CHEMICAL ECOLOGY

Information arms race explains plant-herbivore chemical communication in ecological communities

Pengjuan Zu1*, Karina Boege2, Ek del-Val3, Meredith C. Schuman4,5, Philip C. Stevenson6,7, Alejandro Zaldivar-Riverón8, Serguei Saavedra1

Plants emit an extraordinary diversity of chemicals that provide information about their identity and mediate their interactions with insects. However, most studies of this have focused on a few model species in controlled environments, limiting our capacity to understand plant-insect chemical communication in ecological communities. Here, by integrating information theory with ecological and evolutionary theories, we show that a stable information structure of plant volatile organic compounds (VOCs) can emerge from a conflicting information process between plants and herbivores. We corroborate this information “arms race” theory with field data recording plant-VOC associations and plant-herbivore interactions in a tropical dry forest. We reveal that plant VOC redundancy and herbivore specialization can be explained by a conflicting information transfer. Information-based communication approaches can increase our understanding of species interactions across trophic levels.

Chemical information is an ancient and ubiquitous channel to mediate species interactions (1) (e.g., attracting or repelling individuals) and is regarded as one of the main forces shaping plant-herbivore interaction networks (2, 3). For example, insects have a large number of olfactory receptors with high sensitivity for chemical signals (4, 5), which play essential roles in their fundamental activities (6), including foraging (7) and oviposition (8). Chemically mediated species interactions are based on information transfer, which we define simply as communication regardless of the benefit to the emitter and receiver (following previous work (9)). Understanding the chemical communication between plants and herbivores has been an active field of research (10–12), and most work has focused on how a specific plant (or genus) defends against herbivores directly by using chemical repellents (13) or by producing stronger chemical defenses (14–16), or indirectly by emitting signals to attract herbivores’ enemies (e.g., predators or parasitoids) (17–22). So far, only a few descriptive studies (23, 24) have begun to investigate this chemical communication at the community level. More critically, little is known about how information transfer shapes species interactions. For example, it remains unclear how plants code chemical information to deal with a diversity of potential herbivores and how herbivores decode such olfactory signals to distinguish among plants and identify potential hosts. Even less is known about whether there is any general information structure of plant-herbivore chemical communication and how such a structure can be maintained under an ongoing chemical “arms race” between plants and herbivores.

Information theory (25), which provides a quantitative and scalable way to measure information transfer, has already brought key insights about the structure and emergence of human language (26), and it can be extended to increase our understanding about the “chemical language” in ecological communities (12, 27–29). However, such attempts have only been used to study chemical communication of a single plant species with its interacting insects (30). Here, we used information theory to study plant volatile organic compounds (VOCs) as a communication channel forming plant-insect interaction networks (31). From an information perspective, the relationship between sender and receiver (or speaker and hearer) determines the nature of the signal transmission and the evolution of the information structure shaping the communication pattern between individuals. That is, individuals can either provide clear information that can be decoded easily or spurious information that can be difficult to decode.

Moreover, a conflicting or harmonious communication process between sender and receiver can be modeled by the combinations of maximizing or minimizing uncertainty (or mutual information) of the two parties, respectively (26). Specifically, for a given communication system that is formalized by a matrix describing how a vector of signals (S) is associated with a vector of objects (O), the clearness (or the spuriousness) of the communication can be described by two important information measurements (25). First, mutual information \( I(O|S) \) describes the amount of information that one can obtain about objects by knowing the signals. Second, conditional entropy \( H(O|S) \) describes the uncertainty of correctly identifying an object given that a specific signal has been observed (for a detailed mathematical account of these expressions, see the supplementary materials (32)). Both mutual information and conditional entropy represent the efficiency of the coding (and decoding) strategies and can be related to the fitness of senders and receivers (26).

Tropical dry forests host a large number of interacting and coevolving plant and butterfly species and are regarded as biodiversity hotspots (33). Our fieldwork was conducted in a tropical dry forest of the Chamela-Cuixmala Biosphere Reserve (19°22’ to 19°39’ N; 104°56’ to 105°10’ W) in Jalisco, Mexico. During the rainy season of 2018, we searched comprehensively for lepidopteran larvae on leaves of target plants in our transect plots and reared the larvae in the laboratory with leaves collected from their host plant species to confirm their trophic interaction and to identify the herbivore species. The supplementary materials (32) provide a detailed account of plants and insects, sampling, and identification procedures. We constructed a qualitative (i.e., presence or absence) herbivore (animal)-plant (AP) interaction matrix comprising 28 lepidopteran herbivore species and 20 plant species. The AP matrix was rather sparse (see Fig. 1A and table S1), with a median of one plant species per herbivore (range, one to nine).

Additionally, we sampled the headspace around leaves to retrieve VOCs from each of these plant species. We were able to match 93 analytes from headspace samples to the NIST7 VOC library, of which 56 were likely biogenic (from plants or from plants and microbes), and 31 of these biogenic VOCs could be identified by one or more abundant ions in an untargeted analysis. We used single representative ions from these 31 VOCs for quantification. The supplementary materials (32) provide a detailed account of VOC sampling, identification, and quantification procedures. Figure 1B (and table S2) shows that, contrary to the AP matrix, the plant-VOC (PV) association matrix is dense and

---

*Corresponding author. Email: pengjuan.zu@gmail.com

1Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. 2Instituto de Ecología, Universidad Nacional Autónoma de México, 04510 Ciudad de México, Mexico. 3Instituto de Investigaciones en Ecosistemas y Sustentabilidad, Universidad Nacional Autónoma de México, 58190 Morelia, Michoacán, Mexico. 4Department of Molecular Ecology, Max Planck Institute for Chemical Ecology, DE-07745, Germany. 5Departments of Chemistry and Geography, University of Zurich, CH-8057 Zurich, Switzerland. 6Royal Botanic Gardens, Kew Richmond, Surrey TW9 3AB, UK. 7Natural Resources Institute, University of Greenwich, Chatham, Kent ME4 4TB, UK.

Zu et al., Science 368, 1377–1381 (2020) 19 June 2020
erratic, with a median of 26 VOCs per plant (range, 13 to 30).

The PV matrix provides information about PV associations, whereas the AV matrix (mathematically defined by the matrix multiplication \( AP \times PV \)) provides information about herbivore-VOC associations. These associations describe the patterns of plants emitting information and herbivores receiving it and thus represent plant coding and herbivore decoding strategies. The AV matrix (Fig. 1C) is dense and weighted, indicating that individual VOCs are associated with more than one herbivore. Note that to generate the AV matrix, we assumed that if a herbivore can feed on a plant, it must be able to cope with spatial and temporal variation in VOCs that are used to distinguish suitable hosts despite the effects of these VOCs on the herbivore (attract, deter, or neutral). Altering either the PV or AP matrix can change the AV matrix.

We built the conceptual framework for our study based on a conflicting optimization process between plants and herbivores (see Fig. 2). Specifically, we hypothesized that plants aim to decrease the decoding efficiency of herbivores by changing PV associations, whereas herbivores aim to increase this efficiency by changing AP interactions. Ecologically, the decoding strategy can be linked to the fitness of plants and herbivores. A high decoding efficiency can increase the attack rates and decrease the fitness of plants. By contrast, low efficiency can increase the searching time and decrease the fitness of herbivores (12). Formally, we define the fitness relationships as \( F_p = H(A|V) \in [0,1] \) and \( F_s = 1 - H(V|A) \in [0,1] \), respectively. That is, plant fitness is proportional to the conditional entropy (uncertainty) between VOCs and herbivores, whereas herbivore fitness is negatively proportional to the conditional entropy between herbivores and VOCs. The supplementary materials (32) provide details on the calculations.

To track the evolutionary trajectory of the communication system between plants and herbivores, the coding strategy by plants is characterized by the PV association matrix (green right box). The AP interaction matrix (blue left box) is used to infer how plants’ codes are decoded by herbivores. These matrices undergo mutations by changing any of their elements (e.g., red numbers); zeros and ones correspond to absence and presence, respectively. These mutations will fixate only if the fitness relationship of the corresponding species increases. The fitness relationship of species is defined by herbivores’ decoding efficiency based in the communication system, i.e., the AV matrix that results from the product of AP and PV matrices. Specifically, we defined plant and herbivore fitness relationships by the expressions \( F_p = H(A|V) \) and \( F_s = 1 - H(V|A) \), where \( H(\cdot) \) corresponds to the conditional entropy. To increase this fitness relationship, plants and herbivores go through an alternating optimization process, mimicking their arms race. (For more details about this process, see the main text and the supplementary materials (32).)
herbivores based on our hypothesis, we simulated the optimization process above in the following way. Plants and herbivores can increase their fitness relationships (based on the $AV$ matrix) by modifying the interactions in the $PV$ matrix and the $AP$ matrix, respectively (see Fig. 2).

These modifications come from random mutations of a given number of elements in the $PV$ and $AP$ matrices. We assumed an equal mutation rate for each link ($PV$ link or $AP$ link) in the absence of any prior knowledge. Therefore, the number of random mutations that plants and herbivores could have for each round was proportional to the number of all possible links (i.e., the number of elements in the $PV$ and $AP$ matrices). Mutations are only selected if they increase the corresponding fitness relationships. To mimic a continuous arms race, we only allowed changes by plants and herbivores in an alternating fashion (see Fig. 2). Both the $PV$ and $AP$ matrices can be initialized from any random configuration with the only restriction being matrix size. To make our simulation as simple as possible, we assumed that plants can potentially emit any VOC and herbivores can potentially eat any plant without any phylogenetic or physiological constraints.

We investigated the capacity of our hypothesized mechanism to explain both the (de)coding strategies and the information structure of chemical communication between plants and herbivores observed in the field. Specifically, we determined the ability of our model to explain the conditional entropies $H(P|V)$, $H(A|P)$, and $H(A|V)$, which characterize the $PV$, $AP$, and $AV$ matrices, respectively. Then, we determined its ability to explain the observed fitness relationships $F_P$ and $F_A$. Note that conditional entropy by definition is the average entropy given each specific signal [i.e., individual VOC in the case of $H(P|V)$ and $H(A|V)$]. Furthermore, we tested how accurately our model could generate the cumulative information structure given a combination of VOCs by calculating the mutual information between plants and VOCs [$I(P,V)$] and between herbivores and VOCs [$I(A,V)$] as a function of the number of VOCs.

Additionally, we compared our proposed optimization mechanism against three alternative optimization mechanisms. The fitness relationships of plants and herbivores depend on their capacity to (i) maximize uncertainty [i.e., $F_P = H(A|V)$ and $F_A = H(V|A)$], (ii) minimize uncertainty [i.e., $F_P = 1 - H(A|V)$ and $F_A = 1 - H(V|A)$], and (iii) minimize and maximize uncertainty [i.e., $F_P = 1 - H(A|V)$ and $F_A = H(V|A)$]. We also ran all our simulations over different initial conditions, ranging from highly specialized to highly generalized matrices. The rationale for using these alternative mechanisms and initializations was to illustrate that potential matches between our theoretical expectations and the observed values are not just an artifact of sample size, metrics, any given optimization process, or the high (low) connectivity observed in the $PV$ ($AP$) matrix. The supplementary materials (32) provide all the details about these additional analyses.

Because sampling bias is an important concern in most ecological studies (34), we tested whether our results would be affected by the sampling effort in reporting species, by an incomplete plant VOC profile, and by sampling the interaction network in a different year or place. We found that all of our results were qualitatively equivalent when using

---

**Fig. 3.** Simulated results based on the information arms race theory recover the observed information structures in the field. (A) Simulation over 1000 time steps derived from our proposed optimization mechanism (information arms race). The simulation is randomly initialized. The circles, triangles, and squares correspond to the conditional entropies $H(\text{V}|A)$, $H(P|V)$, and $H(A|P)$, respectively. The solid, dotted, and dashed lines stand for the corresponding conditional entropies from field observations. (B) For the same simulation in (A), the + and × symbols correspond to the fitness relationships of plants ($F_P$) and herbivores ($F_A$), respectively. The solid lines correspond to field observations. (C) Summary of the observed and simulated values (mean ± SD) based on the last 250 steps (the rectangle window in (A) when reaching equilibrium for our proposed mechanism, shown in red; plants aim to maximize ($\text{Max}_P$) and herbivores aim to minimize ($\text{Min}_P$) uncertainty of information; see Fig. 2) and the corresponding ones for the three alternative mechanisms (in green, yellow, and blue; the alternative combinations of the maximization and minimization).
subsamples from the whole dataset (illustrating the scalability of our findings), when changing the VOC profile from 31 to 52 by including potential undefined biogenic VOCs, and when using information of plant-herbivore interaction networks from previous years (from 2007 to 2017) (35). The supplementary materials (32) provide all the details about the robustness of our results to changes in community size, number of VOCs, and identity and number of species interactions.

Figure 3 shows that the proposed conflicting information process is able to explain the coding and decoding patterns observed in the field. In particular, Fig. 3A shows that the conditional entropies $H(P|V)$ (triangles), $H(A|P)$ (squares), and $H(A|V)$ (circles) remain bounded across the simulated time to values close to those observed (see dashed, dotted, and solid lines, respectively). Similarly, Fig. 3B shows that the optimized fitness relationships of plants $F_P$ (+ symbols) and herbivores $F_E$ (× symbols) quickly converge to a stable value close to that observed (solid lines). Figure 3C shows that the alternative optimization mechanisms are not able to explain all the patterns.

Additionally, Fig. 4 shows that of our four tested models, only our proposed conflicting optimization process can recover the information structure observed in the field. Recall that we measured the cumulative information structure by the mutual information $I(P,V)$ where values closer to one imply that the presence or absence of a VOC (or a combination of VOCs) gives more information or has higher probability to tell all the species apart. In fact, the field data revealed that as few as eight VOCs can tell all the plants apart and 15 VOCs can give 97% information about herbivores (i.e., have a 97% probability of correctly telling all the herbivores apart). This structure was closely recovered by our proposed optimization mechanism (nine VOCs yield all the information of plants and 14 VOCs yield 97% of information of herbivores; Fig. 4, red lines). By contrast, the three alternative optimization mechanisms yield very different information structures, where either a large number of VOCs would be needed to uniquely identify species or the identification from VOCs would be less likely. For example, in the case where both parties aim to maximize uncertainty (i.e., confuse their opponents as in a mutually competitive relationship: Fig. 4, blue lines), even by using the combination of all VOCs, it would be impossible to tell all the species apart. All these results are robust to changes in community size, number of VOCs, the identity and number of species interactions, and different initial conditions (see the supplementary materials (32) and figs. S1 to S6).

Although a few recent studies have begun to demonstrate chemical patterns associated with plant-insect communities (24, 31, 36), our study proposes a plausible theoretical framework that explains and recovers patterns of information transfer between plants and herbivores from empirical data. Our work is hypothesis driven and suggests that an information arms race between plants and herbivores can drive plants to produce VOCs that are commonly shared by other species, increasing the difficulty for herbivores to identify suitable plants and potentially pressuring herbivores to specialize on a few plants (as the $AP$ matrix shows in Fig. 1). Indeed, previous studies have shown a substantial overlap among the VOCs emitted by different plant species (31, 37, 38). However, the functional roles of a large number of plant VOCs have been unexplained and overlooked because most mechanistic studies have focused on specific plant species and on the ecological functions of only a few VOCs (32). Here, our information arms race theory suggests that these seemingly redundant VOCs play an important role in confusing herbivores at the community level. Focusing on herbivore-plant interactions, herbivores are commonly regarded as specialists (6), which was also supported by our theory. In fact, it is estimated that <10% of all herbivores feed on plants that span more than three families (39).

Our findings further suggest that a conflicting information process drives the rapid accumulation of information by adding VOCs (Fig. 4), contrary to other optimization processes. This provides additional evidence for an arms race process explaining the large diversity of VOCs in nature, where herbivores have the evolutionary potential to quickly tell all plant species apart by making use of the few most informative VOCs, and plants can in turn respond to this potential by adding more VOCs to their profile. Under the same process, herbivores themselves can also be identified using a set of informative VOCs (Fig. 4B).

This raises the question of how these VOC profiles result in evolutionary trade-offs regulating the attraction of herbivores and their predators and parasitoids (17, 18). Overall, our study suggests that information transfer processes are key drivers of the formation and maintenance of species interactions across trophic levels.

REFERENCES AND NOTES

33. See the supplementary materials.


ACKNOWLEDGMENTS
We thank I. S. Aranda and R. Moore for help with fieldwork; G. Kite for help with gas chromatography–mass spectrometry; R. Pérez-Ishiwara for logistical support; R. Halitschke at the Max Planck Institute for Chemical Ecology for providing silicon tubing for volatile sampling; M. Al-Adwani, S. Cenci, L. Medeiros, R. Pérez-Ishiwara for logistical support; R. Halitschke at the Max Planck Institute for Chemical Ecology for providing silicon tubing for volatile sampling; M. Al-Adwani, S. Cenci, L. Medeiros, J. Tumlinson, M. Erb, and Biodiversity. Author contributions: This work was funded by the Max Planck Society, the NOMIS Foundation, and the University of Zurich Research Priority Program on Global Change and Biodiversity. Author contributions: P.Z. and S.S. designed the study; P.Z. and E.d.-V. collected field data; K.B., E.d.-V., and A.Z.-R., identified herbivores and constructed plant-herbivore networks; P.Z., M.C.S., and P.C.S. performed chemical analyses. P.Z. and S.S. performed theoretical analyses. P.Z. and S.S. wrote the manuscript with contributions by M.C.S. and all other coauthors. Competing interests: The authors declare no competing interests. Data and materials availability: All source data and R code needed to reproduce the results are available in the supplementary materials (32) and on Zenodo (40).

SUPPLEMENTARY MATERIALS
science.sciencemag.org/content/368/6497/1377/suppl/DC1
Materials and Methods
Supplementary Text
Table S1 to S6
References (41–46)
Database S1
MDAR Reproducibility Checklist
5 December 2019; resubmitted 25 February 2020
Accepted 15 April 2020
10.1126/science.aba2965
A plant-herbivore information "arms race"

The consumption of plants by herbivores has driven the evolution of many diverse plant defense chemicals to which herbivores have constantly adapted. The transmission of chemical information at the community level is less known but important given the plethora of plant and herbivore species, especially in tropical communities. Zu et al. propose an information "arms race" approach to explain plant-herbivore chemical communication at the community level (see the Perspective by Solé). To test their conceptual framework, they used field data of herbivore-plant interactions and plant volatile organic compound associations in a tropical dry forest. Their approach provides an understanding of the functioning and persistence of systems where individuals send and receive information in the form of signals to which other individuals react and, in turn, affect the behavior of other participants in these systems.

Science, this issue p. 1377; see also p. 1315