package nlme (Pinheiro et al., 2015) for the analysis of dry weight and richness. Landscape was included as random factor. We checked for temporal autocorrelation within the models. A generalized least squares model with a weights function (to account for variance heterogeneity between dates) and a correlation structure (to account for replication of sites) was applied to analyze diversity. A linear mixed model was applied using the package lme4 (Bates, Maechler, and Bolker, 2015) to analyze log-transformed pollen foraging distances. Site and colony were implemented as random factors, date and landscape diversity as interacting explanatory variables. The importance of explanatory variables was evaluated for each model based on model selection via AICc comparison with the function "dredge" (multi-model inference; applied to the full model) and model averaging (Burnham and Anderson, 2004) over the full set of possible models (R package MuMIn; Barton, 2012). The importance of the variables (based on the

TABLE 6.2: Results of model selection analyzing dry weight, richness and diversity of pollen samples collected by honey bee colonies. The 95 % best models (sum of weight 0.95) per response variable are shown. Explanatory variables are D = date, F = foraging distance, L = landscape diversity, and R = observation round. Weight is the AIC weight compared to all possible models.

Response	Model specification	df	AICc	Δ AIC	weight
Dry weight	D	9	269.4	0.00	0.78
	D + L	10	271.9	2.48	0.22
Richness	D	9	203.1	0.00	0.62
	D + L	10	204.1	0.99	0.37
Diversity	D	15	86.7	0.00	0.76
	D + L	16	89.0	1.96	0.24
Distance	R + L + R:L	17	3222.8	0.00	0.97

sum of weights; Burnham and Anderson (2004)) are given in the text, results of "dredge" in Table 6.2. Only important variables and interactions (importance $>$ 0.5) were included in the final model. Results of post hoc comparisons after Tukey with "BH" corrections (Benjamini and Yekutieli, 2001) are shown in figures. In all cases model diagnostic plots were checked for validity of the full model.

Results

Dry weight of pollen samples

Dry weight of daily pollen samples ranged from 1.7 to 39.3 g with a mean of 14.7 g ($\pm$ 0.8 s.e.m.) over the whole study period. In contrast to expectations the amount of collected pollen per day did not depend on landscape diversity (importance: date 1, landscape diversity 0.23, interaction $<$0.01; Figure 6.2; Table 6.2; Table S1). However, we observed a significant seasonal variation in the amount of collected pollen with highest values in April and May (for results of post hoc test see Figure 6.2). We also checked whether seasonal variation in the amount of collected pollen varied more in simple compared to diverse landscapes, but the interaction between landscape and season was not important.

Richness and diversity of pollen samples

DNA sequencing of the mixed pollen samples allowed us to determine the plant species from which pollen originated. In total, we generated 1529901 quality filtered sequencing reads, with an average throughput of 15936 reads per sample ($\pm$ SD 11775). We detected 80 taxa from 56 genera with an abundance of DNA reads $>$1% per sample (N = 102; Figure 6.3; Table S5). 75 taxa could be determined to species level and five to genus level. The species richness of pollen

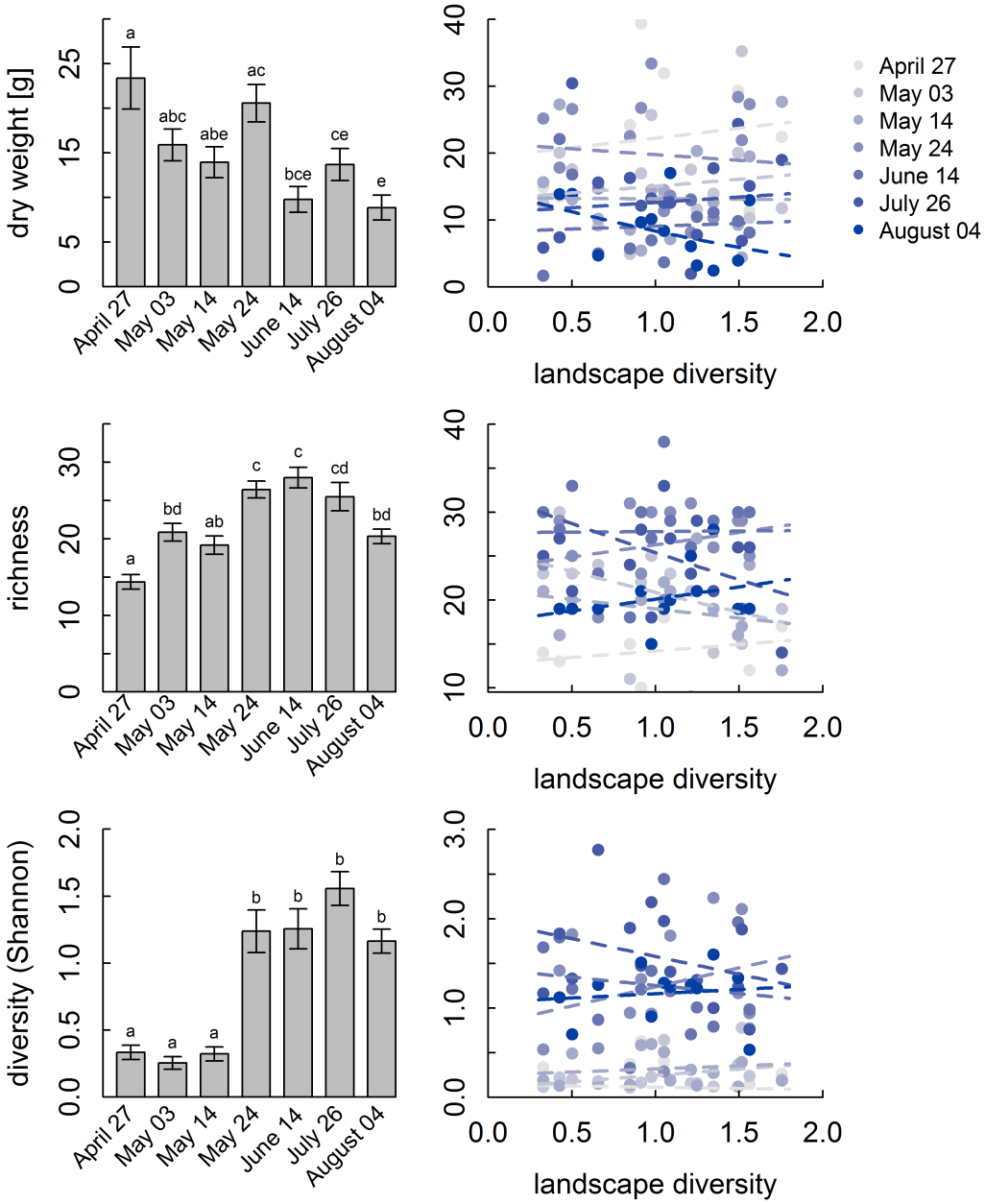


FIGURE 6.2: Dry weight, richness and diversity of pollen samples, collected by honey bee colonies on seven dates in 16 landscapes. Different letters above bars indicate significant differences ($P < 0.05$). The respective relations to landscape diversity in interaction with date are shown on the right and are not significant.

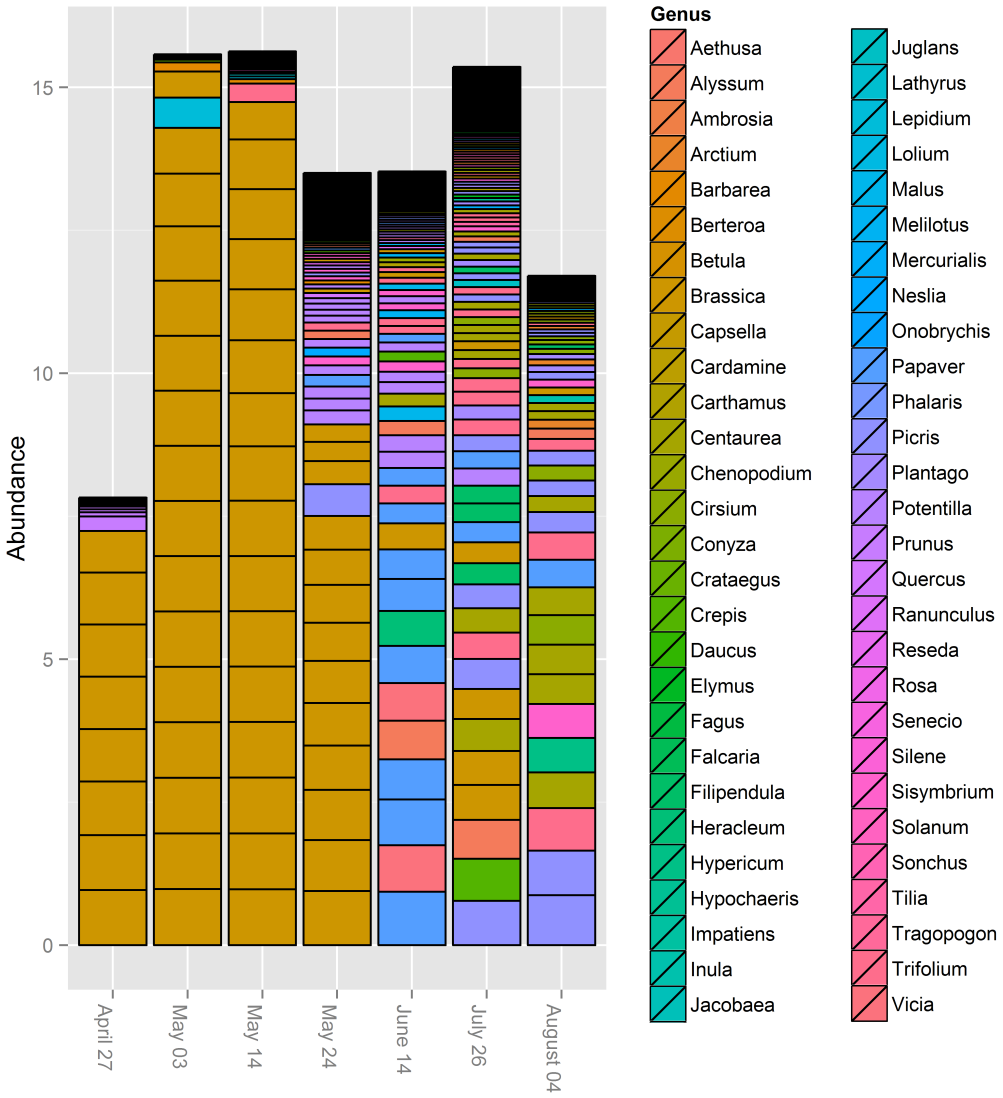


FIGURE 6.3: Genera in pollen samples at 7 dates (figure has to be replaced yet with better readable x axis and legend). Every single box represents one genus within one sample. The y-axis shows the relative abundance over all samples.

samples ranged from 9 to 47 species per sample with a mean of 23 ($\pm$ 0.6 s.e.m.). Pollen richness did not depend on landscape diversity and its interaction with date (Figure 6.2; Table S2). Again, season had a strong effect on pollen richness (importance: date 1, landscape diversity 0.38, interaction $<$0.01; Table 6.2) with significantly lower values in April and first half of May compared to end of May, June and July. Similarly, pollen diversity (Shannon index) was independent from landscape diversity but significantly varied with date (importance: datef 1, landscape diversity 0.22, interacion $<$0.01; Table S3). It ranged from 0.11 - 2.77 and was significantly lower in April and first half of May compared to all other dates.

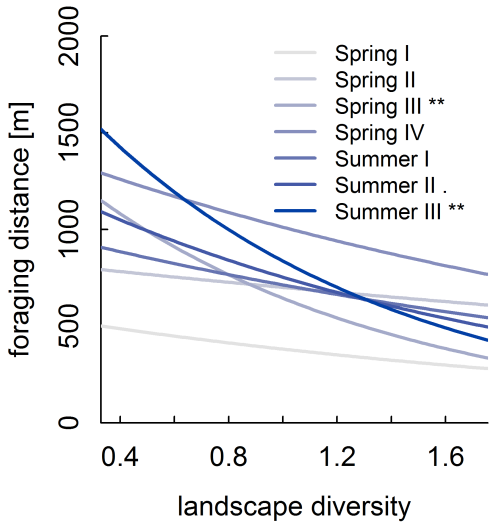


FIGURE 6.4: Pollen foraging distances of honey bees in relation to landscape diversity. Data were collected within eight observation rounds in 16 landscapes. Slopes signicantly different from zero: . $P < 0.1$; ** $P < 0.01$. For details about dates, see Figure 6.1.

Pollen foraging range

We hypothesized that honey bee colonies would increase their foraging range in landscapes with low habitat and thus floral resource diversity to maintain the amount and diversity of incoming pollen constant. Indeed, foraging distances for pollen significantly increased with decreasing landscape diversity, depending on date (importance: date 1, landscape diversity 1, interaction 0.97; for results of post hoc test see Figure 6.4; Figure S1 and Table S4).

Dominant taxa

The relative abundance of sequencing reads within a sample was used as an approximation of estimated pollen quantity of a specific taxon. Overall pollen from *Brassica napus* (48%) was most abundant followed by *Papaver rhoeas* (6%), *Picris hieracioides* (6%), *Centaurea jacea* (5%) and all other species with <5% abundance. On genus level *Brassica* (49%) was followed by *Papaver* (6%), *Picris* (6%), *Trifolium* (5%) and *Centaurea* (5%). On family level *Brassicaceae* (55%) were followed by *Asteraceae* (16%), *Fabaceae* (7%), *Rosaceae* (7%) and *Papaveraceae* (6%). *Brassica napus* was very dominant in April and May (86% $\pm$ 0.06 s.e.m.) and partly accompanied by *Prunus avium* in April or *Potentilla spec.* end of May. In June, July and August similar dominant genera were lacking. *Papaver rhoeas* was the most abundant species in June (35%), *Picris hieracioides* in July (21%) and *Centaurea jacea* in August (25%).

Discussion

DNA metabarcoding of mixed pollen samples collected by honey bee colonies revealed interesting insights into honey bee foraging in agricultural landscapes. We could not find any relationship between dry weight, richness and diversity of pollen samples with landscape diversity and its interaction with date. However, in all cases we detected a dependency on season (date). Importantly, our results indicate that honey bees compensated for less diverse landscapes by increasing their foraging range to maintain the amount and diversity of collected pollen.

Comparing the amount (dry weight) of collected pollen with data from other studies we have to take colony size into account. The average of 15 g pollen per day for colonies with an average size of 4000 workers in our study would approximately correspond to daily 75 g for colonies of 20000 workers if simply multiplied. Such colonies have a daily pollen intake of approximately 30 g (Danner et al., unpublished) and the collected amount of pollen in this study seems to be rather high compared to that. However, the efficiency of pollen traps might vary (Keller, Fluri, and Imdorf, 2005) and data are not always comparable between studies. The observed seasonal differences in dry weight with the highest values in April and May might partly reflect the colony development with higher demands of protein for larvae rearing in spring. In contrast to Requier et al. (2015), we observed a peak of collected pollen mass in April which coincides with the mass-flowering of oilseed rape. However, the very low input in June coincides with the food depletion period reported by Requier et al. (2015). Occurring when colonies are at their maximum population size, it could affect colony growth rates and health (Requier et al., 2015).

Collected pollen came from a wide variety of plant species (80). Our result lies in between previous ones of 46 (Alves and Santos, 2014) and 95 for feral honey bee colonies (Baum et al., 2004), but much under 164 and 228 reported by Koeppler, Vorwohl, and Koeniger (2007) and Requier et al. (2015) for honey bee colonies (*Apis mellifera carnica*). However, if our dataset is restricted to species with a minimum relative abundance of 0.1% per sample instead of 1%, we detect pollen from 149 different plant species. The additional species do not change the patterns of richness and diversity but rise their levels almost equally over all samples (data not shown). We excluded taxa below 1% abundance since their contribution is low in the overall composition of pollens, and also to reduce the risk of including such cross-contaminated by wind and other pollinators or sequencing errors into the analysis in our study.

Richness of pollen was lower in April and first half of May compared to end of May and June while this pattern was even stronger for pollen diversity. Lower values in spring correspond to the mass-flowering of oilseed rape and the higher demand of colonies for pollen. They collected the bulk of their pollen diet from oilseed rape to meet the demand during a stage of naturally strong colony

growth. In contrast, previous studies report a low probability of collecting pollen from oilseed rape (Garbuzov et al., 2015; Requier et al., 2015; Woodcock et al., 2013). In a previous analysis within the same field experiment (Danner et al., 2016) we assessed the role of oilseed rape by calculating foraging frequencies for different land use types based on decoded waggle dances and found that oilseed rape did not dominate. Apparently, waggle dances for specific patches reflect the actual input of their offered pollen disproportionately. Since our abundance estimations are based on an amplicon generation step by PCR, abundances of dominant species may be overrepresented by random processes underlying the process. We accounted for this effect by preparing triplicate samples following a study that directly compared pollen abundances by light-microscopy and metabarcoding, but it may be not removed completely.

Our study supports the hypothesis that honey bees try to collect a diverse pollen diet in order to ensure colony performance (Danner et al., 2016; Hendriksma and Shafir, 2016; Requier et al., 2015; Alaux et al., 2010). First, during mass-flowering of oilseed rape in April and May honey bees still foraged for supplementary pollen from about 20 plant species (± 2.49 s.e.m.; each >1% abundance) per landscape that constituted about 4-25% of the daily input, depending on date. This richness seems to be high if compared with results from Israel, where the number of pollen sources from trapped pollen pellets ranged between 5 and 20 plant species per sampling date per site (Avni, Dag, and Shafir, 2009). It has recently been shown in a laboratory experiment that honey bees are able to balance colony nutrition deficiencies by selectivity for different pollen Hendriksma and Shafir (2016), while our data provides evidence for it under field conditions. Similarly, Requier et al. (2015) report an unexpected high pollen diversity during oilseed rape bloom. Second, landscape diversity (and therefore diversity of resources) did not influence the amount and diversity of collected pollen, as could be assumed if honey bees would forage randomly. Piroux et al. (2014) also report that the diversity of pollen, collected by honey bees in a field experiment, does not reflect landscape diversity.

We could show that foraging distances increase with decreasing landscape diversity. At the same time neither pollen amount nor diversity were influenced by landscape diversity, suggesting that honey bees compensate for a lower resource availability by an increased foraging range to maintain pollen amount and diversity. Similarly, Steffan-Dewenter and Kuhn (2003) report higher pollen foraging distances in simple landscapes compared to complex ones that are supposed to offer more resources, and Danner et al. (2016) found increasing pollen foraging distances when semi-natural habitats close to the hive decreased. The diversity of pollen input, as shown in our study, is most probably a result of nutritional demands that have to be met and apparently can be met regardless of landscape diversity but under the costs of higher foraging distances. In other words, the effectiveness of a colony decreases in less diverse landscapes, as more effort has

to be put into foraging for pollen. If we think of landscapes with even lower diversity (as it is the case in many agricultural regions within Germany but also worldwide) than in our study, a honey bee colony should come to a point where the costs exceed the benefits, i.e. when it is not worthwhile any more to expand the foraging range for an increased diversity of pollen diet. However, foraging distances up to 13.5 km (Frisch, 1967) are reported. The importance of a diverse pollen diet for honey bee health is confirmed by Alaux et al. (2010) who found that it can enhance immune functions of honey bees and that a minimal nutrient diversity may not meet all honey bee needs. Di Pasquale et al. (2013) found positive effects of pollen diet diversity on the lifespan of parasitized honey bees.

Conclusions

In conclusion, the amount and diversity of pollen input were influenced by season but not by landscape diversity, suggesting that the demands of a small honey bee colony can be met in simple and complex agricultural landscapes within our study region. In contrast to results from previous studies oilseed rape was an important pollen resource, resulting in a lower pollen diversity in spring. However, gaps in availability might arise if alternative resources outside mass-flowering periods are missing, that can not be compensated by further extending the foraging range. Agri-environmental schemes aiming to support pollinators should develop adequate floral resources for agricultural landscapes, that fill the temporal and spatial gaps in resource availability and diversity. The successful application of DNA metabarcoding in our study underpins that it is a useful novel method in honey bee foraging ecology that reduces the reliance on specific palynological expert knowledge.

Supplement

Table S1: Dry weight - results from multi-model inference and model averaging.

```
Component models:
df  logLik   AICc delta weight
1        9 -124.65 269.39  0.00   0.78
12      10 -124.64 271.87  2.48   0.22
123     16 -123.45 285.78 16.40   0.00
(Null)   3 -141.94 290.14 20.75   0.00
2        4 -141.93 292.30 22.91   0.00

Term codes:
datef       landdiv datef:landdiv
1             2             3

Model-averaged coefficients:
(full average)
Estimate Std. Error Adjusted SE z value Pr(>|z|)
```

```
(Intercept)             4.721e+00  3.499e-01   3.557e-01  13.274  < 2e-16 ***
datef20120503          -8.315e-01  3.947e-01   4.013e-01   2.072  0.03827 *
datef20120514          -1.100e+00  3.948e-01   4.013e-01   2.741  0.00612 **
datef20120524          -2.783e-01  4.043e-01   4.110e-01   0.677  0.49843
datef20120614          -1.715e+00  4.044e-01   4.111e-01   4.171 3.03e-05 ***
datef20120726          -1.162e+00  3.948e-01   4.013e-01   2.894  0.00380 **
datef20120804          -1.874e+00  4.178e-01   4.247e-01   4.412 1.03e-05 ***
landdiv                 6.522e-03  1.211e-01   1.324e-01   0.049  0.96070
datef20120503:landdiv -9.279e-06  1.291e-02   1.314e-02   0.001  0.99944
datef20120514:landdiv -6.905e-05  1.373e-02   1.395e-02   0.005  0.99605
datef20120524:landdiv -1.064e-04  1.543e-02   1.565e-02   0.007  0.99457
datef20120614:landdiv -3.515e-05  1.414e-02   1.439e-02   0.002  0.99805
datef20120726:landdiv -1.719e-05  1.295e-02   1.318e-02   0.001  0.99896
datef20120804:landdiv -2.616e-04  2.339e-02   2.357e-02   0.011  0.99114
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Relative variable importance:
datef landdiv datef:landdiv
Importance:              1  0.22   <0.01
N containing models:     3     3       1
```

Table S2: Richness - results from multi-model inference and model averaging.

```
Component models:
df logLik   AICc delta weight
1       9 -69.65 159.39  0.00   0.62
12     10 -68.93 160.44  1.05   0.37
123    16 -64.67 168.22  8.84   0.01
(Null)  3 -96.89 200.03 40.65   0.00
2       4 -96.11 200.67 41.28   0.00

Term codes:
datef       landdiv datef:landdiv
1             2             3

Model-averaged coefficients:
(full average)
Estimate Std. Error Adjusted SE z value Pr(>|z|)
(Intercept)             3.8330039  0.2188744   0.2220491  17.262  < 2e-16 ***
datef20120503           0.7696041  0.2377671   0.2414466   3.187  0.00144 **
datef20120514           0.5726951  0.2324015   0.2361646   2.425  0.01531 *
datef20120524           1.3453269  0.2344145   0.2383235   5.645  < 2e-16 ***
datef20120614           1.4920617  0.2351660   0.2390683   6.241  < 2e-16 ***
datef20120726           1.2350380  0.2411290   0.2447579   5.046    5e-07 ***
datef20120804           0.7184713  0.2423483   0.2463891   2.916  0.00355 **
landdiv                -0.0538779  0.1139867   0.1201300   0.448  0.65379
datef20120503:landdiv -0.0053132  0.0739796   0.0744107   0.071  0.94308
datef20120514:landdiv -0.0033283  0.0566645   0.0572262   0.058  0.95362
datef20120524:landdiv  0.0005677  0.0445403   0.0453300   0.013  0.99001
datef20120614:landdiv -0.0013894  0.0479021   0.0486744   0.029  0.97723
datef20120726:landdiv -0.0062581  0.0831567   0.0835405   0.075  0.94029
```

```
datef20120804:landdiv  0.0008079  0.0495681   0.0504349   0.016  0.98722
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Relative variable importance:
datef landdiv datef:landdiv
Importance:          1.00  0.38    0.01
N containing models:    3     3       1
```

Table S3: Diversity - results from multi-model inference and model averaging.

```
Component models:
df logLik   AICc delta weight
1      15 -25.24  86.49  0.00   0.78
12     16 -25.05  88.98  2.49   0.22
123    22 -22.34 102.55 16.06   0.00
(Null)  9 -62.75 145.59 59.10   0.00
2      10 -62.74 148.07 61.58   0.00

Term codes:
datef       landdiv datef:landdiv
1             2             3

Model-averaged coefficients:
(full average)
Estimate Std. Error Adjusted SE z value Pr(>|z|)
(Intercept)            3.250e-01  6.347e-02   6.428e-02   5.056    4e-07 ***
datef20120503         -7.903e-02  6.958e-02   7.054e-02   1.120    0.263
datef20120514         -1.114e-02  7.383e-02   7.485e-02   0.149    0.882
datef20120524          8.996e-01  1.689e-01   1.712e-01   5.255    1e-07 ***
datef20120614          9.276e-01  1.598e-01   1.620e-01   5.725   <2e-16 ***
datef20120726          1.221e+00  1.372e-01   1.391e-01   8.779   <2e-16 ***
datef20120804          8.306e-01  1.033e-01   1.047e-01   7.934   <2e-16 ***
landdiv                8.999e-03  3.434e-02   3.471e-02   0.259    0.795
datef20120503:landdiv  4.681e-05  3.843e-03   3.867e-03   0.012    0.990
datef20120514:landdiv  2.852e-05  3.235e-03   3.269e-03   0.009    0.993
datef20120524:landdiv  1.189e-04  9.994e-03   1.006e-02   0.012    0.991
datef20120614:landdiv -3.535e-05  7.162e-03   7.259e-03   0.005    0.996
datef20120726:landdiv -8.987e-05  7.574e-03   7.625e-03   0.012    0.991
datef20120804:landdiv  3.502e-05  5.024e-03   5.085e-03   0.007    0.995
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Relative variable importance:
datef landdiv datef:landdiv
Importance:             1  0.22   <0.01
N containing models:    3     3       1
```

Table S4: Foraging distances - results from multi-model inference and model averaging.

```
Component models:
```

```
df   logLik    AICc  delta weight
123    17 -1594.17 3222.79   0.00   0.97
12     11 -1603.98 3230.16   7.37   0.02
2      10 -1606.89 3233.95  11.15   0.00
1       5 -1655.62 3321.28  98.49   0.00
(Null)  4 -1657.77 3323.57 100.77   0.00

Term codes:
landdiv         round landdiv:round
1             2             3

Model-averaged coefficients:
(full average)
Estimate Std. Error Adjusted SE z value Pr(>|z|)
(Intercept)     6.35188    0.29654     0.29681  21.401  < 2e-16 ***
landdiv        -0.40674    0.25124     0.25146   1.617 0.105773
round2          0.39624    0.24603     0.24625   1.609 0.107592
round3          0.96788    0.24434     0.24454   3.958 7.56e-05 ***
round4          0.93072    0.26312     0.26336   3.534 0.000409 ***
round6          0.58141    0.25824     0.25847   2.249 0.024488 *
round7          0.82275    0.29082     0.29108   2.827 0.004706 **
round8          1.25001    0.31943     0.31970   3.910 9.23e-05 ***
landdiv:round2  0.21176    0.21843     0.21863   0.969 0.332743
landdiv:round3 -0.44427    0.21663     0.21681   2.049 0.040448 *
landdiv:round4  0.04340    0.22667     0.22688   0.191 0.848298
landdiv:round6  0.04247    0.24654     0.24676   0.172 0.863339
landdiv:round7 -0.14387    0.25577     0.25600   0.562 0.574112
landdiv:round8 -0.46602    0.29370     0.29394   1.585 0.112876
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Relative variable importance:
round landdiv landdiv:round
Importance:          1.00  1.00    0.97
N containing models:    3     3       1
```

Table S5: List of species, sorted by relative abundance over all samples. All species had an abundance of >1% within the samples in which they occurred.

Genus	**Species**	**Abundance [%]**
Brassica	napus	48.38
Papaver	rhoeas	6.43
Picris	hieracioides	6.16
Centaurea	jacea	4.91
Alyssum	murale	2.65
Potentilla	micrantha	2.04
Trifolium	pratense	1.96
Cirsium	palustre	1.67
Trifolium	repens	1.56

Vicia	lathyroides	1.49
Sisymbrium	irio	1.26
Filipendula	ulmaria	1.22
Crepis	vesicaria	1.07
Trifolium	hybridum	0.94
Potentilla	reptans	0.8
Hypericum	perforatum	0.71
Trifolium	striatum	0.63
Plantago	major	0.62
Heracleum	mantegazzianum	0.61
Lepidium	perfoliatum	0.6
Prunus	avium	0.57
Sisymbrium	officinale	0.56
Malus	domestica	0.55
Potentilla	clusiana	0.51
Arctium	tomentosum	0.49
Brassica	nigra	0.4
Barbarea	vulgaris	0.34
Melilotus	officinalis	0.3
Plantago	lanceolata	0.28
Carthamus	spec	0.26
Potentilla	alba	0.25
Sisymbrium	austriacum	0.22
Juglans	regia	0.22
Centaurea	cyanus	0.21
Mercurialis	annua	0.18
Chenopodium	ficifolium	0.18
Inula	germanica	0.17
Potentilla	norvegica	0.17
Neslia	paniculata	0.15
Quercus	robur	0.13
Reseda	lutea	0.11
Potentilla	caulescens	0.11
Tilia	platyphyllos	0.08
Lolium	multiflorum	0.07
Hypochaeris	radicata	0.07
Crataegus	monogyna	0.06
Elymus	repens	0.06
Trifolium	arvense	0.06
Silene	vulgaris	0.05
Tragopogon	pratensis	0.05
Sisymbrium	loeselii	0.05

Sisymbrium	strictissimum	0.04
Berteroa	incana	0.04
Mercurialis	spec	0.04
Phalaris	arundinacea	0.04
Senecio	doronicum	0.04
Elymus	caninus	0.04
Quercus	petraea	0.03
Prunus	domestica	0.03
Lathyrus	pratensis	0.03
Aethusa	cynapium	0.03
Conyza	canadensis	0.02
Rosa	rugosa	0.02
Ambrosia	artemisiifolia	0.02
Fagus	sylvatica	0.02
Cardamine	amara	0.01
Betula	nana	0.01
Crataegus	laevigata	0.01
Solanum	nigrum	0.01
Ranunculus	circinatus	0.01
Jacobaea	spec	0.01
Capsella	bursa-pastoris	0
Sonchus	oleraceus	0
Falcaria	vulgaris	0
Daucus	carota	0
Impatiens	glandulifera	0
Lolium	perenne	0
Hypericum	hirsutum	0
Quercus	spec	0
Onobrychis	spec	0

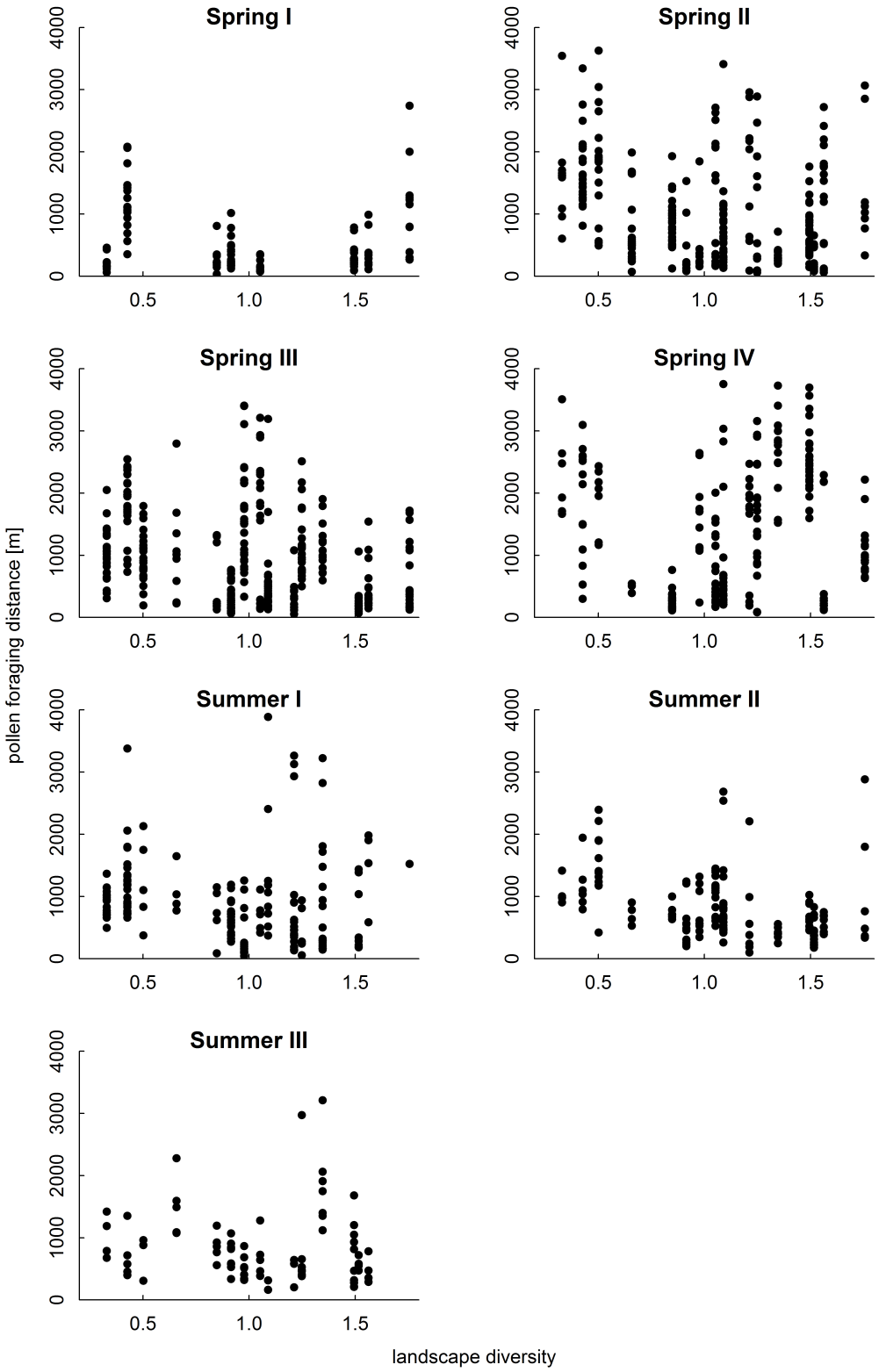


Figure S1: Pollen foraging distances versus landscape diversity. Raw data points for seven observation rounds in 16 landscapes (8 landscapes in the first round).

7 General Discussion

This dissertation provides insight into the pollen foraging of honey bee colonies in agricultural landscapes, the influence of landscape structure, possible implications for honey bee health as well as suggestions to support them. For the first time, foraging distances for a specific pollen type (maize) are provided, allowing for the estimation of the exposure risk to systemic insecticides or genetically modified maize (Chapter 4). Foraging frequencies for different landuse types were analyzed and showed the importance of maize pollen among a variety of used pollen resources in summer. In a second field experiment, focussing on oilseed rape (OSR) and semi-natural habitats (SNH), the analysis of foraging distances for pollen revealed a strong influence of landscape composition, depending on scale (Chapter 5). Pollen foragers did not prefer mass-flowering OSR compared to mixed floral resources in SNH. A strong increase of foraging frequency on SNH in simple agricultural landscapes was observed, indicating its importance as a pollen resource and the potentially resulting resource competition with wild bees. A detailed analysis of pollen samples, collected within the same field experiment, showed that dry weight, richness and diversity of the pollen diet were not influenced by landscape diversity (Chapter 6). However, foraging distances increased with decreasing landscape diversity.

7.1 The influence of landscape structure

Our design of field experiments on a landscape scale allowed us to analyze the influence of landscape structure - more specifically landscape composition - on the pollen foraging of honey bee colonies. Parameters like area of SNH/OSR/maize (Chapters 4 and 5) and landscape diversity (Chapter 6) are measures of landscape composition (McGarigal and Marks, 1994) and assumed to reflect the availability of pollen resources. We could show that landscape composition influenced honey bee foraging for pollen (Chapters 5 and 6). Mass-flowering OSR and SNH influenced pollen foraging distances on different scales and depending on season. Furthermore, the importance of SNH as a pollen resource was influenced by landscape composition, i.e. foraging frequency on SNH increased with decreasing SNH area in the landscape. Decreasing landscape diversity (a more general parameter reflecting resource availability), resulted in increased foraging ranges while pollen diet was unaffected (Chapter 6). However, maize area within 1500m around honey bee hives had no influence on foraging distances, and grassland area as an alternative pollen resource did not modulate the percentage of maize

pollen foragers (Chapter 4). It has been shown earlier that simple structured landscapes can result in increased foraging distances of honey bees (Steffan-Dewenter and Kuhn, 2003). Within our studies its influence was analyzed by incorporating composition gradients in the study design for the first time, which allowed for a comprehensive analysis. Summarizing our results it is clear that a better availability of diverse pollen resources (reflected by measures of landscape composition) supports honey bee colonies, with mass-flowering crops being an important part but obviously not sufficient to meet the nutritional demands. An increased amount of intensively used pollen resources does not necessarily result in lower foraging ranges as shown in the case of maize (Chapter 4). It seems to be the combination of quantity and diversity of pollen resources that needs to reach certain levels to meet the demand of honey bee colonies, i.e. to allow for an efficient foraging with positive influence on colony growth and health.

7.2 The need for diverse pollen resources

In Chapter 4 it is already indicated that honey bees prefer to collect a diverse pollen diet instead of a pure maize diet. They took longer distances under higher costs of energy for the collection of other pollen than maize (Figure 4.2). The design of the follow-up study allowed for a direct comparison between a mass-flowering crop (in this case OSR) and alternative resources (SNH; Chapter 5). The importance of SNH as a pollen resource in contrast to OSR showed that diverse pollen resources are important for honey bees. Similarly, in Chapter 6 we found that landscape diversity did not influence richness and diversity of the pollen diet. Instead, less diverse landscapes were compensated by increasing the foraging range to maintain richness and diversity of the pollen diet. The need for diverse pollen resources to ensure honey bee colony health is confirmed by several studies, e.g. Alaux et al. (2010) report positive effects of diet diversity on the immunity of honey bees. Furthermore, Arien et al. (2015) show that a deficiency of the essential fatty acid omega-3 greatly impaired honey bee learning. At the same time their pollen analysis revealed that many managed colonies are experiencing a shift in available forage toward a higher omega-6:3 ratio (i.e. lower omega 3 content), which may be leading to colony declines. Hendriksma and Shafir (2016) could recently show in a laboratory experiment that honey bees are able to select different pollen in order to balance colony nutrition deficiencies. The unexpected high pollen diversity in a field experiment by Requier et al. (2015) during OSR bloom similarly supports our observations, that honey bees aim for a diverse pollen diet. However, different mass-flowering crops may represent different importances within the diet, depending on the availability of alternative resources. The question, how bees would percept the nutritional status of the colony and the potential need for different pollen has to be clarified yet (Hendriksma and Shafir, 2016) and was not part of this work.

7.3 Environmental risks

Mass-flowering crops like maize and oilseed rape are potential exposure pathways especially for systemic pesticides like neonicotinoids. In North America a remarkably high amount of 98 pesticides and metabolites was detected in mixtures up to 214 ppm in bee pollen and exposure to many of these neurotoxicants elicits acute and sublethal reductions in honey bee fitness (Mullin et al., 2010). In France residues of imidacloprid were among the most frequently detected in pollen loads (Chauzat et al., 2009). Nonlethal exposure of honey bees to thiamethoxam (neonicotinoid systemic pesticide) causes high mortality due to homing failure at levels that could put a colony at risk of collapse (Henry et al., 2012a). According to our results in Chapters 4 and 5, honey bees intensively forage on maize and oilseed rape and therefore can be strongly exposed to pesticides under field conditions. However, foraging on non-cultivated plants in crop-dominated landscapes can also play a major role and exposes honey bees to pesticides across the season (Requier et al., 2015; Long and Krupke, 2016).

In the case of genetically modified crops no harms of honey bees have been proven so far (see Introduction 1.3). However, they should still be part of environmental risk assessments regarding honey bees and other non-target organisms (Steijven, Steffan-Dewenter, and Härtel, 2015). Another point is the potential contamination of honey with GM pollen, which caused problems in the past, since the pollen was first regarded as an ingredient, and GM ingredients in food products have to be authorized as such, otherwise the product could not be marketed. Two years ago, the European Parlament voted for a rule, defining pollen in honey as a constituent rather than an ingredient (**_parliament_????**). That means current EU legislation for labelling applies, requiring an indication of GM pollen if present in quantities over 0.9 % of the honey, which should not be relevant in practice anymore.

Beside the exposure risks there is the risk of nutritional deficiency with its implications for honey bee health (Brodschneider and Crailsheim, 2010). The availability of particular nutrients may be low in particular locations or seasons and severe nutritional deficiencies may occur in especially depleted environments like agricultural monocultures (Avni et al., 2014). For example, maize pollen is lacking the amino acid histidine, which is essential for honey bees (Groot de, 1953). If collected intensively - as shown in Chapter 4 - it may lead to a diet of poor nutritional quality. A reduced availability and/or deficiency in diet quality can directly impact worker longevity and nursing physiology (Pasquale et al., 2016).

7.4 Potential measures

According to our results the potential hazard of maize fields could not be controlled by maize area within the main foraging range (Chapter 4). A reduction

in exposure could be reached by a large distance of 1500m between the hive and the next maize field. However, there was an indication that flower-rich habitats could prevent honey bees from maize pollen foraging and therefore reduce the exposure to pesticides.

We could show that pollen foraging of honey bee colonies - a relevant part of energy effort - can be shaped by landscape composition which offers the possibility to actively support honey bees in the agricultural landscape. Shorter foraging distances should be an advantage for the wellbeing of a colony since it can put more effort in growth and reproduction. They may be reduced by increasing the availability of pollen resources in the landscape. In Chapter 5 we revealed the importance of SNH for honey bee colonies and consequently maintaining SNH would be of great importance for their wellbeing. In fact, every measure or type of landscape management that increases the availability of resources, e.g. the establishment of flower strips, organic farming and diversifying flowering crops are adequate for supporting honey bees - and pollinators in general - in agricutural landscapes. When implementing new measures it should be considered whether resources - and especially pollen resources - are available continuously not only in space but also in time.

Bibliography

Aizen, Marcelo A. et al. (2009). "How Much Does Agriculture Depend on Pollinators? Lessons from Long-Term Trends in Crop Production". In: *Annals of Botany* 103.9, pp. 1579–1588.

Alaux, Cédric et al. (2010). "Diet Effects on Honeybee Immunocompetence". In: *Biology Letters* 6.4, pp. 562–565.

Alves, Rodolfo de Franca and Francisco de Assis Ribeiro dos Santos (2014). "Plant Sources for Bee Pollen Load Production in Sergipe, Northeast Brazil". In: *Palynology* 38.1, pp. 90–100.

Arien, Yael et al. (2015). "Omega-3 Deficiency Impairs Honey Bee Learning". In: *Proceedings of the National Academy of Sciences* 112.51, pp. 15761–15766.

Avni, Dorit, Arnon Dag, and Sharoni Shafir (2009). "Pollen Sources for Honeybees in Israel: Source, Periods of Shortage, and Influence on Population Growth". In: *Israel Journal of Plant Sciences* 57.3, pp. 263–275.

Avni, Dorit et al. (2014). "Nutritional Aspects of Honey Bee-Collected Pollen and Constraints on Colony Development in the Eastern Mediterranean". In: *Journal of Insect Physiology*. Mechanisms of Nutritional Homeostasis in Insects 69, pp. 65–73.

Babendreier, D. et al. (2004). "Pollen Consumption in Honey Bee Larvae: A Step Forward in the Risk Assessment of Transgenic Plants". In: *Apidologie* 35.3, pp. 293–300.

Barton, K. (2012). *MuMIn: Multi-Model Inference*. Version 1.7.2.

Bates, D. M., M. Maechler, and B. Bolker (2011). *lme4: Linear Mixed-Effects Models Using S4 Classes*. Version 0.

— (2015). *lme4: Linear Mixed-Effects Models using 'Eigen' and S4*. Version 1.1-8.

Baum, Kristen A. et al. (2004). "Pollen Selection by Feral Honey Bee (Hymenoptera: Apidae) Colonies in a Coastal Prairie Landscape". In: *Environmental Entomology* 33.3, pp. 727–739.

Beekman, M. and F.L.W. Ratnieks (2000). "Long-Range Foraging by the Honey-Bee, Apis Mellifera L." In: *Functional Ecology* 14.4, pp. 490–496.

Beekman, M. et al. (2004). "Comparing Foraging Behaviour of Small and Large Honey-Bee Colonies by Decoding Waggle Dances Made by Foragers". In: *Functional Ecology* 18.6, pp. 829–835.

Benjamini, Y. and D. Yekutieli (2001). "The Control of the False Discovery Rate in Multiple Testing under Dependency". In: *Annals of Statistics* 29.4, pp. 1165–1188.

Blacquière, Tjeerd et al. (2012). "Neonicotinoids in Bees: A Review on Concentrations, Side-Effects and Risk Assessment". In: *Ecotoxicology* 21.4, pp. 973–992.

Bonmatin, J. M. et al. (2003). "A LC/APCI-MS/MS Method for Analysis of Imidacloprid in Soils, in Plants, and in Pollens". In: *Analytical Chemistry* 75.9, pp. 2027–2033.

Brodschneider, Robert and Karl Crailsheim (2010). "Nutrition and Health in Honey Bees". In: *Apidologie* 41.3, pp. 278–294.

Brown, Mark J. F. and Robert J. Paxton (2009). "The Conservation of Bees: A Global Perspective". In: *Apidologie* 40.3, pp. 410–416.

Burnham, Kenneth P. and David R. Anderson, eds. (2004). *Model Selection and Multimodel Inference*. New York, NY: Springer New York. URL: http://link.springer.com/10.1007/b97636.

Camazine, S. (1993). "The Regulation of Pollen Foraging by Honey-Bees: How Foragers Assess the Colony's Need for Pollen". In: *Behavioral Ecoclogy and Sociobiology* 32.4, pp. 265–272.

Caporaso, J. Gregory et al. (2010). "QIIME Allows Analysis of High-Throughput Community Sequencing Data". In: *Nature Methods* 7.5, pp. 335–336.

Chauzat, Marie-Pierre et al. (2009). "Influence of Pesticide Residues on Honey Bee (Hymenoptera: Apidae) Colony Health in France". In: *Environmental Entomology* 38.3, pp. 514–523.

Chen, Shilin et al. (2010). "Validation of the ITS2 Region as a Novel DNA Barcode for Identifying Medicinal Plant Species". In: *PLOS ONE* 5.1, e8613.

Couvillon, M. J. (2012). "The Dance Legacy of Karl von Frisch". In: *Insectes Sociaux* 59.3, pp. 297–306.

Couvillon, Margaret J., Roger Schürch, and Francis L. W. Ratnieks (2014a). "Dancing Bees Communicate a Foraging Preference for Rural Lands in High-Level Agri-Environment Schemes". In: *Current Biology* 24.11, pp. 1212–1215.

— (2014b). "Waggle Dance Distances as Integrative Indicators of Seasonal Foraging Challenges". In: *PLoS ONE* 9.4, e93495.

Couvillon, Margaret J. et al. (2012). "Intra-Dance Variation among Waggle Runs and the Design of Efficient Protocols for Honey Bee Dance Decoding." In: *Biology open* 1.5, pp. 467–472.

Couvillon, Margaret J. et al. (2015). "Honey Bee Foraging Distance Depends on Month and Forage Type". In: *Apidologie* 46.1, pp. 61–70.

Crawley, M.J. (2007). *The R Book*. Chichester: John Wiley & Sons Ltd.

Danner, Nadja, Stephan Härtel, and Ingolf Steffan-Dewenter (2014). "Maize Pollen Foraging by Honey Bees in Relation to Crop Area and Landscape Context". In: *Basic and Applied Ecology* 15.8, pp. 677–684.

Danner, Nadja et al. (2016). "Season and Landscape Composition Affect Pollen Foraging Distances and Habitat Use of Honey Bees". In: *Ecological Applications* 26.6, pp. 1920–1929.

Decourtye, Axel, Eric Mader, and Nicolas Desneux (2010). "Landscape Enhancement of Floral Resources for Honey Bees in Agro-Ecosystems". In: *Apidologie* 41.3, pp. 264–277.

Di Pasquale, Garance et al. (2013). "Influence of Pollen Nutrition on Honey Bee Health: Do Pollen Quality and Diversity Matter?" In: *PLoS ONE* 8.8, e72016.

Dolezal, Adam G. et al. (2016). "Intensively Cultivated Landscape and Varroa Mite Infestation Are Associated with Reduced Honey Bee Nutritional State". In: *PLOS ONE* 11.4, e0153531.

Duan, Jian J. et al. (2008). "A Meta-Analysis of Effects of Bt Crops on Honey Bees (Hymenoptera: Apidae)". In: *PLoS ONE* 3.1, e1415.

Edgar, Robert C. (2010). "Search and Clustering Orders of Magnitude Faster than BLAST". In: *Bioinformatics (Oxford, England)* 26.19, pp. 2460–2461.

ESRI (2011). *ArcGIS 10*. Redlands, CA.

Fewell, J.H. and M.L. Winston (1992). "Colony State and Regulation of Pollen Foraging in the Honey-Bee, Apis Mellifera L." In: *Behavioral Ecology and Sociobiology* 30.6, pp. 387–393.

Fierer, N. et al. (2008). "The Influence of Sex, Handedness, and Washing on the Diversity of Hand Surface Bacteria". In: *Proceedings of the National Academy of Sciences* 105.46, pp. 17994–17999.

Food and Agriculture Organization of the United Nations (2016). URL: `faostat3.fao.org;accessed6.10.2016`.

Free, J. B. (1970). *Insect Pollination of Crops.* London and New York: Academic Press.

Frisch, Karl von (1967). *The Dance Language and Orientation of Bees*. Belknap Press of Harvard University Press. 582 pp.

Garbuzov, Mihail et al. (2015). "Honey Bee Dance Decoding and Pollen-Load Analysis Show Limited Foraging on Spring-Flowering Oilseed Rape, a Potential Source of Neonicotinoid Contamination". In: *Agriculture, Ecosystems & Environment* 203, pp. 62–68.

Graf, Roland F. et al. (2005). "The Importance of Spatial Scale in Habitat Models: Capercaillie in the Swiss Alps". In: *Landscape Ecology* 20.6, pp. 703–717.

Groot de, A. P. (1953). "Protein and Amino Acid Requirements of the Honeybee (Apis Mellifera L.)" In: *Physiol. Comp. Oecologia* 3, pp. 197–285.

Haydak, M.H. (1970). "Honey Bee Nutrition". In: *Annual Review of Entomology* 15, pp. 143–156.

Hendriksma, Harmen P. and Sharoni Shafir (2016). "Honey Bee Foragers Balance Colony Nutritional Deficiencies". In: *Behavioral Ecology and Sociobiology* 70.4, pp. 509–517.

Hendriksma, Harmen P. et al. (2012). "Effects of Multiple Bt Proteins and GNA Lectin on in Vitro-Reared Honey Bee Larvae". In: *Apidologie* 43.5, pp. 549–560.

Hendriksma, Harmen P. et al. (2013). "Effect of Stacked Insecticidal Cry Proteins from Maize Pollen on Nurse Bees (Apis Mellifera Carnica) and Their Gut Bacteria". In: *Plos One* 8.3, e59589.

Henry, Mickaël et al. (2012a). "A Common Pesticide Decreases Foraging Success and Survival in Honey Bees". In: *Science* 336.6079, pp. 348–350.

— (2012b). "Response to Comment on "A Common Pesticide Decreases Foraging Success and Survival in Honey Bees"". In: *Science* 337.6101, pp. 1453–1453.

Holzschuh, Andrea et al. (2013). "Mass-Flowering Crops Enhance Wild Bee Abundance". In: *Oecologia* 172.2, pp. 477–484.

Härtel, Stephan and Ingolf Steffan-Dewenter (2014). "Ecology: Honey Bee Foraging in Human-Modified Landscapes". In: *Current Biology* 24.11, R524–R526.

Hudewenz, Anika and Alexandra-Maria Klein (2013). "Competition between Honey Bees and Wild Bees and the Role of Nesting Resources in a Nature Reserve". In: *Journal of Insect Conservation* 17.6, pp. 1275–1283.

Keller, A. et al. (2015). "Evaluating Multiplexed next-Generation Sequencing as a Method in Palynology for Mixed Pollen Samples". In: *Plant Biology* 17.2, pp. 558–566.

Keller, I., P. Fluri, and A. Imdorf (2005). "Pollen Nutrition and Colony Development in Honey Bees - Part II". In: *Bee World* 86.2, pp. 27–34.

Klein, Alexandra-Maria et al. (2007). "Importance of Pollinators in Changing Landscapes for World Crops". In: *Proceedings of the Royal Society B-Biological Sciences* 274.1608, pp. 303–313.

Koeppler, Kirsten, Guenther Vorwohl, and Nikolaus Koeniger (2007). "Comparison of Pollen Spectra Collected by Four Different Subspecies of the Honey Bee Apis Mellifera". In: *Apidologie* 38.4, pp. 341–353.

Krupke, Christian H. et al. (2012). "Multiple Routes of Pesticide Exposure for Honey Bees Living Near Agricultural Fields". In: *PLoS ONE* 7.1, e29268.

Lecocq, Antoine et al. (2015). "Weight Watching and the Effect of Landscape on Honeybee Colony Productivity: Investigating the Value of Colony Weight Monitoring for the Beekeeping Industry". In: *PLoS ONE* 10.7, e0132473.

Long, Elizabeth Y. and Christian H. Krupke (2016). "Non-Cultivated Plants Present a Season-Long Route of Pesticide Exposure for Honey Bees". In: *Nature Communications* 7, p. 11629.

MacGarigal, K and B.J. Marks (1994). *Fragstats: Spatial Pattern Analysis Program for Quantifying Landscape Structure*. URL: `http://www.umass.edu/landeco/research/fragstats/fragstats.html`.

Malone, L.A. and E.P.J. Burgess (2009). "Impact of Genetically Modified Crops on Pollinators". In: *Environmental Impact of Genetically Modified Crops*. Ed. by N. Ferry and A.M.R. Gatehouse. Oxfordshire: CAB International.

McMurdie, Paul J. and Susan Holmes (2013). "Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data". In: *PLoS ONE* 8.4, e61217.

Meissle, M. et al. (2009). "Pests, Pesticide Use and Alternative Options in Euopean Maize Production: Current Status and Future Prospects". In: *Journal of Applied Entomology*.

Mullin, Christopher A. et al. (2010). "High Levels of Miticides and Agrochemicals in North American Apiaries: Implications for Honey Bee Health". In: *PLOS ONE* 5.3, e9754.

Nagendra, Harini (2002). "Opposite Trends in Response for the Shannon and Simpson Indices of Landscape Diversity". In: *Applied Geography* 22.2, pp. 175–186.

Nedić, Nebojša et al. (2013). "Melliferous Potential of Brassica Napus L. Subsp. Napus (Cruciferae)". In: *Arthropod-Plant Interactions* 7.3, pp. 323–333.

Neumann, Peter and Norman L. Carreck (2010). "Honey Bee Colony Losses". In: *Journal of Apicultural Research* 49.1, pp. 1–6.

Ockinger, Erik and Henrik G. Smith (2007). "Semi-Natural Grasslands as Population Sources for Pollinating Insects in Agricultural Landscapes". In: *Journal of Applied Ecology* 44.1, pp. 50–59.

Odoux, Jean-Francois et al. (2012). "Territorial Biodiversity and Consequences on Physico-Chemical Characteristics of Pollen Collected by Honey Bee Colonies". In: *Apidologie* 43.5, pp. 561–575.

Odoux, Jean-Francois et al. (2014). "ECOBEE: A Tool for Long-Term Honey Bee Colony Monitoring at the Landscape Scale in West European Intensive Agroecosystems". In: *Journal of Apicultural Research* 53.1, pp. 57–66.

Pasquale, Garance Di et al. (2016). "Variations in the Availability of Pollen Resources Affect Honey Bee Health". In: *PLOS ONE* 11.9, e0162818.

Pernal, Stephen F. and Robert W. Currie (2001). "The Influence of Pollen Quality on Foraging Behavior in Honeybees (Apis Mellifera L.)" In: *Behavioral Ecology and Sociobiology* 51.1, pp. 53–68.

Pettis, Jeffery S. et al. (2012). "Pesticide Exposure in Honey Bees Results in Increased Levels of the Gut Pathogen Nosema". In: *Naturwissenschaften* 99.2, pp. 153–158.

Pettis, Jeffery S. et al. (2013). "Crop Pollination Exposes Honey Bees to Pesticides Which Alters Their Susceptibility to the Gut Pathogen Nosema Ceranae". In: *PLoS ONE* 8.7, e70182.

Pinheiro, J. et al. (2015). *Nlme: Linear and Nonlinear Mixed Effects Models*. Version 3.1-121. URL: `http://CRAN.R-project.org/package=nlme`.

Piroux, M. et al. (2014). "Correlating the Pollens Gathered by Apis Mellifera with the Landscape Features in Western France". In: *Applied Ecology and Environmental Research* 12.2, pp. 423–439.

Potts, Simon G. et al. (2010). "Global Pollinator Declines: Trends, Impacts and Drivers". In: *Trends in Ecology & Evolution* 25.6, pp. 345–353.

R Development Core Team (2011). *R: A Language and Environment for Statistical Computing*. Version 2.13.0. Vienna, Austria.

R Development Core Team (2015). *R: A Language and Environment for Statistical Computing*. Vienna, Austria.

Requier, Fabrice et al. (2015). "Honey Bee Diet in Intensive Farmland Habitats Reveals an Unexpectedly High Flower Richness and a Major Role of Weeds". In: *Ecological Applications* 25.4, pp. 881–890.

Rollin, Orianne et al. (2013). "Differences of Floral Resource Use between Honey Bees and Wild Bees in an Intensive Farming System". In: *Agriculture, Ecosystems & Environment* 179, pp. 78–86.

Rundlöf, Maj et al. (2015). "Seed Coating with a Neonicotinoid Insecticide Negatively Affects Wild Bees". In: *Nature* 521.7550, pp. 77–80.

Schneider, S.S. and L.C. McNally (1993). "Spatial Foraging Patterns and Colony Energy Status in the African Honey-Bee, Apis-Mellifera-Scutellata". In: *Journal of Insect Behavior* 6.2, pp. 195–210.

Schürch, Roger et al. (2013). "Incorporating Variability in Honey Bee Waggle Dance Decoding Improves the Mapping of Communicated Resource Locations". In: *Journal of Comparative Physiology A*, pp. 1–10.

Seeley, T.D. (1995). *The Wisdom of the Hive*. Cambridge: Harvard University Press.

Sickel, Wiebke et al. (2015). "Increased Efficiency in Identifying Mixed Pollen Samples by Meta-Barcoding with a Dual-Indexing Approach". In: *BMC Ecology* 15, p. 20.

Smart, Matthew et al. (2016). "Linking Measures of Colony and Individual Honey Bee Health to Survival among Apiaries Exposed to Varying Agricultural Land Use". In: *PLOS ONE* 11.3, e0152685.

Sponsler, DB and RM Johnson (2015). "Honey Bee Success Predicted by Landscape Composition in Ohio, USA". In: *PeerJ* 3.

Stanley, Dara A., Daryl Gunning, and Jane C. Stout (2013). "Pollinators and Pollination of Oilseed Rape Crops (Brassica Napus L.) in Ireland: Ecological and Economic Incentives for Pollinator Conservation". In: *Journal of Insect Conservation* 17.6, pp. 1181–1189.

Steffan-Dewenter, I. and A. Kuhn (2003). "Honeybee Foraging in Differentially Structured Landscapes". In: *Proceedings of the Royal Society of London, Series B* 270.1515, pp. 569–575.

Steffan-Dewenter, I. and T. Tscharntke (2000). "Resource Overlap and Possible Competition between Honey Bees and Wild Bees in Central Europe". In: *Oecologia* 122.2, pp. 288–296.

Steffan-Dewenter, I. et al. (2002). "Scale-Dependent Effects of Landscape Context on Three Pollinator Guilds". In: *Ecology* 83.5, pp. 1421–1432.

Steijven, Karin, Ingolf Steffan-Dewenter, and Stephan Härtel (2015). "Testing Dose-Dependent Effects of Stacked Bt Maize Pollen on in Vitro-Reared Honey Bee Larvae". In: *Apidologie* 47.2, pp. 216–226.

Stokstad, Erik (2013). "Pesticides Under Fire for Risks to Pollinators". In: *Science* 340.6133, pp. 674–676.

Suchail, Séverine, David Guez, and Luc P. Belzunces (2001). "Discrepancy between Acute and Chronic Toxicity Induced by Imidacloprid and Its Metabolites in Apis Mellifera". In: *Environmental Toxicology and Chemistry* 20.11, pp. 2482–2486.

Toth, Amy L. and Gene E. Robinson (2005). "Worker Nutrition and Division of Labour in Honeybees". In: *Animal Behaviour* 69.2, pp. 427–435.

vanEngelsdorp, Dennis et al. (2008). "A Survey of Honey Bee Colony Losses in the U.S., Fall 2007 to Spring 2008". In: *PLoS ONE* 3.12. Ed. by Nick Gay, e4071.

Visscher, P. K. and T. D. Seeley (1982). "Foraging Strategy of Honeybee Colonies in a Temperate Deciduous Forest". In: *Ecology* 63.6, pp. 1790–1801.

Von Frisch, K. (1965). *Tanzsprache Und Orientierung Der Bienen*. Berlin: Springer.

Waddington, K.D. et al. (1994). "Comparisons of Forager Distributions from Matched Honey-Bee Colocies in Suburban Environments". In: *Behavioral Ecology and Sociobiology* 35.6, pp. 423–429.

White, Tj et al. (1990). "Amplification and Direct Sequencing of Fungal Ribosomal RNA Genes for Phylogenetics". In: *PCR Protocols: A Guide to Methods and Applications*. Ed. by Ma Innis et al. Academic Press, pp. 315–322.

Winston, Mark L. (1987). *The Biology of the Honey Bee*. Harvard University Press. 300 pp.

Woodcock, B.A. et al. (2013). "Crop Flower Visitation by Honeybees, Bumblebees and Solitary Bees: Behavioural Differences and Diversity Responses to Landscape". In: *Agriculture, Ecosystems & Environment* 171, pp. 1–8.

Author contributions

Chapter 3

This chapter is published as: A. Keller, N. Danner, G. Grimmer, M. Ankenbrand, K. von der Ohe, W. von der Ohe, S. Rost, S. Härtel and I. Steffan-Dewenter (2015). Evaluating multiplexed next-generation sequencing as a method in palynology for mixed pollen samples. Plant Biology 17:558-566.

Author contributions

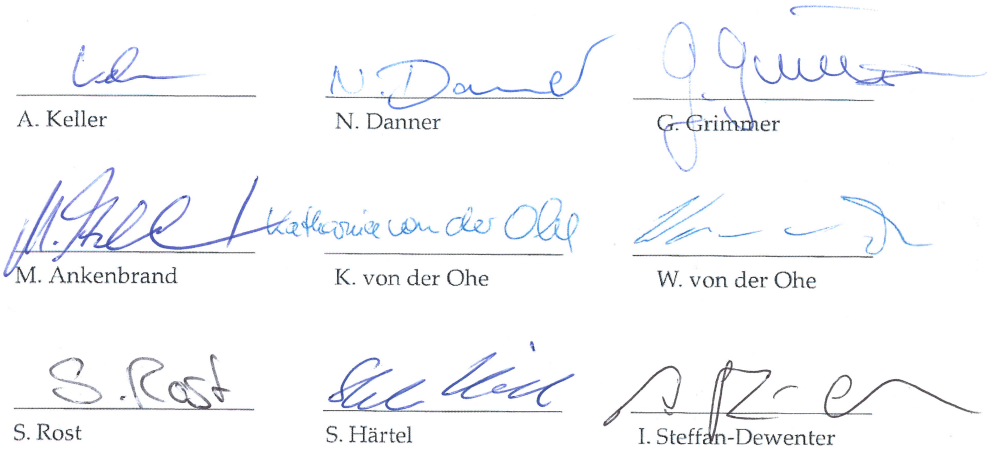
AK designed the overall study design as well as laboratory and bioinformatical parts of the project. ND, SH and ISD designed the field study. ND conducted the field work. ND, KvdO and WvdO conducted the pollen analysis via light microscopy. GG and SR performed the laboratory steps, AK and MA the bioinformatical analysis. AK wrote the first draft of the manuscript, ND contributed to the method section about pollen collection and classical identification, and all authors contributed and approved to the final version of the manuscript.

Acknowledgements

We greatly appreciate the help of staff from the Department of Human Genetics (University of Würzburg, Germany), especially C. Müller-Reible and T. Haaf for placing the 454 device at our disposal for this study. We also thank F. Förster (Department of Bioinformatics, University of Würzburg, Germany) for providing a useful taxonomy Perl module for generation of the training sets, and two anonymous reviewers for their valuable comments.

Funding

This research was partly funded by the German Federal Ministry for Education and Research (BMBF, project 0315215E) and within the project Assessing and Monitoring the Impacts of Genetically modified plants on Agro-ecosystems (AMIGA) funded by the European Commission under the Frame-work programme 7, Grant agreement 289706.



Chapter 4

This chapter is published as: Nadja Danner, Stephan Härtel, Ingolf Steffan-Dewenter (2014). Maize pollen foraging by honey bees in relation to crop area and landscape context. Basic and Applied Ecology 15:677-684.

Author contributions

ND, SH and ISD designed the study. ND conducted the field work. ND analyzed the data. ND wrote the first draft of the manuscript and all authors contributed to the final version of the manuscript.

Acknowledgements

We kindly thank Veronika Burgmayer (University of Bayreuth) for fieldwork support, Thorsten Steinhäuser (University of Bayreuth) for bee keeping support and all land owners for providing sites for observation of honey bee colonies. We thank three anonymous reviewers for their helpful comments on our manuscript.

Funding

This research was funded by the German Federal Ministry for Education and Research (BMBF, project 0315215E) and within the project Assessing and Monitoring the Impacts of Genetically modified plants on Agro-ecosystems (AMIGA) funded by the European Commission under the Frame-work programme 7, Grant agreement 289706.

Nadja Danner Stephan Härtel Ingolf Steffan-Dewenter

Chapter 5

This chapter is published as: Nadja Danner, Anna Maria Molitor, Susanne Schiele, Stephan Härtel, Ingolf Steffan-Dewenter (2016). Season and landscape composition affect pollen foraging distances and habitat use of honey bees. Ecological Applications 26:1920-1929.

Author contributions

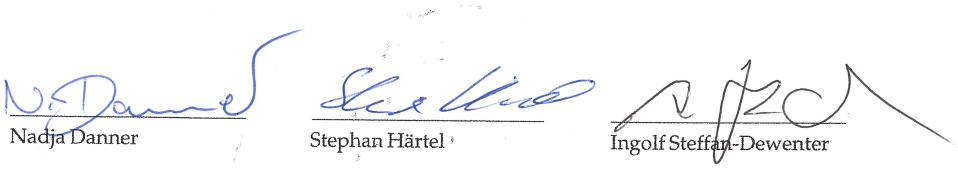
ND, SH and ISD designed the study. ND, AMM and SS conducted the field work. ND analyzed the data. ND wrote the first draft of the manuscript and all authors contributed to the final version of the manuscript.

Acknowledgements

We kindly thank Dirk Ahrens-Lagast and Beatrice Portail for their assistance with field and laboratory work.

Funding

This research was funded by the European Commission under the Framework Programme 7, Grant agreement 289706 (AMIGA: Assessing and Monitoring the Impacts of Genetically modified plants on Agro-ecosystems) and partly by the German Research Foundation (Collaborative Research Center 1047 "Insect Timing", project C2).



Chapter 6

This chapter is in preparation for submission as: Nadja Danner, Alexander Keller, Stephan Härtel, Ingolf Steffan-Dewenter (in prep.). Applying DNA metabarcoding in honey bee foraging ecology – Season but not landscape structure shapes the diversity of collected pollen.

Author contributions

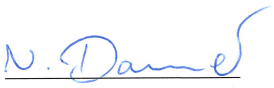
ND, SH and ISD designed the study. AK analyzed pollen samples and ND analyzed the data. ND wrote the first draft of the manuscript, AK contributed the method section about pollen analysis and all authors contributed to the final version of the manuscript.

Acknowledgements

We kindly thank Gudrun Grimmer for laboratory work.

Funding

This research was funded by the European Commission under the Framework Programme 7, Grant agreement 289706 (AMIGA: Assessing and Monitoring the Impacts of Genetically modified plants on Agro-ecosystems).



Nadja Danner

Stephan Härtel

Ingolf Steffan-Dewenter

Alexander Keller

Acknowledgements

My thanks go to Ingolf Steffan-Dewenter, Stephan Härtel, Karin Steijven, Alexander Keller, Susanne Schiele and the whole group of our department Zoologie III, Universität Würzburg, for great support and collaboration. Thanks to Vroni for field work support in Bayreuth and staying in contact, and to Thorsten Steinhäuser for introducing me into a respectful way of beekeeping. Thanks to all the bees for staying (mostly) patient during the field work and for showing me the way...

Affidavit

I hereby declare that my thesis entitled: "Honey bee foraging in agricultural landscapes" is the result of my own work. I did not receive any help or support from commercial consultants. All sources and / or materials applied are listed and specified in the thesis. Furthermore I verify that the thesis has not been submitted as part of another examination process neither in identical nor in similar form.

Eidesstattliche Erklärung
Hiermit erkläre ich an Eides statt, die vorliegende Dissertation: „Honey bee foraging in agricultural landscapes", eigenständig, d.h. insbesondere selbständig und ohne Hilfe eines kommerziellen Promotionsberaters, angefertigt und keine anderen, als die von mir angegebenen Quellen und Hilfsmittel verwendet zu haben. Ich erkläre außerdem, dass die Dissertation weder in gleicher noch in ähnlicher Form bereits in einem anderen Prüfungsverfahren vorgelegen hat.

Würzburg, den 13.10.2016

Nadja Danner