

Scales of bacteria interactions on the leaf surface

Dissertation

zur Erlangung des mathematisch-naturwissenschaftlichen Doktorgrades

„Doctor rerum naturalium“

der Georg-August-Universität Göttingen

im Promotionsprogramm für Agrarwissenschaften (PAG)

der Georg-August University School of Science (GAUSS)

vorgelegt von

Daniel Sebastian Esser

aus Frechen

Göttingen, 2015

Betreuungsausschuss

.....
Prof. Dr. Kerstin Wiegand, Abteilung Ökosystemmodellierung, Fakultät für Forstwissenschaften und Waldökologie, Georg-August-Universität Göttingen

.....
Prof. Dr. Martin Schlather, Lehrstuhl für Stochastik und ihre Anwendung, Institut für Mathematik, Universität Mannheim

.....
Prof. Dr. Johan Leveau, Department of Plant Pathology, University of California, Davis, USA

Mitglieder der Prüfungskommission

Referentin: Prof. Dr. Kerstin Wiegand, Abteilung Ökosystemmodellierung, Fakultät für Forstwissenschaften und Waldökologie, Georg-August-Universität Göttingen

Korreferent: Prof. Dr. Martin Schlather, Lehrstuhl für Stochastik und ihre Anwendung, Institut für Mathematik, Universität Mannheim

Weitere Mitglieder der Prüfungskommission

Prof. Dr. Johan Leveau, Department of Plant Pathology, University of California, Davis, USA

Prof. Dr. Joachim Saborowski, Abteilung Ökoinformatik, Biometrie und Waldwachstum, Fakultät für Forstwissenschaften und Waldökologie, Georg-August-Universität Göttingen

Prof. Dr. Teja Tschardt, Abteilung Agrarökologie, Fakultät für Agrarwissenschaften, Georg-August-Universität Göttingen

Prof. Dr. Stefan Vidal, Abteilung Agrarentomologie, Fakultät für Agrarwissenschaften, Georg-August-Universität Göttingen

Tag der mündlichen Prüfung: 15.02.2016

Es ist nicht genug zu wissen – man muss es auch anwenden.

aus

Wilhelm Meisters Wanderjahre

Johann Wolfgang von Goethe

Table of Contents

Index of Figures	iii
Index of Tables	iv
Abbreviations	iv
Index of the most important symbols	v
Chapter 1 – Introduction	1
How do bacteria perceive their environment? The ecology of leaf-colonizing bacteria	2
Chapter 2 – Methodological approach	7
What is the appropriate spatial scale for observing leaf-colonizing bacteria?	7
How can we assess the spatial scale of bacterial interactions on the leaf surface?	8
Chapter 3 – Are observed bacterial distribution patterns on leaves biased under laboratory conditions?	15
A measure of entropy	16
Segregation index	18
Integrity	18
Comparing pair correlation functions	19
Pattern reconstruction	19
Modeling disturbance	20
Results and Discussion	20
Chapter 4 – Spatial scales of interactions among bacteria and between bacteria and the leaf surface	26
Abstract	26
Introduction	27
Methods	31
Results and Discussion	36
Conclusions	44
Chapter 5 – Biomimetic surfaces help detecting micro-scale effects of topography on bacterial leaf colonization	47
Abstract	47
Introduction	48
Materials and methods	50
Results	55
Discussion	61

Chapter 6 – Scales of interactions between a point pattern and a line pattern: an extension to current methods in spatial statistics	64
Abstract	64
Introduction	65
Materials and Methods	67
Results	73
Discussion	77
Chapter 7 – Consolidated discussion	80
Limitations and future directions	83
References	86
Supplementary Materials	99
Supplementary materials to Chapter 3	99
Supplementary materials to Chapter 4	100
Supplementary materials to Chapter 5	107
Supplementary materials to Chapter 6	110
Abstract	112
Acknowledgments	114
Curriculum vitae	115
Doktoranden-Erklärung der Georg-August-Universität Göttingen	118

Index of Figures

Figure 1: Microbial interactions in the phyllosphere	4
Figure 2: A point pattern illustrating the limitations of numerical summary statistics.	10
Figure 3: Comparison of entropy, segregation and integrity values	21
Figure 4: Leaf colonizing bacteria observed with and without cover slip	23
Figure 5: Explanation of cross-type pair correlation functions	29
Figure 6: Bacterial growth on leaves	36
Figure 7: Intraspecific bacterial interactions on leaves	38
Figure 8: Interspecific bacterial interactions on leaves	40
Figure 9: A typical bacterial colonization pattern on leaves	41
Figure 10: Bacterial interactions with structural elements of the leaf surface	42
Figure 11: Conceptual model of bacteria colonizing plant leaves	46
Figure 12: Artificial bean leaf surfaces made from PDMS	52
Figure 13: Explanation of pair correlation functions	53
Figure 14: Bacterial growth on artificial leaves	56
Figure 15: Intraspecific bacterial interactions on artificial leaves	57
Figure 16: Differences between intraspecific interactions on real and artificial leaves	57
Figure 17: Interspecific bacterial interactions on artificial leaves	58
Figure 18: Differences between interspecific interactions on real and artificial leaves	58
Figure 19: Bacterial interactions with structural elements of the artificial leaf surface	59
Figure 20: Differences between interactions with real and artificial leaf surface	60
Figure 21: Bacteria colonizing the grooves between epidermal cells	66
Figure 22: LPPCFs of artificial point patterns interacting with leaves	74
Figure 23: Differently fine segmentations of a pattern of curves	75
Figure 24: The effect of inadequately segmented curves	76
Figure 25: Testing the performance of LPPCFs compared to other methods	77

Index of Tables

1:	Testing for disturbance in bacterial colonization patterns on bean leaves	21
2:	Overview of experiments on artificial leaf surfaces	52
Table S4.1:	Detailed overview of samples collected on real leaves	100
Table S5.1:	Detailed overview of samples collected on artificial leaves	107

Abbreviations

AHL	N-Acyl Homoserine Lactones. A class of bacterial signaling molecules.
BPS	Biologically Patterned Surface. An artificially designed surface designed to mimic the properties of a biological surface
CSR	Complete Spatial Randomness. A null model where the distribution of points a point process is independent from the location of other points or objects.
DsRed	DiscoSoma red. A red fluorescent protein isolated from the coral <i>Discosoma spec.</i> It was used in our studies to visualize different types of bacteria.
FISH	Fluorescent <i>In-Situ</i> Hybridization. A technology for introducing fluorescent marker proteins such as GFP into living organisms.
GFP	Green Fluorescent Protein. A green fluorescent protein isolated from the jelly fish <i>Aequorea victoria</i> . It was used in our studies to visualize different types of bacteria.
IAA	Indole-3-Acetic Acid. A common plant hormone that also plays a vital role in plant-microbe interactions.
LPPCF	Line-to-Point Pair Correlation Function. Our extension of the PCF for the analysis of spatial correlations between point- and line-like entities
PCF	Pair Correlation Function. The central spatial analysis tool used in this dissertation
PDMS	Polydimethylsiloxane. A liquid plastic used in Chapter 5 of this thesis to produce BPSs mimicking bean leaves.

Index of the most important symbols

$\mathbf{1}(\cdot)$	Indicator function that takes the value 1 if the condition (\cdot) is true and 0 otherwise.
λ_l	The intensity of Φ_l or the line density in ϕ_l .
λ_p	The intensity of Φ_p or the point density in ϕ_p .
π	A constant of value 3.14159 ...
$\varrho(r)$	The product density.
Φ_l	A line process (fibre process)
ϕ_l	A line pattern, i.e. a realization of the point process Φ_l . the number of lines with at least one point in W is $\Phi_l(W)$.
Φ_p	A point process.
ϕ_p	A point pattern, i.e. a realization of the point process Φ_p . The number of points in W is $\Phi_p(W)$.
$d(x_1, x_2)$	The distance between two points x_1 and x_2 .
$g(r)$	A pair correlation function.
$g_{11}(r), g_{12}(r)$	The univariate and partial (=cross-type) pair correlation function.
$g_{lp}(h)$	A line-to-point pair correlation function.
g_{MID}, g_{MIN} g_{MEAN}, g_{POT}	Specific line-to-point pair correlation functions based on different line-to-point pair correlation functions (Chapter 6).
h	The spatial scale or the distance between a point and a line segment or curve.
r	The spatial scale or the distance between two points in a homogeneous pattern
W	An observation window.
W_x	the translated window $\{x + z : z \in W\}$.
$ W $	The area of W
x, x_p, x_q	Points in space.
x_l	A line segment or curve in space.
$ x_l $	The length of the l -th line segment or curve.

Chapter 1 – Introduction

Microbes are ubiquitous inhabitants of our environment. Soils, plants, animals, oceans, and atmosphere are colonized by large quantities of microbial individuals living in complex microbial communities. The global total plant leaf surface, for example, has been estimated at more than 1 billion square kilometers (Vorholt 2012) and the average number only of bacteria on leaves has been estimated at 10,000 to 100,000 individuals cells per square millimeter of leaf surface (Lindow and Brandl 2003). Both figures taken together underline the ubiquity of bacterial activity on a global scale. Nevertheless, we have only limited knowledge of their spatial distribution and the underlying ecological processes at global scale, at the scales of landscapes, hosts, or bacterial individuals. In this dissertation, I studied leaf-colonizing bacteria, their interactions with the environment, and the spatial scale at which these interactions can be observed. The goal of the project was to improve our understanding of microbial leaf surface colonization processes on an individual level. This understanding is crucial to the interpretation of bacterial patterns at larger scales, e.g. at the level of colonies or whole leaves. The research presented in this dissertation is expected to enhance our understanding of how plant-colonizing bacteria establish on plant surfaces and thus to open avenues towards efficient and sustainable measures for biological control of bacterial plant pathogens.

The study of microbial leaf colonization on the level of bacterial individuals raises several fundamental questions: 1) How do bacterial individuals perceive their environment? 2) How can we assess the spatial scale of bacterial interactions on the leaf surface? 3) What is the appropriate spatial scale of observation? 4) Are the observed patterns equivalent to natural patterns or are they biased under laboratory conditions? The introductory chapters 1, 2, and 3 of this dissertation are aimed at discussing these questions in the context of the current state of research. Chapter 1 covers the current knowledge about microbial life on plant leaves. Chapter 2 discusses the above aspects of spatial scale-specific analysis. Chapter 3 presents results of an auxiliary study of how bacterial distribution patterns on leaves observed under the microscope compare to natural patterns. Given the cross-disciplinary nature of the dissertation project, i.e. microbiology and ecology as well as application and theory of spatial statistics, these chapters give a more detailed overview of these topics than was required for the publication of the core research chapters 4, 5, and 6 of this thesis in scientific journals. In Chapter 4, the distribution of leaf-colonizing bacteria

on bean leaves is analyzed. We show how bacteria interact with their environment at different spatial scales and how the interplay of these interactions explains the complex bacterial colonization patterns on plant leaves. Chapter 5 contrasts the underlying processes as being either physical, i.e. driven by leaf topography, or being driven by leaf biological properties. Chapter 6 presents an extension of current methods in spatial statistics that allows the study of the distribution of bacterial individuals along linear structures such as the network of grooves between leaf epidermal cells. The new method was then applied to further explore this important interaction of bacteria with the leaf surface. The concluding Chapter 7 summarizes the most important results of the research project and brings them into a broader context.

How do bacteria perceive their environment? The ecology of leaf-colonizing bacteria

The term 'phyllosphere' was independently coined by F. T. Last (1955) and J. Ruinen (1956). In analogy to the term rhizosphere, it describes the 'characteristic milieu' (Ruinen 1956) that the leaf surface as an environment provides to its microbial colonizers. More precisely, the leaf surface, including a thin surrounding gaseous boundary layer, provides a habitat that differs physically, chemically, and biologically from the surrounding environment, i.e. the leaf tissue and the free atmosphere. The importance of the phyllosphere as an object of study lies, amongst other things, in its role as an active site of many plant pathogens and as a habitat for human pathogens. Both issues have pressing implications for food production but microbial leaf colonization has also been studied in other fields such as forestry (e.g. Peñuelas *et al.* 2012; Kembel *et al.* 2014; Griffin and Carson 2015) or remediation of air pollutants (Bringel and Couée 2015).

The phyllosphere has frequently been described as a harsh environment for microbial colonizers (Lindow and Brandl 2003; Leveau 2006; Vorholt 2012). Other harsh environments are usually defined by the presence of extreme values for an environmental factor such as excessive heat around black smokers on ocean floors (Blöchl *et al.* 1997) or high amount of toxins in contaminated soils (Nies 2000). The phyllosphere, however, is considered as an extreme environment for the wide range and frequent fluctuations in several environmental factors. Compared to other microbial habitats such as the rhizosphere, colonizers of the phyllosphere are exposed to excessive amounts of ultra-violet light. This is expected to be the reason why most leaf-colonizing bacteria are

pigmented and often have well developed DNA-repair abilities (Sundin and Jacobs 1999; Jacobs, Carroll and Sundin 2004; Atamna-Ismaeel *et al.* 2012). Local weather conditions frequently alter light and water availability, which requires adaptations conferring drought tolerance as well as adaptations to high and low temperatures. Rain and wind cause mechanical stress to the leaf surface which may challenge microbial attachment to the leaf surface. The rapid and frequent changes in UV radiation, temperature, and water availability define the phyllosphere as a harsh environment for microbial colonizers compared to more homeostatic habitats such as the leaf interior.

Tied to the environmental factors are chemical properties of the leaf surface. The leaf epidermis is covered by a cuticle that regulates transpiration and gas exchange as well as protects the leaf interior from mechanical stress (Riederer and Müller 2006). The hydrophobic nature of the cuticle impedes microbial colonization and adds to the self-cleaning capability of leaves. Spread across this cuticle is a wealth of metabolic products from microbial and non-microbial sources with nutritional, antimicrobial, regulatory, and communication functions for microbes. From bioreporter studies, it is known that many of these substances are heterogeneously distributed on the leaf surface. Important examples include the distribution of water (Axtell and Beattie 2002), fructose and sucrose (Leveau and Lindow 2001), nitrogen (Parangan-Smith and Lindow 2013), and iron (Joyner and Lindow 2000). Similar biosensors exist for signaling molecules (Deng *et al.* 2014; Rai, Rai and Venkatesh 2015) and environmental contaminants (Stiner and Halverson 2002; Liu *et al.* 2010) but their spatial distribution in the phyllosphere has not been studied yet. Besides this micro-scale heterogeneity the amounts of single compounds may vary considerably between the leaves of a single host plant (Fiala *et al.* 1990; Mercier and Lindow 2000).

Finally, biological factors such as facilitation and competition within microbial communities (reviewed in Meyer and Leveau 2011), as well as plant host responses (Dangl and Jones 2001; Conrath, Pieterse and Mauch-Mani 2002), or predation (O'Rourke *et al.* 2015) shape microbial life in the phyllosphere. Considering the immense microbial diversity found on leaves, extensive networks of inter-microbial and host-microbial interactions may exist (Figure 1) but we are only beginning to grasp their complexity (Ponomarova and Patil 2015). Common ecological processes such as direct competition for nutrients or predation are complemented by more complex processes such as quorum sensing (Hosni *et al.* 2011) or the cooperative interspecific production of extracellular matrix compounds (Morris and Monier 2003). Quorum sensing regulates population

density-dependent changes in microbial physiology induced by microbial signaling molecules (Lv *et al.* 2013). Ambient concentrations of these molecules are increased in regions of high microbial densities and trigger reactions such as locomotion in order to avoid competition or the production of compounds for an extracellular matrix. Such gelatinous matrices that encase groups of microbes are known to improve microbial fitness in the phyllosphere, e.g. by filtering environmental factors such as antimicrobial compounds or drought (Costerton *et al.* 1995; Morris and Monier 2003). For some microbial colonizers it has been shown that the colonizers can change the physiology of their hosts (Yamada 1993), e.g. by the production of growth hormones such as ethylene or indole acidic acid (IAA, Brandl and Lindow 1998). Conversely, plants are able to alter microbial physiology, e.g. by the excretion of specialized compounds that mimic microbial communication molecules such as *N*-acyl homoserine lactones (AHLs) (Teplitski, Robinson and Bauer 2000).

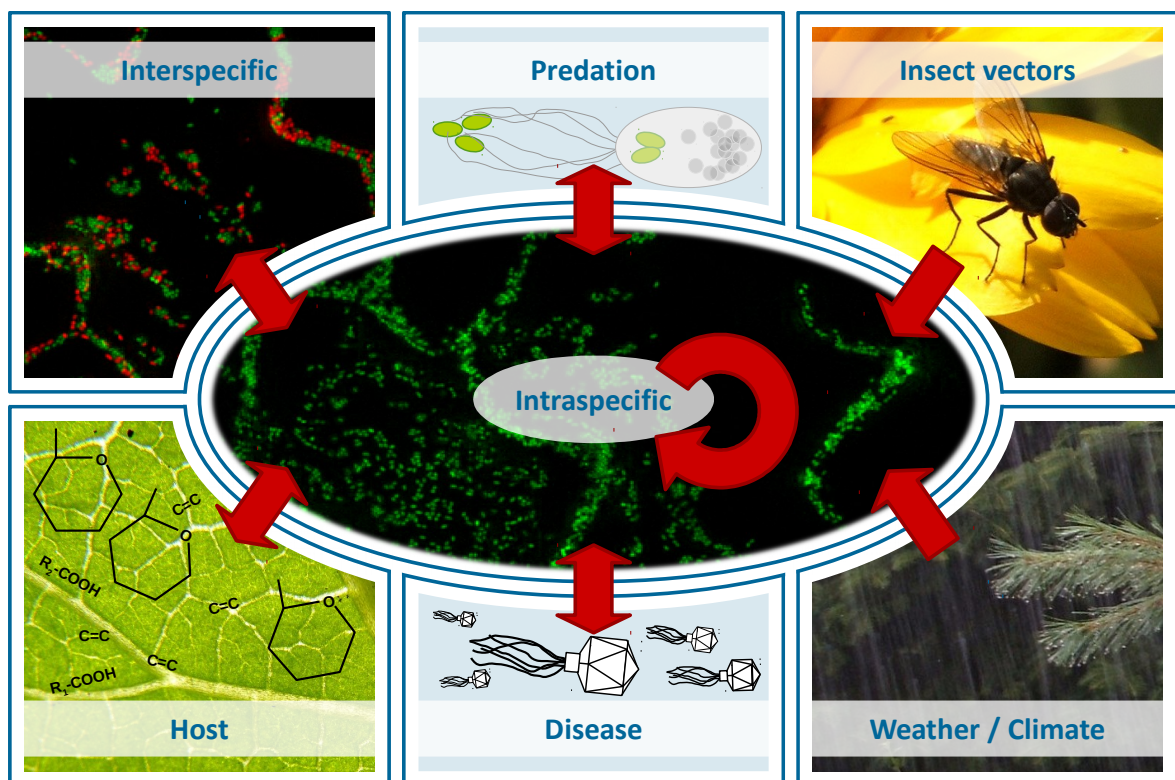


Figure 1: Microbial colonizers of the phyllosphere are affected by and interact with their environment in numerous ways – physically, chemically, biologically. Shown are seven exemplary classes of bacterial interactions on the plant leaf.

The interplay of physical, chemical, and biological factors in the phyllosphere are drivers of an astonishing diversity in microbial taxa, chemical compounds, and interactions, and as consequence, a diversity of phyllosphere research directions. These include but are not

limited to the study (Lindow and Brandl 2003; Vorholt 2012; Berg *et al.* 2014) and modeling (Pérez-Velázquez *et al.* 2012; Kreft *et al.* 2013; van der Wal *et al.* 2013) of all levels of microbial diversity and dynamics, mechanisms and determinants of microbial community composition (Knief *et al.* 2010; Schlaeppi and Bulgarelli 2014), as well as bio-film formation and succession (Blakeman 1985; Yang *et al.* 2001; Whipps *et al.* 2008; Redford and Fierer 2009; Tecon and Leveau 2012; Maignien *et al.* 2014; Copeland *et al.* 2015). Furthermore, the spatial distribution of microbial cells (Monier and Lindow 2004; Redford *et al.* 2010; Remus-Emsermann *et al.* 2014) and processes connected to immigration, growth, and emigration (Dunne 2002; Mattick 2002; Dechesne *et al.* 2010; Yu *et al.* 2014; Wackett 2015) are of general interest in phyllosphere ecology, as well as the exchange of genetic information (Espinosa-Urgel 2004; Pontiroli *et al.* 2009), host-pathogen interactions (Trouvelot *et al.* 2014), and the means of biological control of plant pathogens (Kim *et al.* 2011).

Evidently, the phyllosphere constitutes a diverse and complex environment. Its colonization by microbes is governed by a large family of interrelated processes (Figure 1). The outcome of these concurrent processes are complex *colonization patterns* of microbes on leaves. These patterns are the central study object of this dissertation. My research was built on the premise that an understanding of the complex microbial colonization patterns on leaves and the processes that created them starts at the level of the *bacterial individual*.

Environmental factors such as water or carbon availability are difficult to measure spatially explicit in the phyllosphere. Bacterial bioreporter strains have been genetically engineered that produce fluorescent biomolecules in the presence of a certain target substance. Thus they can be used to visualize the presence and, within limits, the local concentration of these molecules. Such bioreporters are now available for many substances (e.g. Joyner and Lindow 2000; Leveau and Lindow 2001; Axtell and Beattie 2002) but they only measure environmental conditions at the location of microbial individuals. Although these bioreporters provide important information on the heterogeneity in the distribution of environmental factors, explaining bacterial distribution based on this information would be circular. Consequently, this dissertation focuses on the analysis of correlations between the locations of bacterial individuals as well as between bacteria and certain morphological features of the leaf surface. Such spatial correlations were assumed to be indicative of bacterial interactions on the leaf surface. The spatial analysis required an unbiased observation of spatial patterns formed by these individuals under controlled conditions and

during an early phase of plant leaf colonization.

The key goal of this dissertation was a quantitative description of bacterial colonization patterns on plant leaves and the exploration of spatial correlation between conspecific and heterospecific bacterial individuals on the one hand and spatial correlations between bacterial individuals and morphological features of the leaf surface on the other hand.

Chapter 2 – Methodological approach

What is the appropriate spatial scale for observing leaf-colonizing bacteria?

One of the major goals in ecology is to gain a better understanding of the processes and interactions that govern life on Earth. Within the context of this dissertation, *process* refers to a more uni-directional or basic causality such as the increase in nutrient availability around a glandular trichome in the phyllosphere due to excretion of metabolites from the leaf interior. Similarly, the term *interaction* would describe a bi-directional chain or a network of processes on a higher level such as an increase in bacterial population density near glandular trichomes due to increased nutrient availability. In most cases, however, a strict differentiation of the two expressions will be difficult.

Probably all ecological processes can somehow be put into a spatial context. Given time, a process will likely translate into a spatial (and temporal) pattern. Consequently, the analysis of spatial patterns yields the opportunity to 'derive hypotheses on the nature of the underlying processes producing the pattern' (Wiegand and Moloney 2014, p. xvii). Inferring process from pattern has been labeled the 'ultimate goal' (Wiegand and Moloney 2014, p. xix) of spatial pattern analysis in ecology. One particular challenge of such an endeavor is that the ecology of an organism is usually ruled by a variety of facilitating or antagonizing processes that, in the worst case from the viewpoint of the researcher, might be canceling out each other at least at some spatial scale of observation.

The distribution of leaf-colonizing bacteria in space, for example, is controlled by a wide variety of processes that operate at different spatial scales: Monier and Lindow (2004) described the aggregation of bacterial colonizers on leaves at the base of trichomes and in the grooves between undifferentiated epidermal cells. While these patterns, at leaf scale, did not differ qualitatively between the abaxial (bottom) and the adaxial (top) side of leaves (Monier and Lindow 2004), the abaxial surface of leaves tends to be more densely populated (Beattie and Lindow 1999). Furthermore, population densities differ between leaves of a single plant (Hirano *et al.* 1982) and are generally more abundant on lower leaves close to the ground (Andrews and Harris 2000). Finally, abundances of plant pathogens vary on a within-field scale and on landscape scale (Birkhofer *et al.* 2012; Ranjard *et al.* 2013). Such multi-scale complexity in the spatial distribution of bacteria in the landscape can be expected to require multiple processes that operate at different spatial

scales and that produce these patterns. Consequently, a multi-scale spatial analysis of bacterial patterns should help disentangling these diverse processes that rule bacterial life on plants. When viewed from the other direction, an ecological study limited to a single spatial scale comprises the danger of missing out on, or even misjudging the underlying relationships.

How can we assess the spatial scale of bacterial interactions on the leaf surface?

In spatial analyses, the choice of methods depends on the kind of variable to be analyzed. Continuous variables such as temperature can be measured at any location in space whereas discrete entities such as bacterial individuals in space occur at discrete locations only. Although such entities have at least some spatial extent, it is often appropriate to only consider them as individual points in space. All data considered in this dissertation fall into this last category, i.e. spatial point patterns. The location of spheroid bacterial individuals on leaves, for example, can easily be approximated by a set of points with negligible loss of information. The ever-growing field of spatial point pattern analysis continues to develop methods designed to describe point patterns and help to infer the processes that produced them. These methods are summarized in several textbooks targeted at various audiences (Stoyan and Stoyan 1994; Illian *et al.* 2008; Chiu *et al.* 2013; Diggle 2013; Wiegand and Moloney 2014).

The fundamental concept of spatial point pattern analysis is that an observed *point pattern* ϕ is one realization of a *point process* Φ that was observed within a region W , which is called observation window. Multiple point patterns ϕ_i that are all realizations of the same point process Φ are supposed to have the same spatial properties except for some random deviations between the patterns. The point process Φ has an *intensity* λ which is the mean number of points per unit volume. Although leaf surfaces, on the micrometer scale, have a pronounced three-dimensional topography, we only considered point patterns in two-dimensional space (\mathbb{R}^2). Therefore, the intensity is given in points per unit area. Given the fact that leaf topography is rather smooth, the error from not considering the third dimension was expected to be small.

Besides the intensity, the distance between pairs of points is another keystone quantity in spatial point pattern statistics. In our studies, the distance between two points was always the direct (Euclidean) distance. Only in Chapter 6, where correlations between a point

pattern and a pattern of line segments are considered, additional definitions for distance are used. Most summary statistics used to characterize point patterns are either based on the intensity of the point process (estimated from the observed point pattern) or the distributions of inter-point distances, or both. They are compared to the expected values of these quantities, given a predefined null model.

The most common and fundamental null model is *complete spatial randomness* (CSR). Under CSR, each point is placed at random and independently from the location of the other points. The best and most commonly used example of a point process that fulfills the CSR property is the homogeneous Poisson point process, where the number of points in a test set (observation window) placed at arbitrary locations in the pattern follows a Poisson distribution. The CSR null model is an appropriate null model in exploratory analysis because it is the simplest spatial model available and well understood. Nevertheless, more can be learned from complex null models such as models including a gradient or a cluster process (Wiegand and Moloney 2004). These more complex null models can incorporate ecological knowledge about processes or traits of the study organism. The ultimate goal would be to find a null model that generates point patterns that have the same spatial properties as the observed point pattern, based on the different spatial statistical measures. For example, if the dispersal kernel of seeds of a typical tree of a certain species is well known, one could construct a null model based on the parameters of this dispersal. This null model could be used to test if location of seeds of trees of other species, e.g. from the same genus, follows the same dispersal kernel. Such methods have been used, for example, to study the ecology of trees (Yu *et al.* 2009; Fedriani, Wiegand and Delibes 2010), orchids (Jacquemyn *et al.* 2009), and even fossil macro-organisms (Mitchell *et al.* 2015).

The simplest class of summary statistics used to describe point patterns are *numerical summary statistics*. These indices summarize some characteristics of an observed point pattern in a single number. The wide-spread Clark-Evans index R (Clark and Evans 1954), for example, divides the observed mean distance of points to their nearest neighbor \bar{d} by the expected distance \bar{d}_{CSR} , given the points follow CSR. In other words, R is the ratio of the average *nearest neighbor distance* between points in the observed pattern and the expected nearest neighbor distance if the points were randomly distributed. A point for which the distance to its nearest neighbor is about the average nearest neighbor distance in the pattern is considered a *typical point*. This term can also be extended to any other spatial property of the point other than the nearest neighbor distance. Values of $R \approx 1$ indicate a

random distribution of the points whereas values $R > 1$ or $R < 1$ indicate an hyperdispersion or an aggregation of points, respectively. A hyperdispersion of points thus describes a lack (or scarcity) of short distances to the nearest neighbors, compared to randomly distributed points. Such patterns then appear to be more regular, because of the more evenly spaced distribution of the points. Similarly, an aggregation (or underdispersion) of points results in a lack of longer distances to the nearest neighboring points. A considerable limitation of the Clark-Evans index and other numerical summary characteristics such as the intensity is that these measures lose much information about the observed pattern by their attempt to characterize the pattern by a single number. Furthermore, such basic measures are limited by the fact that very different point patterns may still have very similar values for these measures. They are also not scale-specific in the sense that a pattern of regularly-spaced clusters can only be characterized by them to be either clustered or regular (Figure 2).

Functional summary statistics, in contrast, retain much more information about an observed pattern by characterizing the pattern for a range of scales. A wide variety of summary functions have been developed that all capture different characteristics of a point pattern. The nearest-neighbor function $D(r)$ (Hanisch 1984), for example, is the cumulative distribution function of the nearest-neighbor distances d_1 . Clearly, this functional summary statistic contains more information than the Clark-Evans index that only expresses the average nearest neighbor distance. The natural short-sightedness of $D(r)$ may be reduced by looking at multiple distribution functions $D_k(r)$ of the distances d_k to the k -th nearest neighbor.

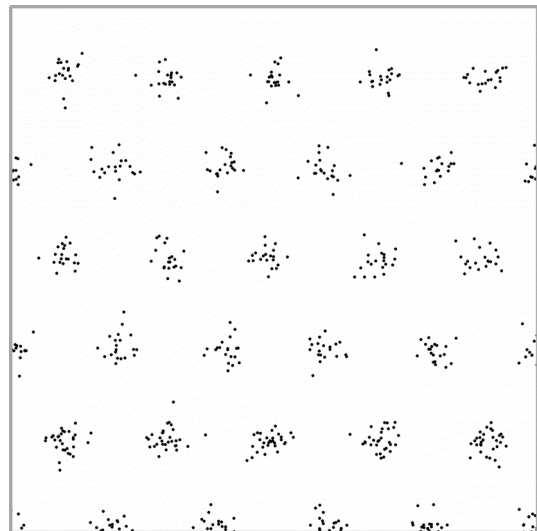


Figure 2: A point pattern illustrating the limitations of numerical summary statistics. Clearly, the points in this pattern are clustered. The clusters are aligned along a lattice resulting in a larger-scale regularity in the pattern. The Clark-Evans index for this pattern is $R = 0.48$, suggesting a regular distribution of points. Numerical summary statistics such as the Clark-Evans index cannot accurately describe such multi-scale distribution patterns.

A similar functional summary statistic is the spherical contact distribution function $H_S(r)$ (Diggle 1983). It too, is a cumulative distribution function of the distances to the closest point but measured from a random location in space that does not need to be part of the point pattern. Compared to the nearest-neighbor function, its reference point is therefore

not a typical point of a point pattern but a *random location* in space. Consequently, it characterizes the open spaces between the points (holes in the point pattern) rather than the clustering of points.

The widely used Ripley's K function $K(r)$ (Ripley 1977) is a normalization of the average number of points within a radius r around a typical point of the point process. For a CSR pattern, $K(r) = r^2$, whereas values $K(r) > r^2$ indicate an aggregation of points at scale r . This means that within a distance r around a typical point of the pattern, one finds more points than expected in a CSR pattern. Similarly, values $K(r) < r^2$ indicate a scarcity of points within a distance r around the typical point. Studying the K -function across a range of distances r allows an advanced analysis of the structure of the point pattern. For example, if all values $K(r) = 1$ for all $r < s$ and all $K(r) > 1$ for all $r > s$, this would suggest that the aggregation of points only sets off at a scale s around a typical point of the pattern.

Ripley's K function, entails a natural 'memory effect' from its cumulative nature. It gives the average number of points *within* a radius r around the typical point. For example, given a strong small-scale (r_1) aggregation, an additional weaker scarcity of points at an intermediate scale (r_2) is difficult to detect because a ring of radius r_2 will include the many points already present within r_1 . The scarcity of points between r_1 and r_2 may fail to compensate the 'memory' of aggregation at scales $< r_1$. The pair correlation function $g(r)$ (PCF, Stoyan and Ohser 1982; Stoyan and Stoyan 1994), which is a normalized derivative of Ripley's K , describes the average number of points *at* a distance r_1 and r_2 separately and therefore has no memory. The same is true for the closely related O-ring function $o(r) = \lambda \cdot g(r)$ (Stoyan and Ohser 1982; Wiegand and Moloney 2004).

All the above mentioned scale-dependent summary statistics characterize different aspects of the observed point pattern and a thorough treatment of a single point pattern should include the study of several of these and other characteristics (Wiegand, He and Hubbell 2013). Given the large amounts of data to be analyzed and the multitude of interactions to be interrogated in this dissertation, only the pair correlation function was considered here. It has been recognized to be the single summary statistic that captures most information from a point pattern and has been recommended as the primary tool in exploratory point pattern analysis (Illian *et al.* 2008; Wiegand, He and Hubbell 2013).

As mentioned above, the pair correlation function $g(r)$ is a normalized derivative of Ripley's K function

$$\text{Eq. 2.1 : } g(r) = \frac{1}{2\pi r} \frac{dK}{dr}.$$

A more formal definition of the pair correlation function includes the product density $\varrho(x_1, x_2)$

$$\text{Eq. 2.2 : } g(x_1, x_2) = \frac{\varrho(x_1, x_2)}{\lambda^2},$$

where $\varrho(x_1, x_2)dx_1dx_2$ gives the probability to find one point of a point pattern in each of two infinitesimally small spheres centered at x_1 and x_2 (Stoyan and Stoyan 1994), respectively. Given the point process is motion invariant, i.e. its properties such as the intensity are both invariant under translation (stationary) and rotation (isotropic) of the observation window, then $\varrho(x_1, x_2)$ and $g(x_1, x_2)$ solely depend on the distance between the points $r = \|x_2 - x_1\|$ and we write

$$\text{Eq. 2.3 : } g(r) = \frac{\varrho(r)}{\lambda^2}.$$

Heuristically speaking, the difference between Ripley's K -function and the pair correlation function is that the former is similar to the average number of points within a *disk* of radius r centered at a typical point of the point pattern, whereas the latter approximates the average number of points on a *ring*, i.e. the margin of a disk, of radius r , centered at a typical point of the point pattern. Clearly, for finite point density λ , the probability to find a point at an exact distance r from another point is zero. This is the reason why the definition of the product density ϱ involves small discs. They ensure positive probabilities while at the same time precluding to find more than one point in a disc. Transferred to the interpretation above, $g(r)$ contains the average number of points on a ring of approximate radius r .

A common challenge not only tied to pair correlation functions arises from the fact that for larger distances r there is a substantial probability that parts of the rings (or discs) lie outside the observation window but no point data is available for these regions. Several so-called *edge-correction* methods have been developed to compensate for this effect (summarized in Illian *et al.* 2008). Edge correction methods introduced by Ripley (isotropic edge correction; 1976), Stoyan & Stoyan (1994), Ohser & Mücklich (2000) and Wiegand & Moloney (2004). include weights in the estimators for the summary statistics. The weights are either based on the fraction of the area of a disc of radius r (Ripley 1976) or ring of radius r and width Δ_r (Wiegand and Moloney 2004) centered at a point x_i in the

observation W , or they are based on the probability that start- and endpoint of a vector of random direction φ and length r both lie within the observation window (Stoyan and Stoyan 1994; Ohser and Mücklich 2000). For our research, we chose the isotropic edge correction which is also the default method used in pair correlation function estimation in the *spatstat* package (Baddeley and Turner 2005).

The concept of the pair correlation function can also be transferred to studies where the relative distribution of two point patterns is of interest. For example, microbiologists might be interested in the question if individuals of two bacterial strains aggregate close to each other, or are avoiding each other, or if there is no spatial correlation between them. Such questions can be targeted with the partial pair correlation function which also goes back to Stoyan and Ohser (1982). If we assume two bacterial strains in a stationary point process with points of type 1 (strain 1) and points of type 2 (strain 2), the partial pair correlation function is defined as

$$\text{Eq. 2.4 : } g_{12}(r) = \frac{\varrho_{12}(r)}{\lambda_1 \lambda_2},$$

where $\varrho_{12}(r)dx_1dx_2$ gives the probability to find a point of type 1 in an infinitesimally small disk of area dx_1 and a point of type 2 in a similar disc of area dx_2 and the distance between the centers of the two discs equals r (Stoyan & Stoyan 1994). The uni-variate pair correlation function (Eq. 2.3) can be considered the special case of the partial pair correlation function where only points of the same type are considered.

In application, the above stationarity assumption is rather necessary than realistic. Ecology often deals with phenomena along gradients in space and the abundances of organisms will often change along these gradients. Small observation windows (or low abundances within the window) furthermore might suggest an absence of stationarity although the point process, on a larger scale, is actually stationary. The stationarity assumption can be loosened by either testing the summary statistic against a non-stationary null model (e.g. inhomogeneous Poisson point process) or by applying the inhomogeneous pair correlation function $g_{\text{inhom}}(r)$ (Baddeley, Møller and Waagepetersen 2000). For the inhomogeneous pair correlation function, one assumes a large-scale heterogeneity (above the largest scale r up to which we want to evaluate our pair correlation function) which is expressed as changes in local intensity $\lambda(x)$ of the point process, e.g. along a gradient in space. This means that the intensity of the point process is no longer a constant but a function in space.

Throughout our studies, we applied the default estimators of the pair correlation function

implemented in the *spatstat* package (Baddeley and Turner 2005) in *R* (R Core Team 2013). They are based on the estimators formulated by Stoyan & Stoyan (1994). For the homogeneous, uni-variate pair correlation function, this is

$$\text{Eq. 2.5 : } \hat{g}(r) = \frac{1}{2\pi r \lambda^2} \sum_{i=1}^n \sum_{j=1, j \neq i}^n \frac{\mathbf{k}_h(r - \|x_j - x_i\|)}{|W_{x_j} \cap W_{x_i}|}.$$

Here, $\mathbf{k}_h(t)$ is the Epanechnikov kernel with $t \geq 0$ and standard deviation $\sigma = h/\sqrt{5}$:

$$\text{Eq. 2.6 : } \mathbf{k}_h(t) = \begin{cases} \frac{3}{4h} \left(1 - \frac{t^2}{h^2}\right) & (-h \leq t \leq h) \\ 0 & \text{otherwise.} \end{cases},$$

The Epanechnikov kernel smooths the results of the pair correlation function which helps to reduce stochasticity in the results. The remaining term in the sum in Eq. 2.5, i.e. $b_{ij}/|W_r|$, with $r = \|x_j - x_i\|$ is Ripley's isotropic edge correction (Ripley 1977), where $|W_r|$ is the area of the translated window W_r .

Analogously, the estimator for the partial pair correlation function is

$$\text{Eq. 2.7 : } \hat{g}_{12}(r) = \frac{1}{2\pi r \lambda_1 \lambda_2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \frac{\mathbf{k}_h(r - \|x_{2,j} - x_{1,i}\|)}{|W_{x_{2,j}} \cap W_{x_{1,i}}|}$$

which is also the default estimator used in the *spatstat* package.

The estimator of the inhomogeneous PCF accounts for heterogeneity in the pattern by using weights based on estimates of the local intensity $\lambda(x)$ around location x .

$$\text{Eq. 2.8 : } \hat{g}_{inhom}(r) = \frac{1}{2\pi r} \sum_{i=1}^n \sum_{j=1, j \neq i}^n \frac{\mathbf{k}_h(r - \|x_j - x_i\|)}{\lambda(x_i)\lambda(x_j) |W_{x_{2,j}} \cap W_{x_{1,i}}|}.$$

The estimator of the inhomogeneous partial pair correlation function follows analogously from Eq. 2.5, Eq. 2.7, and Eq. 2.8.

Chapter 3 – Are observed bacterial distribution patterns on leaves biased under laboratory conditions?

Methods for observing bacterial colonization of the phyllosphere developed rapidly within the last decades. Simple observations of bacterial colonization patterns on leaves were achieved early by leaf printing experiments (Leben 1965, 1970). Here, naturally colonized leaves were pressed gently on a nutritious agar surface and the bacteria that were transferred to the agar surface grew to colonies. The observed patterns are indicative of the bacterial distribution on leaves. Of course, the observed patterns were biased by the possibility that not all bacterial species transfer equally well onto the agar surface and the fact that different species will multiply at different rates. Moreover, most species will not even be culturable (Rastogi *et al.* 2010, 2012). More direct data about the spatial distribution of bacterial individuals on leaves became available with the advent of electron microscopy (Barnes and Neve 1968; Leben 1969; Mansvelt and Hattingh 1987, 1989). These studies already established the existence of bacterial clusters on the leaf surface and that bacterial colonization was especially associated with leaf structural elements such as trichomes, veins, and stomata. A less invasive alternative to electron microscopy was introduced later in the form of fluorescence microscopy (Lichtman and Conchello 2005). Here, the introduction of genes encoding fluorescent proteins such as the Green Fluorescent Protein GFP into microorganisms (Chalfie *et al.* 1994) allowed for the direct light microscopic observation of bacterial leaf colonization.

The observation of colonization patterns of bacterial individuals on leaves is tightly coupled to the question if an observed microbial pattern is equivalent to the real distribution on leaves or if it has been altered during the preparation of the sample. The process of fixation and gold-plating for electron-microscopy possibly alters the distribution of bacterial individuals on leaves but a quantification of such changes is difficult.

Similarly, adding water or a mounting medium and a cover slip to a light-microscopic preparation could wash away bacterial individuals or aggregates. In order to explore this issue, we compared a series of fluorescence micrographs with and without cover slip of bacteria colonizing bean leaves. Therefore, bean leaves were prepared, inoculated, and incubated using the same procedures described in detail in Chapter 4 of this dissertation. Before adding mounting medium and the cover slip, we took micrographs of the

fluorescent bacteria near special landmarks of the leaf surface, e.g. intersections of leaf veins. We used microscope objectives with long working distance (Zeiss EC Plan-Neofluar 10×/0.3, working distance $WD = 5.2$ mm and Zeiss LD Plan-Neofluar 40×/0.6, $WD = 2.9$ mm) that did not require the use of a cover slip. Thereafter, we added mounting medium and fixated a cover slip to the top and took more micrographs at the exact same locations as before using the same 10× and a standard 40× objective (Zeiss EC Plan-Neofluar 40×/0.75).

Such images without cover slip were difficult to obtain. Incubated leaves were covered by water droplets that quickly evaporated once the microscope light was turned on, thereby shifting the plane of focus. Once the water evaporated, the microscope light quickly started damaging the leaf tissue and bacterial colonizers, leaving only a brief window of time for taking a series of pictures in different fluorescent channels and at different planes of focus.

From the 'before' (no cover slip) and 'after' (with cover slip) images, we extracted the location of all green and red bacterial individuals using the free software ImageJ (Schneider, Rasband and Eliceiri 2012). We inspected the observed patterns visually as well as using several statistical tests based on numerical measures that describe the spatial configuration of the leaf colonizers. These measures are presented in the following.

A measure of entropy

For the first measure, we assumed that in a perfectly undisturbed colony, all individuals sit next to each other and are of same identity I . In our studies the identity is the color of the fluorescent marker (DsRed or GFP). All individuals in a colony appear to be descendants of a single colony forming unit (CFU) if they have the same color. Assuming a spherical shape of all individuals, as we had usually observed in our micrographs, each individual may have six adjacent neighbors. For a minimum radius of a cell r_{min} , the minimum distance between the centers of two individuals is $d_{min} = 2 r_{min}$. In spatial statistics this is called a hardcore point process with minimum distance d_{min} (Illian *et al.* 2008).

Our quantity to measure entropy in bacterial distributions is a function of the identity I of an individual i and its six nearest neighbors j and the distances d_{ij} to these six nearest neighbors. For reasons of comparability it is desirable to have a measure that takes only values between 0 and 1. A function of the distances d_{ij} that fulfills this criterion is

$$\text{Eq. 3.1 : } E_D = \frac{1}{n} \sum_{i=1}^n 1 - \frac{6 d_{min}}{\sum_{j=1}^6 d_{ij}},$$

where n is the number of bacterial individuals in the sample. The function is zero if all $d_{ij} = d_{min}$ and it approaches 1 for all $d_{ij} \rightarrow \infty$. Equation 3.1 is a special case of a more general function which accounts for other neighborhood configurations such as the twelve nearest neighbors that are potentially all allowed to touch the focal sphere in a three-dimensional setting. The more general form of Eq. 3.1 can be found in the Supplementary Materials (p. 99).

A suitable measure based on the identity of neighboring cells is the function

$$\text{Eq. 3.2 : } E_I = \frac{1}{n} \sum_{i=1}^n \frac{\sqrt{n_{GFP,i}} + \sqrt{7 - n_{GFP,i}} - \sqrt{7}}{2 + \sqrt{3} - \sqrt{7}}.$$

It only depends on the number $n_{GFP,i}$ of green colored cells within the neighborhood of individual i (also counting i). The entropy E_I is zero if $n_{GFP,i} = 7$ (all green) or if $n_{GFP,i} = 0$ (all red) and peaks for the theoretical case $n_{GFP,i} = 3.5$. The defined cases $n_{GFP,i} = 3$ and $n_{GFP,i} = 4$ are equally representing the maximum achievable entropy. For both cases $E_I = 1$. Equation 3.2 is also a special case of a more general form that can be used for different neighborhoods and more than two identities. The general form of Eq. 3.2 is also given in the Supplementary Materials section (p. 99).

We combined the two functions to one entropy measure using a weighted average which allows to control for the relative influence of the two terms:

$$\text{Eq. 3.3 : } E = \gamma E_D + (1 - \gamma) E_I,$$

where γ is a constant between 0 and 1. It controls the contribution of the distance term to the overall entropy. For $\gamma = 0$, entropy $E = E_I$, for $\gamma = 0.5$ both terms contribute equally to the entropy and for $\gamma = 1$ only the distances between neighbors are considered, i.e. $E = E_D$.

Our combined entropy measure then is

$$\text{Eq. 3.4 : } E = \frac{1}{n} \sum_{i=1}^n \gamma \cdot \frac{6 d_{min}}{\sum_{j=1}^6 d_{ij}} + (1 - \gamma) \cdot \frac{\sqrt{n_{GFP,i}} + \sqrt{7 - n_{GFP,i}} - \sqrt{7}}{2 + \sqrt{3} - \sqrt{7}}.$$

For our study, we chose a parsimonious value of $\gamma = 0.5$, because it requires the least assumptions.

The more general form of equation 3.4 (see Supplementary Materials section, p. 99) can be used for example in experiments where three-dimensional location data is studied. In that

case, each individual would be surrounded by up to 12 direct neighbors. One can also think of an experiment of early colonization stages when only colonies of up to 4 cells are observable. Either way, an adjustment on the number of neighbors n_n is indicated. Similarly, the number of cell types n_c has to be adjusted if more than 2 fluorescent markers (species, strains) are involved in the experiment. The complex equation will always simplify to a form comparable to Eq. 3.4.

Segregation index

Next, we considered the segregation index SI defined by Monier and Lindow (2005). The segregation index is the mean inverse number of colonies in the sample weighted by the relative abundance of the two strains. It takes values between 0 (cells randomly distributed) and 1 (cells highly aggregated). Monier and Lindow (2005) did not exactly specify what constitutes a cluster but it is reasonable to assume that in their study members of a cluster need to be in physical contact with at least one other member of the cluster. For our purposes, we assumed that an individual belongs to a cluster if the distance to the nearest member of the cluster is less than a certain aggregation distance d_{agg} . A value of $d_{agg} = 8 \mu m$ was chosen, as this minimized the variance of SI between samples when we calculated the segregation index for d_{agg} values between 1 and 40 μm in steps of 1 μm . Different from the Monier and Lindow (2005), who determined the number of colonies per aggregate, we counted the number of red and green colonies in each field of view. Thus, we retrieved information about the changes in segregation within the colonization pattern on a larger spatial scale, which appeared to better fit a study of disturbance. Using an aggregation distance to define and count colonies also changed the original purpose of the segregation index. It was designed to quantify the intermingling of differently colored clusters. In our study, we take advantage of its property to compile lots of information about the colonization structure in a single number and of its sensitivity to different types of disturbance in those colonization patterns.

Integrity

Given an aggregation distance d_{agg} that defines the association of a point to a cluster, the integrity I of a point pattern can be defined as

$$Eq. 3.5 : \quad I = \frac{\sum_{i=1}^n \sum_{j=i+1}^n \mathbf{1}(d_{ij} < d_{agg})}{\sum_{i=1}^{n-1} \sum_{j=i+1}^n d_{ij}} .$$

This measure does not require information on the identity of bacterial individuals and thus could also be applied to colonization patterns of general bacteria, e.g. on field-grown leaves treated with a single fluorescent dye such as acridine orange. Heuristically speaking, this measure is the ratio of the sum of all inter-individual distances d_{ij} (denominator) and the sum of the fraction of d_{ij} that is smaller than the predefined aggregation distance d_{agg} (numerator).

Then $I = 1$, if $\max(d_{ij}) < d_{agg}$. This is a problem, when clusters grow to sizes larger than d_{agg} . Then even a perfectly intact (undisturbed) colony will have $I < 1$. This can occur especially during anisotropic colony development. Integrity is absent ($I = 0$) if $\min(d_{ij}) > d_{agg}$. Disturbance to such patterns can only be measured if the disturbance increases the level of aggregation in the pattern.

This illustrates a general problem of disturbance measures: Disturbance can either reduce integrity of a colony or move points to areas where they accumulate, e.g. because of topography, and thus increase integrity. The entropy measure defined above is less prone to this effect, but it still occurs if the different species (colors) are well separated in space.

Comparing pair correlation functions

A further approach considered for the analysis of disturbance involved the direct comparisons of the PCFs of patterns with and without cover slip using either paired t -tests (along spatial scale r), or a Kolmogorov-Smirnov test (across the 21 samples). Both methods were hampered by the instability of the PCF at small scales ($r \rightarrow 0$), where small changes in patterns can cause substantial changes in pair correlation values. An exclusion of these scales, however, is not appropriate because small-scale disturbances are also important to consider. Similar 'butterfly effects' also occur at larger scales if the point pattern violates the stationarity assumption required for PCFs. The addition or removal of single points can heavily change the shape of the PCF (own observation).

Pattern reconstruction

We also consider the use of pattern reconstruction (Nothdurft *et al.* 2010) in order to generate 95% simulation envelopes around the observed pair correlation function of an undisturbed pattern (without cover slip). Deviations of the pair correlation function of the potentially disturbed pattern (with cover slip) would then indicate a significant disturbance of the pattern. This approach, more than any other, illustrates the problem of quantifying

disturbance because it measures disturbance by deviation from a simulation envelope that was derived from a set of fully disturbed patterns, i.e. patterns that have second-order properties similar to the original pattern but often have a completely different distribution of points. Furthermore the definition of stop criteria in pattern reconstruction adds ambiguity to the analysis that is difficult to handle.

Modeling disturbance

A promising but laborious solution to the disturbance problem for future research involves the definition of different disturbance processes such as smearing, disruption or bloating of colonies, or the topography-dependent aggregation of cells. These processes could be implemented in an individual-based model (IBM) that is used to add a defined negligible level of disturbance to the original pattern and derive simulation envelopes from the pair correlation functions of these simulated patterns. Once proper definitions and parametrization exist, the IBM-derived envelopes could be used to quantify significant disturbance added to the bacterial colonization by the addition of mounting medium and a cover slip to a microscopic preparation of leaf-colonizing bacteria.

Results and Discussion

The following 1 summarizes the results of the disturbance analysis using the entropy, segregation, and integrity measures presented above. Shapiro-Wilk tests suggested that entropy, segregation, and integrity values did not follow a normal distribution (all p -values < 0.001). We performed one-sided and two-sided Wilcoxon tests, where for the one-sided tests we used the null hypotheses $E_{\text{after}} > E_{\text{before}}$, $S_{\text{after}} < S_{\text{before}}$, and $I_{\text{after}} < I_{\text{before}}$. All three tests suggested that there were no significant changes in the bacterial colonization patterns when a cover slip was added to a preparation. A plot of the measure values after adding the cover slip against the values before adding the cover slip also did not suggest any systematic changes in the pattern (Figure 3).

Table 1: Testing for disturbance introduced to bacterial colonization patterns on bean leaves by adding a cover slip to microscopic preparations.

Sample	Entropy		Segregation		Integrity	
	No cover slip	Cover slip	No cover slip	Cover slip	No cover slip	Cover slip
58b_11	0.572	0.604	0.096	0.076	3050.743	544.381
71a_01	0.735	0.661	0.230	0.090	3.852	7.234
71a_02	0.942	0.971	0.312	0.147	6.394	7.245
71a_03	0.504	0.549	NA	NA	3.291	3.040
71b_04	0.930	0.539	0.063	0.061	449.152	154.182
72a_01	0.788	0.766	0.051	0.033	173.874	472.519
72a_02	0.809	0.868	0.111	0.055	1541.089	721.622
72a_03	0.568	0.602	0.040	0.037	808.240	1031.911
72b_04	0.826	0.903	0.070	0.059	394.787	213.325
72b_05	0.499	0.477	0.096	0.090	203.745	216.497
73a_05_a	0.485	0.482	NA	NA	135.349	236.682
73a_05_e	0.965	0.961	0.027	0.026	169.258	192.891
73a_05_f	0.961	0.957	0.045	0.049	357.147	346.974
73a_05_g	0.951	0.955	0.251	0.500	290.244	305.869
73a_05_h	0.982	0.928	0.118	0.250	4.332	2.041
73a_05_i	0.949	0.945	0.036	0.038	152.300	138.853
73a_05_j	0.899	0.956	0.040	0.039	71.518	133.555
73a_05_k	0.901	0.926	0.088	0.080	10.539	11.405
73a_05_l	0.945	0.944	0.060	0.053	70.674	65.489
73a_05_m	0.969	0.969	0.036	0.036	163.123	152.540
73a_05_n	0.959	0.945	0.096	0.083	65.375	65.758
73a_05_o	0.968	0.972	0.167	0.154	22.137	23.259
73a_05_p	0.956	0.978	0.341	0.500	5.537	6.023
Wilcoxon-	0.4947 (two-sided)		0.1111 (two-sided)		0.9168 (two-sided)	
test (p-value)	0.7628 (one-sided)		0.9484 (one-sided)		0.4584 (one-sided)	

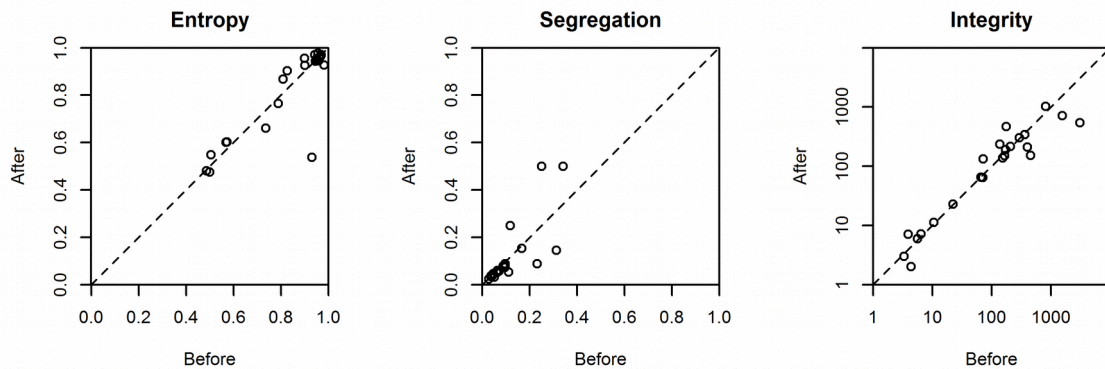


Figure 3: Comparison of entropy, segregation and integrity values of bacterial colonization patterns on bean leaves observed before and after adding a cover slip to the microscopic preparation.

Despite these promising results, it became clear that, at least in theory, even an obviously disturbed pattern could have the same value as the undisturbed pattern for any of the tested indices, e.g. when the whole pattern was shifted or rotated. Therefore, we do not expect our results to allow a final assessment of disturbance. In order to give an impression of how little the observed before/after patterns differed from each other, we refer to the picture pairs in Figure 4. The sum of our results combined with these micrographs suggest that the patterns observed in preparations with cover slip represent the real distribution on leaves sufficiently well. Therefore, we decided for the main studies of this dissertation project (chapters 4-6) to obtain large numbers of micrographs with cover slips rather than few, more difficult to obtain samples without. The increased achievable number of samples was assumed to outweigh a possible bias that might be introduced by adding the cover slip to a preparation.

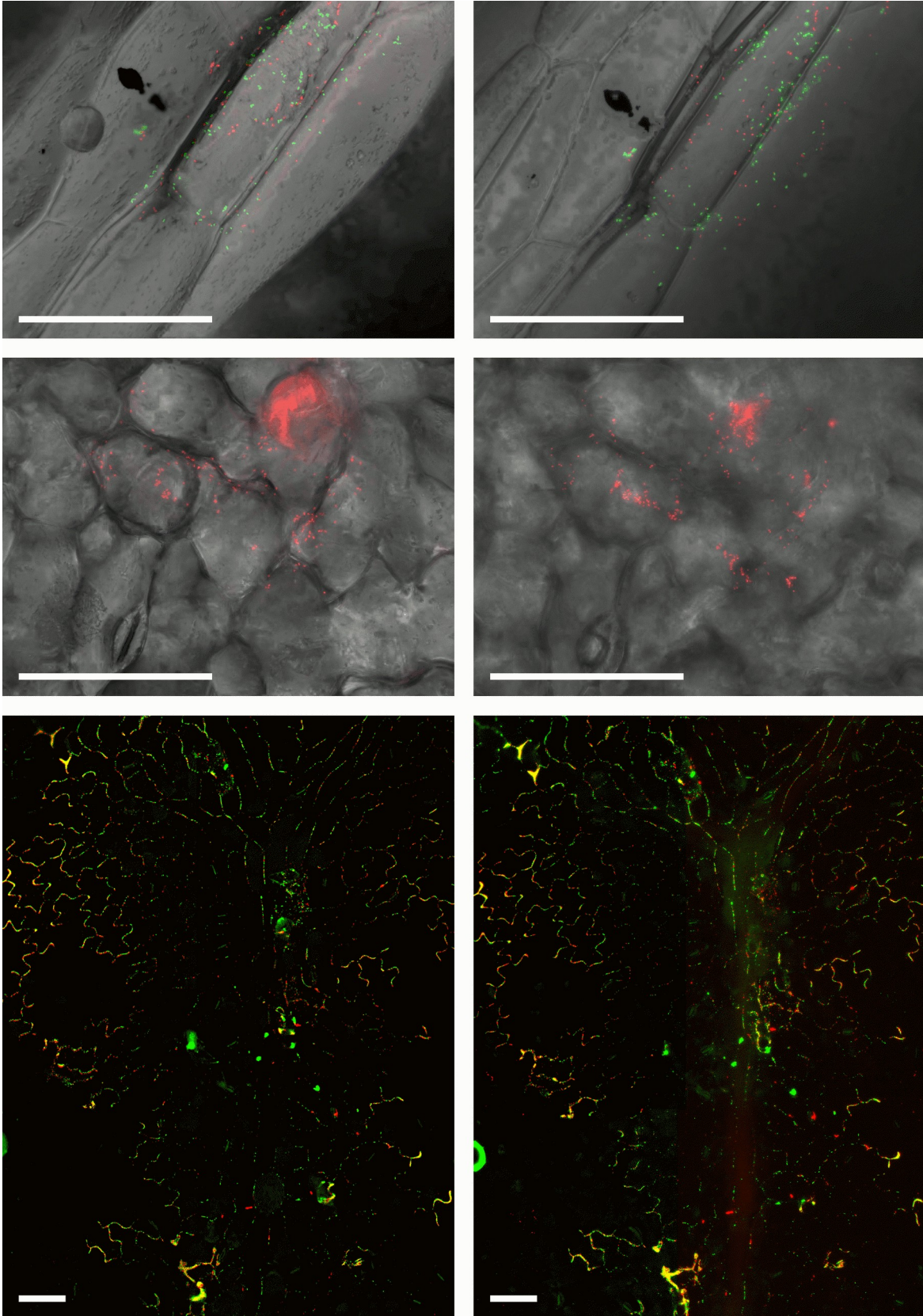


Figure 4: Pages 23 ff.. Micrographs of red and green fluorescent leaf-colonizing bacteria on bean before (left column) and after (right column) adding a cover slip to the microscopic preparation. Bacterial distribution on the leaves is barely effected by this preparation step. Scale bars are 100 μm .

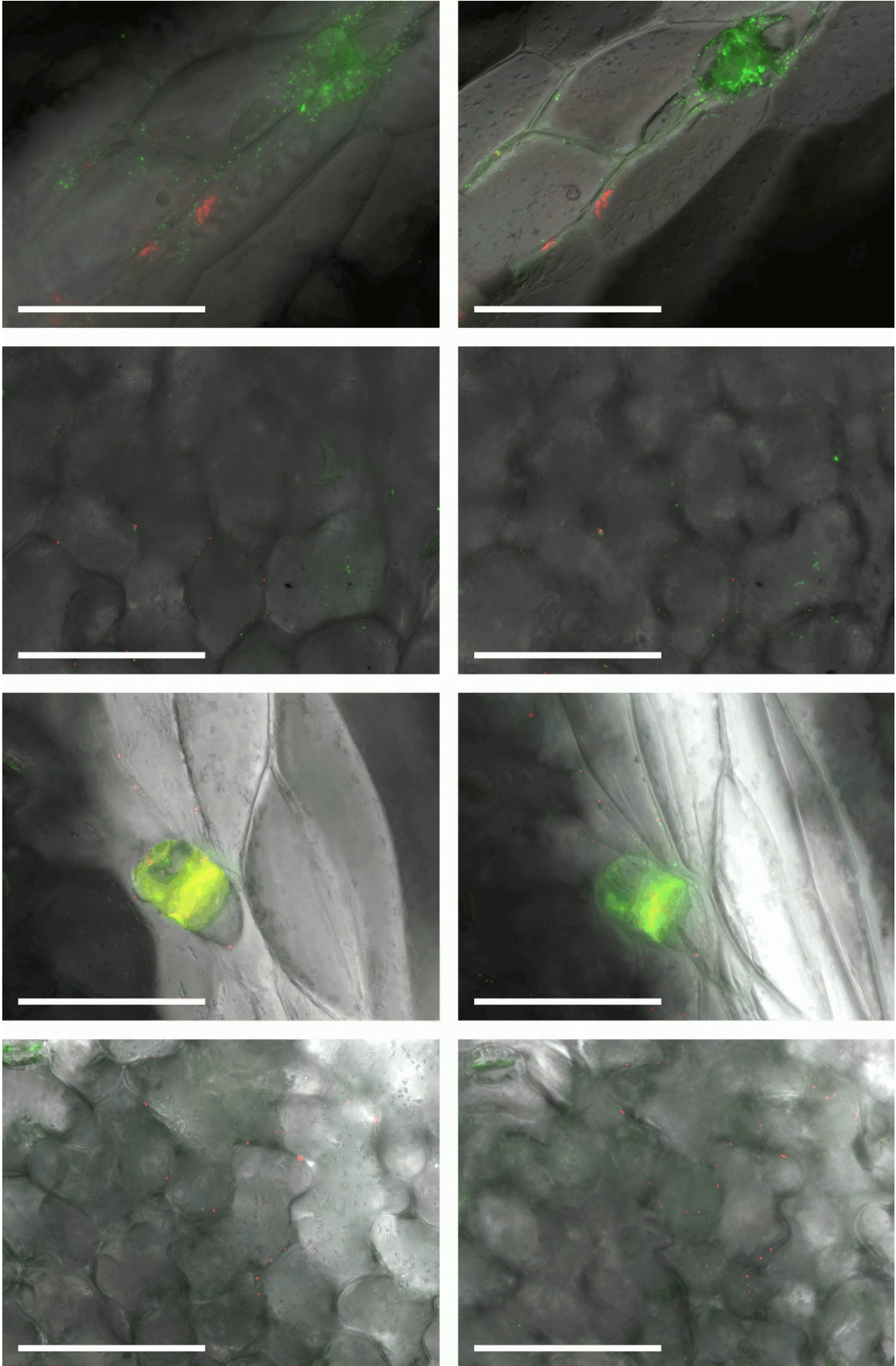


Figure 4 - continuation

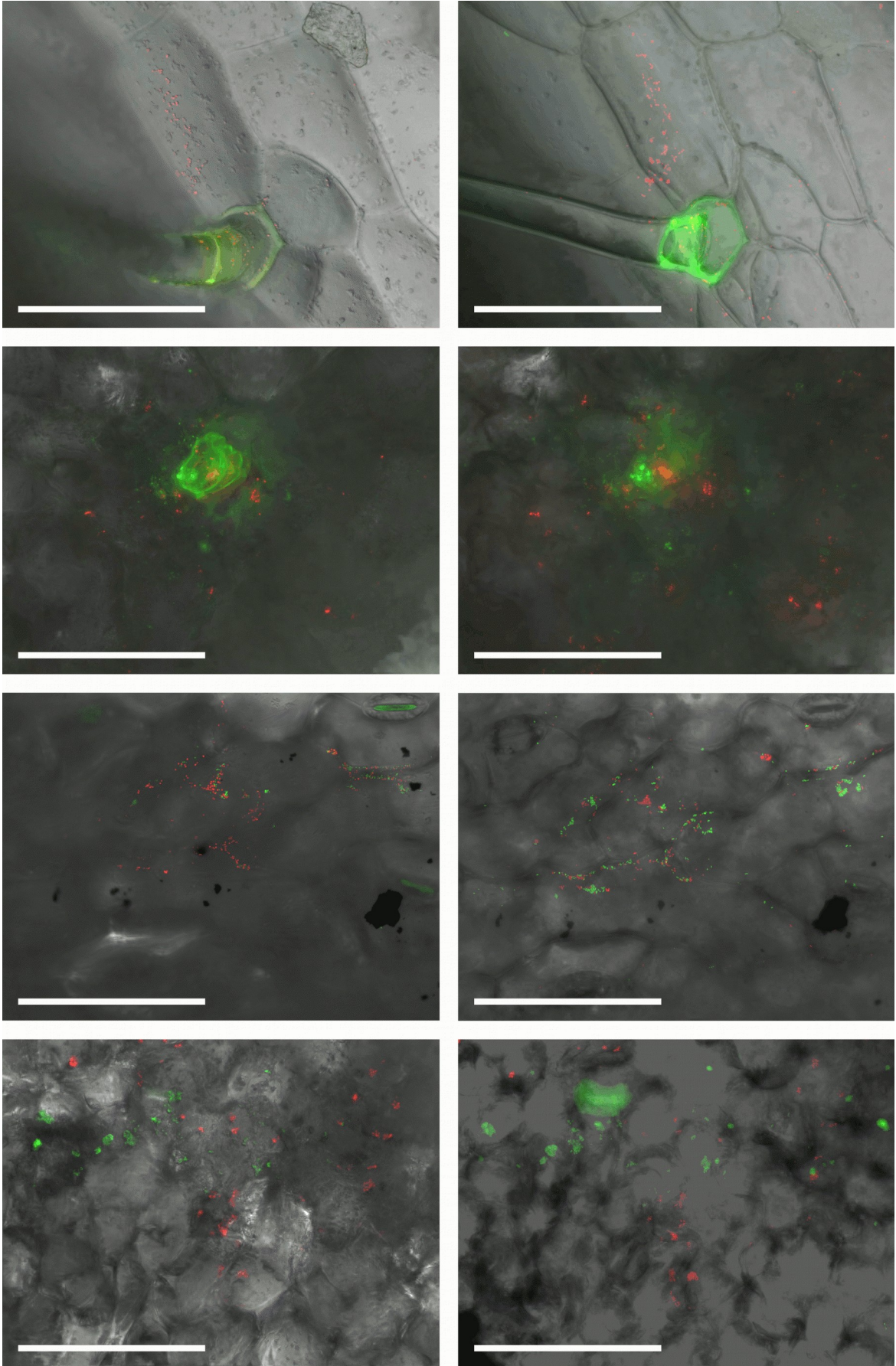


Figure 4 - continuation

Chapter 4 – Spatial scales of interactions among bacteria and between bacteria and the leaf surface

This chapter was published as a research article
Esser D. S., Leveau J. H. J., Meyer K. M., Wiegand K. (2015)
FEMS Microbiology Ecology **91**:fiu034

Abstract

Microbial life on plant leaves is characterized by a multitude of interactions between leaf colonizers and their environment. While the existence of many of these interactions has been confirmed, their spatial scale or reach often remained unknown. In this study we applied spatial point pattern analysis to 244 distribution patterns of *Pantoea agglomerans* and *Pseudomonas syringae* on bean leaves. The results showed that bacterial colonizers of leaves interact with their environment at different spatial scales. Interactions among bacteria were often confined to small spatial scales up to 5-20 μm , compared to interactions between bacteria and leaf surface structures such as trichomes which could be observed in excess of 100 μm . Spatial point pattern analyses prove a comprehensive tool to determine the different spatial scales of bacterial interactions on plant leaves and will help microbiologists to better understand the interplay between these interactions.

Introduction

The plant leaf surface as a microbial habitat, also known as the phyllosphere (Last 1955), is an important arena for plant-microbe interactions. Many of these interactions are in principle well understood biologically (Meyer and Leveau 2011), physically (Hirano and Upper 2000), chemically (Beattie 2011) or genetically (Bailey, Lilley and Diaper 1996; Espinosa-Urgel 2004; Pontiroli *et al.* 2009). The phyllosphere is a complex and heterogeneous environment where microbial colonizers experience temporally and spatially variable competition for resources, facilitative interactions with other microbes, exposure to environmental stresses such as UV radiation, rapidly changing temperatures, and desiccation (Leveau 2006). Furthermore, the leaf surface exhibits a pronounced topography and a variety of structural elements, such as stomata, trichomes, or veins which influence microbial fitness in the phyllosphere in various ways (Timmer, Marois and Achor 1987; Leveau and Lindow 2001; Monier and Lindow 2004; Yadav, Karamanoli and Vokou 2005).

One of the most difficult questions to answer is how all these factors interact with each other and how they rule microbial life in the phyllosphere. Some factors may be more locally confined than others. For example, the competition of microbes for a carbon source may be spatially confined within a few micrometers around the individuals, whereas environmental factors such as temperature will only change along larger distances, e.g. several millimeters or more. This means that different interactions of bacteria with their environment operate at different spatial scales and a good understanding of these scales is a prerequisite for a thorough interpretation of microbial colonization patterns in the phyllosphere. Numerous studies have looked at the spatial distribution of bacteria in the phyllosphere but often stopped at a general description of observed patterns (Blakeman 1985; Mansvelt and Hattingh 1989; Morris, Monier and Jacques 1998; Fett and Cooke 2003; Hong *et al.* 2010; Yu *et al.* 2014). These studies all confirmed the non-random association of microbial colonizers with leaf structures such as stomata or leaf veins and the aggregated character of bacterial colonization patterns on leaves in general. Other studies applied various statistical methods to correlate bacterial success in the phyllosphere to leaf morphological features or to interactions between microbes (Monier and Lindow 2004; Yadav, Karamanoli and Vokou 2005; Hunter *et al.* 2010). None of these studies, however, was spatially explicit in the sense that it quantified the spatial scale (or reach) of

the underlying processes. For instance, the aggregation of *Pseudomonas syringae* near bean leaf trichomes of the glandular type was noted (Monier and Lindow 2004) but the radius around the trichomes within which this process was significant, i.e. the spatial scale of the process, remained undefined. Knowing this spatial scale would be very useful towards conceiving (or rejecting) mechanistic explanations for microbial colonization patterns.

Spatial point pattern analysis (Illian *et al.* 2008; Wiegand and Moloney 2014) provides tools to identify and evaluate interactions between the points in a point pattern, e.g. between individuals in a population. The results can help to understand the processes that formed the pattern and to assess the spatial scale at which these processes operate. In the most basic case, point-pattern analysis is used to test if points in a pattern are randomly distributed. More particularly, the case of *complete spatial randomness* (CSR) is used as a null model against which the observed patterns are compared. Under complete spatial randomness, the location of each point is random and independent from the location of other points. The alternate hypothesis (point distribution does not follow CSR) can be differentiated as points being either aggregated or distributed in a regular pattern which causes a scarcity of certain inter-point distances. The aggregation of points can indicate facilitative interactions between individuals, whereas regular patterns often arise from inhibitory interactions. Both patterns, however, also may have been formed by an unobserved external factor such as heterogeneous nutrient availability – a fact that requires careful consideration when discussing results.

The pair correlation function $g(r)$ (Figure 5) is currently the preferred method to study the distribution of points in a pattern, if fully mapped location data for all points, e.g. all individuals of a species in an area, is available (Illian *et al.* 2008; Wiegand and Moloney 2014). It uses all inter-point distances in a pattern to determine the probability to find points at a certain distance r around a typical point of the pattern. If the points, e.g. bacteria on a leaf surface, are randomly distributed in space, the pair correlation function takes the value 1 at all spatial scales r . Values $g(r) > 1$ indicate an aggregation of points at scale r , whereas values $g(r) < 1$ indicate a scarcity of points (cf. Figure 5). For example, a pattern of randomly placed (circular) clusters of points, where the typical diameter of a cluster is $5\ \mu\text{m}$, will (ideally) result in a pair correlation function with $g(r) > 1$ for radii r smaller than or equal to $5\ \mu\text{m}$ and $g(r) = 1$ for $r > 5\ \mu\text{m}$.

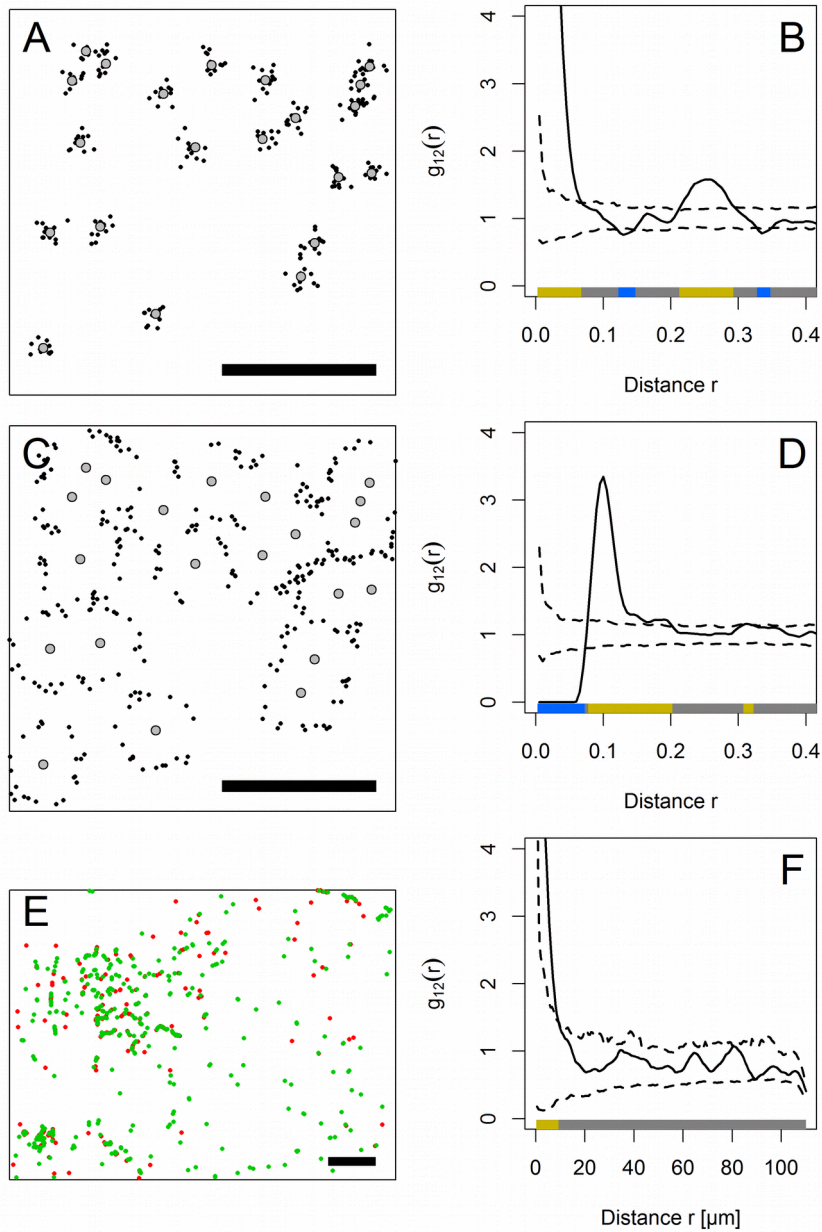


Figure 5: Cross-type pair correlation functions (B, D and F, solid lines) of artificial and observed point patterns (A, C and E). A, C: artificial point patterns in a unit square (scale bars are 0.4 units), E: pattern of DsRed-labeled *Pseudomonas syringae* (red) and GFP-labeled *Pantoaea agglomerans* (green) on a bean leaf (scale bar is 110 μm). If the function exceeds the simulation envelopes (B, D and F, broken lines) at scale r , the points are significantly aggregated at scale r . Values for $g_{12}(r)$ smaller than the lower bound of the simulation envelopes indicate a significant scarcity of points of pattern 2 at distance r of an arbitrary point of pattern 1. The colored bands ('quantum plots') at the bottom of the graphs depict the spatial scales at which the pair correlation functions deviate from the simulation envelopes. In pattern A, the black points were aggregated around the gray points up to a scale of 0.07 units. A second maximum around $r = 0.25$ units suggest a periodicity in the pattern that could arise from a typical distance of 0.25 units between clusters. In pattern C the black points aggregated 0.1 units from the gray points but avoided coming closer than this. The *P. agglomerans* cells in E aggregated around *P. syringae* cells but no significant patterns were found at scales larger than 10 μm .

Two important extensions of the pair correlation function exist. Firstly, the cross-type pair correlation function $g_{12}(r)$ (Lotwick and Silverman 1982) is used to study the interactions between points from two different point patterns, e.g. between individuals of two different species. Secondly, the inhomogeneous pair correlation function $g_{\text{inhom}}(r)$ (Baddeley, Møller and Waagepetersen 2000) was developed to study the interactions between point patterns where the distribution of points is inhomogeneous, e.g. when the point density increases along an unobserved environmental gradient. Both extensions to the pair correlation function can be combined to analyze the interactions between two point patterns of which at least one pattern is inhomogeneous.

Pair correlation functions can exhibit complex behavior, especially when the distribution of points was formed by more than one process. The pattern of black points in Figure 5C for example was produced by two competing processes, one of which attracts the black points towards the grey points whereas the second processes forbids the black points to come closer than 0.1 units to the grey points. One may think of moths being attracted by the light of campfires and repulsed by the intense heat at the same time. Both processes are represented in the corresponding cross-type pair correlation function $g_{12}(r)$ (Figure 5D).

Pair correlation functions are usually evaluated by their deviations from Monte Carlo simulation envelopes (Kenkel 1988). These envelopes are based on a series of simulated point patterns that were generated according to a suitable null model, e.g. complete spatial randomness. But also more complex null models such as a parametrized cluster process are possible (Wiegand and Moloney 2014). By calculating the pair correlation function for a number m of these simulated patterns, and maybe omitting the s most extreme values found at each scale r , one can derive simulation envelopes which delineate the range of values that $g(r)$ takes if the observed points were distributed by the process in the null model. Observed values of $g(r)$ greater than the upper bound of the envelope at scale r indicate a significant aggregation of points at scale r , whereas values of $g(r)$ below the lower bound of the envelope at scale r indicate a scarcity of points at scale r (cf. Figure 5). The level of significance α attached to such a simulation envelope is $2s/(m+1)$. However, note that this level of significance is a guideline only, due to type I error inflation in these simulation envelopes (Loosmore and Ford 2006; Baddeley *et al.* 2014). Nevertheless, for the purposes of exploratory analysis, Monte Carlo simulation envelopes can yield important insights into the point pattern under study (Baddeley *et al.* 2014).

In microbiology, the potential of spatial point pattern analysis has not yet been fully explored. The majority of studies represent landscape-scale epidemiology (Jonsson *et al.* 2010; Rao, Kitron and Weigel 2010; Lin *et al.* 2011). However, at the landscape scale, geostatistical methods are more common (Dandurand, Schotzko and Knudsen 1997; Franklin *et al.* 2002; Franklin and Mills 2003; Brown *et al.* 2004; Gosme and Lucas 2009). The difference between point pattern analysis and geostatistics is that the former studies the spatial associations between objects (points) in an area that is completely mapped while the latter studies continuous processes at selected sampling locations. In spatial point pattern analysis, the location of a finite number of spatially discrete entities such as individuals or colonies is studied. This requires a complete survey of all individuals (colonies, entities) within the observation window and yields information about the spatial relationship between the entities. In geostatistical methodology, a number of independent samples at different locations are considered and values for a spatially continuous variable, e.g. soil moisture or relative abundances in a microbial community, are measured. From that, conclusions about the spatial properties of these variables are drawn. Only few studies have applied point pattern analysis on an individual cell level, for example to quantify micro-scale clustering of bacteria in soils (Nunan *et al.* 2002; Raynaud and Nunan 2014) or the micro-scale inhibition of bacteria and algae in stream biofilms (Augsburger *et al.* 2010). Our study is the first to perform spatial point pattern analysis of bacterial colonization patterns on plant leaves including the interactions between bacteria and leaf structures. The phyllosphere represents an excellent microbial habitat to test the usefulness of point-pattern analyses in surface-based colonization. In this study, we analyzed the colonization patterns of the bacteria *Pantoea agglomerans* and *Pseudomonas syringae* on leaves of common bean. The aim of our study was to reveal how and on which spatial scales these patterns were influenced by the presence of other bacteria and by plant features such as stomata, trichomes, vein cells, or the grooves between epidermal cells and to give an outlook on probable processes that formed the observed patterns.

Methods

Experimental Setup

We inoculated cut sections from primary leaves of two-week old green bean plants (*Phaseolus vulgaris* cv. Blue Lake Bush) with either single or mixed suspensions of the bacterial species *Pantoea agglomerans* 299R and *Pseudomonas syringae* B728a. Our study

therefore included two lines of experiments, a) single-species experiments where leaves were inoculated with one strain (*P. agglomerans*) and b) mixed-species experiments where both strains were co-inoculated in equal quantities. The single-species experiments were used to study spatial patterns of *P. agglomerans*. Mixed inoculation with both strains were used twofold: to study the changes in the colonization patterns of *P. agglomerans* when competing with a second species and to study the interactions between both strains. Single-species and mixed-species setups were further processed the same way. We chose members of *Pseudomonas* and *Pantoea* as they are among the best studied genera in the phyllosphere. The *Pseudomonas* genus is of special economic interest because of its wide variety of plant pathogenic strains (Hirano and Upper 2000; Espinosa-Urgel 2004; Monier and Lindow 2004; Masák *et al.* 2014). Much is known already about the biology of *Pantoea agglomerans*, its interactions with *Pseudomonas* species and its importance as an biocontrol agent (Kempf 1989; Marchi *et al.* 2006; Yu *et al.* 2014).

In our study, single species experiments featured *P. agglomerans* strains 299R (pFRU48) and 299R (pFRU97) (Tecon and Leveau 2012). Plasmids pFRU48 and pFRU97 drive the constitutive expression of fluorescent proteins GFP (green) and DsRed (red), respectively. In two-species experiments we used GFP-producing 299R::JBA28 (Leveau and Lindow 2001) with *P. syringae* B728a (pFRU97) (Monier and Lindow 2004). We cultivated the bacteria separately in 5 ml of Lysogeny Broth (LB) liquid medium with 50 mg kanamycin per liter at 30 °C. Cells were harvested during mid-exponential phase by centrifugation for 10 minutes at 2500 g, washed twice in M9 minimal medium (Sambrook, Fritsch and Maniatis 2001) without carbon source, and diluted in M9 (no carbon) to an approximate concentration of 10^7 individuals/ml.

Sections of 15 mm × 15 mm were cut from bean leaves (mid-leaf, about 5 mm off the central vein) and edge-sealed by dipping briefly into 90-100 °C paraffin wax. Two leaf sections each were placed on agarose gel in a petri dish with the adaxial side facing up. Inoculation with 50 µl of a 10^7 cells per ml suspension was performed using an airbrush Iwata Eclipse HP-CS (ANEST IWATA Corporation, Yokohama, Japan) at 100 kPa pressure. We either inoculated a mixture of red and green *P. agglomerans* or a mixture of red *P. syringae* and green *P. agglomerans* cells. We chose to use mixtures of red and green *P. agglomerans* cells to have more information on *a posteriori* mixing processes also in the single-species experiments (Tecon and Leveau 2012). Inoculation was performed through a hole in the lid of a plastic container 170 mm above the leaf surface. The nozzle of the

airbrush was slightly swirled during inoculation to assure a good spread across the petri dish. The sections were either observed under the microscope directly (time $t=0$) or the petri dishes were sealed to maintain a 100% relative humidity environment and put in an incubator at 28 °C. Incubation times ranged between 10 to 72 hours to cover many stages of early leaf colonization.

Additionally, we inoculated a series of leaves with DsRed-labelled *P. syringae* B728a only, incubated them for 92 hours at room temperature and high humidity. We then cut sections from these leaves, sealed the edges with wax and inoculated these sections with green *P. agglomerans* 299R and incubated for another 0-72 hours at 28 °C.

A full record of incubation times for all samples is given with further information in the Table S4.1 in the Supplementary Materials section (p. 100).

Sample preparation and image acquisition

After incubation, leaf sections were transferred to a microscope slide with the adaxial side facing up and covered with 10-50 μl Aqua Poly/Mount (Polyscience Inc., Warrington PA, USA) medium to ensure a good coverage of the sample. We then carefully added a cover slip which we fixed with strips of adhesive tape to all sides. From each leaf section, we typically took 10 micrographs at random positions using an Axio Imager.M2 fluorescence microscope (Carl Zeiss AG, Oberkochen, Germany) equipped with EC Plan-Neofluar10x/0.3, 20x/0.5 and 40x/0.6 (Zeiss) objectives and a AxioCAM MRn monochrome camera. Image sizes were $895.3 \mu\text{m} \times 670.8 \mu\text{m}$, $447.6 \mu\text{m} \times 335.4 \mu\text{m}$ and $223.8 \mu\text{m} \times 167.7 \mu\text{m}$ for the 10 \times , 20 \times and 40 \times objective, respectively, but in few instances smaller, when out-of-focus areas had to be cropped. For the fluorescence images we used a GFP filter cube (exciter: 470; emitter: 525/50, beam splitter: 495) and a rhodamine filter cube (exciter: 546/12; emitter: 607/80; beam splitter 560). We also took phase-contrast images of all samples to visualize the leaf surface structure. To account for the topography of the leaf surface we took all images as 3-dimensional 'z-stacks', i.e. several shots of the same area at different planes of focus. These were saved in the native Zeiss .zvi format.

Image Processing

The .zvi images were processed using the open-source ImageJ software package (Rasband, W.S., ImageJ, U. S. National Institutes of Health, Bethesda, Maryland, USA,

<http://imagej.nih.gov/ij/>, 1997-2012). We extracted the location (x-/y-coordinates) of all bacterial individuals and all structural elements of the leaf surface, i.e. stomata, trichomes, veins, grooves. The location of stomata was represented by a point within each guard cell, for the location of trichomes we used the center of each base cell, and veins were represented by a point at the center of each vein cell. For the grooves between epidermal cells we marked the intersections where at least three grooves come together. If grooves have an effect on the distribution of bacterial leaf colonizers, these intersections will be of special importance as they represent locations of high groove density.

Spatial Statistics

Our study consists of three groups of analyses: a) the analysis of interactions between the individuals of the same bacterial species = intraspecific interactions, b) the interactions between the two bacterial species = interspecific interactions, and c) the interactions between bacteria (not considering species identity) and the different structural elements of the leaf surface such as stomata or vein cells. The quality and scale of the patterns observed in these analyses can be used to develop hypotheses about the underlying processes/interactions that contributed to the observed distributions. This procedure has to be performed cautiously depending on the null-model against which the observed pair correlation function is tested (Baddeley *et al.* 2014).

To study the spatial distribution of *P. agglomerans*, we calculated the inhomogeneous pair correlation function from the location data of the bacterial cells (intraspecific analysis, *P. agglomerans* only, not considering different colors). We also calculated intraspecific inhomogeneous pair correlation functions of *P. agglomerans* growing with *P. syringae* and vice versa and also for the pattern of ‘pooled’ bacteria, i.e. patterns of all bacterial individuals regardless of their color or species.

To explore the interactions between *P. agglomerans* and *P. syringae*, we calculated the inhomogeneous version of the cross-type pair correlation function $g_{12}(r)$. This was performed separately for the data sets in which both strains were inoculated together at the same time and for the data sets in which *P. agglomerans* was inoculated after *P. syringae* had already grown on the leaf for 92 hours.

For all inhomogeneous pair correlation functions, intraspecific and interspecific, we used a Gaussian smoothing kernel with standard deviation $\sigma = 110 \mu\text{m}$ to estimate the local densities $\lambda(x,y)$. We found that $\sigma = 110 \mu\text{m}$ gave the most stable results for the pair

correlation functions across our whole data set. Moreover, this scale is above the maximum expected interaction distance: based on work by Jeff Chant, Frankin & Mills estimated the interaction distances between bacterial individuals in a solution to be around 10 times their cell size (Franklin and Mills 2007) which translates to about 10 μm in our study. Interactions between bacteria in the rhizoplane were found to steeply decay at scales greater than 5 μm and have not been observed beyond 78 μm (Gantner *et al.* 2006). For the phyllosphere, for which such measurements did not exist until now, we would not expect longer ‘calling distances’.

To study the interactions between the bacterial colonizers and the leaf surface structures (stomata, trichomes, veins, and grooves), we used the cross-type pair correlation function. For stomata, trichomes, and groove nodes we assumed a homogeneous distribution. Vein cells, however, are distributed heterogeneously along linear structures (the veins), and we additionally calculated the inhomogeneous cross-type pair correlation function, again with $\sigma = 110 \mu\text{m}$.

All pair correlation functions $g(r)$ for each sample were tested for significant clumping or scarcity of points using Monte-Carlo simulation envelopes. Throughout our study, we estimated 95%-simulation envelopes from 199 simulations of the null-model. At each spatial scale r we selected the fifth highest and fifth lowest values of $g(r)$ for the upper and lower bound of the envelope, respectively.

The *null models* used in the analyses reflect the biological hypotheses to be tested. For the intraspecific analyses (studying only one bacterial species), we applied the CSR null model where the location of a point is independent from the location of other points. Therefore, we generated random point patterns of the same point density as the observed pattern. For the analyses of cross-type pair correlation functions (interaction between bacterial species and interactions between bacteria and leaf surface structures), we applied the independence null model, i.e. we performed 199 toroidal shifts (Wiegand and Moloney 2014). Here, a bacterial pattern is shifted a random distance into a random direction and points that exit the (rectangular) observation window reappear at the opposite edge of the window. The location of other observed point pattern (either the other bacterial strain or a leaf structure) stayed unchanged. This approach preserves the internal structure of both point patterns (here, the structure of the bacterial pattern and the structure of e.g. the stomatal pattern). This way, the independence null model exclusively tests for independence between the patterns and is unaffected by patterns (e.g. clustering) that may be present within the

individual point patterns. Thus, the toroidal shift null model is especially well suited to study the interactions between two different point patterns as it preserves all other interactions within the first pattern and only considers differences that arise from the dependence (interaction) of the first pattern from the second pattern. Also, it only relies on random shifts of one of the patterns and therefore is not subject to the limitations with respect to interaction distances discussed in (Baddeley *et al.* 2014). All calculations were performed using the *spatstat* package (Baddeley and Turner 2005) in the statistics software *R* (R Core Team 2013).

Data presentation

The results from our analyses using pair correlation functions were summarized in frequency plots. For each of these frequency plots we looked at the pair correlation functions of all samples and plotted the relative frequencies of significant aggregation and scarcity. These were represented as stacked bars for every distance class r at which the pair correlation function was evaluated. Since each pair correlation function was evaluated up to a maximum distance r depending on the size of the observation window of the respective sample, sample size typically decreased with increasing spatial scale r . Not all data sets had stomata, trichomes, or veins in them. This resulted in a reduced number of samples used in the respective analyses.

Results and Discussion

In total, we analyzed point patterns from 244 samples (i.e. images of fields of view under the microscope) from 42 different adaxial leaf sections from 23 independent spray experiments. A rather typical pattern is depicted in Figure 5E. The total area of all observation windows was approximately 75.63 mm² and contained 131,429 bacterial individuals, 2,192 stomata, 135 glandular trichomes, 74 hooked trichomes, and 819 vein cells. We found 43,674 nodes at which three or more grooves between epidermal cells came together. Bacterial population sizes on leaves

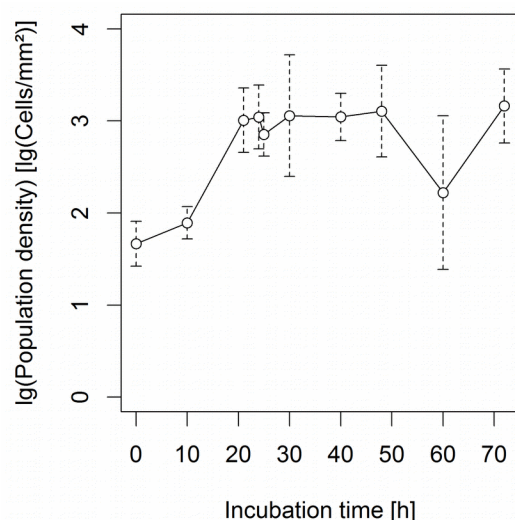


Figure 6: Mean bacterial population density (open circles) and standard deviation (broken lines) over time from all samples excluding samples where *P. agglomerans* was inoculated 92 hours after *P. syringae*.

increased within the first 20 hours after inoculation and remained constant thereafter (Figure 6).

Interactions between bacterial individuals of the same species

When *P. agglomerans* was the only colonizer of the bean leaves, bacterial cells clustered at scales up to 7 μm in 90% of our samples (Figure 7A). This means that within a 7 μm radius around a typical individual we found significantly more bacterial cells than expected if the cells were randomly distributed. The percentage of significant clustering gradually decayed at scales larger than 7 μm and disappeared into a background noise around 85 μm away from a typical *P. agglomerans* cell. This small-scale aggregation of bacteria is in agreement with most published studies on the general spatial distribution of bacteria on leaves (Kinkel 1997; Beattie and Lindow 1999; Leveau and Lindow 2001; Monier and Lindow 2004) but none of these studies were able to report statistically verified information on the scale of the aggregation. Our results are also consistent with the concept of heterogeneously distributed but highly localized availability of resources (Leveau and Lindow 2001; Kinkel, Newton and Leonard 2002) and with the concept of clonal ‘staying-together’ growth behavior (Tecon and Leveau 2012).

When *P. agglomerans* was inoculated in combination with *P. syringae*, small-scale aggregation of *P. agglomerans* cells was observed for more than 98% of our samples, while the noise distance was reduced to about 75 μm (Figure 7B). Comparison of Figure 7A and 7B suggest that the interaction with *P. syringae* increases the level of clustering of *P. agglomerans*. Garbeva et al. (2011) showed that the shape of colonies of *Pseudomonas fluorescens* changes from irregular shapes in isolation to spherical when exposed to *Pedobacter*-born signaling molecules, thereby suggesting an interspecific interaction. Because of the reduced surface-to-volume ratio, cells in a spherical colony will be more spatially aggregated than the same cells in an irregular shaped colony of comparable packing density. Similar mechanisms may explain the increase in small-scale aggregation of our *P. agglomerans* cells when exposed to compete with *P. syringae* cells. Alternatively to the active response mechanism suggested by Garbeva et al. (2011), the increase in small-scale aggregation prevalence could be the result of steric constraints that colonies of the one species imposes on the development of colonies of the other species.

Aggregation of *P. syringae* competing with *P. agglomerans* was found at scales up to at least 5 μm in the majority of the sample images (Figure 7C). At scales larger than 5 μm ,

prevalence of significant aggregation rapidly decayed such that aggregation up to 28 μm was found in every third sample and disappeared in a background noise around 90 μm . Compared to *P. agglomerans* cells in Figure 7B, the *P. syringae* cells were clustered at smaller scales, i.e. more compact than *P. agglomerans*. If we considered local resource availability as the most important determinant of bacterial colonization patterns, the tendency of *P. syringae* being more clustered than *P. agglomerans* could be the result of *P. syringae* using other, more spatially confined resources compared to *P. agglomerans*. Alternatively, if both strains were to use similar resources, *P. syringae* would require higher concentrations of some key resources. Such ideas could be pursued in a future study that combines spatial point pattern analysis with data from bacterial bioreporters for nutrient availability (Leveau and Lindow 2001) or micro-meter-scale metabolic profiling (Fang and Dorrestein 2014) of the phyllosphere, both of which would provide high-resolution spatial information on nutrient availability along the leaf surface.

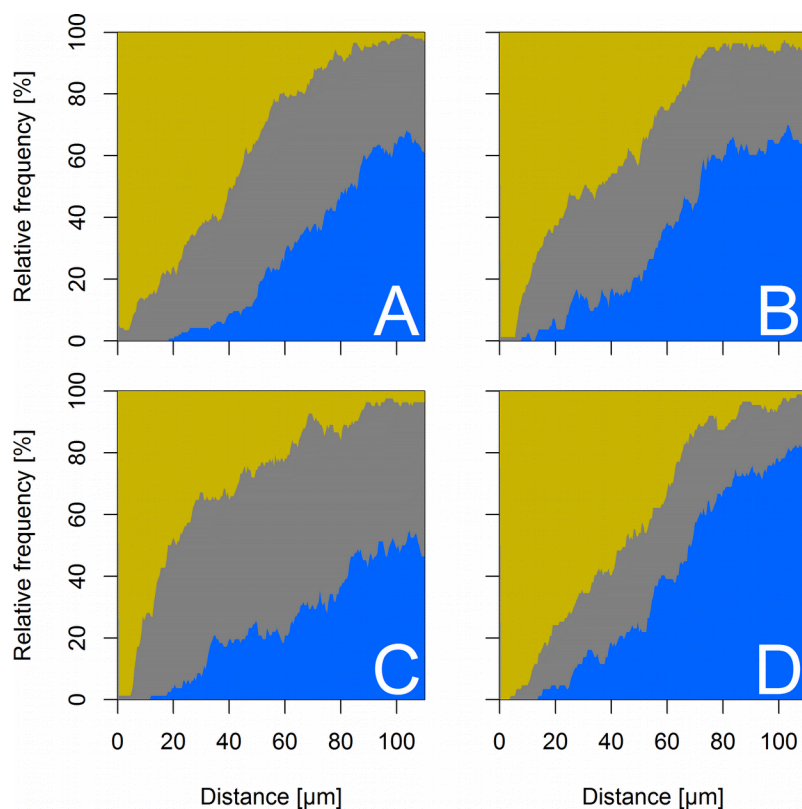


Figure 7: Intraspecific interactions of bacteria in the bean phyllosphere. The graphs give the relative frequency of significant aggregation (yellow/light gray), randomness (medium gray) and scarcity (blue/dark gray) at different scales up to 110 μm as determined by inhomogeneous pair correlation functions. A: Interactions between individuals of *Pantoea agglomerans* when growing as the sole species on the leaf. B: Interactions between individuals of *P. agglomerans* when growing in competition with *Pseudomonas syringae*. C: Interactions between individuals of *P. syringae* when growing in competition with *P. agglomerans*. D: Interactions between bacterial individuals not considering species identity (both species growing together). Numbers of samples were 145 in A, 82 in B, 83 in C and 87 in D.

The spatial distribution of bacteria, not considering their species exhibited small scale aggregation of cells in all samples up to 4 μm followed by an almost linear decay in prevalence that disappeared in a background noise at almost 90 μm (Figure 7D).

The general pattern of scarcity of cells towards larger scales (above 50 μm , Figure 7A-D) is more difficult to explain. A classical interpretation would be a regularity in the environment, e.g. the undifferentiated epidermal cells in our experiments. This is, however, difficult to verify given their complex shape and the probable multitude of additional interfering interactions of bacteria with their environment. A non-biological explanation is that a bias arose from the bandwidth of the Gaussian smoothing kernel used for the inhomogeneous pair correlation functions. The inhomogeneous method applied here is known to be biased downwards and can give unexpected results especially for regular patterns (Baddeley, Møller and Waagepetersen 2000). Although our choice of 110 μm for the bandwidth gave the most stable results for 50 randomly chosen images from our data set, it is possible that it worked better for the sparser *P. syringae* patterns than for the denser *P. agglomerans* patterns. Additionally, the inhomogeneous pair correlation function requires the points to be “second-order intensity reweighed stationary” (Baddeley, Møller and Waagepetersen 2000); an assumption probably violated by our complex bacterial colonization patterns on bean leaves (Baddeley, personal communication). Nevertheless, we recommend the use of the inhomogeneous pair correlation. Given a) its non-cumulative nature which puts it ahead of other methods such as the Ripley’s K function (Ripley 1976; Wiegand and Moloney 2014), b) its ability to look at distance beyond the nearest neighbor of a point, and c) its ability to (within limits) account for a heterogeneity in the point pattern, it is the best available method for the data at hand in our study.

Interactions between bacterial individuals of different species

The interspecific analyses of mix-inoculated *P. agglomerans* and *P. syringae* cells revealed a clustering of *P. agglomerans* cells within 10 μm around individuals of *P. syringae* and vice versa (Figure 8A and B). Clustering was found in more than 60% of our samples. A typical pattern that exhibited this clustering is shown in Figure 5E. In the example pattern of *P. syringae* near *P. agglomerans* pictured in Figure 9A, the inhomogeneous pair correlation was very similar to the overall results in Figure 8B. In this particular sample, *P. syringae* was less abundant than *P. agglomerans* (176 *P. syringae* cells against 1380 of *P. agglomerans*) and the *P. syringae* clusters were small (10 μm in diameter and less,

Figure 9A). Similar patterns were found in many other samples. From the pair correlation function (Figure 9B), we can see that *P. syringae* typically aggregated up to 11 μm near *P. agglomerans* cells, for example in the large *P. agglomerans* colony (left inset in Figure 9A). The aggregation from 18 μm to 23 μm is only very weak and probably mostly due to two colonies at the left edge of the large *P. agglomerans* colony and maybe a colony surrounded by a half moon-shaped *P. agglomerans* colony (right inset in Figure 9A). In summary, the pattern observed in this particular sample captures important aspects of the interactions between the two species that seem to be typical for bean leaves in our experiments.

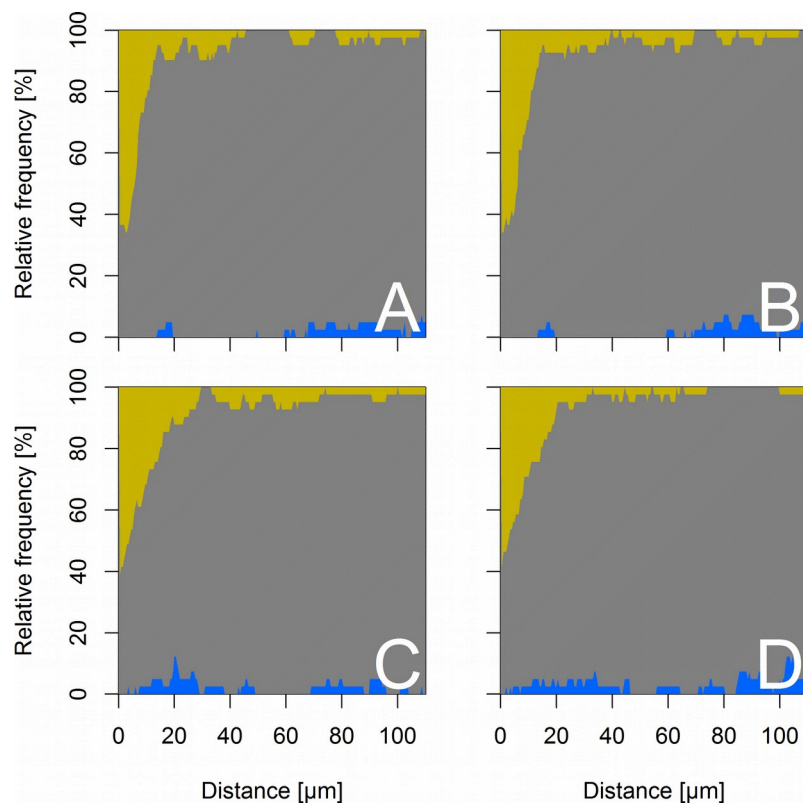


Figure 8: Interspecific interactions of bacteria in the bean phyllosphere. The graphs give the relative frequency of significant aggregation (yellow/light gray), randomness (medium gray) and scarcity (blue/dark gray) at different scales as determined by inhomogeneous cross-type pair correlation functions using random toroidal shifts as a null model. A, C: Aggregation of individuals of *Pantoea agglomerans* around individuals of *Pseudomonas syringae*. B, D: Aggregation of *P. syringae* around *P. agglomerans*. In A and B, both strains were inoculated jointly and incubated for 0-72 hours. In C and D, *P. syringae* was inoculated first and incubated for 92 hours before *P. agglomerans* was inoculated. After inoculation with *P. agglomerans*, these samples were incubated for another 0-72 hours. Number of samples were 41 in all four analyses.

In experiments where *P. syringae* had the opportunity to develop on the bean leaves for 92 hours under 100% humidity at room temperature prior to the arrival of *P. agglomerans*, *P. syringae* was much more abundant than *P. agglomerans*. The late arriving GFP-labeled

P. agglomerans cells only developed poorly (84.2% *P. syringae* versus 15.8% *P. agglomerans*, averaged over all samples). Here, *P. syringae* probably was able to exclude *P. agglomerans* from successful establishment by depleting patches rich in resources prior to the arrival of *P. agglomerans*. Such processes have recently been shown in a study on the reproductive success of *P. agglomerans* on pre-colonized bean leaves (Remus-Emsermann, Kowalchuk and Leveau 2013). In this setting, aggregation of *P. agglomerans* around *P. syringae* was observed in 60% of our samples and up to 20-30 μm from a typical *P. syringae* individual and vice versa (Figure 8C and D).

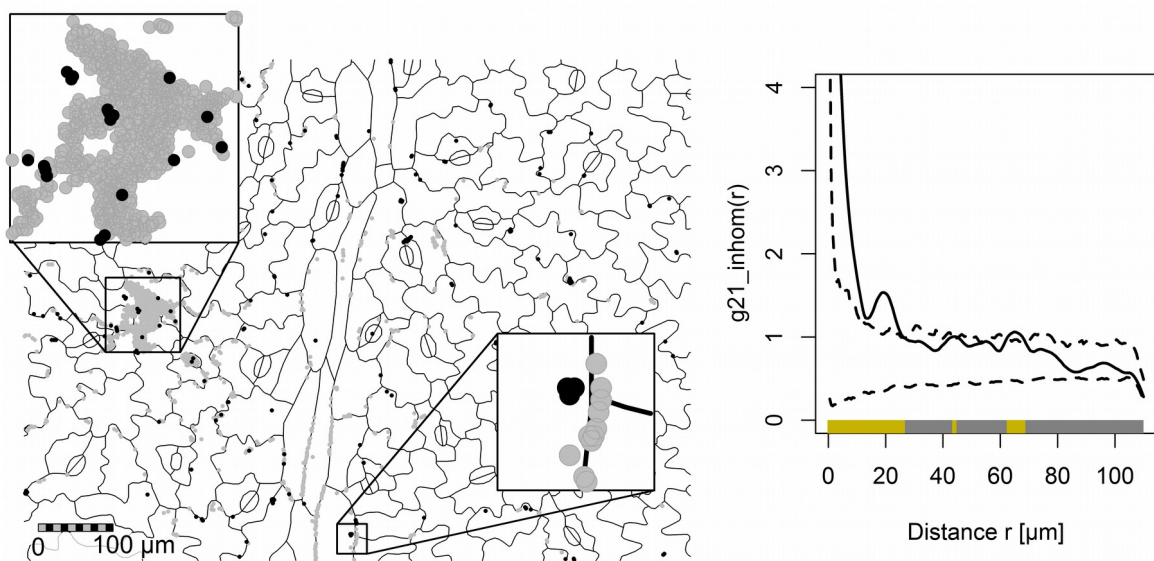


Figure 9: A: Spatial pattern of *P. agglomerans* (gray points) and *P. syringae* (black points) on a bean leaf surface. B: the inhomogeneous pair correlation function (PCF, solid line) and 95% simulation envelope (broken lines) for *P. syringae* cells near *P. agglomerans* cells derived from the pattern (interpretation of the PCF same as in Figure 1). The spatial pattern is typical in the sense that the PCF result well resembles the general results for this interaction shown in Figure 8B.

Summarizing the results from our interspecific studies, the two strains settle in the same regions of a leaf and do not seem to avoid each other. The maximum range of interactions leading to a co-aggregation of the two species lies around 20-30 μm . In the majority of our samples, however, we found no significant interspecific aggregation at scales larger than 10 μm , a value that evidently fits well the value of 10 times the bacterial cell diameter suggested by Frankling and Mills (2007).

Bi-variate analyses of bacteria near bean leaf surface features

Due to the heterogeneous distribution of suitable sites on the leaf surface, interactions between bacteria are usually not the only determinant of bacterial distribution in the phyllosphere.

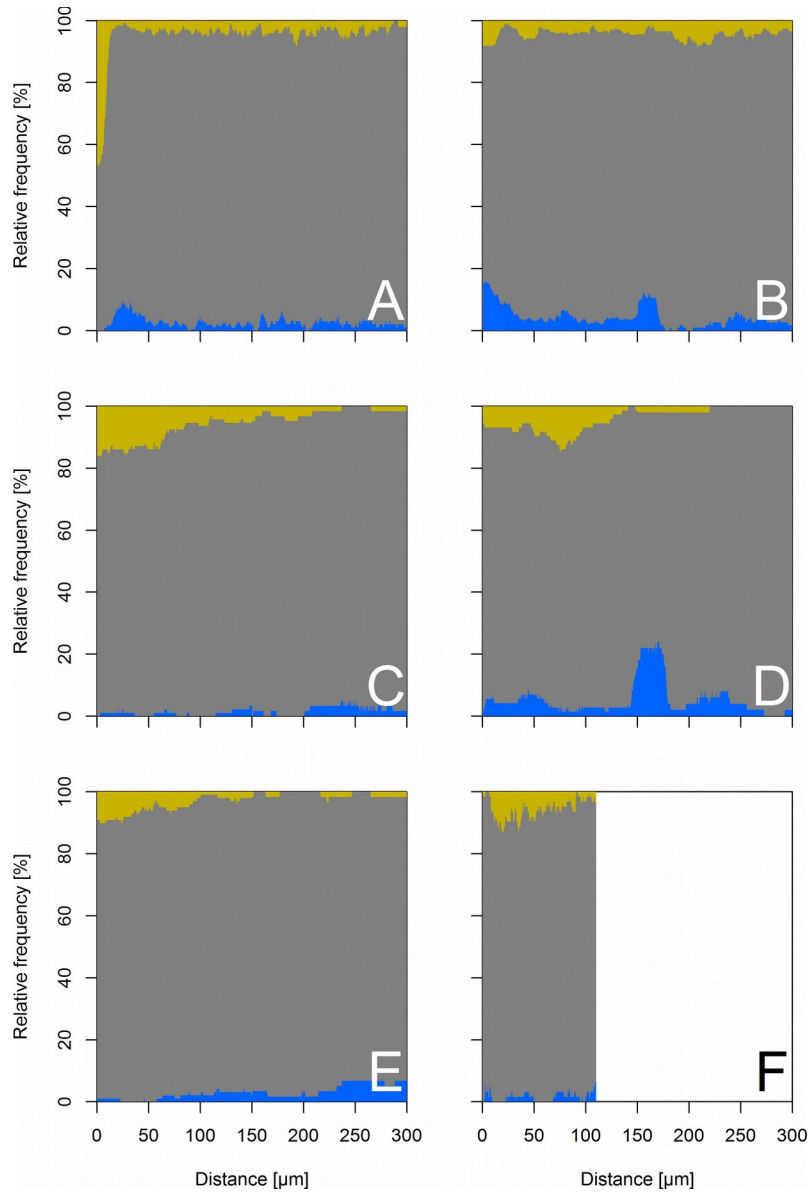


Figure 10: Interactions of bacteria with structural elements of the bean phyllosphere. The graphs give the relative frequency of significant aggregation (yellow/light gray), randomness (medium gray) and scarcity (blue/dark gray) at different scales as determined by cross-type pair correlation functions (g_{12}) using random toroidal shifts (cf. text) as a null model. A-E: homogeneous g_{12} , F: inhomogeneous g_{12} . In F, the local point density was estimated using a moving Gaussian smoothing kernel of bandwidth 110 μm , which limited our analysis up to this scale. A: Interaction of bacteria with grooves between epidermal cells. B: Interaction with stomata. C: Interaction with glandular trichomes. D: Interaction with hooked trichomes. E-F Interactions with vein cells. Number of samples were 244 in A and B, 93 in C, 72 in D, 99 in E and 62 in F.

In our study, the grooves, or more precisely the intersection points of grooves between epidermal cells of the bean leaf were frequently attracting bacteria on a small scale up to 12 μm (Figure 10A). Attraction of bacteria to grooves was apparent in about 45 % of our samples. This aggregation of microbial colonizers of the phyllosphere has been frequently reported (Diem 1974; Blakeman 1985; Davis and Brlansky 1991; Leveau and Lindow

2001; Monier and Lindow 2004). The grooves between epidermal cells have been hypothesized to retain water longest during periods of evaporation (Kinkel 1997; Leveau and Lindow 2001). Considering that the presence of water on leaf surfaces can stimulate leaching of substances from inside the leaf (Tukey 1970) and that nutrients in solution will accumulate in regions that retain evaporating water the longest, it becomes obvious why plant leaf surface features such as the grooves could be a good proxy for explaining microbial success and therefore colonization patterns in the phyllosphere. Nevertheless, it remained unclear how much of the bacterial aggregation on small scales can be accounted to accumulation of nutrients in the grooves and how much is due to other processes such as a physical or gravitational groove effect.

We found no strong signs of bacterial aggregation near stomata which is in line with the study of (Monier and Lindow 2004) who studied *Pseudomonas syringae* on bean leaves. In fact, we rather detected a slight tendency of bacteria to avoid stomata which lasted up to 35 μm (Figure 10B). Then again, stomata have been reported for a long time to be positively correlated with high densities of leaf colonizers (Miles, Daines and Rue 1977; Blakeman 1985; Mew and Vera Cruz 1986; Timmer, Marois and Achor 1987; Mansvelt and Hattingh 1989). In these studies, incubation time was longer (3-30 days) and in the studies of Miles *et al.* (1977), Timmer *et al.* (1987), and Mansvelt and Hattingh (1989) incubation was performed under less humid conditions. We conclude that interactions with stomata may occur on bean leaves but they are less important under the conditions tested in our experiments and that these interactions are more important in less humid environments. Under such conditions, stomatal transpiration has been shown to produce highly localized zones of elevated humidity (Burkhardt *et al.* 1999) which will facilitate bacterial growth under otherwise dry conditions. Further studies of the spatial distribution of bacteria on dry leaves could reveal the spatial scale up to which this localized increase in humidity is effective.

On a larger scale we observed an aggregation of bacteria within a 0-60 μm (less often 100 μm) neighborhood around glandular trichomes in up to 16% of our samples (Figure 10C). This interaction has also been reported before (Blakeman 1985; Leveau and Lindow 2001; Monier and Lindow 2004; Yadav, Karamanoli and Vokou 2005), but here we can for the first time quantify the range of this interaction. In the literature the effect is usually explained by nutritious exudates of the glandular trichomes (Ascensão and Pais 1998; Monier and Lindow 2004) or their ability to retain water droplets (Brewer, Smith

and Vogelmann 1991) which also could increase nutrient leaching from the leaf interior (Schönherr and Baur 1996).

We also found significant aggregation of bacteria within up to 120 μm around hooked trichomes but at any spatial scale aggregation was not observed in more than about 15% of our samples (Figure 10D). Interestingly, aggregation was most prevalent between 60-100 μm around hooked trichomes, suggesting an interplay between an attracting (e.g. increased water availability) and a repulsing process (e.g. gravitational, away from elevated trichomes on veins). Aggregation of bacteria near hooked trichomes has been reported before on bean leaves (Leveau and Lindow 2001; Monier and Lindow 2004). These studies mention a thinner cuticle and the presence of more nutrients compared to undifferentiated epidermal cells as probable properties that facilitate bacterial growth near hooked trichomes.

Bacteria were significantly aggregated around vein cells but usually in not more than 10% of our samples that had vein cells in them (Figure 10E and F). Both the homogeneous and the inhomogeneous pair correlation functions detected aggregation at scales up to 100 μm . The most obvious difference between the two functions was that the inhomogeneous pair correlation function (Figure 10F) found no clear signs of aggregation of bacteria within 8 μm around the centers of vein cells. This small-scale randomness could be explained by a process in which the large-scale attraction of bacteria towards veins is locally overruled by the small-scale aggregation of cells in the grooves between vein cells. However, it remained unclear why this effect is not also visible in Figure 10E. The association of bacterial colonizers of leaves with leaf veins has been reported frequently and has usually been attributed to veins being regions of increased water availability, leakage and nutrient availability (Canny 1990; Leveau and Lindow 2001; Axtell and Beattie 2002; Monier and Lindow 2004). Glandular trichomes, which also caused attraction of bacteria, were often located on veins and may thus be sufficient to explain the attraction of bacteria towards veins. Disentangling such facilitative processes not directly related to the veins should be subject of further studies.

Conclusions

In this study we showed how point-pattern analysis can improve our understanding of micro-ecological processes in the phyllosphere. Most importantly, we determined spatial scales or distances of major importance to processes that shape the spatial distribution of

bacteria on bean leaves. Where the uni-variate analysis (Figure 7) mostly informed about the spatial structure of bacterial colonization patterns, the bi-variate analyses allowed for an estimation of the operational range of interactions between two bacterial strains (Figure 8) or between bacteria and structural elements of the leaf surface (Figure 10). Our study is the first to report such interaction distances for bacteria and plant leaves. These interactions are not static but may develop in time as discussed for example for bacterial interactions with stomata. The strength of the underlying processes ranged from weak effects, such as the tendency of bacteria to avoid stomata, to strong effects such as the aggregation of bacteria in the grooves between epidermal cells. The co-aggregation of individuals of *P. agglomerans* and *P. syringae* suggests that these two strains facilitate each other or at least exploit resources in the phyllosphere in similar ways during early phases of colonization. It is important to consider how all these processes interact with each other, facilitating or canceling each other out, thereby leading to the complex colonization patterns of bacteria observed in the phyllosphere. Based on the results of this study, we developed a conceptual model that gives an impression of how all these processes might come together to create the bacterial colonization patterns observable in the bean phyllosphere (Figure 11). We assert that our findings are primarily valid for *Pantoea agglomerans* and *Pseudomonas syringae* on bean leaves and that the interaction regime of other microbial colonizers on different hosts may differ from our results. This may be especially true for bacterial development under less humid conditions than in our experiments. In conclusion, our study is a starting point of a series of future experiments that will use spatial point pattern analysis to unravel the significance of the different spatial interactions between microbial leaf colonizers and their environment.

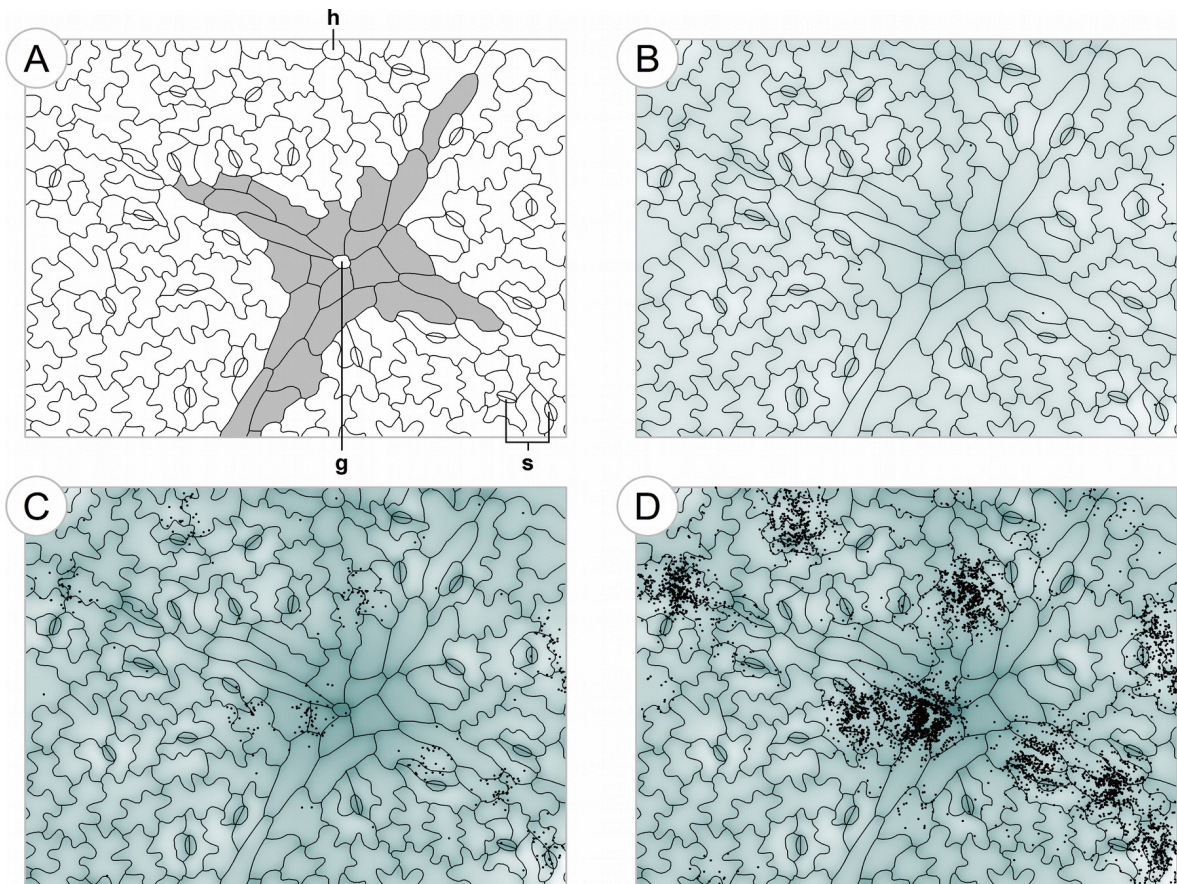


Figure 11: A conceptual model of bacterial development based on the findings of this study. We used a simple simulation model based on the strength and range of bacterial interactions on bean leaves determined in our study in order to determine if these processes were sufficient to generate colonization patterns similar to the observed ones. The simulated colonization pattern in D suggests that the concert of many different interactions at different spatial scales is able to explain the complex bacterial colonization patterns observed on bean leaves. A: The starting 'landscape'. Epidermal structure of a young bean leaf. An x-shaped leaf vein (shaded in grey) locally approaches the leaf surface from the leaf interior with a glandular trichome (g) at the intersection. The vein is surrounded by stomata (s) and undifferentiated epidermal cells ('puzzle pieces'). The base of a hooked trichome (h) can be seen near the middle of the upper boundary. B: Prior to colonization by microbes, resources (shaded areas) gathered on the leaf surface by processes such as epidermal leaching, especially near veins, grooves between epidermal cells, and excretion by glandular trichomes. Single bacterial colonizers arrive on the leaf. C: In the course of time, bacterial colonizers reproduce more successfully at locations rich in resources. Additionally, bacterial cells tend to get trapped near the grooves either by gravitational processes or by the increased density of leaf surface area that decreases cell motility. D: Further bacterial growth. After 20 hours, growth stops in locations where the resource requirements reach the level of leaching of new resources from the leaf interior. Colonization of new resource-rich regions allows further growth of the bacterial population.

Chapter 5 – Biomimetic surfaces help detecting micro-scale effects of topography on bacterial leaf colonization

This chapter was prepared as a manuscript for submission to The ISME Journal

Esser D.S.¹, Meyer K.M.¹, Wiegand K.¹, Tecon R.²,
Gilmore S.F.³, Parikh A.N.³, Leveau J.H.J.²

¹ Department of Ecosystem Modelling, University of Göttingen, Germany,

² Department of Plant Pathology, University of California, Davis, CA, USA

³ Department of Chemical Engineering and Materials Science,
University of California, Davis, CA, USA

Abstract

Microbial life is ruled by a wide variety of interactions that operate at different spatial and temporal scales. Intraspecific interactions co-occur with interspecific interactions with other microbes, with hosts, and with predators, as well as factors and processes of the non-living environment. Often the relative effect of a single process is difficult to isolate from others. For example, it is known that leaf-colonizing bacteria often aggregate in the grooves between epidermal cells of plant leaves but it is unclear if this phenomenon is the result of some process connected to leaf topography or if it is the result of a leaf biological process such as increased leaching of nutrients in the grooves. In our study, we analyzed the spatial distribution of two wide-spread leaf-colonizing bacteria on synthetic replicas of bean leaf surfaces. Comparison of our results to a similar study on real leaves suggests that the frequently reported aggregation of bacteria near grooves between epidermal cells is driven by physical processes connected to leaf topography rather than leaf biological processes. Our study highlights the importance of micro-topographic effects on bacterial surface colonization and illustrates the suitability of soft-lithographically manufactured micro-landscapes for studying these effects.

Introduction

The microbial colonization of surfaces has been intensely studied and has been regularly reviewed from different scientific perspectives such as the general process of biofilm formation (O'Toole, Kaplan and Kolter 2000; Danhorn and Fuqua 2007), particular aspects of adhesion (Katsikogianni and Missirlis 2004), the effect of physical stress on biofilm structure (Otto 2014), the description of chemical compounds and signaling molecules (Petrova and Sauer 2012), surface detection by microbes, changes in gene expression patterns, and cell morphology (Tuson and Weibel 2013; Wozniak and Parsek 2014) as well as the development of anti-fouling surfaces (Salta *et al.* 2010). Surface colonization involves complex interplays of the microbes with the surface (e.g. attachment, modification, detachment), with each other (e.g. competition, antibiosis, collaboration), and with the medium that interfaces the surface (e.g. air or water). Adding to this complexity is the fact that many surfaces feature spatial variability in the physical, chemical and biological conditions at scales that are relevant to microorganisms. One of the major challenges in microbial ecology is to understand this complex interplay as well as the individual and combined impact of these processes on surface colonization. One tried and proven approach to disentangle the complexity of surface colonization is experimental deconstruction, i.e. selectively removing one or more parts of the complexity in order to assess the impact of physical, chemical or biological condition. Examples are the use of single species experiments in order to exclude interspecies interactions or the use of artificial flat surfaces that eliminate potential effects of surface topography.

If, in contrast, such topographical effects, e.g. of biological surfaces, are of interest, these effects could be targeted by comparing the microbial colonization of the biological surface to the colonization patterns on a biomimetically patterned surface (BPS), i.e. a manufactured exact replica of the biological surface. For example, Zhang *et al.* (2014) and colleagues recently showed that soft-lithographic replicas of spinach leaves made from polydimethylsiloxane (PDMS) are well able to replicate spinach leaf surface structure. They also found that wettability of the PDMS replicas was similar to real spinach leaves. Finally, they pointed out that BPS made from agarose gel and supplemented by 'suitably controlled nutrient mixtures' allow studying the effect of surface topography on bacterial colonization (Zhang *et al.* 2014).

The spatial distribution of colonizers of the plant leaf surface, i.e. the colonization of the

phyllosphere (Last 1955), has long been studied (e.g. Leveau and Lindow 2001; Monier and Lindow 2004; Remus-Emsermann *et al.* 2012, 2014). A wide variety of environmental factors and processes was found to contribute to or to hamper the successful establishment of epiphytic microbes (Leben 1970; Roos and Hattingh 1983; Rodenacker *et al.* 2000; Fett and Cooke 2003; Monier and Lindow 2004; Hunter *et al.* 2010; Tecon and Leveau 2012; Remus-Emsermann *et al.* 2014; Esser *et al.* 2015). Many environmental factors are difficult to measure at micrometer-scale resolution but for an increasing number of substances, e.g. carbon, water, or iron, such information became available in the form of bioreporters (Joyner and Lindow 2000; Leveau and Lindow 2001; Axtell and Beattie 2002; Remus-Emsermann and Leveau 2009; Remus-Emsermann *et al.* 2012). They showed the heterogeneous distribution of these factors which affects local success of bacterial colonization in the phyllosphere. Additionally, interactions between the microbial individuals also contribute to the complexity of the observable colonization patterns (Yu *et al.* 2014). Finally, bacterial population density was found to be spatially correlated to the occurrence of certain leaf morphological structures such as veins, stomata, and trichomes (Marcell and Beattie 2002; Monier and Lindow 2004; Hunter *et al.* 2010; Esser *et al.* 2015). For example, glandular trichomes are known to excrete carbon-rich substances (Ascensão and Pais 1998) that can be utilized by bacteria and the crevices (grooves) between epidermal cells have frequently been reported to harbor large numbers of microbial colonizers. Especially for such plant-microbe interactions it is often unclear if they are based solely on some effect of the leaf surface topography or if also a biological effect is active. For example, some trichomes have the ability to better retain water than undifferentiated, smooth leaf areas (Brewer, Smith and Vogelmann 1991). Those areas that retain water the longest are likely to concentrate water-soluble nutrients during times of evaporation. This solely topographical (or physical) process could be amplified by the fact that improved water availability increases nutrient leakage from the leaf interior (Tukey 1970). We would consider such a trans-cuticular process as a leaf biological process, because it would not be present on an artificial biomimetic surface.

The manufacturing of artificial biomimetic leaf surfaces is a growing field of research in materials science. Here, scientists are often interested in producing self-cleaning surface coatings that mimic the superhydrophobic properties of leaves of different plant species such as *Nelumbo nucifera* or *Triticum aestivum* (Sun *et al.* 2005; Koch *et al.* 2007; Schulte *et al.* 2009). Microbial colonization of such artificial surfaces has also become a vivid field

of research. Active fields are the development of sterile surface coatings for medical tools (Ionescu *et al.* 2012; Cerqueira *et al.* 2013; Manabe, Nishizawa and Shiratori 2013; Cole *et al.* 2014; Depan and Misra 2014) or marine applications (Salta *et al.* 2013), or the study of bacterial communities in water treatment facilities (Liu *et al.* 2013; Al Ashhab, Herzberg and Gillor 2014). But only recently, microbial colonization of biomimetic leaf surfaces has been considered for the study of effects of plant leaf topography on bacterial growth and survival on leaves (Zhang *et al.* 2014; Doan and Leveau 2015).

Here, we present the first in-depth spatial analysis of bacterial colonization patterns on such artificial leaf surfaces. The aim of our study was to isolate the effect of topography on bacterial leaf colonization. Our main focus was on detecting differences in the spatial distribution of bacteria growing on bean leaves and on biomimetic bean leaf surfaces made from PDMS. Given the fact that such surfaces have similar physical properties as the biological surface (Zhang *et al.* 2014), we hypothesized that differences between experiments conducted on both surfaces are caused by processes that cannot be explained by surface microstructure alone. Conversely, microbial colonization patterns that do not differ between biological and biomimetic surfaces would suggest that microbial colonization is either surface-independent or at least independent from processes linked to the biology of the surface. We used spatial statistics for a detailed spatial description of bacterial colonization patterns on the artificial surfaces and compared the results to the results of an equivalent study on real bean leaves (Esser *et al.* 2015).

Materials and methods

The experimental setup equaled the setup in Esser *et al.* (2015) except that the bacterial suspension was supplemented with a carbon source to allow bacterial growth and the bacteria were sprayed on artificial leaf surfaces (Doan and Leveau 2015)m. A brief summary is given in the following.

Primary leaves of 2 weeks-old green bean plants (*Phaseolus vulgaris*, variety Blue Lake Bush 274) were washed in deionized water and dried at room temperature. To create a negative mold, one leaf was then placed flat, adaxial surface facing up, in a 145 mm plastic culture dish, and its edges were attached to the dish bottom with masking tape to prevent liquid infiltration underneath. We produced the polydimethylsiloxane (PDMS) mold using a Sylgard elastomer 184 kit (Dow Corning, Flint, MI, USA). We mixed PDMS base and curing agents in a 10:1 ratio (wt/wt), and approximately 40 to 60 g of this mix was poured

into the dish to cover the leaf entirely. After pouring, trapped air bubbles slowly migrated to the surface and were removed by a gentle stream of nitrogen gas. The dish was then covered and placed in an oven at 60 °C for three hours to harden the PDMS. The mold was then cut out of the dish with a scalpel and the real leaf removed. We trimmed the mold to remove excess of PDMS, and we washed it with deionized water containing 1% Triton-X to remove remaining debris. The leaf mold was further placed under a UV light source emitting at 185 and 254 nm (UVP, Upland, CA, USA) for two hours, in order to create a glassy layer of SiO₂ on the surface, and then rinsed again with deionized water. To prepare artificial leaves the cast was placed in a new dish, facing up, and fully covered (1-2 mm above the top) with a new mix of PDMS base and curing agent (10:1 ratio). Air bubbles were removed as described above and the dish was placed in an oven at 60 °C for three hours. Finally, PDMS was removed from the dish and the artificial leaf was manually separated from the mold (which can be washed and reused multiple times), and cleaned prior to experimental use.

In our experiments we used two fluorescent bacterial species: 1) *Pantoea agglomerans* green labeled strains 299R JBA::28, 299R(pFRU48), and red-labeled strain 299R(pFRU97) and 2) *Pseudomonas syringae* red-labeled strain B728a(pFRU97) (Leveau and Lindow 2001; Monier and Lindow 2004; Tecon and Leveau 2012). All strains were grown separately in Lysogeny Broth (LB) with 50 mg kanamycin per liter at 30 °C and harvested in mid-exponential phase by centrifugation and washing twice in M9 minimal medium (Sambrook, Fritsch and Maniatis 2001) with no carbon source. Cells were diluted to an approximate concentration of 10⁷ cells per milliliter. Square sections, about 15×15 mm in size were cut from the artificial PDMS leaves. Two sections each were placed on a plain agarose surface in a petri dish. Inoculation was performed by spraying 50 µl of cell suspension (M9, now complemented with 400 mg/l D-fructose as the only carbon source) onto the artificial leaves using an airbrush Iwata Eclipse HP-CS (ANEST IWATA Corporation, Yokohama, Japan) from about 170 mm above the artificial leaf surface. We moved the nozzle of the airbrush slightly during inoculation to ensure a good cover of the artificial leaf pieces. We either co-inoculated *P. agglomerans* 299R JBA::28 and DsRed-labeled *P. syringae* B728a(pFRU97) or *P. agglomerans* only. Because the use of green- and red-labeled strains proved useful for *a posteriori* assessment of mixing events (Tecon and Leveau 2012; Esser *et al.* 2015), we also used green- and red-labeled strains 299R(pFRU48) and 299R(pFRU97) in these *Pantoea*-only experiments. An exact

overview of the design of spraying experiments is given in Table 2.

Bacteria were observed under the microscope either directly after inoculation ($t = 0$) or first incubated before observation for 24, 48, or 72 hours in sealed petri dishes at 28 °C. In preparation of the microscopic observations, single artificial leaf sections were transferred from the petri dish to microscopic slides and covered with a mounting medium (Aqua Poly/Mount Polyscience Inc., Warrington PA, USA) and a cover slip. Depending on the topography of the PDMS pieces, 25-150 μl of mounting medium were required to ensure a good coverage of the whole surface. The cover slip was then fixed to the glass slide with adhesive tape (Figure 12) to exclude its accidental dislocation which could have introduced a disturbance to the colonization patterns on the artificial leaves.

Table 2: Overview of experiments on artificial leaf surfaces in this study

Section	Red species	Green species	Incubation [h]	10×	20×	40×
80	299R(pFRU97)	299R(pFRU48)	48	2	5	-
81	B728a(pFRU97)	299R JBA::28	0	-	3	2
82	B728a(pFRU97)	299R JBA::28	24	-	19	-
83	B728a(pFRU97)	299R JBA::28	48	-	9	10
84	B728a(pFRU97)	299R JBA::28	72	5	4	-
89	299R(pFRU97)	299R(pFRU48)	0	10	10	-
90	299R(pFRU97)	299R(pFRU48)	24	7	6	-
91	299R(pFRU97)	299R(pFRU48)	48	9	8	-
92	299R(pFRU97)	299R(pFRU48)	72	-	6	-
Total				33	70	12

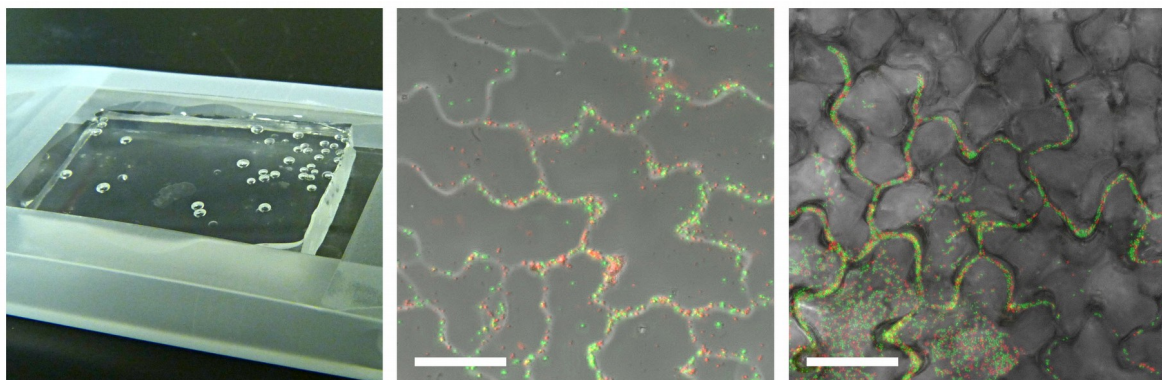


Figure 12: Artificial leaf surface made from PDMS. Left: Photograph of an inoculated section of artificial bean leaf surface. The section has been covered with a cover slip that was fixated by adhesive tape. The air bubbles are inside the PDMS and do not connect to the surface. Center: Micrograph of an inoculated PDMS leaf surface after 24 hours of incubation. Red- and green-labeled fluorescent bacteria mostly settle in the grooves between the puzzle-piece-shaped epidermal cells of the leaf surface. Right: Micrograph of an inoculated real bean leaf after 40 hours of incubation. Scale bars are 50 μm .

From each artificial leaf section we took a series of micrographs (samples) at random locations (Table S5.1, p. 107). The microscope was an Axio Imager.M2 fluorescent microscope (Carl Zeiss AG, Oberkochen, Germany), equipped with Zeiss Neofluar 10x/0.3, 20x/0.5 and 40x/0.6 objectives and an AxioCAM MRn monochrome camera. For fluorescence imaging we used a GFP filter cube (exciter: 470; emitter: 525/50, beam splitter: 495) and a rhodamine filter cube (exciter: 546/12; emitter: 607/80; beam splitter 560). Phase-contrast micrographs were additionally taken to recognize the location of leaf surface structures such as veins and stomata. All images were taken at several planes of focus in order to account for leaf topography and finally saved as z-stacked images in the native .zvi file format.

We used the ImageJ software package (U.S. National Institute of Health, Bethesda, MD, USA) to manually mark the location of bacterial cells and structural elements of the leaf surface, i.e. structures of the artificial leaf surface that could be matched to stomata, hooked and glandular trichomes, as well as veins and the grooves between epidermal cells on the real leaf from which the PDMS mold was made from. Linear grooves between epidermal cells were approximated by a point pattern of groove intersections, i.e. points where three or more grooves joined together. The veins were approximated by the pattern of center points of all vein cells. We saved the locations of all objects as tables of x-/y- coordinates in .csv text files. All analyses were performed using the *spatstat* package (Baddeley and Turner 2005) in R (R Core Team 2013).

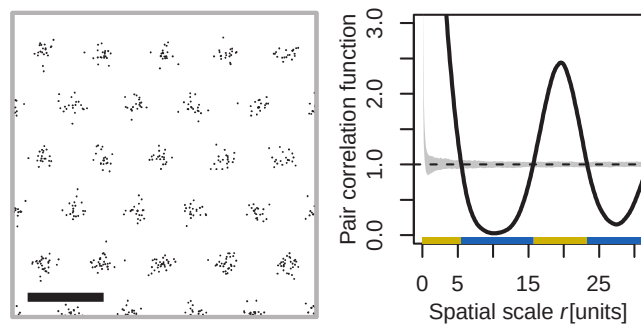


Figure 13: Pair correlation function (PCF) $g(r)$. Left: Artificial pattern of clustered points on a hexagonal grid to illustrate the PCF method. Scale bar is 25 units. Distance between neighboring cluster centers is 10 units. Right: The observed PCF (solid line) of the pattern. The narrow gray band centered at $g(r) = 1$ is an approximately 95% simulation envelope derived from the PCFs of 199 random point patterns of comparable density. Deviations of the observed PCF above the simulation envelope indicate an aggregation of points at spatial scale r , whereas deviations below the bottom of the envelope indicate a scarcity of points at scale r . The colored bar at the bottom of the graph indicates significant aggregation (yellow) and scarcity (blue). Note how the regularity of the hexagonal grid (multitudes of 20 units) is reflected in the PCF.

Equivalently to Esser *et al.* (2015), we estimated pair correlation functions (PCFs; Stoyan and Stoyan 1994; Wiegand and Moloney 2014) from the point pattern data. In general, a

pair correlation function analyzes the distribution of inter-point distances in a point pattern in order to determine if the points are either clustered, randomly or regularly distributed in space (Figure 13). This analysis is performed at a series of different spatial scales r which refer to the inter-point distances. The critical scales, i.e. scales or distances at which the distribution of bacterial cells switches between being clustered, random, or regular, as well as local maxima and minima, can often be matched to typical cluster sizes (diameter, not number of bacterial cells) or distances between cluster centers: For the artificial data in Figure 13, which we specifically designed for illustration purposes, the first transition between the clustered and regular distribution at a little more than 5 units reflects the radius of a typical cluster which was modeled to be about 5 units. Similarly, the minimum at 10 units reflects the empty areas between clusters where a scarcity of points causes a larger scale regularity in the pattern of clusters, where the centers of clusters are typically 20 units apart (reflected by the local maximum at 20 units). PCFs allow an assessment of the nature of the spatial processes that formed the point patterns and the spatial scales at which these processes operate. The choice of pair correlation functions (for homogeneous/inhomogeneous and for single-strain and dual-strain patterns) as well as parameter choices for smoothing kernels and maximum spatial scales up to which we performed the analysis were the same as in Esser *et al.* (2015): Intraspecific interactions were analyzed using the pair correlation for inhomogeneous data (Baddeley, Møller and Waagepetersen 2000). For interspecific interactions between *P. agglomerans* and *P. syringae* we used the cross-type PCF for inhomogeneous point patterns (Stoyan and Ohser 1982; Stoyan and Stoyan 1994). Assuming a homogeneous distribution of leaf structural elements, we also studied the interactions between bacteria and trichomes, stomata and veins using the cross-type PCF for homogeneous data. For interactions with veins, whose cells are aligned along the veins we additionally calculated the inhomogeneous cross-type PCF. More details on these functions can be found in current textbooks (Illian *et al.* 2008; Wiegand and Moloney 2014) and for an application to bacterial data see Esser *et al.* (2015).

We used Monte Carlo simulation envelopes (Ripley 1977; Kenkel 1988; Illian *et al.* 2008) to detect scales at which the observed pair correlation functions deviated from the random null models that represent the absence of interactions between the points. We used two null models: A) We used the Poisson point process for all univariate analyses, i.e. studying the distribution of a single bacterial strain or the combined pattern of all bacteria as a whole.

B) We used a toroidal shift model (Wiegand and Moloney 2014) to test for independence between two point patterns, i.e. between the two bacterial strains or between bacteria and the leaf structural elements. Observed values greater than the upper bound of the simulation envelopes indicate an bivariate association of points at the corresponding spatial scale. Observed values smaller than the lower bound of the simulation envelope indicated a scarcity at the corresponding scale. Scarcity at the smallest spatial scale is often the result of a repulsion between points and may cause a regularity in the pattern as shown in Figure 13.

We summarized the results from all observed samples to detect ubiquitous spatial trends in the patterns that allow assumptions about the processes that created the patterns. We collectively analyzed the results from samples that have been incubated for 24, 48 and 72 hours in order to only find general trends that persist over this early phase of leaf colonization. In our analyses, we were mostly interested in the (empirical) probability to find non-random patterns (aggregation, scarcity) at various spatial scales r , e.g. in the distribution of *Pantoea* cells only or of bacterial cells at distance r around stomata. These probabilities are proxies of how commonly bacteria occur in certain spatial configurations (e.g. aggregates) and how commonly bacteria are associated with (or avoid) leaf structures such as stomata. We compared these probabilities to the respective probabilities reported in Esser et al. (2015) for bacteria colonizing real bean leaves. Differences in empirical probabilities between real and artificial leaf surfaces are used to distinguish physical effects on bacterial colonization patterns, e.g. from leaf topography, from leaf biological effects, such as leaf surface permeability or the secretion of nutrient-rich exudates. Where no differences between the patterns from real and artificial leaves are found, we conversely expected an absence of leaf biological factors and hence physical processes connected to shape, roughness or topography of the leaf surface to be the primary driving forces behind these patterns.

Results

We collected 115 samples from 9 artificial leaf sections (Table S5.1, p. 107). Overall we counted 208,722 bacterial individuals covering a total sampling area of 29.89 mm². After inoculation, bacteria could use the fructose supplied in the bacterial suspension to increase population densities on the artificial leaves (Figure 14). In the single-species experiments, *P. agglomerans* populations increased by more than one magnitude during 72 hours of

incubation (Figure 14). Mixed *P. agglomerans* and *P. syringae* population densities were highest after 48 hours of incubation and slightly declined afterward (Figure 14). Images of samples taken directly after inoculation usually did not contain enough bacteria for our statistical analysis. Visual inspection of these micrographs, however, revealed that the bacterial individuals were well spread across the artificial leaf surfaces suggesting no spatial structure in the bacterial distribution patterns introduced by the spray-inoculation technique using an airbrush system. Pair correlation functions were calculated for incubated samples only. In all experiments, most of the bacterial growth occurred during the first 24 hours of incubation (Figure 14). Thus the pooling of results from samples incubated for 24, 48, and 72 hours is not expected to have added much variability into the results. Not all samples contained trichomes, veins, or stomata which reduced the number of images that could be analyzed with respect to these structures. Final sample sizes are given in the respective figure captions.

General patterns of bacterial distribution on artificial and real leaves

The univariate pair correlation functions describe the distribution pattern of bacteria of the same type. Bacteria in mixed-species experiments were clustered at small scales up to 30 μm , but in some samples aggregation was observed up to 80 μm (Figure 15A). At these larger scales, however, bacteria were more often regularly distributed. In combination, these results indicate that bacteria are distributed in regularly spaced clusters. When considering each species in the mixed-species experiments separately, both species were similarly distributed and consequently the curves (Figure 15B and C) both resemble the curves for the species pool in Figure 15A. When *P. agglomerans* was the only surface colonizer the distribution of cells was qualitatively similar to the patterns described above, however at scales between 10 and 70 μm aggregation was not as commonly found in these samples (Figure 15D).

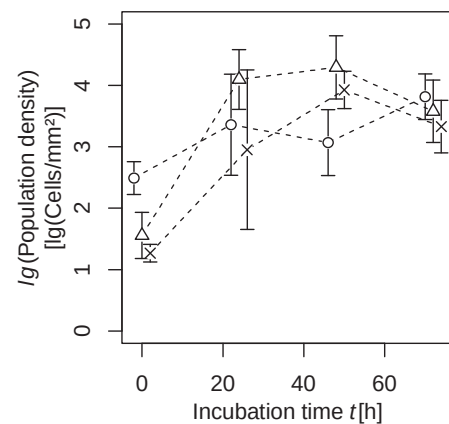


Figure 14: Bacterial growth on artificial leaf surfaces. When *P. agglomerans* was the only species applied to the surface (\circ), population density increased about 21-fold within 72 hours after inoculation. When *P. agglomerans* and *P. syringae* were both applied to the surface at comparable initial concentrations, both strains were similarly successful in reproducing with *P. agglomerans* (\square) increasing about 105-fold and *P. syringae* (\times) by about 115-fold after 72 hours after inoculation.

The patterns of bacterial distribution in dual-strain experiments on artificial leaves qualitatively resembled the distribution on real leaves under comparable conditions reported in Chapter 4: On both surfaces a small-scale aggregation and a larger scale regularity was observed in the patterns. Deviations between the patterns on both surfaces were mostly confined to outside the direct neighborhood of cells, i.e. scales larger than 10 μm (Figure 16). On this intermediate scale ($\sim 10\text{-}70 \mu\text{m}$), the study of bacteria not considering species revealed that bacterial clustering was more frequently observed on artificial leaf surfaces than on real leaves (Figure 16A). For example, on artificial leaves, clustering occurred on scales up to 70 μm in over 55% of the samples (Figure 15A), whereas on real leaves clustering at this scale was found in merely 5% of the samples (Figure 8A). The difference of over 50% can directly be read from the line showing aggregation in Figure 16A. The difference between real and artificial leaf surfaces can be traced back to the *Pantoea* cells in these experiments who exhibited very similar deviations between their distribution on real and artificial

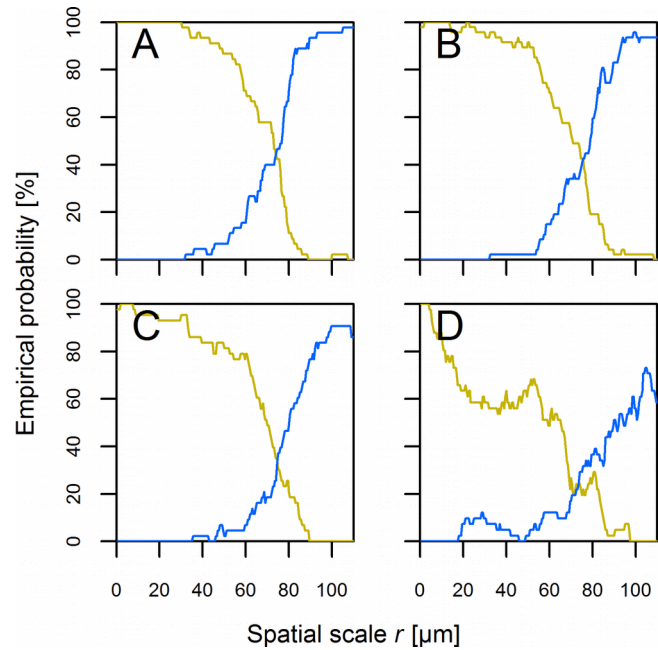


Figure 15: Frequency of intraspecific aggregation (yellow) and scarcity (blue) of bacteria on artificial leaves across spatial scales. (A) Analysis of bacteria irrespective of their species. (B) Analysis of *P. agglomerans* in competition with *P. syringae*. (C) Analysis of *P. syringae* growing in competition with *P. agglomerans*. (D) Analysis of *P. agglomerans* as the sole colonizer of the artificial leaf. Sample sizes were 48 in (A), 51 in (B), 45 in (C) and 61 in (D).

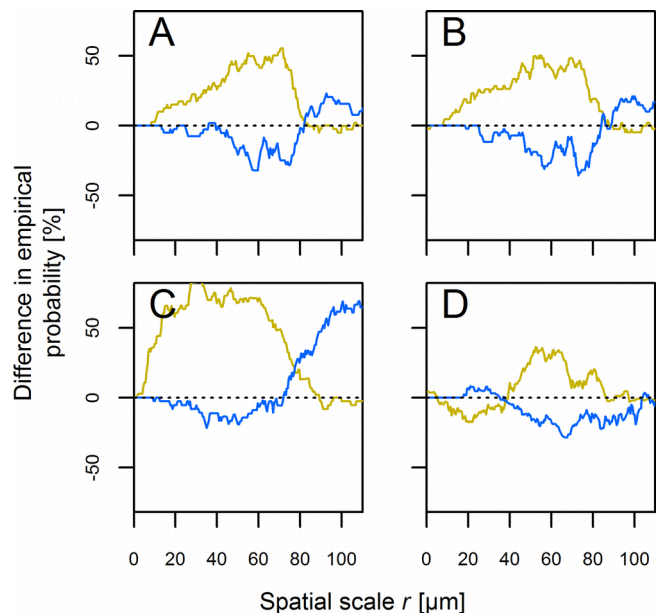


Figure 16: Differences $F_{\text{artificial}} - F_{\text{leaf}}$ in the relative frequency F of intraspecific aggregation (yellow) and scarcity (blue) of bacteria on artificial and real leaves. (A) – (D) and sample sizes were the same as in Figure 15.

leaves (Figure 16B). The *P. syringae* cells, however, exhibited much more aggregation on intermediate scales on artificial surfaces compared to real bean leaves (Figure 16C). This means that, although *P. syringae* was commonly clustered on both surfaces, cluster sizes were smaller on real leaves (5-10 μm) and thus more strongly clustered than under similar conditions on artificial leaves. When *P. agglomerans* was the only colonizer of the artificial leaves, these cells were less aggregated at scales up to 40 μm , compared to real leaves (Figure 16D).

Co-location of *P. agglomerans* and *P. syringae* on artificial & real bean leaves.

The bi-variate inhomogeneous pair correlation functions describe spatial correlations between individuals of two different strains, e.g. *P. syringae* and *P. agglomerans*. On artificial bean leaves, *P. agglomerans* cells aggregated typically within 20 μm around *P. syringae* cells and vice versa (Figure 17). This co-location of individuals of both species was almost symmetrical for *P. syringae* cells aggregating near *P. agglomerans* cells and *P. agglomerans* cells aggregating near *P. syringae* cells (Figure 17A and B). In about 70% of our samples we found significant aggregation at the smallest spatial scale, indicating that cells of both species can be direct neighbors forming mixed colonies and aggregates. We regularly observed

this also directly in the micrographs, e.g. in the center panel of Figure 12. The co-aggregation was usually not found to be significant at scales larger than 20-30 μm . Starting at distances larger than 30 μm , there was a slight tendency of segregation between strains, but this was typically observed in less than 20% of our samples.

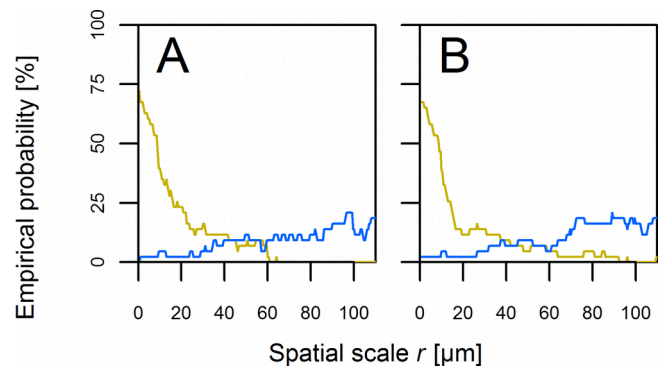


Figure 17: Relative frequency of interspecific co-aggregation (yellow) and scarcity (blue) of bacteria on artificial leaves across spatial scales. (A) Aggregation of *P. agglomerans* around *P. syringae*. (B) Aggregation of *P. syringae* around *P. agglomerans*. Sample sizes were 45 in (A) and (B).

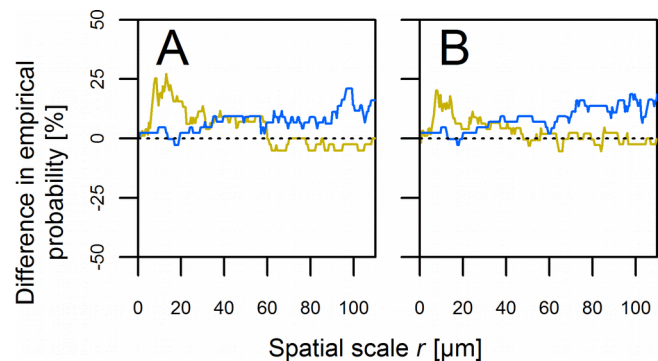


Figure 18: Differences $F_{\text{artificial}} - F_{\text{leaf}}$ in the relative frequency F of interspecific co-aggregation (yellow) and scarcity (blue) of bacteria on artificial and real leaves. (A) and (B) and sample sizes were the same as in Figure 17.

On real leaves, the co-aggregation of *P. agglomerans* and *P. syringae* was observed at the same spatial scale although slightly less frequent than on artificial leaves (Figure 18). Interestingly, the weak segregation of the two strains at scales larger than 30 μm was a phenomenon only observed on artificial leaves (Figures 17 and Figure 18).

Role of structural elements of the leaf surface on bacterial distribution

Grooves between epidermal cells (more precisely, the intersection points of grooves) were the only surface structure of artificial leaves for which we found a clear effect on the distribution of bacteria (Figure 19A). In more than 65% of our samples, we found significant aggregation of bacteria at the groove intersection points suggesting an association of bacteria with the grooves in general. Aggregation was limited to distances of 15 μm away from the groove intersections. On real leaves, we found a strikingly similar distribution of bacteria near grooves, although the aggregation was somewhat more prevalent in samples from artificial leaves (Figure 20A).

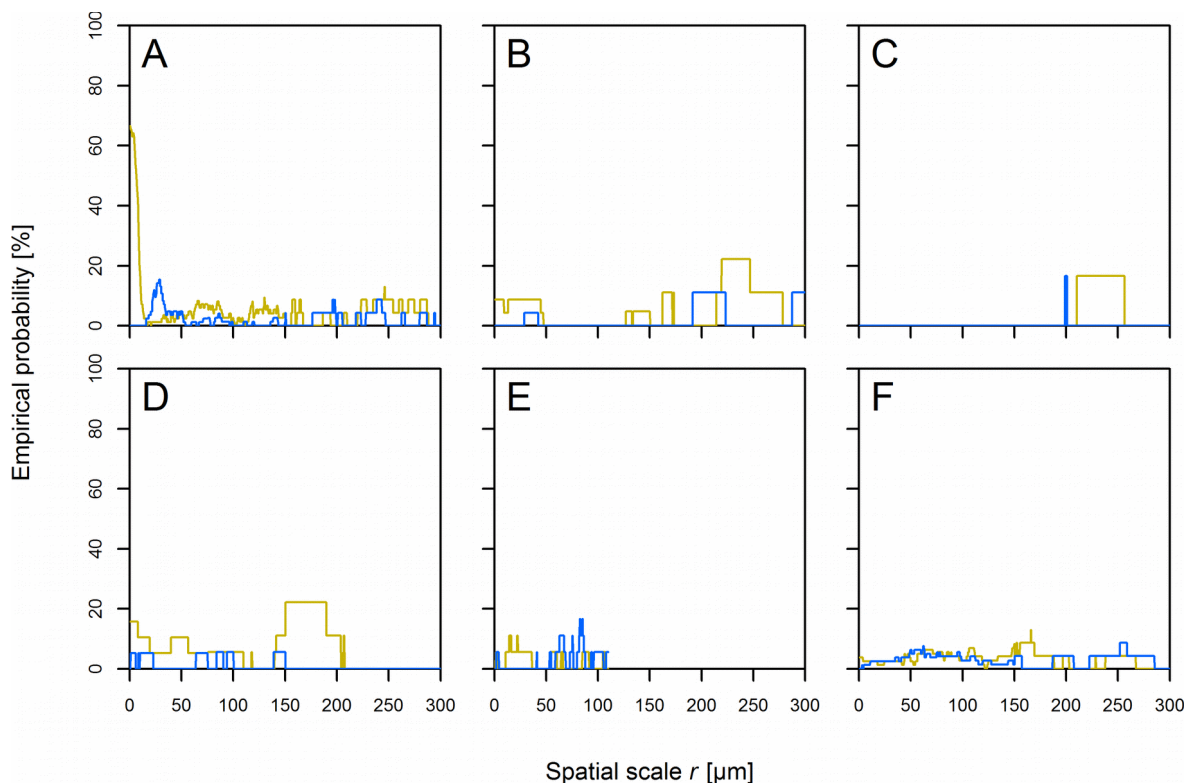


Figure 19: Bacterial aggregation (yellow) and scarcity (blue) around structural elements of the artificial leaf surface. (A) Aggregation of bacteria near intersection points of grooves between epidermal cells. (B) Interactions with glandular trichomes. (C) Interactions with hooked trichomes. (D) and (E) Interactions with vein cells. (F) Interactions with stomata. In (E), we used the inhomogeneous partial pair correlation function whereas in the remaining analyses, we used the homogeneous partial pair correlation function. Sample sizes were 109 in (A), 31 in (B), 17 in (C), 24 in (D), 21 in (E) and 104 in (F).

We found no clear small-scale correlations between the location of bacteria and neither glandular trichomes (Figure 19B) nor hooked trichomes (Figure 19C). This was different from real leaf surfaces, where an aggregation of cells was observed within 120 μm and more around both types of trichomes (Figure 20B and C).

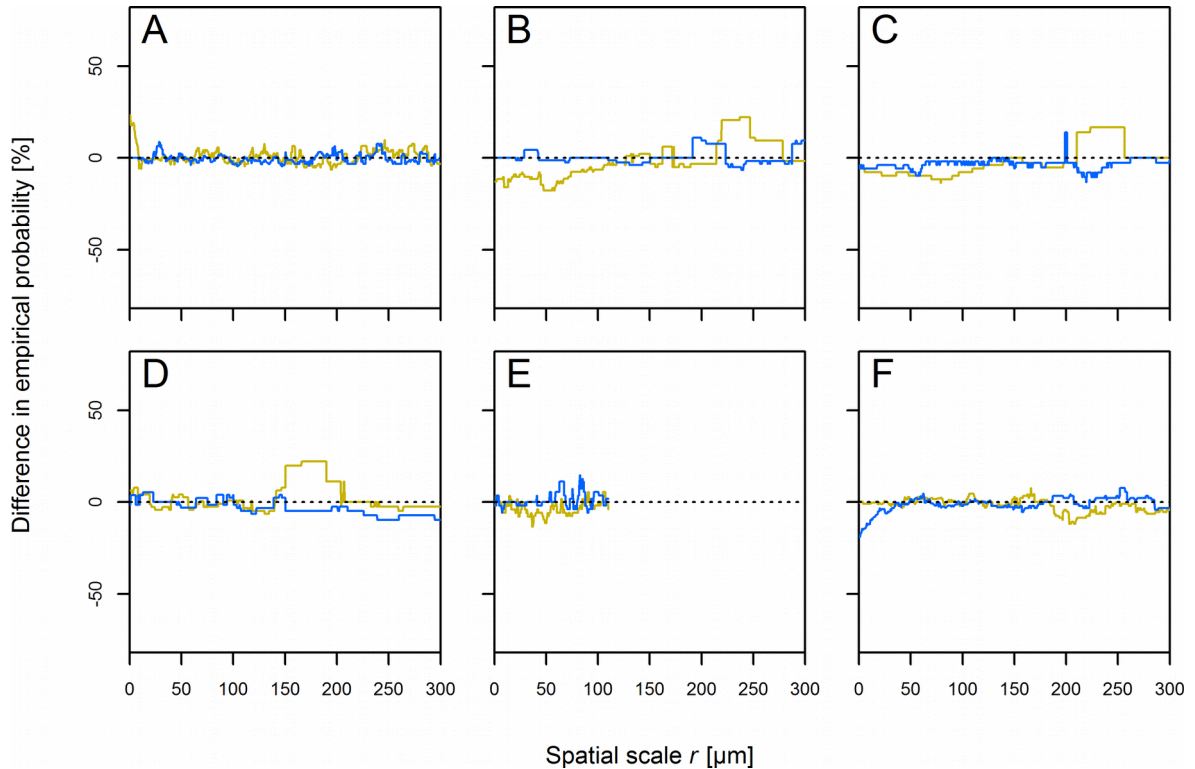


Figure 20: Differences $F_{\text{artificial leaf}} - F_{\text{real leaf}}$ in the relative frequency of bacterial aggregation (yellow) and scarcity (blue) around structural elements of the artificial leaf surface. (A) – (F) and sample sizes were the same as in Figure 19.

In less than 20% of our artificial leaf samples, we found bacteria aggregating near veins (Figure 19D) and no such effect was observed when we accounted for the inhomogeneous distribution of vein cells (Figure 19E). Not considering the inhomogeneous distribution of vein cells allowed no differentiation between bacterial distribution patterns near veins on artificial and real bean leaves (Figure 20D), whereas accounting for this heterogeneity showed that the weak aggregation of bacteria within 100 μm around vein cells found on real leaves was absent from artificial surfaces (Figure 20E).

Stomata had no effect on the distribution of bacteria in our experiments on artificial leaf surfaces (Figure 19F) whereas a repulsion of cells from these structures was observed on real bean leaves (Figures 20F and Figure 10B, p. 42).

Discussion

The distribution patterns of bacteria colonizing our artificial leaf surfaces generally resembled the distribution of bacteria found on real leaves. The bacteria mostly formed loose aggregates with only few areas with colonies of adjacent bacterial individuals and a frequent occurrence of single individuals in the wide areas between aggregates. These general patterns have been frequently reported for bacteria colonizing leaves (Kinkel 1997; Leveau and Lindow 2001; Monier and Lindow 2004; Remus-Emsermann *et al.* 2014; Esser *et al.* 2015). The similarity between bacteria colonizing real and artificial leaves illustrates the ability of micro-molded PDMS surfaces to mimic the physical properties of natural environments such as the topography. Zhang *et al.* (2014) recently showed that artificial spinach leaves made from PDMS not only capture the topography but also closely mimic wettability of the real leaves. This predestines PDMS for the replication of plant leaves and other similarly hydrophobic surfaces but coating of PDMS surfaces with an appropriate surfactant (if not produced by the microbes) may open avenues towards a wide variety of artificial surfaces to be colonized by microbes.

Of all correlations considered in our study, the aggregation of bacteria around the grooves between epidermal cells was the most remarkable. On both surfaces, real and artificial bean leaves, bacteria aggregated closely to these structures and the distance decay of this aggregation away from the grooves was very similar both in shape and range (~15 μm) (Figures 19A, 20A and 10A). This suggests that this important plant microbe interaction that has been frequently reported in the literature (Diem 1974; Blakeman 1985; Davis and Brlansky 1991; Leveau and Lindow 2001; Monier and Lindow 2004; Esser *et al.* 2015) can be explained by the topographic properties of the leaf surface without the need of additional leaf biological processes such as increased nutrient leaching near grooves. Two mechanisms come to mind: It is reasonable to assume that, under dry conditions, the deeper grooves between epidermal cells retain more water from evaporation than the exposed epidermal cell tops, e.g. by capillary forces. Under such conditions, the grooves may serve as a protected habitat and, at the same time, concentrate nutrients from the evaporating solution. The other probable mechanism is based on the fact that small particles in solution move slower near surfaces than in the open liquid phase (Dusenberry 2008). Grooves between epidermal cells constitute locations of high surface density, i.e. there is more leaf surface per liquid volume compared to the top of an epidermal cell or the tip of a hooked trichome. Assuming a random movement of motile bacteria, these cells

would thus get stochastically entrapped within the grooves.

In contrast to the grooves between epidermal cells, the differences in the aggregation patterns of bacteria around glandular trichomes on real and artificial leaves suggest that an additional leaf biological process is required to explain the patterns observed on real leaves. It is important to note that the analysis of real and artificial leaf surfaces with spatial statistical methods only allow an assessment of the nature of the processes that determine bacterial distribution in the phyllosphere, i.e. differentiating physical/topographical effects from leaf biological effects. They do not prove the presence or importance of specific processes such as the secretion of nutrients from trichomes (Ascensão and Pais 1998) or the retention of water droplets to the leaf surface (Brewer, Smith and Vogelmann 1991). Consequently, the absence of spatial correlations between bacteria and hooked trichomes, veins and stomata on artificial leaves (Figure 19C, E, and F) also suggests that such correlations observed on real leaves (Figure 10D, F, and B, respectively) are mediated by leaf biological processes but not topographic effects. The exact leaf biological processes that caused these spatial correlations remain nevertheless unknown. Spatial point pattern analysis of distribution patterns of surface-colonizing microbes is a tool for exploratory studies that can reveal the relative importance and spatial scales of unknown microbial interactions. The true mechanism behind these interactions remain to be revealed by complementary analysis. In our study, for example, we isolated the effect of leaf topography from leaf biological processes by comparing the colonization patterns of bacteria on leaves and biomimetic artificial leaf surfaces. Similarly, knock-out studies could be combined with spatial point pattern analyses to isolate and study the effects and spatial scales of known physiological processes, e.g. biofilm formation or antimicrobial resistance, on bacterial colonization of surfaces.

In our study we established that topography is the main driver of bacterial aggregation near grooves and found that the aggregation of bacteria near trichomes and veins (and the avoidance of stomata) involve additional leaf biological processes. Consequently, the general distribution patterns of bacteria on artificial leaf surfaces, at least around these structures, needs to be different from the distribution on real leaves. On both surfaces the bacteria were aggregated on small scales. However on intermediate scales (10-70 μm), we observed additional aggregation more frequently on artificial leaves than on the real leaves (Figure 18). This was an unexpected result considering the fact that larger scale aggregation patterns around trichomes and veins were only observed on real bean leaves

(Figure 10) and that these patterns should also occur in the intraspecific distribution analysis (Figure 15). Instead, at these scales, we found additional intraspecific aggregation on artificial leaves where the aggregation around trichomes and veins was not active. This is probably due to the fact that population densities on artificial leaves were about 20 times higher than in our experiments on real bean leaves (Esser *et al.* 2015). Probably, the fructose levels added to the M9 medium in the experiments on artificial leaves were at higher concentrations than the sugar availability on natural leaves or more homogeneously distributed. The population densities found on artificial leaf surfaces did not exceed population densities of 10^6 to 10^7 cells/cm² that have been reported for typical leaves (Lindow and Brandl 2003) but it is valid to assume that the higher colonization rates on artificial leaves result in larger bacterial aggregates and thus in larger scales of significant clustering compared to the less densely populated real bean leaves from our earlier our experiments (Esser *et al.* 2015). The fact that the bacterial aggregation near grooves was limited to about 15 μ m around grooves on both surfaces however suggests that the larger colonies stretch out along the grooves rather than growing into more distant regions away from the grooves. This would also explain the differences in the co-aggregation of our two model strains on real and artificial leaf surfaces, i.e. an equal co-aggregation at scales up to 15 μ m but more prevalent co-aggregation at larger scales up to 60 μ m on the artificial surfaces (Figure 18). In our opinion, the development of larger colonies along the grooves is evidence that topographical effects can overrule other bacterial pattern formation processes such as clonal growth or the formation of an extracellular matrix. By extension, this has to be true also for bacterial colonization of some other biological, environmental, or technical surfaces. At least, it illustrates that micro-topographic effects have to be considered to fully understand microbial surface colonization.

Chapter 6 – Scales of interactions between a point pattern and a line pattern: an extension to current methods in spatial statistics

This chapter is under preparation as a manuscript for submission to *Methods in Ecology and Evolution*

Daniel S. Esser¹, Martin Schlather², Katrin M. Meyer¹, Kerstin Wiegand¹

¹ Department of Ecosystem Modelling, University of Göttingen, Germany

² School of Business Informatics and Mathematics, University of Mannheim, 68131 Mannheim, Germany

Abstract

Spatial studies in ecology often involve entities such as individuals, nests, colonies or small resource patches whose locations in space are regularly approximated by a single point coordinate. Spatial correlations and interactions between these point-like objects are frequently studied relying on a rich body of spatial point pattern analysis tools such as the pair correlation function. Nevertheless, there exists a wide variety of spatial objects that are better described by a pattern of lines, e.g. roads, streams, geological faults, or cracks. The analysis of effects of such linear structures on the distribution of point-like objects has received considerably less attention, both in theory and application. Here, we introduce the line-to-point pair-correlation function (LPPCF) as an extension of the pair correlation function. The LPPCF is able to detect effects of linear objects on the location of point-like entities, it identifies the spatial scale at which these effects operate, and determines if they are of an attracting or inhibiting nature. We present several types of LPPCFs which only differ in the definition of the distance between a point and a line. We further propose estimators for the LPPCF based on four different line-to-point distances. We demonstrate the performance of our method by using (a) artificially generated point- and line- patterns and (b) the data sets of bacterial colonizers on bean leaves from Chapter 4. The line-to-point pair-correlation function will help ecologists to better understand how the spatial distributions of organisms are affected by linear structures such as roads, streams, hedges or other types of ecotones.

Introduction

In ecological studies, the distribution of individuals and objects are routinely approximated by spatial point patterns where each point represents the location of an individual or object. The sum of these locations forms a point pattern and each observed point pattern is then seen as one realization of a spatial process that created the pattern. The analysis of such point patterns may therefore allow insights into the mechanics of biotic and abiotic interactions. Consequently, the analysis of spatial point patterns has become increasingly important in ecology. It has been widely used to study the distribution of individuals in a population (Law *et al.* 2009; Miller *et al.* 2010; Murphy and McCarthy 2012) and the interactions between individuals of different age classes (Franklin and Santos 2011; Pillay and Ward 2012) or individuals from different taxonomic groups (Schurr *et al.* 2004; Wiegand *et al.* 2007).

Within the field of spatial point pattern analysis, so-called second-order summary statistics such as Ripley's K function (Ripley 1976), its widely used transformation, the *L*-function (Besag 1977) and its derivative, the pair correlation function (Stoyan and Ohser 1982; Stoyan and Stoyan 1994) constitute a class of especially powerful tools to describe and analyze point patterns (Wiegand, He and Hubbell 2013). All these methods analyze the frequency distribution of the distances between the points of the same type (univariate case) or between points of two different types (bivariate case) and relate them to the expected distribution given a predefined null model. The most common null model used in such analyses is complete spatial randomness (CSR), where all points are randomly distributed in space and independent from the location of other points. The comparison of observed inter-point distances and the distances predicted by the null model allows an assessment of certain classes of inter-point distances being uncommonly abundant or scarce compared to the null model. In case of the CSR-model this then translates to the points in the pattern being aggregated or scarce at the corresponding spatial scales.

Often in ecology, the location of point-like entities such as individuals, colonies, or nests may not so much be affected by the location of other point-like objects of the environment but by the presence of linear structures such as roads, streams, hedges, or cracks in parched soil. Our interest in this topic was motivated by a problem from microbial ecology. Bacterial plant leaf colonizers (Figure 21) are known to settle close to or within grooves between epidermal cells of the leaf surface (Diem 1974; Blakeman 1985; Davis and

Brlansky 1991; Leveau and Lindow 2001; Monier and Lindow 2004; Esser *et al.* 2015), but the exact scales of these interactions are unknown. The grooves can be represented by a pattern of straight line segments that form a complex network.

Compared to spatial point pattern analysis, only little attention has been spent on the analysis of interactions between a point pattern and a line pattern. Nevertheless, there exists some literature on the statistical theory of interactions with line patterns (Mecke and Stoyan 1980; Stoyan, Mecke and Pohlmann 1980; Mecke 1981; Stoyan and Ohser 1982; Stoyan 1984; Heinrich and Schwandtke 1990; Krasnoperov and Stoyan 2004; Weiss, Ohser and Nagel 2011). With some of this work originating in material science, interactions with linear structures were often referred to as *fibre processes*, although this term may not be intuitive

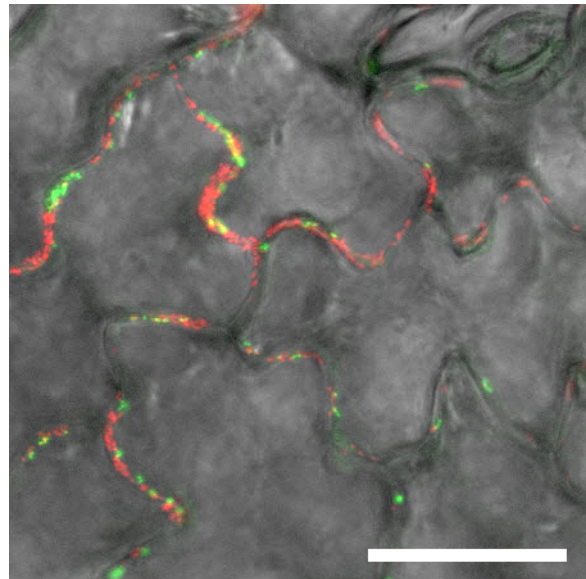


Figure 21: Microscopic view of a bean leaf surface with bacterial colonizers carrying red (*Pseudomonas syringae*) and green (*Pantoea agglomerans*) fluorescent markers. All bacterial cells sit in puzzle-piece shaped grooves that developed at the margins of leaf epidermal cells. Scale bar is 50 μm .

within an ecological framework. These studies presented measures that describe second-order properties of line patterns and in few cases cross-correlations between different line patterns. A comprehensive overview is given in Chiu *et al.* (2013, chapter 8). For example, Stoyan and Ohser (1982) proposed a wide variety of cross-correlation measures to study the interactions between random sets and introduced three estimators which analyze the cross-correlation between a point process and a fibre process. However, these estimators have several drawbacks such as being computationally intensive and difficult to handle in application.

The Berman-Test (Berman 1986) is another tool used to confirm the association of a point pattern with other geometrical structures, such as points, lines or areal objects. It is based on the shortest distance between a point of the point pattern to the nearest spatial object and is therefore 'short-sighted' as it does not include objects that are further away than the nearest neighbor. Similarly the Nearest Neighbor Distance method (Okabe and Fujii 1984) also is a short-range method. Finally, Foxall and Baddeley (2002) introduced an adaption

of the J -function (van Lieshout and Baddeley 1996), a normalized nearest-neighbor distance distribution function. Bedford and van den Berg (1997) argue that the J -function may find a random distribution of points around the line pattern despite of the presence of interactions. Although this is a common problem of summary characteristics, those which have memory, such as the J -function, are more prone to this type of error (Wiegand and Moloney 2004).

In this paper, we propose the *line-to-point pair correlation function* (LPPCF), a method to examine the influence of a linear structure on the distribution of point-like objects in its vicinity. It incorporates the benefits of non-cumulative, second-order summary statistics and also studies spatial scales beyond nearest-neighbor distances. Furthermore, our approach suits the common practice to approximate linear structures by a set of straight line segments. As an extension of the pair-correlation function, the LPPCF evaluates the number of points in a neighborhood of a line segment in the line pattern. We can make inferences about the nature of the underlying processes (aggregation or scarcity of points near linear structure) and about the spatial scale at which these processes operate. We also present several estimators for the LPPCF. We then illustrate the principle of our method in several simulation studies. Finally, we apply the LPPCF to the patterns of leaf-colonizing bacteria on bean from Chapter 4.

Materials and Methods

The line-to-point pair correlation function (LPPCF)

In the following, we introduce the line-to-point pair correlation function for patterns in two-dimensional space (\mathbb{R}^2) which is the most important case in application. Nevertheless, LPPCFs can be derived for higher dimensions, accordingly, using existing estimators of the partial pair correlation function.

Let Φ_p be a stationary point process on \mathbb{R}^2 . The intensity λ_p of the point process equals the expected number of points per unit square. A point pattern φ_p is one realization of this process observed inside an observation window W . Furthermore, x_p denotes the p -th point in φ_p and we write $n_p = \Phi_p(W)$ for the number of points in the observation window. The area of W is denoted by $|W|$.

Similarly, we define Φ_l to be a stationary line process and the line pattern φ_l to be one observed realization of this line process inside the observation window W . Then x_l denotes

the l -th line in the line pattern. Its length is denoted by $|x_l|$. The intensity of the underlying line process is denoted by λ_l and equals the expected total line length per unit square. The number of lines that have at least one point in W is denoted by n_l . The total length of all line segments in W is given by $\Phi_l(W)$.

We first describe the concept of pair correlation functions for a single point pattern ϕ_p . If we consider two infinitesimally small discs centered at locations x_1 and x_2 with areas dx and dy , respectively, then $\varrho(x_1, x_2)dx dy$ gives the probability to find one point of the point pattern in each of the two discs (Stoyan and Stoyan 1994). The term $\varrho(x_1, x_2)$ is called the second-order product density. As Φ_p is assumed to be stationary, the product density is independent of the locations of the points x_1 and x_2 themselves and depends on their distance h , only:

$$\text{Eq. 6.1 : } \varrho(x_1, x_2) = \varrho(h), \quad h = |x_1 - x_2|$$

The pair correlation function (Stoyan and Ohser 1982; Stoyan and Stoyan 1994) is a normalization of the product density and is defined as

$$\text{Eq. 6.2 : } g(h) = \frac{\varrho(h)}{\lambda_p^2}.$$

Interactions between the points of two different point processes Φ_p and Φ_q can be characterized by the partial pair correlation function (also termed cross-type pair correlation function) which is defined as

$$\text{Eq. 6.3 : } g_{pq}(h) = \frac{\varrho_{pq}(h)}{\lambda_p \lambda_q}.$$

Here, $\varrho_{pq}(h)dx dy$ gives the probability to find a point of type p in an infinitesimally small disc of area dx and a point of type q in a similar disc of area dy and the distance between the centers of the two discs is h (Stoyan & Stoyan 1994). Both pair correlation functions (Eq. 6.2 and Eq. 6.3) share the property that $g(h) = 1$ when there are no interactions between the points at scale h . Values $g(h) > 1$ indicate an aggregation of points at scale h whereas values $g(h) < 1$ indicate a scarcity of points.

Our LPPCF is based on the partial pair correlation function (Eq. 6.3). We write:

$$\text{Eq. 6.4 : } g_{lp}(h) = \frac{\varrho_{lp}(h)}{\lambda_l \lambda_p}$$

There are two important aspects to be considered compared to the partial pair correlation function: Firstly, the line density λ_l is given in length per square unit compared to number

of points per square unit for λ_p . Secondly, the product density $q_p(h)$ can only be formulated for a single point on a line but not for the line as a whole. Most importantly, the formulation of a LPPCF depends on how the distance $d(x_l, x_p)$ between the line and the point is defined. This is actually true for any pair correlation function that one would like to formulate for patterns of spatial objects other than points, i.e. lines, areas, or volumes. Once an appropriate distance measure is selected, we can estimate the LPPCF. Analogously to Illian et al. (2008, p. 282), we use the estimator for the partial pair correlation function and write

$$\text{Eq. 6.5 : } \hat{g}_{lp}(h) = \frac{\hat{q}_{lp}(h)}{\lambda_l \lambda_p} = \frac{1}{2\pi h \lambda_l \lambda_p} \sum_{l=1}^{n_l} \sum_{p=1}^{n_p} \frac{k(d(x_l, x_p) - h)}{|W_{\bar{x}_l} \cap W_{x_p}|}$$

for $h > 0$, where $k(\cdot)$ is a kernel function, typically a box kernel or the Epanechnikov kernel (Illian et al. 2008). W_x denotes the translated observation window $W_x = \{z + x : z \in W\}$. The term $|W_{\bar{x}_l} \cap W_{x_p}|$ is therefore the weight of the translational edge correction (Illian et al. 2008), where \bar{x}_l is a midpoint of the l -th line segment. The midpoint is not uniquely defined, e.g. it could be the center of gravity of the line segment or a point halfway running along the line segment. In our examples, where all lines are straight line pieces, both points are identical. For the intensities we used the standard estimator (Illian et al. 2008, p. 189) $\hat{\lambda}_l = \Phi_l(W)/|W|$ and $\hat{\lambda}_p = \Phi_p(W)/|W|$, respectively.

A possible choice for the line-to-point distance $d(x_l, x_p)$ is the distance from the focal point to the mid-point of the line.

$$\text{Eq. 6.6a : } d_{MID}(x_l, x_p) = d(\bar{x}_l, x_p).$$

This way, the line pattern is reduced to a point pattern (of mid-points) and the partial pair correlation function (Eq. 6.3) can be applied. Unfortunately, this approach loses most of the information from the line pattern. If we choose the distance between the point and the nearest point on the line $x_{\min, l, p}$,

$$\text{Eq. 6.6b : } d_{MIN}(x_l, x_p) = d(x_{\min, l, p}, x_p).$$

more information might be preserved because potentially any point $x \in x_l$ can be $x_{\min, l, p}$, depending on the configuration of the points in ϕ_p .

Since the purpose of the LPPCF is to analyze the effect of a line x_l on the location of a point x_p in space, it might be reasonable to assume that any part of the line, i.e. every point $x \in x_l$, affects the point x_p . Furthermore, the strength of this effect may depend on the

distances between the points $x \in x_l$ and x_p . These aspects are not covered by any of the above approaches. To account for all points on a line, we suggest the use of average distances \bar{d}_{lp} between a line and a point. Now, at least one point x'_l on the line segment exists for which $d(x'_l, x_p) \equiv \bar{d}_{lp}$.

$$\text{Eq. 6.6c : } d_{MEAN}(x_l, x_p) = \frac{1}{|x_l|} \int_{x_l} d(x, x_p) dx$$

gives the average distance between the point and the line segment.

Alternatively, we can define a distance-dependent potential $U_{lp}(d(x, x_p))$, which describes the effect of the point $x \in x_l$ on the location of x_p . Then, the average distance between the line and the point is given by

Choosing $U_{lp}(d) = d^{-1}$, the distance measure in Eq. 6.6d gives values closer to d_{MIN} , whereas d_{MEAN} generally gives larger values than d_{MIN} . Therefore, Eq. 6.6d is the distance measure that uses most information from the line pattern while being still biologically interpretable. The explicit formulas for d_{MEAN} and d_{POT} are derived in the Supplementary Materials section S6.1 at the end of this thesis. The implementation of the LPPCF-methods in our R-package `lppcf` is described in the Supplementary Section S6.2.

$$\text{Eq. 6.6d : } d_{POT}^{-1}(x_l, x_p) = \bar{U}_{lp}(x, p) = \int_{x_l} U_{lp}(d(x, x_p))$$

Evaluation of the line-to-point pair correlation function in a simulation study

In a simulation study, we tested the four different types of LPPCFs presented above, i.e. the functions based on d_{MID} , d_{MIN} , d_{MEAN} and d_{POT} . The goals of the simulation study were: (a) to establish that the LPPCFs give the expected results when points are attracted by, or repulsed from a line pattern, (b) to study the behavior of the LPPCFs depending on how well a line pattern is approximated by a set of straight line segments and (c) to compare the behavior of the 4 types of LPPCFs to each other.

The general behavior of the four LPPCFs was tested using an artificial line pattern within a unit square observation window (Figure 22, p. 74, left column). The line pattern forms a network of sine-shaped lines, similar to the grooves between epidermal cells of bean leaves (cf. Figure 21) The distance between parallel sine-waves was 0.2 units such that the cell centers are about 0.1 units away from the closest line. We constructed 5 different point

patterns based on this line pattern representing different probable interactions between the points in the point pattern and the lines: The first point pattern (Figure 22A) was constructed with the points aggregating closely around the line pattern. In the second pattern (Figure 22B), points were designed to aggregate in a buffer zone at distance 0.03-0.05 units from the line pattern. In the third pattern (Figure 22C), the points followed a random distribution that was not affected by the presence of the lines. Pattern 4 (Figure 22D) was constructed such that the points follow a random distribution except for avoiding a buffer zone 0.02 to 0.04 units from the closest point on a line segment. This pattern can be seen as the inverse of the point pattern in Figure 22B, with slightly different parameters for the buffer zone. Finally in pattern 5 (Figure 22E), the points were designed to actively avoid the line pattern.

We calculated the four different LPPCFs for each of the five point patterns and chose to use no smoothing of the results (bandwidth of the Epanechnikov kernel = 0) in order to not lose any deviations between the 4 types of LPPCFs.

Mode of segmentation

The line pattern in our examples (Figures Figure 22A-E) represent sets of continuous (sine) curves that were approximated by sets of straight line pieces or 'segments'. This is a common practice in geographical information system (GIS) applications but often it will be difficult to justify the choice of segment lengths or, more broadly speaking, the resolution that was chosen for the approximation. Nevertheless, one has to consider this segmentation issue when working with LPPCFs. Obviously, the average distance between a point and a line segment d_{MEAN} , for example, will increase with segment length $|x_i|$, even if the shortest distance between the point and the line segment d_{MIN} stays constant. The finer the resolution of the segmentation process, the shorter are the final line segment lengths and the closer the results of the four LPPCF modes will be. For the estimator based on d_{POT} this effect is less severe than for d_{MEAN} because distant points of the line segment get smaller weights when calculating this other type of average distance. Values for d_{POT} will thus always be closer to d_{MIN} in the first place. In order to assess the effect of segmentation, we used different approximations ('inadequate', 'adequate', 'fine', see Figure 23, p. 75, left, center, and right, respectively) of the line pattern used in Figure 22 and calculated the LPPCFs for the point pattern from Figure 22A (aggregation of points near the lines). We chose this point pattern because it has the shortest line-to-point distances (d_{MIN}) and

therefore is most prone to the effect of long line segment lengths. Moreover in our analyses in Figure 22, it yielded the most dynamic LPPCFs and should therefore be suited best for capturing changes due to a coarse segmentation resolution. The segment lengths in the three line patterns were smaller than 0.091 units in the inadequate pattern, smaller than 0.051 units in the adequate pattern, and smaller than 0.004 units in the fine pattern.

Application: Bacteria on plant leaves

In Chapter 4, we considered the spatial scale of the aggregation of leaf-colonizing bacteria near grooves between leaf epidermal cells (see also Esser *et al.* 2015). We combined point location data of the bacteria on leaves and intersection points of the linear grooves in a spatial analysis using the partial pair correlation function. Although this approach proved useful, strictly speaking it only confirmed the aggregation of bacteria near *groove intersections* which is less intuitive and biologically less meaningful than considering the aggregation near the full pattern of *linear grooves*. This was the main motivation behind the development of the LPPCF. It can be discussed with respect to the groove lines instead of only groove intersection points. We ‘field-tested’ the LPPCF by analyzing the bacteria-groove data sets from Chapter 4. We manually extracted the line patterns from the source images using ImageJ software (Schneider, Rasband and Eliceiri 2012) and transformed the data into the *linnet*-format, native to the *spatstat* package for *R* (Baddeley and Turner 2005; R Core Team 2013). The distribution of segment lengths were approximately log-normal (not shown) and we decided not to cut them into smaller pieces. We only calculated g_{MIN} and g_{POT} because we consider these to be the most intuitive and biologically meaningful modes of the line-to-point pair correlation function. We chose a bandwidth of $\text{epan.scale} = 4 \mu\text{m}$ for Epanechnikov kernel smoothing (Eq. 6.5). We compared the results to those of Chapter 4, where we calculated partial pair correlation functions of bacteria near groove intersection points as an approximation of the groove pattern. Additionally, we tested a recommendation of Stoyan and Ohser (1982) to approximate the groove pattern by a pattern of points randomly distributed *on* the line pattern and then calculate the partial pair correlation function for these two patterns. All pair correlation functions were tested for significance using 95% simulation envelopes from 199 simulations (Kenkel 1988). We used toroidal shifts (Dixon 2002) of the bacterial patterns, a null model which is well suited for testing for independence between two patterns (Wiegand and Moloney 2014).

Results

Ability of LPPCFs to correctly describe spatial correlations

We discuss the four methods collectively because the results were very similar for all four methods except for some deviations at very small spatial scale h (Figure 22, right column). Another exception is the *MEAN*-method where $g_{MD}(h)$ always approached 0 as scale h approaches 0. This behavior was expected because the mean distance between a straight line segment and a point sitting *on* the line segment lies between one quarter and one half of the length of the line segment. Consequently, only non-zero point-line distances exist with the MEAN methods and consequently $g_{MD}(h) = 0$ for $h \rightarrow 0$.

The point pattern in Figure 22A was constructed with the points aggregating closely around the line pattern. The aggregation was visible in the respective LPPCFs in Figure 22F up to a scale of about 0.05 units and, because of the regularity in the pattern, again at spatial scales of 0.2 and 0.4 units. The scarcity of points at scales in between indicates the empty areas between the lines, i.e. the centers of the 'plant epidermal cells' (Figure 22A and F).

In the second pattern (Figure 22B), points aggregated in a buffer zone at distance 0.03-0.05 units from the line pattern, indicated by the peaks in Figure 22G around 0.05 units. The expected scarcity at scale $h = 0.1$ units ('epidermal cell' centers) was not observed. At scales 0.16 – 0.28 units, we observed an alternating pattern of aggregation, scarcity and aggregation, which very well described the aggregation near and scarcity at the nearest parallel sine wave, i.e. the periodicity in the pattern. The fact that the scarcity was not highest at 0.2, as one would expect, can be accounted to the dominance of diagonal line-to-point distances, i.e. distances between a line segment to points non-perpendicular to the sine wave of the line segment and aggregating near a parallel sine wave. These distances will always usually be larger than 0.2 units. Even at a scale of about 0.4 units the periodicity of the pattern was visible (Figure 22G).

In the third pattern (Figure 22C), the points followed a random distribution that was not affected by the presence of the lines. As expected, $g_{lp}(h) \approx 1$ for all h for all four methods (Figure 22H). Again, the only exception was the MEAN method at very small scales.

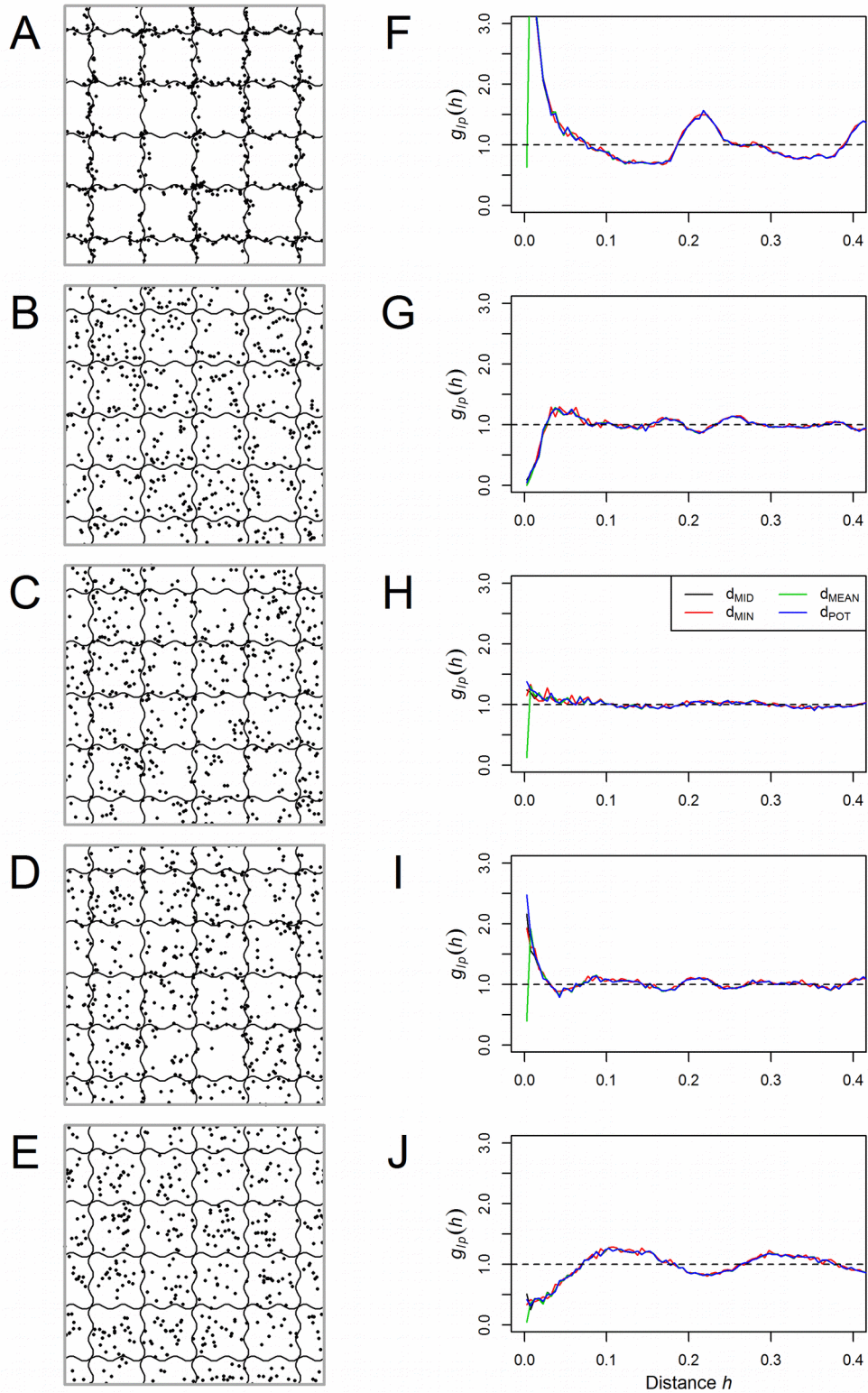


Figure 22: Artificial line patterns and point patterns in a unit square (left column). Points are (A) aggregated close to the lines, (B) aggregated at a certain distance away from the lines in the pattern, (C) randomly distributed, (D) scarce at distances 0.02-0.04 units from the line pattern, and (E) avoiding the line pattern. F-J: line-to-point pair correlation functions of the point patterns in A-E based on distances of the points to the mid-point of each line segment (d_{MID} , black), the shortest distance to each line segment (d_{MIN} , red), the average distance to every point on a line segment (d_{MEAN} , green), and distances based on the potential effect of each point of a line segment (d_{POT} , blue).

In pattern 4, the points were designed to avoid a buffer zone 0.02 – 0.04 units distant from the lines. To the human eye the scarcity of points in this buffer zone is not obvious. All four LPPCFs, nevertheless, suggest a scarcity of points at a spatial scale of about 0.04 units (Figure 22I). The LPPCFs, however, detected an aggregation of points at the lines. This is the consequence of how we constructed the pattern by thinning a CSR pattern only within the buffer zone. This reduced the overall point density of the pattern such that the local point densities in the buffer zone was decreased while at the same time it was increased outside the buffer zone when compared to the overall intensity. A close look at the LPPCFs between scales 0.0 and 0.2 units exactly reveals the designed point distribution with two local minima at the scales of the buffer zones.

Finally, for pattern 5, where the points actively avoided the areas close to the lines, all four LPPCFs capture this behavior very well and also unveiled the regularity in the pattern at scales that are manifolds of 0.2 units (Figure 22J).

Mode of segmentation

The inadequate approximation of the line pattern (Figure 23, left) caused the LPPCFs to overestimate the strength of point aggregation compared to the finer patterns, (Figure 24). This does not need to be a general bias but could be an effect of how the line pattern was changed by this particular (inadequate) approximation. The adequate line pattern (Figure 23, center) produced results very similar to the fine line pattern (Figure 23, right, and Figure 24). Deviations were small and probably the result of the 'fine' LPPCF being more smooth due to the larger number of line segments in this pattern.

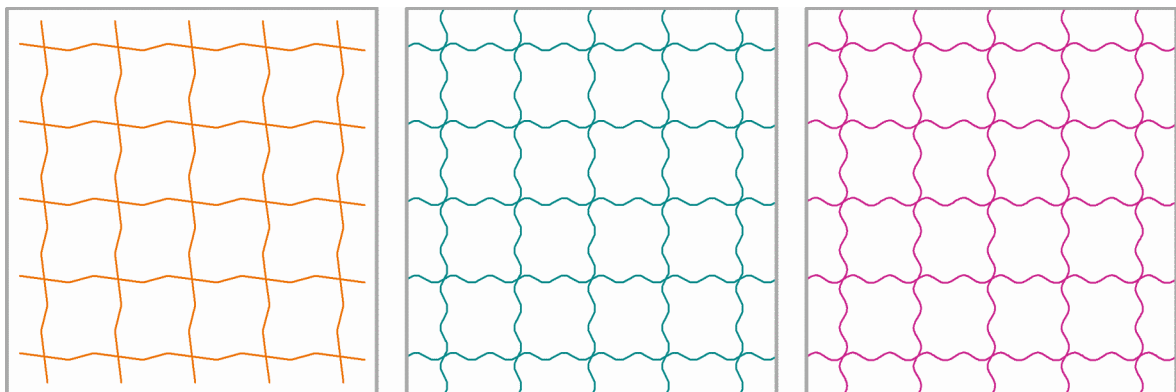


Figure 23: Different approximations of the line pattern in Figure 22. The approximations can be considered as 'inadequate' (left), 'adequate' (center) and 'fine' (right). The effect of the quality of these approximation was assessed with analyses presented in Figure 24.

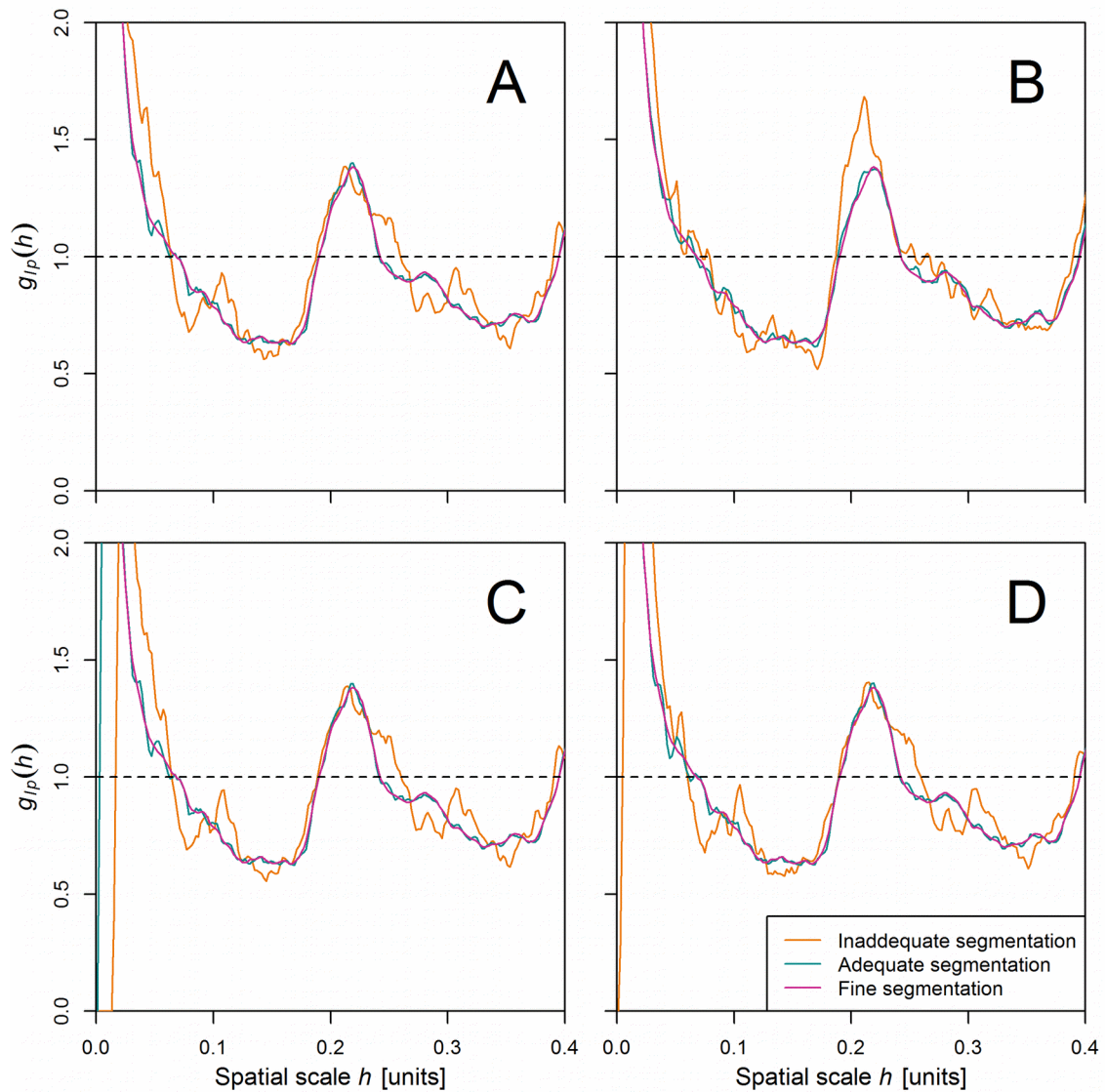


Figure 24: Analysis of the effect of inadequate line pattern approximation using inadequate, adequate, and fine segmentation of the line pattern (Figure 23). A: line-to-point pair correlation functions (LPPCFs) based on d_{MID} . B: LPPCFs based on d_{MIN} . C: LPPCFs based on d_{MEAN} . D: LPPCFs based on d_{POT} .

Bacteria on plant leaves

The performance of our LPPCFs based on the distances d_{MIN} and d_{POT} in comparison to the results from Chapter 4 and the method proposed by Stoyan & Ohser (1982) revealed the superior sensitivity of the LPPCFs over the other methods in detecting spatial correlations between the points in a point pattern and the lines in a line pattern (Figure 25). Whereas our approach in Chapter 4 found significant aggregation of bacteria near grooves only in 47.1% of our samples, the LPPCFs had detection rates of almost 80% (Figure 25, left). The method proposed by Stoyan and Ohser (1982) performed intermediate with a detection rate

of 57.5% (Figure 25, left). Similarly the detection rates for bacterial scarcity on the tops of leaf epidermal cells (scales 15-40 μm) were much higher for the LPPCFs (41.4% for g_{MIN} and 36.5% for g_{POT}) compared to the Stoyan and Ohser method (17.9%) and our results from Chapter 4 (9.6%).

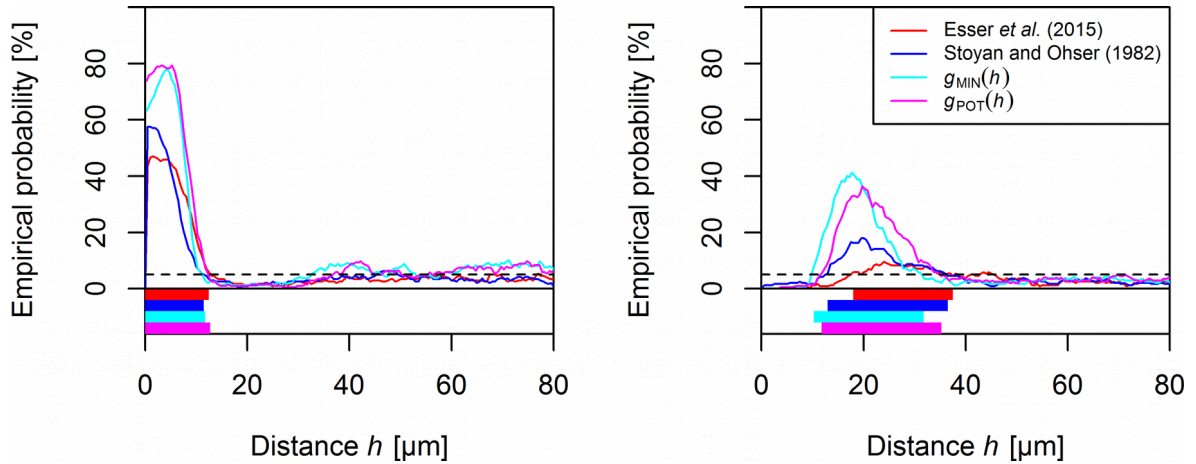


Figure 25: Detection rate of significant aggregation (left) and scarcity (right) of leaf-colonizing bacteria at distance h from grooves between leaf epidermal cells. Our new LPPCFs $g_{\text{MIN}}(h)$ and $g_{\text{POT}}(h)$ are compared to the results in Chapter 4 (Esser *et al.*, 2015) and a method proposed by Stoyan and Ohser (1982, cf. main text). The broken line is a threshold line for the quantum plots (colored bars) that outline the scales at which the 4 methods detected significant aggregation and scarcity in more than 5% of the samples.

Of particular interest are the quantum plots (colored bars) in Figure 25, which indicate at which scales the four methods found significant aggregation (or scarcity) in more than 5% of the samples. Obviously, there is much agreement between the four methods about the maximum spatial scale up to which grooves between leaf epidermal cells attract bacterial individuals. Thus, the major advantage of our LPPCFs is not so much an increased precision in the scales of spatial correlations between points and lines but in the sensitivity for detecting such correlations. Furthermore, the LPPCFs gave some indication of an aggregation of cells at distances larger than 60 μm which is in good agreement with the typical size of leaf epidermal cells. This suggests that LPPCFs are able to detect larger scale aggregation caused by a regularity in the line pattern – information that the two other approaches in Figure 25 did not provide and that can also not be detected by nearest-neighbor distance-based methods such as the J -function (van Lieshout and Baddeley 1999) or the Berman test (Berman 1986).

Discussion

In this paper, we presented an adaption of the partial pair correlation function that can be

applied to study spatial correlations between the points in a point pattern and linear spatial structures. What sets it apart from existing tools such as the J -function (van Lieshout and Baddeley 1999) or the Berman test (Berman 1986) is the fact that it is able to analyze spatial scales beyond the distance to the nearest point on the closest line segment. Also our estimators are inherently more stable than the estimators suggested for Stoyan & Ohser's k_{12} and c_{12} functions (Stoyan and Ohser 1982). We presented four estimators for our line-to-point pair correlation function (LPPCF) that only differed in the definition of the distance between a point and a line. Under homogeneous conditions, all four estimators gave almost identical results (Figure 22). In application, such high degrees of homogeneity cannot be guaranteed and the differences between the estimators become more visible (Figure 25). On the one hand this provides the opportunity to study various aspects of point-line correlations since the four estimators can reveal different information about the patterns. On the other hand it raises the question of how to deal with heterogeneity in line patterns.

Further studies that apply LPPCFs are needed to develop a more in-depth understanding of what the different distance modes tell us in ecological practice. Our function g_{MIN} is probably the most intuitive one because it is based on the distance from a point to the closest points on the line segments. This line-to-point distance is commonly used in practice, e.g. in geographical information systems. The estimation of the function g_{MID} is computationally the most efficient one because the midpoint of a segment is easier to determine than the point closest to a target point in the point pattern which is required for g_{MIN} . The estimation of the remaining functions g_{MEAN} and g_{POT} , while being computationally more demanding, has the advantage that both functions use much more spatial information contained in the line pattern than the other two functions. Although g_{MEAN} has the more intuitive definition compared to g_{POT} , given that d_{MEAN} is the average distance between a point of the point pattern and every point on a line segment, it will usually be biologically less meaningful. It assumes that every point on the line segment has an equally strong effect on the location of points in the point pattern, regardless of its distance to the points. The estimator thus gives overly much weight to distant points on the line segment and therefore d_{MEAN} takes values considerably larger than d_{MIN} . Probably the best estimator for the LPPCF in application is \hat{g}_{POT} . It uses all the spatial information contained in the line pattern but gives most weight to distances close to d_{MIN} . That means that it gives most weight to the points of the line segment that are closest to the point in the

point pattern. We expect these points of the line segment to have the greatest *effect* on the point pattern in most applications and therefore suggest that g_{POT} is the LPPCF that captures correlations between a point pattern and a linear structure the best.

Despite the promising results presented above, the issue of the mode of segmentation (Figures 23 and 24) has not been fully resolved yet. The more straight line segments are used to approximate a curve-like structure in reality, the more similar the four modes of LPPCFs will be. The resolution of the segmentation will be limited by the computational power available and thus is rather arbitrary. One step towards reducing this ambiguity could be to group line segments into well defined segmented curves and average or sum the line-to-point distances across this segmented curve. In our examples (plant leaf-colonizing bacteria) such a well defined segmented curves could be the strings of segments connecting the groove intersection points used in Chapter 4. Then not the distance between a bacterium and a groove segment would be calculated but the distance to the segmented curves. The formulation of such a curve-based line-to-point pair correlation function should be subject to future research.

In summary, we expect the line-to-point pair correlation functions to facilitate the study of spatial correlations between point-like and linear objects in ecology as well as in a wide variety of other fields. In particular, the LPPCFs based on d_{MIN} and d_{POT} will play a central role in such endeavors.

Chapter 7 – Consolidated discussion

The key goal of this dissertation was a quantitative, scale-specific description of bacterial colonization patterns on plant leaves directed at establishing an approach for untangling the multitude of concurrent interactions between bacteria and their environment. While a detailed identification and description of these interactions will require continuous future research efforts, the present study illustrates how spatial point pattern analysis can be used for this task and at the same time presents important fundamental principles of how bacteria interact with their phyllospheric environment.

Most importantly, we showed that the bacterial leaf colonizers interact with their environment at different spatial scales and were able for the first time to quantify these scales. Moreover, the bi-variate spatial point pattern analyses also allowed for the differentiation of concurrent interactions that act at similar spatial scales such as the aggregations of bacteria near glandular and hooked trichomes which both occurred at scales up to 100-120 μm . This showed that the bi-variate spatial point pattern analysis of large numbers of patterns allows for a differentiation of multiple concurrent effects. It is especially this last aspect that somehow challenges the ergodicity assumption that one pattern may suffice for an identification of all relevant processes (Illian *et al.* 2008, p. 39). The analyses in chapters 4, 5, and 6 revealed that there is a substantial probability to miss out on certain correlations that, based on the shape of our empirical probability plots, actually do exist. One prominent example for this is the aggregation of bacteria near glandular trichomes that was clearly identified at scales up to 100 μm , but only in about 16% of our samples that had glandular trichomes in them (Chapter 4). Two possible explanations for this low detection rate could be 1) the low number of glandular trichomes per sample (often only 1) and 2) the low probability that a bacterium lands within 100 μm around a glandular trichome during inoculation. Thus, the ergodicity assumption is of limited use in spatial point pattern analysis unless practical experience with the study system exists that may help to decide if the one pattern under study is of appropriate size. This is especially true for study systems, where multiple processes at different, also larger spatial scales affect the distribution of the points. In our study system (leaf-colonizing bacteria), we found processes operating at small scales (aggregation near grooves), intermediate scales (repulsion from stomata), and larger scales (aggregation near trichomes), such that it is unlikely to find all these interactions in one sample. Although

larger samples might meet the ergodicity assumption, such data sets are very difficult to acquire with fluorescence microscopy. Nevertheless the parallel analysis of large numbers of smaller samples proved to be a more than adequate alternative. A second alternative would have been to combine the replicate samples directly within the estimator of the pair correlation function such that one PCF is estimated across all samples (Wiegand and Moloney 2014, p. 246 ff.). Such practices have not yet been implemented in the *spatstat*-package used for our analyses and were therefore not applied in our studies.

With our high number of samples, the bi-variate analyses isolated single interactions from the point patterns which were created by multiple processes. This in turn has important implications for the uni-variate analysis of bacterial colonization patterns on leaves and by extension for any uni-variate point pattern analysis where multiple processes can be expected to have created the observed pattern. In our studies on leaves (Chapter 4), bacteria were found to be aggregated at scales up to 90 μm , but various processes were identified that could have caused this aggregation: The aggregation of cells in the grooves between epidermal cells, near both types of leaf trichomes, and near veins. But also other processes such as the clonal growth mechanism of bacteria could have added to the observed aggregation patterns. Thus, the uni-variate analysis is really just a description of the bacterial colonization patterns on leaves, whereas spatial correlations identified in the bi-variate analyses allow much more for an assessment of the involved processes.

Since a specific identified interaction such as an aggregation near grooves can still have multiple underlying mechanisms, we also studied bacterial colonization of artificial biomimetic bean leaf surfaces. We found that leaf physical properties such as topography and wettability were sufficient to explain the aggregation of bacteria near the grooves. Leaf biological processes such as cuticle permeability or local excretion of nutrients were thus excluded to be obligate processes. Again, the fundamental physical process that caused the aggregation near the grooves, e.g. a gravitational process or 'stochastic entrapment' (Chapter 5), remained unknown. Our procedures illustrate how spatial point pattern analyses combined with specialized null models and the simplification of environmental conditions (here the exclusion of leaf biological processes) can be applied in a multi-step approach to study the true drivers of bacterial distribution on plant surfaces. We see great potential for such practices for a scale-specific identification of bacterial interactions on plant surfaces especially with respect to the wealth of modern and upcoming molecular methods. For example, the hydrophobicity of artificial leaf surfaces can be controlled by

UV-light treatment as described in Chapter 5. This effect could be used in future studies of the effect of leaf surface hydrophobicity on both, changes in the waterscape (Doan and Leveau 2015), or the spatial distribution of bacterial colonizers. Both studies could be combined by a spatially explicit analysis of water availability based on the response of individual bacteria carrying a fluorescent bioreporter system that responds to water deprivation (Axtell and Beattie 2002).

Another important aspect of this dissertation project arose from the predominant aggregation of bacteria near the grooves between epidermal cells. These grooves are linear by nature. Although the bi-variate analysis of bacteria near groove intersection points (Chapter 4 and 5) produced well interpretable results, we expected to gain more detailed information by using the full information contained in the line patterns. As discussed in Chapter 6, there exists a variety of methods targeted at such point-line problems (Stoyan and Ohser 1982; Berman 1986; Foxall and Baddeley 2002), but these methods are either based on a reduction of information, e.g. representation of line pattern by a point pattern, or they are difficult to handle or even unstable, and did not find their way into broad application. The line-to-point pair correlation function (LPPCF) presented in Chapter 6 was motivated by the idea that the concept of pair correlation functions could be transferred to point-line problems. The implementation of our concept, however, was impeded by the question of what is the exact distance between a point and a line. In application, this issue was usually intuitively neglected and simply the shortest distance from the point to the line was used (Wang 2006). This is also true for the definition of distances between points and lines in the *spatstat* package for R that was used for most of our analyses, e.g. in the function *nncross*. In Chapter 6, we introduced more line-to-point distance measures, two of which are based on distance integrals which can use all information contained in the line pattern. Of course, these non-euclidian distances may be difficult to interpret. Nevertheless, especially the potential-based measure discussed in Chapter 6 captures the interactions between point- and line-like objects in a way that can be assumed to occur in many natural settings, i.e. closer parts of a linear structure interact more strongly with a point-like entity than more distant parts of the same linear structure. We hope for these distance measures to reignite discussions across many fields on how relationships between point-like and linear objects can be evaluated.

A comparison of the analyses of bacterial aggregation near grooves between epidermal leaf cells in Chapter 4 (bi-variate point pattern analysis using groove intersection points) and

Chapter 6 (potential based line-to-point pair correlation function) revealed that even for such complex linear structures, comparable results can be attained. The high quality line-data requirements of the LPPCF (fully connected linear network, no duplicated vertices, no zero-length line segments) are challenging to meet if a large number of line data sets is required, especially if automated picture analysis algorithms are used. Therefore, for many applications, the suggestion by Stoyan and Ohser (1982) to approximate line patterns by point patterns and use the appropriate bi-variate point process estimators should be sufficient. In cases where an appropriate approximation requires a large number of points compared to the number of line segments in the pattern, e.g. in line patterns consisting of a large number of long segments, the LPPCF might still confer a computational advantage over the approach of Stoyan and Ohser (1982). The biggest advantage of the LPPCF, however, is its superior sensitivity that may enable researchers to detect weak line-point correlations where other methods might fail. Furthermore, we expect that the abstract scale space of 'average potential-based distances' between points and lines has interesting properties that might be beneficial for the analysis and description of interactions between points and linear objects in a wide variety of scientific fields such as the cumulative effect of road noise on point-like objects in the landscape or diffusion-dependent concentration of chemicals in the vicinity of vascular systems in physiology. An in-depth description and analysis of this abstract scale space, however, was beyond the scope of this dissertation.

Limitations and future directions

This dissertation project provides of a scheme for exploring and discriminating the multitude of interactions that shape microbial life in the phyllosphere and that can easily be extended to other microbial and macrobiological environments. A thorough spatial point pattern analysis of bacterial colonization patterns with specific null models and the reduction of environmental complexity using artificial leaf surfaces are both methods that have not been established in phyllosphere ecology yet (but see Remus-Emsermann *et al.* 2014). The present work discussed fundamental principles of microbial colonization of this habitat such as the aggregation near the grooves between epidermal leaf cells. However, at this early point, it is challenging to transfer the results from our controlled laboratory environment to field conditions. Here, other abiotic and biotic drivers may play a much more important role than the interactions identified in our studies. Some of the most important deviations from a natural field system include a) the reduced microbial diversity in our experiments, i.e. the restriction to only two bacterial strains, b) the constantly high

humid conditions (~100%) and constant temperature during incubation, and c) the restriction of observing colonization only for three days after a single immigration event (spray inoculation) compared to the seasonal succession with continuous immigration and emigration events under field conditions.

The other major result of this dissertation project is the review and the extension of statistical methodology for studying spatial correlations between point-like and linear objects. Such interactions are omnipresent in natural environments but due to the complexity of their analysis are less often considered or they are reduced to the analysis of the shortest distance between a point and the nearest linear object. This dissertation discussed the available methods and proposed the line-to-point pair correlation function as an alternative method that can flexibly handle various point-to-line distance measures. The integral-based distance measure derived from potential influence fields is an especially interesting candidate that represents the distance of a line from a point as a function of distance-dependent influence of each point of the line piece on the point – a type of interaction that can be assumed for many natural processes. Nevertheless, the analysis of point-line interactions remains challenging and more effort needs to be spent to better understanding the limits and implications of such analyses. These future research efforts should concern both the properties and interpretation of the potential-based distance measure as well as the general question of inhomogeneous line patterns. Especially the last point might be challenging to tackle given the often arbitrary choice of parameters for estimating the local intensity $\lambda(x,y)$ of point and line processes and the fact that heterogeneity in general is not fully understood. For example, various classes of spatial heterogeneity might exist outside the realm of second-order intensity reweighed stationarity (Baddeley, Møller and Waagepetersen 2000), e.g. considering complex anisotropy. The aggregation of bacterial cells in the puzzle-piece shaped grooves between epidermal leaf cells forces the bacteria into a complex spatial distribution pattern which might be described as following a 'scale-dependent isotropy'. At larger scales, the aggregation of bacterial cells follows the homogeneous distribution of the linear network pattern of epidermal grooves. However, at smaller scales of tens of micrometers, the cells are forced into an anisotropic distribution following the meandering structure of the grooves. An investigation of such complex heterogeneity structures is recommended for future studies. Furthermore, these issues also translate into higher dimensional sets, e.g. the spatial correlations between point patterns and areas or between linear structures and

volumes, thereby pointing out the pivotal role of the earlier work by Stoyan and Ohser (1982). Clearly, such techniques could be applied to a wide range of spatial research questions also beyond the field of ecology.

In the field of microbial ecology, we expect our work to be only the starting point of future, more detailed spatial analyses of life in microbial habitats. Multiple extensions come to mind. Our research was limited to the early colonization of bean leaves by artificial communities of *Pantoea agglomerans* and *Pseudomonas syringae* under controlled laboratory conditions. The concepts established in our work should be verified under different conditions, e.g. under less humid or more variable conditions. Also more different microbial species and multiple host plants should be considered. The sum of such efforts will allow the identification of common themes of microbial colonization of leaves. The next step would be the analysis of microbial colonization of plants under field conditions. Remus-Emsermann et al. (2014) introduced a method to observe the distribution of various bacterial taxonomic groups under field conditions. Their approach applied fluorescent *in situ* hybridization (FISH) techniques (Amann and Fuchs 2008) to analyze highly diverse bacterial communities that developed over time under variable conditions with naturally occurring succession and migration events. Similar to our studies, they observed intra- and interspecific aggregation of bacteria on these leaves, but the spatial processes that caused this aggregation remained unknown. We expect fluorescent identification tools such as FISH and the use of bioreporters (Joyner and Lindow 2000; Leveau and Lindow 2001; Axtell and Beattie 2002; Remus-Emsermann and Leveau 2009; Remus-Emsermann *et al.* 2012), when combined with spatial marked point pattern analysis, to vastly improve our understanding of microbial diversity, interactions, and functions on plant surfaces. In particular, bioreporters, i.e. bacteria that were genetically modified to report the presence and relative concentration of various substances such as nutrients, allow the introduction of geostatistical methods such as kriging into phyllosphere research. The results of such studies could then be further processed in marked point pattern analyses involving such methods as the mark-correlation function (Stoyan and Stoyan 1994) or Baddeley's ρ -function (Baddeley *et al.* 2012), both of which may be offering deeper insights into spatial ecological processes and both of which may be more sensitive to certain processes than the pair correlation function applied in our studies. Clearly, we just only turned page one of the spatial analysis of microbial interactions in the phyllosphere.

References

- Al Ashhab A, Herzberg M, Gillor O. Biofouling of reverse-osmosis membranes during tertiary wastewater desalination: microbial community composition. *Water Res* 2014;**50**:341–9.
- Amann R, Fuchs BM. Single-cell identification in microbial communities by improved fluorescence *in situ* hybridization techniques. *Nat Rev Microbiol* 2008;**6**:339–48.
- Andrews JH, Harris RF. The ecology and biogeography of microorganisms of plant surfaces. *Annu Rev Phytopathol* 2000;**38**:145–80.
- Ascensão L, Pais MS. The leaf capitate trichomes of *Leonotis leonurus*: histochemistry, ultrastructure and secretion. *Ann Bot* 1998;**81**:263–71.
- Atamna-Ismaeel N, Finkel OM, Glaser F *et al.* Microbial rhodopsins on leaf surfaces of terrestrial plants. *Environ Microbiol* 2012;**14**:140–6.
- Augspurger C, Karwautz C, Mussmann M *et al.* Drivers of bacterial colonization patterns in stream biofilms. *FEMS Microbiol Ecol* 2010;**72**:47–57.
- Axtell CA, Beattie GA. Construction and characterization of a *proU-gfp* transcriptional fusion that measures water availability in a microbial habitat. *Appl Environ Microbiol* 2002;**68**:4604–12.
- Baddeley A, Chang Y-M, Song Y *et al.* Nonparametric estimation of the dependence of a spatial point process on spatial covariates. *Stat Interface* 2012;**5**:221–36.
- Baddeley A, Diggle PJ, Hardegen A *et al.* On tests of spatial pattern based on simulation envelopes. *Ecol Monogr* 2014;**84**:477–89.
- Baddeley AJ, Turner R. spatstat: an R package for analyzing spatial point patterns. *J Stat Softw* 2005;**12**:1–42.
- Baddeley A, Møller J, Waagepetersen R. Non- and semi-parametric estimation of interaction in inhomogeneous point patterns. *Stat Neerlandica* 2000;**54**:329–50.
- Bailey MJ, Lilley AK, Diaper JP. Gene transfer between microorganisms in the phyllosphere. In: Morris CE, Nicot PC, Nguyen-The C (eds.). *Aerial Plant Surface Microbiology*. New York, NY: Plenum Press, 1996, 103–23.
- Barnes G, Neve NFB. Examination of plant surface microflora by the scanning electron microscope. *Trans Br Mycol Soc* 1968;**51**:811IN17–812.
- Beattie GA. Water relations in the interaction of foliar bacterial pathogens with plants. In: VanAlfen NK, Bruening G, Leach JE (eds.). *Annual Review of Phytopathology*. Vol 49. Palo Alto: Annual Reviews, 2011, 533–55.
- Beattie GA, Lindow SE. Bacterial colonization of leaves: A spectrum of strategies. *Phytopathology* 1999;**89**:353–9.

- Bedford T, van den Berg J. A remark on the van Lieshout and Baddeley J-Function for point processes. *Adv Appl Probab* 1997;**29**:19–25.
- Berg G, Grube M, Schlöter M *et al.* Unraveling the plant microbiome: looking back and future perspectives. *Front Microbiol* 2014;**5**, DOI: 10.3389/fmicb.2014.00148.
- Berman M. Testing for spatial association between a point process and another stochastic process. *J R Stat Soc Ser C Appl Stat* 1986;**35**:54–62.
- Besag JE. Contribution to the discussion of Dr. Ripley's paper. *J R Stat Soc Ser B Methodol* 1977;**39**:193–5.
- Birkhofer K, Schoening I, Alt F *et al.* General relationships between abiotic soil properties and soil biota across spatial scales and different land-use types. *Plos One* 2012;**7**, DOI: 10.1371/journal.pone.0043292.
- Blakeman JP. Ecological succession of leaf surface microorganisms in relation to biological control. In: Windel CE, Lindow SE (eds.). *Biological Control on the Phylloplane*. The American Phytopathological Society, 1985, 6–30.
- Blöchl E, Rachel R, Burggraf S *et al.* *Pyrolobus fumarii*, gen. and sp. nov., represents a novel group of archaea, extending the upper temperature limit for life to 113°C. *Extremophiles* 1997;**1**:14–21.
- Brandl MT, Lindow SE. Contribution of Indole-3-Acetic Acid production to the epiphytic fitness of *Erwinia herbicola*. *Appl Environ Microbiol* 1998;**64**:3256–63.
- Brewer CA, Smith WK, Vogelmann TC. Functional interaction between leaf trichomes, leaf wettability and the optical properties of water droplets. *Plant Cell Environ* 1991;**14**:955–62.
- Bringel F, Couée I. Pivotal roles of phyllosphere microorganisms at the interface between plant functioning and atmospheric trace gas dynamics. *Terr Microbiol* 2015:486.
- Brown PE, Christensen OF, Clough HE *et al.* Frequency and spatial distribution of environmental *Campylobacter spp.* *Appl Environ Microbiol* 2004;**70**:6501–11.
- Burkhardt J, Kaiser H, Goldbach H *et al.* Measurements of electrical leaf surface conductance reveal re-condensation of transpired water vapour on leaf surfaces. *Plant Cell Environ* 1999;**22**:189–96.
- Canny MJ. Fine veins of dicotyledon leaves as sites for enrichment of solutes of the xylem sap. *New Phytol* 1990;**115**:511–6.
- Cerqueira L, Oliveira JA, Nicolau A *et al.* Biofilm formation with mixed cultures of *Pseudomonas aeruginosa/Escherichia coli* on silicone using artificial urine to mimic urinary catheters. *Biofouling* 2013;**29**:829–40.
- Chalfie M, Tu Y, Euskirchen G *et al.* Green fluorescent protein as a marker for gene expression. *Science* 1994;**263**:802–5.
- Chiu SN, Stoyan D, Kendall WS *et al.* *Stochastic Geometry and Its Applications*. 3.

- edition. Chichester, United Kingdom: John Wiley & Sons, 2013.
- Clark PJ, Evans FC. Distance to nearest neighbor as a measure of spatial relationships in populations. *Ecology* 1954;**35**:445–53.
- Cole SJ, Records AR, Orr MW *et al.* Catheter-associated urinary tract infection by *Pseudomonas aeruginosa* is mediated by exopolysaccharide-independent biofilms. *Infect Immun* 2014;**82**:2048–58.
- Conrath U, Pieterse CMJ, Mauch-Mani B. Priming in plant–pathogen interactions. *Trends Plant Sci* 2002;**7**:210–6.
- Copeland JK, Yuan L, Layeghifard M *et al.* Seasonal community succession of the phyllosphere microbiome. *Mol Plant Microbe Interact* 2015;**28**:274–85.
- Costerton JW, Lewandowski Z, Caldwell DE *et al.* Microbial Biofilms. *Annu Rev Microbiol* 1995;**49**:711–45.
- Dandurand LM, Schotzko DJ, Knudsen GR. Spatial patterns of rhizoplane populations of *Pseudomonas fluorescens*. *Appl Environ Microbiol* 1997;**63**:3211–7.
- Dangl JL, Jones JDG. Plant pathogens and integrated defence responses to infection. *Nature* 2001;**411**:826–33.
- Danhorn T, Fuqua C. Biofilm formation by plant-associated bacteria. *Annu Rev Microbiol* 2007;**61**:401–22.
- Davis CL, Brlansky RH. Use of immunogold labelling with scanning electron microscopy to identify phytopathogenic bacteria on leaf surfaces. *Appl Environ Microbiol* 1991;**57**:3052–5.
- Dechesne A, Wang G, Gulez G *et al.* Hydration-controlled bacterial motility and dispersal on surfaces. *Proc Natl Acad Sci* 2010;**107**:14369–72.
- Deng X, Zhuang G, Ma A *et al.* Construction of a dual fluorescence whole-cell biosensor to detect N-acyl homoserine lactones. *J Environ Sci-China* 2014;**26**:415–22.
- Depan D, Misra RDK. On the determining role of network structure titania in silicone against bacterial colonization: mechanism and disruption of biofilm. *Mater Sci Eng C-Mater Biol Appl* 2014;**34**:221–8.
- Diem HG. Micro-organisms of the leaf surface: estimation of the mycoflora of the barley phyllosphere. *J Gen Microbiol* 1974;**80**:77–83.
- Diggle PJ. *Statistical Analysis of Spatial Point Patterns*. London [u.a.]: Academic Pr., 1983.
- Diggle PJ. *Statistical Analysis of Spatial and Spatio-Temporal Point Patterns*. 3rd edition. Chapman and Hall/CRC, 2013.
- Dixon PM. Ripley's K function. *Encyclopedia of Environmetrics*. Vol 3. John Wiley & Sons, Ltd, 2002, 1796–803.

- Doan H, Leveau J. Artificial surfaces in phyllosphere microbiology. *Phytopathology* 2015, DOI: 10.1094/PHYTO-02-15-0050-RVW.
- Dunne WM. Bacterial adhesion: Seen any good biofilms lately? *Clin Microbiol Rev* 2002;**15**:155 – +.
- Dusenberry D. *Living at Micro Scale*. Princeton University Press, 2008.
- Espinosa-Urgel M. Plant-associated *Pseudomonas* populations: molecular biology, DNA dynamics, and gene transfer. *Plasmid* 2004;**52**:139–50.
- Esser DS, Leveau JHJ, Meyer KM *et al*. Spatial scales of interactions among bacteria and between bacteria and the leaf surface. *FEMS Microbiol Ecol* 2015;**91**:fiu034.
- Fang J, Dorrestein PC. Emerging mass spectrometry techniques for the direct analysis of microbial colonies. *Curr Opin Microbiol* 2014;**19**:120–9.
- Fedriani JM, Wiegand T, Delibes M. Spatial pattern of adult trees and the mammal-generated seed rain in the Iberian pear. *Ecography* 2010;**33**:545–55.
- Fett WF, Cooke PH. Scanning electron microscopy of native biofilms on mung bean sprouts. *Can J Microbiol* 2003;**49**:45–50.
- Fiala V, Glad C, Martin M *et al*. Occurrence of soluble carbohydrates on the phylloplane of maize (*Zea mays* L.): variations in relation to leaf heterogeneity and position on the plant. *New Phytol* 1990;**115**:609–15.
- Foxall R, Baddeley A. Nonparametric measures of association between a spatial point process and a random set, with geological applications. *J R Stat Soc Ser C-Appl Stat* 2002;**51**:165–82.
- Franklin J, Santos EV. A spatially explicit census reveals population structure and recruitment patterns for a narrowly endemic pine, *Pinus torreyana*. *Plant Ecol* 2011;**212**:293–306.
- Franklin RB, Blum LK, McComb AC *et al*. A geostatistical analysis of small-scale spatial variability in bacterial abundance and community structure in salt marsh creek bank sediments. *FEMS Microbiol Ecol* 2002;**42**:71–80.
- Franklin RB, Mills AL. Multi-scale variation in spatial heterogeneity for microbial community structure in an eastern Virginia agricultural field. *FEMS Microbiol Ecol* 2003;**44**:335–46.
- Franklin R, Mills A eds. *The Spatial Distribution of Microbes in the Environment*. 1st edition. Dordrecht, the Netherlands: Springer, 2007.
- Gantner S, Schmid M, Dürr C *et al*. *In situ* quantitation of the spatial scale of calling distances and population density-independent *N*-acylhomoserine lactone-mediated communication by rhizobacteria colonized on plant roots. *FEMS Microbiol Ecol* 2006;**56**:188–94.
- Garbeva P, Tyc O, Remus-Emsermann MNP *et al*. No apparent costs for facultative

- antibiotic production by the soil bacterium *Pseudomonas fluorescens* Pf0-1. *PLoS ONE* 2011;**6**:e27266.
- Gosme M, Lucas P. Disease spread across multiple scales in a spatial hierarchy: effect of host spatial structure and of inoculum quantity and distribution. *Phytopathology* 2009;**99**:833–9.
- Gradstein IS, Ryshik IW. *Summen, Produkt- und Integraltafeln / Tables of series, products, and integrals*. 5. ed. Thun: Harri Deutsch, 1981.
- Griffin EA, Carson WP. The Ecology and Natural History of Foliar Bacteria with a Focus on Tropical Forests and Agroecosystems. *Bot Rev* 2015:1–45.
- Hanisch K-H. Some remarks on estimators of the distribution function of nearest neighbour distance in stationary spatial point processes. *Ser Stat* 1984;**15**:409–12.
- Heinrich L, Schwandtke A. Some remarks on second-order-analysis of planar fibre processes by line intersections. *Math Nachrichten* 1990;**146**:247–57.
- Hirano SS, Nordheim EV, Arny DC *et al*. Lognormal distribution of epiphytic bacterial populations on leaf surfaces. *Appl Environ Microbiol* 1982;**44**:695–700.
- Hirano SS, Upper CD. Bacteria in the leaf ecosystem with emphasis on *Pseudomonas syringae* - a pathogen, ice nucleus, and epiphyte. *Microbiol Mol Biol Rev* 2000;**64**:624–53.
- Hong HG, Lee HJ, Bae JY *et al*. Spatial and temporal distribution of a biocontrol bacterium *Bacillus licheniformis* N1 on the strawberry plants. *Plant Pathol J* 2010;**26**:238–44.
- Hosni T, Moretti C, Devescovi G *et al*. Sharing of quorum-sensing signals and role of interspecies communities in a bacterial plant disease. *ISME J* 2011;**5**:1857–70.
- Hunter PJ, Hand P, Pink D *et al*. Both leaf properties and microbe-microbe interactions influence within-species variation in bacterial population diversity and structure in the lettuce (*Lactuca* species) phyllosphere. *Appl Env Microbiol* 2010;**76**:8117–25.
- Illian J, Penttinen A, Stoyan H *et al*. *Statistical Analysis and Modelling of Spatial Point Patterns: From Spatial Data to Knowledge*. 1st Edition. Chichester, United Kingdom: John Wiley & Sons, 2008.
- Ionescu A, Wutscher E, Brambilla E *et al*. Influence of surface properties of resin-based composites on *in vitro* *Streptococcus mutans* biofilm development. *Eur J Oral Sci* 2012;**120**:458–65.
- Jacobs JL, Carroll TL, Sundin GW. The role of pigmentation, ultraviolet radiation tolerance, and leaf colonization strategies in the epiphytic survival of phyllosphere bacteria. *Microb Ecol* 2004;**49**:104–13.
- Jacquemyn H, Wiegand T, Vandepitte K *et al*. Multigenerational analysis of spatial structure in the terrestrial, food-deceptive orchid *Orchis mascula*. *J Ecol* 2009;**97**:206–16.

- Jonsson ME, Norstrom M, Sandberg M *et al.* Space-time patterns of *Campylobacter spp.* colonization in broiler flocks, 2002-2006. *Epidemiol Infect* 2010;**138**:1336–45.
- Joyner DC, Lindow SE. Heterogeneity of iron bioavailability on plants assessed with a whole-cell GFP-based bacterial biosensor. *Microbiology* 2000;**146**:2435–45.
- Katsikogianni M, Missirlis YF. Concise review of mechanisms of bacterial adhesion to biomaterials and of techniques used in estimating bacteria-material interactions. *Eur Cell Mater* 2004;**8**:37–57.
- Kembel SW, O'Connor TK, Arnold HK *et al.* Relationships between phyllosphere bacterial communities and plant functional traits in a neotropical forest. *Proc Natl Acad Sci* 2014;**111**:13715–20.
- Kempf H-J. *Erwinia herbicola* as a biocontrol agent of *Fusarium culmorum* and *Puccinia recondita* f. sp. *tritici* on wheat. *Phytopathology* 1989;**79**:990.
- Kenkel NC. Pattern of self-thinning in jack-pines: testing the random mortality hypothesis. *Ecology* 1988;**69**:1017–24.
- Kim YC, Leveau JHJ, Gardener BBM *et al.* The multifactorial basis for plant health promotion by plant-associated bacteria. *Appl Environ Microbiol* 2011;**77**:1548–55.
- Kinkel LL. Microbial population dynamics on leaves. *Annu Rev Phytopathol* 1997;**35**:327–47.
- Kinkel LL, Newton MR, Leonard KJ. Resource aggregation in the phyllosphere: implications for microbial dynamics across spatial scales. In: Lindow SE, Hecht-Poinar EI, Elliott VJ (eds.). *Phyllosphere Microbiology*. American Phytopathological Society, 2002, 317–40.
- Knief C, Ramette A, Frances L *et al.* Site and plant species are important determinants of the *Methylobacterium* community composition in the plant phyllosphere. *ISME J* 2010;**4**:719–28.
- Koch K, Dommissé A, Barthlott W *et al.* The use of plant waxes as templates for micro- and nanopatterning of surfaces. *Acta Biomater* 2007;**3**:905–9.
- Krasnoperov RA, Stoyan D. Second-order stereology of spatial fibre systems. *J Microsc* 2004;**216**:156–64.
- Kreft J-U, Plugge CM, Grimm V *et al.* Mighty small: Observing and modeling individual microbes becomes big science. *Proc Natl Acad Sci* 2013;**110**:18027–8.
- Last FT. Seasonal incidence of *Sporobolomyces* on cereal leaves. *Trans Br Mycol Soc* 1955;**38**:221–39.
- Law R, Illian J, Burslem DFRP *et al.* Ecological information from spatial patterns of plants: insights from point process theory. *J Ecol* 2009;**97**:616–28.
- Leben C. Epiphytic microorganisms in relation to plant disease. *Annu Rev Phytopathol* 1965;**3**:209–30.

- Leben C. Colonization of soybean buds by bacteria: observations with the scanning electron microscope. *Can J Microbiol* 1969;**15**:319–20.
- Leben C. Colonization and movement of *Pseudomonas syringae* on healthy bean seedlings. *Phytopathology* 1970;**60**:677.
- Leveau JHJ. Microbial communities in the phyllosphere. In: Riederer M, Müller C (eds.). *Biology of the Plant Cuticle*. John Wiley & Sons, 2006, 334–67.
- Leveau JHJ, Lindow SE. Appetite of an epiphyte: quantitative monitoring of bacterial sugar consumption in the phyllosphere. *Proc Natl Acad Sci* 2001;**98**:3446–53.
- Lichtman JW, Conchello J-A. Fluorescence microscopy. *Nat Methods* 2005;**2**:910–9.
- van Lieshout MNM, Baddeley AJ. A nonparametric measure of spatial interaction in point patterns. *Stat Neerlandica* 1996;**50**:344–61.
- van Lieshout MNM, Baddeley AJ. Indices of dependence between types in multivariate point patterns. *Scand J Stat* 1999;**26**:511–32.
- Lindow SE, Brandl MT. Microbiology of the phyllosphere. *Appl Env Microbiol* 2003;**69**:1875–83.
- Lin H, Shin S, Blaya JA *et al.* Assessing spatiotemporal patterns of multidrug-resistant and drug-sensitive tuberculosis in a South American setting. *Epidemiol Infect* 2011;**139**:1784–93.
- Liu X, Germaine KJ, Ryan D *et al.* Whole-Cell Fluorescent Biosensors for Bioavailability and Biodegradation of Polychlorinated Biphenyls. *Sensors* 2010;**10**:1377–98.
- Liu Y, Rosenfield E, Hu M *et al.* Direct observation of bacterial deposition on and detachment from nanocomposite membranes embedded with silver nanoparticles. *Water Res* 2013;**47**:2949–58.
- Loosmore NB, Ford ED. Statistical inference using the G or K point pattern spatial statistics. *Ecology* 2006;**87**:1925–31.
- Lotwick H, Silverman B. Methods for analyzing spatial processes of several types of points. *J R Stat Soc Ser B-Methodol* 1982;**44**:406–13.
- Lv D, Ma A, Tang X *et al.* Profile of the culturable microbiome capable of producing acyl-homoserine lactone in the tobacco phyllosphere. *J Environ Sci* 2013;**25**:357–66.
- Maignien L, DeForce EA, Chafee ME *et al.* Ecological Succession and Stochastic Variation in the Assembly of *Arabidopsis thaliana* Phyllosphere Communities. *mBio* 2014;**5**:e00682–13.
- Manabe K, Nishizawa S, Shiratori S. Porous surface structure fabricated by breath figures that suppresses *Pseudomonas aeruginosa* biofilm formation. *ACS Appl Mater Interfaces* 2013;**5**:11900–5.
- Mansvelt EL, Hattingh MJ. Scanning electron microscopy of colonization of pear leaves

- by *Pseudomonas syringae* pv. *syringae*. *Can J Bot* 1987;**65**:2517–22.
- Mansvelt EL, Hattingh MJ. Scanning electron microscopy of invasion of apple leaves and blossoms by *Pseudomonas syringae* pv. *syringae*. *Appl Environ Microbiol* 1989;**55**:533–8.
- Marcell LM, Beattie GA. Effect of leaf surface waxes on leaf colonization by *Pantoea agglomerans* and *Clavibacter michiganensis*. *Mol Plant Microbe Interact* 2002;**15**:1236–44.
- Marchi G, Sisto A, Cimmino A *et al.* Interaction between *Pseudomonas savastanoi* pv. *savastanoi* and *Pantoea agglomerans* in olive knots. *Plant Pathol* 2006;**55**:614–24.
- Masák J, Čejková A, Schreiberová O *et al.* *Pseudomonas* biofilms: possibilities of their control. *FEMS Microbiol Ecol* 2014;**89**:1–14.
- Mattick JS. Type IV pili and twitching motility. *Annu Rev Microbiol* 2002;**56**:289–314.
- Mecke J. Formulas for stationary planar fibre processes III – intersections with fibre systems. *Ser Stat* 1981;**12**:201–10.
- Mecke J, Stoyan D. Formulas for stationary planar fibre processes I – general theory. *Ser Stat* 1980;**11**:267–79.
- Mercier J, Lindow SE. Role of leaf surface sugars in colonization of plants by bacterial epiphytes. *Appl Environ Microbiol* 2000;**66**:369–74.
- Mew TW, Vera Cruz CM. Epiphytic colonization of host and non-host plants by phytopathogenic bacteria. In: Fokkema NJ, van den Heuvel J (eds.). *Microbiology of the Phyllosphere*. New York: Cambridge University Press, 1986, 269–82.
- Meyer KM, Leveau JHJ. Microbiology of the phyllosphere: a playground for testing ecological concepts. *Oecologia* 2011;**168**:621–9.
- Miles WG, Daines RH, Rue JW. Presymptomatic egress of *Xanthomonas pruni* from infected peach leaves. *Phytopathology* 1977;**77**:895.
- Miller BP, Perry GLW, Enright NJ *et al.* Contrasting spatial pattern and pattern-forming processes in natural vs. restored shrublands. *J Appl Ecol* 2010;**47**:701–9.
- Mitchell EG, Kenchington CG, Liu AG *et al.* Reconstructing the reproductive mode of an Ediacaran macro-organism. *Nature* 2015;**524**:343–6.
- Monier J-M, Lindow SE. Frequency, size, and localization of bacterial aggregates on bean leaf surfaces. *Appl Environ Microbiol* 2004;**70**:346–55.
- Monier J-M, Lindow SE. Spatial organization of dual-species bacterial aggregates on leaf surfaces. *Appl Environ Microbiol* 2005;**71**:5484–93.
- Morris CE, Monier J-M. The ecological significance of biofilm formation by plant-associated bacteria. *Annu Rev Phytopathol* 2003;**41**:429–53.

- Morris CE, Monier J-M, Jacques M-A. A technique to quantify the population size and composition of the biofilm component in communities of bacteria in the phyllosphere. *Appl Environ Microbiol* 1998;**64**:4789–95.
- Murphy SJ, McCarthy BC. Evidence for topographic control of tree spatial patterning in an old-growth, mixed mesophytic forest in Southeastern Ohio, USA. *J Torrey Bot Soc* 2012;**139**:181–93.
- Nies DH. Heavy metal-resistant bacteria as extremophiles: molecular physiology and biotechnological use of *Ralstonia* sp. CH34. *Extremophiles* 2000;**4**:77–82.
- Nothdurft A, Saborowski J, Nuske RS *et al.* Density estimation based on k-tree sampling and point pattern reconstruction. *Can J For Res* 2010;**40**:953–67.
- Nunan N, Wu K, Young IM *et al.* *In situ* spatial patterns of soil bacterial populations, mapped at multiple scales, in an arable soil. *Microb Ecol* 2002;**44**:296–305.
- Ohser J, Mücklich F. *Statistical Analysis of Microstructures in Materials Science*. Auflage: 1. Auflage. Chichester England ; New York: John Wiley & Sons, 2000.
- Okabe A, Fujii A. The statistical analysis through a computational method of a distribution of points in relation to its surrounding network. *Environ Plan A* 1984;**16**:107–14.
- O’Rorke R, Cobian GM, Holland BS *et al.* Dining local: the microbial diet of a snail that grazes microbial communities is geographically structured. *Environ Microbiol* 2015;**17**:1753–64.
- O’Toole G, Kaplan HB, Kolter R. Biofilm formation as microbial development. *Annu Rev Microbiol* 2000;**54**:49–79.
- Otto M. Physical stress and bacterial colonization. *FEMS Microbiol Rev* 2014;**38**:1250–70.
- Parangan-Smith A, Lindow S. Contribution of nitrate assimilation to the fitness of *Pseudomonas syringae* pv. *syringae* B728a on plants. *Appl Environ Microbiol* 2013;**79**:678–87.
- Peñuelas J, Rico L, Ogaya R *et al.* Summer season and long-term drought increase the richness of bacteria and fungi in the foliar phyllosphere of *Quercus ilex* in a mixed Mediterranean forest. *Plant Biol* 2012;**14**:565–75.
- Pérez-Velázquez J, Schlicht R, Dulla G *et al.* Stochastic modeling of *Pseudomonas syringae* growth in the phyllosphere. *Math Biosci* 2012;**239**:106–16.
- Petrova OE, Sauer K. Sticky situations: key components that control bacterial surface attachment. *J Bacteriol* 2012;**194**:2413–25.
- Pillay T, Ward D. Spatial pattern analysis and competition between *Acacia* karroo trees in humid savannas. *Plant Ecol* 2012;**213**:1609–19.
- Ponomarova O, Patil KR. Metabolic interactions in microbial communities: untangling the Gordian knot. *Curr Opin Microbiol* 2015;**27**:37–44.

- Pontiroli A, Rizzi A, Simonet P *et al.* Visual evidence of horizontal gene transfer between plants and bacteria in the phytosphere of transplastomic tobacco. *Appl Environ Microbiol* 2009;**75**:3314–22.
- Rai N, Rai R, Venkatesh KV. Quorum Sensing Biosensors. In: Kalia VC (ed.). *Quorum Sensing vs Quorum Quenching: A Battle with No End in Sight*. Springer India, 2015, 173–83.
- Ranjard L, Dequiedt S, Chemidlin Prevost-Boure N *et al.* Turnover of soil bacterial diversity driven by wide-scale environmental heterogeneity. *Nat Commun* 2013;**4**, DOI: 10.1038/ncomms2431.
- Rao S, Kitron U, Weigel RM. Spatial and genotypic clustering of *Salmonella* over time in a swine production unit. *Prev Vet Med* 2010;**97**:90–9.
- Rastogi G, Sbodio A, Tech JJ *et al.* Leaf microbiota in an agroecosystem: spatiotemporal variation in bacterial community composition on field-grown lettuce. *ISME J* 2012;**6**:1812–22.
- Rastogi G, Tech JJ, Coaker GL *et al.* A PCR-based toolbox for the culture-independent quantification of total bacterial abundances in plant environments. *J Microbiol Methods* 2010;**83**:127–32.
- Raynaud X, Nunan N. Spatial ecology of bacteria at the microscale in soil. *PLoS ONE* 2014;**9**:e87217.
- R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2013.
- Redford AJ, Bowers RM, Knight R *et al.* The ecology of the phyllosphere: geographic and phylogenetic variability in the distribution of bacteria on tree leaves. *Environ Microbiol* 2010;**12**:2885–93.
- Redford AJ, Fierer N. Bacterial succession on the leaf surface: a novel system for studying successional dynamics. *Microb Ecol* 2009;**58**:189–98.
- Remus-Emsermann MNP, Kowalchuk GA, Leveau JHJ. Single-cell versus population-level reproductive success of bacterial immigrants to pre-colonized leaf surfaces. *Environ Microbiol Rep* 2013;**5**:387–92.
- Remus-Emsermann MNP, Leveau JHJ. Linking environmental heterogeneity and reproductive success at single-cell resolution. *ISME J* 2009;**4**:215–22.
- Remus-Emsermann MNP, Lückner S, Müller DB *et al.* Spatial distribution analyses of natural phyllosphere-colonizing bacteria on *Arabidopsis thaliana* revealed by fluorescence *in situ* hybridization. *Environ Microbiol* 2014;**16**:2329–40.
- Remus-Emsermann MNP, Tecon R, Kowalchuk GA *et al.* Variation in local carrying capacity and the individual fate of bacterial colonizers in the phyllosphere. *ISME J* 2012;**6**:756–65.
- Riederer M, Müller C. Introduction: biology of the plant cuticle. In: Riederer M, Müller C

- (eds.). *Biology of the Plant Cuticle*. Oxford, UK: Blackwell Publishing, 2006, 1–10.
- Ripley BD. 2nd-order analysis of stationary point processes. *J Appl Probab* 1976;**13**:255–66.
- Ripley BD. Modelling spatial patterns. *J R Stat Soc Ser B Methodol* 1977;**39**:172–212.
- Rodenacker K, Brühl A, Hausner M *et al*. Quantification of biofilms in multi-spectral digital volumes from confocal laser-scanning microscopy. *Image Anal Stereol* 2000;**19**:151–6.
- Roos IMM, Hattingh MJ. Scanning electron microscopy of *Pseudomonas syringae* pv. *morsprunorum* on sweet cherry leaves. *J Phytopathol* 1983;**108**:18–25.
- Ruinen J. Occurrence of *Beijerinckia* species in the phyllosphere. *Nature* 1956;**177**:220–1.
- Salta M, Wharton JA, Blache Y *et al*. Marine biofilms on artificial surfaces: structure and dynamics. *Environ Microbiol* 2013;**15**:2879–93.
- Salta M, Wharton JA, Stoodley P *et al*. Designing biomimetic antifouling surfaces. *Philos Trans R Soc Lond Math Phys Eng Sci* 2010;**368**:4729–54.
- Sambrook J, Fritsch EF, Maniatis T. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 2001.
- Schlaeppli K, Bulgarelli D. The plant microbiome at work. *Mol Plant Microbe Interact* 2014;**28**:212–7.
- Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 2012;**9**:671–5.
- Schönherr J, Baur P. Cuticle permeability studies. In: Morris CE, Nicot PC, Nguyen-The C (eds.). *Aerial Plant Surface Microbiology*. New York and London: Plenum Press, 1996, 1–24.
- Schulte AJ, Koch K, Spaeth M *et al*. Biomimetic replicas: transfer of complex architectures with different optical properties from plant surfaces onto technical materials. *Acta Biomater* 2009;**5**:1848–54.
- Schurr FM, Bossdorf O, Milton SJ *et al*. Spatial pattern formation in semi-arid shrubland: *a priori* predicted versus observed pattern characteristics. *Plant Ecol* 2004;**173**:271–82.
- Stiner L, Halverson LJ. Development and characterization of a Green Fluorescent Protein-based bacterial biosensor for bioavailable toluene and related compounds. *Appl Environ Microbiol* 2002;**68**:1962–71.
- Stoyan D. Further stereological formulae for spatial fibre processes. *Ser Stat* 1984;**15**:421–8.
- Stoyan D, Mecke J, Pohlmann S. Formulas for stationary planar fibre processes II – partially oriented-fibre systems. *Ser Stat* 1980;**11**:281–6.

- Stoyan D, Ohser J. Correlations between planar random structures, with an ecological application. *Biom J* 1982;**24**:631–47.
- Stoyan D, Stoyan H. *Fractals, Random Shapes and Point Fields: Methods of Geometrical Statistics*. John Wiley & Sons, 1994.
- Sundin GW, Jacobs JL. Ultraviolet radiation (UVR) sensitivity analysis and UVR survival strategies of a bacterial community from the phyllosphere of field-grown peanut (*Arachis hypogaea* L.). *Microb Ecol* 1999;**38**:27–38.
- Sun M, Luo C, Xu L *et al*. Artificial lotus leaf by nanocasting. *Langmuir* 2005;**21**:8978–81.
- Tecon R, Leveau JHJ. The mechanics of bacterial cluster formation on plant leaf surfaces as revealed by bioreporter technology. *Environ Microbiol* 2012;**14**:1325–32.
- Teplitski M, Robinson JB, Bauer WD. Plants secrete substances that mimic bacterial N-acyl homoserine lactone signal activities and affect population density-dependent behaviors in associated bacteria. *Mol Plant Microbe Interact* 2000;**13**:637–48.
- Timmer LW, Marois JJ, Achor D. Growth and survival of xanthomonads under conditions nonconducive to disease development. *Phytopathology* 1987;**77**:1341.
- Trouvelot S, Héloir M-C, Poinssot B *et al*. Carbohydrates in plant immunity and plant protection: roles and potential application as foliar sprays. *Front Plant Sci* 2014;**5**, DOI: 10.3389/fpls.2014.00592.
- Tukey HB. The leaching of substances from plants. *Annu Rev Plant Physiol* 1970;**21**:305–24.
- Tuson HH, Weibel DB. Bacteria–surface interactions. *Soft Matter* 2013;**9**:4368.
- Vorholt JA. Microbial life in the phyllosphere. *Nat Rev Microbiol* 2012;**10**:828–40.
- Wackett LP. Microbial adhesion. *Environ Microbiol Rep* 2015;**7**:164–5.
- van der Wal A, Tecon R, Kreft J-U *et al*. Explaining bacterial dispersion on leaf surfaces with an individual-based model (PHYLLOSIM). *Plos One* 2013;**8**:e75633.
- Wang F. *Quantitative Methods and Applications in GIS*. CRC Press, 2006.
- Weiss V, Ohser J, Nagel W. Second moment measure and *K*-function for planar STIT tessellations. *Image Anal Stereol* 2011;**29**:121–31.
- Whipps JM, Hand P, Pink D *et al*. Phyllosphere microbiology with special reference to diversity and plant genotype. *J Appl Microbiol* 2008;**105**:1744–55.
- Wiegand T, Gunatilleke CVS, Gunatilleke IAUN *et al*. How individual species structure diversity in tropical forests. *Proc Natl Acad Sci U S A* 2007;**104**:19029–33.
- Wiegand T, He F, Hubbell SP. A systematic comparison of summary characteristics for quantifying point patterns in ecology. *Ecography* 2013;**36**:92–103.
- Wiegand T, Moloney KA. Rings, circles, and null-models for point pattern analysis in

- ecology. *Oikos* 2004;**104**:209–29.
- Wiegand T, Moloney KA. *Handbook of Spatial Point-Pattern Analysis in Ecology*. 1st edition. Boca Raton: CRC Press, 2014.
- Wozniak DJ, Parsek MR. Surface-associated microbes continue to surprise us in their sophisticated strategies for assembling biofilm communities. *F1000Prime Rep* 2014;**6**, DOI: 10.12703/P6-26.
- Yadav RKP, Karamanoli K, Vokou D. Bacterial colonization of the phyllosphere of mediterranean perennial species as influenced by leaf structural and chemical features. *Microb Ecol* 2005;**50**:185–96.
- Yamada T. The Role of Auxin in Plant-Disease Development. *Annu Rev Phytopathol* 1993;**31**:253–73.
- Yang C-H, Crowley DE, Borneman J *et al*. Microbial phyllosphere populations are more complex than previously realized. *Proc Natl Acad Sci* 2001;**98**:3889–94.
- Yu H, Wiegand T, Yang X *et al*. The impact of fire and density-dependent mortality on the spatial patterns of a pine forest in the Hulun Buir sandland, Inner Mongolia, China. *For Ecol Manag* 2009;**257**:2098–107.
- Yu Q, Ma A, Cui M *et al*. Immigrant *Pantoea agglomerans* embedded within indigenous microbial aggregates: a novel spatial distribution of epiphytic bacteria. *J Environ Sci-China* 2014;**26**:398–403.
- Zhang B, Luo Y, Pearlstein AJ *et al*. Fabrication of biomimetically patterned surfaces and their application to probing plant–bacteria Interactions. *ACS Appl Mater Interfaces* 2014;**6**:12467–78.

Supplementary Materials

Supplementary materials to Chapter 3

General form of the entropy measure

The more general form of Eq. 3.1 is

$$\text{Eq. S3.1 : } E_D(n_n) = \frac{1}{n} \sum_{i=1}^n 1 - \frac{(n_n - 1) d_{min}}{\sum_{j=1}^{n_n-1} d_{ij}}$$

Here the neighborhood size n_n can be chosen depending on the experiment. For example in three-dimensional colonies $n_n = 13$. It is also useful in experiments which only consider early developing colonies, e.g. $n_n = 4$ for 4-cell states. Once n_n is chosen equation Eq. S3.1 simplifies to an equation similar to Eq. 3.1. The more general form of Eq. 3.2 is

$$\begin{aligned} \text{Eq. S3.2 : } E_I(n_n, n_c) \\ = \frac{1}{n} \sum_{i=1}^n \frac{\sqrt{n_{GFP,i}} + \sqrt{n_n - n_{GFP,i}} - \sqrt{n_n}}{(n_n \bmod n_c) \sqrt{[n_n n_c + 0.5]} + (n_c (n_n \bmod n_c)) \sqrt{[n_n n_c - 0.5]} - \sqrt{n_n}}. \end{aligned}$$

Here, mod stands for the modulo operation and brackets stand for the round-to-closest-integer operation. The denominator gives the maximum achievable entropy given the number of neighbors n_n-1 and the number of possible identities n_c (fluorescent dyes, species, etc.). The number of individuals in the neighborhood of the focal individual i which have the same color as i is given by $n_{GFP,i}$. Then $n_n - n_{GFP,i}$ gives the number of differently colored individuals without considering the entropy within this subset. This entropy is still mapped by considering every point i in the pattern.

For completeness we give here the combined general measure for entropy:

$$\begin{aligned} \text{Eq. S3.3 : } E(n_n, n_c) \\ = \frac{1}{n} \sum_{i=1}^n \gamma \left(1 - \frac{(n_n - 1) d_{min}}{\sum_{j=1}^{n_n-1} d_{ij}} \right) + \\ (1 - \gamma) \frac{\sqrt{n_{GFP,i}} + \sqrt{n_n - n_{GFP,i}} - \sqrt{n_n}}{(n_n \bmod n_c) \sqrt{[n_n n_c + 0.5]} + (n_c (n_n \bmod n_c)) \sqrt{[n_n n_c - 0.5]} - \sqrt{n_n}} \end{aligned}$$

Supplementary materials to Chapter 4

Table_S4 1: Overview of all samples analyzed in Chapter 4

Sample	Microscope objective	Area [μm^2]	Incubation time [h]	DsRed species	GFP species	ndSRred	nGFP	Stomata	Trichomes	Glandular	Hooked Trichomes	Vein cells	Nodes
61a_01	20	150022	21	299R	299R	129	162	2	0	1	1	112	
61a_02	20	149986	21	299R	299R	92	44	1	0	0	0	84	
61a_03	20	149878	21	299R	299R	43	80	0	0	1	0	75	
61a_04	20	150130	21	299R	299R	69	101	1	1	0	0	107	
61a_05	20	149986	21	299R	299R	59	79	3	0	1	0	83	
61a_11	10	597953	21	299R	299R	362	391	7	0	1	0	327	
61a_12	10	597953	21	299R	299R	602	795	6	0	0	0	291	
61a_13	10	529781	21	299R	299R	146	123	6	1	1	0	294	
61a_14	10	598528	21	299R	299R	141	172	11	0	1	0	323	
61a_15	10	598528	21	299R	299R	30	37	22	2	0	38	437	
61a_21	40	37410	21	299R	299R	91	57	0	0	0	6	16	
61b_06	20	149878	21	299R	299R	116	126	6	0	1	0	110	
61b_07	20	63797	21	299R	299R	51	41	2	1	0	17	79	
61b_09	20	150022	21	299R	299R	2	280	4	0	1	0	104	
61b_10	20	150022	21	299R	299R	38	78	10	1	0	5	123	
61b_16	20	598528	21	299R	299R	20	22	8	0	0	0	117	
61b_16	10	149878	21	299R	299R	560	603	20	0	1	0	397	
61b_17	10	597953	21	299R	299R	690	886	24	1	1	2	428	
61b_18	10	597953	21	299R	299R	193	334	11	0	2	0	336	
61b_19	10	597953	21	299R	299R	229	394	15	1	1	0	347	
61b_22	40	37464	21	299R	299R	21	2	1	0	0	0	18	
62a_02	20	150130	40	299R	299R	125	147	0	0	1	0	76	
62a_03	20	150130	40	299R	299R	114	15	4	0	0	0	82	
62a_04	20	150130	40	299R	299R	89	52	1	0	0	0	62	
62a_06	10	399300	40	299R	299R	110	239	10	0	0	5	224	
62a_08	10	436020	40	299R	299R	393	855	8	2	0	0	224	
62a_09	10	559538	40	299R	299R	125	131	11	1	0	0	335	
62a_10	10	328210	40	299R	299R	246	148	6	1	0	0	178	
63a_01	20	150130	48	299R	299R	177	43	0	0	0	0	74	
63a_05	20	146776	48	299R	299R	33	66	5	0	0	0	104	
63a_06	10	598528	48	299R	299R	414	685	6	1	0	10	335	
63a_07	10	598528	48	299R	299R	423	500	15	1	0	13	379	
63a_08	10	598528	48	299R	299R	194	185	14	2	2	21	393	
63a_09	10	506642	48	299R	299R	117	117	12	2	0	30	355	
63a_10	10	600540	48	299R	299R	106	75	4	0	1	6	351	
63b_16	10	598528	48	299R	299R	230	30	16	1	0	8	270	
63b_16	10	379033	48	299R	299R	151	176	11	0	0	13	252	

Sample	Microscope objective	Area [μm^2]	Incubation time [h]	DsRed species	GFP species	ndsRed	nGFP	Stomata	Glandular Trichomes	Hooked Trichomes	Vein cells	Nodes
63b_16	10	598528	48	299R	299R	14	62	10	2	0	23	295
64b_16	10	562672	10	299R	299R	6	22	15	0	0	0	254
64b_17	10	598528	10	299R	299R	2	62	19	2	0	15	277
64b_18	10	597953	10	299R	299R	20	34	14	0	0	6	267
65a_02	20	150130	25	299R	299R	106	0	5	0	0	0	64
65a_12	10	533028	25	299R	299R	108	291	21	1	1	0	337
65a_14	10	598528	25	299R	299R	23	148	18	1	0	0	317
65a_15	10	598528	25	299R	299R	326	514	24	1	0	17	335
65b_06	20	150130	25	299R	299R	42	89	5	0	0	0	81
65b_07	20	149914	25	299R	299R	215	1	4	1	0	0	74
65b_17	10	598528	25	299R	299R	296	116	16	0	0	3	253
65b_19	10	516690	25	299R	299R	87	95	13	0	0	0	264
65b_20	10	598528	25	299R	299R	248	189	15	1	0	0	294
66a_11	10	598528	30	299R	299R	8	78	9	0	2	0	241
66a_12	10	598528	30	299R	299R	108	245	8	2	0	14	328
66a_13	10	598528	30	299R	299R	239	587	21	2	0	39	330
66b_07	20	149770	30	299R	299R	535	555	4	0	0	0	61
66b_09	20	149878	30	299R	299R	0	96	5	0	0	0	76
66b_16	10	598528	30	299R	299R	568	41	23	1	1	29	351
66b_18	10	597953	30	299R	299R	251	120	14	1	1	0	284
66b_19	10	597953	30	299R	299R	115	23	16	2	0	0	279
66b_21	20	149878	30	299R	299R	3416	1	4	0	0	0	67
66b_22	20	150130	30	299R	299R	85	223	5	1	0	8	76
67a_01	20	149986	72	299R	299R	532	424	6	0	0	4	96
67a_03	20	150130	72	299R	299R	223	128	3	0	0	0	75
67a_11	10	598528	72	299R	299R	164	330	10	0	1	0	356
67a_12	10	598528	72	299R	299R	61	30	10	0	1	0	309
67a_14	10	598528	72	299R	299R	741	1007	20	1	0	19	395
67a_15	10	598528	72	299R	299R	196	109	12	0	1	0	330
67b_06	20	150130	72	299R	299R	113	127	2	1	0	1	67
67b_07	20	150130	72	299R	299R	162	135	6	1	0	8	122
67b_08	20	150130	72	299R	299R	279	288	5	0	0	0	89
67b_09	20	150130	72	299R	299R	89	121	6	0	0	0	98
67b_10	20	150130	72	299R	299R	117	114	5	0	0	0	91
67b_16	10	598528	72	299R	299R	731	821	16	0	1	0	368
67b_17	10	598528	72	299R	299R	479	402	20	0	2	0	386
67b_18	10	598528	72	299R	299R	305	331	28	1	0	13	409
67b_19	10	598528	72	299R	299R	237	302	14	2	0	7	397
68a_01	20	149878	60	299R	299R	1	0	0	0	0	2	89
68a_02	20	149878	60	299R	299R	72	0	8	0	0	5	99

Sample	Microscope objective	Area [μm^2]	Incubation time [h]	DsRed species	GFP species	ndsRed	nGFP	Stomata	Glandular Trichomes	Hooked Trichomes	Vein cells	Nodes
68a_03	20	149878	60	299R	299R	4	0	5	1	0	2	81
68a_04	20	149734	60	299R	299R	59	18	3	0	0	0	71
68a_05	20	149878	60	299R	299R	1	0	3	0	0	0	76
68a_11	10	598528	60	299R	299R	3	18	16	0	1	0	310
68a_12	10	598528	60	299R	299R	149	3	16	0	2	0	312
68a_14	10	547139	60	299R	299R	0	26	16	1	1	0	300
68a_15	10	597953	60	299R	299R	75	0	16	0	2	0	346
68b_06	20	149878	60	299R	299R	21	3	6	0	2	0	91
68b_07	20	150130	60	299R	299R	189	148	4	0	0	1	78
68b_08	20	150022	60	299R	299R	8	29	3	0	0	0	68
68b_09	20	150022	60	299R	299R	184	406	3	1	0	0	70
68b_10	20	149878	60	299R	299R	44	14	4	0	0	0	71
68b_16	10	598528	60	299R	299R	42	4	15	1	0	0	309
68b_17	10	597953	60	299R	299R	455	797	22	3	0	7	384
70a_12	10	598528	0	299R	299R	0	31	26	0	0	0	439
70a_13	10	598528	0	299R	299R	0	25	18	0	1	0	321
70a_15	10	598528	0	299R	299R	0	20	26	1	1	1	410
70b_06	20	149878	0	299R	299R	0	3	4	0	0	0	52
71a_01	40	37500	24	299R	299R	21	62	0	0	0	0	18
74a_03	20	149458	0	299R	299R	0	7	5	0	0	2	93
74a_05	20	149590	0	299R	299R	2	6	2	0	1	0	76
74b_06	20	149734	0	299R	299R	10	5	4	0	1	0	87
74b_07	20	149589	0	299R	299R	3	10	7	0	1	0	116
75a_01	20	150130	24	299R	299R	94	133	4	0	0	0	82
75a_02	20	150130	24	299R	299R	42	55	1	0	0	0	55
75a_04	20	150130	24	299R	299R	9	75	2	0	0	0	59
75a_05	20	150130	24	299R	299R	262	266	4	0	0	0	76
75a_11	10	598528	24	299R	299R	215	143	12	2	0	0	279
75a_13	10	598528	24	299R	299R	162	106	17	1	1	0	310
75a_14	10	598528	24	299R	299R	309	475	21	1	0	3	349
75a_15	10	598528	24	299R	299R	398	682	23	1	1	6	340
75b_16	10	598528	24	299R	299R	43	167	20	0	0	0	326
75b_17	10	598528	24	299R	299R	317	265	10	0	0	0	256
75b_18	10	374972	24	299R	299R	63	59	14	0	0	0	195
75b_19	10	376196	24	299R	299R	305	237	18	1	0	8	228
75b_20	10	598528	24	299R	299R	95	45	16	1	0	0	281
76a_01	20	150022	48	299R	299R	94	132	3	0	0	0	75
76a_02	20	150130	48	299R	299R	130	174	5	0	0	0	81
76a_04	20	150022	48	299R	299R	49	25	3	0	0	0	63
76a_11	10	321465	48	299R	299R	36	48	6	0	0	0	140

Sample	Microscope objective	Area [μm^2]	Incubation time [h]	DsRed species	GFP species	ndsRed	nGFP	Stomata	Glandular Trichomes	Hooked Trichomes	Vein cells	Nodes
76a_12	10	564730	48	299R	299R	100	63	14	0	0	0	273
76a_13	10	503100	48	299R	299R	121	89	5	1	0	0	217
76a_14	10	597953	48	299R	299R	377	622	22	2	0	48	369
76a_15	10	597953	48	299R	299R	45	128	14	0	1	0	354
76b_06	20	149878	48	299R	299R	58	151	8	0	0	6	108
76b_07	20	149878	48	299R	299R	57	48	6	0	0	0	81
76b_08	20	149986	48	299R	299R	155	177	4	1	2	39	99
76b_09	20	150022	48	299R	299R	110	32	5	0	0	0	83
76b_10	20	150130	48	299R	299R	26	241	5	1	0	0	80
76b_16	10	546825	48	299R	299R	37	85	21	0	0	0	344
76b_19	10	598528	48	299R	299R	331	191	22	2	0	0	355
76b_20	10	598528	48	299R	299R	433	601	23	1	0	5	394
77a_01	20	149878	72	299R	299R	290	111	4	2	0	9	85
77a_02	20	106330	72	299R	299R	397	210	0	0	0	0	44
77a_03	20	149986	72	299R	299R	96	108	2	0	0	0	68
77a_05	20	149986	72	299R	299R	121	158	8	0	0	6	99
77a_11	10	598528	72	299R	299R	641	731	8	0	1	0	275
77a_12	10	383590	72	299R	299R	505	666	13	4	1	11	231
77a_13	10	598528	72	299R	299R	383	482	14	4	1	19	362
77a_14	10	598528	72	299R	299R	145	157	3	1	0	0	247
77a_15	10	597953	72	299R	299R	1127	1519	20	1	0	18	353
77a_21	10	416689	72	299R	299R	1046	634	0	0	1	0	18
77b_07	20	127977	72	299R	299R	0	6	2	0	0	0	75
77b_08	20	149878	72	299R	299R	134	19	4	0	1	0	99
77b_09	20	141341	72	299R	299R	14	39	7	0	0	0	74
77b_10	20	149878	72	299R	299R	24	143	6	0	0	0	88
77b_16	10	598528	72	299R	299R	249	810	18	1	0	1	388
77b_17	10	597953	72	299R	299R	141	187	13	1	0	0	349
77b_18	10	598528	72	299R	299R	581	799	23	3	1	0	405
77b_19	10	598528	72	299R	299R	330	506	28	1	0	23	378
77b_20	10	598528	72	299R	299R	417	380	28	1	0	5	367
85a_15	10	598528	0	B728a	299R	28	10	12	1	2	0	445
85b_06	20	149986	0	B728a	299R	0	2	5	0	1	0	111
85b_08	20	149878	0	B728a	299R	7	2	6	0	1	0	114
85b_09	20	149878	0	B728a	299R	8	3	7	1	0	2	130
85b_10	20	149986	0	B728a	299R	4	2	6	1	0	5	129
86a_05	20	150022	24	B728a	299R	29	559	4	1	1	0	102
86a_11	10	597953	24	B728a	299R	16	400	15	2	0	19	398
86a_12	10	598528	24	B728a	299R	95	602	16	1	0	3	342
86a_13	10	598528	24	B728a	299R	42	606	11	0	1	0	376

Sample	Microscope objective	Area [μm^2]	Incubation time [h]	DsRed species	GFP species	ndsRed	nGFP	Stomata	Glandular Trichomes	Hooked Trichomes	Vein cells	Nodes
86a_21	10	598528	24	B728a	299R	204	431	18	2	1	14	434
86a_22	10	318630	24	B728a	299R	275	1222	14	2	0	9	226
86a_23	10	598528	24	B728a	299R	94	1385	18	2	1	2	406
86a_24	10	598528	24	B728a	299R	146	725	17	2	0	13	364
86a_25	10	598528	24	B728a	299R	69	1204	25	3	0	33	391
86b_07	20	149878	24	B728a	299R	9	108	5	2	1	0	127
87a_01	20	150130	48	B728a	299R	112	456	4	0	0	0	81
87a_02	20	150130	48	B728a	299R	68	188	3	0	0	0	73
87a_03	20	150130	48	B728a	299R	109	507	4	0	0	0	73
87a_04	20	150022	48	B728a	299R	94	463	7	1	0	4	123
87a_05	20	150130	48	B728a	299R	82	1187	4	0	0	0	83
87a_11	40	37374	48	B728a	299R	6	80	0	0	0	0	19
87a_12	40	37473	48	B728a	299R	21	170	1	0	0	0	12
87a_13	40	37473	48	B728a	299R	28	128	2	0	0	0	25
87a_14	40	37473	48	B728a	299R	37	58	2	0	0	0	32
87b_06	20	150022	48	B728a	299R	15	424	6	1	0	9	101
87b_07	20	149878	48	B728a	299R	1	982	5	0	0	0	96
87b_08	20	150022	48	B728a	299R	10	876	6	2	0	12	118
87b_09	20	150130	48	B728a	299R	7	179	10	0	0	5	98
87b_10	20	149878	48	B728a	299R	13	423	5	0	0	7	78
87b_16	40	37473	48	B728a	299R	2	156	1	0	0	0	20
87b_18	40	37509	48	B728a	299R	2	250	0	0	0	0	25
87b_19	40	31779	48	B728a	299R	3	71	2	0	0	0	18
87b_20	40	37446	48	B728a	299R	2	0	1	0	0	0	16
88a_01	20	150130	72	B728a	299R	466	301	6	0	0	0	73
88a_02	20	150130	72	B728a	299R	6	264	4	0	0	0	73
88a_04	20	150130	72	B728a	299R	21	357	8	0	0	0	89
88a_05	20	150130	72	B728a	299R	58	1309	8	1	1	9	93
88a_11	10	598528	72	B728a	299R	86	233	20	0	1	0	329
88a_12	10	598528	72	B728a	299R	262	752	26	0	0	0	329
88a_14	10	598528	72	B728a	299R	83	420	31	1	0	0	367
88a_21	10	598528	72	B728a	299R	176	1380	28	2	0	9	380
88b_06	20	150130	72	B728a	299R	0	156	3	0	0	0	94
88b_07	20	150130	72	B728a	299R	30	447	3	0	1	0	90
88b_08	20	150130	72	B728a	299R	2	96	4	0	0	0	80
88b_09	20	150130	72	B728a	299R	4	84	4	0	0	0	82
88b_10	20	150130	72	B728a	299R	14	119	0	1	0	0	64
88b_16	10	598528	72	B728a	299R	6	525	8	2	1	0	341
88b_17	10	598528	72	B728a	299R	3	843	10	2	1	0	348
88b_18	10	598528	72	B728a	299R	20	1856	10	0	1	0	288

Sample	Microscope objective	Area [μm^2]	Incubation time [h]	DsRed species	GFP species	ndsRed	nGFP	Stomata	Trichomes	Glandular Trichomes	Hooked Trichomes	Vein cells	Nodes
88b_19	10	598528	72	B728a	299R	87	698	10	2	2	0	312	
93a_01	20	149986	92+0	B728a	299R	664	68	8	1	0	0	122	
93a_03	20	78762	92+0	B728a	299R	116	37	5	0	0	0	56	
93a_04	20	98671	92+0	B728a	299R	642	25	7	1	0	0	79	
93a_05	20	149878	92+0	B728a	299R	5214	48	6	1	0	0	88	
93b_06	20	149986	92+0	B728a	299R	998	2	9	2	0	0	112	
93b_07	20	49152	92+0	B728a	299R	0	5	3	0	0	0	43	
93b_08	20	114754	92+0	B728a	299R	1235	10	9	1	0	12	109	
93b_09	20	49473	92+0	B728a	299R	1481	8	6	0	0	0	49	
94a_01	20	149878	92+24	B728a	299R	2173	0	6	0	0	0	84	
94a_02	20	149878	92+24	B728a	299R	493	8	5	0	1	6	83	
94a_03	20	137512	92+24	B728a	299R	1265	20	4	0	0	0	79	
94a_04	20	149986	92+24	B728a	299R	8390	24	4	0	0	0	76	
94a_05	20	150130	92+24	B728a	299R	2896	170	4	0	0	0	69	
94a_11	20	150022	92+24	B728a	299R	874	6	4	1	0	6	81	
94b_06	20	150022	92+24	B728a	299R	1265	520	5	1	0	0	96	
94b_07	20	150022	92+24	B728a	299R	612	171	1	1	1	0	90	
94b_08	20	150022	92+24	B728a	299R	528	72	5	0	0	0	95	
94b_09	20	123393	92+24	B728a	299R	311	75	4	2	0	7	93	
94b_10	20	106025	92+24	B728a	299R	344	14	4	0	0	0	58	
94b_12	20	150022	92+24	B728a	299R	455	61	4	1	0	0	78	
94b_13	20	149878	92+24	B728a	299R	422	1315	4	0	0	0	57	
95a_01	20	149878	92+48	B728a	299R	277	585	7	0	0	4	80	
95a_02	20	150022	92+48	B728a	299R	120	211	6	0	1	3	73	
95a_03	20	149878	92+48	B728a	299R	53	657	4	0	0	0	79	
95a_04	20	149878	92+48	B728a	299R	10	205	5	0	0	2	78	
95a_05	20	149986	92+48	B728a	299R	84	275	4	0	0	0	79	
95b_06	20	149878	92+48	B728a	299R	170	32	3	1	0	4	87	
95b_07	20	149878	92+48	B728a	299R	102	35	4	0	0	0	61	
95b_08	20	149986	92+48	B728a	299R	2	97	3	0	0	0	80	
95b_09	20	149878	92+48	B728a	299R	119	268	6	0	0	4	82	
95b_10	20	107739	92+48	B728a	299R	82	219	6	0	0	0	79	
96a_01	20	150022	92+72	B728a	299R	305	3	10	0	0	0	111	
96a_02	20	149878	92+72	B728a	299R	582	19	8	0	0	0	102	
96a_03	20	149878	92+72	B728a	299R	270	207	6	0	1	0	115	
96a_04	20	138291	92+72	B728a	299R	528	359	11	1	0	4	138	
96a_05	20	150022	92+72	B728a	299R	523	229	8	0	0	0	85	
96a_06	20	149878	92+72	B728a	299R	308	175	7	0	0	0	82	
96a_07	20	149878	92+72	B728a	299R	784	194	7	0	1	0	96	
96a_08	20	150022	92+72	B728a	299R	647	22	7	0	0	0	98	

Sample	Microscope objective	Area [μm^2]	Incubation time [h]	DsRed species	GFP species	nDsRed	nGFP	Stomata	Glandular Trichomes	Hooked Trichomes	Vein cells	Nodes
96b_09	20	149878	92+72	B728a	299R	90	299	10	0	0	0	125
96b_10	20	149986	92+72	B728a	299R	90	47	8	0	0	0	114
96b_11	20	150022	92+72	B728a	299R	147	79	7	1	0	0	115
96b_13	20	150022	92+72	B728a	299R	121	43	8	0	0	0	114
96b_14	20	150022	92+72	B728a	299R	37	33	8	0	0	0	111
96b_15	20	149878	92+72	B728a	299R	87	118	12	0	0	4	156
					Σ	69313	62116	2192	135	74	819	43674

Supplementary materials to Chapter 5

Table_S5 1: Overview of all samples analyzed in Chapter 5

Sample	Microscope objective	Area [μm^2]	Incubation time [h]	DsRed species	GFP species	ndsRed	nGFP	Stomata	Trichomes	Glandular	Hooked Trichomes	Vein cells	Nodes
80_01	20	150130	48	299R	299R	38	0	20	1	0	8	80	
80_02	20	150130	48	299R	299R	48	95	6	1	0	0	61	
80_03	20	150130	48	299R	299R	42	1	12	1	1	6	76	
80_06	10	598528	48	299R	299R	58	0	58	1	0	13	381	
80_07	10	598528	48	299R	299R	63	0	28	0	0	0	287	
81b_06	40	37437	0	B728a	299R	0	6	6	0	0	0	29	
81b_07	40	37473	0	B728a	299R	0	2	4	0	0	0	27	
81b_08	20	149878	0	B728a	299R	2	5	10	0	0	0	101	
81b_09	20	150022	0	B728a	299R	4	0	6	0	1	0	91	
81b_10	20	150130	0	B728a	299R	2	5	12	0	0	0	98	
81b_11	20	149986	0	B728a	299R	4	9	10	1	0	0	91	
81b_12	20	149986	0	B728a	299R	4	4	10	0	0	0	87	
82a_01	20	149986	24	B728a	299R	2997	7963	0	0	0	0	112	
82a_02	20	149878	24	B728a	299R	1967	4490	10	0	0	0	89	
82a_03	20	130749	24	B728a	299R	1703	6247	8	1	0	0	86	
82a_04	20	131832	24	B728a	299R	2100	4993	4	0	0	0	67	
82a_05	20	149878	24	B728a	299R	718	1593	8	0	0	0	80	
82a_11	20	149878	24	B728a	299R	1932	3465	17	0	0	5	92	
82a_12	20	149878	24	B728a	299R	2668	8477	12	0	0	0	107	
82a_13	20	149878	24	B728a	299R	1381	5074	10	0	0	0	105	
82a_14	20	135265	24	B728a	299R	1194	1879	8	0	0	0	86	
82a_15	20	149986	24	B728a	299R	569	2916	12	0	0	0	87	
82b_06	20	130742	24	B728a	299R	4	2100	12	1	0	0	85	
82b_07	20	149878	24	B728a	299R	1	485	8	0	0	0	90	
82b_08	20	150022	24	B728a	299R	201	1511	12	0	0	0	96	
82b_09	20	117773	24	B728a	299R	36	1024	14	1	0	4	92	
82b_10	20	149878	24	B728a	299R	2	226	12	0	0	0	89	
82b_16	20	149878	24	B728a	299R	2	379	8	0	0	0	85	
82b_17	20	150022	24	B728a	299R	75	1567	8	0	0	0	88	
82b_18	20	149986	24	B728a	299R	10	541	10	0	1	0	92	
82b_20	20	150022	24	B728a	299R	2	421	16	0	0	0	93	
83a_01	20	150130	48	B728a	299R	914	2453	8	1	0	0	107	
83a_02	20	150130	48	B728a	299R	465	90	6	0	0	0	74	
83a_03	20	150130	48	B728a	299R	1832	997	8	2	0	0	88	
83a_04	20	128148	48	B728a	299R	1326	1753	6	0	0	0	68	
83a_05	20	149842	48	B728a	299R	1942	4756	10	0	0	0	89	

Sample	Microscope objective	Area [μm^2]	Incubation time [h]	DsRed species	GFP species	ndSRed	nGFP	Stomata	Trichomes	Glandular Trichomes	Hooked Trichomes	Vein cells	Nodes
83a_11	40	37356	48	B728a	299R	400	1030	4	0	0	0	0	26
83a_12	40	37536	48	B728a	299R	835	428	2	0	0	0	0	20
83a_13	40	37509	48	B728a	299R	934	468	6	0	0	0	0	32
83a_14	40	37536	48	B728a	299R	80	120	2	0	0	0	0	26
83a_15	40	37536	48	B728a	299R	911	405	2	1	0	0	0	27
83b_06	20	150130	48	B728a	299R	1000	5881	8	0	0	0	0	73
83b_07	20	150130	48	B728a	299R	1279	5019	10	0	0	0	6	86
83b_08	20	150130	48	B728a	299R	1278	6787	6	0	0	0	0	77
83b_09	20	150022	48	B728a	299R	1264	5266	0	0	0	0	0	81
83b_16	40	37536	48	B728a	299R	433	1876	0	0	0	0	0	13
83b_17	40	37509	48	B728a	299R	86	2627	4	0	0	0	0	26
83b_18	40	37500	48	B728a	299R	179	1989	0	0	0	0	5	10
83b_19	40	37536	48	B728a	299R	416	1447	0	0	0	0	0	16
83b_20	40	37536	48	B728a	299R	272	2464	2	1	0	0	0	21
84b_01	10	484043	72	B728a	299R	743	2852	6	0	0	0	0	248
84b_02	10	598528	72	B728a	299R	1726	1877	12	1	1	1	7	342
84b_03	10	597953	72	B728a	299R	2722	1137	30	0	0	0	8	326
84b_04	10	598528	72	B728a	299R	839	3531	14	0	0	0	0	339
84b_05	10	597953	72	B728a	299R	919	5094	8	1	3	3	0	365
84b_06	20	150022	72	B728a	299R	1066	1815	12	0	0	0	11	110
84b_07	20	149842	72	B728a	299R	935	349	2	0	1	1	0	85
84b_09	20	149878	72	B728a	299R	43	1662	4	0	0	0	0	86
84b_10	20	149878	72	B728a	299R	253	43	2	0	0	0	0	97
89a_01	20	150022	0	299R	299R	13	58	4	0	0	0	0	88
89a_02	20	149878	0	299R	299R	17	31	6	0	1	1	0	118
89a_03	20	150022	0	299R	299R	34	28	10	0	1	1	0	110
89a_04	20	149986	0	299R	299R	118	62	6	0	0	0	0	94
89a_05	20	150022	0	299R	299R	55	87	16	1	0	0	0	114
89a_11	10	597953	0	299R	299R	52	91	46	1	0	0	0	432
89a_12	10	598528	0	299R	299R	69	93	41	0	1	1	0	358
89a_13	10	598528	0	299R	299R	144	143	38	0	0	0	0	385
89a_14	10	598528	0	299R	299R	62	62	34	0	1	1	0	388
89a_15	10	597953	0	299R	299R	84	80	40	0	1	1	0	418
89b_06	20	149878	0	299R	299R	65	32	16	1	0	0	3	107
89b_07	20	149878	0	299R	299R	26	10	8	0	0	0	0	88
89b_08	20	149878	0	299R	299R	28	8	8	0	0	0	0	81
89b_09	20	149986	0	299R	299R	32	19	12	0	0	0	0	94
89b_10	20	149878	0	299R	299R	9	30	6	0	0	0	0	77
89b_16	10	598528	0	299R	299R	36	45	62	1	0	0	15	422
89b_17	10	598528	0	299R	299R	72	149	40	2	0	0	7	395

Sample	Microscope objective	Area [μm^2]	Incubation time [h]	DsRed species	GFP species	ndsRed	nGFP	Stomata	Trichomes	Glandular Trichomes	Hooked Trichomes	Vein cells	Nodes
89b_18	10	597953	0	299R	299R	21	37	54	1	0	0	0	427
89b_19	10	597953	0	299R	299R	42	79	32	1	0	0	2	401
89b_20	10	598528	0	299R	299R	44	54	47	0	0	0	10	408
90a_02	20	150130	24	299R	299R	0	54	10	0	0	0	0	72
90a_03	20	150130	24	299R	299R	412	1293	8	0	0	0	0	74
90a_04	20	149878	24	299R	299R	240	63	8	0	0	0	0	70
90a_05	20	150022	24	299R	299R	1161	768	8	1	0	0	0	67
90a_11	10	598528	24	299R	299R	1732	2380	33	0	0	0	0	356
90a_12	10	598528	24	299R	299R	3978	3123	16	0	0	0	11	258
90a_13	10	598528	24	299R	299R	1372	1329	40	2	0	0	5	367
90a_14	10	598528	24	299R	299R	770	3515	32	3	0	0	0	363
90b_08	20	150130	24	299R	299R	70	724	4	1	0	0	0	75
90b_09	20	150130	24	299R	299R	57	673	8	0	0	0	0	78
90b_16	10	600540	24	299R	299R	30	59	30	1	0	0	0	338
90b_17	10	598528	24	299R	299R	82	866	30	0	0	0	0	305
90b_21	10	361502	24	299R	299R	7	3	10	0	0	0	0	227
91b_01	20	78319	48	299R	299R	21	344	14	0	0	0	0	72
91b_02	20	149878	48	299R	299R	679	6	12	0	0	0	0	97
91b_03	20	100615	48	299R	299R	322	1	4	0	0	0	0	59
91b_04	20	143286	48	299R	299R	85	89	8	0	0	0	0	88
91b_05	20	150022	48	299R	299R	186	24	16	0	0	0	0	99
91b_06	20	150022	48	299R	299R	20	630	12	0	0	0	0	83
91b_08	20	150022	48	299R	299R	736	81	7	0	1	0	0	78
91b_10	20	149986	48	299R	299R	285	127	15	0	0	0	0	108
91b_11	10	597953	48	299R	299R	1361	578	54	0	0	0	43	438
91b_12	10	597953	48	299R	299R	136	29	56	0	6	77	451	
91b_13	10	473290	48	299R	299R	763	11	26	1	0	0	0	284
91b_14	10	597953	48	299R	299R	362	266	48	1	0	0	10	391
91b_15	10	597953	48	299R	299R	676	763	38	0	0	0	0	353
91b_16	10	597953	48	299R	299R	295	674	38	0	1	0	0	368
91b_17	10	598528	48	299R	299R	346	178	37	0	1	0	0	363
91b_19	10	598528	48	299R	299R	1508	251	67	1	0	0	38	446
91b_20	10	598528	48	299R	299R	34	187	32	0	2	0	0	372
92a_01	20	149878	72	299R	299R	1233	840	10	1	0	0	7	103
92a_03	20	150130	72	299R	299R	833	116	12	0	0	0	0	86
92a_04	20	150130	72	299R	299R	303	241	10	0	0	0	2	85
92a_05	20	150022	72	299R	299R	1146	990	14	0	0	0	0	89
92b_06	20	115080	72	299R	299R	2	1198	10	1	0	0	5	77
92b_07	20	103429	72	299R	299R	3	168	4	0	1	0	0	63
					Σ	65188	143534	1852	36	25	308	18396	

Supplementary materials to Chapter 6

S6.1 – Derivation of the line-to-point distance measures d_{MEAN} und d_{POT}

In order to calculate the average distances d_{MEAN} (Eq. 6.6c) and d_{POT} (Eq. 6.6d) and the average potential U_{lp} , we need to solve the integrals

$$\text{Eq. S6.1 : } \int d(x_l, x_p) dx$$

and

$$\text{Eq. S6.2 : } \int \frac{1}{d(x_l, x_p)} dx.$$

To do so, we consider a point $x_p = P(p_1, p_2)$ and a line segment $x_l = AB$ which is defined by two endpoints (vertices) $A(a_1, a_2)$ and $B(b_1, b_2)$ in \mathbb{R}^2 . Any point $X(x_1, x_2) \in AB$ can be expressed by a control variable $\tilde{x} \in [0, 1]$ by

$$X = \begin{pmatrix} b_1 - a_1 \\ b_2 - a_2 \end{pmatrix} \tilde{x} + \begin{pmatrix} a_1 \\ a_2 \end{pmatrix},$$

such that $X = A$ if $\tilde{x} = 0$ and $X = B$ if $\tilde{x} = 1$. Now,

$$\text{Eq. S6.3 : } \bar{d} = \int_0^1 \sqrt{\xi_0 - \xi_1 \tilde{x} + \xi_2 \tilde{x}^2} dx,$$

where

$$\xi_0 = (p_1 - a_1)^2 + (p_2 - a_2)^2,$$

$$\xi_1 = -2((p_1 - a_1)(b_1 - a_1)) + ((p_2 - a_2)(b_2 - a_2)), \text{ and}$$

$$\xi_2 = (b_1 - a_1)^2 + (b_2 - a_2)^2.$$

Analogously to Gradstein & Ryshik (1981), we define $R = \xi_0 + \xi_1 \tilde{x} + \xi_2 \tilde{x}^2$ and $\Delta = 4\xi_0\xi_2 - \xi_1^2$ and their formula 2.262 (1.):

$$\text{Eq. S6.4 : } d_{\text{MEAN}} = \frac{(2\xi_2 \tilde{x} + \xi_1)\sqrt{r}}{4\xi_2|x_l|} + \frac{\Delta}{\xi_2} \cdot |x_l| d_{\text{POT}}$$

with 2.261 T(127) from Gradstein & Ryshik (1981)

$$\text{Eq. S6.5 : } d_{\text{POT}} = \frac{1}{|x_l|\sqrt{\xi_2}} \cdot \ln \left(2\sqrt{\xi_2 R} + 2\xi_2 \tilde{x} + \xi_1 \right).$$

S6.2 – Implementation

Our R-package *lppcf* contains an implementation of the four types of line-to-point pair correlation functions (LPPCFs). The main function `segment.pcf()` processes point patterns of type *ppp* and line patterns of type *linnet*, both native formats of the *spatstat* package (Baddeley and Turner 2005) for R (R Core Team 2013). A *ppp*-object contains information about the point pattern such as *x*/*y*-coordinates of the points, the dimensions of the observation window, and the unit of measure. Similarly, *linnet*-objects contain all relevant information about a line pattern (**linear network**) such as the coordinates of nodes (points) and a mapping of the connections between nodes. Thus Line patterns in the *lppcf*-package are approximated by a network of straight line pieces (segments) even if the observed pattern actually consists of a set of continuous curves. This is a common practice, e.g. in geographical information systems, but it has to be noted that the precision at which this approximation is performed (expressed in number of line segments) may introduce a bias into the data, while at the same time it has great influence on calculation performance.

The *segment.pcf*-function in our *lppcf*-package calculates the distances d_{MID} , d_{MIN} , d_{MEAN} , and d_{POT} (modes MID, MIN, MEAN, and POT, respectively) between the points and the lines. Based on these distances, the function estimates the corresponding LPPCFs at all spatial scales $h \in a \cdot \Delta_h$, $a \in [1, 2, \dots, h_{\text{max}}/\Delta_h]$, where h_{max} is maximum spatial scale up to which the LPPCF will be evaluated and Δ_h is the step-width between the scales at which the LPPCF will be evaluated. The function *segment.pcf* uses the estimator

$$\text{Eq. S6.6 : } \hat{g}_{lp}(h) = \frac{1}{2\pi h \lambda_l \lambda_p} \sum_{l=1}^{n_l} \sum_{p=1}^{n_p} \frac{\mathbf{1}(-0.5\Delta_h < d(x_l, x_p) - h < 0.5\Delta_h)}{|W_{\tilde{x}_l} \cap W_{x_p}|}$$

Here, λ_l and λ_p are estimators of the intensity of the line process and the point process, respectively, n_l and n_p are the number of lines and points, respectively. The indicator function $\mathbf{1}(\cdot)$ takes a value of 1 if the statement (\cdot) is true and 0 otherwise. The denominator $|W_{\tilde{x}_l} \cap W_{x_p}|$ is the translational edge correction factor (c.f. section Material and Methods in Chapter 6). The resulting function values can be smoothed in the *lppcf*-package using a moving Epanechnikov kernel. The bandwidth of the kernel is controlled by the parameter *epan.scale*. Line patterns with very different segment lengths can be dissected into smaller segments of approximate length Δl using the parameter *deltal*. This procedure will produce more precise results by reducing the segmentation bias discussed in Chapter 6 at the cost of increased calculation times. The default setting (*deltal*=Inf) will perform no dissections.

Abstract

Microbial organisms are ubiquitous to habitats on Earth and they are important components in cycling of matter. This is true for the micro-scale and at global scale, and all spatial scales in between. In soils, aquatic environments, in the atmosphere, as well as inside and on higher organisms, they are highly active in the activation and recycling of organic and mineral nutrients. They regulate population dynamics as pathogens or increase host fitness by reducing the effects of pathogens and toxic compounds. Despite their high abundance in nature and their important role for the environment, little is known about the spatial distribution and interactions of microbes, especially at the micrometer scale.

In my dissertation “*Scales of bacterial interactions on the leaf surface*”, I studied the spatial distribution of leaf-colonizing bacteria. My study system consisted of artificial single- and two-species communities of two common leaf-colonizing bacteria, *Pantoea agglomerans* and *Pseudomonas syringae*, colonizing bean leaves (*Phaseolus vulgaris*). At the center of my studies were fluorescence microscopic records which allowed the full measurement of the location of individual bacteria on the leaf. Additional phase-contrast micrographs revealed the location of leaf structural elements that were reported before to influence the spatial distribution of bacteria. These structural elements were glandular and hooked trichomes, leaf veins, stomata, and the network of crevices between epidermal cells (“grooves”). The resulting data set of bacterial and leaf structural point patterns were then analyzed using modern spatial statistical methods and here the pair correlation function (PCF) in particular. I analyzed the intraspecific and interspecific aggregation of both bacterial species as well as the spatial correlations between the bacterial colonizers and the leaf structural elements (Chapter 4). The correlations found by the PCF were generally interpreted to be bacterial interactions with their leaf environment. Additionally, the PCFs allowed an estimation of the spatial scale at which these interactions operate.

The experiments were furthermore doubled on artificial biomimetic leaf surfaces made from PDMS (Polydimethylsiloxane) by micro molding techniques (Chapter 4). This allowed to study how much of the observed bacterial distribution patterns on leaves can be explained by leaf topography and which interactions require additional leaf physiological processes.

The third main component of the dissertation involved the development of the line-to-

point-pair correlation function (LPPCF). The LPPCF extends the concepts of the existing bivariate PCF to interactions between point- and line-like objects such as bacteria and the grooves between epidermal cells (Chapter 6). I present four different LPPCFs that differ in the definition of the distance between a point and a line.

The dissertation makes multiple important contributions, both methodologically and with respect to the biology of leaf-colonizing bacteria. The ten most important contributions comprise the fields of microbial ecology, spatial ecology and spatial statistics.

Methodology

1. Introduction of PCFs in single-cell microbiology on plant leaves (Chapter 4).
2. Introduction of a new method involving artificial biomimetic leaves for differentiating leaf topographical and leaf physiological effects on the spatial distribution of bacteria on leaves (Chapter 5).
3. Introduction of a new extension of the pair correlation function for studying spatial correlations between point- and line-like objects (Chapter 6). The new line-to-point pair correlation function reignites an overdue discussion about how linear structures affect point-like objects in ecology (Chapter 6).
4. New definitions for the distance between a point and a line are introduced. The potential-based of these distance measure may be of particular interest to many research problems in ecology (Chapter 6).
5. Line patterns can be analyzed with satisfactory precision by approximating them by nodes (if forming a network) or by random points distributed on lines (Chapter 6).

Biology of leaf-colonizing bacteria

6. Bacterial individuals interact with their environment in different ways and the underlying processes operate at different spatial scales (Chapter 4).
7. The aggregation of bacteria near grooves near epidermal leaf cells is mainly driven by leaf topography (Chapters 4 and 5).
8. This aggregation near grooves operates at short distances ($<15 \mu\text{m}$; Chapter 4 & 5)
9. The aggregation of bacteria near trichomes and veins as well as the scarcity of bacteria near stomata cannot solely be explained by leaf topography (Chapter 4 & 5).
10. The effects of leaf trichomes, veins and stomata can operate also at larger scales ($100 \mu\text{m}$ and more; Chapter 4).

Acknowledgments

Although this thesis has my name written on the cover it is clear that it can never be the achievement of one person. Many people greatly supported me in my work as supervisors, coauthors, colleagues, and moral support. I want to thank all these people for their dedication.

I am most grateful to my executive supervisor **Katrin Meyer** for her constant support, her numerous good ideas, and her always insightful and encouraging comments during the design, conduction and publication of my research. Her profound knowledge of ecological theory proved an invaluable source of information. Early, she encouraged me to present my work at conferences where she also introduced me to many interesting scientist and thus afforded me to develop my scientific self-confidence. I am thankful for the splendid and enduring support in the last five years.

I am also very grateful to my supervisor **Kerstin Wiegand**, who provided an extraordinarily fruitful work environment that provided me with all the support I needed. In many discussions, she often added the one point of view that I was missing out on and thus brilliantly complemented the theoretical foundation of my work. During paper drafting, it was often fascinating to observe how she disclosed that the words I wrote, did not convey their intended meaning. I particularly appreciate the trust that she put into my work by funding my position for two more years.

Much gratitude appertains to **Martin Schlather** who fostered my interest in the theoretical background of spatial statistics. He gently pushed me to attain insights into the theory of pair correlation functions that I had not even hoped to reach. Without his continuous and tremendously patient support, the line-to-point pair correlation function would not have become what it is today. I do not know any mathematician or statistician with such a fine sense for what he can expect from applied scientists like me. Thank you for getting me there.

Similarly, I am very grateful to **Johan Leveau** who opened to me the field of phyllosphere microbiology. He kindly welcomed in his lab during my 6 month of data collection in the US. He introduced me very efficiently into his field of research which for the most part was new to me. He proved invaluable as an insightful and refreshingly critical co-author who, provided me with new points of view on my research.

I also would like my other co-authors, **Robin Tecon**, **Sean Gilmore**, and **Atul Parikh**, who provided me with the artificial leaf surfaces and assisted me in the experiments. My special thanks go to **Robin Tecon** who refreshed my memory and kindly assisted me in bench work-related things.

I sincerely thank my colleagues at the Ecosystem Modelling Group at the University of Göttingen for the endless fruitful discussions and for patiently listening to and commenting on my talks about the progress of my work. Of all these people, I would like to particularly thank **Joachim Saborowski**, **Matthias Fritsch**, **Jan Salecker**, and **Clara van Waveren** who have all been long-time colleagues and friends that I could rely on both in the scientific context and beyond. I also thank the other of the Leveau Lab, **Isaac Greenhut**, **Nilesh Maharaj**, **Gurdeep Rastogi**, **Jeness Scott**, and **Jan Tech**, who all helped me to become more of a microbiologists and who gave me a very good time in California. I also want to thank **Mitja Remus-Emsermann** for the great discussions, especially towards the end of my thesis writing.

Many thanks are due to my student helper force **Janika Heyden**, **Katja Karmrodt**, **Joanna Kohnke**, and **Alejandra Sarmiento**. These four brave women completed the heroic task of manually marking the location of at least half a million bacterial individuals and drawing almost 400 line patterns of the leaf surface. I am sorry that not all data could be included in this thesis.

I also sincerely thank the RTG 1644 coordination office **Monika Carlsson**, **Dörte Dede**, and **Barbara Strauss**. They were of incredible help with all administration-related things and always tried extra hard to come up with unusual and helpful solutions. The same is true for **Ilona Watteler-Spang** from the Ecological Modelling Group office. I thank her for her great work and the many encouraging talks.

My work was funded by the Research Training Group 1644 *Scaling Problems in Statistics* at the Georg-August-University of Göttingen, which was funded by the German Research Foundation DFG. After three years of funding, my position was extended by the Ecosystem Modelling Group at the Georg-August-University of Göttingen. Thank you to all.

Finally, I am grateful for all the moral support that I attained from myfamily and friends. I particularly thank my patient girlfriend **Janika Heyden**, who I love very deeply, and my parents **Willi** and **Karin Esser** who have permanently supported me so graciously all my life.

Curriculum vitae

Name: Daniel Sebastian Esser
Address: Hannoversche Straße 156 B, 37077 Göttingen
Born: 30. July 1982, Frechen (Germany)
Nationality: German
Highest degree: Diplom Biology (M.Sc.-equivalent)

Research Interests

My scientific interests orbit around the **analysis and modeling of ecological patterns** and the processes that created them. In my current PhD thesis, I study the distribution of **bacterial colonizers on plant leaf surfaces**. My research method relies equally on theoretical work, application and interdisciplinary exchange.

Key words: Microbial ecology, single-cell microbiology, plant-microbe interactions, environmental and population heterogeneity, spatial analysis, statistics, ecological modeling, plant community ecology

Professional experience

Since 03/2014 Scientific employee at the Department of Ecosystem Modelling,
University of Göttingen, Germany

10/2013 – 02/2014 Scientific assistant at the Department of Ecosystem Modelling,
University of Göttingen, Germany

10/2010 – 09/2013 Scientific employee as part of the research training group 1644
'Scaling Problems in Statistics' at **University of Göttingen**, Germany

11/2007 – 07/2010 Student assistant at the HRZ university computer center of the
University of Bonn, server division

Education

- Since 10/2010 PhD student in DFG Research Training Group 1644 – '**Scaling Problems in Statistics**' (Ecosystem Modelling Group), University of Göttingen, Germany. Thesis topic: *Scales of Bacterial Interactions on Plant Leaf Surfaces*
- 04/2006 – 03/2010 Advanced study period of biology at the University of Bonn, Germany. Main field of study: **Theoretical Biology** (Ecology, Geobotany), minor subjects: Zoology, Geography, Ecology and Environmental Sciences. Degree: **Diplom**, (M.Sc.-equivalent, grade: 1.0)
- Diploma thesis: *Species Area Relationship Analysis in Grassland Communities of Different Land Use Intensities by the Example of the Eifel* (grade: 1.0)
- 10/2003 – 03/2006 Basic study period of **Biology** at the University of Cologne (Germany), Degree: Vordiplom, grade: 1.1
- 09/1993 – 05/2002 Geschwister-Scholl-Gymnasium Pulheim (secondary school), School leaving certificate: Abitur, grade: 1,9
-

Publications

Published Articles

- Esser D.S.**, Leveau J.H.J., Meyer K.M., Wiegand K. (2015) Spatial scales of interactions among bacteria and between bacteria and the leaf surface. *FEMS Microbiol Ecol* 91 (3) fii034.
- Esser D.S.**, Meyer K.M., Leveau J.H.J. (2015) Modeling microbial growth and dynamics. *Applied Microbiology and Biotechnology* 99(21) 8831-8846

Articles in preparation

- Esser D.S.**, Leveau J.H.J., Meyer K.M., Wiegand K. (in prep.) Biomimetic leaf surfaces help detecting micro-scale effects of topography on spatial patterns of leaf-colonizing bacteria. *The ISME Journal*
- Esser D.S.**, Meyer K.M., Schlather M., Wiegand K. (in prep.) Scales of interactions between a point pattern and a line pattern – an extension to current methods in spatial statistics. *Methods in Ecology and Evolution*

Accepted Abstracts

Esser D.S., Meyer K.M., Schlather, M., Wiegand K. (2013) Interaction scales revealed with spatial statistics, talk at **GfÖ annual meeting**, Potsdam, Germany

Esser D.S., Leveau J.H.J., Meyer K.M., M., Wiegand K. (2013) Scales of bacterial interactions in the phyllosphere, poster presentation at **FEMS Microbiology congress**, Leipzig, Germany

Esser D.S., Meyer K.M., Leveau J.H.J., Schlather, M., Wiegand K. (2013) Scales of bacterial interactions with structural elements of the phyllosphere, invited talk at **Terrestrial Ecology Workgroup, University of Cologne**, Cologne, Germany

Esser D.S., Meyer K.M., Leveau J.H.J., Schlather, M., Wiegand K. (2012) Interactions between spatial point-and line-processes in ecology, talk at **GfÖ annual meeting**, Lüneburg, Germany

Esser D.S. (2012) Bacterial colonization patterns on plant surfaces: Adapting methods from spatial statistics to understand mixed point-line patterns, talk at **International Statistical Ecology Conference**, Sundvollen, Norway

Esser D.S., Meyer K.M., Leveau J.H.J., Wiegand K. (2011) Scales of bacterial interactions on plant leaf surfaces, poster presentation at **WE Heraeus Seminar 'Biothermodynamics of Metabolic and Ecological Networks'**, Bad Honnef, Germany

Workshops and conference sessions

Esser D.S., Steudel B., Remus-Emsermann M.N.P., (2015) The role of microbes in ecology, conference session during **GfÖ annual meeting**, Göttingen, Germany

Meyer K.M., **Esser D.S.** (2014) – Spatial patterns and ecological processes, conference session during **GfÖ annual meeting**, Hildesheim, Germany

Prices and honors

2nd price for best student presentation at International Statistical Ecology Conference (ISEC 2012), Sundvollen, Norway

Signature _____

Doktoranden-Erklärung der Georg-August-Universität

Göttingen

Name **Esser, Daniel**

(Name, Vorname)

Anschrift **Bonner Talweg 87, D-53113 Bonn**

(Straße, PLZ, Wohnort)

Ich habe eine Dissertation zum Thema **Scales of bacterial interactions on the leaf surface**

an der Georg-August-Universität Göttingen angefertigt. Dabei wurde ich von Frau/Herrn Prof. ... **Kerstin Wiegand** betreut.

Ich gebe folgende Erklärung ab:

1. Die Gelegenheit zum vorliegenden Promotionsvorhaben ist mir nicht kommerziell vermittelt worden. Insbesondere habe ich keine Organisation eingeschaltet, die gegen Entgelt Betreuerinnen und Betreuer für die Anfertigung von Dissertationen sucht oder die mir obliegenden Pflichten hinsichtlich der Prüfungsleistungen für mich ganz oder teilweise erledigt.

2. Hilfe Dritter wurde bis jetzt und wird auch künftig nur in wissenschaftlich vertretbarem und prüfungsrechtlich zulässigem Ausmaß in Anspruch genommen. Insbesondere sind alle Teile der Dissertation selbst angefertigt; fremde Hilfe habe ich dazu weder unentgeltlich noch entgeltlich entgegengenommen und werde dies auch zukünftig so halten.

Die Kapitel 4-6 meiner Dissertation basieren auf Manuskripten, die in wissenschaftlichen Zeitschriften publiziert wurden (Kapitel 4), oder für eine baldige Einreichung in ebensolchen Zeitschriften vorgesehen sind (Kapitel 5 und 6). Die Schaffung dieser Kapitel wurden im typischen wissenschaftlichen Rahmen durch die angegebenen Mitautoren, gerade mit Blick auf Aufbau und Sprache, unterstützt. Der Absatz über die Herstellung künstlicher Blattoberflächen in Kapitel 4 (Seite 50) wurde vollständig vom genannten Mitautor Robin Tecon geschrieben.

Des Weiteren ist mir bekannt, dass Unwahrhaftigkeiten hinsichtlich der vorstehenden Erklärung die Zulassung zur Promotion ausschließen bzw. später zum Verfahrensabbruch oder zur Rücknahme des erlangten Grades berechtigen.

....., den

(Ort)

..... (Unterschrift)