Crop pests and predators exhibit inconsistent responses to surrounding landscape composition

Daniel S. Karp[a,1], Rebecca Chaplin-Kramer[b], Timothy D. Meehan[c], Emily A. Martin[d], Fabrice DeClerck[e], Heather Grab[f], Claudio Gratton[g], Lauren Hunt[h], Ashley E. Larsen[i], Alejandra Martínez-Salinas[j], Megan E. O'Rourke[k], Adrien Rusch[l], Katja Poveda[f], Mattias Jonsson[m], Jay A. Rosenheim[n], Nancy A. Schellhorn[o], Teja Tscharntke[p], Stephen D. Wratten[q], Wei Zhang[r], Aaron L. Iverson[f], Lynn S. Adler[s], Matthias Albrecht[t], Audrey Alignier[u], Gina M. Angelella[k], Muhammad Zubair Anjum[v], Jacques Avelino[w], Péter Batáry[p], Johannes M. Baveco[x], Felix J. J. A. Bianchi[y], Klaus Birkhofer[z], Eric W. Bohnenblust[aa], Riccardo Bommarco[m], Michael J. Brewer[bb], Berta Caballero-López[cc], Yves Carrière[dd], Luísa G. Carvalheiro[ee], Luis Cayuela[ff], Mary Centrella[f], Aleksandar Ćetković[gg], Dominic Charles Henri[hh], Ariane Chabert[ii], Alejandro C. Costamagna[jj], Aldo De la Mora[kk], Joop de Kraker[ll], Nicolas Desneux[mm], Eva Diehl[nn], Tim Diekötter[oo], Carsten F. Dormann[pp], James O. Eckberg[qq], Martin H. Entling[rr], Daniela Fiedler[ss], Pierre Franck[tt], F. J. Frank van Veen[uu], Thomas Frank[vv], Vesna Gagic[o], Michael P. D. Garratt[ww], Awraris Getachew[xx], David J. Gonthier[yy], Peter B. Goodell[zz], Ignazio Graziosi[aaa], Russell L. Groves[g], Geoff M. Gurr[bbb], Zachary Hajian-Forooshani[ccc], George E. Heimpel[ddd], John D. Herrmann[oo], Anders S. Huseth[eee], Diego J. Inclán[fff], Adam J. Ingrao[ggg], Phirun Iv[hhh], Katja Jacot[t], Gregg A. Johnson[qq], Laura Jones[o], Marina Kaiser[gg], Joe M. Kaser[ddd], Tamar Keasar[iii], Tania N. Kim[jjj], Miriam Kishinevsky[kkk], Douglas A. Landis[ggg], Blas Lavandero[lll], Claire Lavigne[tt], Anne Le Ralec[mmm], Debissa Lemessa[nnn], Deborah K. Letourneau[ooo], Heidi Liere[jjj], Yanhui Lu[ppp], Yael Lubin[qqq], Tim Luttermoser[f], Bea Maas[rrr], Kevi Mace[sss], Filipe Madeira[ttt], Viktoria Mader[nn], Anne Marie Cortesero[uuu], Lorenzo Marini[vvv], Eliana Martinez[www], Holly M. Martinson[xxx], Philippe Menozzi[yyy], Matthew G. E. Mitchell[zzz], Tadashi Miyashita[aaaa], Gonzalo A. R. Molina[bbbb], Marco A. Molina-Montenegro[cccc], Matthew E. O'Neal[dddd], Itai Opatovsky[eeee], Sebaastian Ortiz-Martinez[lll], Michael Nash[ffff], Örjan Östman[gggg], Annie Ouin[hhhh], Damie Pak[iiii], Daniel Paredes[jjjj], Soroush Parsa[kkkk], Hazel Parry[o], Ricardo Perez-Alvarez[f], David J. Perović[bbb], Julie A. Peterson[ddd], Sandrine Petit[llll], Stacy M. Philpott[ooo], Manuel Plantegenest[mmm], Milan Plećaš[gg], Therese Pluess[mmmm], Xavier Pons[ttt], Simon G. Potts[ww], Richard F. Pywell[nnnn], David W. Ragsdale[oooo], Tatyana A. Rand[pppp], Lucie Raymond[mmm], Benoît Ricci[llll], Chris Sargent[h], Jean-Pierre Sarthou[qqqq], Julia Saulais[mmm], Jessica Schäckermann[rrrr], Nick P. Schmidt[dddd], Gudrun Schneider[d], Christof Schüepp[mmmm], Frances S. Sivakoff[ssss], Henrik G. Smith[tttt], Kaitlin Stack Whitney[uuuu], Sonja Stutz[vvvv], Zsofia Szendrei[ggg], Mayura B. Takada[wwwww], Hisatomo Taki[xxxx], Giovanni Tamburini[m], Linda J. Thomson[yyyy], Yann Tricault[zzzz], Noelline Tsafack[aaaaa], Matthias Tschumi[t], Muriel Valantin-Morison[bbbbb], Mai Van Trinh[ccccc], Wopke van der Werf[ddddd], Kerri T. Vierling[eeeee], Ben P. Werling[fffff], Jennifer B. Wickens[ww], Victoria J. Wickens[ww], Ben A. Woodcock[nnnn], Kris Wyckhuys[ggggg,hhhhh], Haijun Xiao[iiiii], Mika Yasuda[jjjjj], Akira Yoshioka[kkkkk], and Yi Zou[lllll]

The idea that noncrop habitat enhances pest control and represents a win–win opportunity to conserve biodiversity and bolster yields has emerged as an agroecological paradigm. However, while noncrop habitat in landscapes surrounding farms sometimes benefits pest predators, natural enemy responses remain heterogeneous across studies and effects on pests are inconclusive. The observed heterogeneity in species responses to noncrop habitat may be biological in origin or could result from variation in how habitat and biocontrol are measured. Here, we use a pest-control database encompassing 132 studies and 6,759 sites worldwide to model natural enemy and pest abundances, predation rates, and crop damage as a function of landscape composition. Our results showed that although landscape composition explained significant variation within studies, pest and enemy abundances, predation rates, crop damage, and yields each exhibited different responses across studies, sometimes increasing and sometimes decreasing in landscapes with more noncrop habitat but overall showing no consistent trend. Thus, models that used landscape-composition variables to predict pest-control dynamics demonstrated little potential to explain variation across studies, though prediction did improve when comparing studies with similar crop and landscape features. Overall, our work shows that surrounding noncrop habitat does not consistently improve pest management, meaning habitat conservation may bolster production in some systems and depress yields in others. Future efforts to develop tools that inform farmers when habitat conservation truly represents a win–win would benefit from increased understanding of how landscape effects are modulated by local farm management and the biology of pests and their enemies.

agroecology | biodiversity | biological control | ecosystem services | natural enemies

Ecologists increasingly consider maintaining or restoring non-crop habitat in the landscapes surrounding farm fields to be a

Significance

Decades of research have fostered the now-prevalent assumption that noncrop habitat facilitates better pest suppression by providing shelter and food resources to the predators and parasitoids of crop pests. Based on our analysis of the largest pest-control database of its kind, noncrop habitat surrounding farm fields does affect multiple dimensions of pest control, but the actual responses of pests and enemies are highly variable across geographies and cropping systems. Because noncrop habitat often does not enhance biological control, more information about local farming contexts is needed before habitat conservation can be recommended as a viable pest-suppression strategy. Consequently, when pest control does not benefit from noncrop vegetation, farms will need to be carefully comanaged for competing conservation and production objectives.

win–win for farmers and for conservation. Beyond well-documented benefits for crop pollinators (1), surrounding noncrop habitat is also generally thought to benefit predators and parasitoids and thereby enhance biological control of crop pests. Indeed, both qualitative (e.g., refs. 2 and 3) and quantitative (e.g., refs. 4 and 5) syntheses have concluded that, on average, enemy abundance, diversity, and activity (pest consumption) increase in landscapes with more noncrop habitat. The benefits of noncrop habitat are likely multifaceted: Just as for pollinators (1), noncrop habitat can provide enemies with supplemental food resources, nesting locations, and/or overwintering sites (2, 4, 6). However, the assumption that more noncrop habitat in farming landscapes universally increases biocontrol belies significant variation among studies in how enemies respond to landscape composition (i.e., the relative proportions of crop and noncrop land-use types in the landscapes surrounding farm fields) (4, 7). Moreover, the effect of landscape composition on pests themselves remains inconclusive, as many crop pests also benefit from nearby noncrop habitat (4, 7). As a result, programs focused on improving the conservation value of farming landscapes could, in some cases, precipitate increased yield losses to crop pests (7).

The great variation among studies in pest and enemy responses, coupled with a research bias for studies in temperate systems, has also precluded our ability to use relationships derived in prior quantitative syntheses to model pest control globally. As a result, despite presumed significant contributions to agricultural production—estimated at approximately US$4.5 billion/y in avoided crop damage in the United States alone (8)—and repeated recognition in ecosystem-service frameworks (9, 10), natural biological control is rarely taken into account in landscape-level planning and environmental decision making. Existing decision-support tools and models focus on pest-control dynamics as a function of on-farm activities (11, 12). The few landscape-level models that do exist and could inform decision making require a thorough understanding of the life histories of target pests and their natural enemies (e.g., attack rates, growth rates, etc.) (13–15). At present, such models can be applied only by experts in very data-rich environments and may not be generally applicable.

Because there is no established standard for measuring either pest control or landscape composition, it is possible that the diverse responses of pests and enemies to surrounding landscape composition result from researchers measuring pests, enemies, their interactions, and the landscapes they inhabit in different ways. Alternatively, the heterogeneity in species responses could be real, resulting from shifting interactions between different species of pests and enemies with varying traits in different landscapes. Here we standardize, reanalyze, and quantitatively synthesize published and unpublished studies to (*i*) assess the role of surrounding noncrop habitat in providing pest-control

services to farmers, and (*ii*) determine to what extent landscape information alone can be used to model and predict variation in pest control across systems.

First, we compiled an extensive pest-control database composed of 18,094 total observations of pest and enemy abundance, enemy activity (predation rates), crop damage, and yields across 6,759 sites in 31 countries (Fig. 1). We then used a model-averaging approach to explore to what extent surrounding forest, grassland, scrubland, annual crops, and perennial crops (i.e., landscape composition), at multiple spatial scales, could explain spatial patterns in the pest-control variables recorded in each study. From prior observations (4, 7), we predicted that landscapes with more noncrop habitat would generally increase enemy abundance and activity (pest predation) but that pest abundance, damage, and crop yield would exhibit significant variation in how they respond to noncrop habitat across studies. Nevertheless, we predicted that landscape models would be able to explain variation in independent datasets, provided that the independent dataset and the dataset from which the landscape model was constructed shared geographic, crop, and/or landscape features.

Results

Pest-Control Models. We constructed separate models ($n = 359$) that explored how surrounding landscape composition affected each unique pest-control variable reported in each study (hereafter referred to as unique "pest-control responses"). We report a summary of the 359 pest-control responses, focal crops, target pests, and sample sizes (Dataset S1) and include the entire biological control database (Dataset S2). Pest-control responses were made up of censuses of dominant pests, all pests, and dominant enemies ($n = 76, 27,$ and 78 responses, respectively), exclosure and sentinel pest experiments that measured enemy activity ($n = 23$ and 68, respectively), crop-damage surveys that measured pest activity ($n = 34$), and crop-yield data ($n = 53$). Though 31 countries were represented, there was a strong temperate bias, with 39% of the 132 studies from Europe, 23% from North America, 14% from Central and South America, 12% from Asia, 8% from Australia and New Zealand, and 4% from Africa. Similarly, though the database encompassed many distinct crops ($n = 41$ crops, 38 species), the vast majority of sites were in annual (90%) versus perennial (10%) agriculture.

Overall, we found that there was significant variation among studies in the effects of different landscape variables on pest-control responses. Contrary to our first prediction, we observed no consistent effects of "natural" landscape variables (forest, grassland, scrubland; aggregate natural habitat category) or "agricultural" landscape variables (perennial crops, annual crops; aggregate crop category) on enemy and pest abundance and activity. That is, χ^2 tests indicated that surrounding noncrop habitat and cropland affected pest-control variables positively and negatively in a roughly equivalent number of models (all $P > 0.05$; Fig. 2). Effects of natural landscape variables were often even inconsistent within studies. For example, in the majority of enemy-abundance models, enemies responded positively to one natural land-use type (e.g., forest) but negatively to another (e.g., grassland). Restricting our focus to only the most predictive land-use variable present in each model did not change these conclusions. We still found that most pest-control variables responded positively and negatively to landscape variables in a roughly equivalent number of models (*SI Appendix*, Fig. S1; all $P > 0.05$). An exception was crop yields, which tended to decline with surrounding natural land uses more often than not ($\chi^2 = 9.8$, $P < 0.01$). Similarly, no single spatial scale tended to be most predictive, though more pest-abundance and enemy-activity (cage experiments) responses were best predicted by models that placed greater weight on areas closer to sampling sites (*SI Appendix*, Fig. S2).
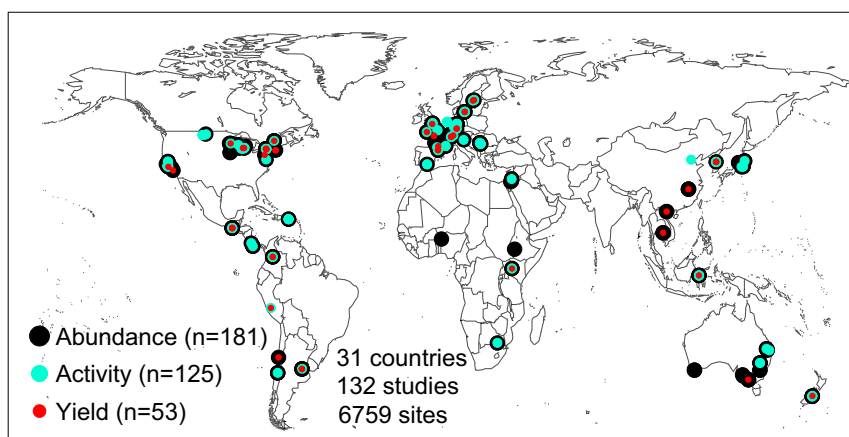
Fig. 1. Map of study locations. We collected pest-control data from 132 studies across 6,759 sites and 31 countries. Pest-control data included abundances of dominant pests, all pests, and natural enemies (black dots; 181 responses), pest and predator activity data from crop-damage surveys, sentinel pest experiments, and exclosure experiments (cyan dots; 125 responses), and yield data (red dots; 53 responses).

Model Performance. Despite observed heterogeneity, we did find that landscape composition explained variation in pest-control responses within studies (average R^2 0.14 to 0.20; Fig. 3 and *SI Appendix*, Table S1). The average correlation across individual studies between observed data and model predictions was significantly larger than 0, and ranged from 0.37 (all enemies) to 0.45 (sentinel pest experiments). With respect to individual datasets, correlations between model predictions and observed data were significant and positive ~50% of the time (*SI Appendix*, Table S1). Interestingly, model performance varied across biogeographic realms (*SI Appendix*, Fig. S3). Models tended to explain more variation in pest variables (pest abundance and damage) when studies were located in the Palearctic (42% of responses; $\chi^2 = 5.6$, $P = 0.02$) and, to a lesser extent, in the Nearctic (28% of responses; $\chi^2 = 3.0$, $P = 0.08$) compared with other areas. While landscape relationships with natural enemy variables (enemy abundance and activity) did not vary geo-

graphically (all $P > 0.05$), more variation in yield data was explained in the Nearctic than in other areas ($\chi^2 = 4.4$, $P = 0.04$).

To actually assess our models' predictive power, we examined how well observations in each dataset correlated with predictions generated from the full suite of models from the other (independent) datasets. As hypothesized, field observations did not correlate with average predictions across all independent models (Fig. 4 and *SI Appendix*, Table S2). That is, correlations between average model predictions and field observations were not statistically distinguishable from 0, which was not surprising given such high observed variation among studies in the directionality of landscape-composition responses.

Using Model Subsets to Improve Predictive Power. To improve predictive power, we used a full-factorial sensitivity analysis (*Methods*) to explore multiple scenarios of filtering and applying subsets of models to field observations, rather than using the
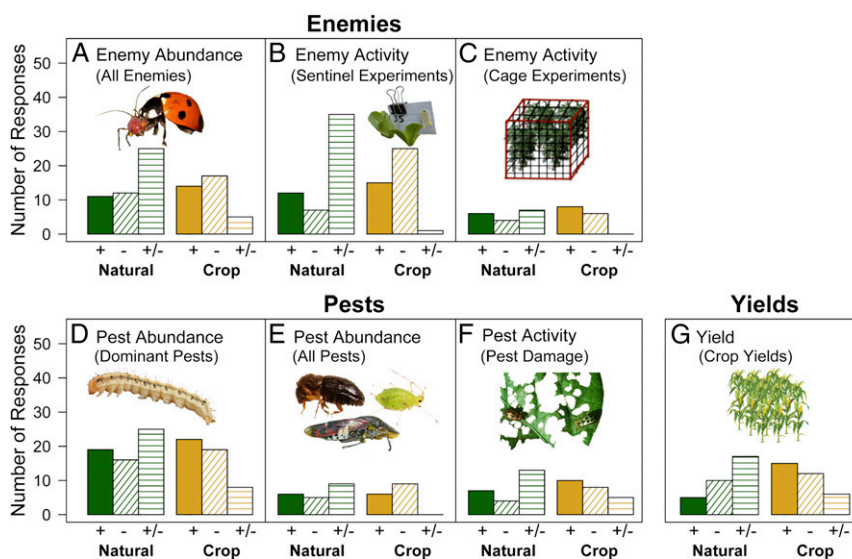
Fig. 2. Landscape effects on pest-control variables. After selecting the most predictive model for each pest-control response (N = 367) and redefining land-cover variables as natural (forest, grassland, and scrubland; green bars) versus crop (annual and perennial; orange bars), we tallied the number of pest-control responses for which models had either positive (solid), negative (diagonal hashed), or mixed (horizontal) estimates of the effect of each landscape predictor. Panels represent the seven pest-control variables, including abundance (*A*) and activity (*B* and *C*) of natural enemies; abundance (*D* and *E*) and activity (*F*) of pests; and crop yields (*G*). χ^2 tests indicated that pest-control response variables showed heterogeneous patterns of association with the extent of surrounding natural habitat and cropland—with roughly equivalent numbers of pest-control responses having models with positive and negative effects (all $P > 0.05$).

Fig. 3. Explanatory power of landscape pest-control models. After selecting the most predictive spatial scale (*Methods*), model predictions were correlated with observed data. Gray dots are both Pearson correlations between model predictions and observed data and R^2 values (square of Pearson's r). Filled circles and empty circles indicate significant ($P < 0.05$) and nonsignificant correlations, respectively. Black dots indicate the mean correlation across all datasets between observed and predicted values. Black lines correspond to 95% confidence intervals.

average predictions generated from all landscape models (*SI Appendix*, Table S3). While this approach suggested several model-selection methods that improved pr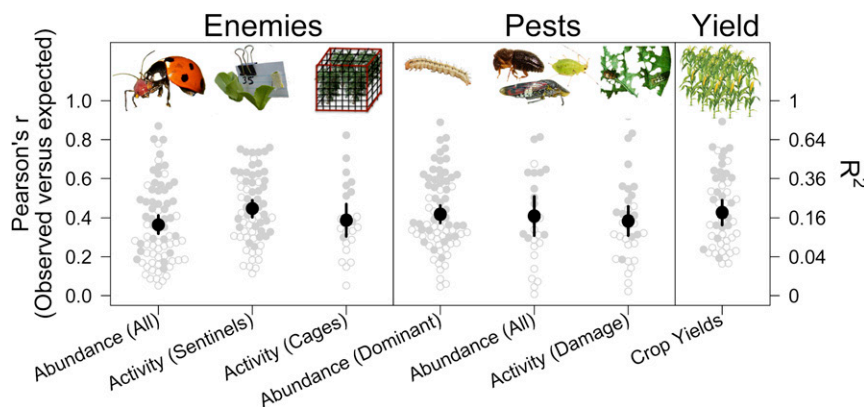edictive power, correlations between model predictions and field observations remained low. For example, correlations between field observations and average model predictions tended to be higher when we only used models constructed from studies that focused on the same crop as the field observations. Prediction was also enhanced when we only included models that shared the same land-cover variables as those present around the field observations or, even more stringently, when land-cover variable values associated with the field observations were within the range of values present in the model data (i.e., the landscape-composition gradients matched). By strategically selecting models from studies that better matched the field observations, we could increase correlations between average model predictions and observed pest-abundance, enemy-abundance, and enemy-activity data such that correlations were, on average, positive across all field observations (Fig. 4, *Bottom* and *SI Appendix*, Table S2). Critically, however, model predictions rarely explained significant variation among individual pest-control responses: In ~12 and ~5% of responses, correlations between field observations and model predictions were significantly positive and negative, respectively.

Explorations of other strategies for improving predictive power yielded mixed results (*SI Appendix*, Table S3). Only using models from studies in the same biogeographic realm as the field observations marginally improved predictive power for pest-damage responses but reduced predictive power for all other pest-control variables. When models and field observations shared the same biome, predictive power significantly increased for some variables (e.g., enemy activity measured in cage experiments) but decreased for others (e.g., enemy activity measured in sentinel experiments, crop yield). Similarly, predictive power sometimes increased (e.g., all pests, crop yield) and sometimes decreased (e.g., dominant pests, all enemies, sentinel experiments, cage experiments) when we only used models from studies that were located within 500 km of field observations.

Discussion

Despite heterogeneity in environmental conditions, farm management, and species assemblages, we found that every dimension of pest control, from enemy abundance to crop yields, tended to show some associations with the surrounding landscape composition. Our work thus confirms prior observations that habitat composition of the broader farmscape can modulate on-farm pest control (2–4). Interestingly, explanatory power varied

geographically. For example, models from Nearctic and Palearctic realms generally explained more variation in pest abundance and activity than models from other areas. This difference may result from latitudinal variation in data quality. Fewer regional land-use maps exist in data-poor tropical areas, and informal inspection of the global land-use map suggested that land-cover classifications were often inaccurate in tropical landscapes.
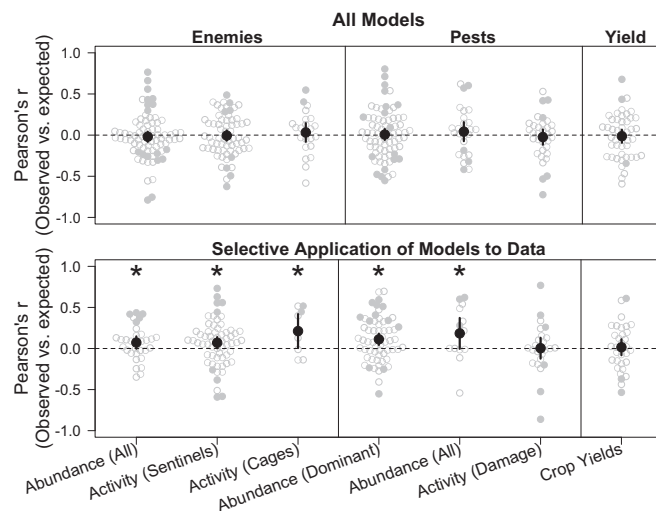
Fig. 4. Testing landscape models. (*Top*) Correlating average predictions across all possible landscape models (*Methods*) against independent field observations resulted in low predictive power. Each gray circle is the observed correlation for one dataset (set of field observations); filled circles are significant correlations ($P < 0.05$). Black circles are average correlations across all tested datasets; lines are confidence intervals. (*Bottom*) More selective application of models to independent field observations caused correlations to be on average positive for all pest-control variables except pest damage and crop yields (asterisks indicate $P < 0.05$). Specifically, this panel demonstrates that predictive power was higher when a more selective subset of models was applied to the independent field observations, subject to several of the following constraints: (*i*) Field observations and the data from which models were constructed (model data) shared the same crop; (*ii*) the same land-cover variables were present in model data and field observations; (*iii*) landscape values in field observations were within the range of landscape values in the model data; and (*iv*) models explained significant variation in their own data ($r > 0.25$). Dominant pests, pest damage, and crop yields were subject to constraints (*i*) and (*ii*); all pests to (*i*), (*ii*), and (*iii*); sentinel experiments to (*i*), (*ii*), and (*iv*); and all enemies and cage experiments to (*i*), (*ii*), (*iii*), and (*iv*).

Though landscape-composition variables often explained variation in pest-control responses within studies, we observed substantial variation among studies in how pests, enemies, predation rates, crop damage, and yields responded to different landscape variables. Specifically, we did not observe a consistent increase or decrease in response to surrounding noncrop habitat in any variable. Although increased biological control can cause pests to decline in landscapes with more noncrop habitat, pests too can benefit from noncrop habitat (7). This variability also pertained to enemies, despite meta-analyses showing that enemy abundance (2, 4) and activity (5) increase in landscapes with more noncrop habitat. Key to this finding is acknowledging that meta-analyses take the effect sizes reported by individual studies (often including interactions with other variables and other idiosyncrasies in analysis), not in the standardized, uniform way presented here that isolates the effect of noncrop habitat. These results contrast prior pollinator studies, which tend to report more consistently positive effects of noncrop habitat (1, 16). One explanation for this difference may be that the greater diversity of organisms involved in pest control (e.g., birds, bats, spiders, beetles, flies, etc.) may underlie their more diverse landscape responses. Even more importantly, complex, tritrophic interactions between enemies, pests, and crops may make for more context-specific dynamics than in comparatively simpler pollinator–plant systems.

Sources of Variation. Ultimately, such context dependence in pest and enemy responses to landscape composition made it difficult to generalize effects of land use-change decisions on pest control. While we did identify some strategies that could marginally increase predictive power (e.g., by only applying models to datasets focused on the same crop), models were still only able to predict very limited amounts of variation in pest-control variables when applied to independent datasets. Surprisingly, predictive power did not even increase when we only used models that were built from studies near to (<500 km), in the same biome as, or in the same biogeographic realm as field observations. Below, we detail five possible explanations for the remarkable variation that we observed in landscape responses; the first two focus on potential problems with underlying datasets, and the latter three focus on unaccounted-for sources of variation that could be addressed in future analyses.

First, discordance between our findings and the broader pest-control literature (4) may reflect underlying publication biases. Prior reported trends may lack reproducibility because authors are incentivized to continue interrogating their data and conducting new experiments until they find significant results (17). We do not find this explanation particularly likely, however, because prior meta-analyses do report significant cross-study variation in pest responses (4).

Second, our broad quantitative synthesis may have missed key temporal dynamics that may influence landscape effects but are rarely measured in pest-control studies. Specifically, if predator and pest abundances fluctuate and respond to each other in an oscillatory system (18), then snapshot estimates of species abundances or interactions may not reflect the longer-term impact of surrounding landscape composition. For example, one study found that, near natural habitats, syrphid fly larvae suppress week-to-week population growth of aphids, contributing to lower densities at harvest, but aggregating data into annual average abundances masked landscape effects (19).

Third, pests and enemies may respond more strongly to aspects of the surrounding landscape that we did not measure, such as landscape configuration (i.e., the spatial arrangement of surrounding land-use types) or diversity (i.e., the number of surrounding land-use types present) (7, 20, 21). Still, our approach of analyzing the proportional area of different land-use types represents the most common landscape metric used in pest-control studies (4, 20).

Fourth, discordance with prior findings may be attributable to additional sources of variability tested in the original studies but not addressed by this analysis. Indeed, landscape composition by itself explained on average only 14 to 20% of the variation in pest-control variables, leaving substantial variation unaccounted-for after correlating model predictions with field observations. Landscape-composition effects could easily be present but masked by variation in local farm-management factors or local climate variables that published studies usually treat as covariates. Indeed, crop diversification, insectary plantings, and insecticide treatments all influence pest control and, in some cases, can modify landscape effects (22–24).

Finally, different pest and enemy species often respond differently to landscape variables (7). Because most of the variables we modeled were not species-specific (e.g., abundance of all possible pests), differences in the species that were modeled versus the species present in the tested datasets likely lessened the accuracy of our predictions. Indeed, while many natural enemies need noncrop habitats to complete their life cycles (6), some enemies are more dependent on cropland than they are on noncrop vegetation (7). Moreover, if pests increase in simple landscapes, then enemies may increase too, as they start to exploit this burgeoning resource.

A Path Forward. If we hope to move toward more sustainable land and pest management, decision makers need to be informed about tradeoffs among multiple ecosystem services of their land-use decisions. For example, a variety of stakeholders interested in agricultural development are concerned with the consequences of habitat loss that accompany land-use change, including governments (25, 26), multilateral development banks (27), and corporations (28). These actors have typically been focusing on impacts of agriculture on biodiversity, carbon storage, and water, but agricultural expansion or intensification also has massive implications for ecosystem services to agriculture, like pest suppression (29, 30).

High cross-system variability impeded this initial attempt to develop simple landscape models that could deliver accurate pest-control predictions, but has also opened the door to new avenues of exploration. We suggest three promising paths forward that leverage our newly created, open-access pest-control database.

First, future efforts could increase model complexity, incorporating landscape configuration and diversity predictors as well as indicators of local farming practices. A key consideration will be choosing variables for which sufficient data exist to model pest control in data-poor environments. Fortunately, some global, wall-to-wall datasets that could be used to characterize local farming practices already exist and others are being developed. For example, the International Institute for Applied Systems Analysis has created a global crop field size dataset (31), and EarthStat maintains a global crop-composition dataset (32) that can be used to construct crop-diversity indices.

Second, models could be made more dynamic to account for temporal changes in predator and pest relationships. Many studies in our database resurveyed pests, enemies, and their interactions multiple times and averaged them throughout the growing season. Indeed, few published studies have reported relationships between landscapes and insects at multiple time intervals, despite evidence that these relationships can change both between years and during the growing season (33). As noted above, aggregate measures of pest or enemy abundance may mask important effects on pest growth or densities at harvest (19).

Third, including trait data as species-level covariates in pest-control models could help identify groups of species that react similarly to landscape variables. Traits have been used to explain variation in tropical bird community responses to land-use intensification worldwide (34), and similar approaches may help make sense of the context dependence of these results based on

life history, mobility, habitat preference, and other characteristics (35). For example, specialist pests that depend on crop resources are likely to exhibit more negative associations with surrounding noncrop vegetation than generalist pests (36). Such trait-driven models may be especially important when using future predictive models to identify tradeoffs, for example, exploring when surrounding noncrop habitat would depress one pest's abundance but likely benefit another. An alternative approach to increasing predictive power would be redefining land-use types into more functionally relevant classifications that are tailored to specific pests and/or enemies (37). For example, with sufficient natural history information, landscapes could theoretically be classified into species-specific foraging, nesting, roosting, and unused sites.

Conclusions

By compiling and systematically analyzing the largest pest-control database of its kind to date, we have demonstrated that landscape composition alone can explain variation in the abundance and activity of natural enemies and crop pests. Critically, however, we found remarkable variability in how pests and enemies respond to different landscape metrics, preventing the prediction of pest-control responses by simple empirical models constructed from independent datasets. Ultimately, these results suggest that natural habitat conservation cannot be considered a panacea. While habitat conservation is known to enhance pest control in many cropping systems globally (4), there are also areas where conservation may bolster biodiversity—even of natural enemies—but still reduce crop production by simultaneously enhancing pests. Our study thus highlights the need to examine multiple services (e.g., pest control, pollination, soil conservation) and disservices when designing or assessing policies such as Europe's agrienvironment schemes, which incentivize habitat creation in farmland (38).

Looking forward, generalizable decision-support tools for pest control are needed to help farmers understand when habitat conservation represents a true win–win and when conservation activities will need to be carefully comanaged to reduce the risk of damaging pest outbreaks. Future efforts to develop such models could recognize and account for system-specific variability by incorporating local management factors, temporal variation, and life-history variation among species into predictive models. To this end, we envision the supporting pest-control data that we have assembled to be a living database. Through ongoing efforts to address species traits, compile data on local management practices, and develop infrastructure to continue adding new studies that focus on different crops and geographies, the reach of our approach will continue extending to novel contexts, increasing predictive power and offering remarkable opportunities to answer critical questions about the ecology of pest control.

Methods

Database Compilation. We formed a working group of 19 pest-control experts and practitioners to compile a database of spatial pest-control observations. First, we developed a list of potential data contributors, leveraging past syntheses (2–5), literature searches, and our professional networks. Potential contributors were contacted and asked to complete a data-entry form. Studies were included if pest-control observations were obtained across at least five distinct sampling locations (mean 50 sites), all within crop fields and across a gradient of surrounding landscape composition. Authors were required to report the following information for each sampling site: spatial coordinates, crop type, farm name, and study year. For predator, parasitoid, and pest-abundance data, we also requested information on species taxonomy, sampling dates, sampling methods (pan trap, pitfall trap, sweep net, etc.), and the number and duration of censuses. We also required each author to identify each animal as a dominant (economically important) pest, secondary pest, predator, or parasitoid.

Pest activity was measured as the fraction or amount of each crop consumed, infested, or damaged. Predator activity was measured through three

types of field observations. First, some researchers reported data from experiments in which pests were collected, placed in chambers, and monitored to quantify the fraction of emerging parasitoids. Second, in sentinel pest experiments, researchers placed pests in crop fields and returned later to measure the fraction that had been parasitized or consumed (e.g., ref. 39). Third, researchers used field exclosures to exclude natural enemies and quantify differences in pest abundances and/or crop damage between plants with and without natural enemies (e.g., ref. 5). Activity data were accompanied by information on the date, duration, type, and number of in-field experiments. Crop yield data were also collected if available. Data on total crop quantity or quality were supplemented with information regarding crop species, measure type (fruit weight, biomass/area, etc.), and crop quality (marketable versus total yield). Before analysis, all pest-control data were standardized to increase comparability across diverse data types (see *SI Appendix* for more details).

Landscape Composition. To quantify surrounding landscape composition, we used a hierarchical approach to acquire land-cover maps that encompassed all areas within 2 km of each study site (*SI Appendix*). Specifically, we asked authors to submit high-resolution land-use maps of their study regions when available; otherwise, we used either regional maps or, as a last resort, a 30-m global land-cover product (40). Land-cover maps were classified into seven categories: (*i*) forest and tree plantations, (*ii*) grassland, (*iii*) scrubland, (*iv*) annual cropland, (*v*) perennial cropland, (*vi*) urban areas, and (*vii*) other. We then used a distance-weighting function (with multiple spatial scales) to quantify landscape composition around each study site and value each landscape sector as a function of distance to the study site (15, 39) (*SI Appendix*).

Model Averaging. We modeled pest-control data as a function of landscape-composition variables, creating separate models for each pest-control response (i.e., each unique pest-control variable measured in each study). Because landscape variables were often highly collinear within studies, we first iteratively excluded individual landscape variables until variance inflation factors (VIFs) were less than 2.5, dropping the landscape variable with the highest VIF first (41). Subsequent reanalysis that included all variables, regardless of collinearity, did not indicate any systemically higher predictive power. We then used a model-averaging procedure in which the full suite of all possible landscape models (i.e., every combination of landscape predictors) was used to quantify effects of each landscape-composition variable on the pest-control responses (MuMIn package in R). Specifically, nonshrinkage variance estimates were obtained for each fixed effect, where only models that included a given fixed effect contributed to its ultimate predicted value. For each pest-control response variable, all landscape variables associated with one spatial scale were included as fixed effects. Different spatial scales were analyzed separately. Along with landscape variables, we also included sampling method, study year, and crop type as fixed effects. For datasets in which crop yields were measured at multiple times throughout the growing season, we also included "elapsed time since planting" as an additional fixed effect. In all cases, "farm identity" was included as a random effect when multiple sampling sites were situated on the same farm.

Pest-control variables were often log-, square root-, or fourth root-transformed to abide by model assumptions (normality and homoscedasticity). Transformations did not resolve normality and heteroscedasticity concerns in 47 and 63 of 367 analyses, respectively. Subsequent sensitivity analyses indicated that excluding those models did not affect any of our conclusions or systematically improve our ability to predict variation in pest-control responses in independent datasets (see below). Therefore, these models were retained.

To determine whether pest-control variables exhibited consistent responses to landscape composition, we first reclassified landscape variables into natural habitat variables (forest, grassland, or scrubland) and crop variables (annual or perennial crops). Then, because three models (one for each spatial scale) were constructed for each dataset, in this and every subsequent analysis we restricted our focus to the one spatial scale in which landscape predictors explained the most variation. Next, we tallied the number of models in which natural habitat variables or crop variables were predicted to have consistently positive, consistently negative, or mixed effects on each pest-control variable. For example, if pest abundance increased with surrounding forest cover but decreased with grassland cover, then we categorized pest abundance as exhibiting mixed responses to surrounding natural habitat. In another analysis, we isolated the land-use variable with the highest-ranked importance for each model. We then tallied the number of models for which important natural habitat variables or crop variables increased or decreased pest-control metrics. In both cases, we used χ^2 tests to assess whether each pest-control variable increased or decreased in response to surrounding natural habitat or cropland.

Analysis of Model Fit. We calculated Pearson correlation coefficients and R^2 values between model predictions and observed data to assess model fit, as these metrics are directly comparable across models. We then used t tests to determine whether the average Pearson correlation coefficient (across all datasets) differed from 0. We also tested whether models explained more variation in pest-control responses in certain regions. To do so, we implemented linear mixed models that predicted correlation coefficients from a variable identifying the biogeographic realm of the study. We used likelihood ratio tests to compare models with and without the biogeographic realm predictor.

Prediction Analysis. To quantify our models' predictive power, we correlated the observed pest-control variable values in each of the 367 pest-control responses ("field observations," hereafter) with the average of all predictions generated across our many landscape models, using the following procedure. First, we selected which landscape models to use. Models built from datasets ("model data," hereafter) that shared >25% of their study sites with the field observations were excluded so that models and field observations could be considered totally independent. We also excluded models when model data had more finely resolved land-cover classes than independent field observations. Models constructed from one pest-control variable could be applied to observations of a different pest-control variable, provided that they were either both enemy variables or pest variables. Yield models were always applied to only yield data.

Once a set of candidate models was obtained for each set of independent field observations, we used the model-averaged estimates of the effects of each landscape variable to predict pest-control variable values. Predictions from each model were scaled by subtracting the mean value and dividing by the SD, and then averaged across all models. Finally, we calculated the Pearson correlation between the field observations and the multimodel-averaged predictions. Again, we used t tests to determine whether the average Pearson correlation coefficient (across all tested datasets) differed from 0.

We also conducted a sensitivity analysis to determine if there were strategies for filtering and applying subsets of models to independent field observations so as to increase predictive power. We repeated the above model-testing procedure but after choosing to only include models derived from studies that (*i*) focused on the same crop as the independent field observations, (*ii*) were in a same biome (42), (*iii*) were in the same biogeographic realm, (*iv*) were located <500 km away from the field observa-

tions, and (*v*) had the same landscape variable values as the values present in the field observations. Specifically, we tried only including models constructed from studies in regions with the same land-use classes as the field observations and, even more stringently, experimented with only including models when landscape variable values associated with field observations were within the range of the variable values in the model data (i.e., the landscape gradients matched). The sensitivity analysis was conducted in a full-factorial manner, whereby every combination of the five inclusion procedures was analyzed independently. Finally, independent from the full-factorial sensitivity analysis, we also experimented with excluding models that did not explain any variation in their own datasets (Pearson's r between model predictions and observations was <0.25). For each scenario and for each tested dataset, we calculated the Pearson correlation coefficients between model predictions and testing data.

We then used linear mixed models to evaluate whether each model restriction procedure improved or diminished overall predictive power across the independent field observations. Specifically, for each class of pest-control variable, mixed models were used to relate the Pearson correlation coefficients (response) to a series of fixed effects that indicated whether each exclusion procedure outlined above had taken place (e.g., exclude models that were focused on a different crop, in a different biome, etc.). Because the same field observation datasets appeared in each exclusion scenario, "field observation dataset name" was included as a random effect. We also included a random effect of "unique scenario combination" to identify all of the studies that were subject to the same set of exclusion procedures (e.g., exclude models that were focused on a different crop but allow models to be in different biogeographic realms and biomes). Then, we used likelihood ratio tests to compare the full mixed model with models in which we iteratively removed each fixed effect (indicator variables of each exclusion procedure).

[a]Department of Wildlife, Fish, and Conservation Biology, University of California, Davis, CA 95616; [b]Natural Capital Project, Woods Institute for the Environment, Stanford University, Stanford, CA 94618; [c]National Audubon Society, Boulder, CO 80305; [d]Department of Animal Ecology and Tropical Biology, Biocenter, University of Würzburg, 97074 Würzburg, Germany; [e]Biodiversity and Ecosystem Services, Bioversity International, 34397 Montpellier, France; [f]Department of Entomology, Cornell University, Ithaca, NY 14853; [g]Department of Entomology, University of Wisconsin, Madison, WI 53706; [h]Department of Entomology, University of Maryland, College Park, MD 20742; [i]Bren School of Environmental Science & Management, University of California, Santa Barbara, CA 93106-5131; [j]Tropical Agricultural Research and Higher Education Center (CATIE), Turrialba, Costa Rica 30501; [k]Department of Horticulture, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061; [l]Institut National de la Recherche Agronomique (INRA), UMR 1065 SAVE, ISVV, Université de Bordeaux, Bordeaux Sciences Agro, F-33883 Villenave d'Ornon, France; [m]Department of Ecology, Swedish University of Agricultural Sciences, SE-75007 Uppsala, Sweden; [n]Department of Entomology and Nematology, University of California, Davis, CA 95616; [o]Commonwealth Scientific and Industrial Research Organisation (CSIRO), Brisbane, QLD 4001, Australia; [p]Agroecology, University of Goettingen, 37077 Goettingen, Germany; [q]Bio-Protection Research Centre, Lincoln University, Lincoln 7647, New Zealand; [r]Environment and Production Technology Division, International Food Policy Research Institute, Washington, DC 20005; [s]Department of Biology, University of Massachusetts, Amherst, MA 01003; [t]Agroecology and Environment, Agroscope, CH-8046 Zürich, Switzerland; [u]INRA, UMR 0980 BAGAP, F-35042 Rennes, France; [v]Department of Biology and Zoology, Faculty of Sciences, PMAS-Arid Agriculture University Rawalpindi, 46000 Rawalpindi, Pakistan; [w]French Agricultural Research Centre for International Development (CIRAD), UPR Bioagresseurs Analyse et Maîtrise du Risque, F-34398 Montpellier, France; [x]Wageningen Environmental Research (Alterra), Wageningen University and Research, 6700 AA Wageningen, The Netherlands; [y]Farming Systems Ecology, Wageningen University and Research, 6700 AK Wageningen, The Netherlands; [z]Department of Ecology, Brandenburg University of Technology Cottbus-Senftenberg, 03046 Cottbus, Germany; [aa]Department of Entomology, The Pennsylvania State University, University Park, PA 16802; [bb]Department of Entomology, Texas A&M AgriLife Research, Corpus Christi, TX 78406; [cc]Arthropods Department, Natural Sciences Museum of Barcelona, 08003 Barcelona, Spain; [dd]Department of Entomology, University of Arizona, Tucson, AZ 85721-0036; [ee]Departamento de Ecologia, Universidade Federal de Goiás, Goiânia, Brazil 74690-900; [ff]Department of Biology, Geology, Physics and Inorganic Chemistry, Universidad Rey Juan Carlos, E28933 Móstoles, Madrid, Spain; [gg]Faculty of Biology, University of Belgrade, 11000 Belgrade, Serbia; [hh]School of Environmental Sciences, University of Hull, Hull HU6 7RX, United Kingdom; [ii]INRA, UMR AGIR, Toulouse-Occitanie, Chemin de Borderouge, 31326 Castanet-Tolosan, France; [jj]Department of Entomology, University of Manitoba, Winnipeg, MB R3T 2N2, Canada; [kk]Department of Plant and Soil Science, The University of Vermont, Burlington, VT 05405; [ll]Faculty of Management, Science & Technology, Open Universiteit, 6401 DL Heerlen, The Netherlands; [mm]INRA, University of Côte d'Azur, CNRS, UMR 1355-7254, 06903 Sophia Antipolis, France; [nn]Department of Animal Ecology, Justus Liebig University, 35392 Giessen, Germany; [oo]Department of Landscape Ecology, Kiel University, D-24118 Kiel, Germany; [pp]Biometry and Environmental System Analysis, University of Freiburg, D-79106 Freiburg, Germany; [qq]Department of Agronomy and Plant Genetics, University of Minnesota, St. Paul, MN 55108; [rr]Institute for Environmental Sciences, University of Koblenz-Landau, D-76829 Landau, Germany; [ss]Department of Biology Education, Leibniz Institute for Science and Mathematics Education (IPN) at Kiel University, 24118 Kiel, Germany; [tt]UR1115 Plantes et Système de Cultures Horticoles, INRA, F-84000 Avignon, France; [uu]Centre for Ecology and Conservation, University of Exeter, Penryn TR10 9FE, United Kingdom; [vv]Institute of Zoology, University of Natural Resources and Life Sciences, A-1180 Vienna, Austria; [ww]Centre for Agri-Environmental Research, School of Agriculture, Policy and Development, University of Reading, Reading RG6 6AR, United Kingdom; [xx]Key Laboratory of Pollinating Insect Biology, Ministry of Agriculture, Institute of Apicultural Research, Chinese Academy of Agricultural Sciences, 100193 Beijing, China; [yy]Department of Entomology, University of Kentucky, Lexington, KY 40546; [zz]Statewide IPM Program, Kearney Agricultural Center, Parlier, CA 93648; [aaa]World Agroforestry Center, Nairobi 00100, Kenya; [bbb]Charles Sturt University, Orange, NSW 2800, Australia; [ccc]Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI 48109; [ddd]Department of Entomology, University of Minnesota, St. Paul, MN 55108; [eee]Department of Entomology and Plant Pathology, North Carolina State University, Raleigh, NC 27605-7630; [fff]Instituto Nacional de Biodiversidad–INABIO & Facultad de Ciencias Agícolas, Universidad Central del Ecuador, Quito 170129, Ecuador; [ggg]Department of Entomology, Michigan State University, East Lansing, MI 48912; [hhh]Ministry of Agriculture, Forestry and Fisheries, Phnom Penh 255 370, Cambodia; [iii]Department of Biology & Environment, University of Haifa–Oranim, 36006 Tivon, Israel; [jjj]Great Lakes Bioenergy Research Center, University of Wisconsin, Madison, WI 53726; [kkk]Department of Evolutionary and Environmental Biology, University of Haifa, 3498838 Haifa, Israel; [lll]Laboratorio de Control Biológico, Instituto de Ciencias Biológicas, Universidad de Talca, Talca 3465548, Chile;

Karp et al.

mmmUMR IGEPP, Agrocampus–Ouest, F-35042 Rennes Cedex, France; nnnForest and Range Land Plant Biodiversity Directorate, Ethiopian Biodiversity Institute, 30726 Addis Ababa, Ethiopia; oooDepartment of Environmental Studies, University of California, Santa Cruz, CA 95060; pppInstitute of Plant Protection, Chinese Academy of Agricultural Sciences, 100193 Beijing, China; qqqBlaustein Institutes for Desert Research, Ben-Gurion University of the Negev, Midreshet Ben-Gurion 8499000, Israel; rrrDepartment of Botany and Biodiversity Research, Division of Conservation Biology, Vegetation Ecology and Landscape Ecology, University of Vienna, 1030 Vienna, Austria; sssOffice of Environmental Farming and Innovation, California Department of Food and Agriculture, Sacramento, CA 95814; tttDepartment of Crop & Forest Sciences, Agrotecnio Centre, Universitat de Lleida, 25198 Lleida, Spain; uuuUMR Institute for Genetics, Environment and Plant Protection, University of Rennes 1, F-35042 Rennes Cedex, France; vvvDepartment of Agronomy, Food, Natural Resources, Animals and Environment, University of Padova, 35020 Padua, Italy; wwwNariño, Corpoica, Centro de Investigación Obonuco, 520003 Pasto, Colombia; xxxDepartment of Biology, McDaniel College, Westminster, MD 21157; yyyPerformance of Tropical Production and Processing Systems, CIRAD, UPR 115 AIDA, TA B-115/02, 34398 Montpellier Cedex 5, France; zzzInstitute for Resources, Environment and Sustainability, University of British Columbia, Vancouver, BC V6T 1Z4, Canada; aaaaDepartment of Ecosystem Studies, School of Agriculture & Life Sciences, The University of Tokyo, 113-0032 Tokyo, Japan; bbbbInstituto de Investigaciones Fisiológicas y Ecológicas Vinculadas a la Agricultura, Facultad de Agronomía, Universidad de Buenos Aires, CONICET, C1417DSE CABA, Argentina; ccccCentro de Estudios Avanzados en Zonas Áridas (CEAZA), Universidad Católica del Norte, Coquimbo 1780000, Chile; ddddEntomology, Iowa State University, Ames, IA 50011; eeeeRegional Agricultural Research and Development Center, Southern Branch, Besor 85400, Israel; ffffSchool of Agriculture, Food and Wine, University of Adelaide, Adelaide, SA 5005, Australia; ggggDepartment for Aquatic Resources, Swedish Agricultural University, 742 42 Öregrund, Sweden; hhhhUMR Dyanfor (INRA-INPT), Université de Toulouse, 31326 Auzeville Tolosane, France; iiiiDepartment of Biology, The Pennsylvania State University, State College, PA 16801; jjjjDepartamento de Protección Ambiental, Estación Experimental de Zaidín (CSIC), 18008 Granada, Spain; kkkkOffice of National Programs, US Department of Agriculture–Agricultural Research Service, Beltsville, MD 20705; llllAgroécologie, AgroSup Dijon, INRA, Université Bourgogne Franche-Comté, F-21000 Dijon, France; mmmmInstitute of Ecology and Evolution, University of Bern, CH-3012 Bern, Switzerland; nnnnNERC Centre for Ecology & Hydrology, Wallingford, Oxfordshire OX10 8BB, United Kingdom; ooooDepartment of Entomology, Texas A&M University, College Station, TX 77845; ppppNorthern Plains Agricultural Research Laboratory, US Department of Agriculture–Agricultural Research Service, Sidney, MT 59270; qqqqINRA AGIR, University of Toulouse, INPT, F-31326 Castanet-Tolosan, France; rrrrChair of Nature Conservation and Landscape Ecology, University of Freiburg, 79106 Freiburg, Germany; ssssDepartment of Entomology, The Ohio State University, Columbus, OH 43210; ttttCentre for Environmental & Climate Research, Lund University, 223 62 Lund, Sweden; uuuuEnvironmental Science Program, Rochester Institute of Technology, Rochester, NY 14623; vvvvCentre for Agriculture and Biosciences International (CABI), 2800 Delémont, Switzerland; wwwwInstitute for Sustainable Agro-Ecosystem Services, The University of Tokyo, 188-0002 Tokyo, Japan; xxxxForestry and Forest Products Research Institute, Tsukuba, 305-8687 Ibaraki, Japan; yyyyDepartment of BioSciences, University of Melbourne, Parkville, VIC 3010, Australia; zzzzUMR IGEPP, Agrocampus–Ouest, F-49045 Angers Cedex 1, France; aaaaaDepartment of Practaculture, School of Agriculture, Ningxia University, 750021 Yinchuan, China; bbbbbUMR 0211 INRA AgroParisTech Agronomie, 78850 Thiverval-Grignon, France; cccccInstitute for Agricultural Environment, Phu Do Ward 100000, Vietnam; dddddCentre for Crop Systems Analysis, Wageningen University and Research, 6708 PB Wageningen, The Netherlands; eeeeeDepartment of Fish and Wildlife Sciences, University of Idaho, Moscow, ID 83844-1136; fffffMichigan State University Extension, Hart, MI 49420; gggggCGIAR Program on Roots, Tubers and Banana (CRP-RTB), International Center for Tropical Agriculture (CIAT), Hanoi 100000, Vietnam; hhhhhSchool for Biological Sciences, University of Queensland, Brisbane, QLD 4000, Australia; iiiiiInstitute of Entomology, Jiangxi Agricultural University, 330045 Nanchang, China; jjjjjNational Agriculture and Food Research Organisation, 305-8517 Tsukuba, Japan; kkkkkFukusima Branch, National Institute for Environmental Studies, 963-770 Fukushima, Japan; and lllllDepartment of Health and Environmental Sciences, Xi'an Jiaotong–Liverpool University, 215123 Suzhou, China

1. Kennedy CM, et al. (2013) A global quantitative synthesis of local and landscape effects on wild bee pollinators in agroecosystems. Ecol Lett 16:584–599.
2. Bianchi FJJA, Booij CJH, Tscharntke T (2006) Sustainable pest regulation in agricultural landscapes: A review on landscape composition, biodiversity and natural pest control. Proc Biol Sci 273:1715–1727.
3. Veres A, Petit S, Conord C, Lavigne C (2013) Does landscape composition affect pest abundance and their control by natural enemies? A review. Agric Ecosyst Environ 160:110–117.
4. Chaplin-Kramer R, O'Rourke ME, Blitzer EJ, Kremen C (2011) A meta-analysis of crop pest and natural enemy response to landscape complexity. Ecol Lett 14:922–932.
5. Rusch A, et al. (2016) Agricultural landscape simplification reduces natural pest control: A quantitative synthesis. Agric Ecosyst Environ 221:198–204.
6. Gurr GM, Wratten SD, Landis DA, You M (2017) Habitat management to suppress pest populations: Progress and prospects. Annu Rev Entomol 62:91–109.
7. Tscharntke T, et al. (2016) When natural habitat fails to enhance biological pest control––Five hypotheses. Biol Conserv 204:449–458.
8. Losey JE, Vaughan M (2006) The economic value of ecological services provided by insects. Bioscience 56:311–323.
9. Millennium Ecosystem Assessment (2005) Ecosystems & Human Well-Being: Synthesis (Island, Washington, DC).
10. The Economics of Ecosystems and Biodiversity (2015) TEEB for agriculture and food interim report (UN Environment Program, Geneva). Available at img.teebweb.org/wp-content/uploads/2016/01/TEEBAgFood_Interim_Report_2015_web.pdf?utm_source=website&utm_medium=report&utm_campaign=TeebAgriFoodInterimReport. Accessed July 3, 2018.
11. Kean J, Wratten S, Tylianakis J, Barlow N (2008) The population consequences of natural enemy enhancement, and implications for conservation biological control. Ecol Lett 11:604–612.
12. Zhang W, Swinton SM (2009) Incorporating natural enemies in an economic threshold for dynamically optimal pest management. Ecol Modell 220:1315–1324.
13. Bianchi FJJA, Honĕk A, van der Werf W (2007) Changes in agricultural land use can explain population decline in a ladybeetle species in the Czech Republic: Evidence from a process-based spatially explicit model. Landsc Ecol 22:1541–1554.
14. Bianchi FJJA, Schellhorn NA, Buckley YM, Possingham HP (2010) Spatial variability in ecosystem services: Simple rules for predator-mediated pest suppression. Ecol Appl 20:2322–2333.
15. Jonsson M, et al. (2014) Ecological production functions for biological control services in agricultural landscapes. Methods Ecol Evol 5:243–252.
16. Ricketts TH, et al. (2008) Landscape effects on crop pollination services: Are there general patterns? Ecol Lett 11:499–515, and erratum (2008) 11:1121.
17. Open Science Collaboration (2015) Estimating the reproducibility of psychological science. Science 349:aac4716.
18. Levins R, Schultz BB (1996) Effects of density dependence, feedback, and environmental sensitivity on correlations among predators, prey, and plant resources: Models and practical implications. J Anim Ecol 65:802–812.
19. Chaplin-Kramer R, de Valpine P, Mills NJ, Kremen C (2013) Detecting pest control services across spatial and temporal scales. Agric Ecosyst Environ 181:206–212.
20. Perovic DJ, Gurr GM, Raman A, Nicol HI (2010) Effect of landscape composition and arrangement on biological control agents in a simplified agricultural system: A cost-distance approach. Biol Control 52:263–270.
21. Martin EA, Seo B, Park CR, Reineking B, Steffan-Dewenter I (2016) Scale-dependent effects of landscape composition and configuration on natural enemy diversity, crop herbivory, and yields. Ecol Appl 26:448–462.
22. Jonsson M, et al. (2012) Agricultural intensification drives landscape-context effects on host-parasitoid interactions in agroecosystems. J Appl Ecol 49:706–714.
23. Letourneau DK, et al. (2011) Does plant diversity benefit agroecosystems? A synthetic review. Ecol Appl 21:9–21.
24. Chaplin-Kramer R, Kremen C (2012) Pest control experiments show benefits of complexity at landscape and local scales. Ecol Appl 22:1936–1948.
25. Mitchell MGE, et al. (2015) The Monteregie connection: Linking landscapes, biodiversity, and ecosystem services to improve decision making. Ecol Soc 20:15.
26. Ouyang Z, et al. (2016) Improvements in ecosystem services from investments in natural capital. Science 352:1455–1459.
27. Mandle L, Bryant B, Ruckelshaus M, Pfaff A (2015) Entry points for considering ecosystem services within infrastructure planning: How to integrate conservation with development in order to aid them both. Conserv Lett 9:221–227.
28. Nepstad D, et al. (2013) More food, more forests, fewer emissions, better livelihoods: Linking REDD+, sustainable supply chains and domestic policy in Brazil, Indonesia and Colombia. Carbon Manag 4:639–658.
29. Chaplin-Kramer R, et al. (2015) Spatial patterns of agricultural expansion determine impacts on biodiversity and carbon storage. Proc Natl Acad Sci USA 112:7402–7407.
30. Meehan TD, Werling BP, Landis DA, Gratton C (2011) Agricultural landscape simplification and insecticide use in the midwestern United States. Proc Natl Acad Sci USA 108:11500–11505.
31. Fritz S, et al. (2015) Mapping global cropland and field size. Glob Change Biol 21:1980–1992.
32. Monfreda C, Ramankutty N, Foley JA (2008) Farming the planet: Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. Global Biogeochem Cycles 22:1–19.
33. Plecas M, et al. (2014) Landscape composition and configuration influence cereal aphid–hyperparasitoid interactions and biological control differentially across years. Agric Ecosyst Environ 183:1–10.
34. Newbold T, et al. (2013) Ecological traits affect the response of tropical forest bird species to land-use intensity. Proc Biol Sci 280:20122131.
35. Segoli M, Rosenheim JA (2012) Should increasing the field size of monocultural crops be expected to exacerbate pest damage? Agric Ecosyst Environ 150:38–44.
36. Root RB (1973) Organization of a plant-arthropod association in simple and diverse habitats: The fauna of collards (Brassica oleracea). Ecol Monogr 43:95–124.
37. Fahrig L, et al. (2011) Functional landscape heterogeneity and animal biodiversity in agricultural landscapes. Ecol Lett 14:101–112.
38. Holland JM, et al. (2017) Semi-natural habitats support biological control, pollination and soil conservation in Europe. A review. Agron Sustain Dev 37:31.
39. Karp DS, et al. (2016) Farming practices for food safety threaten pest-control services to fresh produce. J Appl Ecol 53:1402–1412.
40. Jun C, Ban Y, Li S (2014) Open access to Earth land-cover map. Nature 514:434.
41. Zuur AF, Ieno EN, Walker NJ, Saveliev AA, Smith GM (2009) Mixed Effects Models and Extensions in Ecology with R (Springer, New York).
42. Olson DM, et al. (2001) Terrestrial ecoregions of the world: A new map of life on Earth. Bioscience 51:933–938.

Supplementary Methods

Overview

We compiled a global pest-control database to model impacts of landscape composition on pest-control services (Fig. S4). First, we cleaned and transformed pest-control response data and land-use maps to help standardize predictor (landscape) and response (pest control) variables across diverse studies. Then, we used a model averaging procedure to assess the extent to which landscape composition alone could predict spatial variation in pest and enemy abundances, predator activity, plant damage, and crop yields. Finally, to assess predictive power, we applied landscape models to independent field observations, exploring different strategies for matching models to field observations. Below, we discuss these steps in greater detail.

Data Processing

We used a multi-faceted strategy to standardize data within and across studies. Abundance data were processed as follows. First, because authors reported presences (but not absences) in their survey data, we assumed that if at least one individual of any species was censused at a given location, on a given date, and using a given sampling method, then any other species not listed on that location-date-method combination were absent. Second, because some sampling methods are ineffective for particular taxa (*e.g.,* sticky traps for epigeal predators), we only included data from sampling methods that contributed at least 5% of the total observations (across sampling methods). Third, because sampling effort sometimes differed among locations, we divided the total number of individuals in each functional group (pest, predator, etc.) by the number and/or duration of surveys conducted. If a mean abundance across surveys was reported, then accounting for variable effort was not necessary. Finally, to allocate equal importance to observations taken in different years or with different sampling methods, we standardized effort-corrected data across sites within a given year and sampling method by subtracting the mean and dividing by the standard deviation. Prior to standardization,

we log, square-root, and fourth-root transformed the data to adhere to model assumptions (see below).

Predator and pest activity and yield data were processed similarly to abundance data, accounting for variable effort and standardizing data within years. However, unlike abundance data, which was always expressed as animal counts, different activity and yield sampling methods produced distinct data types; therefore, unique sampling methods were analyzed separately. Moreover, additional processing was required to analyze data resulting from exclosure experiments. Specifically, we used exclosure data to create a biocontrol index (1). First, when before-after-control-impact data were reported, we subtracted pest density or damage values obtained before the experiments began from final values recorded upon the experiment's conclusion. Second, when necessary, we averaged across all exclusion replicates present at the same site. Third, we computed the biocontrol index either by dividing values observed in the exclusion treatment from those in the control treatment (provided that pests were detected in both treatments) or subtracting values (when any data other than pest counts were reported).

Land-use maps

We used a hierarchical approach to obtain land-cover maps for each study (Appendix 1). Because many data contributors had already obtained or created high-resolution maps of their study regions, we asked authors to submit land-cover maps if available. When contributors could not provide us with maps, we used either regional maps or, as a last resort, a 30 m global land cover product (Jun *et al.* 2014; http://www.globallandcover.com/). Regional maps included: Cropland Data Layer (United States; https://nassgeodata.gmu.edu/CropScape/), Land Use 2010 (Canada; http://open.canada.ca/data/en/dataset/9e1efe92-e5a3-4f70-b313-68fb1283eadf), CORINE land cover (Europe; http://www.eea.europa.eu/publications/COR0-landcover), Catchment Scale Land Use Map (Australia; https://data.gov.au/dataset/catchment-scale-land-use-of-australia-update-march-2015), and Land Cover Database (New Zealand;

). When possible, land-cover maps were chosen so that they matched the dates that samples were collected. All maps were projected into WGS 1984, rasterized, and reclassified into the following categories: (1) forest and tree plantations, (2) grassland, (3) scrubland, (4) annual cropland, (5) perennial cropland, (6) urban areas, and (7) other. Some maps reported less finely-resolved land-cover categories, necessitating two aggregate categories: (8) all natural/semi-natural habitats and (9) all cropland. If both aggregate and finely-resolved categories were reported in the same map (*e.g.,* forest, grassland, scrubland, and natural habitats), then we reclassified all finely-resolved categories into the aggregate category (*e.g.,* forest, grassland, and scrubland into natural habitat).

In some instances, authors provided land-cover maps that did not extend to 2 km around their study sites. In this case, we overlaid their local land-cover map with a regional map (when possible) or the global product. Doing so sometimes resulted in a mixture of aggregate and finely-resolved land-cover categories. If all areas <700 m from the study locations were finely resolved, then we assumed that the same land cover proportions extended beyond to 2 km and disaggregated the broader classes. For example, if a local map indicated that annual and perennial agriculture occupied 10% and 30% of the area within 700 m of study sites, then we assumed that any areas beyond 700 m that were delineated as "agriculture" were composed of 25% perennial and 75% annual agriculture. If, on the other hand, any aggregate land-cover category was present within 700 m of a study site, then we reclassified all finely-resolved areas across the entire map into that aggregate land-cover category.

Landscape composition

We used a distance-weighting procedure to quantify landscape composition. First, we calculated the area (m^2) of each land cover class in 20 concentric rings around each study site, with outer radii at 100 m intervals between 100 m and 2 km. Next, we used a Gaussian decay function to assign weights to each ring, where rings closer to the focal site were weighted more than those further away (3):

$$W = \exp(-O^2/(2*d^2)) \qquad\qquad\qquad\qquad \text{equation 1}$$

where W is the weight, O is out the outer edge distance of the ring, and d is a decay rate that determines how rapidly weightings decline with distance. We used three decay rates (250, 750, 1250) to later assess landscape composition at multiple spatial scales. We then used the weightings to calculate a weighted sum for the total area of each land cover class. Finally, we divided that number by the weighted sum of all land covers to ultimately obtain a proportional representation of each land-cover class around each study site.

References

1.   Chaplin-Kramer R, Kremen C (2012) Pest control experiments show benefits of complexity at landscape and local scales. *Ecol Appl* 22:1936–1948.

2.   Jun C, Ban Y, Li S (2014) Open access to Earth land-cover map (correspondance). *Nature* 514:434.

3.   Karp DS, et al. (2016) Farming practices for food safety threaten pest-control services to fresh produce. *J Appl Ecol* 53:1402–1412.

Supplementary Tables and Figures

Table S1: Correlations between landscape model predictions and observed values (corresponding to Fig. 3).

	N	N sig. models	r	R^2	T	95% C.I.	P
Dominant pests	76	42	0.42	0.18	19.30	(0.37, 0.46)	<0.001
All pests	27	14	0.41	0.17	8.45	(0.31, 0.51)	<0.001
All enemies	78	41	0.36	0.13	16.00	(0.32, 0.41)	<0.001
Sentinel experiments	68	42	0.45	0.20	21.00	(0.40, 0.49)	<0.001
Cage experiments	23	9	0.39	0.15	9.84	(0.31, 0.47)	<0.001
Pest damage	34	17	0.38	0.14	10.60	(0.31, 0.46)	<0.001
Crop yield	53	25	0.43	0.18	13.70	(0.36, 0.49)	<0.001

N= number of tested datasets
N sig.= number of datasets where predictions and observed data were significantly correlated
r= mean Pearson correlation coefficient comparing predictions and observed data
R^2= mean R^2 across all tested datasets
t= t-statistic assessing whether the mean Pearson correlation coefficient across datasets is different from 0
95% C.I.= 95% confidence interval around the mean Pearson correlation coefficient
P= p-value associated with the t-statistic

Table S2: Correlations between landscape model predictions and testing data (corresponding to Fig. 4).

	All models							Shared crop species						
	N	N Sig. Models (+)	N Sig. Models (-)	r	T	C.I.	P	N	N Sig. Models (+)	N Sig. Models (-)	r	T	C.I.	P
Dominant pests	75	9	8	0.01	0.22	(-0.06, 0.07)	0.83	61	9	3	0.11	3.25	(0.04, 0.18)	**<0.01**
All pests	27	2	3	0.04	0.75	(-0.07, 0.16)	0.46	14	3	0	0.18	2.12	(0, 0.37)	**0.05**
All enemies	77	8	9	-0.02	-0.63	(-0.08, 0.04)	0.53	32	5	0	0.07	1.96	(0, 0.15)	**0.05**
Sentinel experiments	66	4	4	-0.01	-0.18	(-0.07, 0.06)	0.86	62	5	4	0.07	2.04	(0, 0.14)	**0.05**
Cage experiments	23	2	0	0.03	0.59	(-0.08, 0.15)	0.56	9	2	0	0.21	2.36	(0.01, 0.42)	**0.05**
Pest damage	34	3	3	-0.02	-0.52	(-0.12, 0.07)	0.61	25	3	3	0.00	0.06	(-0.12, 0.13)	0.95
Crop yield	47	2	1	-0.01	-0.33	(-0.09, 0.07)	0.75	33	1	2	0.02	0.32	(-0.08, 0.11)	0.75

N= number of tested datasets
N sig. += number of datasets where predictions and observed data were significantly, positively correlated
N sig. -= number of datasets where predictions and observed data were significantly, negatively correlated
r= mean Pearson's correlation coefficient across all tested datasets
t= t-statistic assessing whether the mean correlation across datasets is different from 0
95% C.I.= 95% confidence interval around the mean Pearson correlation coefficient
P= p-value associated with the t-statistic

Table S3: Sensitivity analysis of procedures that selectively applied models to data. In each cell, we indicated whether the procedure improved (+) or reduced (-) correlations between model predictions and testing data. Numbers in parentheses are χ^2 statistics and associated *P* values from likelihood-ratio tests (see Methods). Bold indicates significance (*P*<.0.05).

| Models and field observations…	Dominant pests	All pests	All enemies	Sentinel experiments	Cage experiments	Pest damage	Crop yield
Are in the same biogeographic realm	- (1.3, 0.26)	- (3.1, 0.08)	- (2.8, 0.1)	**- (38.1, <0.01)**	**- (3.9, 0.05)**	+ (3.2, 0.08)	**- (5.4, 0.02)**
Are in the same biome	+ (0.5, 0.46)	- (0, 0.89)	+ (0.5, 0.48)	**- (6.1, 0.01)**	**+ (37, <0.01)**	+ (1.9, 0.16)	**- (9.6, <0.01)**
Share the same crop species	**+ (55.6, <0.01)**	**+ (10.6, <0.01)**	+ (1.5, 0.22)	**+ (34.1, <0.01)**	**+ (10.7, <0.01)**	**+ (38.5, <0.01)**	- (0.9, 0.34)
Are <500km apart	**- (35.5, <0.01)**	**+ (25.1, <0.01)**	**- (11.7, <0.01)**	**- (88.2, <0.01)**	**- (108.8, <0.01)**	+ (2.6, 0.1)	**+ (37.2, <0.01)**
Share the same land-use classes (or ranges of land use classes)	+ (3.2, 0.2)	**+ (117.3, <0.01)**	**+ (46.5, <0.01)**	**+ (55.4, <0.01)**	+ (1.2, 0.55)	+ (2.2, 0.34)	**+ (138.4, <0.01)**
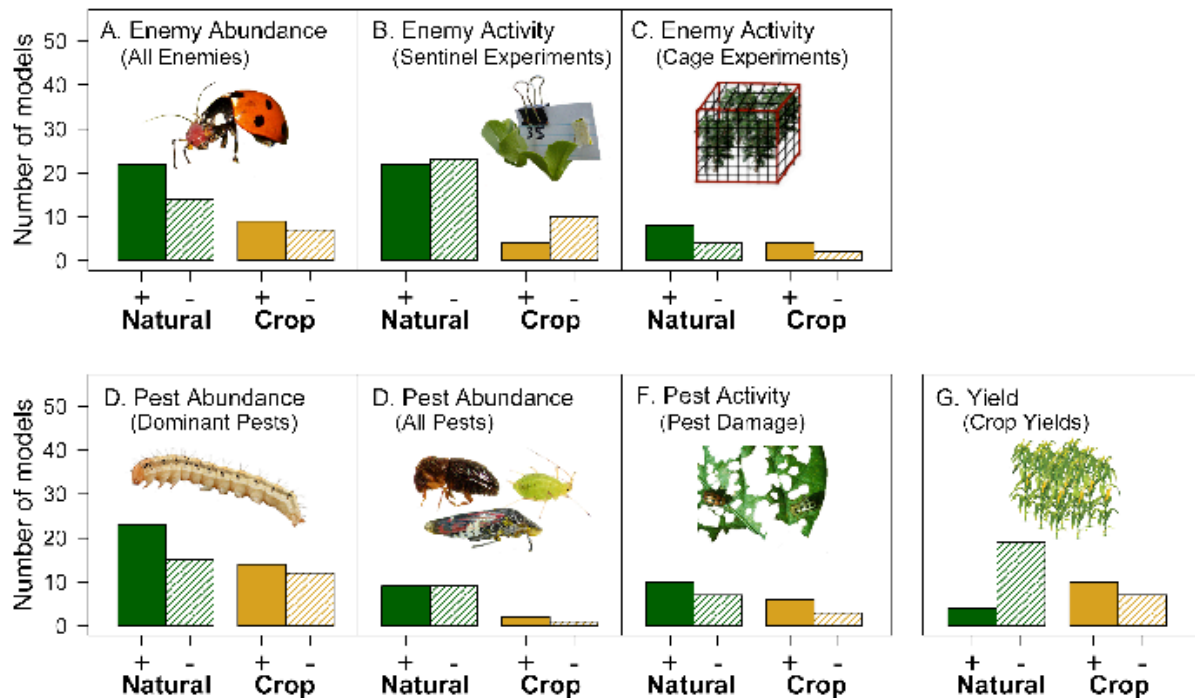
Fig. S1: *Landscape effects on most important pest-control variable in each model.* After (1) selecting the most predictive model for each pest-control response, (2) isolating the most important land-use predictor in each model, and (3) reclassifying those predictors as natural (forest, grassland, and scrubland; green bars) versus crop (annual and perennial; orange bars), we tallied the number of pest-control responses for which models had either positive (solid) or negative (diagonal hashed) estimates of the effect of each landscape predictor. Panels represent the seven pest-control variables: including abundance (A) and activity (B, C) of natural enemies; abundance (D, E) and activity (F) of pests, and crop yields (G). χ^2 tests indicated that most pest-control responses variables showed heterogeneous patterns of association with the extent of surrounding natural habitat and cropland— with roughly equivalent numbers of pest-control responses having models with positive and negative effects (all *P>* 0.05). Crop yields, however, tended to be negatively associated with natural land covers.
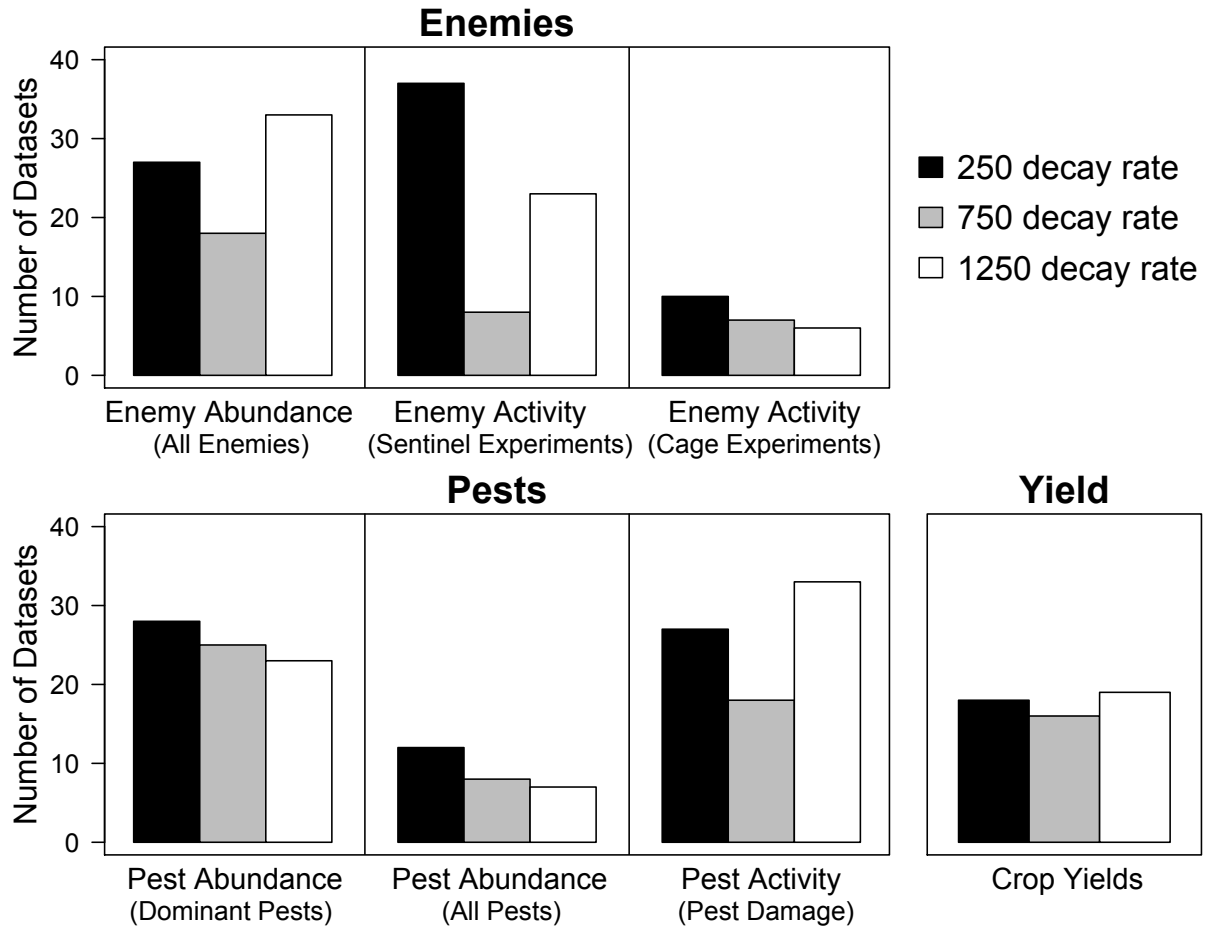
Figure S2: *The most predictive spatial scales associated with each dataset.* Landscape composition was calculated using three Gaussian decay rates that gave increasing weight to areas further from sampling sites: 250 (black bars), 750 (gray bars), and 1250 (white bars). For each dataset, the most predictive scale was defined as the scale for which model predictions best correlated with observed data. Bars indicate the number of datasets that exhibited highest correlations for each spatial scale. Pest abundance responses, for example, tended to be more predicted by landscape variables that gave higher weights to areas closer to the sampling sites.
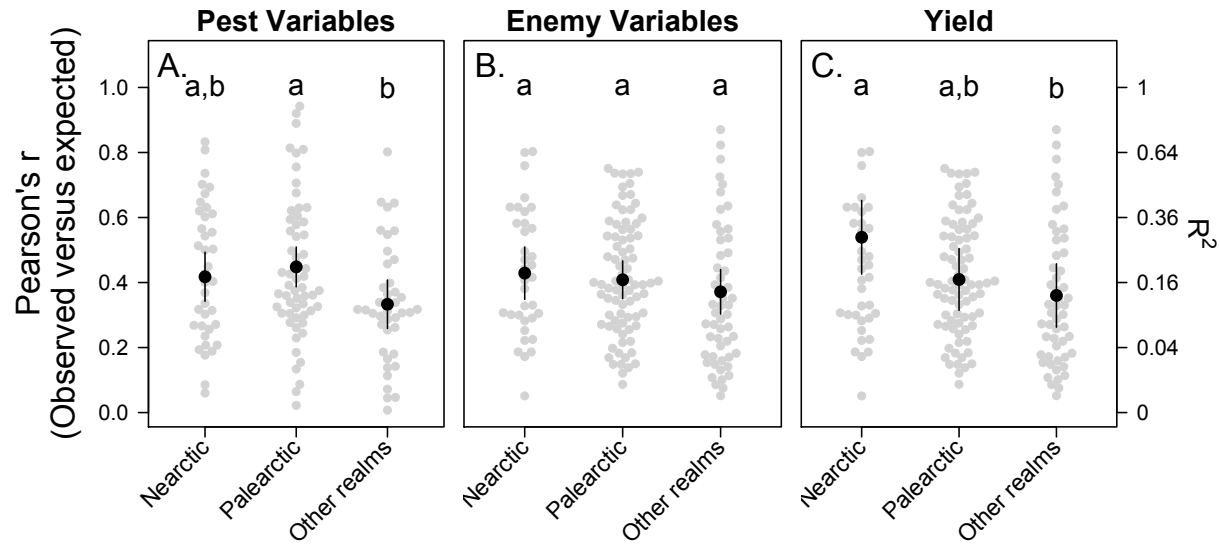
Figure S3: *Variation in explanatory power of landscape models across biogeographic realms.* Generalized Linear Mixed Models were used to assess the effect of the biogeographic realm of the dataset (Nearctic, Palearctic, or other realms) on the correlation between landscape model predictions and observed data. Separate analyses were conducted on pest variables (Panel A; abundance and crop damage), enemy variables (Panel B; abundance and activity), and crop yields (Panel C). Gray circles represent both Pearson correlations between model predictions and observed data and R^2 values (square of Pearson's r). Black circles are the mean correlation across all datasets in a given biogeographic realm between observed and predicted values. Black lines are 95% confidence intervals. Lowercase letters denote significance under likelihood ratio tests. Panel A shows that, for pest variables, explanatory power of landscape models was highest in Nearctic and Palearctic realms. Though no differences were observed in enemy variables (Panel B), yield models better fit observations in the Nearctic compared to other realms (Panel C).
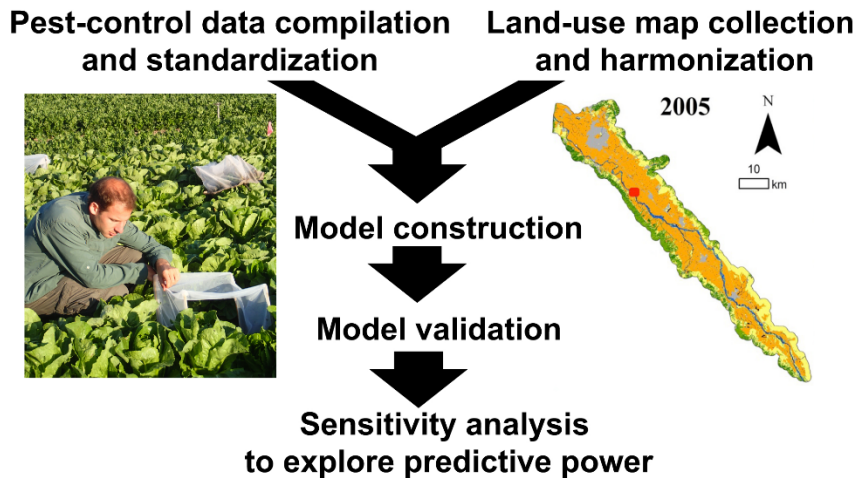

Fig. S4: *Methods diagram.* Flowchart indicating the sequence of research activities, from data compilation to model testing.

Dataset Descriptions

Dataset S1: *Study attributes.* Table detailing the pest-control responses analyzed. For each response (row), we include information on the study names (column a), the country of the study (column b), the target crop (column c), the broad response category (abundance, activity, or yield; column d), the specific type of pest-control response (abundance of all enemies, all pests, or dominant pests, cage experiments measuring changes in crop damage, crop yields, or infestation, sentinel pest experiments, crop damage, or crop yields; column e), the target species or measure type (column f), the number of sites analyzed (column g), the source of the land-use data (column h), and the citation corresponding to the target study (column i).

Dataset S2: *Biocontrol database.* We include the biological control database, consisting of 5 worksheets. Below, we detail worksheet:
Metadata worksheet: The metadata worksheet provides a description of each study in the database, as well relevant reference material. Columns include:
1. Study_ID: The unique identifier for each included study.
2. Study_Year: The years in which the study was conducted.
3. Country: The country in which the study was conducted.
4. Region: The region where the study was conducted. Authors were asked to describe a fine resolution descriptor of the region in which the study took place.
5. Land_Use_Change: Study authors reported whether or not their study area has undergone a significant amount of land-use change in recent years.
6. Covariates: Study authors reported whether there were any other factors that varied across their study sites that were important predictors of pest-control responses (*e.g.,* local diversity, soil moisture, organic status, *etc.*).
7. Sampling_Design: Study authors were asked to describe how animal sampling was performed in their study.
8. Exclosure_Treatments: Study authors were asked to state if their study involved exclosures, and, if so, to describe the treatments that were included and the number of replicates.
9. Function_Methods: Study authors were asked to describe how enemy activity, pest activity, and crop yield data were collected.
10. Land.use.Source: The source of the land-use map that was used to calculate landscape composition around the study sites (see Karp *et al.* 2018 for more details).
11. Citations: References to any publications that have resulted from study as of May 2018.

Sitedata worksheet: The sitedata worksheet provides a description of each study site and key covariates recorded at those sites.
1. Study_ID: The unique identifier for each included study.
2. Site: A unique identifier for the study site (*i.e.,* one of the places were data was collected).
3. Farm: A unique identifier for the different farms where studies took places. Multiple sites could be located on the same farm.
4. Study_Year: The year (or years) in which the site was sampled.
5. Crop_Species: The crop species that were present at the study sites.
6. Annual_Perennial: Whether the site hosted annual, perennial agriculture, or unknown.
7. Organic: Whether the site was certified organic (yes), conventional (no), or unknown.
8. Local_Diversity: Whether the site was locally diverse (yes), not diverse (no), or unknown. Examples of local diversity include flowering strips, weedy margins, hedgerows, smaller field sizes, more diverse plantings, *etc*.

9. Insecticide_Plot: This column indicates whether the study site was treated with insecticides (yes), not treated (no), or unknown. Insecticides could be applied through sprays, seed coating, planting transgenic crops, *etc.*
10. Insecticide_Farm: This column indicates whether the farm was treated with insecticides (yes), not treated (no), or unknown. Insecticides could be applied through sprays, seed coating, planting transgenic crops, *etc.*
11. Confidence_Insecticide_Farm: This column details the percent confidence the authors had in their assessment that insecticides were or were not present on the farm, using a scale from 0% to 100%. Authors who wrote 'unknown' did not score their confidence.
12. Tilling: Whether the site was tilled (yes), not tilled (no), or unknown during the year of the study.
13. Site_Covariate1: This space was provided if a different site-level covariate was found to be important in an author's study system. Authors included information in this column at their discretion.
14. Site_Covariate2: This space was provided if a second unique site-level covariate was found to be important in an author's study system. Authors included information in this column at their discretion.

Landusedata worksheet: The landusedata worksheet reports the landscape composition around each study site. Nine different land use classes (detailed below) were examined, and we reported the Gaussian-decay-weighted estimates of landscape composition around each site.
1. Study_ID: The unique identifier for each included study.
2. Site: A unique identifier for the study site (*i.e.,* one of the places were data was collected).
3. X: The x coordinate of the study site, in decimal degrees (WGS 1984, unprojected). Some coordinates were masked to protect grower anonymity.
4. Y: The y coordinate of the study site, in decimal degrees (WGS 1984, unprojected). Some coordinates were masked to protect grower anonymity.
5. LC1Gau250: The fraction of forest around the study site, using 250 for the decay rate. When the land-use map did not resolve differences between forests, grassland, and shrubland, 'natural habitat' was used as the category and all sites in the study received a 0 in this category.
6. LC2Gau250: The fraction of grassland around the study site, using 250 for the decay rate. When the land-use map did not resolve differences between forests, grassland, and shrubland, 'natural habitat' was used as the category and all sites in the study received a 0 in this category.
7. LC3Gau250: The fraction of shrubland around the study site, using 250 for the decay rate. When the land-use map did not resolve differences between forests, grassland, and shrubland, 'natural habitat' was used as the category and all sites in the study received a 0 in this category.
8. LC4Gau250: The fraction of annual agriculture around the study site, using 250 for the decay rate. When the land-use map did not resolve differences between annual and perennial agriculture, 'agriculture' was used as the category and all sites in the study received a 0 in this category.
9. LC5Gau250: The fraction of perennial agriculture around the study site, using 250 for the decay rate. When the land-use map did not resolve differences between annual and perennial agriculture, 'agriculture' was used as the category and all sites in the study received a 0 in this category.
10. LC6Gau250: The fraction of other land uses around the study site, using 250 for the decay rate.
11. LC7Gau250: The fraction of urban areas around the study site, using 250 for the decay rate.

12. LC8Gau250: The fraction of 'natural habitats' around the study site, using 250 for the decay rate. This was used when the land use map did not sufficiently resolve forests grasslands, and/or shrublands. Otherwise, it is listed as 0.
13. LC9Gau250: The fraction of 'agriculture' around the study site, using 250 for the decay rate. This was used when the land use map did not sufficiently resolve annual and perennial agriculture. Otherwise, it is listed as 0.
14. LC1Gau750: The fraction of forest around the study site, using 750 for the decay rate. When the land-use map did not resolve differences between forests, grassland, and shrubland, 'natural habitat' was used as the category and all sites in the study received a 0 in this category.
15. LC2Gau750: The fraction of grassland around the study site, using 750 for the decay rate. When the land-use map did not resolve differences between forests, grassland, and shrubland, 'natural habitat' was used as the category and all sites in the study received a 0 in this category.
16. LC3Gau750: The fraction of shrubland around the study site, using 750 for the decay rate. When the land-use map did not resolve differences between forests, grassland, and shrubland, 'natural habitat' was used as the category and all sites in the study received a 0 in this category.
17. LC4Gau750: The fraction of annual agriculture around the study site, using 750 for the decay rate. When the land-use map did not resolve differences between annual and perennial agriculture, 'agriculture' was used as the category and all sites in the study received a 0 in this category.
18. LC5Gau750: The fraction of perennial agriculture around the study site, using 750 for the decay rate. When the land-use map did not resolve differences between annual and perennial agriculture, 'agriculture' was used as the category and all sites in the study received a 0 in this category.
19. LC6Gau750: The fraction of other land uses around the study site, using 750 for the decay rate.
20. LC7Gau750: The fraction of urban areas around the study site, using 750 for the decay rate.
21. LC8Gau750: The fraction of 'natural habitats' around the study site, using 750 for the decay rate. This was used when the land use map did not sufficiently resolve forests grasslands, and/or shrublands. Otherwise, it is listed as 0.
22. LC9Gau750: The fraction of 'agriculture' around the study site, using 750 for the decay rate. This was used when the land use map did not sufficiently resolve annual and perennial agriculture. Otherwise, it is listed as 0.
23. LC1Gau1250: The fraction of forest around the study site, using 1250 for the decay rate. When the land-use map did not resolve differences between forests, grassland, and shrubland, 'natural habitat' was used as the category and all sites in the study received a 0 in this category.
24. LC2Gau1250: The fraction of grassland around the study site, using 1250 for the decay rate. When the land-use map did not resolve differences between forests, grassland, and shrubland, 'natural habitat' was used as the category and all sites in the study received a 0 in this category.
25. LC3Gau1250: The fraction of shrubland around the study site, using 1250 for the decay rate. When the land-use map did not resolve differences between forests, grassland, and shrubland, 'natural habitat' was used as the category and all sites in the study received a 0 in this category.
26. LC4Gau1250: The fraction of annual agriculture around the study site, using 1250 for the decay rate. When the land-use map did not resolve differences between annual and

perennial agriculture, 'agriculture' was used as the category and all sites in the study received a 0 in this category.

27. LC5Gau1250: The fraction of perennial agriculture around the study site, using 1250 for the decay rate. When the land-use map did not resolve differences between annual and perennial agriculture, 'agriculture' was used as the category and all sites in the study received a 0 in this category.
28. LC6Gau1250: The fraction of other land uses around the study site, using 1250 for the decay rate.
29. LC7Gau1250: The fraction of urban areas around the study site, using 1250 for the decay rate.
30. LC8Gau1250: The fraction of 'natural habitats' around the study site, using 1250 for the decay rate. This was used when the land use map did not sufficiently resolve forests grasslands, and/or shrublands. Otherwise, it is listed as 0.
31. LC9Gau1250: The fraction of 'agriculture' around the study site, using 1250 for the decay rate. This was used when the land use map did not sufficiently resolve annual and perennial agriculture. Otherwise, it is listed as 0.

Abundancedata worksheet: The abundancedata worksheet details the abundances of all pests and/or natural enemies recorded at the study sites.
1. Study_ID: The unique identifier for each included study.
2. Site: A unique identifier for the study site (*i.e.,* one of the places were data was collected).
3. Study_Year: The year (or years) in which the site was sampled.
4. Sampling_Date: A unique identifier that differs for each site when it was visited on a different day within the same year.
5. Species: The scientific name of the species that was surveyed.
6. Family: The family of the species that was surveyed.
7. Order: The order of the species that was surveyed.
8. Functional_Group: The functional group of the species that was surveyed. Possibilities include: dominant pest (economically damaging), secondary pest (not strongly damaging but still herbivorous), dominant enemy (a predator or parasitoid that is known to control a dominant pest), parasitoid, or predator.
9. Abundance: The recorded abundance of the species. This value differs across studies, with some reporting the number of individual recorded and others reporting average abundances/counts across multiple surveys.
10. Number_Censuses: The number of sub-samples that were summarized in the abundance column (*e.g.,* three pan traps averaged per plot would be recorded as 3).
11. Sampling_Method: The method used to sample organisms (*e.g.,* pan traps, pitfall traps, mist nets, sticky cards, *etc.*)
12. Abundance_Duration: The number of days that the sampling lasted (*e.g.,* how many days the sticky trap was left in the field).
13. Notes: Any notes that the study's author felt important to include.

Activitydata worksheet**:** The activitydata worksheet details the activity of pests (*e.g.,* crop damage) and enemies (*e.g.,* percent pests parasitized) that was recorded at the study sites.
1. Study_ID: The unique identifier for each included study.
2. Site: A unique identifier for the study site (*i.e.,* one of the places were data was collected).
3. Study_Year: The year (or years) in which the site was sampled.
4. Sampling_Date: A unique identifier that differs for each site when it was visited on a different day within the same year.
5. Function_Type: The broad category of data being collected. Can be one of the categories:

a. Cage experiment (change in infestation): pest abundance recorded within an exclosure experiment
b. Cage experiment (change in crop damage): crop damage recorded within an exclosure experiment
c. Cage experiment (change in crop yield): crop yield recorded within an exclosure experiment
d. Sentinel experiments (enemy activity): the percent predation or parasitism observed after placing sentinel pests in the field or, for parasitism, the fraction of parasitized pests observed naturally in the field.
e. Pest damage: crop damage incurred by pest feeding.

6. Pest: The name of the pest species which is being studied (*i.e.,* the pest that is being consumed or parasitized or the pest that is exacting damage on the crop). If the pest is not differentiated, then 'all' is written.
7. Exclosure_Treatment: If an exclosure is present, this column indicates whether the data reported are from the 'open' treatment (allowing predators to be present) or the 'closed' treatment (excluding predators).
8. Initial_Density: If the study involved exclosures, this column indicates the level of function data before the exclosures were erected (*e.g.,* the starting abundances of the pests).
9. Function_Data: The value of the measured function in the exclosure treatment and/or at the site.
10. Unit: The unit of measurement for the function data (*e.g.,* percentage of pests consumed).
11. Number_Censuses: The number of sub-samples that were summarized in activity column (*e.g.,* three cage traps averaged per site would be recorded as 3).
12. Activity_Duration: The number of days that the sampling lasted (*e.g.,* how many days the sentinel pests was left in the field).
13. Notes: Any notes that the study's author felt important to include.

Yielddata worksheet**:** The yielddata worksheet details the crop yields that were recorded at the study sites.
1. Study_ID: The unique identifier for each included study.
2. Site: A unique identifier for the study site (*i.e.,* one of the places were data was collected).
3. Study_Year: The year (or years) in which the site was sampled.
4. Sampling_Date: A unique identifier that differs for each site when it was visited on a different day within the same year.
5. Function_Data: The reported value of crop yield at a given site.
6. Unit: The unit of measurement for the yield data (*e.g.,* kg/ha).
7. Measure_Type: A variable indicating whether the yield was recorded as the total yield of the plant (*e.g.,* total biomass) or within-plant-yield.
8. Sample_Size: The number of plants or sampled area that went in to calculating the yield estimate.
9. Marketable_or_Total: A variable indicated whether the estimate of yield reported corresponded to the total yield of the crop or just the fraction that was marketable (*e.g.,* due to quality concerns).
10. Notes: Any notes that the study's author felt important to include.
11. CropType: The crop species for which the yield estimate was calculated.
12. WeeksSincePlanting: The number of weeks since planting when the yield estimate obtained in the field.