



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2023

Plant-insect chemical communication in ecological communities: an information theory perspective

Zu, Pengjuan ; García-García, Reinaldo ; Schuman, Meredith Christine ; Saavedra, Serguei ; Melián, Carlos J

Abstract: Cross-species communication, where signals are sent by one species and perceived by others, is one of the most intriguing types of communication that functionally links different species to form complex ecological networks. Global change and human activity can affect communication by increasing fluctuations in species composition and phenology, altering signal profiles and intensity, and introducing noise. So far, most studies on cross-species communication have focused on a few specific species isolated from ecological communities. Scaling up investigations of cross-species communication to the community level is currently hampered by a lack of conceptual and practical methodologies. Here, we propose an interdisciplinary framework based on information theory to investigate mechanisms shaping cross-species communication at the community level. We use plants and insects, the cornerstones of most ecosystems, as a showcase; and focus on chemical communication as the key communication channel. We first introduce some basic concepts of information theory, then we illustrate information patterns in plant-insect chemical communication, followed by a further exploration of how to integrate information theory into ecological and evolutionary processes to form testable mechanistic hypotheses. We conclude by highlighting the importance of community-level information as a means to better understand the maintenance and workings of ecological systems, especially during rapid global change.

DOI: <https://doi.org/10.1111/jse.12841>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-217381>

Journal Article

Accepted Version

Originally published at:

Zu, Pengjuan; García-García, Reinaldo; Schuman, Meredith Christine; Saavedra, Serguei; Melián, Carlos J (2023). Plant-insect chemical communication in ecological communities: an information theory perspective. *Journal of Systematics and Evolution*, 61(3):445-453.

DOI: <https://doi.org/10.1111/jse.12841>

Pengjuan Zu ORCID iD: 0000-0002-3222-598X

Plant-insect chemical communication in ecological communities: an information theory perspective

Pengjuan Zu^{1,2*}, Reinaldo García-García³, Meredith Schuman⁴, Serguei Saavedra⁵, Carlos J. Melián^{2,6}

¹Department of Environmental Systems Science, ETH Zurich, Schmelzbergstrasse 9, CH-8092, Zurich, Switzerland. ORCID: 0000-0002-3222-598X

²Department Fish Ecology & Evolution, Swiss Federal Institute of Aquatic Science and Technology (Eawag), Seestrasse 79, CH-6047, Kastanienbaum, Switzerland

³Departamento de Física y Matemática Aplicada, Facultad de Ciencias, Universidad de Navarra, Irunlarrea 1, 31008, Pamplona, Navarra, Spain. ORCID: 0000-0002-5491-6036

⁴Departments of Chemistry and Geography, University of Zurich, Winterthurerstrasse 190, Zurich CH-8057, Switzerland. ORCID: 0000-0003-3159-3534

⁵Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, 77 Massachusetts Av., 02139 Cambridge, MA, USA. ORCID: 0000-0003-1768-363X

⁶Institute of Ecology and Evolution, University of Bern, Baltzerstrasse 6, Bern, CH-3012, Switzerland. ORCID: 0000-0003-3974-6515

* Correspondence to: pengjuan.zu@gmail.com

Key words: Chemical communication, ecological networks, functional traits, information theory, plant-insect interaction, volatile organic compounds, Zipf's law

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/jse.12841.

This article is protected by copyright. All rights reserved.

Accepted Article

Abstract

Cross-species communication, where signals are sent by one species and perceived by others, is one of the most intriguing types of communication that functionally links different species to form complex ecological networks. Global change and human activity can affect communication by increasing fluctuations in species composition and phenology, altering signal profiles and intensity, and introducing noise. So far, most studies on cross-species communication have focused on a few specific species isolated from ecological communities. Scaling up investigations of cross-species communication to the community level is currently hampered by a lack of conceptual and practical methodologies. Here, we propose an interdisciplinary framework based on information theory to investigate mechanisms shaping cross-species communication at the community level. We use plants and insects, the cornerstones of most ecosystems, as a showcase; and focus on chemical communication as the key communication channel. We first introduce some basic concepts of information theory, then we illustrate information patterns in plant-insect chemical communication, followed by a further exploration of how to integrate information theory into ecological and evolutionary processes to form testable mechanistic hypotheses. We conclude by highlighting the importance of community-level information as a means to better understand the maintenance and workings of ecological systems, especially during rapid global change.

1 Introduction

Communication is prevalent in nature. For example, honeybees waggle to send information that guide other bees in the colony (Von Frisch, 1974); birds sing to mate or alert others in the flock (Freeberg, 2008); flowers exhibit colors and scent that can attract pollinators or deter herbivores (Schoonhoven et al., 2005). Cross-kingdom (Plantae and Animalia) plant-insect communication is of extreme interest not only because of its ubiquity and fundamental roles in both natural and agricultural systems (Isbell et al., 2011; Ollerton, 2021; Potts et al., 2016; Seastedt and Crossley Jr, 1984; Strong et al., 1984), but also due to the extraordinarily diverse communication

mediators shaped by hundreds of millions of years of co-evolution (Ehrlich and Raven, 1964).

Chemical communication is one of the most ancient and pivotal means for plants and insects. In fact, chemistry underlies color (pigments), shape (genetic encoding in nucleic acids and chemical inducers such as hormones), and scent (volatile organic compounds, or VOCs). Yet, chemical mediators of communication are most often considered from among the vast number of secondary metabolites produced by plants, with an estimate of 200,000 compounds including VOCs that have been extracted and identified (Kessler and Kalske, 2018; Wink, 2010), that are not strictly required for plant growth and development, but may have important defensive and attractive functions (Ehrlich and Raven, 1964; Fraenkel, 1959; Pichersky and Gershenzon, 2002). Insects, with their numerous chemoreceptors, are very sensitive to many plant chemicals (Hansson and Stensmyr, 2011; Kaupp, 2010; Schoonhoven et al., 2005) and depend on chemical signals for fundamental activities such as foraging (Schiestl, 2010), mating (Alexander et al., 1997), and oviposition (Renwick and Chew, 1994). Presumably, plant-feeding insects and their natural enemies have been under strong selective pressure to perceive and interpret plant chemicals which are relevant for these essential activities. So far, most plant-insect chemical communication studies have focused on specific species, isolating them from the community context in which they are naturally embedded. However, scaling up plant-insect communication from the species to the community level can be possible through the lens of information theory and communication systems (Shannon, 1948).

Communication can be considered as a process of information transfer from a sender who encodes the information (e.g., visual, vocal, or olfactory), to a receiver who decodes the information (Fig. 1). Depending on how well the information is reproduced at the receiver's end, we can quantify the clarity or ambiguity of the communication (Shannon, 1948). That is, the better a message is reproduced, the greater amount of information (clarity) it contains, and the lower its entropy (ambiguity, uncertainty). This mathematical theory of communication is called information theory. Fig. 1 summarizes key information quantities for a vector of signals (entropy and self-information of each

signal) and for an object-signal matrix (conditional entropy and mutual information). Information theory was originally developed for man-made communication systems such as telecommunication, and later applied to a wide range of communication-related fields, for example, cybernetics (Gabor, 1954), cryptography (Ahlsvede and Csiszár, 1993), linguistics (Ferrer-i-Cancho and Solé, 2003; Zipf, 1949) and neurobiology (Sharpee et al., 2014). Information theory brings at least two important new features that make it a promising tool for various systems and scales. First, information theory translates diverse communication signals into information. Thus, it unites different signals by extracting their information content without being overwhelmed by their different identities. Second, information theory measures the reproducibility of information between two states based on probabilistic concepts (Boso and Tartakovsky, 2018). This nonparametric perspective allows the scaling of information across different dimensions.

Not surprisingly, information theory has long been proposed as a conceptual and quantitative framework to study information transfer across ecological systems (Margalef, 1968; O'Connor et al., 2019; Ulanowicz, 2001). However, concrete applications to plant-insect communication are just starting to emerge (Zu et al., 2020). In this paper, we illustrate the application of information theory as a means to increase our understanding of plant-insect communication. In particular, we propose to consider plant VOCs as communication signals, and explore the emerging information patterns of plant-insect chemical communication. We then discuss potential future directions to uncover the underlying mechanisms of plant-insect communication by the integration of information theory with ecological and evolutionary processes. Note that here we use “signals” in an information theory framework to simply refer to *all* plant VOCs that form the repository (regardless whether they have demonstrated attractant or deterrent functions, or not).

2 Patterns of plant-insect chemical communication

2.1 Zipf's law

Zipf's law (Zipf, 1932, 1949) in linguistics describes that the frequency of a word decays rapidly proportional to its rank, or frequency of usage, following a power-law distribution ($p(r) = ar^{-k}$, where r denotes rank). Taking the English language for example, "the" is the most frequent word *in corpora*, occurring at a rate of 6% in British National Corpus, followed by "of", which only occurs half as often (3%) (Leech et al., 2014). And there is a large number of words that occur in very low frequency, leading to a heavy tail distribution. The power-law decay distribution is also related to the Pareto principle (or 80-20 rule) known in social sciences, i.e., roughly 80% of outcomes result from 20% of causes (Pareto, 1964). Importantly, the power-law exponent describing the speed of decay (i.e., the shape of this tailed distribution) and the redundancy of information in communication systems, is around $k = 1$ in many human languages (Piantadosi, 2014; Zipf, 1949). Zipf's law has also been explored in various biological systems (Semple et al., 2021).

If we consider a plant chemical "language", we can think of plants emitting different chemical "words" to their environment. Studies on plant VOCs have revealed that some VOCs are more frequent than others (Farré-Armengol et al., 2020; Knudsen et al., 2006), but the detailed information structures have not been explicitly studied. To explore the redundancy of plant chemical language and whether the frequency of its vocabulary (VOCs) follows Zipf's law, we gathered the data from the only four community-level plant VOC studies so far, three of which focused on floral VOCs in plant-pollinator networks (Burkle and Runyon, 2019; Filella et al., 2013; Kantsa et al., 2017) and one on leaf VOCs in a plant-herbivore (caterpillar) network (Zu et al., 2020). Additionally, we also use a review study by Farré-Armengol et al. (2020), who compiled a floral VOC dataset from 305 species. In the review study, the authors also categorized chemical groups of VOCs, families of plant species, and pollination systems of these plant species.

By analysing the distributions of VOCs in these studies and in different categories (chemical groups, plant families, pollination groups), we found that most cases follow a heavy-tail distribution where a few VOCs are predominant whereas many other VOCs occur rarely (Fig. 2). In 5 cases (out of the 16 examined cases and categories), a power-law distribution fits either the whole data (in one community study PV_k , and the group of N-, S-containing compounds, Fig. 2, Table 1) or part of the data (in Orchidaceae, insect-pollinated flowers, and the overall review data, Fig. 2, Table 1). The fitting was performed using the R package “powerLaw” (Clauset et al., 2009; Gillespie, 2017). Many of the remaining cases appear to have power-law like distributions, but a fit cannot be statistically supported (Table S1). Indeed, although power-law distributions are thought to be common, it is not easy to achieve a strict fit to many empirical data (Clauset et al., 2009) due to limited sample size and sample bias on abundant vs. rare signals.

In the 5 cases in which a full or partial fit can be made, the slopes of the distribution range from $k = 1.94$ to $k = 2.68$, which is steeper than observed in human languages. These patterns suggest that plant VOC languages may be more redundant than human language. Following information theory, the entropy of VOCs in all these cases is very high, ranging from 0.78 to 0.99 ($H(V)$ in Table 1a).

Note that throughout the whole paper, we used standardised entropy so that it ranges between 0 and 1, where the higher the value (closer to 1), the higher the uncertainty. Entropy calculation formulas can be found in Fig. 1.

2.2 Coding, decoding and interactions

In the previous section, we only focus on the putative signals (VOCs) themselves. In this section, we place these into the context of communication: the coding and decoding of signals, and species interactions. Following the work by Zu et al. (2020), we can use the associations of VOCs (V) with plants (P) in a community to describe plant coding process (PV-matrix), the associations of VOCs with insects (or animals, A) to describe insect decoding process (AV-matrix), and associations of animals with plants (AP-matrix) for insect-plant interactions. Empirically, to screen how each insect reacts to

each of the VOCs in the community requires extensive experimental manipulations under both controlled and natural conditions. Therefore, it is challenging to generate the whole insect decoding matrix from field and experimental work.

From the four community-level plant VOC studies which documented a PV-matrix, three of them (Burkle and Runyon, 2019; Kantsa et al., 2019; Zu et al., 2020) also document insect-plant interaction networks (AP-matrix) in the same community. We found that plant coding patterns vary from community to community (Fig. 3, Table 1). In the case of the leaf VOC community (PV_z), VOCs are more commonly shared among plants ($H(P|V) = 0.94$), whereas in two cases of flower VOCs (PV_f, PV_k), there are more unique VOCs ($H(P|V) = 0.41, 0.51$ respectively), and the third case of flower VOCs is in between ($H(P|V) = 0.75$). We also analyzed plant coding patterns based on the review study, and found these to vary from category to category (Table 1a), with the overall (from the whole dataset) value of around 0.5 indicating some specialization ($H(P|V) = 0.53, H(V|P) = 0.41$, Table 1a).

We observe specialised patterns for all these insect-plant interactions as revealed by entropy values much less than 1 ($H(A|P)$ around 0.5, $H(P|A)$ around 0.2, Table 1b, Fig. 4). The asymmetric values of $H(A|P)$ and $H(P|A)$ indicate that insects are more specialized on plants than plants specialized on insects ($H(P|A)$ are lower). In addition, the interaction patterns seem to be indistinguishable between mutualistic (pollinator-plant) and antagonistic (herbivore-plant) networks based on these few studies. Indeed, a global analysis of more than 250 mutualistic and antagonistic networks have revealed that the network structures (e.g., generality, nestedness) *per se* cannot separate mutualistic from antagonistic networks (Michalska-Smith and Allesina, 2019) unless one includes environmental factors (Song and Saavedra, 2020).

3 From patterns to mechanisms

We have shown the information structures of plant-insect chemical communications as revealed by the apparent power-law behavior in some cases (but where these conclusions are limited by limited data availability), and by the ubiquitous decrease in

entropy of VOC associations to plants as opposed to total entropy of VOCs. The latter is not surprising, as not all plants produce all VOCs. Interestingly, the conditional entropy of VOCs ($H(P|V)$) involved in plant-herbivore interactions seems to be generally higher than the entropy of those involved in plant-pollinator interactions; but this is not reflected in the conditional entropy of plant-animal associations ($H(A|P)$) - which is similar among these interaction types. This suggests that plants try to communicate more clearly with their pollinators than with their herbivores, while herbivores must try harder than pollinators to decode information from their host plants. While we recognize that further conclusions are likewise currently limited by a shortage of data, we can still ask why plant-insect chemical communication is structured in the way it is, and not otherwise. In this section, we highlight two approaches to investigate the potential mechanisms driving these chemical communication patterns between plants and insects. Recall the difficulties of building a complete decoding matrix (how each insect decodes each of the VOCs) empirically. Here, we theoretically construct insect decoding from two approaches: 1) using a top-down logic assuming that the insect can decode all the VOCs from a plant as long as they can feed on that plant, that is, $AV = AP \times PV$; and 2) use a bottom-up logic by hypothesizing the relationships between VOC abundance and functionality (i.e., effects on insects). In both directions, we integrate information theory into ecological and evolutionary theories to generate testable hypotheses.

3.1 Top-down: Interactions shape communication

This direction follows the inspiration of studies exploring how Zipf's law patterns emerge in linguistics. Various hypotheses have been proposed to explain this intriguing pattern in human languages. Among these, the "least effort" hypothesis (Ferrer-i-Cancho and Solé, 2003; Zipf, 1949) is a compelling hypothesis that nicely recovers the Zipf's law distribution as a natural consequence from the conflicting interests between speakers (who are thought to aim for "brevity and phonological reduction" to minimize their effort to speak) and listeners (who should desire "explicitness and clarity" to minimize their effort to understand). The successful application of information theory in studying the structure and emergence of human language has inspired studies on animal

vocal communication (e.g., reviews in Kershenbaum et al. 2021; McCowan et al. 2008; McCowan et al. 1999), and indicates its potential to be extended in studying the chemical “language” in plant-insect communities (Zu et al., 2020).

Indeed, Zu et al. (2020) aimed to borrow the framework to test whether plant-herbivore chemical communication patterns can be explained by conflicting interests of plants (speakers) and herbivores (listeners). It is well-known that the chemical arms race between plants and herbivores has been playing out for hundreds of millions of years, where herbivores must constantly adapt to plant defense chemicals while plants must keep producing novel chemicals (Ehrlich and Raven, 1964). Zu and colleagues (Zu et al., 2020) translated this chemical arms race into an information arms race: plants code the (chemical) information in a way (changes in PV-matrix) to make the decoding (AV given by $AP \times PV$) difficult (high entropy) for herbivores, whereas herbivores interact with plants in a way (changes in AP-matrix) to make the decoding easier (low entropy). With repeated cycles of information optimisation by plants and herbivores, they found that an equilibrium stage is reached in which a plant-VOC redundancy matrix and herbivore-plant specialization matrix have emerged. These information patterns at the equilibrium stage matched their field data collected in a tropical dry forest.

This is a successful case of integrating other known evolutionary and ecological theories into information theory to build a framework that aims to disentangle mechanisms driving plant-insect interactions and communication at the community level. In general, there are three steps to construct communication frameworks for a given system. First, we must define who are the speakers and who are the listeners. Second, we must define proper fitness functions for both parties based on their ecological and evolutionary relationships. Third, we must define the rules of optimization for each party in the information game. Both the second and the third steps require very careful examination by integrating evolutionary and ecological processes. We also want to emphasize that the “fitness” function at the community level, which defines success in terms of passing into the next “round” in a game, is not the same as the Darwinian fitness proxies which biologists normally use to assess individual success in survival and reproduction, but rather refers to broader benefits which can be assessed

at the level of communities rather than individuals. This corresponds as well to the broader use of the term “signal” introduced at the beginning of this paper. This approach is compatible with the evaluation of individual-level fitness, but aims at criteria which can be used at the community level. We hope more studies will be inspired by this information perspective to provide insights into other plant-insect chemical communication types (e.g., plant-pollinator, plant-herbivore-parasitoid).

3.2 Bottom-up: Connect VOC frequency with “information functionality”

A different, bottom-up approach would connect VOC frequency with information functionality to study plant-insect communication. Given that there are rare and common VOCs (frequency), it is natural to ask whether rare VOCs or common VOCs can be decoded more easily by insects (functionality) and how much information can be gained by decoding the VOCs.

Shannon’s original mathematical theory of communication (Shannon, 1948) was mostly based on the “Inverse Relationship Principle” (D’Alfonso, 2011) which states that the less probable a signal is, the more information it bears. In our plant-insect communication systems, insects face an environment with VOCs in different frequencies. On the one hand, specializing in decoding rare VOCs brings advantages if substantial information is thereby gained to access a niche that can be exploited by specialists. However, specialization comes with risks because rare VOCs (emitted by only one plant species or a few phenotypes) can go extinct more easily, locally or globally. On the other hand, decoding common VOCs is likely to be safer but less informative, meaning that insects might need to decode more signals to gather enough information to identify their plant hosts. Which strategy do insects use and how will different strategies affect the interaction network structure? For example, will positive-frequency dependent decoding structure, i.e., preferentially decoding common VOCs, lead to more generalized plant-insect interactions; whereas negative-frequency dependence, preferentially decoding rare VOCs, may result in more specialized interactions? How could this be understood in terms of the co-evolution of plant VOCs and insect olfaction?

Within this framework, one can hypothesize different specific relationships between VOC frequency and information functionality (VA-matrix), to test how these give rise to plant-insect interaction network structures (using $PA = PV \times VA$). Specifically, information functionality can be modelled with two elements: one, how easily a VOC can be decoded (VA_{func}); two, how much information can be gained by decoding the VOC (V_{info}). Therefore, the information functionality matrix (VA) can be treated as an information-weighted decoding matrix, generated by ($VA = V_{info} \times VA_{func}$). We can use a drift process mimicking frequency-independent decoding which might act as a test for the strength of frequency-dependent decoding occurring in the empirical data.

Conclusion

We have found from the limited studies that plant VOC frequencies follow a heavy-tail distribution with a few predominant compounds and many rare compounds; all the three documented plant-insect interaction matrices are highly specialized, whereas plant-VOC coding patterns seem to vary from case to case, with different plant communities and perhaps, interaction types, resulting in signals that are encoded more or less redundantly. We also provide a top-down and a bottom-up direction for constructing theoretical frameworks to integrate information theory in studies of plant-insect chemical communication. By proposing both top-down and bottom-up frameworks, we hope we can draw attention to, and inspire more studies on the information perspective of trophic networks in ecosystems.

In the face of more serious and more frequent threats to ecological systems, plant-insect interactions are vulnerable, yet particularly important aspects of ecosystem functioning: insect pollinators are vital for agricultural productivity as well as for the survival of wild plant communities (Ollerton, 2021; Wei et al., 2021); herbivores support all higher trophic levels which depend on plants (Harvey et al., 2003; Moreira et al., 2016; Price et al., 1980), and their predators and parasitoids perform pest control services (Schmidt et al., 2003). We still struggle to control these interactions with relatively crude tools which carry heavy collateral damage, such as physically uprooting and transporting

communities of pollinators to agricultural fields, and broad prescribed spraying of pesticides. Integrating information theory has the potential to help us gain a better understanding of the underlying mechanisms, and thus promises novel insights to better understand and manage these interactions. For example, novel chemicals of exotic plants have been suggested to be a main mechanism for invading the local communities (Cappuccino and Arnason, 2006; Macel et al., 2014). From the perspective of information theory, we can ask what are the effects of invasive species on the information landscape of the current community, and whether some information structures are more resilient than others. We can furthermore ask how climate change, pollution (including pesticide) and habitat fragmentation will affect communication signal profiles, dynamics, and signal-to-noise ratios and thus rewire plant-insect networks. Overall, understanding the effects of communication on species interactions, and *vice versa*, has the potential to improve our ability to assess impacts of anthropogenic disturbances on ecosystems, and assist conservation and restoration practices. For instance, instead of focusing on taxonomic diversity, chemical diversity and chemical information structure may be shown to be more crucial and informative for ecosystem biodiversity (Schuman et al., 2016).

Information approaches can provide the scalability to integrate further dimensions to achieve a better and potentially more unified understanding of information flows and their effects in ecosystems (Deng et al., 2021). For instance, we can expand these frameworks by including different sets of VOCs (e.g., from flowers, leaves and roots); different sets of partners and trophic levels (e.g., pollinators, microbes, parasitoids and predators); and even different information forms (e.g., visual, vocal). Information theory abstracts all kinds of signals (or functional traits) as information, and thus has the potential to gather the multi-layer functional maps into a unified framework. In summary, integrating information theory to existing theories in ecology and evolution has the potential to unveil central biological mechanisms driving the formation and maintenance of the functionality of species interactions and thus of entire ecosystems (Kessler and Kalske, 2018; Margalef, 1968; O'Connor et al., 2019), as has recently been argued more generally for plant chemistry (Walker et al., 2022).

Bibliography

- Ahlsvede, R. and Csiszár, I. 1993. Common randomness in information theory and cryptography. i. secret sharing. *IEEE Transactions on Information Theory* 39:1121–1132.
- Alexander, R. D., Marshall, D. C., and Cooley, J. R. 1997. Evolutionary perspectives on insect mating. *The evolution of mating systems in insects and arachnids*: 4–31.
- Boso, F. and Tartakovsky, D. M. 2018. Information-theoretic approach to bidirectional scaling. *Water Resources Research* 54:4916–4928.
- Burkle, L. A. and Runyon, J. B. 2019. Floral volatiles structure plant–pollinator interactions in a diverse community across the growing season. *Functional Ecology* 33:2116–2129.
- Cappuccino, N. and Arnason, J. T. 2006. Novel chemistry of invasive exotic plants. *Biology letters* 2:189–193.
- Clauset, A., Shalizi, C. R., and Newman, M. E. 2009. Power-law distributions in empirical data. *SIAM review* 51:661–703.
- Deng, J., Angulo, M. T., and Saavedra, S. 2021. Generalizing game-changing species across microbial communities. *ISME Communications* 1:1–8.
- D’Alfonso, S. 2011. On quantifying semantic information. *Information* 2:61–101.
- Ehrlich, P. R. and Raven, P. H. 1964. Butterflies and plants: a study in coevolution. *Evolution* 18:586–608.
- Farré-Armengol, G., Fernández-Martínez, M., Filella, I., Junker, R. R., and Peñuelas, J. 2020. Deciphering the biotic and climatic factors that influence floral scents: a systematic review of floral volatile emissions. *Frontiers in plant science* 11:1154.
- Ferrer-i-Cancho, R. and Solé, R. V. 2003. Least effort and the origins of scaling in human language. *Proceedings of the National Academy of Sciences* 100:788–791.

- Filella, I., Primante, C., Llusia, J., González, A. M. M., Seco, R., Farré-Armengol, G., Rodrigo, A., Bosch, J., and Penuelas, J. 2013. Floral advertisement scent in a changing plant-pollinators market. *Scientific Reports* 3:3434.
- Fraenkel, G. S. 1959. The raison d'etre of secondary plant substances. *Science*: 1466–1470.
- Freeberg, T. M. 2008. Complexity in the chick-a-dee call of carolina chickadees (poecile carolinensis): associations of context and signaler behavior to call structure. *The Auk* 125:896–907.
- Gabor, D. 1954. Communication theory and cybernetics. *Transactions of the IRE Professional Group on Circuit Theory* Pages 19–31.
- Gillespie, C. S. 2017. Estimating the number of casualties in the american indian war: a bayesian analysis using the power law distribution. *The Annals of Applied Statistics* 11:2357–2374.
- Hansson, B. S. and Stensmyr, M. C. 2011. Evolution of insect olfaction. *Neuron* 72:698–711.
- Harvey, J. A., Van Dam, N. M., and Gols, R. 2003. Interactions over four trophic levels: foodplant quality affects development of a hyperparasitoid as mediated through a herbivore and its primary parasitoid. *Journal of Animal Ecology* 72:520–531.
- Isbell, F., Calcagno, V., Hector, A., Connolly, J., Harpole, W. S., Reich, P. B., Scherer-Lorenzen, M., Schmid, B., Tilman, D., Van Ruijven, J., et al. 2011. High plant diversity is needed to maintain ecosystem services. *Nature* 477:199–202.
- Kantsa, A., Raguso, R. A., Dyer, A. G., Sgardelis, S. P., Olesen, J. M., and Petanidou, T. 2017. Community-wide integration of floral colour and scent in a mediterranean scrubland. *Nature Ecology & Evolution* 1:1502.
- Kantsa, A., Raguso, R. A., Lekkas, T., Kalantzi, O.-I., and Petanidou, T. 2019. Floral volatiles and visitors: A meta-network of associations in a natural community. *Journal of Ecology* 107:2574–2586.
- Kaupp, U. B. 2010. Olfactory signalling in vertebrates and insects: differences and commonalities. *Nature Reviews Neuroscience* 11:188.

- Kershenbaum, A., Demartsev, V., Gammon, D. E., Geffen, E., Gustison, M. L., Ilany, A., and Lameira, A. R. 2021. Shannon entropy as a robust estimator of zipf's law in animal vocal communication repertoires. *Methods in Ecology and Evolution* 12:553–564.
- Kessler, A. and Kalske, A. 2018. Plant secondary metabolite diversity and species interactions. *Annual Review of Ecology, Evolution, and Systematics* 49:115–138.
- Knudsen, J. T., Eriksson, R., Gershenzon, J., and Ståhl, B. 2006. Diversity and distribution of floral scent. *The Botanical Review* 72:1.
- Leech, G., Rayson, P., et al., 2014. *Word frequencies in written and spoken English: Based on the British National Corpus*. Routledge.
- Macel, M., de Vos, R. C., Jansen, J. J., van der Putten, W. H., and van Dam, N. M. 2014. Novel chemistry of invasive plants: exotic species have more unique metabolomic profiles than native congeners. *Ecology and Evolution* 4:2777–2786.
- Margalef, R., 1968. *Perspectives in Ecological Theory*. University of Chicago Press, Chicago.
- McCowan, B., Doyle, L., Hanser, S., Kaufman, A., and Burgess, C., 2008. *Detection and estimation of complexity and contextual flexibility in nonhuman animal communication systems*. MIT Press: Cambridge, MA, USA.
- McCowan, B., Hanser, S. F., and Doyle, L. R. 1999. Quantitative tools for comparing animal communication systems: information theory applied to bottlenose dolphin whistle repertoires. *Animal behaviour* 57:409–419.
- Michalska-Smith, M. J. and Allesina, S. 2019. Telling ecological networks apart by their structure: A computational challenge. *PLoS computational biology* 15:e1007076.
- Moreira, X., Abdala-Roberts, L., Rasmann, S., Castagneyrol, B., and Mooney, K. A. 2016. Plant diversity effects on insect herbivores and their natural enemies: current thinking, recent findings, and future directions. *Current opinion in insect science* 14:1–7.

- O'Connor, M. I., Pennell, M., Altermatt, F., Matthews, B., Melian, C., and Gonzalez, A. 2019. Principles of ecology revisited: integrating information and ecological theories for a more unified science. *Frontiers in Ecology and Evolution* 7:219.
- Ollerton, J., 2021. *Pollinators and Pollination: Nature and Society*. Pelagic Publishing Ltd.
- Pareto, V., 1964. *Cours d'économie politique*, volume 1. Librairie Droz.
- Piantadosi, S. T. 2014. Zipf's word frequency law in natural language: A critical review and future directions. *Psychonomic bulletin & review* 21:1112–1130.
- Pichersky, E. and Gershenzon, J. 2002. The formation and function of plant volatiles: perfumes for pollinator attraction and defense. *Current opinion in plant biology* 5:237–243.
- Potts, S. G., Imperatriz-Fonseca, V., Ngo, H. T., Aizen, M. A., Biesmeijer, J. C., Breeze, T. D., Dicks, L. V., Garibaldi, L. A., Hill, R., Settele, J., et al. 2016. Safeguarding pollinators and their values to human well-being. *Nature* 540:220–229.
- Price, P. W., Bouton, C. E., Gross, P., McPherson, B. A., Thompson, J. N., and Weis, A. E. 1980. Interactions among three trophic levels: influence of plants on interactions between insect herbivores and natural enemies. *Annual review of Ecology and Systematics* 11:41–65.
- Renwick, J. and Chew, F. 1994. Oviposition behavior in lepidoptera. *Annual Review of Entomology* 39:377–400.
- Schiestl, F. P. 2010. The evolution of floral scent and insect chemical communication. *Ecology Letters* 13:643–656.
- Schmidt, M. H., Lauer, A., Purtauf, T., Thies, C., Schaefer, M., and Tschamtkke, T. 2003. Relative importance of predators and parasitoids for cereal aphid control. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 270:1905–1909.
- Schoonhoven, L. M., Van Loon, B., van Loon, J. J., and Dicke, M., 2005. *Insect-plant biology*. Oxford University Press on Demand.

- Schuman, M. C., van Dam, N. M., Beran, F., and Harpole, W. S. 2016. How does plant chemical diversity contribute to biodiversity at higher trophic levels? *Current opinion in insect science* 14:46–55.
- Seastedt, T. and Crossley Jr, D. 1984. The influence of arthropods on ecosystems. *Bioscience* 34:157–161.
- Semple, S., Ferrer-i-Cancho, R., and Gustison, M. L. 2021. Linguistic laws in biology. *Trends in Ecology & Evolution* 37(1):53-56.
- Shannon, C. E. 1948. A mathematical theory of communication. *Bell System Technical Journal* 27:379–423.
- Sharpee, T. O., Calhoun, A. J., and Chalasani, S. H. 2014. Information theory of adaptation in neurons, behavior, and mood. *Current opinion in neurobiology* 25:47–53.
- Song, C. and Saavedra, S. 2020. Telling ecological networks apart by their structure: An environment-dependent approach. *PLoS computational biology* 16:e1007787.
- Strong, D. R., Lawton, J. H., Southwood, S. R., et al., 1984. *Insects on plants. Community patterns and mechanisms*. Blackwell Scientific Publications.
- Ulanowicz, R. E. 2001. Information theory in ecology. *Computers & Chemistry* 25:393–399.
- Von Frisch, K. 1974. Decoding the language of the bee. *Science* 185:663–668.
- Walker, T. W., Alexander, J. M., Allard, P.-M., Baines, O., Baldy, V., Bardgett, R. D., Capdevila, P., Coley, P. D., David, B., Defosse, E., et al., 2022. Functional traits 2.0: The power of the metabolome for ecology. *Journal of Ecology*, 110(1), 4-20.
- Wei, N., Kaczorowski, R. L., Arceo-Gómez, G., O’Neill, E. M., Hayes, R. A., and Ashman, T.-L. 2021. Pollinators contribute to the maintenance of flowering plant diversity. *Nature* 579(7878):688-692.
- Wink, M., 2010. *Annual plant reviews, functions and biotechnology of plant secondary metabolites*, volume 39. John Wiley & Sons.

Zipf, G. K. 1932. *Selected studies of the principle of relative frequency in language*. Cambridge: Harvard University Press.

Zipf, G. K. 1949. *Human behavior and the principle of least effort: an introduction to human ecology*. Cambridge: Addison-Wesley Press.

Zu, P., Boege, K., Del-Val, E., Schuman, M. C., Stevenson, P. C., Zaldivar-Riverón, A., and Saavedra, S. 2020. Information arms race explains plant-herbivore chemical communication in ecological communities. *Science* 368:1377–1381.

Acknowledgments

PZ acknowledges the Swiss National Science Foundation (SNSF) Spark grant (CRSK-3 196506) and SNSF Prima grant (PR00P3 193237). MCS acknowledges the University of Zurich University Research Priority Program on Global Change and Biodiversity and membership in the US NSF-funded Biology Integration Institute BII-Implementation: The causes and consequences of plant biodiversity across scales in a rapidly changing world (Award Number:2021898). Funding to SS was provided by NSF grant No. DEB-2024349.

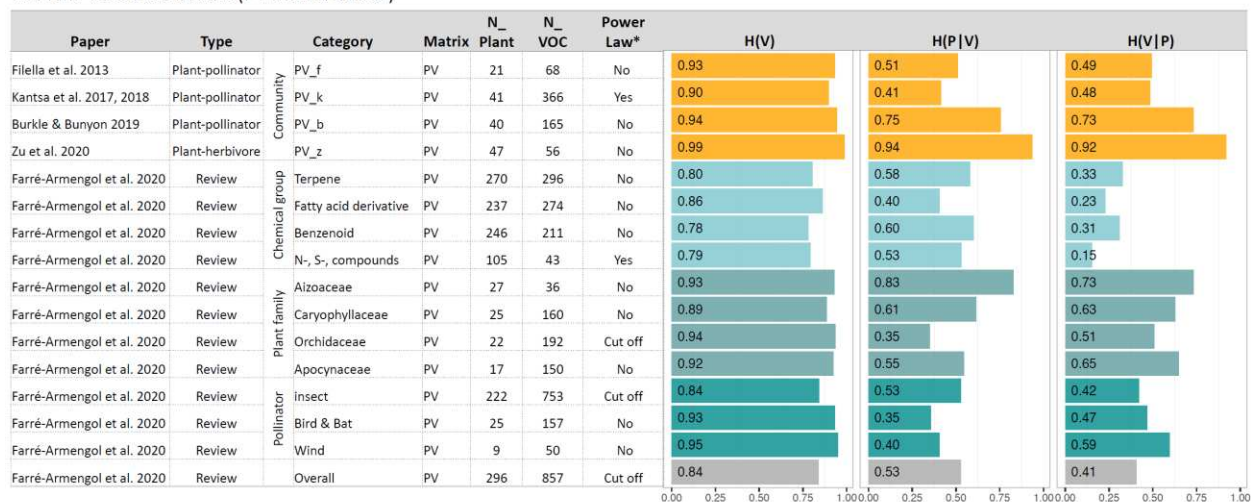
Competing financial interests The authors declare no competing financial interests.

Author contributions PZ initiated and performed the study. All authors contributed with ideas and wrote the manuscript.

Data accessibility The data and R code supporting the results can be found at https://github.com/PengjuanZu/Zu_etal_Plant-insect_InformationTheory.

Table 1. Entropy measures based on **a)** Plant-VOC associations (PV-matrix based) and **b)** Insect-plant networks (AP-matrix based) in the four community-level studies papers and a review paper that compiled plant-VOC publications. $H(V)$ entropy of VOCs. $H(\cdot|\cdot)$ conditional entropy. Calculation formulas see Fig. 1. * Details of statistical tests for power law see Table S1.

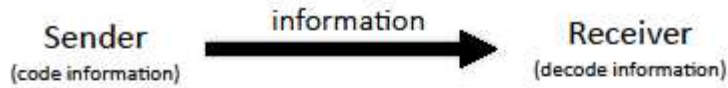
A. Plant-VOC associations (PV-matrix based)



B. Insect-plant networks (AP-matrix based)



Figure legends



Vector

For a vector S with n signals: $S = (S_1, S_2, \dots, S_n)$.

Probability of each signal is: (p_1, p_2, \dots, p_n) .

Using information theory, we can calculate:

Entropy (H) of the information, which describes the uncertainty (or the ambiguity) of the signals:

$$H = - \sum_{i=1}^n p_i \log_2 p_i$$

The information each signal contains is called

Self information (I):

$$I(S_i) = \log_2 \frac{1}{p_i}$$

Matrix

	Signal 1	...	Signal n
Object 1	0.6	0.1	0.3
...	0.1	0.5	0.4
Object m	0.6	0.1	0.3

(Object-Signal matrix)

For a matrix with n signals and the corresponding m objects, we can calculate:

Conditional entropy quantifies the uncertainty (or the amount of information needed) about one variable given another variable.

$$H(S_1|S_2) = - \sum_{i=1}^n p(S_1) \log_2 p(S_1|S_2)$$

Where $p(S_1) = \frac{\sum_{j=1}^m p_{1j}}{\sum_{j=1}^m p_{1j} + \dots + p_{nj}}$

$$H(S_1|S_2) = - \sum_{j=1}^m p(S_1|S_2) \log_2 p(S_1|S_2)$$

$$p(S_1|S_2) = \frac{p_{1j}}{\sum_{i=1}^n p_{ij}}$$

Mutual information quantifies mutual dependence between two variables. In our example of object-signal matrix, it quantifies the amount of information obtained about the object by observing the signals.

$$I(S, O) = H(S, O) - H(S|O) - H(O|S)$$

Where $H(S, O) = - \sum_{i=1}^n \sum_{j=1}^m p_{ij} \log_2 p_{ij}$

$$H(S) = \frac{1}{n}$$

$$I(S, O) = H(S, O)$$

Fig. 1. Information transferring process in communication and formula quantifying information in communication systems.

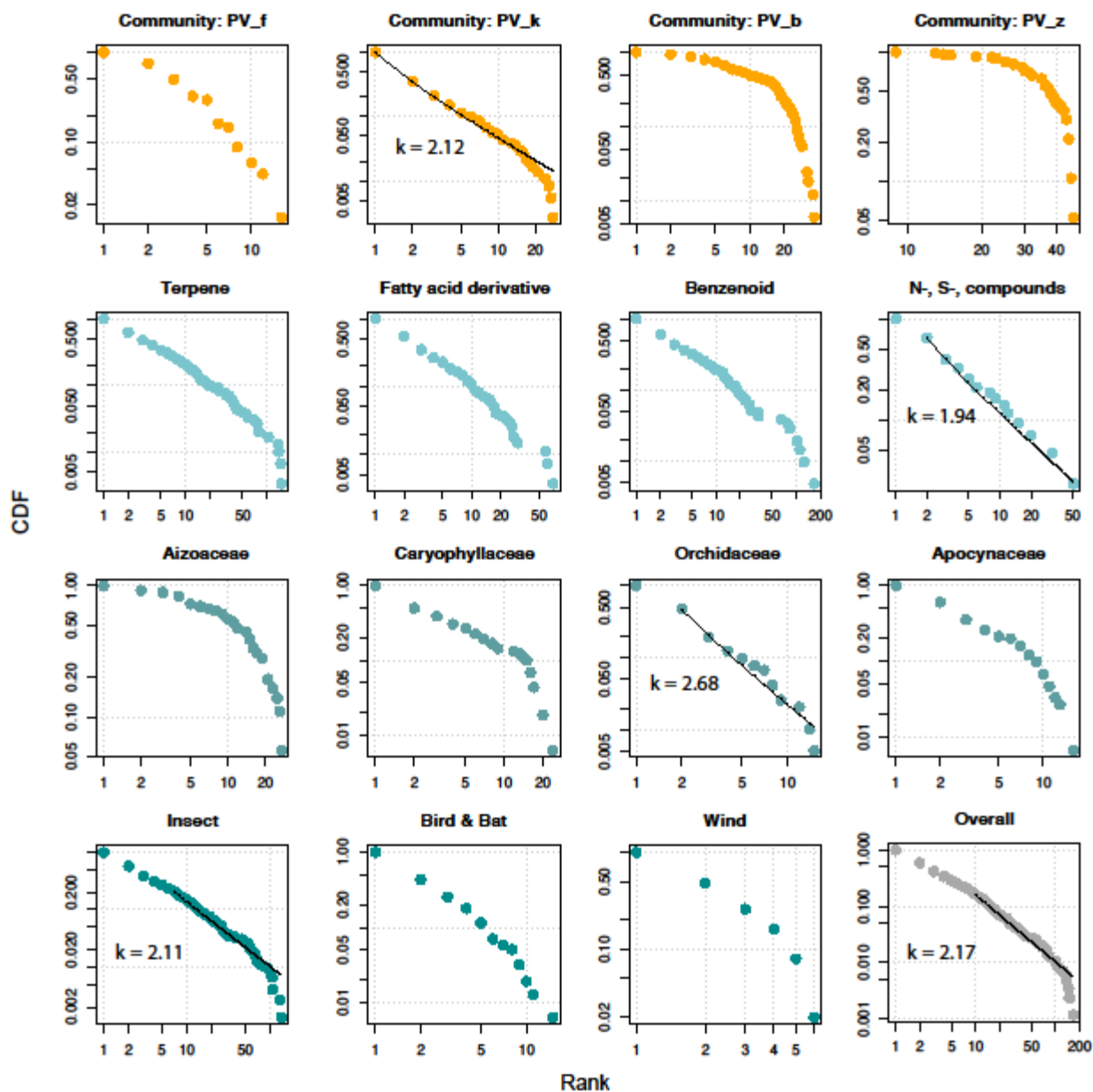


Fig. 2: VOC frequency vs rank (log-log scale) from four community-level study papers (top row) and a review paper (2nd to 4th rows) that summarised VOC frequency in different chemical groups (2nd row), four representatives of plant families (3rd row), and plants with different pollination groups and the overall review data (4th row). Black lines in the cases of “Community: PV_k”, “N-, S-, compounds”, “Orchidaceae”, “Overall” indicate that data can be described by power law distribution (k values represent the exponent of the power-law distribution $p(r) = ar^{-k}$), although log normal distribution can describe the data similarly well (tests see Supplementary Table S1). N-, S-, compounds: nitrogen or sulphur containing volatile compounds. Number of sampled

plant species (N_plant) and number of VOCs (N_VOC) in each case can be found in Table 1.

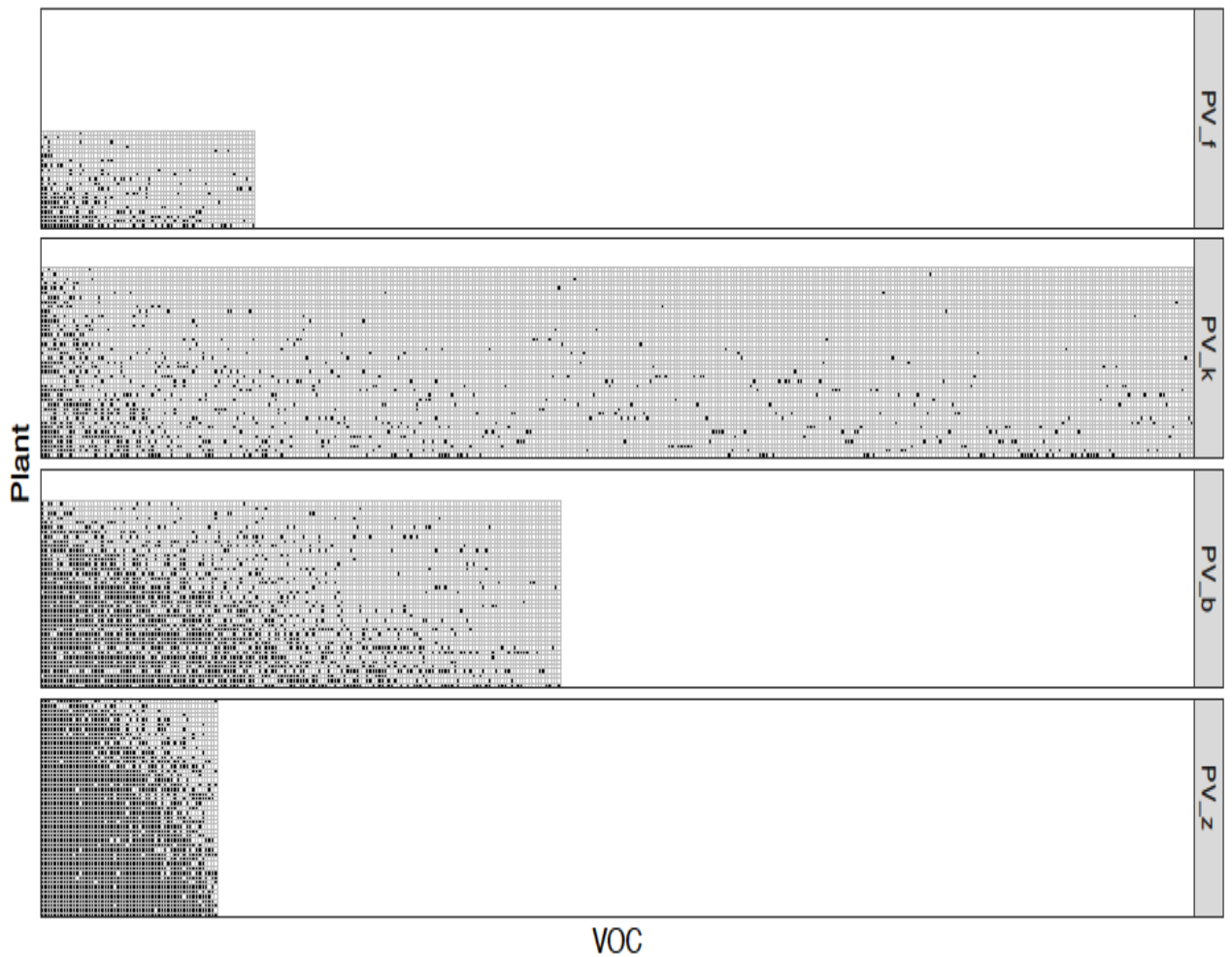


Fig. 3: Patterns of plant-VOC associations (PV-matrix) in the four community studies (details see top four rows in Table 1a). Each row and column represents plant species and VOCs respectively. Black and white squares show that a given VOC was present or absent in a given plant, respectively.

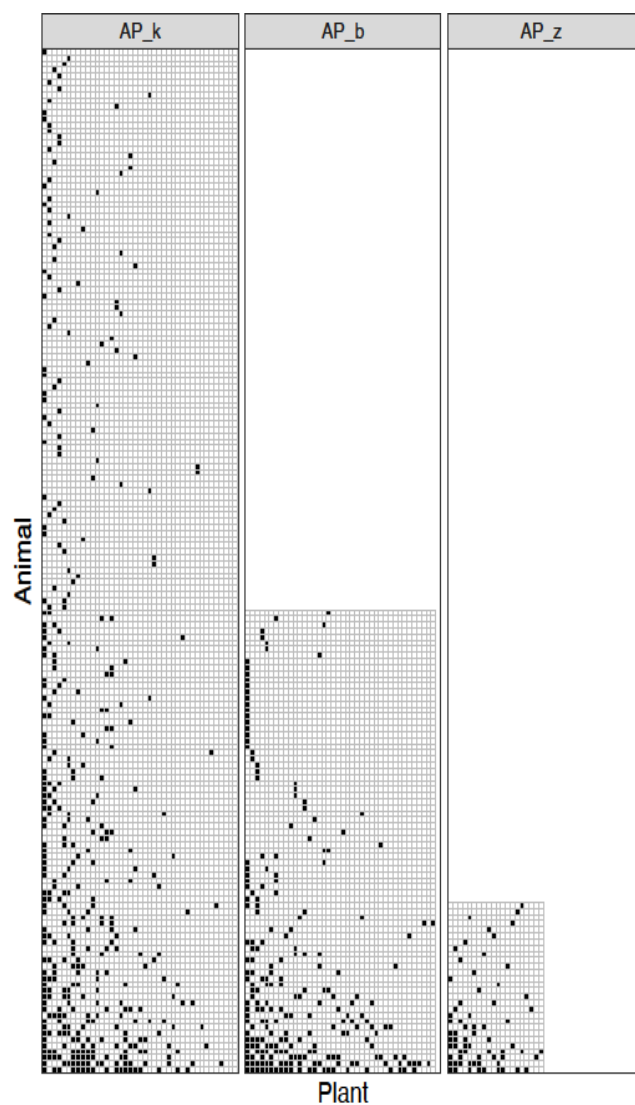


Fig. 4: Patterns of insect-plant interactions (AP-matrix) in the three community studies (details see Table 1b). Each row and column represents insect species and plant species respectively. Black and white squares show that a given insect was, or was not observed feeding on a given plant, respectively