

IMPACT OF PLANT COMMUNITY DIVERSITY ON AGRICULTURALLY IMPORTANT
ECOSYSTEM SERVICES IN CASH AND COVER CROP SYSTEMS

A Dissertation

Presented to the Faculty of the Graduate School
of Cornell University

In Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy

by

Emily Rose Reiss

May 2018

© 2018

Emily Rose Reiss

IMPACT OF PLANT COMMUNITY DIVERSITY ON AGRICULTURALLY IMPORTANT ECOSYSTEM SERVICES IN CASH AND COVER CROP SYSTEMS

Emily Rose Reiss, Ph. D.

Cornell University 2018

Agroecological theory suggests that environmental impacts of agriculture can be reduced while maintaining yields by increasing diversity in agroecosystems. This increased diversity may support enhanced ecosystem services, reducing the need for inputs such as pesticides and fertilizers. Agricultural researchers have begun to test in agroecosystems what is already well established in unmanaged ecosystems; that more diverse communities, especially those with greater functional trait diversity, are more productive and tend to be higher functioning. Using data collected in the field of cover crop mixtures and a meta-analysis approach, I set out to test the effect of different levels of diversity (intraspecific, interspecific, and functional group), on a set of ecosystem services such as yield, weed suppression, and nitrogen fixation among others. Additionally, I attempted to evaluate the effect of environment on the relationship between diversity and ecosystem service performance, again using multiple methods including a mother-daughter design on working farms.

The meta-analysis of cultivar mixtures showed an overall 2.2% relative yield benefit in food crops, however increased intraspecific diversity had only a minimal effect in cover crop mixtures in the two year field experiment at the research farm or on farms. There was a significant reduction in weed biomass in cultivar mixtures overall. Ultimately, there was no negative impact on any ecosystem service. Cover crop mixtures of multiple species, especially those with both grass and legume functional groups, often had better biomass production

compared to monocultures. The grass/legume mixtures buffered the effects of soil nitrogen fertility much better than either functional group alone for some services. In both the meta-analysis and field trials, more stressful environmental conditions tended to accentuate the diversity effect. In other words, the benefit of more diverse communities was greater in higher stress environments, lending some support to the stress-gradient hypothesis. The on-farm results showed that while functional group mixtures were always as good as or better than monocultures, the precise relationship between these different levels of diversity for the different services varied depending on the location. Overall, my results indicate benefits from all levels of diversity, with more functionally diverse mixtures delivering greater ecosystem service performance, especially in higher stress environments.

BIOGRAPHICAL SKETCH

Emily Rose Reiss was born on December 22, 1985 in Flemington, NJ where she lived until attending the University of Rochester for her undergraduate education. She received her B.S. in Environmental Science with minors in Biology and Spanish in 2008. After working in the international education and energy efficiency fields in New York City, she returned to Upstate New York in 2011 to pursue her PhD under Dr. Laurie E. Drinkwater in the Field of Horticulture at Cornell University. She now lives in Rochester, NY with her husband and young son.

This work is dedicated to my family, who dedicated themselves to my success through this journey.

ACKNOWLEDGMENTS

This dissertation truly would not have been possible without the help of many people. I am grateful to everyone I was fortunate enough to meet along the way, and I apologize to the people I will inevitably leave out.

First, thank you to my family for their unwavering support. You kept me going in more ways than you know, and I hope to be able to express my gratitude and appreciation to you in the years to come.

Thank you to my chair Laurie Drinkwater for introducing me to the field of agroecology, encouraging me to dream big research questions, and for continuing to demand my best from me. Thank you also to my two committee members Katja Poveda and Matthew Ryan. I have appreciated your thoughtful advice, insightful suggestions, and kindness throughout this process.

The current and past graduate students and post-docs in the Drinkwater lab provided invaluable morale, intellectual and material support. Thank you to fellow student and office mate, Bryan Emmett who endured extensive complaints, statistical questions, and distraction from me. Thank you also to Heather Scott the master of the lab, who ensured I had all the materials I needed as well as the mental and emotional fortitude to press on. Brian Caldwell and Carri Marschner were early teachers of field research, as well as great guides to the world of organic farming in general.

Thank you to the numerous undergraduates in the Drinkwater lab that made this project something impossible to do on my own to simply daunting. Their enthusiasm and manpower helped whittle down the piles of biomass grinding, shorten the to-do list, and keep up morale. Thank you especially to Brooke P., Yerae S., Nicole S., Sam C., and Rachel S.

I am thankful for comradery of other graduate students in the Departments of Ecology

and Evolutionary Biology, Entomology, Soil and Crop Sciences and the Soil and Water Lab in Biological and Environmental Engineering, especially Chelsea Morris, Jeff Liebert, and Sonam Sherpa.

Thank you to Paul, Wes and Jeff at Musgrave Research Farm for their help establishing and ensuring the success of my field experiments. I learned a lot from them and will remember their dedication to the farm and research. Thank you also to the amazing people in the Cornell Statistical Consulting Unit, Francois, Erika, and Jay were invaluable to my work and I learned so much from all of you.

A huge debt of gratitude goes out to the many farmers who taught me so much about farming, organic and otherwise. These have been some of the most valuable lessons of my time here at Cornell. Thank you especially to Lou Johns, Melissa Madden, Aaron Munzer and Kara Cusolito, and Thor Oschener for their contributions to my research directly.

Thank you to everyone in the Horticulture field, department, section (which is it now?). I was always glad to have a friendly, cheerful home base, made even better by the free veggies and flowers. Thank you to the staff in general and especially the three graduate field assistants who helped me during my degree, Dolores Higareda, Carol Grove, and Bridget Cristelli. Thank you also to Craig Cramer who taught me a lot about technology, communication, cover crops and how to always keep everything in perspective.

I also would like to acknowledge the financial support from the Graduate School, the Horticulture section for TA positions, and to NESARE for my graduate student grant that supported my research along with support from a USDA-NIFA Organic Transitions Program grant (Project number NYC-145521).

TABLE OF CONTENTS

BIOGRAPHICAL SKETCH	v
ACKNOWLEDGMENTS	vii
PREFACE	x
CHAPTER 1	2
CHAPTER 2	63
CHAPTER 3	129

PREFACE

As agricultural producers are increasingly asked to increase yields while lowering environmental impact, there are numerous suggestions for how to achieve these dual goals. Ecological intensification and similar paradigms contend that more diverse systems support improved ecosystem function and associated services. This allows for reduced input use and thus environmental impact while still producing necessary yields. As modern agroecosystems tend to have low biodiversity, even small increases in diversity have the potential to make a substantial impact on ecosystem functioning. Research in unmanaged ecosystems has produced extensive evidence for a strong relationship between diversity and ecosystem functioning. This relationship is strongest where diversity is measured, typically in plants, in terms of the functional trait diversity present. Additionally, the stress-gradient hypothesis suggests that this positive diversity effect is stronger in more stressful environments where facilitation between individuals is more common and results in even greater productivity and function.

Agriculture research and practice has pushed these results forward. Originating with farmers, the idea of cover crop “cocktails” contends that the more species the better the cover crop’s performance. In recent years there has been a concerted research effort addressing this question of how mixtures of cover crop species can improve the corresponding ecosystem service outcomes. The ecosystem services possible from cover crops cover are numerous, from erosion control, to nutrient management, to pest regulation. Research in this field commonly focuses on total productivity of the cover crop typically measured as aboveground biomass, which is correlated with other services with varying strength. For my research, I chose to measure cover crop biomass and weed suppression, both typically included in studies, as well as nitrogen retention, and nitrogen fixation by legumes which are not commonly measured as part

of cover crop mixture trials.

In my dissertation I continued the line of research connecting diversity and ecosystem function in agroecosystems, focusing on several areas. First, though intraspecific diversity is abundant in agriculture due to crop breeding, it has not received much attention in agroecosystem research. I examined the impact of increased intraspecific diversity in food crops, as well as empirically in cover crop mixtures. Additionally, I wanted to compare the strength of the diversity effect from intraspecific to functional group mixtures, in this case only using cover crop species in a field experiment. Finally, I was also interested in how environmental conditions could influence the diversity response. This has theoretical implications in terms of how the relationship in agroecosystems might be similar or different to unmanaged systems. Additionally, it is critical to understand how consistent and broadly applicable research results may be as this helps farmers make management decisions for their operations.

With the power from over 3500 observations in my meta-analysis of cultivar mixtures in food crops, I was able to detect a 2.2% yield benefit from increased intraspecific diversity. I was not able to observe this effect as strongly in my field experiments, which may have been due to less stressful environments or insufficient trait diversity in the treatments. It was in the meta-analysis that I first noted the increased yield response of diverse mixtures under stress, specifically disease stress. While it was not as strong, I did observe this positive diversity response under stress in the field experiment as well with respect to soil N availability.

In my main field experiment done over two seasons at the research station, I was able to trial a set of treatments more expansive than previous research. I used six species of legumes and grasses along with multiple cultivars of each to extend from simple monocultures to extremely diverse multi-species and multi-cultivar mixtures. As expected, species mixtures had more cover

crop biomass with increasing species richness compared to monocultures. Functional group mixtures had the greatest positive impact on ecosystem services, and helped to buffer environmental differences such as N soil fertility. By putting a subset of these treatments on three local farms as well, I was able to examine the interaction between diversity and environment. While the vetch/wheat mixture was as good as or better than the monocultures across the sites, the treatments responded differently at each farm.

In simplest terms, I would summarize my results as: 1) There is minimal risk to increasing diversity in agroecosystems with potential yield and ecosystem service benefits 2) The environment does impact this relationship, generally with more stressful conditions resulting in greater benefits from diversity.

If I could do it all again without too much realistic consideration for time, money and effort, I would make some changes. First, I would have larger trial plots to allow for establishment of a subsequent crop after the cover crop to determine the actual impact of the cover crop on yield. Additionally, with this initial data I would be more strategic with which treatments I included, and I would include additional ones to fill in gaps. While a full factorial design is impossible, I think a strategic inclusion of certain treatments would make a big difference. I would also have more sites in a given year, both on farms and at other research locations, while also replicating over more than just two years at a selection of those sites. Stability and resilience over time and space is more and more critical for farmers given climate change and shrinking land base.

CHAPTER 1

CULTIVAR MIXTURES: A META-ANALYSIS OF THE EFFECT OF INTRASPECIFIC DIVERSITY ON CROP YIELD

Emily R. Reiss and Laurie E. Drinkwater

Abstract

Extensive research has shown that greater plant community diversity leads to higher levels of productivity and other ecosystem services, and such increased diversity has been suggested as a way to improve yield and agricultural sustainability. Increasing intraspecific diversity with cultivar mixtures is one way to increase diversity in agricultural systems. We examined the relationship between intraspecific diversity and yield in cultivar mixtures using a meta-analysis of 91 studies and >3600 observations. Additionally, we investigated how environmental and management factors might influence this relationship, and if the yield stability of cultivar mixtures differed from that of monocultures. We found that the yield increased by 2.2% overall in cultivar mixtures relative to their monoculture components. Mixtures with more cultivars and those with more functional trait diversity showed higher relative yields. Under biotic stressors, such as disease pressure, and abiotic stressors, such as low levels of soil organic matter and nutrient availability, this diversity effect was stronger, resulting in higher relative yields. Finally, cultivar mixtures generally showed higher yield stability compared to monocultures, especially in response to annual weather variability at a site over time. This practice of mixing cultivars can be integrated into intensified cropping systems where species monocultures dominate, as well as in smallholder cropping systems where low-cost improvements are in demand. Overall, these results suggest that cultivar mixtures are a viable

strategy to increase diversity in agroecosystems, promoting increased yield and yield stability, with minimal environmental impact.

Introduction

There is increasing pressure on global agriculture to increase yields and feed a growing population (Godfray et al. 2010, Tilman et al. 2011). Simultaneously, there is demand to reduce the environmental impact of agricultural production (Tilman et al. 2011). Ecological intensification may be one way to achieve both of these goals (Bommarco et al. 2013, Garibaldi et al. 2016). Ecological intensification uses biological understanding to replace inputs and restore ecosystem functions to agroecosystems and maintain or increase yields (Petersen and Snapp 2015). Increasing diversity within and across agricultural systems is a key principle of ecological intensification proposed to improve agroecosystem performance and minimize the need for external inputs (Bommarco et al. 2013). Agricultural intensification has decreased both spatial and temporal diversity, and as a rule, agricultural systems have reduced plant species diversity within fields and across landscapes (Meyer et al. 2013). Furthermore, because of crop-breeding goals, which aim to optimize varieties for specific environments and agricultural markets, the genetic diversity within fields tends to be very low.

Experiments in unmanaged ecosystems have shown that increased diversity, typically measured as species richness, is positively related to the overall ecosystem functioning of the community, often measured as total productivity (Hooper et al. 2005, Cadotte et al. 2008, 2009, Cardinale et al. 2011, Grace et al. 2016). A high-functioning community is often a more productive one, where all available resources are utilized by the diverse set of individuals present (Hooper et al. 2005). We are still isolating the specific mechanisms responsible for this increase in functioning, but they likely vary by site and community (Grace et al. 2016). However, the

diversity of functional traits present in the community is a very good predictor of ecosystem function (Cadotte et al. 2011). Functional traits ultimately relate to how an organism extracts and utilizes resources from the environment (McGill et al. 2006). Diversity of these traits in a community increases partitioning of the ecosystem resources by organisms, resulting in more complete resource utilization (Cadotte et al. 2011).

Though the types and levels of diversity in agricultural systems may differ from unmanaged systems, we would expect the underlying ecological principles to remain the same, allowing us to enhance agroecosystem functions through increased diversity, which we can do through a number of strategies (Jackson et al. 2007, Costanzo and Barberi 2014, Martin and Isaac 2015, Wood et al. 2015). Many annual crops are currently grown in continuous monocultures or in rotations that include only two plant species. Temporal plant diversity can be increased by expanding crop rotations to include additional cash crop species, as well as species that serve other functions such as cover crops (Davis et al. 2012). Spatial diversity can also be increased within fields, farms, or across landscapes. Hedgerows or other between-field vegetation, as well as the cultivation of multiple crops in a region, can increase landscape-scale diversity (Staley et al. 2013). Intercropping (increasing interspecific diversity by growing two or more crops together in a field) has been shown to increase productivity (Vandermeer 1989, Lithourgidis et al. 2011). Intercropping is not a common practice in mechanized systems due to the greater complexity of managing two or more species, but there are many examples of intercropping in smallholder systems (Lithourgidis et al. 2011). Cultivar mixtures, the simultaneous cultivation of multiple cultivars of the same species, can increase the intraspecific diversity of monocultures by increasing the genetic diversity present in a species and field. This use of intraspecific diversity is well suited to mechanized systems, which are designed to manage

one species at a time, and can provide benefits ranging from reduced disease, weed, and insect pressure as well as improvements in yield and quality (Newton et al. 2009, Grettenberger and Tooker 2015).

Currently, the use of cultivar mixtures in industrial, mechanized agriculture is on the rise and has already become a standard of practice in several regional production systems, mainly for grain production. Cultivar mixtures have been successfully used on a large scale, over 3,000 hectares, to reduce disease and increase yield in rice in China and in wheat in East Germany (Wolfe 1985, Zhu et al. 2000, Newton et al. 2009). In each Poland, Denmark and Switzerland up to 90,000 hectares are planted in cereal cultivar mixtures each year (Newton and Swanston 1999, Newton et al. 2009).

Exploitation of the intraspecific diversity present within crop species has a long history in agricultural research, and crop breeding has led to the development of increasingly specialized crop cultivars (Duvick 2005, Fu 2015). Crop breeding selects for desired characteristics and traits to improve yields, and against traits that limit productivity or are not compatible with agricultural management regimes (e.g., mechanical harvest) or consumer expectations. Thus, while a limited number of crop species dominate food production worldwide, there are many cultivars of these species, which are functionally distinct. For example, the domesticated apple has over 7,500 cultivars. While most are selected for their taste and other palatability traits, disease resistance and climate sensitivity also vary widely (Elzebroek 2008). Even in a primarily vegetatively propagated crop like potato, there remains a diversity of several thousand cultivars and landraces (Brush 1995). Despite this diversity, it was the use of a single-cultivar clone that was responsible for the potato late blight epidemic and ensuing famine in 19th-century Ireland, an extreme example of the consequences of low diversity (Machida-Hirano 2015).

Experiments comparing cultivar mixtures (particularly of small grains) to single cultivars dates back to the first half of the 20th century (Frankel 1939). This work has been concentrated in North America, with substantive contributions from around the world. The basic experimental design has changed very little since early work was done, allowing for comprehensive data collection and comparison. Two reviews and one meta-analysis in the last two decades have shown some of the benefits of cultivar mixtures. The earliest review (Smithson and Lenne 1996) provided qualitative conclusions along with a simple quantitative analysis, and showed a slight yield benefit overall across multiple crops, as well as some disease reduction benefits. More recent studies have focused exclusively on small grains. One review (Mundt 2002) highlighted how cultivar mixtures can successfully reduce disease in small grains. A recent meta-analysis (Kiær et al. 2009) was done on a small dataset of wheat and barley cultivar mixtures (26 studies, all located in temperate regions). As part of the selection criteria, this study only included studies that reported a measure of experimental variation. Overall, it showed a slight yield benefit for cultivar mixtures in these two crops.

In this study, we conducted a global meta-analysis using an extensive database of 94 studies to assess the impact of intraspecific diversity in a wider range of food crops and growing conditions. First, we compared cultivar mixtures to their component monocultures to determine the effect of mixing on yield, a critical ecosystem service of agricultural systems. We also investigated the impact of environmental factors and growing conditions, as well as how experimental and mixture design might influence yield. We expected that cultivar mixtures would have a greater impact on yields under stressful growing conditions (e.g., nutrient or water limitation). Finally, we tested our prediction that cultivar mixtures would have greater yield stability compared to single-cultivar monocultures over multiple seasons and across sites. Our

meta-analysis is the first to examine the yield response and stability of increased intraspecific diversity through cultivar mixtures in a wide range of food crops and environments.

Methods

Using the *Web of Science* database we searched the literature for a variety of search terms to target cultivar mixtures of important annual food crops (excluding rice due to its specialized and varied cultivation) and limited our search to journal articles published in English that fell within the *Web of Science* categories related to ecology or agriculture. While all the results included the matching search terms, there were also papers that did not test cultivar mixtures as the terms were not used in the context we intended (e.g., individual “cultivar” trials treated with a “mixture” of herbicides). As further search term revisions eliminated some of the cultivar mixtures studies that met our criteria, we reviewed individual titles and abstracts to eliminate nearly 90% of the papers that were clearly not studies of cultivar mixtures. We also reviewed relevant studies referenced in the selected studies, including all digitally accessible papers from two previous reviews (Smithson and Lenne 1996, Kiær et al. 2009). We then individually assessed the remaining >200 papers related to the mixing of cultivars. We included papers in the meta-analysis if they met all of the following specific selection criteria: (1) the study was field-based (not conducted in a greenhouse or microcosm) and conducted for at least one full growing season; (2) the study reported either actual yields for all treatments, or relative yield of cultivar mixtures compared to component monocultures; (3) the study included only simultaneous plantings of cultivar mixtures and monocultures, with only one harvest (i.e., not relay planting or multiple cuttings for forage); and (4) a replacement series experimental design was used. Ultimately, 91 papers published between 1939 and 2014 met our four criteria and were included in the meta-analysis (see Appendix S1 for references and full search details). Our database has an

additional 77 papers compared to the previous 2009 meta-analysis (Kiær et al. 2009), expanding the range of crops beyond wheat and barley, as well as the geographic extent beyond the temperate region.

Building the database

We built a database of the yield performance of cultivar mixtures along with information on management and environmental conditions that could influence productivity (Tables 1 and 2). For each observation we recorded year, crop species, crop type (legume, non-legume), cultivar count in mixture (two to nine), evenness of cultivar proportions in seeded mixture (even, uneven), and location of the experimental site. Other variables were not reported consistently in all studies, but where available, we included the following: soil organic matter (converted to percentage if not reported as such), N, P, K fertilizer applied, row spacing, seeding rate (kg/ha or number of seeds/m²), elevation (kilometers above sea level), soil N, P, K levels and soil pH, water management, and disease management. We used DataThief v.1.6 (Tummers 2006) to extract data reported in figures.

We modified and categorized some of the raw data to facilitate analysis of the role played by environmental and management conditions in determining cultivar mixture performance. Latitude was converted to the absolute value, effectively measuring distance from the equator. We categorized climate zone as tropical where the latitude fell between the Tropics of Cancer and Capricorn at 23° and –23° latitude. We approximated the potential of abiotic and biotic stressors using the reported information on soil pH, fertilizer application, water management, and disease management. Soil pH: As optimal soil pH levels are crop specific we categorized site pH as *high*, *optimal*, or *low* using available extension information on recommended growing conditions. For the four crops where pH was reported (seven studies), the optimal ranges were

defined as follows: wheat, 6.0-7.0 (Vitosh 1998); barley, 6.0-6.5 (Mallory and Kersbergen 2013); sugarbeet, 6.5-7.0 (Steinke 2014¹); and field pea, 5.5-7.0 (Pavek 2012). Fertilizer application: While an appropriate indicator of nutrient availability is actual soil nutrient levels, less than 10% of observations reported this information. However, nearly half of the studies provided information about fertilizer use so we designated two levels of nutrient availability for these studies as fertilizer applied: *yes* or *no*. Water management: We grouped studies as *sufficient rain* or *irrigated*, where rainfall was specifically noted by the authors as sufficient or irrigation was used, and *rainfed*, where it was stated that the experiment was rainfed, but without an indication of rainfall amount or actual water deficit. Disease management: Disease pressure was classified as *high* when reported as moderate to high (typically due to no control methods or inoculation with a pathogen), and *low* when reported as absent or minimal, or when a control (i.e., fungicide) was used.

To assess the impact of experimental design features on mixture performance, we compiled information from the introduction and methods sections of the papers. Here authors outlined the experimental purpose and selection basis for the cultivars they used in mixtures, and we based our categories for both on author terminology. First, we noted the goals of the experiment in terms of what the authors hoped cultivar mixtures would do compared to the monocultures (for example, improve yield or reduce disease). We then characterized their rationale for the particular cultivar mixtures they tested. When the authors described specific traits or general basis for cultivar selection and mixing, we categorized that as *specified* and noted the particular characteristics used to construct the mixtures. We classified studies where there was no discussion of mixture rationale as *unspecified*. For those mixtures where the authors

¹ Sugarbeet Soil Fertility, <http://www.soils.msu.edu/wp-content/uploads/2014/05/Sugarbeet-Soil-Fertility-and-Health.pdf>

explained a rationale, classified as *specified*, we also categorized the type of characteristics used to construct mixtures from the component cultivars as either *disease* or *physical*, or *both*. Mixtures created based on physical characteristics included breeding history, heading date, height, lodging susceptibility, growth habit, maturity group, phenology, phenotype, yield potential, and competitive ability against weeds. Anywhere the authors noted the disease response of a cultivar, such as susceptibility or resistance, we categorized the mixtures as having a disease basis. A sizable percentage (~25%) of mixtures considered both disease and physical traits for selection of cultivars.

Few studies reported the complete set of information on experimental design, management practices, and growing conditions necessary for testing all of our hypotheses. When data on yield or critical variables were missing, we contacted the authors and incorporated the data received from these inquiries into the dataset. Despite these measures, it was still not possible to collect the full range of variables for all studies, so we conducted some analyses on subsets of the data. Tables 1 and 2 show all the variables collected and analyzed, categorical and continuous, respectively. Also shown are the results for the measures of within-group (Q_w) and between-group (Q_b) heterogeneity. These are similar to partitioning of variation in an ANOVA; specifically, model sums of squares and error sums of squares, respectively (McDaniel et al. 2014).

Meta-analysis calculations

In meta-analyses, an effect size is calculated to compare evaluate the treatment relative to the control, allowing quantification of trends across a range of experiments and environments. The response ratio, r , which is commonly used as the effect size in meta-analyses is calculated as the mean of the experimental treatment over the mean of the control treatment (Koricheva and

Gurevitch 2014). For cultivar mixtures, the experimental treatment is the actual mixture yield, and the control is the expected mixture yield calculated based on the component monoculture yields. Consequently, r is the same as the relative yield (RY) of the cultivar mixtures.

Relative yield (RY) is the metric most commonly used in competition or mixture experiments to compare the productivity of plants grown as monocultures and those grown in combination with others (Weigelt and Jolliffe 2003). This is a useful index for cultivar mixture trials as it indicates when a mixture is more or less productive than expected based on the mixture components in monoculture. This measure automatically accounts for area and seeding proportions when the monocultures and mixtures are grown at the same seeding density (replacement series design). This design is by far the most commonly used in this field. Relative yield for each mixture was calculated as

$$RY = Y_{mx} / (Y_{mo1} P_1 + Y_{mo2} P_2, \text{ etc.})$$

where Y_{mx} is the total yield of the mixture, Y_{mo1} is the yield of cultivar 1 in monoculture, and P_1 is the proportion of cultivar 1 in the mixture. A $RY > 1$ indicates a yield benefit from mixing, a $RY < 1$ indicates a yield penalty from mixing, and a $RY = 1$ indicates no change in yield from mixing compared to the component monoculture yields. Converting RY to percentage change in yield for the mixture compared to the component monocultures is calculated as $(RY - 1) \cdot 100$.

We calculated RY using the above formula for the majority of the studies. A small percentage, less than 15% of studies, reported only RY without reporting the corresponding component monoculture yields. For these we took the RY directly as reported in the study. Our final database includes 3,612 mixture treatments, “observations”, from 94 experiments reported in 91 published papers.

With RY as the effect size, the measure of the magnitude of the effect of mixing, we used

MetaWin version 2.1 software (Rosenberg et al. 2000) to explore the mean response of RY to a variety of environmental and experimental variables (Tables 1 and 2). As $RY = r$, and RY is already normally distributed (Figure 1), no transformation was needed to change it to a normally distributed effect size (Tonitto et al. 2006). Very few papers reported study variance. Consequently, we performed an unweighted meta-analysis, assigning a variance of 1 to all observations. To improve the statistical significance of our results without experimental variance, we calculated the mean RY and a bias-corrected 95% confidence interval using a bootstrapping method with 9999 iterations (Tonitto et al. 2006, McDaniel et al. 2014). Means are considered significant if the 95% confidence interval does not cross 1 (Adams et al. 1997). Significant Q_b values indicate significant differences between classes, where nonoverlapping confidence intervals can be used as an informal evaluation to distinguish significant contrasts (Scheiner and Gurevitch 1993). However, classes are considered significantly different where Q_b is significant even when confidence intervals overlap (see Figure 5b, for example, and Table 1). Following convention, we acknowledge the lack of independence between mixture observations from the same experiment, from studies in the same journal, and studies by the same author. We have not modified the dataset, but rather used a more conservative significance level ($P < 0.03$; Gurevitch et al. 1992). We ran regression analyses of RY and the continuous variables (Table 2) with study as a random effect. As this effect was not significant, we removed it from the final regression results. All regression analyses were performed using R Version 3.1.2 (R Core Team 2014).

Before conducting further analyses of environmental and management effects using subsets of the data, we checked for bias, outliers, and confounding variables. The vast majority (80%) of studies had fewer than 50 observations, with only six studies containing more than 100

observations. We used funnel plots to test for bias, with number of observations in the study as the explanatory variable. Overall, the shape of the scatter remained consistent over the range of observation values, and did not suggest any bias in the dataset (Philibert et al. 2012, McDaniel et al. 2014). We identified two observations as outliers in the data for RY. They both came from the same study and were the results of near-complete crop failure of the monoculture plots for three of the four cultivars in one year at one site. Due to extremely low monoculture yields, the RYs of the mixtures were skewed high and did not accurately reflect the effect of mixtures as represented by all other experimental results. We removed these two points (greater than eight standard deviations away from the mean) from the dataset. Finally, when we assessed each variable for its effect on RY, we examined possible confounding variables, especially when there was a small sample size. Where there may have been confounding factors, we made a note in the figure legend or the text.

Yield stability analysis

A subset of the meta-analysis dataset was used to assess the yield stability in monocultures and mixtures in response to varying environmental conditions (e.g., soils, precipitation, rainfall) across spatial and temporal scales. In crop production, stability is the maintenance of yield or productivity at a consistent amount in the face of differing environmental conditions (Newton et al. 2009). For this analysis, we included studies comparing monocultures and mixtures over multiple years at a single site, as well as those across multiple sites for one year.

For the yield stability analysis, we used actual yields rather than relative yields to compare the monocultures and mixtures. We calculated the coefficient of variation (CV) for each cultivar monoculture or mixture either across multiple years for each site (stability over time), or

across multiple sites for each year (stability across different environments). Where studies included both multiple sites and multiple years, data were included in both analyses. Additionally, where there were multiple treatments at one site (for example, crops inoculated with disease and disease controlled with fungicide), we separated these treatments for the stability calculation so that this additional treatment did not add to the variability. We averaged the CV for all the monocultures and all the mixtures across sites or years for that study and treatment. We then compared the average monoculture CV to that of the mixtures in a study and classified it as either higher or lower. We counted these classifications to assess the proportion (χ^2 test) of experiments where the variability of the monocultures was higher or lower than that of the mixtures. We also did a Wilcoxon signed rank test on the average CV for monocultures and mixtures for stability across space and across time.

Results

We found studies that tested cultivar mixtures across a wide variety of crops and agricultural regions, but small grains in temperate regions dominated the dataset. Small grains accounted for 80% of the mixture observations (wheat, 43%; barley, 20%; oats, 17%), with soybean and corn accounting for 10.5% and 3.5% of the observations, respectively. For the common vegetable crops that we searched, we found no cultivar mixture studies that met our criteria. Most experiments were conducted in temperate regions, with North America accounting for 80% of the observations, with Europe (7.8%) and Asia (6.8%) making up the next largest groups. Oceania, Africa, and South America accounted for very small percentages of the observations (2.7%, 2.2%, 0.7%, respectively).

Many studies of cultivar mixtures share similar characteristics and research goals. Replacement series designs and relative yield have been used from the earliest study in 1939

through to the present. Interestingly, the work on cultivar mixtures has not evolved over this time to include substantive investigations into the mechanisms underlying differences in ecosystem services between cultivar monocultures and mixtures. The experimental purpose varied across studies and often included multiple goals. Yield improvement, particularly in terms of quantity and/or stability, was the most common purpose (91%) for testing cultivar mixtures, with disease control and grain quality improvement accounting for the remainder. Though not included in the dataset because they did not report yield, an increasing number of studies in the last 20 years have investigated the ability of cultivar mixtures to provide supporting ecosystem services such as disease reduction, insect pest and weed suppression, and improved water use efficiency. Only a handful of studies specifically evaluated yield stability, though more than half of the studies reported results from either multiple years or multiple sites or both. There was very little overlap in the cultivars used for these experiments so we could not reach any conclusions about whether specific cultivars are better suited for use in mixtures. For example, the 42 studies of wheat mixtures used a total of 77 cultivars with only 13 cultivars used in more than one study.

Relative yield response and impact of growing conditions

Overall, mixtures yielded 2.2% (RY = 1.0217) more than expected based on their monoculture yields (Figure 1). We observed a significant yield increase for all crops tested in three or more studies, with the exception of sorghum (Figure 2). Though they have different

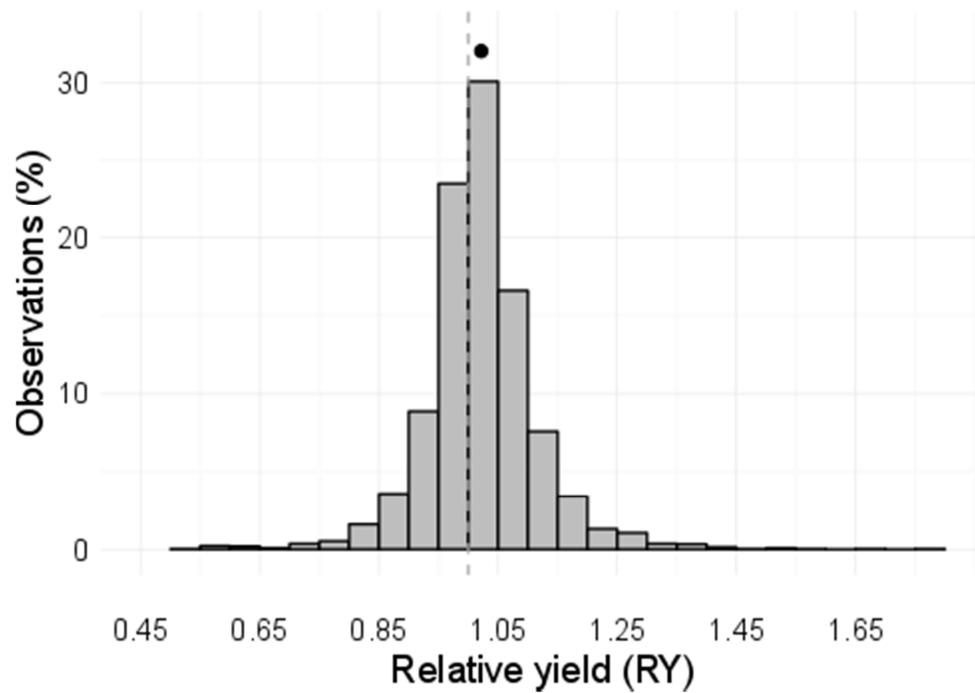


Figure 1. Distribution of cultivar mixture observations by relative yield (RY). Overall mean shown at RY = 1.0217 (black dot) with 95% bias-corrected confidence interval (1.0187 – 1.0247), though not visible due to point size. Dashed line at RY = 1, indicating no change in yield from mixing.

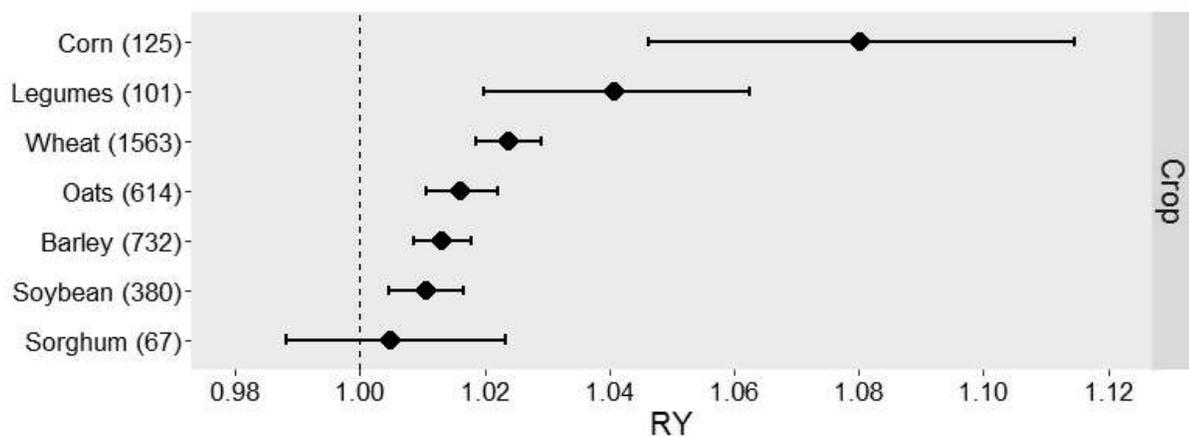


Figure 2. Effect of crop on RY of mixtures. Mean values and 95% bias-corrected confidence intervals for the RY are shown, with the number of mixture observations in parentheses. Dashed line at $RY = 1$, indicating no change in yield from mixing. The legumes group includes common bean, common vetch, cowpea, field pea, moth bean, all represented by one study each. Not shown here are one study of rye and sugarbeet each. All other crops or groups are represented by three or more studies.

nutrient acquisition and utilization traits, legume and non-legume crops did not differ in their RY (Table 1). Overall, the RYs of mixtures in the dataset closely follow a normal distribution. Less than 7% of the mixtures had RY reductions of greater than 10%, which made it difficult to detect patterns leading to such yield losses. In contrast, 14% had RY increases greater than 10%. While most studies only examined two- or three-cultivar mixtures, resulting in mean RYs of 1.02, mixtures with four or more cultivars had a much higher mean RY (RY = 1.05; Figure 3a). Relative yields did not change with year over the range of the dataset, showing the consistency of the effect of cultivar mixtures across time, even with substantial changes in management, such as breeding and input use, over the seven decades (Table 2).

Experimental design

Some decisions around experimental design and management can have a substantial impact on RY of mixtures, while others are insignificant. The experimental purpose and/or intent of the study did not have a significant impact on the RY of the mixtures in the trials (Table 1). However, mixtures constructed with a specified basis for selecting particular cultivars had higher RYs compared to mixtures where no rationale was stated (Figure 3b). Planned mixtures based on both disease and physical characteristics were significantly better than those mixtures based on either a physical or a disease basis alone (Figure 3c). Management decisions related to seeding of cultivars in mixtures (row spacing and seeding rates) had no effect on the RY outcome. The evenness of the mixture did not have a significant impact on the RY, though even mixtures had a greater mean and were almost four times more frequent in the dataset (Table 2).

Table 1. Number of observations (n), between-group heterogeneity (Q_b), and P values for relative yield (RY) of categorical variables analyzed as part of meta-analysis of cultivar mixtures. Significant Q_b values indicate significant differences between classes (Scheiner and Gurevitch 1993).

Variable	Levels	RY		
		n	Q_b	P
Crop	Barley, Corn, Legumes, Oats, Sorghum, Soybean, Wheat	3582	0.6065	<<0.001
Crop type	Legume, Non-legume	3612	0.0117	0.254
Number of cultivars in mixture	Two, Three, Four or more	3612	0.1169	0.002
Mixture intention	Specified basis, Unspecified basis	3612	0.0486	0.020
Mixture composition basis	Disease, Physical, Both	2554	0.3615	<<0.001
Soil pH	High, Optimal, Low	303	0.2317	<<0.001
Fertilizer applied	No, Yes	1432	0.0476	0.020
Water management	Rainfed, Sufficient rain or irrigated	1553	0.000	0.950
Disease pressure	High, Low	1315	0.0494	0.015
Mixture evenness	Even, Uneven	3612	0.0064	0.406
Purpose of experiment	Disease, Height, Lodging, Management, Maturity, Mixing ability, Seeding proportions, Spacing, Stability, Yield	3612	0.024	0.971
Climate zone	Temperate, Tropical	3612	0.6522	<<0.001

Note: The legumes group includes common bean, common vetch, cowpea, field pea, moth bean, all represented by one study each.

Table 2. Number of observations (n) and regression results (intercept, slope, slope P values, and adjusted R^2) for relative yield (RY) and continuous variables analyzed as part of meta-analysis of cultivar mixtures.

Variable	n	Intercept*	Slope	P	Adj. R^2
Soil organic matter (%)	240	1.075	-0.028	<<0.001	0.075
Fertilizer applied, N (kg/ha)	1148	1.004	0.000	0.018	0.004
Fertilizer applied, P (kg/ha)	1013	1.018	0.000	0.048	0.003
Fertilizer applied, K (kg/ha)	957	1.017	0.000	0.085	0.002
Row spacing (cm)	2827	1.008	0.000	<<0.001	0.007
Seeding rate (kg/ha)	236	0.991	0.000	0.014	0.022
Seeding rate (no. seeds/m²)	2624	1.030	0.000	0.016	0.002
Year	3180	1.685	0.000	0.016	0.002
Latitude	3332	1.070	-0.001	<<0.001	0.012
Kilometers above sea level	459	0.991	0.038	0.005	0.015
Soil N (ppm)	259	1.021	0.000	0.079	0.008
Soil P (ppm)	275	1.017	0.000	0.972	-0.004
Soil K (ppm)	86	1.030	0.000	0.561	-0.008

Note: * indicates $P << 0.001$ for all.

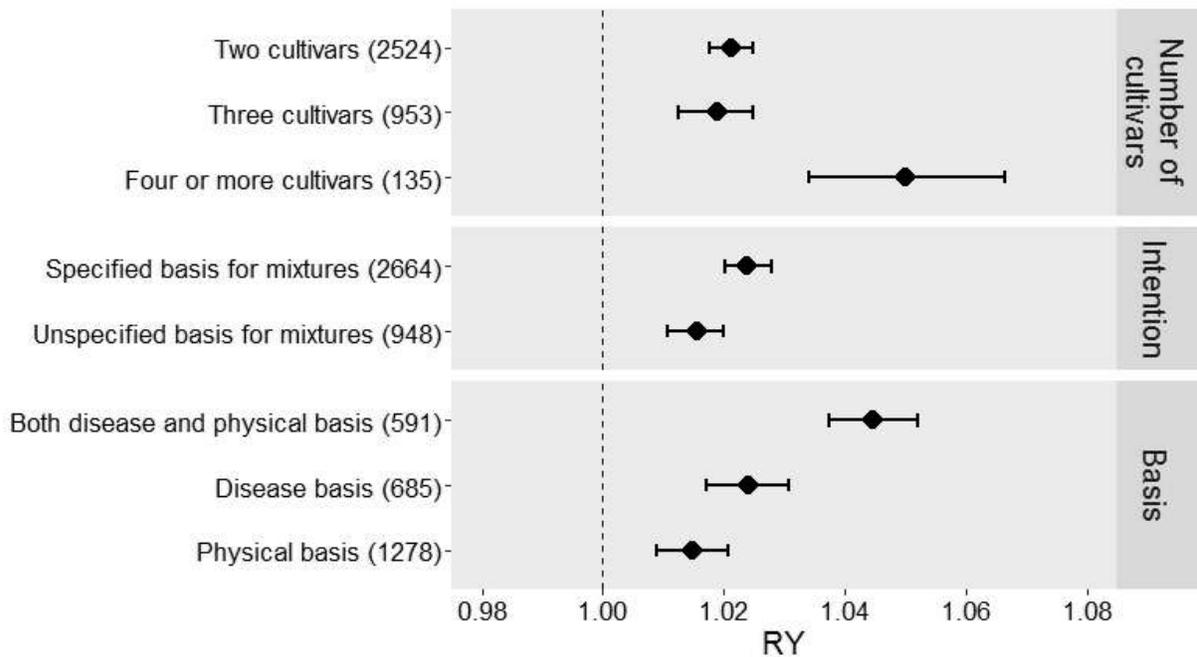


Figure 3. Effect of mixture composition on RY of mixtures. (a) Number of cultivars in mixture (The four or more cultivar group contains observations from 20 studies, across eight crops.). (b) Mixture intention, *specified* or *unspecified* by study author. (c) Mixture composition basis, characteristics of cultivars in mixtures with a specified basis. Mean values and 95% bias-corrected confidence intervals for the RY are shown, with the number of mixture observations in parentheses.

Effect of abiotic and biotic stressors

We found that a variety of environmental characteristics and related biotic and abiotic stressors influenced the performance of cultivar mixtures and RY. Reporting of certain variables was inconsistent across studies, making it necessary to use subsets of the data to test hypotheses about the role of environmental and biotic stress (see *Data limitations and further research*). We found a negative correlation between soil organic matter (SOM) and RY (Figure 4), suggesting that in environments where nutrient supply from organic matter mineralization may be more limited, mixtures resulted in greater yield benefits. Soil pH levels that were below crop-optimum levels positively affected RY, while soils with pH levels above crop optima substantially reduced the RY of mixtures (Figure 5a), possibly due to reduced availability of nutrients such as phosphorus. Following the trend suggested by the impacts of SOM content and pH on RY, we found that RY was greater in studies where no fertilizers were applied (Figure 5b). When we used actual fertilizer rates as a continuous variable, the slope of the relationship between amount of fertilizer applied and RY was near zero for N, P, and K (Table 2). Soil nutrient content (N, P, and K) had no detectable relationship with RY, similar to fertilizer applications (Table 2). We also did not detect any differences in the RY of irrigated mixtures compared to those grown under rainfed conditions (Figure 5c).

Disease pressure was the only biotic stress we were able to quantify in the database. Mixtures in environments with high disease pressure had greater RY compared to those grown under conditions with little or no disease pressure (Figure 5d). However, when mixtures were intentionally constructed based on disease characteristics, there was no difference in RY under high or low disease pressure (Figure 6a). This may be due to poor selection of cultivars, or limited resistance. It is possible that mixtures reduced disease compared to the monocultures,

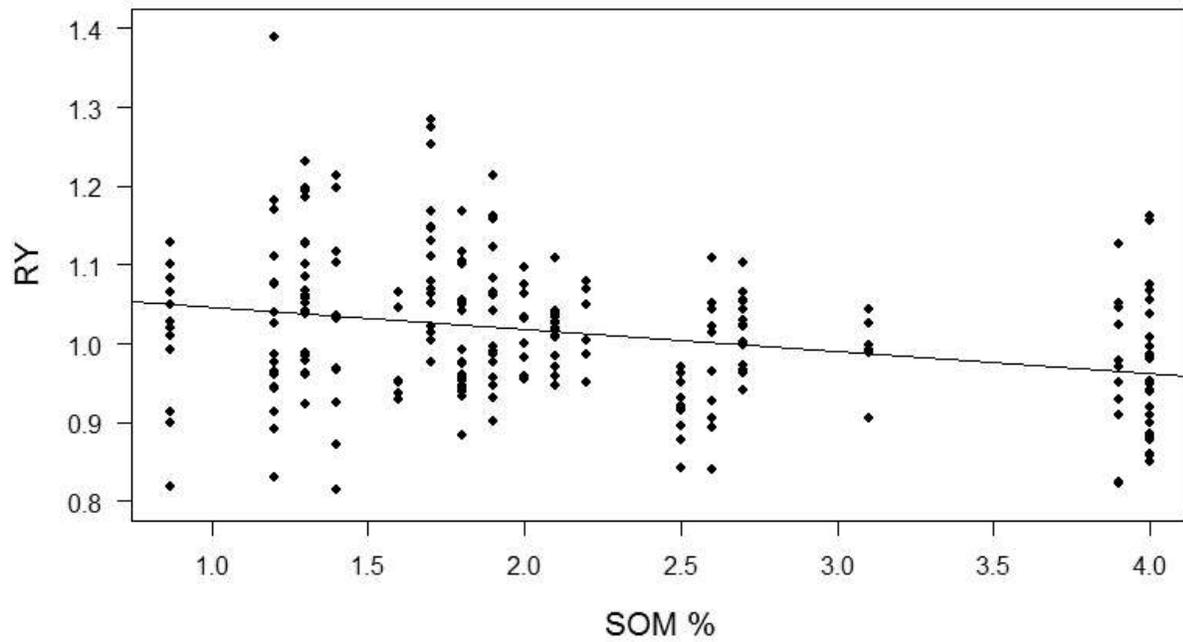


Figure 4. The relationship between percent soil organic matter (SOM %) and RY shows an increase in RY with lower SOM % (black line, $R^2=0.0748$, $F_{1, 238}=20.32$, $P<<0.001$, $y=1.075 - 0.0282x$). Data from six papers (240 observations), encompassing sites with 17 unique SOM levels.

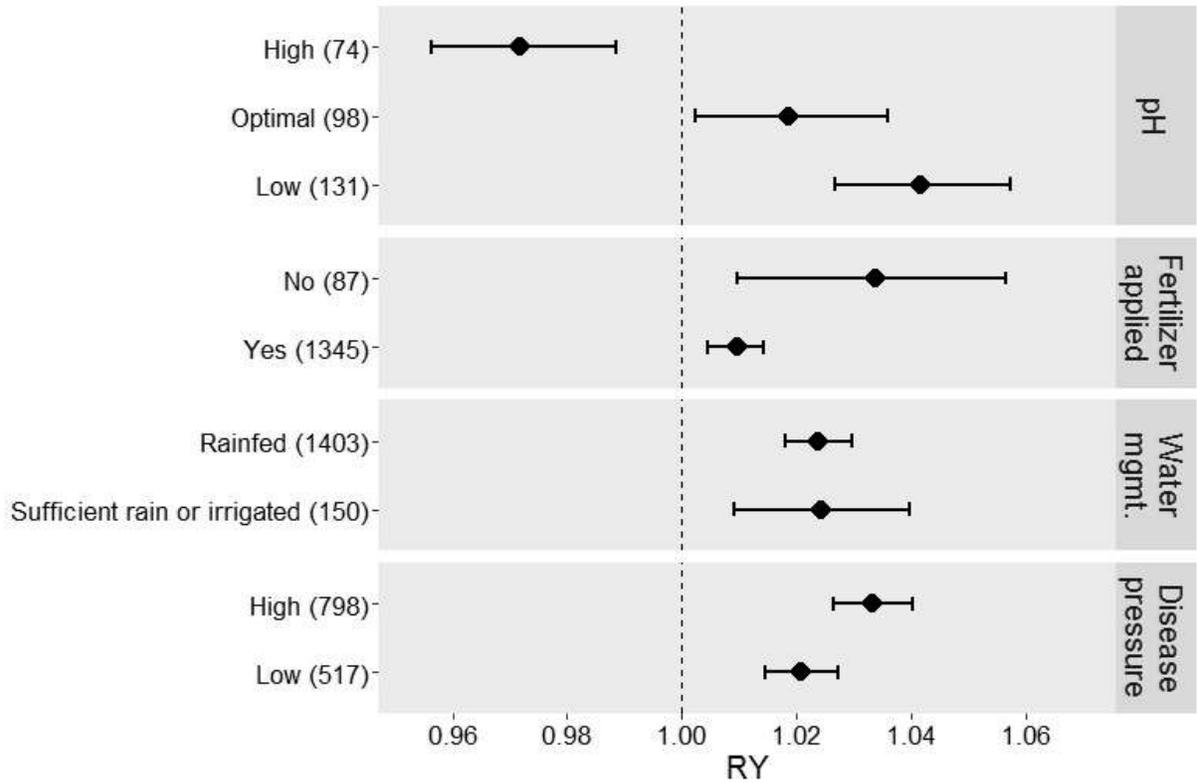


Figure 5. Effect of environmental stress on RY of mixtures. (a) The pH of soil categorized by suitability for individual crops. Both the high and low pH categories consist of observations from three unique studies, while one study containing multiple sites provided observations in all three categories. (b) Nutrient stress evaluated as fertilizer applied or not. The group with no fertilizer applied included six studies of three crops and three continents). (c) Water stress evaluated as water management: only rainfed, or sufficient rain or irrigated. (d) Disease stress evaluated as disease pressure: high (when disease was present and/or no control was used) or low (when control was applied and/or no disease was present). Mean values and 95% bias-corrected confidence intervals for the RY are shown, with the number of mixture observations in parentheses.

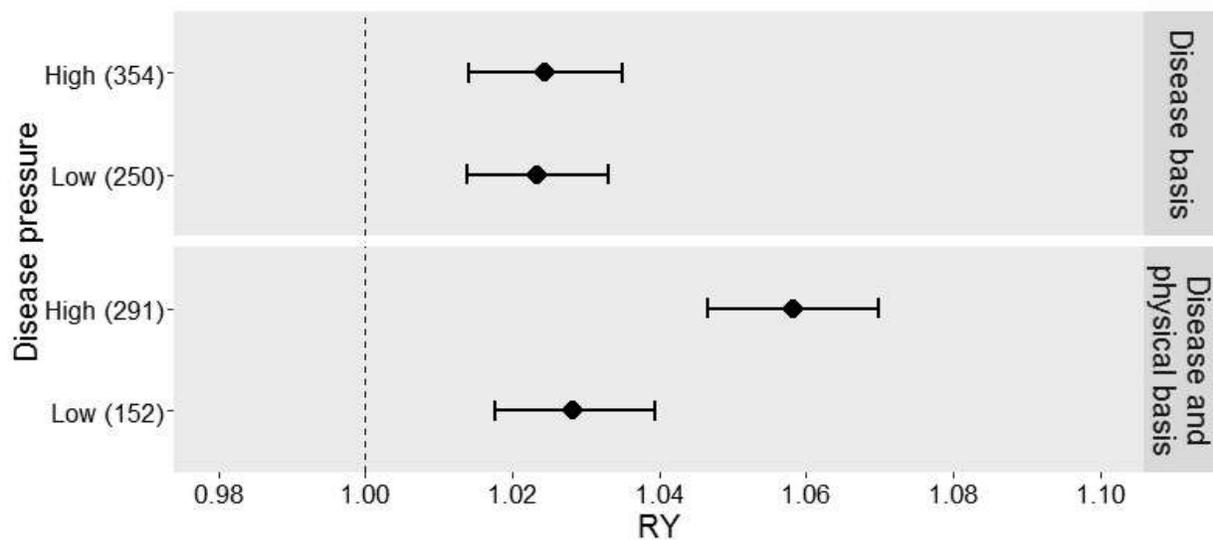


Figure 6. Effect of mixture basis on RY of mixtures under high and low disease pressure. (a) Mixtures based on disease characteristics of component monocultures. (b) Mixtures based on both disease and physical characteristics. Mean values and 95% bias-corrected confidence intervals for the RY are shown, with the number of mixture observations in parentheses.

without a corresponding yield increase. However, where mixtures were constructed based on both disease and physical traits, there was a significant RY increase under high disease pressure conditions. In environments with low disease pressure, RY did not differ between mixtures constructed with only a disease basis and those with a combined disease and physical basis (Figure 6b).

The effects of larger scale abiotic conditions on RY, such as those dependent on latitude, had a greater effect than those driven by elevation. Those studies conducted closer to the equator in the tropical region had significantly higher RYs than those studies conducted in the temperate region, which constituted the vast majority of observations (Figure 7). Latitude as a continuous variable follows the trend suggested by the climate zones with a weak, but significant, negative relationship with RY (Table 2). The weak and insignificant relationship between the elevation of the experimental site and RY, suggests that latitudinal position has a stronger influence on RY than the effects of elevation (Table 2).

Yield stability analysis

Overall, we found that, compared to mixtures, monocultures tended to have greater yield variability, as measured by average CV of yield (Figure 8). However, the yield stabilizing effect of cultivar mixtures in response to seasonal fluctuations in weather in one location over time was stronger compared to the response to variation in one season from environmental conditions across a study area (Figure 8). In fact, there was no difference between average CV of monocultures and mixtures over multiple sites in the same year (Figure 8). The percentage of trials where the monoculture is more variable than the mixture was significant only when examined over time (61%, $\chi^2 = 5.24$, $df = 1$, $P = 0.022$), and not over space (59%, $\chi^2 = 1.59$, $df = 1$, $P = 0.208$). Using both average CV and the percentage of trials showing increased variability

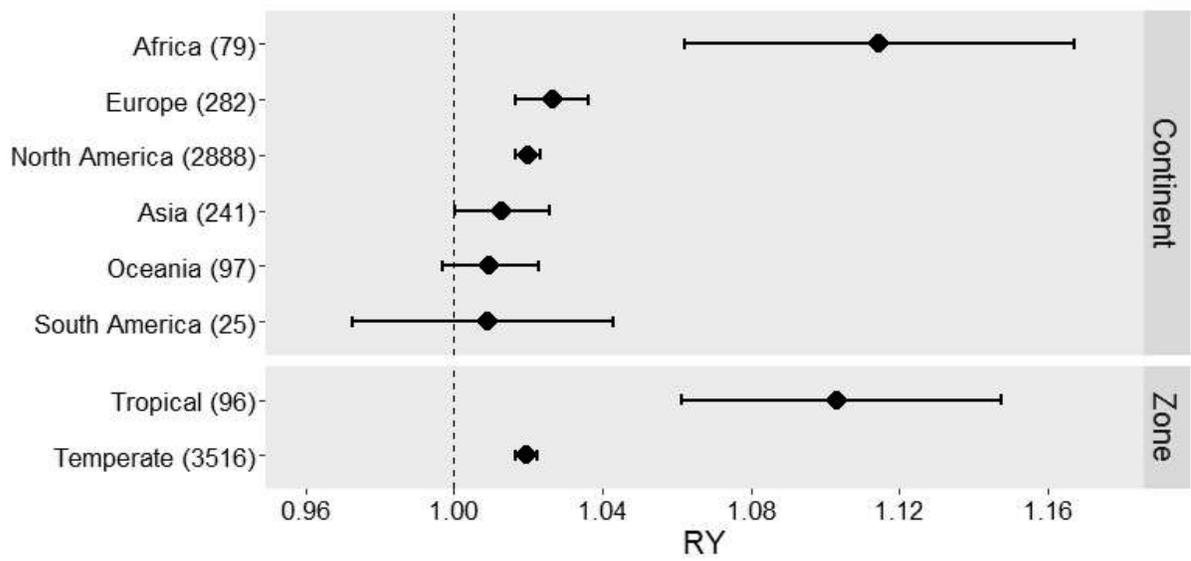


Figure 7. Effect of climate zone on RY of mixtures. Mean values and 95% bias-corrected confidence intervals for the RY are shown, with the number of mixture observations in parentheses.

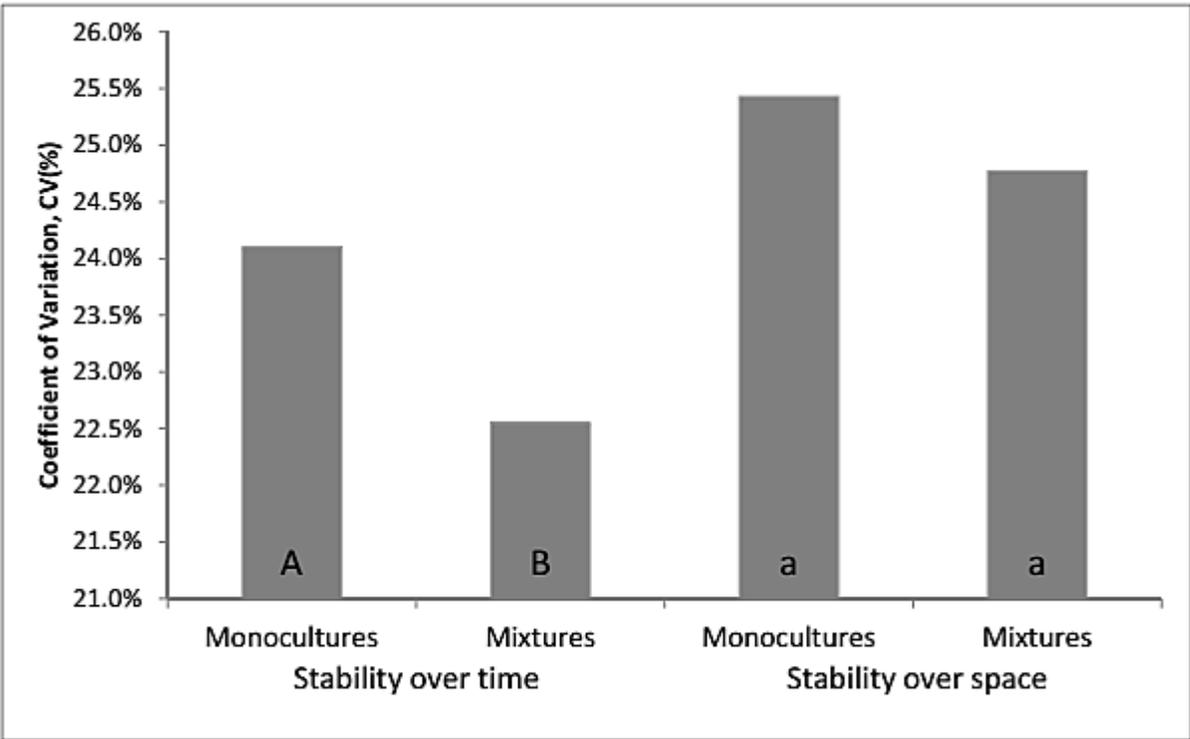


Figure 8. Yield stability over time (multiple years at the same site, $n=2,191$ observations) and over space (multiple sites in the same year, $n=1,412$ observations). Bars show the average coefficient of variation (CV), with monocultures more variable than mixtures when examined over time only (paired Wilcoxon signed rank test, over time, $V = 3198$, $P = 0.035$; and over space: $V = 757$, $P = 0.380$).

in monocultures, mixtures have a stronger stabilizing effect on yield over multiple growing seasons, compared to their weaker effect on yield stabilization over a geographic area in one growing season (Figure 8).

Discussion

Reincorporating diversity into agroecosystems to promote ecosystem services is one viable approach for reducing environmental impacts while maintaining and even increasing yields (Kremen and Miles 2012). The practice of planting cultivar mixtures, which increases intraspecific diversity in monoculture fields where diversity is very low, contributes to increased overall diversity from the field to landscape scale. This additional diversity promotes ecosystem services, including increasing and stabilizing yield. The 2.2% overall yield increase we found for cultivar mixtures compared to the expected yield from their component monocultures is small, but comparable to the average annual rate of yield gain due to plant breeding improvements of between 1% and 3% (Fernandez Cornejo 2004). The prospect of this small yield gain from breeding regularly drives farmers to purchase the newest cultivars. The fact that the RY benefit from mixtures has not changed over the seven decades in this review shows that the practice of mixing cultivars is robust and compatible with the consistent improvement of plant genetics and other changes in management (Table 2). In addition to the overall yield benefit that we found, stressful environments appear to strengthen the positive diversity response, fitting with the stress-gradient hypothesis (Li et al. 2007, He et al. 2013, Tang et al. 2016). Finally, we found that the yield stability of mixtures from one growing season to the next is generally higher than that of monocultures, a potentially important factor for farmers as climate-influenced environment conditions become more variable.

Increased yield

While we are limited in our ability to isolate the specific mechanisms responsible for the RY increase observed in cultivar mixtures, the ability of the community as a whole to maximize available resource use through distinct functional traits is likely a major driver. Functional traits characterize an organism's response to the environment and/or effect on the ecosystem functioning, which relates to resource use by individuals and ultimately the community as a whole (Díaz and Cabido 2001). Functional diversity encompasses the values and range of these traits in the community (Díaz and Cabido 2001). A more functionally diverse community has greater resource partitioning, potentially utilizing the available resources more efficiently and ultimately increasing overall productivity and function (Cadotte et al. 2011). Recent reviews of studies in unmanaged ecosystems, both aquatic and terrestrial, have confirmed an overall relationship between various measures of diversity, including species richness and functional diversity, as well as productivity response (Weiss et al. 1990, Cadotte et al. 2011, Cardinale et al. 2011). As cultivars are specifically bred for desirable and distinct functional traits, we would expect mixtures of cultivars to be functionally diverse, and display a similar relationship between functional diversity and productivity. Specifically, we would expect cultivar richness, as well as targeted trait selection in mixtures, to increase functional diversity and ultimately productivity.

Three lines of evidence from this study demonstrate the positive relationship between functional diversity in cultivar mixtures and increases in RY. First, we saw that higher cultivar richness, specifically mixtures with four or more cultivars, results in higher RY compared to mixtures with a richness of two or three, and all had increased RY compared to the monocultures (Figure 3a). Second, beyond the simple richness effect, the impact of cultivar mixtures is increased when they are specifically designed to include greater functional diversity. The

specification of a basis or rationale for selection of specific cultivars in the mixtures by the authors suggests an awareness and intentionality related to traits that may result in more functionally diverse mixtures. We do in fact see greater RYs in mixtures constructed with a specified basis for the selection of component cultivars compared to mixtures without such a specified basis (Figure 3b). Third, where mixture composition was intentionally based on both disease and physical traits, we see a substantial increase in RY (Figure 3c). We would expect these mixtures to have greater functional trait diversity, compared to mixtures based on either disease or physical traits in isolation, as the traits for disease resistance or physical characteristics rarely overlap, providing a larger range of functional traits. This independence of disease and physical traits is supported by the fact that there was no difference in the RYs of mixtures where cultivar selection was based on these traits alone.

Stress and environmental conditions

The response of cultivar mixtures to certain environmental conditions and stress may be responsible for the wide range of RY responses we found in the dataset (Figure 1). The insurance hypothesis predicts that ecosystem function will be maintained in more diverse communities due to the divergent responses of species (or other organismal groupings) to environmental conditions (Naeem and Li 1997, Yachi and Loreau 1999, Jackson et al. 2007). Though the mechanisms responsible for this maintenance of function may differ slightly at the cultivar level, there is likely overlap with those suggested for higher levels of diversity (e.g., species). Additionally, landraces (heterogeneous populations of locally adapted crop species) have long been known to be more successful than modern cultivars in stressful conditions (Newton et al. 2010, Dwivedi et al. 2016). The genetic and phenotypic diversity in these landraces is very high, resulting in great functional diversity (Dwivedi et al. 2016). Therefore, we expected that under

stressful environments cultivar mixtures might provide greater benefits in terms of yield outcomes. Fundamentally, stressful conditions may make the improved function and associated productivity of a diverse community more apparent, as efficient use and sharing of resources is more important. As an example, one experiment found no relationship between species richness and productivity of bryophytes under constant conditions, but under drought conditions, higher species richness increased survivorship of all species, due to facilitation (Mulder et al. 2001). The results of a recent meta-analysis of global plant communities support the stress-gradient hypothesis and showed that most plant interactions respond to stress, and typically shift towards facilitation and reduced competition (He et al. 2013).

Across the range of abiotic stresses we were able to quantify in the database, we found a clear trend towards higher RY outcomes under more stressful conditions, but the strength varied between the different stresses. The lack of consistent reporting of environmental conditions in the studies reduced our ability to make strong conclusions (see *Data limitations and further research*). However, we were able to analyze the impact of two key soil characteristics: soil organic matter (SOM) and pH. SOM correlates with higher water-holding capacity, improved aeration, and greater aggregate stability of soils. SOM is not only a key source of plant nutrients during decomposition, but it also helps to retain nutrients in the soil, and subsequently enhance their availability to plants by increasing cation exchange capacity (Hudson 1994, Reeves 1997). We observed a clear trend of increasing RY with decreasing SOM levels, which would indicate conditions that are more stressful for plant growth (Figure 4). We tried to isolate the effect of nutrient stress specifically in two ways: fertilizer application and soil nutrient status. The small set of unfertilized mixtures had a substantially higher mean RY compared to those where fertilizer was applied (Figure 5b), and even with the large range, was significantly different from

the fertilized set (Table 1). The lower RYs for mixtures under more ideal, fertilized conditions suggest that there is less benefit of mixtures when resources are not limiting. This may be because the differential and more complete root exploitation by cultivar mixtures is not as necessary in these better conditions. Alternately, mixtures could yield as well as monocultures with less inputs, reducing associated environmental impacts and costs (Elser et al. 2014). Soil nutrient status showed no significant trends for N, P, or K, likely due in large part to the under-reporting of these conditions, which were available for less than 10% of observations in dataset (Table 2).

Our results showing a significant impact of soil pH on RY also support this relationship between plant nutrient stress and diversity. Mixtures performed substantially better in more acidic conditions than basic ones (Figure 5a). Macronutrient (N, P, K, S, Ca) concentrations tend to be less in acidic soils (Brady and Weil 2008). In contrast, micronutrient cations are often more available at lower pH, but can sometimes become too available and toxic, especially manganese and aluminum (Barak et al. 1997, Fageria et al. 2002, Brady and Weil 2008). Where pH is low, the coexistence of multiple cultivars may be able to ameliorate the potentially toxic conditions. First, multiple cultivars are able to more completely exploit the soil profile with different root architectures, increasing the overall production and distribution of root exudates. These exudates promote microbial decomposition, the products of which form complexes with aluminum, reducing its toxicity (De la Fuente-Martínez and Herrera-Estrella 1999, Brady and Weil 2008).

We observed an enhanced yield response under the biotic stress of high disease pressure, especially in the most functionally diverse mixtures. Overall, mixtures experiencing disease stress had higher RYs, perhaps because of mortality and compensation by other cultivars, or some form of facilitation reducing the spread of disease (Figure 5d; Brophy and Mundt 1991).

Mixtures constructed with a disease-trait basis had the same RY response regardless of the disease pressure (Figure 6a). However, the more functionally diverse mixtures had significantly higher RY responses under high disease compared to low disease pressure (Figure 6b). These functionally diverse mixtures were those constructed with both disease and physical traits in mind. The benefit of increased functional diversity in these mixtures was most important in this high-stress (high disease) environment. From these two comparisons we might conclude that not only is the diversity effect strengthened under stressful conditions, but also that the inverse is true: We may not see a stress response if the community is not sufficiently diverse (as in the mixtures with a disease basis). With disease pressure expected to increase with climate change, this benefit of mixtures may be particularly relevant (Tripathi et al. 2016).

Yield stability

Stability (low variation of yield over time or space) under stressful or less than ideal conditions is gaining in importance for breeding, but is rarely a goal on its own, as achieving acceptable yields when conditions are favorable remains paramount. Landraces have persisted for hundreds to thousands of years, not because of their high yields, but rather their ability to adapt to environmental conditions and maintain adequate yields (Newton et al. 2010, Dwivedi et al. 2016). In environments where uniform conditions are achieved with inputs and where there is a favorable climate, modern cultivars excel and outperform the genotypically and functionally diverse landraces (Mohammadi et al. 2014, Yahiaoui et al. 2014). As we found in this study, cultivar mixtures with their increased intraspecific diversity generally showed improved yield stability compared to monoculture (Figure 8). Cultivar mixtures may be a way to integrate the yield benefits of modern breeding with the stability from genetic diversity similar to landraces.

Ecologically, our understanding of the relationship between stability and diversity is still

developing. A recent review found that both productivity and stability over time increased with increasing diversity across a range of unmanaged ecosystems; however, the effects were independent of one another (Cardinale et al. 2013). Mechanistically, there is not yet a clear understanding of what drives this diversity–stability relationship. However, the insurance hypothesis again provides a useful framework. For example, cultivar-specific mortality in an early drought allows better adapted cultivars to exploit this additional space, compensating for the loss of individuals and maintaining productivity (Cadotte et al. 2012). This may be particularly true for small grains with the ability to tiller and fill space left by less successful cultivars. Facilitation may also play a role, with some cultivars providing a more hospitable growing environment for others that might otherwise fail under the given conditions (Mulder et al. 2001).

In our study, we observed a much stronger yield-stabilizing effect of diversity in response to weather variability as opposed to the broader environmental variability of sites in a region. We assessed yield stability by splitting the dataset in two ways, each capturing a different type of variability. Yield stability of the cultivars and mixtures over multiple seasons at one site primarily reflects the response to weather variability, specifically, tolerance or resistance to annual variations in rainfall and temperature, keeping constant other environmental conditions (i.e., other ecosystem state factors such as soil properties). In contrast, yield stability across multiple sites in one season reflects the response to the variability of environmental conditions across an experimental region *in addition to* localized weather variability. While we might expect mixtures to increase the yield stability across sites, as different cultivars might each thrive in different sites and conditions, we are not able to separate this variation from that of the weather. Additionally, the interactions between weather variability and site conditions may have

inhibited our ability to detect a stronger diversity effect across sites (Figure 8). Where only annual weather variation is concerned, there is a clear stability-over-time advantage of mixtures compared to monocultures. We see this in terms of both average CV and the percentage of trials where monocultures were more variable (Figure 8). Recent climate modeling of crop performance has shown that, as annual weather variation increases, so does yield variability (Porter and Semenov 2005). More specifically, when temperature variation is increased and mean temperature held constant, yield variability responds much more dramatically than absolute yield losses (Porter and Semenov 2005). Weather variability is a powerful driver of yield-stability outcomes, and mixtures appear to be able to buffer some of that variability.

Data limitations and further research

As our technical ability to analyze large data sets through meta-analyses and other methods increases, it is critical that all researchers report complete experimental methods and site conditions in as much detail as possible. This additional information will allow us to advance our understanding of the mechanisms underlying the effect of intraspecific diversity on yield and other ecosystem functions. Specifically, authors should prioritize the reporting of basic soil and climate conditions for the site, along with standard management practices including fertilizer and pesticide applications and water management. With these data and an improved understanding of mechanisms, we can better isolate the conditions where cultivar mixtures will be either beneficial or detrimental to RY. For example, with more extensive reporting of soil pH and SOM our findings would be applicable to a larger range of conditions. Additionally, we could have investigated interactions between these basic soil characteristics with management practices such as fertilizer application, a critical assessment for practical applications.

A specific challenge to advancing our understanding of mechanisms in cultivar mixtures

is the difficulty of separating the component cultivars in the mixture after harvest for measurement and analysis. The few studies in this meta-analysis that were able to separate cultivars used characteristics such as seed color, or developed clever methods of planting and hand-harvesting to track cultivars in the mixture and maintain separation at harvest (Brophy and Mundt 1991, Finckh et al. 1999, Worster and Mundt 2007, Fang et al. 2014). More work along these lines will help us to understand the mechanisms behind improved, or depressed, RY in cultivar mixtures.

In addition to a more detailed understanding of the drivers behind yield increases in cultivar mixtures, we should broaden our scope to include other important ecosystem services. Though still limited, more papers are looking at the relationship between intraspecific diversity and water use efficiency (Song et al. 2010, Haghshenas et al. 2013, Fang et al. 2014, Adu-Gyamfi et al. 2015) and insect pest regulation (Weiss et al. 1990, Vera et al. 2013, Pan and Qin 2014). Specifically, compared to mono-cultivar planting, cultivar mixtures can reduce the abundance of herbivore pests such as aphids and whiteflies, which have similar characteristics to pathogens and thus may be controlled well with intraspecific diversity (Tooker and Frank 2012). However, there is less data on the potential for cultivar mixtures to affect natural enemy populations and this area particularly warrants additional attention (Jones et al. 2011, Tooker and Frank 2012). Other services relevant to agroecosystems that may respond to increased diversity include nutrient retention and use efficiency, soil organic matter accumulation, weed suppression, and crop pollination.

Management implications

Diversifying our monoculture-dominated landscape with cultivar mixtures is a tractable first step toward ecological intensification for farmers globally. In the long term, increasing

species diversity and other larger scale spatial and temporal diversification strategies aimed at increasing agricultural diversity will likely be more important and effective in enhancing ecosystem services (Davis et al. 2012). Cultivar mixtures can be integrated into mechanized and input-intensive systems where single-crop cultivation is the norm, and increased yield potential a constant goal. Specifically, mixtures may be one tool to reduce external inputs, such as fertilizers and pesticides, which are often fossil-fuel intensive and known to have negative effects on surrounding environments (Tilman et al. 2011, Schipanski et al. 2016). This may result in increased profitability, as there are fewer costs, with a similar or increased yield. Additionally, cultivar mixtures may help to manage pest resistance as demonstrated by the use of Bt and non-Bt corn to create integrated refuges. As of 2014, almost half of all growers surveyed in 2014 exclusively planted seed mixtures for this purpose (*Grettenberger and Tooker 2015*). Of course, proper selection of cultivars is important to ensure a similar maturation time and compatibility with existing mechanical management such as combine height settings for harvesting beans.

Wider acceptance of cultivar mixtures in the marketplace would likely increase adoption of mixtures by growers. Modern malting operations prefer single cultivars as there is the assumption of greater homogeneity, however this view is increasingly challenged with mixtures delivering equivalent quality for a range of metrics (Newton and Swanston 1999, Newton et al. 2009). Additionally, for other uses such as alcohol production for distilling or biofuels, single cultivars are still the norm, but there may be no benefit for pure batches, and mixtures may in fact allow for greater total production (Newton et al. 2009). Cultivar mixtures can be indistinguishable from single cultivars in terms of baking and end-loaf quality for bread wheat (Manthey and Fehrmann 1993, Mille et al. 2006). While more research demonstrating the viability of cultivar mixtures for a range of end uses would be helpful, collaborations that ensure

growers will have a buyer for their cultivar mixtures will likely do more to advance cultivar mixture acreage.

For small-scale farmers or those in low-input systems where stress may be more intense, the diversity benefits for yield would have an even greater impact. It is not uncommon for farmers in these systems to grow both modern varieties as well as local varieties or landraces (Kolech et al. 2015). Additionally, as farmers are already familiar with cultivating landraces tailored to specific regional and farm-level conditions, creating cultivar mixtures with available modern varieties already grown in the area might be an accessible next step. This intraspecific diversity increase is also compatible with existing technologies and practices to reduce the yield gap, as the fundamental structure of the plant community is not different from existing monoculture production. As these smallholder and low-input systems occupy more heterogeneous environments compared to more typical, mechanized systems, increased reliance on participatory breeding, rather than traditional, centralized breeding, will be more important (Dawson et al. 2008).

Our results showing increased yield stability with cultivar mixtures have important implications for management strategy, especially as climate change is projected to result in more variable weather and environmental extremes. Mixtures alone may not substantially reduce the variability associated with multiple fields on a farm, but they can reduce the variability that might be experienced over years within one field. For farmers concerned about yield stability, planting a portion of their farm in cultivar mixtures may help to provide consistency in production in the face of weather variability. The importance and impact of yield stability may differ for farmers around the world. For farmers in the developed world, consistent and predictable yields influence planting decisions, cash flow, and long-term farm viability (Koesling

et al. 2004). In the developing world, widely variable yields can have a direct impact on food security, as well as market prices and cash flow (Sinebo 2005, Asrat et al. 2010). Cultivar mixtures are a practical way for all farmers potentially to reduce their yield variability over time.

Acknowledgments

This research was supported by USDA-NIFA Organic Transitions Program (Project number NYC-145521). We would like to thank Jennifer Blesh for feedback on the manuscript, as well as two anonymous reviewers for their comments.

Literature Cited

- Adams, D. C., J. Gurevitch, and M. S. Rosenberg. 1997. Resampling tests for meta-analysis of ecological data. *Ecology* 78:1277–1283.
- Adu-Gyamfi, P., T. Mahmood, and R. Trethowan. 2015. Can wheat varietal mixtures buffer the impacts of water deficit? *Crop and Pasture Science* 66:757.
- Asrat, S., M. Yesuf, F. Carlsson, and E. Wale. 2010. Farmers' preferences for crop variety traits: Lessons for on-farm conservation and technology adoption. *Ecological Economics* 69:2394–2401.
- Barak, P., B. O. Jobe, A. R. Krueger, L. A. Peterson, and D. A. Laird. 1997. Effects of long-term soil acidification due to nitrogen fertilizer inputs in Wisconsin. *Plant and Soil* 197:61–69.
- Bommarco, R., D. Kleijn, and S. G. Potts. 2013. Ecological intensification: harnessing ecosystem services for food security. *Trends in Ecology & Evolution* 28:230–238.
- Brady, N. C., and R. R. Weil. 2008. *The Nature and Properties of Soils*. 14th edition. Pearson, Prentice Hall.

- Brophy, L. S., and C. C. Mundt. 1991. Influence of plant spatial patterns on disease dynamics, plant competition and grain yield in genetically diverse wheat populations. *Agriculture, ecosystems & environment* 35:1–12.
- Brush, S. B. 1995. In situ conservation of landraces in centers of crop diversity. *Crop Science* 35:346–354.
- Cadotte, M. W., B. J. Cardinale, and T. H. Oakley. 2008. Evolutionary history and the effect of biodiversity on plant productivity. *Proceedings of the National Academy of Sciences* 105:17012–17017.
- Cadotte, M. W., K. Carscadden, and N. Mirotchnick. 2011. Beyond species: functional diversity and the maintenance of ecological processes and services. *Journal of Applied Ecology* 48:1079–1087.
- Cadotte, M. W., J. Cavender-Bares, D. Tilman, and T. H. Oakley. 2009. Using Phylogenetic, Functional and Trait Diversity to Understand Patterns of Plant Community Productivity. *PLoS ONE* 4:e5695.
- Cadotte, M. W., R. Dinnage, and D. Tilman. 2012. Phylogenetic diversity promotes ecosystem stability. *Ecology* 93:S223–S233.
- Cardinale, B. J., K. Gross, K. Fritschie, P. Flombaum, J. W. Fox, C. Rixen, J. van Ruijven, P. B. Reich, M. Scherer-Lorenzen, and B. J. Wilsey. 2013. Biodiversity simultaneously enhances the production and stability of community biomass, but the effects are independent. *Ecology*.
- Cardinale, B. J., K. L. Matulich, D. U. Hooper, J. E. Byrnes, E. Duffy, L. Gamfeldt, P. Balvanera, M. I. O'Connor, and A. Gonzalez. 2011. The functional role of producer diversity in ecosystems. *American Journal of Botany* 98:572–592.

- Costanzo, A., and P. Bàrberi. 2014. Functional agrobiodiversity and agroecosystem services in sustainable wheat production. A review. *Agronomy for Sustainable Development* 34:327–348.
- Davis, A. S., J. D. Hill, C. A. Chase, A. M. Johanns, and M. Liebman. 2012. Increasing Cropping System Diversity Balances Productivity, Profitability and Environmental Health. *PLoS ONE* 7:e47149.
- Dawson, J. C., K. M. Murphy, and S. S. Jones. 2008. Decentralized selection and participatory approaches in plant breeding for low-input systems. *Euphytica* 160:143–154.
- De la Fuente-Martínez, J. M., and L. Herrera-Estrella. 1999. Advances in the Understanding of Aluminum Toxicity and the Development of Aluminum-Tolerant Transgenic Plants. *Advances in Agronomy* 66:103–120.
- Díaz, S., and M. Cabido. 2001. Vive la difference: plant functional diversity matters to ecosystem processes. *Trends in Ecology & Evolution* 16:646–655.
- Duvick, D. N. 2005. The contribution of breeding to yield advances in maize (*Zea mays* L.). *Advances in agronomy* 86:83–145.
- Dwivedi, S. L., S. Ceccarelli, M. W. Blair, H. D. Upadhyaya, A. K. Are, and R. Ortiz. 2016. Landrace Germplasm for Improving Yield and Abiotic Stress Adaptation. *Trends in Plant Science* 21:31–42.
- Elser, J. J., T. J. Elser, S. R. Carpenter, and W. A. Brock. 2014. Regime Shift in Fertilizer Commodities Indicates More Turbulence Ahead for Food Security. *PLOS ONE* 9:e93998.
- Elzebroek, A. T. G. 2008. *Guide to Cultivated Plants*. CABI.

- Fageria, N. K., V. C. Baligar, and R. B. Clark. 2002. Micronutrients in Crop Production. Pages 185–268 *in* D. L. Sparks, editor. *Advances in Agronomy*. Academic Press.
- Fang, Y., B. Xu, L. Liu, Y. Gu, Q. Liu, N. C. Turner, and F. M. Li. 2014. Does a mixture of old and modern winter wheat cultivars increase yield and water use efficiency in water-limited environments? *Field Crops Research* 156:12–21.
- Fernandez Cornejo, J. 2004. The Seed Industry in U.S. Agriculture: An Exploration of Data and Information on Crop Seed Markets, Regulation, Industry Structure, and Research and Development. Page 81. *Agriculture Information Bulletin*, ERS USDA GOV, Washington, DC.
- Finckh, M. R., E. S. Gacek, H. J. Czembor, and M. S. Wolfe. 1999. Host frequency and density effects on powdery mildew and yield in mixtures of barley cultivars. *Plant Pathology* 48:807–816.
- Frankel, O. H. 1939. Analytical yield investigations on New Zealand wheat: IV. Blending varieties of wheat. *The Journal of Agricultural Science* 29:249–261.
- Fu, Y.-B. 2015. Understanding crop genetic diversity under modern plant breeding. *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik* 128:2131–2142.
- Garibaldi, L. A., L. G. Carvalheiro, B. E. Vaissière, B. Gemmill-Herren, J. Hipólito, B. M. Freitas, H. T. Ngo, N. Azzu, A. Sáez, J. Åström, J. An, B. Blochtein, D. Buchori, F. J. C. García, F. O. da Silva, K. Devkota, M. de F. Ribeiro, L. Freitas, M. C. Gaglianone, M. Goss, M. Irshad, M. Kasina, A. J. S. P. Filho, L. H. P. Kiill, P. Kwapong, G. N. Parra, C. Pires, V. Pires, R. S. Rawal, A. Rizali, A. M. Saraiva, R. Veldtman, B. F. Viana, S.

- Witter, and H. Zhang. 2016. Mutually beneficial pollinator diversity and crop yield outcomes in small and large farms. *Science* 351:388–391.
- Godfray, H. C. J., J. R. Beddington, I. R. Crute, L. Haddad, D. Lawrence, J. F. Muir, J. Pretty, S. Robinson, S. M. Thomas, and C. Toulmin. 2010. Food Security: The Challenge of Feeding 9 Billion People. *Science* 327:812–818.
- Grace, J. B., T. M. Anderson, E. W. Seabloom, E. T. Borer, P. B. Adler, W. S. Harpole, Y. Hautier, H. Hillebrand, E. M. Lind, M. Pärtel, J. D. Bakker, Y. M. Buckley, M. J. Crawley, E. I. Damschen, K. F. Davies, P. A. Fay, J. Firn, D. S. Gruner, A. Hector, J. M. H. Knops, A. S. MacDougall, B. A. Melbourne, J. W. Morgan, J. L. Orrock, S. M. Prober, and M. D. Smith. 2016. Integrative modelling reveals mechanisms linking productivity and plant species richness. *Nature* 529:390–393.
- Grettenberger, I. M., and J. F. Tooker. 2015. Moving beyond resistance management toward an expanded role for seed mixtures in agriculture. *Agriculture, Ecosystems & Environment* 208:29–36.
- Gurevitch, J., L. L. Morrow, A. Wallace, and J. S. Walsh. 1992. A Meta-Analysis of Competition in Field Experiments. *The American Naturalist* 140:539–572.
- Haghshenas, A., Y. Emam, H. Ghadiri, S. A. Kazemeini, and A. A. Kamgar-Haghighi. 2013. Effect of mixed cropping of an early- and a middle-ripening wheat cultivar on mitigation of competition during post-anthesis moisture stress. *Journal of Agricultural Science and Technology* 15:491–503.
- He, Q., M. D. Bertness, and A. H. Altieri. 2013. Global shifts towards positive species interactions with increasing environmental stress. *Ecology Letters* 16:695–706.

- Hooper, D. U., F. S. Chapin, J. J. Ewel, A. Hector, P. Inchausti, S. Lavorel, J. H. Lawton, D. M. Lodge, M. Loreau, S. Naeem, B. Schmid, H. Setälä, A. J. Symstad, J. Vandermeer, and D. A. Wardle. 2005. Effects of biodiversity on ecosystem functioning: A consensus of current knowledge. *Ecological Monographs* 75:3–35.
- Hudson, B. D. 1994. Soil organic matter and available water capacity. *Journal of Soil and Water Conservation* 49:189–194.
- Jackson, L. E., U. Pascual, and T. Hodgkin. 2007. Utilizing and conserving agrobiodiversity in agricultural landscapes. *Agriculture, Ecosystems & Environment* 121:196–210.
- Jones, T. S., E. Allan, S. A. Härril, J. Krauss, C. B. Müller, and F. J. F. van Veen. 2011. Effects of genetic diversity of grass on insect species diversity at higher trophic levels are not due to cascading diversity effects. *Oikos* 120:1031–1036.
- Kiær, L. P., I. M. Skovgaard, and H. Østergård. 2009. Grain yield increase in cereal variety mixtures: A meta-analysis of field trials. *Field Crops Research* 114:361–373.
- Koesling, M., M. Ebbesvik, G. Lien, O. Flaten, P. S. Valle, and H. Arntzen. 2004. Risk and risk management in organic and conventional cash crop farming in Norway. *Acta Agriculturae Scandinavica, Section C — Food Economics* 1:195–206.
- Kolech, S. A., D. Halseth, W. D. Jong, K. Perry, D. Wolfe, F. M. Tiruneh, and S. Schulz. 2015. Potato Variety Diversity, Determinants and Implications for Potato Breeding Strategy in Ethiopia. *American Journal of Potato Research* 92:551–566.
- Koricheva, J., and J. Gurevitch. 2014. Uses and misuses of meta-analysis in plant ecology. *Journal of Ecology* 102:828–844.

- Kremen, C., and A. Miles. 2012. Ecosystem Services in Biologically Diversified versus Conventional Farming Systems: Benefits, Externalities, and Trade-Offs. *Ecology and Society* 17.
- Li, L., S.-M. Li, J.-H. Sun, L.-L. Zhou, X.-G. Bao, H.-G. Zhang, and F.-S. Zhang. 2007. Diversity enhances agricultural productivity via rhizosphere phosphorus facilitation on phosphorus-deficient soils. *Proceedings of the National Academy of Sciences* 104:11192–11196.
- Lithourgidis, A. S., C. A. Dordas, C. A. Damalas, and D. N. Vlachostergios. 2011. Annual intercrops: an alternative pathway for sustainable agriculture. *Australian Journal of Crop Science* 5:396–410.
- Machida-Hirano, R. 2015. Diversity of potato genetic resources. *Breeding Science* 65:26–40.
- Mallory, E., and R. Kersbergen. 2013. Bulletin #1027, Growing Organic Barley in Maine. University of Maine Cooperative Extension, Orono, ME.
- Manthey, R., and H. Fehrmann. 1993. Effect of cultivar mixtures in wheat on fungal diseases, yield and profitability. *Crop Protection* 12:63–68.
- Martin, A. R., and M. E. Isaac. 2015. Plant functional traits in agroecosystems: a blueprint for research. *Journal of Applied Ecology*.
- McDaniel, M. D., L. K. Tiemann, and A. S. Grandy. 2013. Does agricultural crop diversity enhance soil microbial biomass and organic matter dynamics? a meta-analysis. *Ecological Applications*.
- McGill, B. J., B. J. Enquist, E. Weiher, and M. Westoby. 2006. Rebuilding community ecology from functional traits. *Trends in Ecology & Evolution* 21:178–185.

- Meyer, S., K. Wesche, B. Krause, and C. Leuschner. 2013. Dramatic losses of specialist arable plants in Central Germany since the 1950s/60s - a cross-regional analysis. *Diversity and Distributions* 19:1175–1187.
- Mille, B., M. B. Fraj, H. Monod, and C. Vallavieille-Pope. 2006. Assessing Four-Way Mixtures of Winter Wheat Cultivars from the Performances of their Two-Way and Individual Components. *European Journal of Plant Pathology* 114:163–173.
- Mohammadi, R., R. Haghparast, B. Sadeghzadeh, H. Ahmadi, K. Solimani, and A. Amri. 2014. Adaptation Patterns and Yield Stability of Durum Wheat Landraces to Highland Cold Rainfed Areas of Iran. *Crop Science* 54:944.
- Mulder, C. P. H., D. D. Uliassi, and D. F. Doak. 2001. Physical stress and diversity-productivity relationships: The role of positive interactions. *Proceedings of the National Academy of Sciences* 98:6704–6708.
- Mundt, C. C. 2002. Use of Multiline Cultivars and Cultivar Mixtures for Disease Management. *Annual Review of Phytopathology* 40:381–410.
- Naeem, S., and S. Li. 1997. Biodiversity enhances ecosystem reliability. *Nature* 390:507–509.
- Newton, A. C., T. Akar, J. P. Baresel, P. J. Bebeli, E. Bettencourt, K. V. Bladenopoulos, J. H. Czembor, D. A. Fasoula, A. Katsiotis, K. Koutis, M. Koutsika-Sotiriou, G. Kovacs, H. Larsson, M. A. A. P. de Carvalho, D. Rubiales, J. Russell, T. M. M. Dos Santos, and M. C. Vaz Patto. 2010. Cereal landraces for sustainable agriculture. A review. *Agronomy for Sustainable Development* 30:237–269.
- Newton, A. C., G. S. Begg, and J. S. Swanston. 2009. Deployment of diversity for enhanced crop function. *Annals of Applied Biology* 154:309–322.

- Newton, A. C., and J. S. Swanston. 1999. Cereal variety mixtures reduce inputs and improve yield and quality-why isn't everybody growing them? Scottish Crop Research Institute:55.
- Pan, P., and Y. Qin. 2014. Genotypic diversity of soybean in mixed cropping can affect the populations of insect pests and their natural enemies. *International Journal of Pest Management* 60:287–292.
- Pavek, P. L. S. 2012. Plant fact sheet for pea (*Pisum sativum* L.). USDA - Natural Resources Conservation Service, Pullman, WA.
- Petersen, B., and S. Snapp. 2015. What is sustainable intensification? Views from experts. *Land Use Policy* 46:1–10.
- Philibert, A., C. Loyce, and D. Makowski. 2012. Assessment of the quality of meta-analysis in agronomy. *Agriculture, Ecosystems & Environment* 148:72–82.
- Porter, J. R., and M. A. Semenov. 2005. Crop responses to climatic variation. *Philosophical Transactions of the Royal Society B: Biological Sciences* 360:2021–2035.
- R Core Team. 2014. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Reeves, D. W. 1997. The role of soil organic matter in maintaining soil quality in continuous cropping systems. *Soil and Tillage Research* 43:131–167.
- Rosenberg, M. S., D. C. Adams, and J. Gurevitch. 2000. MetaWin: Statistical software for meta-analysis.
- Scheiner, S. M., and J. Gurevitch, editors. 1993. Design and analysis of ecological experiments. Chapman and Hall, New York.

- Schipanski, M. E., G. K. MacDonald, S. Rosenzweig, M. J. Chappell, E. M. Bennett, R. B. Kerr, J. Blesh, T. Crews, L. Drinkwater, J. G. Lundgren, and C. Schnarr. 2016. Realizing Resilient Food Systems. *BioScience* 66:600–610.
- Sinebo, W. 2005. Trade off between yield increase and yield stability in three decades of barley breeding in a tropical highland environment. *Field Crops Research* 92:35–52.
- Smithson, J. B., and J. M. Lenne. 1996. Varietal mixtures: a viable strategy for sustainable productivity in subsistence agriculture. *Annals of Applied Biology* 128:127–158.
- Song, L., D.-W. Zhang, F.-M. Li, X.-W. Fan, Q. Ma, and N. C. Turner. 2010. Soil Water Availability Alters the Inter- and Intra-Cultivar Competition of Three Spring Wheat Cultivars Bred in Different Eras. *Journal of Agronomy & Crop Science* 196:323–335.
- Staley, J. T., J. M. Bullock, K. C. R. Baldock, J. W. Redhead, D. A. P. Hoofman, N. Button, and R. F. Pywell. 2013. Changes in hedgerow floral diversity over 70 years in an English rural landscape, and the impacts of management. *Biological Conservation* 167:97–105.
- Tang, X., S. A. Placella, F. Daydé, L. Bernard, A. Robin, E.-P. Journet, E. Justes, and P. Hinsinger. 2016. Phosphorus availability and microbial community in the rhizosphere of intercropped cereal and legume along a P-fertilizer gradient. *Plant and Soil* 407:119–134.
- Tilman, D., C. Balzer, J. Hill, and B. L. Befort. 2011. Global food demand and the sustainable intensification of agriculture. *Proceedings of the National Academy of Sciences* 108:20260–20264.
- Tonitto, C., M. B. David, and L. E. Drinkwater. 2006. Replacing bare fallows with cover crops in fertilizer-intensive cropping systems: A meta-analysis of crop yield and N dynamics. *Agriculture, Ecosystems & Environment* 112:58–72.

- Tooker, J. F., and S. D. Frank. 2012. Genotypically diverse cultivar mixtures for insect pest management and increased crop yields. *Journal of Applied Ecology* 49:974–985.
- Tripathi, A., D. K. Tripathi, D. K. Chauhan, N. Kumar, and G. S. Singh. 2016. Paradigms of climate change impacts on some major food sources of the world: A review on current knowledge and future prospects. *Agriculture, Ecosystems & Environment* 216:356–373.
- Tummers, B. 2006. DataThief III. <http://datathief.org/>.
- Vandermeer, J. H. 1989. *The Ecology of Intercropping*. Cambridge University Press.
- Vera, C. L., S. L. Fox, R. M. DePauw, M. A. H. Smith, I. L. Wise, F. R. Clarke, J. D. Procnier, and O. M. Lukow. 2013. Relative performance of resistant wheat varietal blends and susceptible wheat cultivars exposed to wheat midge, *Sitodiplosis mosellana* (Géhin). *Canadian Journal of Plant Science* 93:59–66.
- Vitosh, M. L. 1998. *Wheat Fertility and Fertilization*. Extension Bulletin E-2526, Michigan State University.
- Weigelt, A., and P. Jolliffe. 2003. Indices of plant competition. *Journal of Ecology* 91:707–720.
- Weiss, M. J., N. R. Riveland, L. L. Reitz, and T. C. Olson. 1990. Influence of resistant and susceptible cultivar blends of hard red spring wheat on wheat stem sawfly (Hymenoptera: Cephidae) damage and wheat quality parameters. *Journal of economic entomology* 83:255–259.
- Wolfe, M. S. 1985. The Current Status and Prospects of Multiline Cultivars and Variety Mixtures for Disease Resistance. *Annual Review of Phytopathology* 23:251–273.
- Wood, S. A., D. S. Karp, F. DeClerck, C. Kremen, S. Naeem, and C. A. Palm. 2015. Functional traits in agriculture: agrobiodiversity and ecosystem services. *Trends in Ecology & Evolution* 30:531–539.

- Worster, C. A., and C. C. Mundt. 2007. The effect of diversity and spatial arrangement on biomass of agricultural cultivars and native plant species. *Basic and Applied Ecology* 8:521–532.
- Yachi, S., and M. Loreau. 1999. Biodiversity and ecosystem productivity in a fluctuating environment: The insurance hypothesis. *Proceedings of the National Academy of Sciences* 96:1463–1468.
- Yahiaoui, S., A. Cuesta-Marcos, M. P. Gracia, B. Medina, J. M. Lasa, A. M. Casas, F. J. Ciudad, J. L. Montoya, M. Moralejo, J. L. Molina-Cano, and E. Igartua. 2014. Spanish barley landraces outperform modern cultivars at low-productivity sites. *Plant Breeding* 133:218–226.
- Zhu, Y., H. Chen, J. Fan, Y. Wang, Y. Li, J. Chen, J. Fan, S. Yang, L. Hu, H. Leung, T. W. Mew, P. S. Teng, Z. Wang, and C. C. Mundt. 2000. Genetic diversity and disease control in rice. *Nature* 406:718–722.

Appendix S1

Meta-analysis search terms

Search in *Web of Science* with the following settings for citation indexes and timespan (results updated January 2016):

Indexes=SCI-EXPANDED, SSCI, A&HCI Timespan=All years

Search terms and syntax:

TOPIC: ("grain crops" or maize or corn or wheat or oat or barley or rice or sorghum or millet or "dry beans" or soybean or cowpea or legume or "vegetable crops" or tomato or potato or cucurbit* or squash)*

Refined by: TOPIC: (variet or cultivar* or landrace) AND TOPIC: (mixture or intercrop* or blend or multi-)*

Then, refined by *Web of Science* categories:

Excluding: (FOOD SCIENCE TECHNOLOGY OR BIOPHYSICS OR CHEMISTRY PHYSICAL OR CHEMISTRY APPLIED OR BIOTECHNOLOGY APPLIED MICROBIOLOGY OR STATISTICS PROBABILITY OR IMAGING SCIENCE PHOTOGRAPHIC TECHNOLOGY OR PUBLIC ENVIRONMENTAL OCCUPATIONAL HEALTH OR CHEMISTRY MEDICINAL OR IMMUNOLOGY OR CHEMISTRY ANALYTICAL OR INSTRUMENTS INSTRUMENTATION OR BIOCHEMICAL RESEARCH METHODS OR AUTOMATION CONTROL SYSTEMS OR BEHAVIORAL SCIENCES OR COMPUTER SCIENCE ARTIFICIAL INTELLIGENCE OR FORESTRY OR ENGINEERING CHEMICAL OR FISHERIES OR VETERINARY SCIENCES OR PHYSICS APPLIED OR THERMODYNAMICS OR PHARMACOLOGY PHARMACY OR ENDOCRINOLOGY METABOLISM OR GASTROENTEROLOGY HEPATOLOGY OR POLYMER SCIENCE OR CHEMISTRY ORGANIC OR METEOROLOGY ATMOSPHERIC

SCIENCES OR MECHANICS OR SPECTROSCOPY OR NEUROSCIENCES)

*Refined by: (AGRONOMY OR PLANT SCIENCES OR AGRICULTURE MULTIDISCIPLINARY
OR HORTICULTURE OR ECOLOGY OR SOIL SCIENCE OR ENTOMOLOGY)*

Finally, refined selection by:

LANGUAGE: (English) and NOT Topic= (rice)

References included in the meta-analysis

- Akanda, S. I., and C. C. Mundt. 1997. Effect of two-component cultivar mixtures and yellow rust on yield and yield components of wheat. *Plant Pathology* 46:566–580.
- Aslam, M., and G. Fischbeck. 1993. Development of stripe and leaf rusts in wheat cultivar mixtures. *Journal of Agronomy and Crop Science* 171:49–54.
- Bacon, R., F. Collins, and D. Marx. 1987. Comparison of wheat blends with pure-line cultivars. *Cereal Research Communications* 15:183–186.
- Baker, R. J. 1977. Yield of pure and mixed stands of two spring wheat cultivars sown at five rates of seeding. *Canadian Journal of Plant Science* 57:1005–1007.
- Bebawi, F. F., and A. H. Abdelaziz. 1983. Effects of Cultivar Mixtures, Fertilizer, and Plant Density on Grain Sorghum (*Sorghum bicolor*)/*Striga hermonthica* Relations. *Weed Science* 31:552–556.
- Brophy, L. S., and C. C. Mundt. 1991. Influence of plant spatial patterns on disease dynamics, plant competition and grain yield in genetically diverse wheat populations. *Agriculture, Ecosystems & Environment* 35:1–12.
- Caviness, C. E. 1966. How selected soybean blends perform in Arkansas. *Arkansas Farm Research* 15:2.
- Caviness, C. E., and C. Smeller. 1992. Performance of blends of determinate and indeterminate

- soybeans. *Arkansas Farm Research* 41:3.
- Chapman, S. R., R. W. Allard, and J. Adams. 1969. Effect of planting rate and genotypic frequency on yield and seed size in mixtures of two wheat varieties. *Crop Science* 9:575–576.
- Clay, R. E., and R. W. Allard. 1969. A comparison of the performance of homogeneous and heterogeneous barley populations. *Crop science* 9:407–412.
- Cowger, C., and R. Weisz. 2008. Winter wheat blends (mixtures) produce a yield advantage in North Carolina. *Agronomy Journal* 100:169–177.
- Cox, C. M., K. A. Garrett, R. L. Bowden, A. K. Fritz, S. P. Dendy, and W. F. Heer. 2004. Cultivar mixtures for the simultaneous management of multiple diseases: tan spot and leaf rust of wheat. *Phytopathology* 94:961–969.
- Dai, J., J. J. Wiersma, and D. L. Holen. 2012. Performance of Hard Red Spring Wheat Cultivar Mixtures. *Agronomy Journal* 104:17.
- Darras, S., R. H. McKenzie, M. A. Olson, and C. J. Willenborg. 2014. Influence of genotypic mixtures on field pea yield and competitive ability. *Canadian Journal of Plant Science* 95:315–324.
- Davidson, J. L., D. B. Jones, and K. R. Christian. 1990. Winter feed production and grain yield in mixtures of spring and winter wheats. *Crop and Pasture Science* 41:1–18.
- Dubin, H. J., and M. S. Wolfe. 1994. Comparative behavior of three wheat cultivars and their mixture in India, Nepal and Pakistan. *Field Crops Research* 39:71–83.
- Erskine, W. 1977. Adaptation and competition in mixtures of cowpea (*Vigna unguiculata* (L.) Walp.). *Euphytica* 26:193–202.
- Fang, Y., B. Xu, L. Liu, Y. Gu, Q. Liu, N. C. Turner, and F. M. Li. 2014. Does a mixture of old

- and modern winter wheat cultivars increase yield and water use efficiency in water-limited environments? *Field Crops Research* 156:12–21.
- Fehr, W. R. 1980. Intergenotypic Competition Between Determinate and Indeterminate Soybean Cultivars in Blends and Alternate Rows. *Crop Science* 20:251–254.
- Fehr, W. R., and S. R. de Cianzio. 1980. Relationship of component frequency to compensatory response in soybean blends. *Crop Science* 20:392–393.
- Fehr, W. R., and S. R. Rodriguez. 1974. Effect of row spacing and genotypic frequency on the yield of soybean blends. *Crop science* 14:521–525.
- Finckh, M., and C. Mundt. 1992. Stripe rust, yield, and plant competition in wheat cultivar mixtures. *Phytopathology* 82:905–913.
- Finckh, M. R., E. S. Gacek, H. J. Czembor, and M. S. Wolfe. 1999. Host frequency and density effects on powdery mildew and yield in mixtures of barley cultivars. *Plant Pathology* 48:807–816.
- Finckh, M. R., and C. C. Mundt. 1992. Plant Competition and Disease in Genetically Diverse Wheat Populations. *Oecologia* 91:82–92.
- Frankel, O. H. 1939. Analytical yield investigations on New Zealand wheat: IV. Blending varieties of wheat. *The Journal of Agricultural Science* 29:249–261.
- Frey, K. J., and U. Maldonado. 1967. Relative productivity of homogeneous and heterogeneous oat cultivars in optimum and suboptimum environments. *Crop Science* 7:532–535.
- Gallandt, E. R., S. M. Dofing, P. E. Reisenauer, and E. Donaldson. 2001. Diallel analysis of cultivar mixtures in winter wheat. *Crop Science* 41:792–796.
- Gizlice, Z., T. E. Carter, and J. W. Burton. 1989. The impact of maturity and genotype on blend performance in group V and group VII soybean cultivars. *Agronomy Journal* 81:559–

562.

- Grafius, J. E. 1966. Rate of Change of Lodging Resistance, Yield, and Test Weight in Varietal Mixtures of Oats, *Avena sativa* L. *Crop Science* 6:369–370.
- Haghshenas, A., Y. Emam, H. Ghadiri, S. A. Kazemeini, and A. A. Kamgar-Haghighi. 2013. Effect of mixed cropping of an early- and a middle-ripening wheat cultivar on mitigation of competition during post-anthesis moisture stress. *Journal of Agricultural Science and Technology* 15:491–503.
- Harveson, R. M., and C. M. Rush. 2002. The influence of irrigation frequency and cultivar blends on the severity of multiple root diseases in sugar beets. *Plant Disease* 86:901–908.
- Helland, S. J., and J. B. Holland. 2001. Blend response and stability and cultivar blending ability in oat. *Crop Science* 41:1689–1696.
- Helm, J. H., P. E. Juskiw, and D. F. Salmon. 2000a. Competitive Ability in Mixtures of Small Grain Cereals. *Crop Science* 40:159.
- Helm, J. H., P. E. Juskiw, and D. F. Salmon. 2000b. Forage Yield and Quality for Monocrops and Mixtures of Small Grain Cereals. *Crop Science* 40:138.
- Hoekstra, G. J., L. W. Kannenberg, and B. R. Christie. 1985. Grain yield comparison of pure stands and equal proportion mixtures for seven hybrids of maize. *Canadian journal of plant science* 65:471–479.
- Jackson, L. F., and R. W. Wennig. 1997. Use of wheat cultivar blends to improve grain yield and quality and reduce disease and lodging. *Field Crops Research* 52:261–269.
- Jacques, S., R. K. Bacon, and L. D. Parsch. 1997. Comparison of single cropping, relay cropping and double cropping of soyabeans with wheat using cultivar blends. *Experimental Agriculture* 33:477–486.

- Jedel, P. E., J. H. Helm, and P. A. Burnett. 1998. Yield, quality and stress tolerance of barley mixtures in central Alberta. *Canadian journal of plant science* 78:429–436.
- Jokinen, K. 1991a. Assessment of competition and yield advantage in addition series of barley variety mixtures. *Journal of Agricultural Science in Finland* 63:307–320.
- Jokinen, K. 1991b. Yield and competition in barley variety mixtures. *Journal of Agricultural Science in Finland* 63:287–305.
- Juskiw, P. E., J. H. Helm, and P. A. Burnett. 2001. Three-component barley mixtures: Ratio effects in replacement series. *Canadian Journal of Plant Science* 81:651–656.
- Kannenberg, L. W., and R. B. Hunter. 1972. Yielding ability and competitive influence in hybrid mixtures of maize. *Crop science* 12:274–277.
- Karjalainen, R., and P. Peltonensainio. 1993. Effect of oat cultivar mixtures on disease progress and yield reduction caused by barley yellow dwarf virus. *Zeitschrift Fur Pflanzenkrankheiten und Pflanzenschutz-Journal of Plant Diseases and Protection* 100:58–68.
- Kathju, S., B. K. Garg, S. P. Vyas, and A. N. Lahiri. 2003. Sustainable production of moth bean through genotype management under arid environments. *Journal of Arid Environments* 53:137–143.
- Kaut, A. H. E. E., H. E. Mason, A. Navabi, J. T. O'Donovan, and D. Spaner. 2008. Performance and stability of performance of spring wheat variety mixtures in organic and conventional management systems in western Canada. *The Journal of Agricultural Science* 147:141.
- Kiær, L. P., I. M. Skovgaard, and H. Østergård. 2012. Effects of inter-varietal diversity, biotic stresses and environmental productivity on grain yield of spring barley variety mixtures. *Euphytica* 185:123–138.

- Li, N., S. Jia, X. Wang, X. Duan, Y. Zhou, Z. Wang, and G. Lu. 2012. The effect of wheat mixtures on the powdery mildew disease and some yield components. *Journal of Integrative Agriculture* 11:611–620.
- Mahmood, T., D. Marshall, and M. McDaniel. 1991. Effect of winter-wheat cultivar mixtures on leaf rust severity and grain yield. *Phytopathology* 81:470–474.
- Mengistu, N., P. S. Baenziger, L. A. Nelson, K. M. Eskridge, R. N. Klein, D. D. Baltensperger, and R. W. Elmore. 2010. Grain Yield Performance and Stability of Cultivar Blends vs. Component Cultivars of Hard Winter Wheat in Nebraska. *Crop Science* 50:617.
- Mercer-quarshie, H. 1979. Yield of local sorghum (*Sorghum vulgare*) cultivars and their mixtures in northern Ghana. *Tropical Agriculture* 56:125–133.
- Mille, B., and B. Jouan. 1997. Influence of varietal associations on the development of leaf and glume blotch and brown leaf rust in winter bread wheat. *Agronomie* 17:247–251.
- Mumaw, C. R., and C. R. Weber. 1957. Competition and natural selection in soybean varietal composites. *Agronomy Journal* 49:154–160.
- Mundt, C. C. 2002. Performance of wheat cultivars and cultivar mixtures in the presence of *Cephalosporium* stripe. *Crop protection* 21:93–99.
- Mundt, C. C., L. S. Brophy, and S. C. Kolar. 1996. Effect of genotype unit number and spatial arrangement on severity of yellow rust in wheat cultivar mixtures. *Plant Pathology* 45:215–222.
- Mundt, C. C., L. S. Brophy, and M. S. Schmitt. 1995a. Choosing crop cultivars and cultivar mixtures under low versus high disease pressure: A case study with wheat. *Crop Protection* 14:509–515.
- Mundt, C. C., L. S. Brophy, and M. S. Schmitt. 1995b. Disease severity and yield of pure-line

- wheat cultivars and mixtures in the presence of eyespot, yellow rust, and their combination. *Plant Pathology* 44:173–182.
- Mundt, C. C., P. M. Hayes, and C. C. Schön. 1994. Influence of barley variety mixtures on severity of scald and net blotch and on yield*. *Plant Pathology* 43:356–361.
- Newton, A. C., and D. C. Guy. 2009. The effects of uneven, patchy cultivar mixtures on disease control and yield in winter barley. *Field Crops Research* 110:225–228.
- Palvakul, M., V. C. Finkner, and D. L. Davis. 1973. Blendability of phenotypically similar and dissimilar winter barley cultivars. *Agronomy Journal* 65:74–77.
- Pandey, H., B. Malik, and M. Rao. 1978. Wheat varietal mixtures. *Indian Journal of Genetics and Plant Breeding* 38:277–280.
- Panse, A., J. H. C. Davis, and G. Fischbeck. 1989. Compensation-Induced Yield Gains in Mixtures of Common Bean (*Phaseolus vulgaris* L.). *Journal of Agronomy and Crop Science* 162:347–353.
- Patterson, F. L., J. F. Schafer, R. M. Caldwell, and L. E. Compton. 1963. Comparative standing ability and yield of variety blends of oats. *Crop Science* 3:558–560.
- Paynter, B. H., and A. L. Hills. 2008. Mixing feed barley cultivars did not decrease leaf disease or increase grain yield. *Australasian Plant Pathology* 37:626.
- Peltonen-Sainio, P., and R. Karjalainen. 1991. Agronomic evaluation of growing oat cultivar mixtures under various stress conditions in Finland. *Acta Agriculturae Scandinavica* 41:47–53.
- Pfahler, P. L. 1965. Environmental variability and genetic diversity within populations of oats (cultivated species of *Avena*) and rye (*Secale cereale* L.). *Crop Science* 5:271–275.
- Pfahler, P. L., and H. F. Linskens. 1979. Yield stability and population diversity in oats (*Avena*

- sp.). *Theoretical and Applied Genetics* 54:1–5.
- Pradhanang, P. M., and B. R. Sthapit. 1995. Effect of cultivar mixtures on yellow rust incidence and grain yield of barley in the hills of Nepal. *Crop Protection* 14:331–334.
- Pridham, J. C., M. H. Entz, R. C. Martin, and P. J. Hucl. 2007. Weed, disease and grain yield effects of cultivar mixtures in organically managed spring wheat. *Canadian Journal of Plant Science* 87:855–859.
- Probst, A. H. 1957. Performance of variety blends in soybeans. *Agronomy Journal* 49:148–150.
- Qualset, C. O., and R. M. Granger. 1970. Frequency dependent stability of performance in oats. *Crop Science* 10:386–389.
- Ram, B., A. Redhu, and S. Singh. 1989. Development of rusts and powdery mildew in mixtures of wheat-varieties. *Cereal Research Communications* 17:195–201.
- Rao, B., and R. Prasad. 1982. Studies on productivity of seed blends of 2 spring wheat cultivars under rainfed conditions. *Zeitschrift Fur Acker und Pflanzenbau-Journal of Agronomy and Crop Science* 151:17–23.
- Rao, B. R., and R. Prasad. 1984. Intergenotypic competition in mixed stands of spring wheat genotypes. *Euphytica* 33:241–247.
- Rasmusson, D. C. 1968. Yield and stability of yield of barley populations. *Crop Science* 8:600–602.
- Ross, W. M. 1965. Yield of grain sorghum (*Sorghum vulgare* Pers.) hybrids alone and in blends. *Crop Science* 5:593–594.
- Sage, G. C. M. 1971. Inter-varietal competition and its possible consequences for the production of F1 hybrid wheat. *The Journal of Agricultural Science* 77:491–498.
- Sammons, D. J., and P. S. Baenziger. 1985. Performance of four winter wheat cultivars in

- blended populations. *Field Crops Research* 10:135–142.
- Sarandon, S. J., and R. Sarandon. 1995. Mixture of Cultivars: Pilot Field Trial of an Ecological Alternative to Improve Production or Quality of Wheat (*Triticum aestivum*). *Journal of Applied Ecology* 32:288–294.
- Schweitzer, L. E., W. E. Nyquist, J. B. Santini, and T. M. Kimes. 1986. Soybean cultivar mixtures in a narrow-row, noncultivable production system. *Crop Science* 26:1043–1046.
- Sharma, R. C., and H. J. Dubin. 1996. Effect of wheat cultivar mixtures on spot blotch (*Bipolaris sorokiniana*) and grain yield. *Field Crops Research* 48:95–101.
- Stolen, O., J. Hermansen, and J. Lohde. 1980. Varietal mixtures of barley and their ability to reduce powdery mildew and yellow rust diseases. *Arsskrift Den Kongelige Veterinaer- Og Landbohøjskole*:109–116.
- Stützel, H., and W. Aufhammer. 1989. Effects of winter barley cultivar mixtures on lodging. *The Journal of Agricultural Science* 112:47–55.
- Swanston, J. S., A. C. Newton, J. M. Brosnan, A. Fotheringham, and E. Glasgow. 2005. Determining the spirit yield of wheat varieties and variety mixtures. *Journal of Cereal Science* 42:127–134.
- Tilahun, A. 1995. Yield gain and risk minimization in maize (*Zea mays*) through cultivar mixtures in semi-arid zones of the Rift Valley in Ethiopia. *Experimental Agriculture* 31:161–168.
- Vlachostergios, D., A. Lithourgidis, A. Korkovelos, D. Baxevanos, T. Lazaridou, A. Khah, and A. Mavromatis. 2011a. Mixing ability of conventionally bred common vetch (*Vicia sativa* L.) cultivars for grain yield under low-input cultivation. *Australian Journal of Crop*

- Science 5:1588.
- Vlachostergios, D. N., A. S. Lithourgidis, C. A. Dordas, and D. Baxevasos. 2011b. Advantages of Mixing Common Vetch Cultivars Developed from Conventional Breeding Programs when Grown under Low-Input Farming System. *Crop Science* 51:1274.
- Walsh, E. J., and M. G. Noonan. 1998. Agronomic and quality performance of variety mixtures in spring wheat (*Triticum aestivum* L.) under Irish conditions. *Cereal Research Communications* 26:427–432.
- Weiss, M. J., N. R. Riveland, L. L. Reitz, and T. C. Olson. 1990. Influence of resistant and susceptible cultivar blends of hard red spring wheat on wheat stem sawfly (Hymenoptera: Cephidae) damage and wheat quality parameters. *Journal of economic entomology* 83:255–259.
- White, E. M. 1982. The effects of mixing barley cultivars on incidence of powdery mildew (*Erysighe graminis*) and on yield in Northern Ireland. *Annals of Applied Biology* 101:539–545.
- Zeller, S. L., O. Kalinina, D. F. Flynn, and B. Schmid. 2012. Mixtures of genetically modified wheat lines outperform monocultures. *Ecological Applications* 22:1817–1826.
- Zhou, K. Q., G. D. Wang, Y. H. Li, X. B. Liu, S. J. Herbert, and M. Hashemi. 2014. Assessing variety mixture of continuous spring wheat (*Triticum aestivum* L.) on grain yield and flour quality in Northeast China. *International Journal of Plant Production* 8:91–105.

CHAPTER 2

PROMOTING ENHANCED ECOSYSTEM SERVICES FROM COVER CROPS USING INTRA- AND INTERSPECIFIC DIVERSITY

Emily R. Reiss and Laurie E. Drinkwater

Abstract

Farmers are increasingly looking to cover crops to provide ecosystem services to support yields, while reducing off-farm inputs. Research evidence in both unmanaged and agricultural ecosystems suggests that increased plant diversity can improve ecosystem service outcomes. The use of diverse cover crops as a practice to increase ecosystem service delivery has potential, though knowledge gaps remain regarding the types and levels of diversity that are most effective, and how environment may impact the desired outcomes. Using six species of legumes and grasses (hairy vetch, field pea, crimson clover, cereal rye, common wheat, ryegrass) along with multiple cultivars of each, we conducted a field experiment over two years during the winter fallow season. We tested the effect of a gradient of diversity in intraspecific mixtures, interspecific mixtures, and functional group (grass and legume) mixtures. We measured aboveground biomass production, weed suppression, soil-derived and fixed nitrogen in the shoots, and the C/N ratio of the aboveground biomass to evaluate corresponding ecosystem services. Species mixtures, especially those composed of both grasses and legumes tended to have the most significant ecosystem service benefits suggesting that functional diversity is more significant than species richness. Where soil nutrients were less available, the diversity effect tended to be stronger, though this was significant in only two instances. These results lend some support to the stress-gradient hypothesis in that in more stressful conditions the diversity benefit was greater. Diverse mixtures may not consistently produce substantial service improvements,

but there is little risk, and they may be most useful in marginal or depleted field conditions.

Introduction

Ecological intensification posits that increasing biodiversity can help to enhance ecosystem services reducing the need for anthropogenic inputs (Bommarco et al., 2013; Petersen and Snapp, 2015). As farmers seek to reduce their use of inputs without incurring yield reductions, interest is growing in managing biodiversity to foster ecosystem services. However, the low levels of diversity and ecosystem services in most agroecosystems present challenges for ecological intensification. Farmers are increasingly turning to cover crops as a means of diversifying their cropping systems, both spatially and temporally, which can help address the challenge of maintaining yields with reduced inputs (Dunn et al., 2016). Cover crops are not harvested, and are planted between periods of cash crop production when a field would typically be fallow (Carlson and Stockwell, 2013; Schipanski et al., 2014; Snapp et al., 2005). Cover crops can provide a range of ecosystem services, but farmers tend to prioritize certain ecosystem services over others, with a heavy emphasis on soil health benefits and nutrient management, including symbiotic nitrogen fixation (SNF) by legumes, in addition to pest control (Wayman et al., 2016). Overall productivity of the cover crop is highly valued, especially as many of these ecosystem services are related to biomass production of the cover crop (Finney et al., 2016).

Productivity has also been used as one of the main measures of ecosystem function in unmanaged ecosystems. In these non-agricultural settings evidence suggests that diverse plant communities may improve the productivity and ecosystem service outcomes of these ecosystems (Hooper et al., 2005). These experiments tend to use species richness as the indicator of diversity and have used aboveground primary productivity as a proxy for ecosystem functioning (Cadotte et al., 2008, 2009; Cardinale et al., 2011; Grace et al., 2016). However, this relationship between

plant species diversity and productivity is not always evident, and it is increasingly apparent that the diversity of functional traits within the community is often a better predictor of productivity than species richness alone (Cadotte et al., 2011). These traits, which are mainly related to resource utilization (light, nutrient and water capture), ultimately influence how the whole community functions (McGill et al., 2006). A more functionally diverse community has greater resource partitioning, potentially utilizing the available resources more efficiently and ultimately increasing overall productivity (Cadotte et al., 2011).

Identifying and then linking important functional traits to specific ecosystem functions is complex and challenging, which has led to the common use of simplified functional groups, especially in agroecosystems. The way a plant affects and responds to the full range of ecosystem processes is determined by a large number of functional traits, and only a subset of those have been identified as strong influences on specific ecosystem functions (Eviner and Chapin III, 2003). Additionally, there are likely other key traits yet to be isolated or measured, making it difficult to fully account for and characterize species in terms of measurable functional traits (Cadotte et al., 2009). Functional groupings using single traits known to be major drivers of ecosystem processes (i.e. nutrient acquisition method, phenology, plant architecture) is a way to simplify the relationship between species identity and ecosystem function (Eviner and Chapin III, 2003; Petchey and Gaston, 2002). This simplified approach of functional groups is commonly used in the management of agricultural systems, for example in rotations or companion planting, such as legumes or grasses, and winter or spring grains (Finney et al., 2016; Storkey et al., 2013). When constructing cover crop mixtures this simplification is quite useful as these groupings can help to easily characterize complementary species combinations. However, they fail to account for the full complexity of plant functional and genetic diversity including the

phylogenetic diversity of species within a functional group and the genetic diversity within species.

A common example of the use of functional groups in agriculture to build complementary systems with minimal tradeoffs is grass-legume mixes. These mixtures have a long tradition in agriculture, from forages to intercropped grain systems (Bateman and Keller, 1956; Boyd and Brennan, 2006; Cooper et al., 1960). Consequently they are a very appealing cover crop choice for farmers as they can provide a suite of ecosystem services taking advantage of the strengths of the two functional groups (Baba et al., 2011; Bedoussac et al., 2015; Nyfeler et al., 2009; Osman and Diek, 1982). Legumes fix nitrogen (N) contributing to long term N supply, as well as a lower carbon to nitrogen (C/N) ratio providing N in the short term for the subsequent crop (Brainard et al., 2011). Alternately, grasses tend to produce greater biomass contributing more to organic matter accumulation as well as suppressing weeds better than legumes (Kruidhof et al., 2008). Grasses also excel at scavenging soil N, reducing losses of N from the field (Aronsson et al., 2016). Since farmers are rarely interested in one ecosystem service alone, and tend to prioritize weed suppression and N fixation equally, by mixing these two contrasting functional groups together it is possible to achieve multiple ecosystem-service goals while minimizing tradeoffs (Schipanski et al., 2014; Wayman et al., 2016). In a grass-legume mix, grasses can suppress weeds while stimulating N fixation in the legumes, resulting in greater overall N content of residues, as well as equal or greater biomass production compared to grasses alone. However, if growing conditions favor grasses, they can outcompete the legumes, reducing their benefit. Consequently, managing the functional groups, and selecting appropriate species and cultivars for the given environmental conditions is a critical challenge.

Agricultural systems present the opportunity to draw on intraspecific diversity within

species in functional groups, as cultivars, as well as to enhance cover crop performance and the delivery of ecosystem services, while increasing agroecosystem diversity overall. In agricultural systems, cultivars have been developed that exhibit significant variation in functional traits (Elzebroek, 2008; Fu, 2015). Cultivar mixtures, which increase spatial intraspecific diversity, have been successfully used to increase yield and other services, such as disease management, in cash crops (Kaut et al., 2008; Kiær et al., 2009; Mundt, 2002; Reiss and Drinkwater, 2018).

Given the accumulating evidence from a variety of ecosystems, increasing intraspecific diversity in cover crop mixtures could also prove to be advantageous. Intraspecific diversity in cover crops could be increased alone or in conjunction with species and functional group mixtures. As individual cultivars can vary substantially in their performance, even under common conditions, cultivar selection, for mixture or sole cultivation, can have a major influence on the outcome (Andrew et al., 2015; Duvick, 2005). For example, certain cultivars may excel under drought conditions, or be more competitive against weeds, in addition to having differences in yield (Andrew et al., 2015). Increasing the attention paid to how cultivar identity and specific cover crop traits impact performance is a critical knowledge gap for optimizing cover crop use.

In addition to supporting ecological intensification and broadening ecosystem services provided by cover crop mixtures, diversity can also be a key tool for farmers looking to mitigate the effects of environmental variation. Environmental conditions can influence the diversity-ecosystem function relationship, with more stressful conditions accentuating the generally positive diversity effect, as described by the stress-gradient hypothesis (Dwivedi et al., 2016; He et al., 2013; Newton et al., 2010; Reiss and Drinkwater, 2018; Yu et al., 2015). In a comprehensive meta-analysis, we found that more stressful environments tend to favor positive

plant-plant interactions such as facilitation over competition (He et al. 2013). Facilitation occurs when neighbors alter biotic or abiotic conditions, which promotes success of individuals and may lead to overall greater success of the community (Bertness and Callaway, 1994; He et al., 2013). For cover crops, this interaction between environmental conditions and diversity may have useful management implications, particularly given the increased weather variability and greater frequency of extreme events due to climate change (Gaudin et al., 2015).

To build on all of this existing work, we conducted a field study to evaluate the performance of six species of cover crops as monocultures, and as cultivar and species mixtures. We included three species in each of two functional groups, grasses and legumes. We measured aboveground biomass production, weed suppression, soil-derived and fixed N in the shoots, and the C/N ratio of the aboveground biomass to evaluate the corresponding ecosystem services (Table 1). We were interested in comparing the ecosystem service benefits resulting from intraspecific, interspecific, and functional diversity as well as the variation between cultivars of cover crop species in terms of the outcomes of the associated services. We expected species mixtures, and specifically those including both functional groups, to provide a broader range of ecosystem service outcomes compared to less diverse plantings. We were also interested in testing if the addition of multiple cultivars would improve or stabilize the ecosystem service benefits of cover crops. Lastly, we hypothesized that more stressful conditions, such as those related to soil fertility, would strengthen the diversity effect on the delivery of these services.

Table 1 Measured variables to evaluate corresponding ecosystem service

Collected metric	Ecosystem service
Soil N accrual in cover crop biomass (N_S)	N retention
Shoot fixed N (N_F)	N supply, long-term
C/N ratio	N supply short-term for the following crop
Weed biomass	Weed suppression
Cover crop biomass	C accrual

Methods

Site and experimental design

We established a field experiment for two consecutive seasons at Cornell University Musgrave Research Farm in Aurora, NY (42.73' N, 76.66' W). The region, in USDA plant hardiness zone 6a, has a humid temperate climate with a mean annual precipitation of 88 cm, and a mean annual maximum and minimum air temperature of 14°C and 3°C, respectively. The site consists mostly of moderately well-drained Honeoye (fine-loamy, mixed, semiactive, mesic Glossic Hapludalfs) and Lima (fine-loamy, mixed, semiactive, mesic Oxyaquic Hapludalfs) silt loams. Experimental design and management was consistent in both experimental seasons (September-June). For clarity, results from the 2013-2014 season are reported as from Field 1, and results from the 2014-2015 season are reported as from Field 2. The fields used in the two years were less than 300 meters from one another, but Field 1 was primarily a Honeoye silt loam, whereas Field 2 was dominated by Lima silt loam. The two fields also differed in terms of soil N, which was higher in Field 2, and soil P, which was higher in Field 1 (Table 2). Cover crop treatments were applied on the experimental plots, and we used a randomized, complete block, replacement series, split-plot (for fertilizer treatment) design with four replicates in experimental plots of 2.4m x 2.4m. Experimental plots were established in fields that had been conventionally managed under a typical corn-soybean-wheat rotation. The Field 1 plots were established following soybeans in 2012 and winter wheat harvested in summer 2013. The Field 2 plots were established following corn in 2013 and a winter/spring fallow. Fertilizer was applied to wheat and corn according to local recommendations and soil tests. We applied appropriate bacterial inoculant to all legume species unless they were pretreated by the supplier (N-Dure brand, Verdesian, Cary, NC). We planted all treatments in

Table 2 Field and environmental conditions for the two experiment fields. For p value results, "ns", not significant; "nt", not tested.

Field	Inorg N (mgN/kg soil)	PMN (mgN/kg soil/week)	N (%)	C (%)	C:N	pH	P (ppm)	K (ppm)	CEC	Sand (%)	Clay (%)	Soil Textural Class	Precip. (cm)	GDD (0°C base temp.)
1	5.8 (1.1)	10.9 (3.6)	0.16 (0.01)	2.1 (0.1)	13.5 (0.80)	7.7 (0.1)	23.5 (3.5)	73.8 (10.3)	13.7 (1)	42.7	23.7	Loam	66.4	2068
2	11.0 (1.4)	8.8 (2)	0.18 (0.02)	2.0 (0.1)	11.0 (0.69)	7.7 (0.1)	16.5 (1)	75.8 (8.4)	14.8 (0.6)	44.6	22.7	Loam	60.8	1770
p value	0.0013	ns	0.0271	ns	<0.0017	ns	0.0086	ns	ns	nt	nt	nt	nt	nt

mid-September (September 14-15 in 2013, September 15 in 2014). After the field was disked, plots were broadcast seeded by hand, and then the whole field was rolled with a cultipacker. We applied 80 kg N/ha as pelletized ammonium nitrate fertilizer by hand to the soil surface in a randomized, split-plot design during the first week of May in both 2014 and 2015. The experimental plots received no fertilizers or pesticides other than this application, and were rainfed. Actual precipitation was 66.4 cm and 60.8 cm for the duration of the experiment (planting to sampling) in the two years respectively. Additionally, the growing degree days (GDD) were 2068 and 1770 for the two experiments using 0°C as the base temperature (Table 2).

We selected six annual cover crop species (3 legumes, 3 grasses) commonly used by farmers in the upstate New York region for this experiment. We intentionally selected winter annual species to minimize the potential confounding effect of differing life histories and growth rates. All of these annual species planted in late summer/early fall germinate and grow into late fall/early winter after which they are dormant for the duration of the winter. In spring they break dormancy, and produce the majority of their biomass during April-May. We selected hairy vetch (*Vicia villosa*), winter field pea (*Pisum sativa*) and crimson clover (*Trifolium incarnatum*) for the legume species. Cereal rye (*Secale cereale*), common wheat (*Triticum aestivum*), and ryegrass (*Lolium multiflorum*) constituted our three grass species.

We selected at least two cultivars from each species, and generally used at least one ‘variety not stated’ (VNS) line for each species and one named cultivar (Table 1S). For vetch and wheat, we expanded the number of cultivars to five and four respectively, and used these two as focal species for our mixture treatment structure. Where trait information was available about cultivars we attempted to select contrasting cultivars or lines, though this was difficult for many

species that have not undergone intensive breeding (i.e. vetch).

We constructed treatments to create a diversity gradient allowing for assessment of multiple levels of diversity. These treatments ranged from one species/one cultivar treatments to multiple species/multiple cultivars treatments. These treatments can be grouped into three categories with two subcategories each: A) Species monocultures A1) each cultivar of each species planted alone and A2) all cultivars of each species planted together (2-5 cultivars depending on species), B) Bicultures of wheat or rye with vetch or pea, B1) single cultivar of each species and B2) multiple cultivars of each species, C) Three or more species, C1) single cultivar of each species, C2) all cultivars of each species. For mixtures of legumes and grasses together, each functional group composed 50% of the mixture by number of seeds. Where there were multiple species in one functional group, each species was seeded equally. Actual species richness differed in some plots from the planted species richness, likely due to winter mortality, but as fall emergence and establishment was not assessed, it is not possible to confirm this. Actual species richness was used in all analyses. In addition to the cover crop treatment plots, three plots per block remained unplanted without any cover crops to allow for assessment of the background weed pressure across the field.

We used a replacement series design with a consistent seeding rate of 285 live seeds/m² (adjusted for germination rate) in all plots. This design is consistent with the biodiversity-ecosystem function work in unmanaged ecosystems, where plant density is held constant so that changes in plant community composition can be assessed independently of density (Tilman et al., 1996). This design is different compared with other cover crop diversity research in agroecosystems, where a proportional replacement design, or additive design is more commonly used (Bybee-Finley et al., 2016; Finney et al., 2016; Poffenbarger et al., 2015b). These alternate

designs often use agronomically appropriate seeding rates on a kg/ha basis. With six species of different seed sizes and growth habits, it was not realistic to achieve relevant seeding rates for all species while maintaining the same density. Instead, we pegged our target seeding density to the recommended seeding rate in kg/ha for wheat, one of our focal species (Table 2S) (Clark, 2007; Thomas Björkman, n.d.). Due to seed size, some species were over-seeded (e.g. the larger seeded legumes), whereas others were under-seeded compared to recommended rates. This is partly due to the fact that agronomic seeding rates for legumes tend to be lower, perhaps due to higher seed costs. Despite these differences in seeding rate on a kg/ha basis, we generally achieved our objective of consistent plant density in the plots. The overall harvest densities were on average at least a third of the seeded densities, suggesting that seeding density may have had minimal impact on final plant densities at harvest (Figure 1S).

Data collection and processing

Plants

We sampled the plants at the same physiological stage in both Field 1 and Field 2. When the vetch monoculture plots were at approximately 50% flowering, we harvested aboveground biomass for all treatments (June 16-19 2014, June 9-10 2015). At this developmental stage vetch is unlikely to regrow after field incorporation, yet viable seeds have not been set. This is when most farmers would terminate such a cover crop in order to ensure mortality of the cover crop, and prevent introduction of new seeds from the cover crop which might become problematic. The grasses were all post-anthesis and some were starting grain-fill. The pea and clover were also at peak or just post-peak flowering. In Field 1, winter mortality resulted in patchy stands in some plots. To compensate for this, we used two quadrats (0.125 m² each) in some plots, with the total area equal to the larger quadrat (0.25 m²) used in the remaining plots. Areas for sampling

with the smaller quadrats were selected to represent the planted composition of the plot and combined for analysis. In Field 2, all plots were sampled with the larger quadrat. Biomass was cut 9 cm above the soil surface, and we recorded plant count for all component cover crop species present and weeds. We separated this biomass by cover crop species, and weeds (all species together), which was then oven-dried for at least 48 hours at 60°C before weighing to the nearest 0.01g. Biomass from weed control plots was sampled in the same way as the cover crop plots. All cover crop samples were ground to at least 2 mm. Grass cover crop samples from legume mixture treatments were analyzed for total C and N on combustion using a LECO TruMac CN analyzer (Leco Corporation, St. Joseph, Missouri, USA). Additionally, all legume samples and grass monoculture samples were processed for mass spectrometer isotopic analysis. Specifically, dried plant biomass samples from the field (legumes and reference grasses) and the greenhouse *B* value experiment were first coarsely ground and then ground to 0.5mm with a propeller mill (Cyclotec™ Sample Mill, Foss, Hillerød, Denmark). Samples were analyzed for ¹⁵N natural abundance, total N content and total C content using a continuous flow Isotope Ratio Mass Spectrometer (Stable Isotope Facility, UC Davis).

Using the ¹⁵N natural abundance method, we estimated the N fixation by legumes in monoculture and mixtures in unfertilized plots (Shearer and Kohl, 1986). For this method, we used the monoculture plots of the non-legume, grass species in the experiment as the reference plants, for the following calculation:

$$\% \text{ N from fixation} = 100 \times \left(\frac{\delta^{15}\text{N grass} - \delta^{15}\text{N leg}}{\delta^{15}\text{N grass} - B} \right)$$

The *B* value, used to account for the isotopic fractionation of the nitrogen during fixation in the legumes, was determined through a greenhouse experiment for each cultivar of legume and each seed lot, if there were more than 1 (Table 3S). In the greenhouse we propagated plants

in N-free, autoclaved, calcined clay media. Seeds were sterilized with hypochlorite (Parr, 2010; Somasegaran and Hoben, 1994), and inoculated with the same inoculant used for the field experiment plus native rhizobia present in soil from Field 1 (Unkovich et al., 1994). We prepared the soil slurry by combining a representative sample of approximately 300ml of 2mm sieved soil in 6.5L of distilled and deionized water. Both inoculants were added at planting, and the soil slurry applied again one week after planting. Nitrogen addition from the slurry was negligible. Plants were grown to the same physiological point as they were harvested in the field, and then cut and dried as in the field experiment. Overall, the *B* values we obtained in this experiment were within generally accepted literature values (Parr, 2010; Unkovich et al., 2008).

In Field 1, we used the average of the three reference species in each replicate as the reference value to calculate the %N from fixation for each legume sample. In this field a few samples (less than 5 %) were greater than 100%, and we adjusted those samples to 100% for a realistic result (Unkovich et al., 2008). When we used this approach in Field 2, where there was considerable spatial variability in total N and ^{15}N signatures, we found that 31 of 157 samples had negative nitrogen fixation rates. We attempted to resolve these negative results in a number of ways including using actual field $\delta^{15}\text{N}$ values instead of the greenhouse *B* values, which had no effect on the total number of negative fixation results (Brainard et al., 2012). Ultimately, we used spatial statistics, Ordinary Kriging, to estimate the reference ^{15}N signature more accurately across the field. The remaining samples with negative fixation after using the kriged results, if available, were paired with a neighboring plot. All remaining samples (n=28) after this, that still had negative results, were adjusted to 0% N from fixation.

Soils

To characterize soil conditions, we collected composite soil samples (ten soil cores to 20

cm depth) from each of the four replicates prior to planting in the fall. A subsample of each composite sample was sieved to 2mm, and extracted for total inorganic N (Inorg N) with 2M KCl, and incubated anaerobically for seven days and extracted with 2.67M KCl for potentially mineralizable nitrogen (PMN). Total NH_4^+ and NO_3^- were analyzed via a colorimetric microplate technique (Ringuelet et al., 2011). Gravimetric water content of each sample was determined after being dried in the oven for 7 days at 65C. All samples were analyzed for total C and N on combustion (Leco Corporation, St. Joseph, Missouri, USA). All fall soil samples were analyzed for water pH, CEC, Mehlich buffer lime requirement, and for P, K, Mg, Ca, Zn, Cu, and S by the Mehlich 3 (ICP) test, and particle size (Agricultural Analytical Services Laboratory, Penn State University, University Park, PA) (Table 2).

Data Analysis

Ecosystem service calculations

Total grams of biomass N per m^2 (N_T) was calculated for all plants as %N x g of cover crop biomass/ m^2 . Nitrogen derived from the atmosphere through symbiotic nitrogen fixation (N_F) in legumes was partitioned using the %N from fixation result from the ^{15}N natural abundance method. The grams of fixed N in legumes was calculated as %SNF x N_T . The soil N accrual (N_S) for legumes was calculated as $N_T - N_F$. For grasses, all biomass N is derived from the soil so $N_S = N_T$. These calculations were only done for unfertilized plots as the natural abundance method of fixed N estimation is not applicable when fertilizer is applied.

C/N ratio was calculated as the sum of the cover crop biomass C content (gC/m^2) divided by the sum of biomass N content (gN_T/m^2) for all the cover crop species present in a plot. For consistency with N accrual and long-term supply, these calculations were also only done for unfertilized plots.

The land equivalent ratio (LER), often used for intercropping systems, quantifies the land area needed to achieve the same production in sole crops as in mixtures and is useful for comparing biomass produced in mixtures to productivity of cover crop monocultures. We calculated the total LER (LER_T) as the sum of the partial LERs for each species in the mixture. The partial LER is calculated as:

$$\text{Partial LER}_i = Y_{xi}/Y_{mi}$$

where Y_{xi} is the biomass of species i in mixture, and Y_{mi} is the biomass of species i in monoculture.

To approximate the spatial distribution of plant available soil N across Field 2, we used the %N in the biomass for the three grasses growing as monocultures. The %N was standardized within each species using a z-distribution. With these known standardized %N values, we used Ordinary Kriging to estimate the plant available N across the field for all plots.

Statistical analysis

Mixed models were used to test the effects of diversity on ecosystem service outcomes. In all models, all possible interactions of field (1 and 2) and fertilizer (+N/-N) and the main model factor were tested and included in the model if significant. If the interaction was not significant, data were pooled across fields and fertilizer treatments. Block was included as a random effect, and nested in field when both fields were included. To avoid pseudoreplication, treatments of the same diversity level (i.e. cultivar monoculture, cultivar mixture), were pooled and averaged within block before further analysis. In other words, though vetch has five cultivar monocultures and pea only two, there is no influence of this imbalance in the analysis. All cultivar monocultures for a species are averaged within block first, so that each species only has one data point for cultivar monocultures for each block. Tukey's honestly significant difference

(HSD) at $\alpha=0.05$ was used to test for differences between multiple levels (i.e. species richness), while two sample t-tests were only used to compare low and high diversity levels (i.e. cultivar monoculture vs. cultivar mixture). One sample t-tests were used to assess differences from reference values for total and partial LER. Data were transformed where evaluation of model residuals indicated non-normality, or heterogeneous variance was observed, as noted in figures. In particular, weed biomass data were transformed using natural log after a constant (1) was added in order to meet model assumptions. Unless otherwise noted, untransformed data is presented in figures. All analyses except Ordinary Kriging (Bivand et al., 2013; Gräler et al., 2016; Pebesma, 2004; Pebesma and Bivand, 2005; R Core Team, 2016) were conducted using JMP v.11 software (SAS Institute Inc. Cary, NC).

Results

The field plots we used for this two year experiment were located on the same soil type within the Cornell University Musgrave Research Farm, and we expected weather-related variability to be the primary source of environmental variation between the two years the experiment was conducted. However, while there were small differences in precipitation and temperature, weather conditions were relatively similar with only slightly less total precipitation over the cover crop growth period for Field 2 as well as about 300 fewer GDD (Table 2).

In contrast, there were significant differences in soil fertility and weed pressure between the fields used in the two years. Field 1 had greater extractable P, but lower total soil N and standing inorganic N pools in the spring suggesting that overall plant available N was lower in Field 1 compared to Field 2 (Table 2). Background field weed biomass and weed density were both lower in Field 2 (128 g/m² vs. 58 g/m², $p=0.0028$ compared to 371 plants/m² vs. 39 plants/m², respectively, $p<.0001$).

Variation among and within species

We found considerable variation across the six species in their performance for several of the ecosystem services we evaluated (Figure 1, Tables 4S and 6S). Rye and wheat generally produced greater biomass in conjunction with greater soil N uptake compared to ryegrass and the legume species (Figure 1, Table 6S). Crimson clover survival was substantially impaired by harsh winters in both years, reducing biomass production and, as a result, other related services were also reduced. Plant density at harvest for clover in monoculture was typically less than half as much as the average density of other species for both Field 1 and 2 (Figure 1S). The two functional groups differed in the amount of variability for biomass performance. Specifically, the coefficient of variation for mean cover crop biomass production across all species was 62% (pooling the two fields). Compared to legumes (CV=31%) the grass biomass variability was greater (CV=56%).

We observed much less variation within species than among them. We observed significant differences in the performance of a few individual cultivars for vetch and wheat in terms of biomass and the rate of N fixation (Figure 2S and 3S and Table 7S). Within-species weed suppression and soil N accrual were consistent across cultivars (with one exception for weed suppression in Field 2 where ‘Ernst Vetch’ was significantly different than the highly weed suppressive ‘Purple Bounty’). Within the six species across the two years, variation in biomass was generally small, between 8% and 27%. The two exceptions were pea and clover where the CV for biomass production was 75% and 40% respectively. This may have been due to differential responses by certain cultivars in the two fields. For example, ‘Austrian’ winter pea had twice the density in Field 2 compared to Field 1, which corresponded to a seven fold increase in biomass.

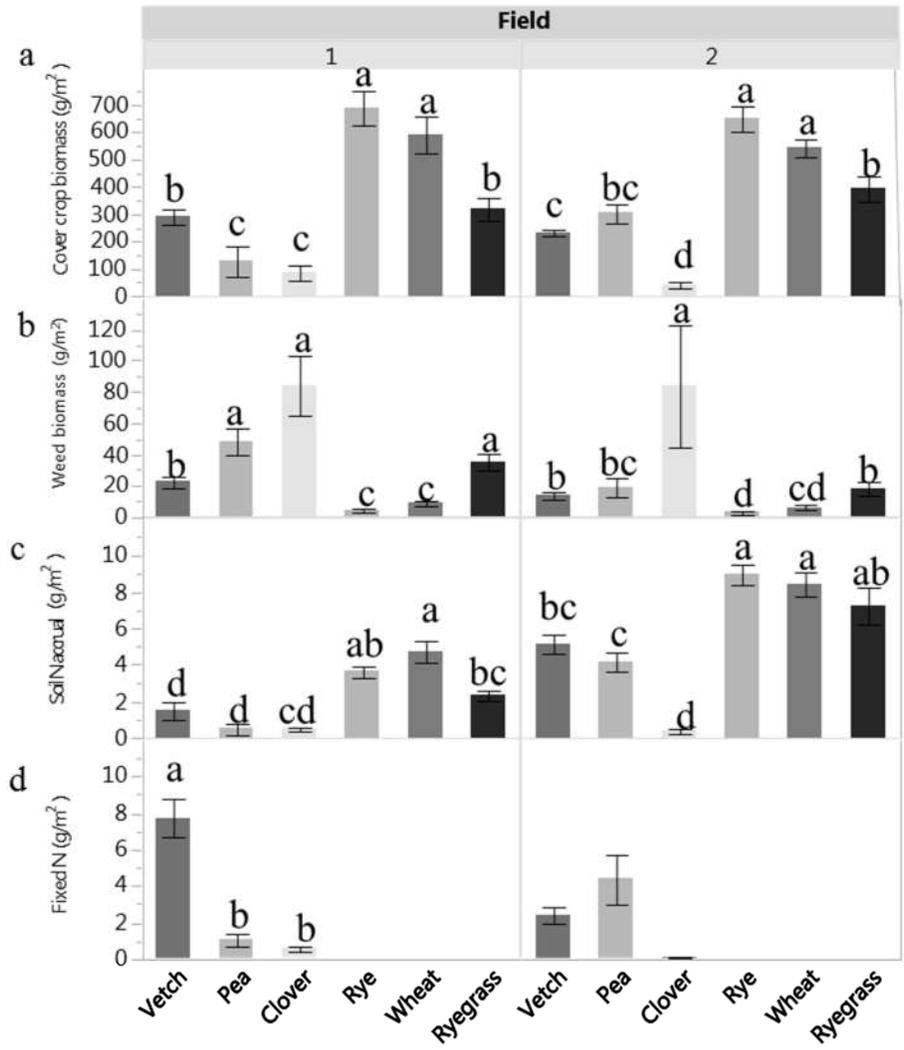


Figure 1 Average a) aboveground cover crop biomass b) weed biomass c) soil N accrual and d) fixed nitrogen of species monocultures for six species in Field 1 and Field 2. Different letters within an ecosystem service and field indicate significant differences between species ($p < 0.05$, Tukey HSD analysis on transformed data (ln for weed biomass; Square root for biomass, soil N accrual, and fixed N), but untransformed data presented). Both fertilized treatments are pooled for a and b with no interactions, while unfertilized data presented for c and d due to methodology. Error bars represent one standard error from the mean.

Intraspecific diversity

The effect of increased intraspecific diversity was inconsistent across the ecosystem services we measured. Overall, we observed a significant increase in weed suppression in cultivar mixtures compared to the average cultivar monoculture performance, but saw no response in any of the other ecosystem services evaluated (Figure 4S, Table 5S and 9S). In addition to the overall effect of intraspecific diversity, the increased weed suppression with cultivar mixtures was specifically significant for three of the five species assessed, vetch, pea, and ryegrass (Figure 4Sb, Table 9S).

To analyze the effects of intraspecific diversity more fully, we compared the performance of cultivar mixtures to the individual monocultures of vetch or wheat cultivars. Even though there were some significant differences across cultivars, our results reinforced our findings from the comparison using averages across monocultures where we found no instances where the cultivar mixture differed from all the cultivar monocultures (Figure 2S and Table 7S).

Increasing species richness and functional diversity

The ecosystem service response to increased species richness was certainly greater than for intraspecific diversity, though again it was strongest for certain instances of biomass production, weed suppression and the rate of N fixation. We observed a significant increase in biomass production with increasing species richness in grass/legume mixtures, and for grass species mixtures compared to their monocultures under unfertilized conditions (Figure 2 and Table 8S). While biomass was influenced by species richness, the weed biomass levels were low for all the grass-legume mixtures, with no significant differences between the different species richness levels (Figure 5Sa and Table 8S). However, increasing the species richness of sole-grass

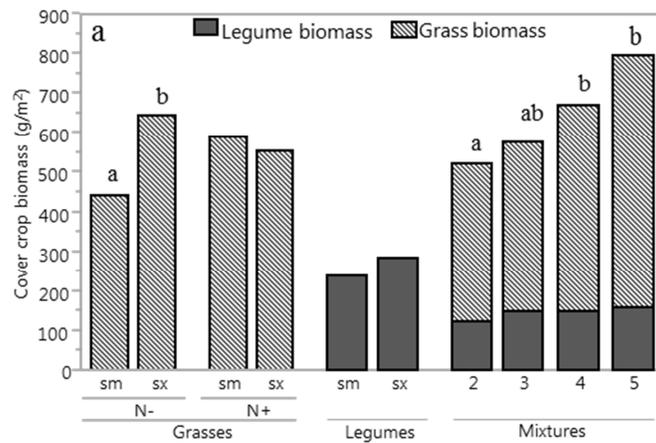


Figure 2 Average effect of interspecific diversity on cover crop biomass. Comparison of species monocultures (sm) and species mixtures (sx) for single functional group treatments (grasses and legumes) and of species richness levels (2-5) for grass/legume mixtures. Both fields and fertilizer treatments are pooled and presented except for grasses, which is presented for both fertilizer treatments (+/- N applied) due to an interaction. Different letters within functional group categories indicate significant differences between interspecific diversity levels ($p < 0.05$, Tukey HSD analysis on transformed data (square root), but untransformed data presented).

mixtures did result in a significant reduction in weed biomass. Increased species richness did not have a significant effect on total N or on soil N accrual (Figure 5Sb and Table 8S). The rate of SNF increased for both vetch and peas in grass/legume mixtures (5-11% in both fields and species, Table 5S). However, with a 50% reduction in the seeding rate compared to the monocultures, the decrease in biomass was a larger driver on total fixed N than the rate of fixation in these mixtures. Consequently, the total amount of N from fixation did not improve with species mixtures.

Relationship between cover crop biomass, ecosystem services and environmental conditions

Weed suppression measured as weed biomass in the cover crop plots had a weak negative correlation with cover crop biomass, but the strength of this correlation varied with functional group composition and field (Figure 3a). Reduction in weed biomass was most strongly correlated with biomass in grasses. Notably, biomass explained somewhat less of the variation in weed biomass for the grass-legume mixtures compared to grasses and legumes alone, more so in Field 2. This suggests that the diversity in the functional group mixtures may be contributing to the weed suppression beyond increased cover crop biomass.

Total N, soil N, and fixed N in plant biomass were all correlated with aboveground biomass, though the relationships among these N metrics varied also across functional groups and field. In contrast to grasses, biomass total N for legumes and grass-legume mixtures was not affected by the differences in soil N availability between Fields 1 and 2 (Figure 3b-e). However, the relative contributions of N from soil and N from fixation changed as legumes acquired more soil N and down-regulated N fixation under the increased soil N fertility conditions in Field 2 (Figure 3b-e). In general, total cover crop biomass was correlated with the amount of N fixed,

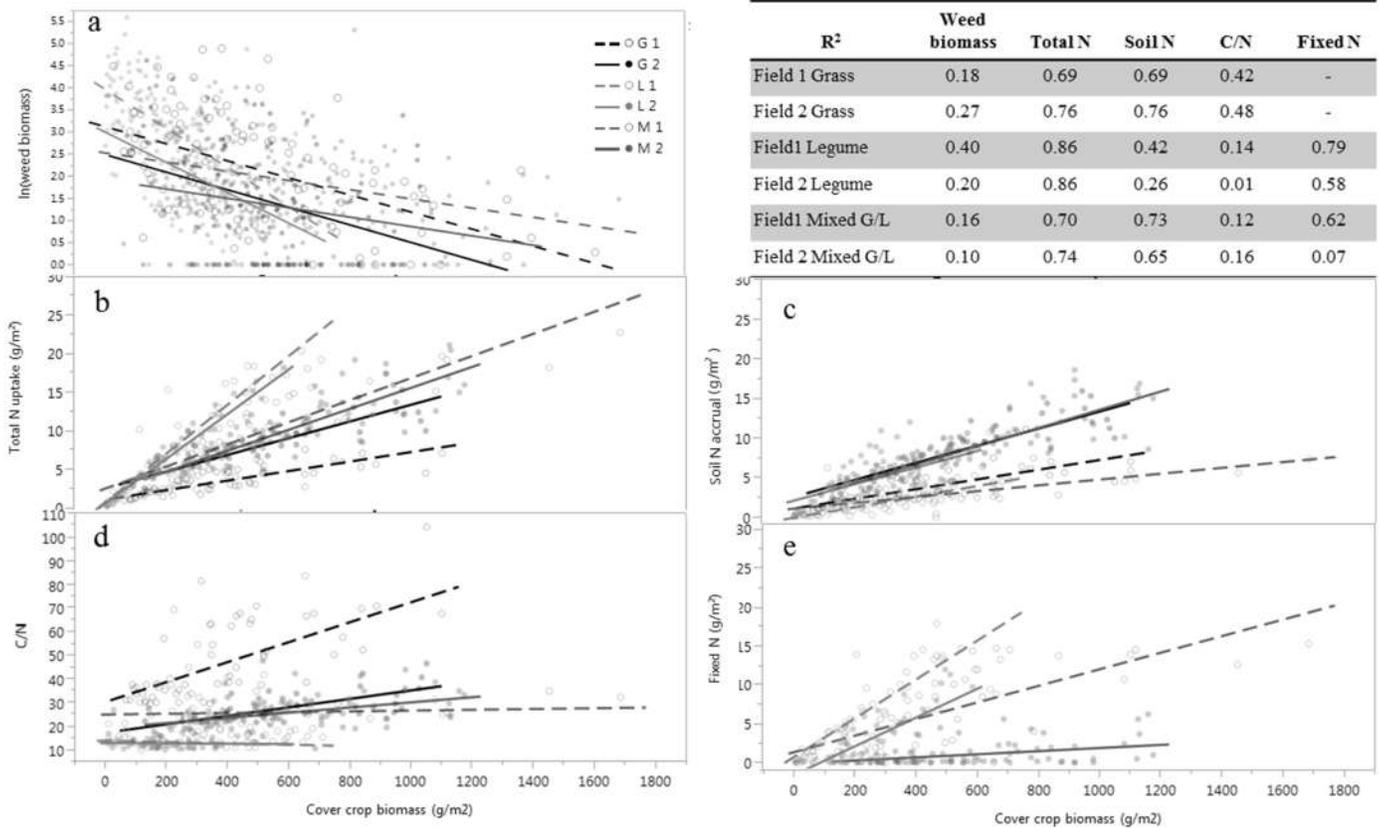


Figure3 a-e) Relationships between cover crop biomass by a) weed biomass (natural log transformed) b) total aboveground biomass nitrogen c) soil N accrual d) C/N ratio of biomass and e) fixed nitrogen in biomass for each field (1,2) and plot composition (G, grass only; L, legume only; M, grass/legume mix). Only data from unfertilized plots presented for b-e, and no interaction with fertilizer for weed biomass (a). R² values for regression models on transformed data. Data transformed where necessary (Square root for total N uptake and soil N accrual; natural log for weed biomass; Log10 for C/N). P<<0.05 for cover crop biomass for all fields, compositions and ecosystem services.

however, in Field 2 there was no relationship between biomass and fixed N in the mixtures. This was likely due to the lower amounts of fixed N in the mixtures where the highest values were ~ 5 g/m² compared to 10-15g/m² in the legumes in both fields and the mixtures in Field 1. We also observed very consistent performance by legumes and the grass-legume mixtures across the two soil N fertility environments for C/N, while grasses were distinctly different in the two fields (Figure 3d). Interestingly, biomass explained less of the variation in C/N for legumes and mixes compared to grasses, potentially suggesting an additional mechanism may be driving this other than biomass alone. As grasses only have one source of N, overall N content in biomass is strongly influenced by site differences. The ability to source N from soil and the atmosphere allows the legumes and mixtures to respond more dynamically to soil conditions.

We found that grass N acquisition varied with location in Field 2 indicating that soil N availability was spatially heterogeneous across the field. We used this variation to examine the response of pea and vetch N fixation in grass/legume mixtures to variations in soil N availability. As our soil tests were composite samples aggregated by block we could not use them to analyze spatial variability at the plot scale, so we used the shoot N concentration of nearby grasses as a proxy for available soil N. Both species down-regulated their N fixation in response to higher levels of available soil N (Figure 4). Due to the small number of points for each treatment, we were not able to detect cultivar or intra/interspecific diversity effects on this relationship. Removal of the data points with undetectable N fixation did not change the regression equation or significance of this relationship (data not shown).

We used the land equivalent ratio (LER) to evaluate the possible benefits, on a land area basis, of growing species together compared to the components in monoculture. Nearly all of the species mixtures had average total LERs for biomass of greater than 1 in both fields indicating a

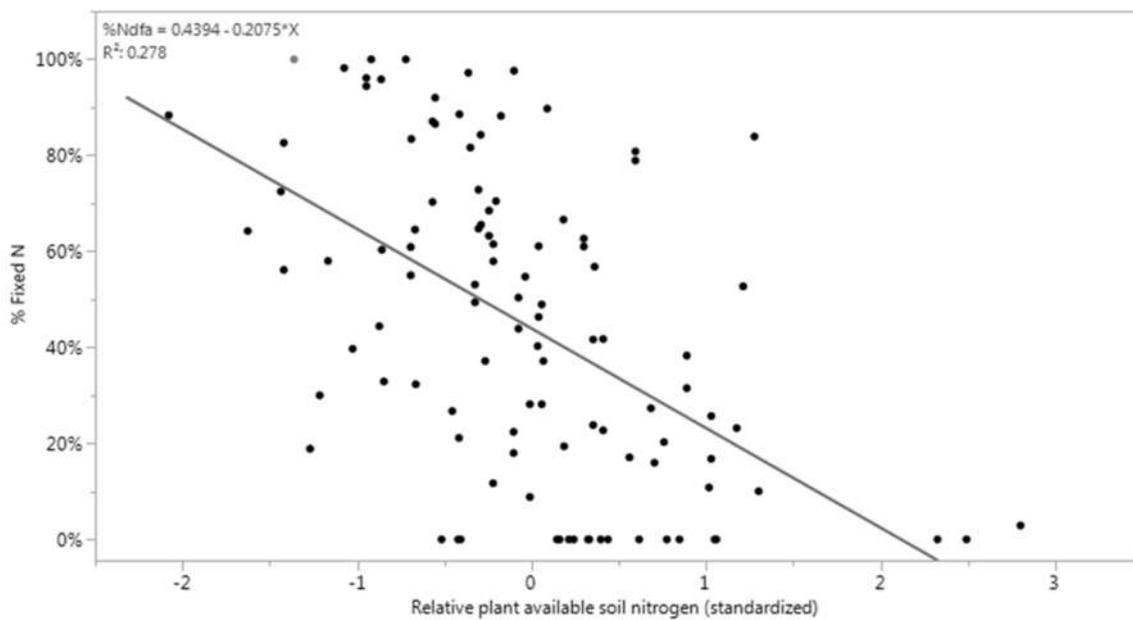


Figure 4 Effect of plant available soil nitrogen on the % of N fixed by pea and vetch in grass/legume mixtures in Field 2. The nitrogen content of grass monocultures was standardized and used as a proxy for plant available soil nitrogen content. $P < 0.001$ for all parameters.

benefit to increased species richness, except for species mixtures composed of only legumes or grasses and the 3Legumes1Grass mix in Field 1 (Figure 5 and Table 10S). The relative success of legumes and grasses in the mixtures differed in each field. Under the higher N fertility conditions in Field 2, LERs greater than 1 were largely due to the success of the grasses in that field, while in the less fertile Field 1 legumes contributed more biomass. In fact, in Field 2, the partial LERs for grasses were all significantly higher than the 0.5 reference, while legumes were all significantly lower.

In general, the imposed available soil N treatment (+/- N fertilizer) did not have a consistent effect on ecosystem service benefits provided by increased species richness. However, there were several instances where diversity effects were noticeable under low N fertility, but not under high N fertility conditions. First, we observed a significant increase in aboveground biomass with grass species mixtures only under unfertilized conditions (Figure 2). Second, in Field 1 we found a significant reduction in the proportion of mixtures exhibiting transgressive overyielding (where mixture yield is greater than the most productive monoculture component) in fertilized plots (32%) compared to 47% in unfertilized plots (Fisher's Exact Test $p=0.0440$, data not shown). However, there were no other differences in occurrence of transgressive overyielding in the other fields or fertilizer treatments, with approximately 50% of the plots showing transgressive overyielding overall. Though not significantly different, species richness accounted for more of the variation in biomass under unfertilized conditions, and this was more pronounced in the lower fertility Field 1 (Figure 6Sa). Finally, in both fields the fertilized plots had lower average LERs than the unfertilized plots, though this difference is also not significant (Figure 6Sb).

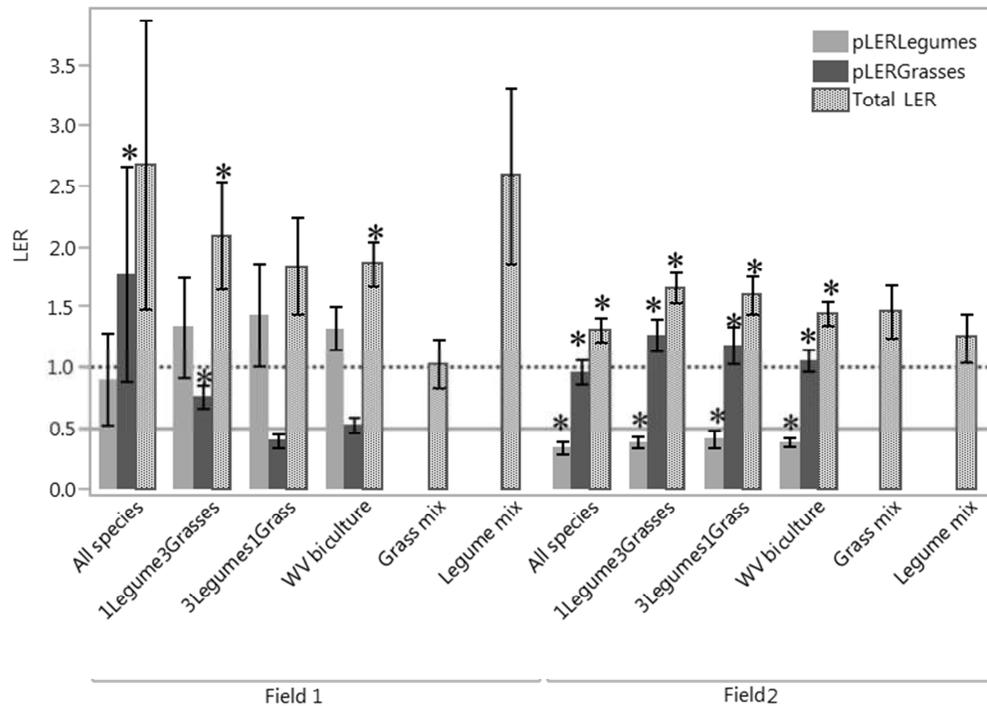


Figure 5 Total and partial LERs of cover crop mixtures by field; pooled fertilizer treatments. Significant difference from reference indicated by * ($p < 0.05$, t-test on log transformed data, but untransformed data presented). Reference is 1 for total LER (dashed line) and 0.5 for partial LER (solid line). All species, all six species (legumes: vetch, pea, clover; grasses: rye, wheat, ryegrass) together; 1Legume3Grasses, vetch and the three grasses; 3Legumes1Grass, the three legumes and wheat; WV biculture, wheat and vetch mixtures; Grass mix, the three grasses only; Legume mix, the three legumes only. Error bars represent one standard error from the mean.

Discussion

Incorporating cover crops into a conventional crop rotation where cash crops normally alternate with bare fallows is an effective way to increase diversity in agroecosystems, and provide ecosystem services beyond improvements in cash crop yield (Davis et al., 2012; King and Blesh, 2018; McDaniel et al., 2014; Smith et al., 2008; Tonitto et al., 2006). Diverse cover crop plantings have the potential to extend these benefits even further, particularly with a better understanding of the relationship between diversity and ecosystem services in agricultural systems. We found that mixtures with the two functional groups, grasses and legumes, provided the greatest increases in measured ecosystem services. These mixtures minimized the tradeoffs of using grasses or legumes alone, and lessened the impact of variation in soil N availability. However, increased intraspecific diversity had a minimal impact on ecosystem services. In some cases, reduced available soil N increased the strength of the diversity effect on biomass production, lending some support to the stress-gradient hypothesis (He et al., 2013).

Intraspecific diversity: limited benefits for ecosystem services

In general, greater intraspecific diversity did not consistently enhance the ecosystem services we measured, though we did observe an overall benefit for weed suppression in cultivar mixtures. In choosing cultivars and populations, we attempted to select contrasting cultivars to provide an increase in the overall functional trait diversity in the mixture. Functional trait diversity has been shown to be highly correlated with ecosystem function in unmanaged ecosystems (Cadotte et al., 2011; Flynn et al., 2011; Roscher et al., 2012). However, there were three barriers to achieving this goal of selecting contrasting cultivars with a wide trait range. First, detailed information about the characteristics of cultivars and populations was not always

available. Second, for several species used in this study there has been little or no effort to develop cultivars. For example, there has been very little cultivar selection for vetch, resulting in few recognized cultivars (Maul et al., 2011), two of which we did use ('Purple Prosperity' and 'Purple Bounty'). Third, out-crossing species such as vetch tend to have significant variation within populations further reducing the potential increase from adding another cultivar or line (Maul et al., 2011; Yeater et al., 2004).

Additionally, for four of the six species we used, our cultivar/population mixtures consisted of only two different lines. Our meta-analysis found that significantly greater benefits accrue with four or more cultivars, consequently this may have been a key limiting factor for species with fewer populations in mixtures (Reiss and Drinkwater, 2018). While we did not observe a strong relationship between intraspecific diversity and ecosystem services in this study, the responses we did detect (e.g. weed suppression for three species) were all positive improvements in ecosystem services. This is generally consistent with other results in cultivar mixtures, though important caveats remain regarding the use of refuges (cultivar mixtures) and resistant genes in plant populations to ensure durability of pest control (Grettenberger and Tooker, 2015; Mundt, 2002; Vlachostergios et al., 2011).

Species richness and functional diversity enhance ecosystem service provision

The grass/legume mixtures in our study, even at low species richness levels, were comparably or more productive than either functional group alone. The increase in functional diversity likely drives the gains in ecosystem services we observed with species mixtures. This is especially true in the grass-legume mixtures where functional groups with contrasting N acquisition strategies were present (Cadotte et al., 2011; Petchey and Gaston, 2002). There is strong evidence to suggest that functional diversity is a better descriptor of the components of a

community that influence ecosystem function compared to simple measures such as species richness, though they are often correlated. This is true in both unmanaged and agricultural ecosystems (Cadotte et al., 2011; Ebeling et al., 2014; Finney et al., 2016; Martin and Isaac, 2015; Roscher et al., 2012). Grasses and legumes are obviously distinguished by their different N acquisition strategies, as well as the clear morphology differences as monocots and dicots. These differences in N uptake as well as growth habits help to distinguish the ecological niches of grasses and legumes, ultimately impacting other plant characteristics such as C/N ratio and seedling vigor (Longo et al., 2013).

In our study the grasses tended to benefit more than the legumes when grown in mixtures and the increase in biomass in mixtures was driven largely by increased grass biomass. In these mixtures the grasses take up the bulk of available soil N while legumes are able to maintain their productivity through SNF (Bedoussac et al., 2015; Temperton et al., 2006). In this case, the differences in functional traits defining N acquisition are the strongest, though other trait differences within these functional groups may also contribute, such as growth habit and height (Eviner and Chapin III, 2003; Roscher et al., 2012). We used CV of biomass as a proxy for trait diversity for the six species in this study. The variation across the functional groups was nearly double that of the legumes alone.

The fact that legumes can down-regulate their symbiotic nitrogen fixation in response to available soil N also has important implications for how legumes can moderate the effect of environmental variability, as well as the tradeoffs of legumes or grasses alone. In contrast to studies with artificial application of N or multiple site conditions, we used the large amount of *in situ* plant-available soil N heterogeneity across Field 2. Specifically, we were able to explain more than a quarter of the variation in N fixation rates for vetch and pea in grass/legume

mixtures by the plant available soil N. In these mixtures, the legumes clearly downregulated N fixation when plant available soil N was high. The regulation of symbiotic nitrogen fixation by soil nitrate has been documented across sites with different soils or in treatments receiving different rates of inorganic fertilizer application (Kiers et al., 2003; Schipanski et al., 2010; Unkovich et al., 2008; Waterer and Vessey, 1993). For example, in a long-term cropping systems experiment, soybeans and clover down-regulated their rates of N fixation (15% and 19% respectively) in soils with a history of compost application and consequently higher soil N levels (Snapp et al., 2017). From a cover crop management perspective, it would be advantageous to capitalize on this self-regulation by selecting legume species and cultivars that can prioritize soil N uptake and retention when available, and still fix N when it is less available (Rengel, 2002).

In our experiment, the rate of N fixation increased for legumes in species mixtures with grasses compared to monocultures, although legume biomass was generally not strongly suppressed by the grasses, the overall amount of fixed N did not increase in mixtures. While we would expect for the rate to increase in such mixtures, the decrease in legume biomass in the mixtures due to the seeding proportions was the main driver in our experiment for overall fixed N, not the rate (Schipanski and Drinkwater, 2011). The legume biomass in mixtures was on average half that of the monoculture, corresponding to the seeding proportions. This suggests that there was not substantial competition by the grasses suppressing legume growth, though there is some evidence of this in the higher fertility Field 2 (White et al., 2017). The absence of an interaction between the rate of N fixation and the field suggests that the increased fixation response is a robust trait supported by the mixing of functional groups. However, the outcome of the total amount of fixed N is dependent on field conditions and the plant community effects on total biomass outcomes.

The grass/legume mixtures successfully moderated some of the tradeoffs of each functional group alone, and buffered the impact of site conditions such as soil N fertility. There was no impact of the soil conditions in the two fields on C/N for the mixtures, which also had an intermediate ratio between the two functional groups. The grasses efficiently took up any available soil N, while the legumes compensated when soil N was low through fixation, maintaining overall N in the aboveground biomass (White et al., 2017). It is also worth noting that aboveground biomass and C/N were not strongly related for the mixtures and the legumes; both had relatively low C/N at small and large biomass levels, a valuable consideration in terms of organic matter contribution and nutrient availability. In contrast, larger biomass production in grasses was significantly correlated with increases in C/N, increasing the risk that N immobilization could negatively impact crop growth with a large stand of cover crop biomass (Jensen, 1997).

The value of the moderate 22:1 C/N of the mixtures is also important for subsequent crop nutrition. When net N mineralization is not well synchronized with crop acquisition, the potential for environmental N losses is greater (Quemada and Cabrera, 1995; Schweizer et al., 1999). When relying on cover crops or other organic matter as a primary crop nutrient source, synchronizing N availability is critical (Poffenbarger et al., 2015a). Grasses, with average C/N of 33:1 might immobilize N when the crop needs it, while legumes have a very narrow C/N, commonly about 10:1, and often decompose quickly before the crop's peak demand (Snapp et al., 2005). In addition to the consistency of C/N in the mixtures and across biomass levels, the 22:1 ratio of the mixtures is more appropriate for adequate nutrient release for a subsequent crop compared to either grasses or legumes grown alone (Ruffo and Bollero, 2003). The consistency of an appropriate C/N in the functional group mixtures across the two fields suggests that these

mixtures are a robust choice across very different soil conditions.

Finally, weed suppression was generally good across all the treatments, regardless of composition, with nearly two thirds of plots with 5g/m^2 or less weed biomass. We observed a tendency for grass-legume mixtures to have lower weed biomass compared to only grasses or legumes at lower levels of cover crop biomass ($< 200\text{g/m}^2$). This suggests that perhaps other mechanisms beyond direct competition for resources reflected in biomass production were involved in weed suppression. Similarly, we did not see significantly greater aboveground cover crop biomass in cultivar mixtures of vetch and ryegrass, even though weed control was significantly greater in these mixtures. More complete and efficient resource use by the species mixtures may have limited the available resources for weed growth (Christensen, 1995; Kruidhof et al., 2008; Masilionyte et al., 2017).

Our study included only a small number of species in mixtures resulting in a maximum of 5 species in our most diverse plots ($n=14$). In unmanaged ecosystems the consensus is when overall species richness is low that the addition of each species has a substantial impact on overall functioning (Cardinale et al., 2011). While the low species richness in our experiment is typical for agricultural systems and research, it may have been too low and restricted our ability to detect more ecosystem service changes, even at these low levels where we would have expected to see a change. We were able to quantify a significant increase for certain ecosystem services such as aboveground biomass and weed suppression, which we were able to relate directly to species richness. For other services, we were able to explain a sizable portion of the variation through cover crop biomass. Expanding the species pool to a larger number and a greater representation of higher diversity mixtures may help to highlight the ecosystem service impacts.

Our approach to seeding rate and density, while consistent with our objectives to evaluate the effect of plant community diversity on ecosystem services, may also have influenced our results and outcomes. The density of a stand can have substantial impacts on the total biomass as well as the performance of the components in that community (Connolly, 1986). If we had selected a different seeding density, one that corresponded with lower legume rates for instance, the dynamics of the grass/legume mixtures evaluated may have been more dominated by grasses. Of course, there are many factors other than initial seeding density or proportions that determine ultimate stand composition, even in short lived annuals, such as time of seeding (Mirsky et al., 2017). Plant architecture and growth habits (e.g. viney legumes and tillering small grains) can also contribute to both stand composition through competition as well as overall biomass outcomes. At a given seeding density, a tillering grass may be able to exploit open space and produce high biomass, when seeded at a lower rate in mixtures. Finally, while we did not consider seed costs while developing our seeding rates, economics likely influence the recommended rates for cover crops. Cheaper grass seed is generally favored over more expensive legume seed, resulting in skewed proportions that may not be best suited to the given desired outcomes.

Enhanced diversity effect under more stressful conditions

In some instances, in our experiment, the effect of species richness was increased under less favorable conditions, providing some support for the stress-gradient hypothesis (SGH) under nutrient stress conditions. The SGH suggests that the greater facilitation between plants in high stress environments accentuates the diversity effect in these conditions, while under lower stress the diversity benefit is not as important or detectable (He et al., 2013). Few studies of the SGH have investigated the impact of the specific resource stress of nitrogen availability or even

nutrients in general. One study in Glacier Bay, Alaska suggested that variations in nutrients across the landscape could shift the balance towards facilitation and away from competition between a conifer and a nitrogen-fixing tree (Callaway and Walker, 1997), though this is obviously influenced by the nutrient acquisition traits of these species. Interestingly, the strongest evidence in our study for the SGH was in the grass monocultures and mixtures, where there was no change in biomass in the species mixtures when fertilized, but a significant increase in biomass when unfertilized. For mixtures including grasses and legumes we also observed slightly, though not always significantly, increased biomass under lower fertility conditions, such as rates of transgressive overyielding and LER and response to increasing species richness.

Complementary rooting structures, or preferences for different soil N forms (ammonia vs. nitrate) may explain some of the facilitation allowing the species mixtures to perform better than expected under these more nutrient limited conditions (Bedoussac et al., 2015; Hauggaard-Nielsen and Jensen, 2005; Kiær et al., 2013). Additionally, in agricultural systems, legumes and grasses have been shown to facilitate growth beyond basic nutrient uptake differences, but also through cultivation of beneficial microbial communities (Duchene et al., 2017). Both grasses and herbs tend to be more competitive under lower stress conditions, but can shift dramatically towards facilitation with increased stress compared to trees and shrubs (He et al., 2013). These characteristics make annual cover crop species ideal for responding to variations in resource stresses across fields and farms. Identifying the threshold for where nutrient stress conditions may begin to impact the diversity response would be important in targeting both further research of the SGH in agricultural systems, as well as practical application of diversity where it would have the greatest benefit.

Lower fertility conditions (both imposed experimentally as well as background

conditions) interacted with plant community diversity to drive several ecosystem service outcomes we measured. This highlights the fact that the design and conditions of cover crop mixture experiments may influence the conclusions from these trials with regard to the strength of the effect of diversity. Some recent cover crop diversity experiments have found inconsistent outcomes from more diverse mixtures in terms of transgressive overyielding (Finney et al., 2016; Smith et al., 2014; Wortman et al., 2012). This may be explained by generally good conditions and management practices in small-scale experiments conducted on research farms. In our experiment alone, we found inconsistent results for transgressive overyielding, with increased rates in one field and fertility treatment, but not others. Additionally, while our experimental design included a basic plus/minus fertilizer treatment to test the effect of N availability on these cover crop monocultures and mixtures, many more studies of environment-cover crop interactions are needed. In particular, treatments designed to specifically test how soil fertility and other stressors influence the diversity-ecosystem function relationship need to be included in research aimed at optimizing cover crop mixtures. Ultimately, more targeted experimental trials could help expand the stress-gradient hypothesis into agroecosystems so that we can develop reliable mixtures for a wide range of environmental conditions.

Conclusions and practical outcomes

We observed significant benefits from increasing species richness, especially when greater species richness coincided with increased functional diversity. Specifically, species mixtures had stronger effects on biomass production compared to intraspecific mixtures. When expanding the conclusions from this experiment to broader applications, it is important to note again that the seeding rates used are not typical for on-farm plantings of cover crops, and factors

such as management constraints or cost-limitations would be additional factors to consider. Mixtures of grasses and legumes, which contain the greatest functional diversity, were particularly effective at increasing ecosystem service performance, while also buffering the impact of environmental variation such as soil fertility. Consequently, diverse mixtures may be especially useful in poorer soils or in fields where there is substantial soil spatial heterogeneity. Additionally, as climate change increases the frequency of extreme weather events such as drought and excessive rainfall, the resulting unpredictable growing conditions will increase the need for planting cover crops composed of species that tolerate these differing conditions.

REFERENCES

- Andrew, I.K.S., Storkey, J., Sparkes, D.L., 2015. A review of the potential for competitive cereal cultivars as a tool in integrated weed management. *Weed Res* 55, 239–248.
<https://doi.org/10.1111/wre.12137>
- Aronsson, H., Hansen, E.M., Thomsen, I.K., Liu, J., Øgaard, A.F., Kähkönen, H., Ulén, B., 2016. The ability of cover crops to reduce nitrogen and phosphorus losses from arable land in southern Scandinavia and Finland. *Journal of Soil and Water Conservation* 71, 41–55.
<https://doi.org/10.2489/jswc.71.1.41>
- Baba, M., Halim, R.A., Alimon, A.R., Abubakar, I., 2011. Grass-legume mixtures for enhanced forage production: Analysis of dry matter yield and competition indices. *African Journal of Agricultural Research* 6, 5242–5250.
- Bateman, G.Q., Keller, W., 1956. Grass-legume mixtures for irrigated pastures for dairy cows.
- Bedoussac, L., Journet, E.-P., Hauggaard-Nielsen, H., Naudin, C., Corre-Hellou, G., Jensen, E.S., Prieur, L., Justes, E., 2015. Ecological principles underlying the increase of productivity achieved by cereal-grain legume intercrops in organic farming. A review. *Agronomy for Sustainable Development* 35, 911–935. <https://doi.org/10.1007/s13593-014-0277-7>
- Bertness, M.D., Callaway, R., 1994. Positive interactions in communities. *Trends in Ecology & Evolution* 9, 191–193.
- Bivand, R.S., Pebesma, E., Gomez-Rubio, V., 2013. *Applied spatial data analysis with R*, Second edition. Springer, NY.

- Bommarco, R., Kleijn, D., Potts, S.G., 2013. Ecological intensification: harnessing ecosystem services for food security. *Trends in Ecology & Evolution* 28, 230–238.
<https://doi.org/10.1016/j.tree.2012.10.012>
- Boyd, N.S., Brennan, E.B., 2006. Weed Management in a Legume-Cereal Cover Crop with the Rotary Hoe 1. *Weed Technology* 20, 733–737. <https://doi.org/10.1614/WT-05-157R.1>
- Brainard, D., Henshaw, B., Snapp, S., 2012. Hairy Vetch Varieties and Bi-Cultures Influence Cover Crop Services in Strip-Tilled Sweet Corn. *Agronomy Journal* 104, 629.
<https://doi.org/10.2134/agronj2011.0360>
- Brainard, D.C., Bellinder, R.R., Kumar, V., 2011. Grass-Legume Mixtures and Soil Fertility Affect Cover Crop Performance and Weed Seed Production. *Weed Technology* 25, 473–479. <https://doi.org/10.1614/WT-D-10-00134.1>
- Bybee-Finley, K.A., Mirsky, S.B., Ryan, M.R., 2016. Functional Diversity in Summer Annual Grass and Legume Intercrops in the Northeastern United States. *Crop Science* 56, 2775.
<https://doi.org/10.2135/cropsci2016.01.0046>
- Cadotte, M.W., Cardinale, B.J., Oakley, T.H., 2008. Evolutionary history and the effect of biodiversity on plant productivity. *Proceedings of the National Academy of Sciences* 105, 17012–17017.
- Cadotte, M.W., Carscadden, K., Mirotchnick, N., 2011. Beyond species: functional diversity and the maintenance of ecological processes and services. *Journal of Applied Ecology* 48, 1079–1087. <https://doi.org/10.1111/j.1365-2664.2011.02048.x>

- Cadotte, M.W., Cavender-Bares, J., Tilman, D., Oakley, T.H., 2009. Using Phylogenetic, Functional and Trait Diversity to Understand Patterns of Plant Community Productivity. *PLoS ONE* 4, e5695. <https://doi.org/10.1371/journal.pone.0005695>
- Callaway, R.M., Walker, L.R., 1997. Competition and facilitation: a synthetic approach to interactions in plant communities. *Ecology* 78, 1958–1965. [https://doi.org/10.1890/0012-9658\(1997\)078\[1958:CAFASA\]2.0.CO;2](https://doi.org/10.1890/0012-9658(1997)078[1958:CAFASA]2.0.CO;2)
- Cardinale, B.J., Matulich, K.L., Hooper, D.U., Byrnes, J.E., Duffy, E., Gamfeldt, L., Balvanera, P., O'Connor, M.I., Gonzalez, A., 2011. The functional role of producer diversity in ecosystems. *Am. J. Bot.* 98, 572–592. <https://doi.org/10.3732/ajb.1000364>
- Carlson, S., Stockwell, R., 2013. Research Priorities for Advancing Adoption of Cover Crops in Agriculture-intensive Regions. *Journal of Agriculture, Food Systems, and Community Development* 1–5. <https://doi.org/10.5304/jafscd.2013.034.017>
- Christensen, S., 1995. Weed suppression ability of spring barley varieties. *Weed Research* 35, 241–247. <https://doi.org/10.1111/j.1365-3180.1995.tb01786.x>
- Clark, A. (Ed.), 2007. *Managing cover crops profitably*, 3rd ed. ed, Handbook series. Sustainable Agriculture Research & Education (Program), College Park, MD.
- Connolly, J., 1986. On Difficulties with Replacement-Series Methodology in Mixture Experiments. *Journal of Applied Ecology* 23, 125–137. <https://doi.org/10.2307/2403086>
- Cooper, C.S., Eslick, R.F., Stitt, R.E., 1960. Yield performance of simple irrigated grass - legume pasture mixtures at Bozeman, Montan. *Bull Montana Agric Expt Sta* 53, 3–12.
- Davis, A.S., Hill, J.D., Chase, C.A., Johanns, A.M., Liebman, M., 2012. Increasing Cropping System Diversity Balances Productivity, Profitability and Environmental Health. *PLoS ONE* 7, e47149. <https://doi.org/10.1371/journal.pone.0047149>

- Duchene, O., Vian, J.-F., Celette, F., 2017. Intercropping with legume for agroecological cropping systems: Complementarity and facilitation processes and the importance of soil microorganisms. A review. *Agriculture, Ecosystems & Environment* 240, 148–161. <https://doi.org/10.1016/j.agee.2017.02.019>
- Dunn, M., Ulrich-Schad, J.D., Prokopy, L.S., Myers, R.L., Watts, C.R., Scanlon, K., 2016. Perceptions and use of cover crops among early adopters: Findings from a national survey. *Journal of Soil and Water Conservation* 71, 29–40. <https://doi.org/10.2489/jswc.71.1.29>
- Duvick, D.N., 2005. The contribution of breeding to yield advances in maize (*Zea mays* L.). *Advances in agronomy* 86, 83–145.
- Dwivedi, S.L., Ceccarelli, S., Blair, M.W., Upadhyaya, H.D., Are, A.K., Ortiz, R., 2016. Landrace Germplasm for Improving Yield and Abiotic Stress Adaptation. *Trends in Plant Science* 21, 31–42. <https://doi.org/10.1016/j.tplants.2015.10.012>
- Ebeling, A., Pompe, S., Baade, J., Eisenhauer, N., Hillebrand, H., Proulx, R., Roscher, C., Schmid, B., Wirth, C., Weisser, W.W., 2014. A trait-based experimental approach to understand the mechanisms underlying biodiversity–ecosystem functioning relationships. *Basic and Applied Ecology* 15, 229–240. <https://doi.org/10.1016/j.baae.2014.02.003>
- Elzebroek, A.T.G., 2008. *Guide to Cultivated Plants*. CABI.
- Eviner, V.T., Chapin III, F.S., 2003. Functional Matrix: A Conceptual Framework for Predicting Multiple Plant Effects on Ecosystem Processes. *Annual Review of Ecology, Evolution, and Systematics* 34, 455–485.

- Finney, D.M., White, C.M., Kaye, J.P., 2016. Biomass Production and Carbon/Nitrogen Ratio Influence Ecosystem Services from Cover Crop Mixtures. *Agronomy Journal* 108, 39. <https://doi.org/10.2134/agronj15.0182>
- Flynn, D.F.B., Mirotchnick, N., Jain, M., Palmer, M.I., Naeem, S., 2011. Functional and phylogenetic diversity as predictors of biodiversity–ecosystem-function relationships. *Ecology* 92, 1573–1581. <https://doi.org/10.1890/10-1245.1>
- Fu, Y.-B., 2015. Understanding crop genetic diversity under modern plant breeding. *Theor Appl Genet* 128, 2131–2142. <https://doi.org/10.1007/s00122-015-2585-y>
- Gaudin, A.C.M., Tolhurst, T.N., Ker, A.P., Janovicek, K., Tortora, C., Martin, R.C., Deen, W., 2015. Increasing Crop Diversity Mitigates Weather Variations and Improves Yield Stability. *PLOS ONE* 10, e0113261. <https://doi.org/10.1371/journal.pone.0113261>
- Grace, J.B., Anderson, T.M., Seabloom, E.W., Borer, E.T., Adler, P.B., Harpole, W.S., Hautier, Y., Hillebrand, H., Lind, E.M., Pärtel, M., Bakker, J.D., Buckley, Y.M., Crawley, M.J., Damschen, E.I., Davies, K.F., Fay, P.A., Firn, J., Gruner, D.S., Hector, A., Knops, J.M.H., MacDougall, A.S., Melbourne, B.A., Morgan, J.W., Orrock, J.L., Prober, S.M., Smith, M.D., 2016. Integrative modelling reveals mechanisms linking productivity and plant species richness. *Nature* 529, 390–393. <https://doi.org/10.1038/nature16524>
- Gräler, B., Pebesma, E., Heuvelink, G., 2016. Spatio-Temporal Interpolation using gstat. *The R Journal* 8, 204–218.
- Grettenberger, I.M., Tooker, J.F., 2015. Moving beyond resistance management toward an expanded role for seed mixtures in agriculture. *Agriculture, Ecosystems & Environment* 208, 29–36. <https://doi.org/10.1016/j.agee.2015.04.019>

- Hauggaard-Nielsen, H., Jensen, E.S., 2005. Facilitative root interactions in intercrops. *Plant & Soil, Plant Ecophysiology* 274, 237–250.
- He, Q., Bertness, M.D., Altieri, A.H., 2013. Global shifts towards positive species interactions with increasing environmental stress. *Ecology Letters* 16, 695–706.
<https://doi.org/10.1111/ele.12080>
- Hooper, D.U., Chapin, F.S., Ewel, J.J., Hector, A., Inchausti, P., Lavorel, S., Lawton, J.H., Lodge, D.M., Loreau, M., Naeem, S., Schmid, B., Setälä, H., Symstad, A.J., Vandermeer, J., Wardle, D.A., 2005. Effects of biodiversity on ecosystem functioning: A consensus of current knowledge. *Ecological Monographs* 75, 3–35. <https://doi.org/10.1890/04-0922>
- Jensen, E.S., 1997. Nitrogen immobilization and mineralization during initial decomposition of¹⁵N-labelled pea and barley residues. *Biol Fert Soils* 24, 39–44. <https://doi.org/10.1007/BF01420218>
- Kaut, A.H.E.E., Mason, H.E., Navabi, A., O'Donovan, J.T., Spaner, D., 2008. Performance and stability of performance of spring wheat variety mixtures in organic and conventional management systems in western Canada. *The Journal of Agricultural Science* 147, 141.
<https://doi.org/10.1017/S0021859608008319>
- Kiær, L.P., Skovgaard, I.M., Østergård, H., 2009. Grain yield increase in cereal variety mixtures: A meta-analysis of field trials. *Field Crops Research* 114, 361–373.
<https://doi.org/10.1016/j.fcr.2009.09.006>
- Kiær, L.P., Weisbach, A.N., Weiner, J., 2013. Root and shoot competition: A meta-analysis. *Journal of Ecology* n/a–n/a. <https://doi.org/10.1111/1365-2745.12129>
- Kiers, E.T., Rousseau, R.A., West, S.A., Denison, R.F., 2003. Host sanctions and the legume-rhizobium mutualism. *Nature* 425, 78–81. <https://doi.org/10.1038/nature01931>

- King, A.E., Blesh, J., 2018. Crop rotations for increased soil carbon: perenniality as a guiding principle. *Ecol Appl* 28, 249–261. <https://doi.org/10.1002/eap.1648>
- Kruidhof, H.M., Bastiaans, L., Kropff, M.J., 2008. Ecological weed management by cover cropping: effects on weed growth in autumn and weed establishment in spring. *Weed Res.* 48, 492–502. <https://doi.org/10.1111/j.1365-3180.2008.00665.x>
- Longo, G., Seidler, T.G., Garibaldi, L.A., Tognetti, P.M., Chaneton, E.J., 2013. Functional group dominance and identity effects influence the magnitude of grassland invasion. *Journal of Ecology* n/a–n/a. <https://doi.org/10.1111/1365-2745.12128>
- Martin, A.R., Isaac, M.E., 2015. REVIEW: Plant functional traits in agroecosystems: a blueprint for research. *J Appl Ecol* 52, 1425–1435. <https://doi.org/10.1111/1365-2664.12526>
- Masilionyte, L., Maiksteniene, S., Kriauciuniene, Z., Jablonskyte-Rasce, D., Zou, L., Sarauskis, E., 2017. Effect of cover crops in smothering weeds and volunteer plants in alternative farming systems. *Crop Protection* 91, 74–81. <https://doi.org/10.1016/j.cropro.2016.09.016>
- Maul, J., Mirsky, S., Emche, S., Devine, T., 2011. Evaluating a Germplasm Collection of the Cover Crop Hairy Vetch for Use in Sustainable Farming Systems. *Crop Science* 51, 2615. <https://doi.org/10.2135/cropsci2010.09.0561>
- McDaniel, M.D., Tiemann, L.K., Grandy, A.S., 2014. Does agricultural crop diversity enhance soil microbial biomass and organic matter dynamics? A meta-analysis. *Ecological Applications* 24, 560–570. <https://doi.org/10.1890/13-0616.1>
- McGill, B.J., Enquist, B.J., Weiher, E., Westoby, M., 2006. Rebuilding community ecology from functional traits. *Trends in Ecology & Evolution* 21, 178–185. <https://doi.org/10.1016/j.tree.2006.02.002>

- Mirsky, S.B., Ackroyd, V.J., Cordeau, S., Curran, W.S., Hashemi, M., Reberg-Horton, S.C., Ryan, M.R., Spargo, J.T., 2017. Hairy Vetch Biomass across the Eastern United States: Effects of Latitude, Seeding Rate and Date, and Termination Timing. *Agronomy Journal* 109, 1510. <https://doi.org/10.2134/agronj2016.09.0556>
- Mundt, C.C., 2002. Use of Multiline Cultivars and Cultivar Mixtures for Disease Management. *Annual Review of Phytopathology* 40, 381–410. <https://doi.org/10.1146/annurev.phyto.40.011402.113723>
- Newton, A.C., Akar, T., Baresel, J.P., Bebeli, P.J., Bettencourt, E., Bladenopoulos, K.V., Czembor, J.H., Fasoula, D.A., Katsiotis, A., Koutis, K., Koutsika-Sotiriou, M., Kovacs, G., Larsson, H., de Carvalho, M.A.A.P., Rubiales, D., Russell, J., Dos Santos, T.M.M., Vaz Patto, M.C., 2010. Cereal landraces for sustainable agriculture. A review. *Agronomy for Sustainable Development* 30, 237–269. <https://doi.org/10.1051/agro/2009032>
- Nyfeler, D., Huguenin-Elie, O., Suter, M., Frossard, E., Connolly, J., Luescher, A., 2009. Strong mixture effects among four species in fertilized agricultural grassland led to persistent and consistent transgressive overyielding. *Journal of Applied Ecology* 46, 683–691. <https://doi.org/10.1111/j.1365-2664.2009.01653.x>
- Osman, A., Diek, A., 1982. Effects of defoliation on yield and forage quality of some tropical grasses, legumes and their mixtures. *Experimental Agriculture* 18, 157–166. <https://doi.org/10.1017/S0014479700013636>
- Parr, M., 2010. Nitrogen Fixation and Mineralization Potential of Winter Annual Legume Cover Crops for Reduced-tillage Organic Corn Production in North Carolina. North Carolina State University.

- Pebesma, E.J., 2004. Multivariable geostatistics in S: the gstat package. *Computers & Geosciences* 30, 683–691.
- Pebesma, E.J., Bivand, R.S., 2005. Classes and methods for spatial data in R. *R News* 5, 9–13.
- Petchey, O.L., Gaston, K.J., 2002. Functional diversity (FD), species richness and community composition. *Ecology Letters* 5, 402–411. <https://doi.org/10.1046/j.1461-0248.2002.00339.x>
- Petersen, B., Snapp, S., 2015. What is sustainable intensification? Views from experts. *Land Use Policy* 46, 1–10. <https://doi.org/10.1016/j.landusepol.2015.02.002>
- Poffenbarger, H.J., Mirsky, S.B., Weil, R.R., Kramer, M., Spargo, J.T., Cavigelli, M.A., 2015a. Legume Proportion, Poultry Litter, and Tillage Effects on Cover Crop Decomposition. *Agronomy Journal* 107, 2083. <https://doi.org/10.2134/agronj15.0065>
- Poffenbarger, H.J., Mirsky, S.B., Weil, R.R., Maul, J.E., Kramer, M., Spargo, J.T., Cavigelli, M.A., 2015b. Biomass and Nitrogen Content of Hairy Vetch–Cereal Rye Cover Crop Mixtures as Influenced by Species Proportions. *Agronomy Journal* 107, 2069. <https://doi.org/10.2134/agronj14.0462>
- Quemada, M., Cabrera, M., 1995. Carbon and Nitrogen Mineralized from Leaves and Stems of 4 Cover Crops. *Soil Sci. Soc. Am. J.* 59, 471–477.
- R Core Team, 2016. A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Reiss, E.R., Drinkwater, L.E., 2018. Cultivar mixtures: a meta-analysis of the effect of intraspecific diversity on crop yield. *Ecol Appl* 28, 62–77. <https://doi.org/10.1002/eap.1629>

- Reiss, E.R., Drinkwater, L.E., 2017. Cultivar mixtures: A meta-analysis of the effect of intraspecific diversity on crop yield. *Ecological Applications*.
- Rengel, Z., 2002. Breeding for better symbiosis. *Plant and Soil* 245, 147–162.
<https://doi.org/10.1023/A:1020692715291>
- Ringuet, S., Sassano, L., Johnson, Z.I., 2011. A suite of microplate reader-based colorimetric methods to quantify ammonium, nitrate, orthophosphate and silicate concentrations for aquatic nutrient monitoring. *J. Environ. Monit.* 13, 370–376.
<https://doi.org/10.1039/C0EM00290A>
- Roscher, C., Schumacher, J., Gubsch, M., Lipowsky, A., Weigelt, A., Buchmann, N., Schmid, B., Schulze, E.-D., 2012. Using Plant Functional Traits to Explain Diversity–Productivity Relationships. *PLoS ONE* 7, e36760. <https://doi.org/10.1371/journal.pone.0036760>
- Ruffo, M.L., Bollero, G.A., 2003. Modeling rye and hairy vetch residue decomposition as a function of degree-days and decomposition-days. *Agronomy Journal* 95, 900–907.
- Schipanski, M., Drinkwater, L., 2011. Nitrogen fixation of red clover interseeded with winter cereals across a management-induced fertility gradient. *Nutrient Cycling in Agroecosystems* 90, 105–119. <https://doi.org/10.1007/s10705-010-9415-z>
- Schipanski, M., Drinkwater, L., Russelle, M., 2010. Understanding the variability in soybean nitrogen fixation across agroecosystems. *Plant and Soil* 329, 379–397.
<https://doi.org/10.1007/s11104-009-0165-0>
- Schipanski, M.E., Barbercheck, M., Douglas, M.R., Finney, D.M., Haider, K., Kaye, J.P., Kemanian, A.R., Mortensen, D.A., Ryan, M.R., Tooker, J., White, C., 2014. A framework for evaluating ecosystem services provided by cover crops in agroecosystems. *Agricultural Systems* 125, 12–22. <https://doi.org/10.1016/j.agsy.2013.11.004>

- Schweizer, M., Fear, J., Cadisch, G., 1999. Isotopic (C-13) fractionation during plant residue decomposition and its implications for soil organic matter studies. *Rapid Communications in Mass Spectrometry* 13, 1284–1290.
[https://doi.org/10.1002/\(SICI\)1097-0231\(19990715\)13:13<1284::AID-RCM578>3.3.CO;2-S](https://doi.org/10.1002/(SICI)1097-0231(19990715)13:13<1284::AID-RCM578>3.3.CO;2-S)
- Shearer, G., Kohl, D.H., 1986. N₂ fixation in field settings: estimates based on natural ¹⁵N abundance. *Australian Journal of Plant Physiology* 13, 699–756.
- Smith, R.G., Atwood, L.W., Warren, N.D., 2014. Increased Productivity of a Cover Crop Mixture Is Not Associated with Enhanced Agroecosystem Services. *PLOS ONE* 9, e97351. <https://doi.org/10.1371/journal.pone.0097351>
- Smith, R.G., Gross, K.L., Robertson, G.P., 2008. Effects of Crop Diversity on Agroecosystem Function: Crop Yield Response. *Ecosystems* 11, 355–366.
<https://doi.org/10.1007/s10021-008-9124-5>
- Snapp, S., Wilke, B., Gentry, L.E., Zoellner, D., 2017. Compost Legacy Down-Regulates Biological Nitrogen Fixation in a Long-Term Field Experiment. *Agronomy Journal* 109, 2662. <https://doi.org/10.2134/agronj2017.03.0152>
- Snapp, S.S., Swinton, S.M., Labarta, R., Mutch, D., Black, J.R., Leep, R., Nyiraneza, J., O’Neil, K., 2005. Evaluating cover crops for benefits, costs and performance within cropping system niches. *Agron. J.* 97, 322–332.
- Somasegaran, P., Hoben, H.J., 1994. *Handbook for Rhizobia*. Springer New York, New York, NY.
- Storkey, J., Brooks, D., Haughton, A., Hawes, C., Smith, B.M., Holland, J.M., 2013. Using functional traits to quantify the value of plant communities to invertebrate ecosystem

- service providers in arable landscapes. *Journal of Ecology* 101, 38–46.
<https://doi.org/10.1111/1365-2745.12020>
- Temperton, V.M., Mwangi, P.N., Scherer-Lorenzen, M., Schmid, B., Buchmann, N., 2006. Positive interactions between nitrogen-fixing legumes and four different neighbouring species in a biodiversity experiment. *Oecologia* 151, 190–205.
<https://doi.org/10.1007/s00442-006-0576-z>
- Thomas Björkman, n.d. NY Cover Crop Guide [WWW Document]. NY Cover Crop Guide. URL <http://covercrop.org/> (accessed 5.5.18).
- Tilman, D., Wedin, D., Knops, J., 1996. Productivity and sustainability influenced by biodiversity in grassland ecosystems. *Nature* 379, 718–720.
<https://doi.org/10.1038/379718a0>
- Tonitto, C., David, M.B., Drinkwater, L.E., 2006. Replacing bare fallows with cover crops in fertilizer-intensive cropping systems: A meta-analysis of crop yield and N dynamics. *Agriculture, Ecosystems & Environment* 112, 58–72.
<https://doi.org/10.1016/j.agee.2005.07.003>
- Unkovich, M., Pate, J., Sanford, P., Armstrong, E., 1994. Potential precision of the $\delta^{15}\text{N}$ natural abundance method in field estimates of nitrogen fixation by crop and pasture legumes in south-west Australia. *Aust. J. Agric. Res.* 45, 119–132.
- Unkovich, M.J., Herridge, D., Peoples, M., Cadisch, G., Boddey, R., Giller, K., Alves, B., Chalk, 2008. Measuring plant-associated nitrogen fixation in agricultural systems (No. ACIAR Monograph No. 136). Australia, Canberra.
- Vlachostergios, D.N., Lithourgidis, A.S., Dordas, C.A., Baxevanos, D., 2011. Advantages of Mixing Common Vetch Cultivars Developed from Conventional Breeding Programs

- when Grown under Low-Input Farming System. *Crop Science* 51, 1274.
<https://doi.org/10.2135/cropsci2010.07.0385>
- Waterer, J.G., Vessey, J.K., 1993. Effect of low static nitrate concentrations on mineral nitrogen uptake, nodulation, and nitrogen fixation in field pea. *Journal of Plant Nutrition* 16, 1775–1789. <https://doi.org/10.1080/01904169309364649>
- Wayman, S., Kissing Kucek, L., Mirsky, S.B., Ackroyd, V., Cordeau, S., Ryan, M.R., 2016. Organic and conventional farmers differ in their perspectives on cover crop use and breeding. *Renewable Agriculture and Food Systems* 1–10.
<https://doi.org/10.1017/S1742170516000338>
- White, C.M., DuPont, S.T., Hautau, M., Hartman, D., Finney, D.M., Bradley, B., LaChance, J.C., Kaye, J.P., 2017. Managing the trade off between nitrogen supply and retention with cover crop mixtures. *Agriculture, Ecosystems & Environment* 237, 121–133.
<https://doi.org/10.1016/j.agee.2016.12.016>
- Wortman, S.E., Francis, C.A., Lindquist, J.L., 2012. Cover Crop Mixtures for the Western Corn Belt: Opportunities for Increased Productivity and Stability. *Agronomy Journal* 104, 699.
<https://doi.org/10.2134/agronj2011.0422>
- Yeater, K.M., Bollero, G.A., Bullock, D.G., Rayburn, A.L., Rodriguez-Zas, S., 2004. Assessment of genetic variation in hairy vetch using canonical discriminant analysis. *Crop Science* 44, 185–189.
- Yu, Y., Stomph, T.-J., Makowski, D., van der Werf, W., 2015. Temporal niche differentiation increases the land equivalent ratio of annual intercrops: A meta-analysis. *Field Crops Research* 184, 133–144. <https://doi.org/10.1016/j.fcr.2015.09.010>

Supplemental information

Table 1S Cultivars used, by species, and seed source

	Hairy vetch (<i>Vicia villosa</i>)		Winter pea (<i>Pisum sativum</i>)		Crimson clover (<i>Trifolium incarnatum</i>)	
	Cultivar	Source	Cultivar	Source	Cultivar	Source
Legumes	AL Vetch	Albert Lea Seed	Austrian Pea	Lakeview Organic Grain	Dixie Clover	Adams-Briscoe Seed Company
	Vetch VNS	Albert Lea Seed	Windham (Year 1)	Byron Seeds, LLC	Clover VNS	Ernst Conservation Seeds
	Ernst Vetch VNS	Ernst Conservation Seeds	Whistler (Year 2)	Hancock Seed Company		
	Purple Bounty	Ernst Conservation Seeds				
	Purple Prosperity	USDA NRCS National Plant Materials Center, Beltsville, MD				
Grasses	Cereal rye (<i>Secale cereale</i>)		Wheat (<i>Triticum aestivum</i>)		Ryegrass (<i>Lolium multiflorum</i>)	
	Cultivar	Source	Cultivar	Source	Cultivar	Source
	Aroostook	Ernst Conservation Seeds	Cayuga	Cornell Small Grains Breeding & Genetics Program	Ryegrass VNS	Ernst Conservation Seeds
	Rye VNS	Lakeview Organic Grain	Caledonia		Jumbo	Byron Seeds, LLC
			Houser		MO-1 (Year 1)	Byron Seeds, LLC
			Red Fife (Year 1)			
		Pride of Genesee (Year 2)				

Table 2S Experimental seeding rate for species monocultures and comparison to recommended rates

Species	Experimental seeding rate (kg/ha)	Experimental rate as a % of recommended rate	Seed size (g/seed)
Vetch	85	233%	0.03
Crimson clover	17	58%	0.005
Winter pea	431	405%	0.15
Rye	68	49%	0.02
Wheat	116	99%	0.04
Ryegrass	8	29%	0.003

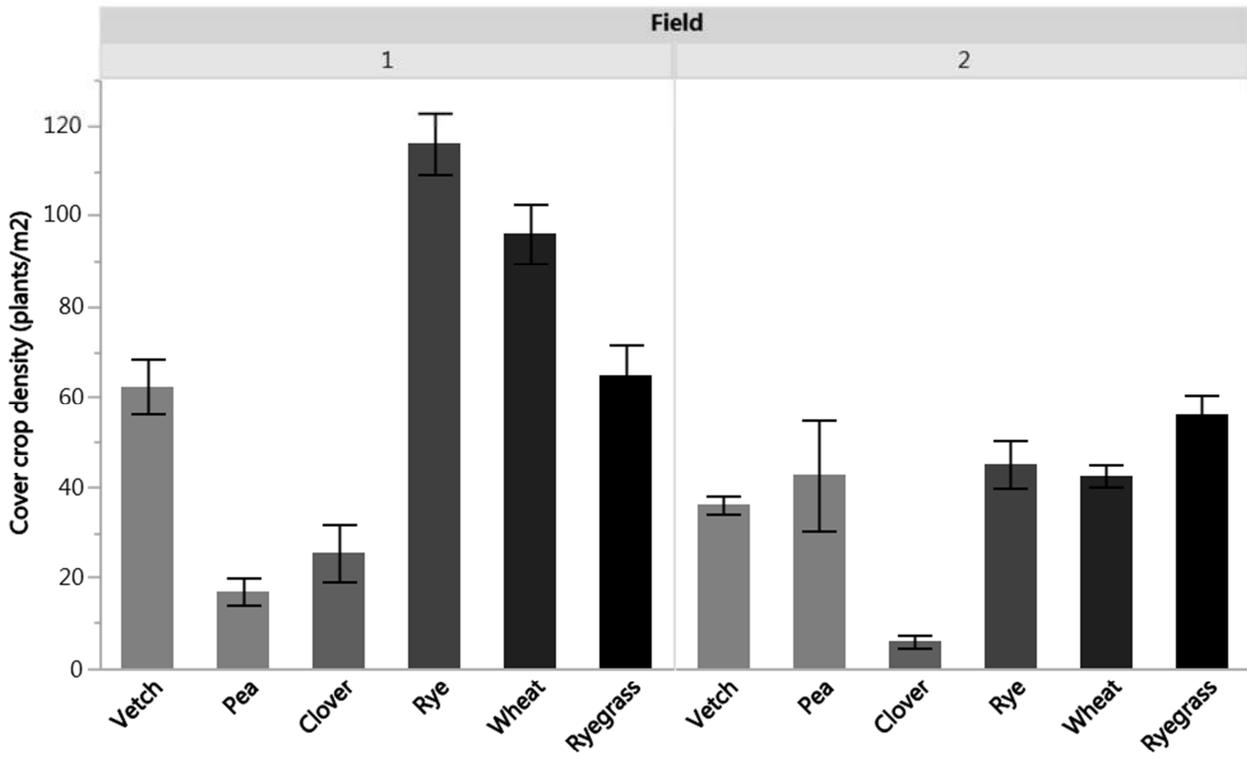


Figure 1S Mean cover crop density for species monocultures for the two fields. Error bars represent one standard error from the mean.

Table 3S *B* values of legume cultivars in study

Cultivar name (year if applicable)	Shoot $\delta^{15}\text{N}$ (‰)
Vetch	
Purple Bounty (2013)	-0.69
Purple Bounty (2014)	-0.72
Vetch VNS	-0.47
Ernst Vetch	-0.60
AL Vetch	-0.15
Purple Prosperity (2013)	-0.49
Purple Prosperity (2014)	-0.68
Crimson clover	
VNS clover	-1.34
Dixie clover	-1.21
Pea	
Windham winter pea	-0.55
Whistler winter pea	-0.77
Austrian winter pea (2013)	-0.89
Austrian winter pea (2014)	-0.88

Table 4S Summary of mean (SD) ecosystem service outcomes by cover crop treatment (species monocultures and mixtures) and field, split by a) unfertilized and, b) fertilized

	Cover crop biomass (g/m ²)		Legume biomass (g/m ²)		Grass biomass (g/m ²)		Percent weed control (%)		Soil N (g/m ²)		Fixed N (g/m ²)		Total N (g/m ²)	
	Field 1	Field 2	Field 1	Field 2	Field 1	Field 2	Field 1	Field 2	Field 1	Field 2	Field 1	Field 2	Field 1	Field 2
a) Unfertilized														
Vetch	297 (39)	225 (15)	297 (39)	225 (15)	-	-	81 (4)	62 (8)	1.5 (0.5)	5.2 (0.5)	7.7 (1.1)	2.4 (0.5)	9.2 (1.3)	7.5 (0.5)
Pea	133 (91)	289 (49)	133 (91)	289 (49)	-	-	53 (12)	69 (11)	0.5 (0.3)	4.2 (0.5)	1 (0.4)	4.4 (1.4)	1.5 (0.7)	8.5 (1.4)
Clover	35 (13)	22 (7)	35 (13)	22 (7)	-	-	37 (21)	48 (18)	0.5 (0.1)	0.4 (0.1)	0.5 (0.2)	0.1 (0)	1 (0.2)	0.4 (0.2)
Rye	544 (53)	651 (59)	-	-	544 (53)	651 (59)	98 (0)	96 (1)	3.7 (0.3)	9 (0.6)	-	-	3.7 (0.3)	9 (0.6)
Wheat	456 (66)	517 (48)	-	-	456 (66)	517 (48)	94 (1)	88 (4)	4.7 (0.6)	8.4 (0.7)	-	-	4.7 (0.6)	8.4 (0.7)
Ryegrass	189 (24)	381 (72)	-	-	189 (24)	381 (72)	70 (7)	58 (11)	2.3 (0.3)	7.2 (1)	-	-	2.3 (0.3)	7.2 (1)
WV biculture	387 (35)	529 (39)	196 (24)	79 (10)	192 (24)	450 (36)	92 (2)	84 (3)	2.8 (0.3)	9 (0.6)	6.7 (0.8)	0.9 (0.2)	9.5 (0.9)	9.9 (0.7)
Legume mix	212 (39)	331 (58)	212 (39)	331 (58)	-	-	79 (7)	67 (15)	1.9 (0.7)	5.5 (1)	7.4 (1.6)	4 (1.3)	9.2 (1.8)	9.6 (1.8)
Grass mix	546 (108)	741 (76)	-	-	546 (108)	741 (76)	97 (0)	95 (1)	3.9 (0.7)	10.1 (1.1)	-	-	3.9 (0.7)	10.1 (1.1)
Vetch and 3 grasses	649 (116)	748 (74)	243 (68)	98 (23)	406 (71)	650 (55)	93 (5)	94 (3)	3 (0.5)	9.3 (0.5)	6.8 (2.3)	1.8 (0.8)	9.9 (2.7)	11.1 (0.9)
3 legumes and wheat	389 (51)	638 (84)	184 (46)	119 (41)	205 (24)	519 (71)	95 (2)	84 (6)	2.7 (0.2)	10.4 (1.2)	6.4 (1.8)	1.8 (0.7)	9.2 (1.8)	12.2 (1.7)
All species mix	678 (192)	600 (103)	201 (69)	89 (16)	477 (133)	512 (110)	96 (1)	84 (10)	3.4 (0.7)	7.3 (1.3)	6.6 (2.2)	1.8 (0.5)	10 (2.7)	9 (1.2)
b) Fertilized														
Cover crop biomass (g/m²)														
Vetch	289 (41)	240 (16)	289 (41)	240 (16)	-	-	84 (3)	83 (5)	-	-	-	-	-	-
Pea	121 (72)	320 (51)	121 (72)	320 (51)	-	-	66 (8)	76 (9)	-	-	-	-	-	-
Clover	137 (48)	79 (3)	137 (48)	79 (3)	-	-	46 (15)	1 (0)	-	-	-	-	-	-
Rye	832 (95)	647 (76)	-	-	832 (95)	647 (76)	96 (1)	90 (5)	-	-	-	-	-	-
Wheat	723 (110)	563 (45)	-	-	723 (110)	563 (45)	92 (2)	87 (4)	-	-	-	-	-	-
Ryegrass	447 (67)	403 (60)	-	-	447 (67)	403 (60)	75 (4)	76 (6)	-	-	-	-	-	-
WV biculture	502 (41)	504 (30)	226 (32)	80 (9)	276 (29)	424 (25)	86 (4)	89 (2)	-	-	-	-	-	-
Legume mix	249 (68)	321 (44)	249 (68)	321 (44)	-	-	75 (9)	82 (5)	-	-	-	-	-	-
Grass mix	573 (177)	577 (83)	-	-	573 (177)	577 (83)	94 (1)	94 (2)	-	-	-	-	-	-
1Legume3Grasses	704 (133)	751 (113)	217 (67)	101 (18)	487 (105)	650 (119)	94 (2)	97 (1)	-	-	-	-	-	-
3Legumes1Grass	490 (81)	536 (72)	145 (45)	130 (33)	346 (51)	407 (70)	90 (4)	92 (3)	-	-	-	-	-	-
All species	783 (160)	577 (104)	137 (35)	99 (25)	646 (134)	477 (111)	94 (2)	95 (2)	-	-	-	-	-	-
Legume biomass (g/m²)														
Vetch	289 (41)	240 (16)	289 (41)	240 (16)	-	-	84 (3)	83 (5)	-	-	-	-	-	-
Pea	121 (72)	320 (51)	121 (72)	320 (51)	-	-	66 (8)	76 (9)	-	-	-	-	-	-
Clover	137 (48)	79 (3)	137 (48)	79 (3)	-	-	46 (15)	1 (0)	-	-	-	-	-	-
Rye	832 (95)	647 (76)	-	-	832 (95)	647 (76)	96 (1)	90 (5)	-	-	-	-	-	-
Wheat	723 (110)	563 (45)	-	-	723 (110)	563 (45)	92 (2)	87 (4)	-	-	-	-	-	-
Ryegrass	447 (67)	403 (60)	-	-	447 (67)	403 (60)	75 (4)	76 (6)	-	-	-	-	-	-
WV biculture	502 (41)	504 (30)	226 (32)	80 (9)	276 (29)	424 (25)	86 (4)	89 (2)	-	-	-	-	-	-
Legume mix	249 (68)	321 (44)	249 (68)	321 (44)	-	-	75 (9)	82 (5)	-	-	-	-	-	-
Grass mix	573 (177)	577 (83)	-	-	573 (177)	577 (83)	94 (1)	94 (2)	-	-	-	-	-	-
1Legume3Grasses	704 (133)	751 (113)	217 (67)	101 (18)	487 (105)	650 (119)	94 (2)	97 (1)	-	-	-	-	-	-
3Legumes1Grass	490 (81)	536 (72)	145 (45)	130 (33)	346 (51)	407 (70)	90 (4)	92 (3)	-	-	-	-	-	-
All species	783 (160)	577 (104)	137 (35)	99 (25)	646 (134)	477 (111)	94 (2)	95 (2)	-	-	-	-	-	-
Grass biomass (g/m²)														
Vetch	289 (41)	240 (16)	289 (41)	240 (16)	-	-	84 (3)	83 (5)	-	-	-	-	-	-
Pea	121 (72)	320 (51)	121 (72)	320 (51)	-	-	66 (8)	76 (9)	-	-	-	-	-	-
Clover	137 (48)	79 (3)	137 (48)	79 (3)	-	-	46 (15)	1 (0)	-	-	-	-	-	-
Rye	832 (95)	647 (76)	-	-	832 (95)	647 (76)	96 (1)	90 (5)	-	-	-	-	-	-
Wheat	723 (110)	563 (45)	-	-	723 (110)	563 (45)	92 (2)	87 (4)	-	-	-	-	-	-
Ryegrass	447 (67)	403 (60)	-	-	447 (67)	403 (60)	75 (4)	76 (6)	-	-	-	-	-	-
WV biculture	502 (41)	504 (30)	226 (32)	80 (9)	276 (29)	424 (25)	86 (4)	89 (2)	-	-	-	-	-	-
Legume mix	249 (68)	321 (44)	249 (68)	321 (44)	-	-	75 (9)	82 (5)	-	-	-	-	-	-
Grass mix	573 (177)	577 (83)	-	-	573 (177)	577 (83)	94 (1)	94 (2)	-	-	-	-	-	-
1Legume3Grasses	704 (133)	751 (113)	217 (67)	101 (18)	487 (105)	650 (119)	94 (2)	97 (1)	-	-	-	-	-	-
3Legumes1Grass	490 (81)	536 (72)	145 (45)	130 (33)	346 (51)	407 (70)	90 (4)	92 (3)	-	-	-	-	-	-
All species	783 (160)	577 (104)	137 (35)	99 (25)	646 (134)	477 (111)	94 (2)	95 (2)	-	-	-	-	-	-
Percent weed control (%)														
Vetch	289 (41)	240 (16)	289 (41)	240 (16)	-	-	84 (3)	83 (5)	-	-	-	-	-	-
Pea	121 (72)	320 (51)	121 (72)	320 (51)	-	-	66 (8)	76 (9)	-	-	-	-	-	-
Clover	137 (48)	79 (3)	137 (48)	79 (3)	-	-	46 (15)	1 (0)	-	-	-	-	-	-
Rye	832 (95)	647 (76)	-	-	832 (95)	647 (76)	96 (1)	90 (5)	-	-	-	-	-	-
Wheat	723 (110)	563 (45)	-	-	723 (110)	563 (45)	92 (2)	87 (4)	-	-	-	-	-	-
Ryegrass	447 (67)	403 (60)	-	-	447 (67)	403 (60)	75 (4)	76 (6)	-	-	-	-	-	-
WV biculture	502 (41)	504 (30)	226 (32)	80 (9)	276 (29)	424 (25)	86 (4)	89 (2)	-	-	-	-	-	-
Legume mix	249 (68)	321 (44)	249 (68)	321 (44)	-	-	75 (9)	82 (5)	-	-	-	-	-	-
Grass mix	573 (177)	577 (83)	-	-	573 (177)	577 (83)	94 (1)	94 (2)	-	-	-	-	-	-
1Legume3Grasses	704 (133)	751 (113)	217 (67)	101 (18)	487 (105)	650 (119)	94 (2)	97 (1)	-	-	-	-	-	-
3Legumes1Grass	490 (81)	536 (72)	145 (45)	130 (33)	346 (51)	407 (70)	90 (4)	92 (3)	-	-	-	-	-	-
All species	783 (160)	577 (104)	137 (35)	99 (25)	646 (134)	477 (111)	94 (2)	95 (2)	-	-	-	-	-	-
Soil N (g/m²)														
Vetch	289 (41)	240 (16)	289 (41)	240 (16)	-	-	84 (3)	83 (5)	-	-	-	-	-	-
Pea	121 (72)	320 (51)	121 (72)	320 (51)	-	-	66 (8)	76 (9)	-	-	-	-	-	-
Clover	137 (48)	79 (3)	137 (48)	79 (3)	-	-	46 (15)	1 (0)	-	-	-	-	-	-
Rye	832 (95)	647 (76)	-	-	832 (95)	647 (76)	96 (1)	90 (5)	-	-	-	-	-	-
Wheat	723 (110)	563 (45)	-	-	723 (110)	563 (45)	92 (2)	87 (4)	-	-	-	-	-	-
Ryegrass	447 (67)	403 (60)	-	-	447 (67)	403 (60)	75 (4)	76 (6)	-	-	-	-	-	-
WV biculture	502 (41)	504 (30)	226 (32)	80 (9)	276 (29)	424 (25)	86 (4)	89 (2)	-	-	-	-	-	-
Legume mix	249 (68)	321 (44)	249 (68)	321 (44)	-	-	75 (9)	82 (5)	-	-	-	-	-	-
Grass mix	573 (177)	577 (83)	-	-	573 (177)	577 (83)	94 (1)	94 (2)	-	-	-	-	-	-
1Legume3Grasses	704 (133)	751 (113)	217 (67)	101 (18)	487 (105)	650 (119)	94 (2)	97 (1)	-	-	-	-	-	-
3Legumes1Grass	490 (81)	536 (72)	145 (45)	130 (33)	346 (51)	407 (70)	90 (4)	92 (3)	-	-	-	-	-	-
All species	783 (160)	577 (104)	137 (35)	99 (25)	646 (134)	477 (111)	94 (2)	95 (2)	-	-	-	-	-	-
Fixed N (g/m²)														
Vetch	289 (41)	240 (16)	289 (41)	240 (16)	-	-	84 (3)	83 (5)	-	-	-	-	-	-
Pea	121 (72)	320 (51)	121 (72)	320 (51)	-	-	66 (8)	76 (9)	-	-	-	-	-	-
Clover	137 (48)	79 (3)	137 (48)	79 (3)	-	-	46 (15)	1 (0)	-	-	-	-	-	-
Rye	832 (95)	647 (76)	-	-	832 (95)	647 (76)	96 (1)	90 (5)	-	-	-	-	-	-
Wheat	723 (110)	563 (45)	-	-	723 (110)	563 (45)	92 (2)	87 (4)	-	-	-	-	-	-
Ryegrass	447 (67)	403 (60)	-	-	447 (67)	403 (60)	75 (4)	76 (6)	-	-	-	-	-	-
WV biculture	502 (41)	504 (30)	226 (32)	80 (9)	276 (29)	424 (25)	86 (4)	89 (2)	-	-	-	-	-	-
Legume mix	249 (68)	321 (44)	249 (68)	321 (44)	-	-	75 (9)	82 (5)	-	-	-	-	-	-
Grass mix	573 (177)	577 (83)	-	-	573 (177)	577 (83)	94 (1)	94 (2)	-	-	-	-	-	-
1Legume3Grasses	704 (133)	751 (113)	217 (67)	101 (18)	487 (105)	650 (119)	94 (2)	97 (1)	-	-	-	-	-	-
3Legumes1Grass	490 (81)	536 (72)	145 (45)	130 (33)	346 (51)	407 (70)	90 (4)	92 (3)	-	-	-	-	-	-

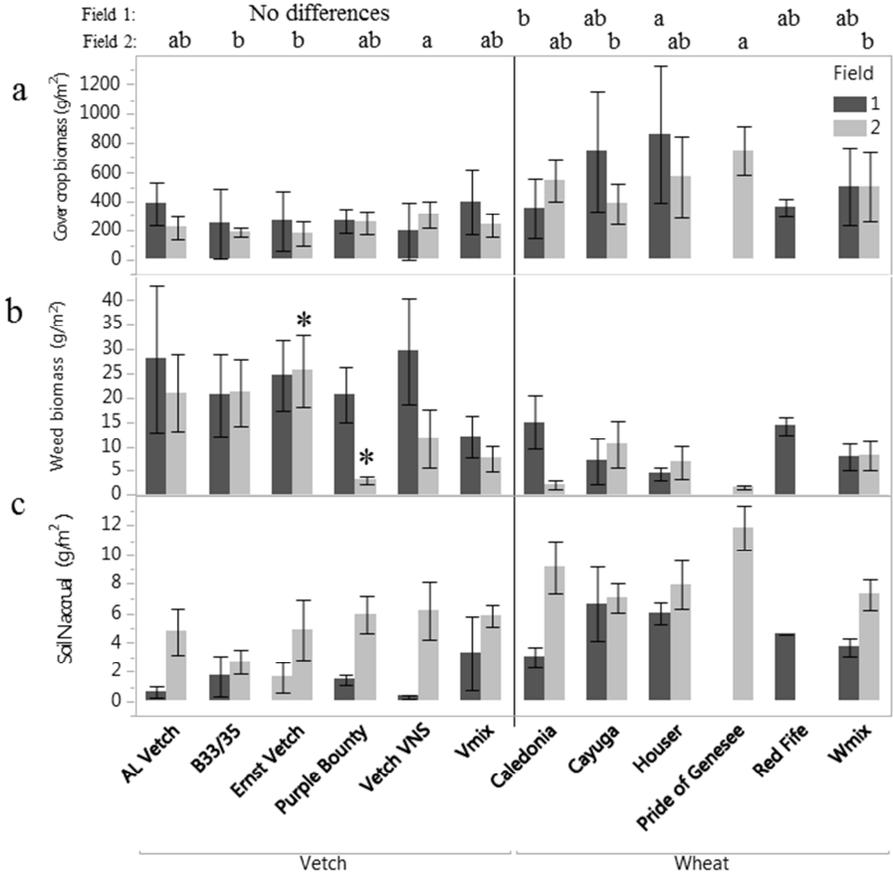


Figure 2S Average a) aboveground cover crop biomass b) weed biomass c) soil N accrual of vetch and wheat cultivars and cultivar mixtures by field (dark grey= Field 1, light grey=Field 2); pooled fertilizer treatments except for c. In Field 2 'Red Fife' was replaced with 'Pride of Genesee'. Different letters above cover crop biomass represent significant differences between cultivars in that field ($p < 0.05$). * indicates a significant difference between weed biomass for 'Ernst Vetch' and 'Purple Bounty' in Field 2, with no other significant differences for weed biomass in either field or species. Analysis on transformed data (Square root for Cover crop biomass and soil N accrual; ln for weed biomass), but untransformed data presented. Where no letters or symbols are present, no significant differences exist. Error bars represent one standard error from the mean.

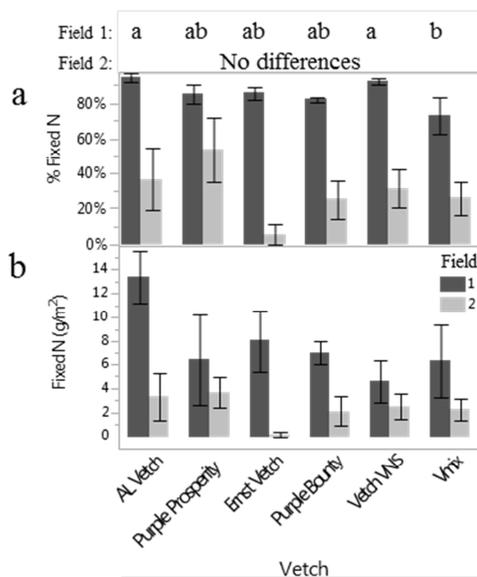


Figure 3S Average of a) percent of nitrogen in above ground biomass from fixation and b) total grams of fixed N in aboveground biomass of vetch cultivars and cultivar mixture by field, data from unfertilized plots. Different letters above % Fixed N represent significant differences between cultivars in that field ($p < 0.05$, Tukey HSD analysis, analysis on transformed data (Square root), but untransformed data presented). Where no letters are present, no significant differences exist. Error bars represent one standard error from the mean.

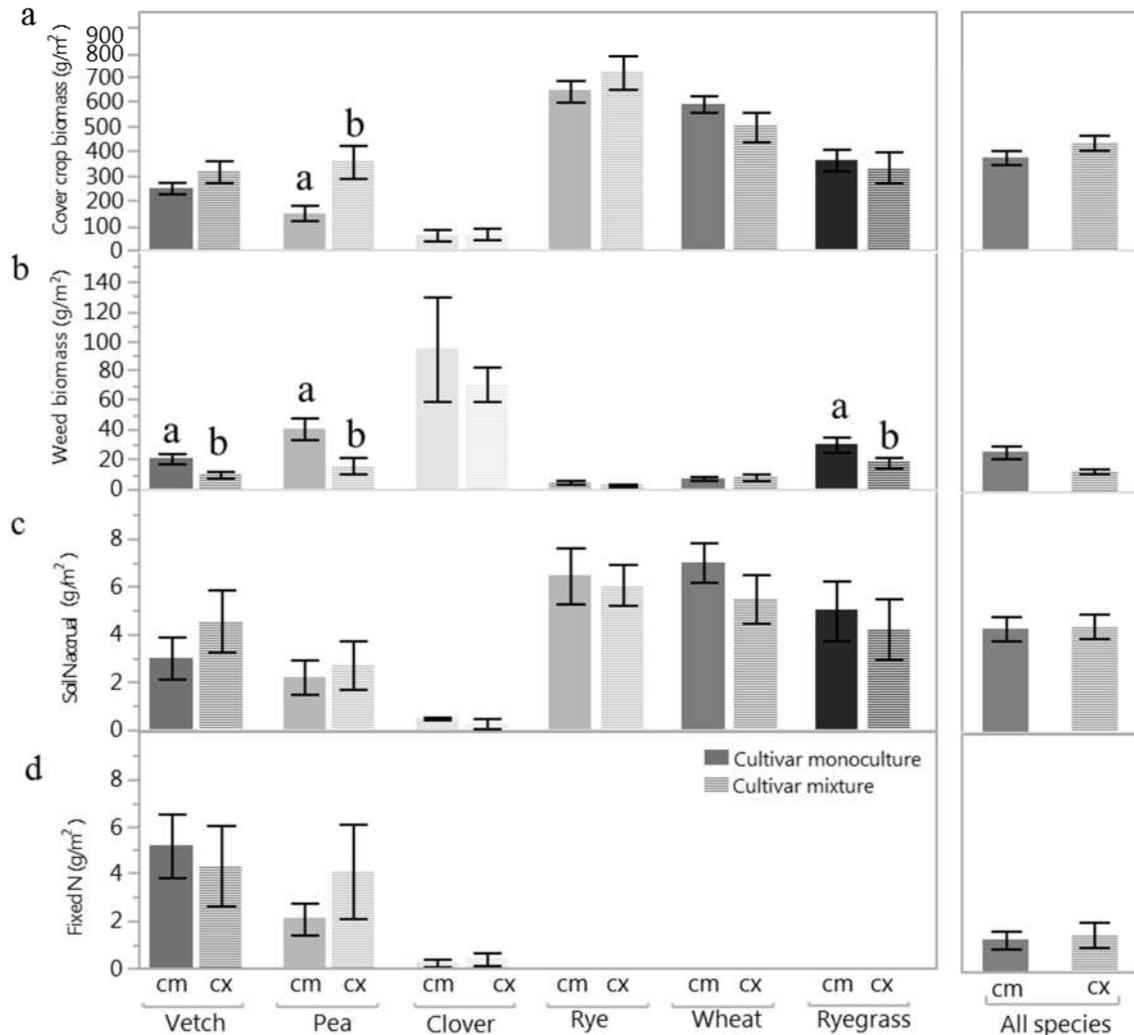


Figure 4S Average a) aboveground cover crop biomass b) weed biomass c) soil N accretion and d) fixed nitrogen of cultivar monocultures (solid bars) and cultivar mixtures (shaded bars) for six species. Different letters within a species indicate significant differences between intraspecific diversity levels ($p < 0.05$, Tukey HSD analysis on transformed data (\ln for weed biomass; Square root for soil N accretion and fixed N), but untransformed data presented). Both fertilizer treatments are pooled for a and b with no interactions, while data from unfertilized plots is presented for c and d. No interaction with field, so data from both fields are pooled. Error bars represent one standard error from the mean.

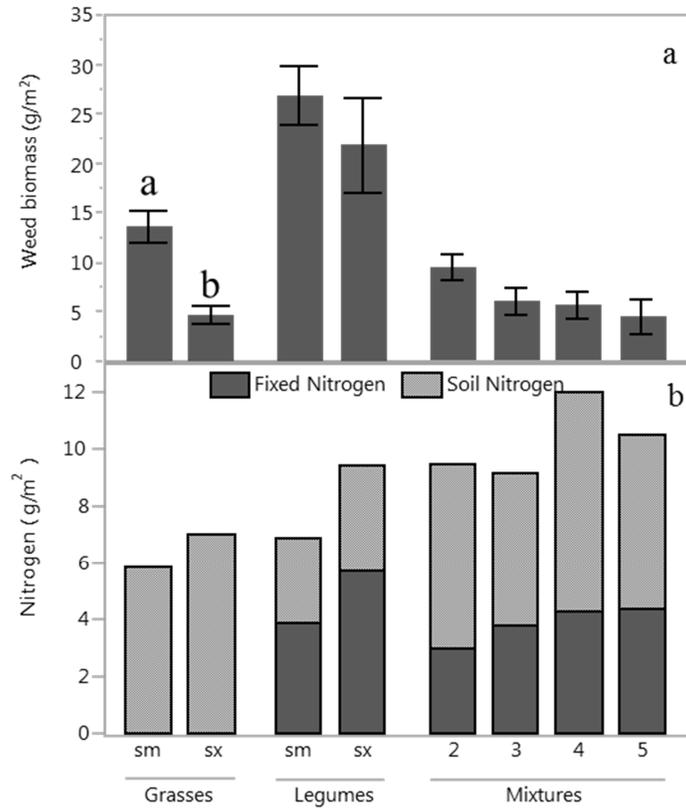


Figure 5S Average effect of interspecific diversity on a) weed biomass and b) soil and fixed nitrogen in aboveground cover crop biomass. Comparison of species monocultures (sm) and species mixtures (sx) for single functional group treatments (grasses and legumes) and of species richness levels (2-5) for grass/legume mixtures (Mean +/- SE). Both fields and fertilizer treatments are pooled and presented. Due to methodology, only unfertilized treatments are included in b. Different letters within functional group categories indicate significant differences between interspecific diversity levels ($p < 0.05$, Tukey HSD analysis on transformed data (Square root for Soil N, and Fixed N; ln for weed biomass), but untransformed data presented. See Table 8S for more details.

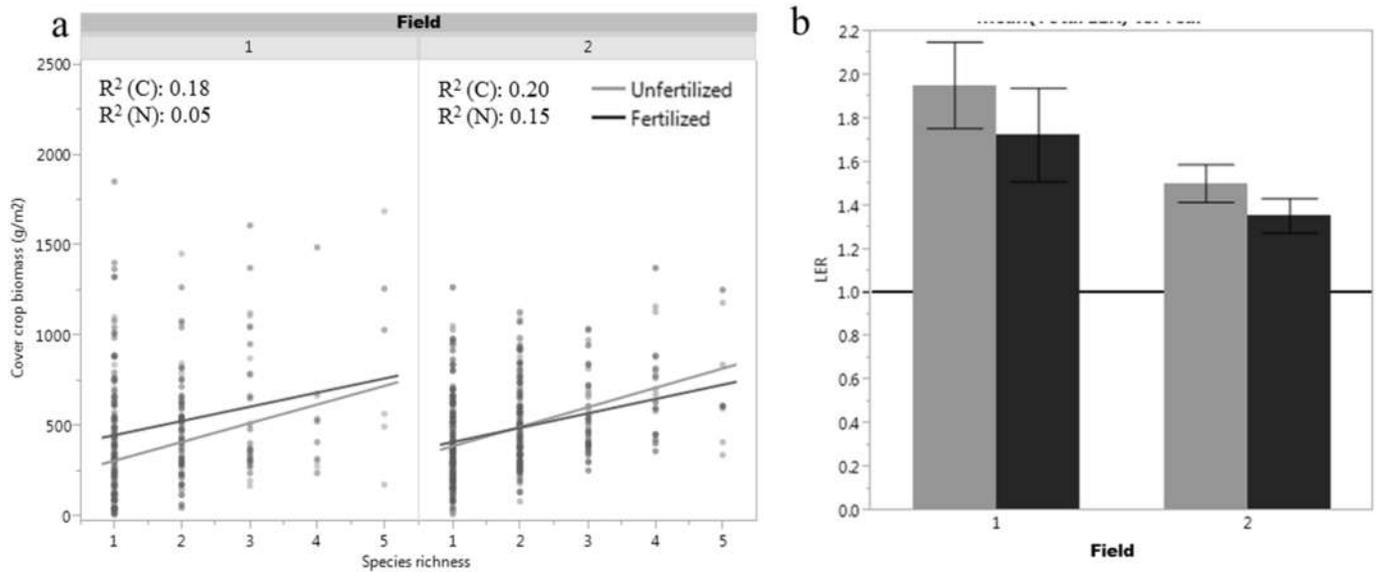


Figure 6S Effect of fertilizer application on a) relationship between species richness and cover crop biomass by field and fertilizer treatment. R² from model of square root transformed data, but untransformed data presented. Fertilizer treatments within fields are not significantly different (Field 1: p=0.4652 ,Field 2: p=0.3158); b) Total average LER by field and fertilizer treatment. Fertilizer treatments within fields are not significantly different (Analysis on log transformed data, but untransformed data presented. Field 1: p=0.1825; Field 2: p=0.2873.) Error bars represent one standard error from the mean.

Table 5S a) Vetch and pea biomass, amount of nitrogen from fixation (g Ndfa), and rate of SNF (%Ndfa) by plot composition and Field 1 and 2, b) model results for effect of increased species richness on %Ndfa and gNdfa for vetch and pea. The interaction between species richness and field was not significant for %Ndfa, and was not included in the final model

a	N	Vetch biomass (g/m ²)	Pea biomass (g/m ²)	Vetch g Ndfa	Pea g Ndfa	Vetch %Ndfa	Pea %Ndfa
Legumes alone, species monocultures and mixtures, both fields	187	239	176	7.2	5.1	55.8%	52.3%
Legume species monocultures, both Fields	155	260	234	7.8	5.8	55.0%	53.5%
<i>Field 1</i>	70	293	128	8.2	1.0	86.8%	77.3%
<i>Field 2</i>	85	232	304	7.5	8.5	30.0%	39.7%
Cultivar monocultures	114	250	170	7.7	4.8	59.3%	50.1%
<i>Field 1</i>	55	271	45	8.1	1.1	88.9%	74.7%
<i>Field 2</i>	59	230	254	7.3	7.2	31.2%	34.7%
Cultivar mixtures	41	292	358	8.3	7.8	39.6%	61.1%
<i>Field 1</i>	15	397	293	8.6	0.7	73.8%	83.9%
<i>Field 2</i>	26	239	399	8.2	11.3	26.8%	49.7%
Legume species mixtures, both Fields	32	170	107	5.2	4.0	58.2%	50.1%
<i>Field 1</i>	16	213	18	6.0	0.8	80.2%	80.5%
<i>Field 2</i>	16	127	196	4.3	5.2	36.2%	38.7%
Grass/legume mixtures, all species richness levels, both Fields	267	122	52	4.1	1.5	64.6%	66.4%
2 species, both Fields	181	126	46	4.1	1.5	59.7%	48.7%
<i>Field 1</i>	69	196	-	6.6	-	94.3%	-
<i>Field 2</i>	112	75	65	2.2	1.5	35.4%	48.7%
3 species, both Fields	40	117	45	3.9	1.5	69.2%	65.2%
<i>Field 1</i>	21	178	28	6.1	0.8	93.7%	91.6%
<i>Field 2</i>	19	49	55	1.4	1.7	39.2%	58.6%
4 species, both Fields	30	109	78	4.4	2.6	72.8%	77.4%
<i>Field 1</i>	10	159	61	7.5	2.5	95.9%	87.8%
<i>Field 2</i>	20	84	93	2.9	2.6	61.3%	67.0%
5 species, both Fields	14	107	48	4.0	0.9	83.3%	77.0%
<i>Field 1</i>	6	195	23	7.1	0.2	94.6%	86.9%
<i>Field 2</i>	8	41	66	1.5	1.5	74.3%	67.2%

b	species richness	field	species richness : field	Transformation
%Ndfa Vetch	$F_{(1,162)}=6.2439$, $p=0.0135$	$F_{(1,2.657)}=967.3284$, $p=0.0002$		square root
%Ndfa Pea	$F_{(1,51.99)}=2.202$, $p=0.1439$	$F_{(1,11.22)}=25.1454$, $p=0.0004$		square root
gNdfa Vetch	$F_{(1,177)}=45.5627$, $p=<.0001$	$F_{(1,6.357)}=2.4572$, $p=0.1653$	$F_{(1,177)}=13.6325$, $p=0.0003$	square root
gNdfa Pea	$F_{(1,54.36)}=10.8512$, $p=0.0017$	$F_{(1,36.76)}=59.8008$, $p=<.0001$	$F_{(1,54.36)}=8.9773$, $p=0.0041$	square root

Table 6S Model results for species differences in monoculture by ecosystem service metric (see Figure 1 for more details)

	Field 1	Field 2	field:species	Transformation
Cover crop biomass	$F_{(5,157.7)}=23.9186, p<.0001$	$F_{(5,170.3)}=39.7987, p<.0001$	$F_{(5,315.8)}=5.4485, p<.0001$	square root
Weed biomass	$F_{(5,157.5)}=22.4607, p<.0001$	$F_{(5,172.8)}=10.2319, p<.0001$	$F_{(5,318.4)}=2.1547, p=0.0589$	ln
Soil N accrual	$F_{(5,74.64)}=17.0838, p<.0001$	$F_{(5,81.3)}=19.5527, p<.0001$	$F_{(5,155.9)}=3.9174, p=0.0023$	square root
Fixed N	$F_{(2,30.7)}=16.8074, p<.0001$	$F_{(2,38.17)}=2.7739, p=0.075$	$F_{(2,70.06)}=9.0569, p=0.0003$	square root

Table 7S Model results for cultivar differences in vetch and wheat (see Figure 2S and 3S for more details). No model results for a given factor indicate that factor was no significant and was not included in the final model.

Cover crop biomass	cultivar identity	fert	cultivar:fert	Transformation
Vetch, field 1	$F_{(5,38.03)}=2.7410, p=0.0328$			square root
Vetch, field 2	$F_{(5,47)}=3.3324, p=0.0117$			square root
Wheat, field 1	$F_{(4,24.18)}=3.8337, p=0.0150$	$F_{(1,23.07)}=6.4464, p=0.0183$		square root
Wheat, field 2	$F_{(4,40)}=3.7734, p=0.0107$			square root
Weed biomass	cultivar identity	fert	cultivar:fert	Transformation
Vetch, field 1	$F_{(5,5)}=38.03, p=0.3667$			$\ln(\text{weed biomass}+1)$
Vetch, field 2	$F_{(5,46)}=3.0462, p=0.0186$	$F_{(1,46)}=4.9068, p=0.0317$		$\ln(\text{weed biomass}+1)$
Wheat, field 1	$F_{(4,23.68)}=1.8259, p=0.1572$	$F_{(1,22.89)}=4.3242, p=0.0489$		$\ln(\text{weed biomass}+1)$
Wheat, field 2	$F_{(4,40)}=2.2383, p=0.0820$			$\ln(\text{weed biomass}+1)$
Soil N accrual	cultivar identity	fert		Transformation
Vetch, field 1	$F_{(5,14.1)}=1.2450, p=0.34$			square root
Vetch, field 2	$F_{(5,19)}=1.4379, p=0.2561$	*fertilizer not applicable for N metrics due to methodology		square root
Wheat, field 1	$F_{(4,6.722)}=1.2141, p=0.3875$			square root
Wheat, field 2	$F_{(4,16)}=1.9742, p=0.1472$			square root
Fixed N	cultivar identity	fert		Transformation
Vetch, field 1	$F_{(5,14.13)}=2.1695, p=0.1160$	*fertilizer not applicable for N metrics due to methodology		square root
Vetch, field 2	$F_{(5,19)}=1.0689, p=0.4082$			square root
%N fixed	cultivar identity	fert		Transformation
Vetch, field 1	$F_{(5,13.38)}=3.632, p=0.0272$	*fertilizer not applicable for N metrics due to methodology		none
Vetch, field 2	$F_{(5,19)}=1.4853, p=0.241$			none

Table 8S Model results for effect of species richness by functional group and mixtures (see Figure 2 and 5S for more details). No model results for a given factor indicate that factor was no significant and was not included in the final model.

Cover crop biomass	species richness	field	fertilizer	species richness*fert	fert*field
Grasses	$F_{(1,214,1)}=2,6617, p=0,1043$			$F_{(1,214,1)}=5,2833, p=0,0225$	
Legume	$F_{(1,177,9)}=2,8518, p=0,0930$				
Mixes	$F_{(3,253,1)}=6,145, p=0,0005$				$F_{(1,252)}=4,2585, p=0,0401$
Weed biomass	species richness	field	fertilizer	species richness*fert	fert*field
Grasses	$F_{(1,214,3)}=5,6267, p=0,0186$	$F_{(1,6,134)}=16,6042, p=0,0062$	$F_{(1,214,2)}=11,5042, p=0,0008$		
Legume	$F_{(1,178,1)}=0,0657, p=0,7980$				
Mixes	$F_{(3,253,9)}=1,9838, p=0,1169$	$F_{(1,5,651)}=14,0976, p=0,0099$	$F_{(1,251,9)}=13,2157, p=0,0003$		$F_{(1,251,8)}=8,2541, p=0,0044$
Soil N accrual	species richness	field			
Grasses	$F_{(1,109)}=2,8093, p=0,0968$	$F_{(1,6)}=81,3892, p=0,0001$			
Legume	$F_{(1,86,12)}=3,4378, p=0,0671$	$F_{(1,5,756)}=60,7947, p=0,0003$	*fertilizer not applicable for N metrics due to methodology		
Mixes	$F_{(3,123,3)}=0,4022, p=0,7517$	$F_{(1,6,121)}=60,5761, p=0,0002$			
Fixed N	species richness	field			
Legumes	$F_{(1,86,59)}=3,2149, p=0,0765$	$F_{(1,6,36)}=7,9843, p=0,0283$	*fertilizer not applicable for N metrics due to methodology		
Mixes	$F_{(3,123,5)}=1,4615, p=0,2284$	$F_{(1,5,579)}=26,2067, p=0,0027$			
Cover crop biomass	Transformation	Unfertilized	Fertilized		
Grasses	square root	t-test: $F_{(1,214)}=7,5324, p=0,0066$	t-test: $F_{(1,214)}=0,2283, p=0,6333$		
Legume	square root				
Mixes	square root				
Weed biomass	Transformation				
Grasses	ln				
Legume	ln				
Mixes	ln				
Soil N accrual	Transformation				
Grasses	square root				
Legume	square root				
Mixes	square root				
Fixed N	Transformation				
Legumes	square root				
Mixes	square root				

Table 9S Model results for intraspecific diversity in species monocultures (see Figure 4S for more details). No model results for a given factor indicate that factor was not significant and was not included in the final model.

Cover crop biomass	Intraspecific diversity (ID)	field	fertilizer	ID:field	ID:fert	field:fert	ID:field:fert	Transformation
Vetch	$F_{1,1,11} = 3.66, p = 0.0682$							none
Pea	$F_{1,1,20,21} = 12.7151, p = 0.0018$							none
Clover	Inadequate number of data points for analysis							
Rye	$F_{1,1,12} = 0.8956, p = 0.3542$		$F_{1,1,12} = 4.69, p = 0.0414$					none
Wheat	$F_{1,1,12} = 4.1118, p = 0.0555$	$F_{1,1,6} = 0.0908, p = 0.0908$	$F_{1,1,12} = 12.7704, p = 0.0018$			$F_{1,1,12,13} = 5.6468, p = 0.0271$	none	
Ryegrass	$F_{1,1,11} = 0.3496, p = 0.5606$	$F_{1,1,6} = 0.4116, p = 0.5449$	$F_{1,1,12} = 9.8728, p = 0.0049$			$F_{1,1,12,13} = 7.9459, p = 0.0103$	none	
All species	$F_{1,1,10,13} = 1.8212, p = 0.1791$							none
Weed biomass	Intraspecific diversity (ID)	field	fertilizer	ID:field	ID:fert		ID:field:fert	Transformation
Vetch	$F_{1,1,11} = 23.4680, p < 0.0001$							$\ln(\text{weed biomass}+1)$
Pea	$F_{1,1,12,13} = 6.3146, p = 0.0196$	$F_{1,1,6,10,11} = 13.9195, p = 0.0095$						$\ln(\text{weed biomass}+1)$
Clover	Inadequate number of data points for analysis							
Rye	$F_{1,1,12} = 1.4353, p = 0.2437$		$F_{1,1,12} = 4.7423, p = 0.0402$					$\ln(\text{weed biomass}+1)$
Wheat	$F_{1,1,12} = 0.0359, p = 0.8515$		$F_{1,1,12} = 4.4829, p = 0.0458$					$\ln(\text{weed biomass}+1)$
Ryegrass	$F_{1,1,12} = 7.4159, p = 0.0121$	$F_{1,1,6} = 6.9983, p = 0.0383$						$\ln(\text{weed biomass}+1)$
All species	$F_{1,1,10,13} = 11.4772, p = 0.0009$	$F_{1,1,6,10,11} = 6.5370, p = 0.0429$						$\ln(\text{weed biomass}+1)$
Soil N accrual	Intraspecific diversity (ID)	field	fertilizer	ID:field				Transformation
Vetch	$F_{1,1,7} = 1.4101, p = 0.2738$							square root
Pea	$F_{1,1,7,8,9} = 0.058, p = 0.8049$	$F_{1,1,6,11} = 65.6925, p = 0.0002$						square root
Clover	Inadequate number of data points for analysis		*Fertilizer not applicable for N metrics due to methodology					square root
Rye	$F_{1,1,7} = 0.1249, p = 0.7342$	$F_{1,1,6} = 63.2491, p = 0.0002$						square root
Wheat	$F_{1,1,7} = 4.0759, p = 0.0833$	$F_{1,1,6} = 16.5057, p = 0.0066$						square root
Ryegrass	$F_{1,1,7} = 1.7463, p = 0.2279$	$F_{1,1,6} = 15.8551, p = 0.0073$						square root
All species	$F_{1,1,7,8,9} = 0.0068, p = 0.9347$							none
Fixed N	Intraspecific diversity (ID)	field	fertilizer	ID:field				Transformation
Vetch	$F_{1,1,7} = 0.9646, p = 0.3588$	ns not included						square root
Pea	$F_{1,1,7,8,9} = 0.2383, p = 0.6398$	$F_{1,1,5,8,9,11} = 12.2526, p = 0.0134$	*Fertilizer not applicable for N metrics due to methodology					square root
Clover	Inadequate number of data points for analysis							
All species	$F_{1,1,7,8,9,11} = 0.1753, p = 0.6766$							none

Table 10 S Results from t-test for deviation of total and partial LER from reference values of 1 and 0.5 respectively. All data were log transformed, and transformed results presented. See Figure 5 for more details

Total LER

Field 1	mean	df	t statistic	p abs t	p > t
1L3G	0.46957	15	2.547	0.0224	0.0112
3L1G	0.26866	14	1.1976	0.251	0.1255
All species	0.29145	13	1.0491	0.3133	0.1566
WV biculture	0.38639	52	3.9723	0.0002	0.0001
Grass mix	-0.1835	15	-1.0624	0.3048	0.8476
Legume mix	0.18361	15	0.5332	0.6017	0.3009
Field 2	mean	df	t statistic	p abs t	p > t
1L3G	0.46243	15	5.6819	<.0001	<.0001
3L1G	0.38777	15	3.3844	0.0041	0.002
All species	0.22762	15	2.8712	0.0117	0.0058
WV biculture	0.21937	87	3.8024	0.0003	0.0001
Grass mix	0.23706	15	1.7427	0.1018	0.0509
Legume mix	0.02256	15	0.1332	0.8958	0.4479

Partial LER for legumes

Field 1	mean	df	t statistic	p abs t	p > t
1L3G	-0.4469	15	0.7165	0.4847	0.2423
3L1G	-0.5426	14	0.3695	0.7173	0.3587
All species	-0.901	13	-0.6275	0.5412	0.7294
WV biculture	-0.2206	52	3.0771	0.0033	0.0017
Field 2	mean	df	t statistic	p abs t	p > t
1L3G	-1.0762	15	-2.6561	0.018	0.991
3L1G	-1.0968	15	-2.3172	0.035	0.9825
All species	-1.3246	15	-2.8678	0.0117	0.9941
WV biculture	-1.2338	86	-6.2071	<.0001	1

Partial LER for grasses

Field 1	mean	df	t statistic	p abs t	p > t
1L3G	-0.4059	15	1.9488	0.0703	0.0351
3L1G	-0.5426	14	-1.3443	0.2037	0.8981
All species	-0.1589	13	1.9552	0.0724	0.0362
WV biculture	-0.7345	45	-0.3666	0.7156	0.6422
Field 1	mean	df	t statistic	p abs t	p > t
1L3G	0.14565	15	7.0998	<.0001	<.0001
3L1G	0.02109	15	4.7662	0.0002	0.0001
All species	-0.1161	15	5.3127	<.0001	<.0001
WV biculture	-0.1489	87	8.3867	<.0001	<.0001

Appendix

Nitrification rate (mg N/kg soil/day)

-N/+N	Treatment	Mean (SE)	Standard Deviation
Unfertilized (-N)	AL Vetch monoculture	11.259 (2.38)	4.779
	Houser wheat monoculture	11.786 (1.12)	2.248
	All vetch cultivars	12.014 (1.73)	3.479
	All wheat cultivars	13.192 (2.74)	5.496
	AL Vetch: Houser	11.105 (1.60)	3.218
	All cultivars of vetch : all cultivars of wheat	12.372 (1.15)	2.319
	Weed control	13.587 (1.12)	2.248
Fertilized (+N)	AL Vetch monoculture	13.714 (1.31)	2.621
	Houser wheat monoculture	12.816 (1.40)	2.817
	AL Vetch: Houser	13.080 (0.78)	1.569
	Weed control	14.916 (2.55)	5.119

CHAPTER 3

CONTEXT DEPENDENT DELIVERY OF ECOSYSTEM SERVICES BY COVER CROPS

Emily R. Reiss and Laurie E. Drinkwater

Abstract

While cover crops and their mixtures are increasingly used to provide ecosystem services in agroecosystems, some fundamental questions remain about the amount of variability in cover crop performance under different conditions. These knowledge gaps limit the ability to make broader conclusions and recommendations to growers about optimal diverse cover crop management for ecosystem services. In the overwintering season of 2014-2015, we conducted a field experiment at a university research farm, with an embedded experiment which included a subset of treatments in three working farm fields. We selected two cover crop species commonly used by farmers in the region, hairy vetch (*Vicia villosa*), a legume, and common wheat (*Triticum aestivum*) a grass, and multiple cultivars of each. We looked at the effect of cover crop composition spanning intraspecific and grass/legume mixtures on five ecosystem services: cover crop productivity, weed suppression, total biomass nitrogen (N), N retention, and long-term N supply via legume fixed N. Specifically, we were interested in evaluating if the effect of cover crop composition on these services was context dependent. We did not find intraspecific diversity to have an effect on any ecosystem services we measured, nor was that response context dependent. This may have been due to the lack of trait and phenotypic diversity in the cultivars we selected. We did observe significant ecosystem service improvements in the grass/legume mixture, though this was context dependent and the performance of the mixture varied relative to the monocultures at different farm sites. The response of cover crop biomass and soil N accrual were the most context dependent. Regardless of this interaction however, the

grass/legume mixture was as good as or better than either monoculture for all services and sites with one exception (soil N accrual at one site). Consequently, as a practical recommendation, increasing complexity in cover crops, through either intraspecific or grass/legume mixtures, is a low risk practice that has the potential to deliver improved ecosystem service outcomes regardless of the growing context.

Introduction

Cover crops are increasingly used by a wide range of farmers to support various ecosystem services, from erosion control to pest regulation, with nutrient management and soil health as high priorities (Schipanski et al., 2014; Dunn et al., 2016; Wayman et al., 2016). Cover crops can be any plant species established when ground is usually fallow between cash crops, with the goal of promoting certain ecosystem services to support cash crop yields or reducing externalities (Ritter et al., 1998; Doltra and Olesen, 2013). Research on the benefits and management of cover crops has a long history, and has also been increasing rapidly in recent years along with grower adoption (Clark et al., 1994; Creamer et al., 1997; Schipanski et al., 2014). Given the wide range of possible cover crop practices and on-farm context of soil, climate and management conditions, providing generalized guidelines and consistent delivery of desired ecosystem services is a major management challenge (Myers and Watts, 2015).

Mixtures of cover crops, specifically those with grasses and legumes, are commonly used by farmers because of the multiple benefits they can provide (Snapp et al., 2005; Valkama et al., 2015; Myers and Watts, 2015; Appelgate et al., 2017). When appropriately implemented these mixtures can deliver the strengths of each functional group simultaneously, minimizing the tradeoffs of each group growing alone (Brainard et al., 2011; Schipanski et al., 2014; White et al., 2017). Because of these benefits, grass/legume mixtures, often referred to as intercrops, are

widely used in agricultural systems from cover crops to pastures, and forage and grain production (Osman and Diek, 1982; Vandermeer, 1989; Nyfeler et al., 2009; Baba et al., 2011; Bedoussac et al., 2015). In symbiosis with root rhizobia, legumes fix nitrogen (N) contributing to long term N supply, and lowering the carbon to nitrogen (C/N) ratio for fast residue decomposition and nutrient availability for subsequent crops (Ruffo and Bollero, 2003; Brainard et al., 2011). Conversely, grasses tend to take up soil N more efficiently, and have greater aboveground biomass, which results in more organic matter accumulation, and better weed suppression than legumes (Sainju et al., 2005; Kruidhof et al., 2008).

While grass/legume mixtures do generally deliver these multiple benefits and exhibit synergy when combined, the performance of these mixtures can be variable depending on environmental conditions and plant community composition (Ranells and Wagger, 1997; Sainju et al., 2005; Poffenbarger et al., 2015b). Critically, environmental and management conditions (e.g. manure applications) can each influence the dynamics of a grass/legume mixture ultimately affecting the ecosystem service outcomes (Schipanski and Drinkwater, 2012; Poffenbarger et al., 2015a). Given a certain context and growth conditions, the ecosystem service delivery of a species may be altered when in a mixture (Ranells and Wagger, 1997; Murrell et al., 2017). For example, the rate of symbiotic nitrogen fixation by hairy vetch was less in monoculture compared to when mixed with a grass, but this was not consistent across cultivars (Brainard et al., 2012).

Understanding the potential interaction between community composition and the on-farm context for growth is critical for successfully managing cover crop mixtures on a farm, as well as understanding how to transfer research results to a wider range of conditions. While frequently included in plant breeding research, evaluating the ‘treatment x environment’ interaction, or

impact of context, is not as common an objective in other fields (Helland and Holland, 2003; Sinebo, 2005). Background soil types and characteristics, weather, as well as management histories and current practices can vary across landscapes and farms (Asrat et al., 2010; Schipanski and Drinkwater, 2011; Drinkwater, 2016). All of these factors influence the growing environment and form the context for crop growth, including that of cover crop mixtures.

While grass/legume mixtures take advantage of different plant functional groups to balance potential tradeoffs, a similar moderating effect may be achieved by increasing the genetic and phenotypic diversity within a species. Mixing cultivars increases the intraspecific diversity within a monoculture of one species. The limited research on the effect of intraspecific diversity in agroecosystems suggests that there could be benefits for cover crops and their associated ecosystem services (Smithson and Lenne, 1996; Mundt, 2002; Kiær et al., 2009; Grettenberger and Tooker, 2015). For example, in cash crops, cultivar mixtures have been found to increase yields as well as reduce the spread of disease (Mundt, 2002; Reiss and Drinkwater, 2018). Cultivars have long been developed and used in food crops to enhance specific desirable traits, and to tailor a crop to regional conditions (Elzebroek, 2008; Fu, 2015). Given the wide range of cultivars available and the varying responses of these cultivars to different conditions, the identity of the cultivar used is likely a large driver of ultimate crop performance (Helland and Holland, 2001; Mengistu et al., 2010; Brainard et al., 2012; Ramirez-Garcia et al., 2015; Canali et al., 2017). However, this cultivar diversity is not commonly utilized for cover crops.

We were interested in addressing these outstanding questions about the site context impact on cover crop performance and ways to moderate this impact, while building practical management recommendations. To do this, we designed a nested set of experiments to evaluate the role of cultivars/populations and grass/legume mixtures in maintaining cover crop

performance across a range of contexts. We selected nine cultivars of two cover crops species, hairy vetch, a legume (*Vicia villosa*) and common wheat, a grass (*Triticum aestivum*). We tested this set of cover crops at a university research farm as part of a larger experiment addressing diversity in cover crop mixtures, and also established these treatments at three working farm sites. We looked at the effect of cover crop identity and composition in three ways on five ecosystem services: cover crop productivity, weed suppression, total biomass nitrogen (N), N retention, and long-term N supply using corresponding measurable metrics (Table 1). We first evaluated the effect of increased intraspecific diversity by comparing the performance of cultivar mixtures to the mean of cultivar monocultures on the above mentioned ecosystem services. Second, we tested for differences in the performance of the individual cultivars to each other and to the cultivar mixture. Finally, we evaluated the effect of mixing vetch and wheat compared to each alone. We tested for an interaction between farm and each of these three characterizations of cover crop composition on the five ecosystem services. Overall, we expected there to be interactions between farm site conditions and cover crop identity, but that the strength and specific impact would be dependent on the context and the services evaluated.

Methods

Plot establishment and management

We established a nested set of experimental trials at three organic, mixed vegetable farms (Farms 1-3) in the Finger Lakes region of New York State, as part of a larger trial at Cornell University Musgrave Research Farm (Farm 4) in Aurora, NY (42.73' N, 76.66' W). Details of the farm sites and management can be found in Table 2 and Table 1S. Additional details about the research farm site and experiment can be found in Reiss and Drinkwater 2018 (in prep.). Even though all the farm sites were within 50 miles of one another, some edaphic soil factors as well

Table 1 Measured variables to evaluate corresponding ecosystem service

Collected metric	Ecosystem service
Soil N accrual in cover crop biomass (N_s)	N retention
Shoot fixed N (N_f)	N supply, long-term
Total biomass N (N_T)	Potential N supply for the following crop
Weed biomass	Weed suppression
Cover crop biomass	C accrual

Table 2 Site conditions at the four farm sites. Mean (standard deviation, range for weed biomass). Different upper case letters for a given variable indicate significant differences among farms based on Tukey's HSD test ($p < 0.05$). Farms 1-3 are small-scale working farms, and Farm 4 is a university-managed research farm. *See note in text regarding uneven distribution of large *Brassica* weeds in sampling area

	Farm 1	Farm 2	Farm 3	Farm 4
Inorganic N (mg N/kg soil)	46.5 (3.0) A	13 (3.4) B	6.5 (1.1) B	11 (1.4) B
N mineralization potential (mg N/kg soil/week)	15.4 (2.9) B	17.6 (1.3) B	33.4 (1.8) A	8.8 (2.0) C
Soil Nitrogen %	0.197 (0.009) AB	0.164 (0.004) B	0.248 (0.036) A	0.187 (0.017) B
Soil Carbon %	1.76 (0.07) BC	1.48 (0.05) C	2.40 (0.40) A	2.05 (0.08) AB
C/N	8.9 (0.2) A	9.0 (0.1) A	9.7 (0.5) AB	11.0 (0.7) B
pH	6.9 (0) B	6.5 (0.1) C	5.8 (0.2) D	7.7 (0.1) A
P (mg/kg)	437 (48) A	50 (4) B	37 (8) B	17 (1) B
K (mg/kg)	196 (39) A	113 (19) B	70 (21) B	76 (8) B
Soil Textural Class	Loam	Clay Loam	Loam	Loam
GDD (Spring, base 0°C)	821	830	800	883
Background weed biomass (g/m²)	191 (25-356) *	29 (13-53)	158 (97-219)	46 (11-87)

as management histories differed, resulting in varied conditions across the sites. Each farm was situated on a different soil series, though they were all classified as silt loams. When actual soil samples from each farm were evaluated for texture all but Farm 2 (clay loam) were loam. The different background soil conditions along with varying management histories serve to create a diverse set of site conditions across these four farms.

Farm 1 was a high fertility site, particularly in terms of inorganic N, P and K likely due to long-term compost additions over the past two decades of organic vegetable production. However, other important soil characteristics such as total carbon, N mineralization potential and CEC were not notably higher than other sites. Background weed pressure was generally low, though there was substantial variability in sampled biomass across the field due to large brassica weeds (*Barbarea vulgaris*) in the sampling area of one cover crop control plot that were not evenly dispersed through the field (range 25-356 g/m² in control plots).

Farm 2 had been in a conventional corn and soybean rotation for over 20 years prior to conversion to hay and mixed vegetables five years prior to the experiment establishment. This may explain the low soil carbon and less soil nutrients overall compared to Farm 1. Weed biomass in control plots was comparable to typical levels at Farm 1 (13-63 g/m²).

Farm 3 had been hayed for 20 years prior to conversion to mixed vegetables two years before we established the experimental plots. This history is reflected in the high organic matter content specifically and in the moderate levels of soil nutrients (Gregory et al., 2016). The very high background weed pressure (97-219 g/m²) may also be due to this history. The low pH at this site (5.8) may inhibit legume growth such as vetch cover crops (Clark, 2007).

Farm 4, the research farm, is notable as the only conventionally managed site with no vegetable production. Additionally, this is a relatively high pH soil, with low N mineralization

potential and lower soil phosphorus levels. Together the lower phosphorus and higher pH may have resulted in reduced plant-available phosphorus, which can also reduce cover crop growth (Clark, 2007). The background weed pressure was similar to that at Farms 1 and 2 (11-87 g/m²).

We selected a subset of treatments from the larger research farm trial to establish at these farms (Table 3). This included two species, a legume and a grass, and multiple cultivars of each. For the legume, hairy vetch (*Vicia villosa*), we included five cultivars or lines (treatments 1-5), and four cultivars of the grass, common winter wheat (*Triticum aestivum*) (treatments 6-9). With these two species and their cultivars, we created mixtures of cultivars to examine the impact of intraspecific diversity (treatments 10 and 11) and mixtures of the two species to examine the impact of combining a grass and legume together (treatments 12-15). We selected the cultivars and lines based on available trait and morphology information such as flowering time for vetch, wheat height, and fall vigor. When constructing the grass/legume mixtures (treatments 12-15), we selected cultivars or lines with contrasting traits where possible. ‘Purple Bounty’ is an early flowering vetch bred in a moderate climate, whereas ‘AL Vetch’ is later flowering vetch bred in a climate with lower average and minimum temperatures (Maul et al., 2011). ‘Houser’ wheat tends to have greater fall vigor compared to ‘Caledonia’ (M. Sorrells, personal communication). Where treatments included multiple cultivars or species, the number of seeds was always split evenly, first by species (i.e. 50% wheat: 50% vetch), and then by cultivar (i.e. each vetch cultivar was 1/5 of the total seeding rate when all 5 cultivars were mixed). In addition to the cover crop treatment plots, one plot per replicate remained unplanted without any cover crops, (treatment 16) as a control, to allow for assessment of the background weed pressure and variation across the fields (Table 2).

On all four farms we used a randomized, complete block, replacement series with three

Table 3 Description of the cover crop treatments applied at the four farm sites

Description	Treatment #	Cultivar or line
Vetch cultivar monocultures	1	'Purple Bounty'
	2	'Vetch VNS'
	3	'Ernst Vetch'
	4	'AL Vetch'
	5	'Purple Prosperity'
Wheat cultivar monocultures	6	'Cayuga'
	7	'Caledonia'
	8	'Houser'
	9	'Pride of Genesee'
Cultivar mixtures	10	all vetch cultivars planted together
	11	all wheat cultivars planted together
Grass/legume mixtures (1 vetch cultivar : 1 wheat cultivar)	12	Vetch ('Purple Bounty') : Wheat ('Caledonia')
	13	Vetch ('Purple Bounty') : Wheat ('Houser')
	14	Vetch ('Albert Lea') : Wheat ('Caledonia')
Cover crop control	15	Vetch ('Albert Lea') : Wheat ('Houser')
	16	No cover crop planted, used to assess background weeds

blocks of the sixteen treatments listed in Table 3. All plots were 2.5m x 2.5m, with the exception of Farm 1 where the permanent bed size restricted the dimensions to 1.83m x 3.25m. Regardless of plot dimensions, plots were planted at the same seeding density of 285 seeds/m², per the replacement series design. This seeding density is an appropriate seeding rate on a kg/ha basis for a wheat cover crop. We then adjusted the seeding rate on a weight basis to ensure the same number of plants per treatment for the other species in the main research farm experiment. The seeding rates as kg/ha varied by species and cultivar per differences in seed size and weight, but were approximately 85 and 115 kg/ha for vetch and wheat monocultures respectively, and half that for each in the biculture composition (Table 2S). This resulted in a higher than recommended rate for vetch (233% the recommended rate). The design of this experiment is more similar to work on the biodiversity-ecosystem function relationship in unmanaged ecosystems, than to proportional replacement series or additive designs used in crop diversity research in agroecosystems (Tilman et al., 1996; Poffenbarger et al., 2015b).

We applied appropriate inoculant to all vetch seeds unless they were pretreated by the producer (N-Dure brand, Verdesian, Cary, NC). We planted all sites in mid-September 2014 within 10 days of one another (September 16-24). The field was prepared by the farmers according to their typical practice for establishing cover crops, after which seeds were hand broadcast in the experimental plots. Following seeding, the field was treated according to farmer-practice for seed incorporation. All sites were managed organically, and received no inputs or other interventions such as irrigation during the trial.

Data collection

We harvested aboveground biomass for all treatments once the vetch monocultures were at approximately 50% flowering, which occurred in mid-June 2015 at all sites. Farmers

commonly terminate vetch at this time as it is unlikely to regrow, and at which time viable seed have not yet set. Biomass was cut 9 cm above the soil surface within a randomly selected quadrant area of 0.25m². We recorded plant count for vetch, wheat, and weeds individually and separated biomass into paper bags, which were oven-dried for at least 48 hours at 60 C before weighing to the nearest 0.01g. We sampled weed biomass from cover crop control plots similarly, but did not separate by species for weeds. All cover crop samples were ground to at least 2 mm. Wheat samples from the grass/legume mixtures were analyzed for total C and N on combustion using a LECO TruMac CN analyzer (Leco Corporation, St. Joseph, Missouri, USA). Additionally, all vetch samples and wheat monoculture samples were processed for mass spectrometer isotopic analysis by first coarsely grinding, and then grinding to 0.5mm with a propeller mill (Cyclotec™ Sample Mill, Foss, Hillerød, Denmark). Samples were analyzed for ¹⁵N natural abundance, total N content, and total C content using a continuous flow Isotope Ratio Mass Spectrometer (Stable Isotope Facility, UC Davis).

Using the ¹⁵N natural abundance method, we estimated the symbiotic nitrogen fixation (SNF) by legumes in monoculture and mixtures in unfertilized plots (Shearer and Kohl, 1986). For the following calculation for each legume sampled ($\delta^{15}\text{N}_{\text{leg}}$), we used the average ¹⁵N signature of the wheat monoculture plots averaged by block, as the reference plant, ($\delta^{15}\text{N}_{\text{grass}}$), where the *B* value, the isotopic fractionation of the nitrogen during fixation in the vetch, was determined as part of a previous study for each vetch cultivar (Table 3S, Reiss and Drinkwater, in prep.).

$$\% \text{ N from fixation} = 100 \times \left(\frac{\delta^{15}\text{N}_{\text{grass}} - \delta^{15}\text{N}_{\text{leg}}}{\delta^{15}\text{N}_{\text{grass}} - B} \right)$$

To characterize background soil conditions we collected ten soil cores to 20 cm depth for composite samples from each of the four replicates in the fall. A subsample of each composite

sample was sieved to 2mm, and extracted for total inorganic N with 2M KCl, and incubated anaerobically for seven days and extracted with 2.67M KCl for N mineralization. Total NH_4^+ and NO_3^- were analyzed via a colorimetric microplate technique (QuikChem, Lachat Instruments, Loveland, CO, Ringuet et al., 2011). We dried each sample in the oven for 7 days at 65°C to determine gravimetric water content. All samples were analyzed for total C and N on combustion (Leco Corporation, St. Joseph, Missouri, USA). All fall soil samples were analyzed for water pH, CEC, Mehlich buffer lime requirement, and for P, K, Mg, Ca, Zn, Cu, and S by the Mehlich 3 (ICP) test, and particle size (Agricultural Analytical Services Laboratory, Penn State University, University Park, PA) (Table 2).

Spring growth period growing degree days were calculated from the daily minimum and maximum air temperature as recorded by temperature loggers (Thermochron iButtons, Maxim, Sunnyvale, CA, <http://www.maxim-ic.com/>) at each site from March 20, 2015 to May 31, 2015, using 0°C as the base temperature (Table 2).

Data Analysis

Ecosystem service calculations

We measured aboveground cover crop biomass (dry matter g/m^2 as described above), which directly contributes to soil organic matter levels and soil carbon accrual. Aboveground biomass also indirectly influences a range of other services. We also evaluated nitrogen use and contribution as total biomass N, and then partitioned this into soil derived N (N retention) and fixed N by legumes (long-term N supply). Total g of biomass N per m^2 (N_T) was calculated for all plants as $\%N \times \text{g of cover crop biomass/ m}^2$. Nitrogen derived from the atmosphere through biological nitrogen fixation (N_F) in vetch was partitioned using the $\%N$ from fixation result from the ^{15}N natural abundance method. The grams of fixed N in vetch was calculated as $\%SNF \times N_T$.

The soil N accrual (N_S) for vetch was calculated as $N_T - N_F$. For wheat, all biomass nitrogen is derived from the soil so $N_S = N_T$. We evaluated weed suppression as the total weed biomass harvested from the cover crop treatment plot.

Statistical analysis

Differences in soil characteristics by farm (Table 2) were tested with a mixed model (block as a random effect nested in farm and farm as a fixed effect) with differences in LSM assessed by Tukey's Honestly Significant Difference (HSD) at $\alpha=0.05$. Variables were assessed for homogeneity of variance and other assumptions for ANOVA, and no transformation was needed. Mixed models were used to test the effects of composition on ecosystem service outcome. We determined the effect of composition with three separate cover crop effect estimates. The three composition/diversity effects are as follows: 1) overall increased intraspecific diversity; 2) individual cultivar and cultivar mixture performance; 3) functional group diversity in the grass/legume mixture. First, we evaluated the overall effect of increased intraspecific diversity by comparing the mean of the ecosystem service performance of cultivar monocultures to that of the cultivar mixture for each of the two species. Specifically, the mean of treatments 1-5 was compared to treatment 10 for vetch and the mean of treatments 6-9 compared to treatment 11 for wheat (Table 3). Second, we assessed the effect of the individual cultivar and cultivar mixture performance by comparing each cultivar monoculture to the others as well as to the cultivar mixture for each species. In other words, treatments 1-5 and 10 were compared to one another for vetch and treatments 6-9 and 11 for wheat. Third and finally, for the grass/legume mixture effect we compared the performance of the four vetch-wheat mixes (treatments 12-15) to the monocultures of each of the vetch and wheat cultivars alone (treatments 1 and 4 for vetch and 7 and 8 for wheat). Treatment means by replicate were used to avoid

pseudoreplication for the separate diversity/composition levels used in the analysis. For example, the mean of treatments 1 and 4 together was calculated for each block representing monoculture vetch, and the mean of treatments 12-15 together by block was also calculated for biculture of vetch and wheat. Consequently, for means presented in figures, the n for each treatment type/diversity level is equal to the number of blocks (3 at Farms 1-3, 4 at Farm 4). The mixed models included the main effect being tested (e.g. grass/legume mixture), the farm and their interaction along with block nested in farm as a random effect. Tukey's HSD at $\alpha=0.05$ was used to test for differences between multiple levels, while two sample t-tests were used to compare two levels. Data was transformed to meet model assumptions. Variables and models where data was transformed are identified in figure legends and supplemental materials. For most analyses unless otherwise stated, a constant (1) was added to weed biomass and then transformed using natural log. All analyses were conducted using JMP v.11 software (SAS Institute Inc. Cary, NC).

Results and Discussion

Cover crop performance by farm

When we compared the delivery of the five ecosystem services among the four farms, we clearly saw differences among them for vetch and wheat monocultures (Figure 1). Overall, Farm 1 had high or the highest ecosystem services performance measured by the performance of the vetch and wheat monocultures. Cover crop seed at this farm was incorporated with a shallow rototill pass, which resulted in very consistent and high density establishment (Figure 1S). Farm 2 had relatively high cover crop biomass and ecosystem services performance, often equivalent to Farm 1, even though seeds were not mechanically incorporated after hand broadcasting, per typical farm practice. Farm 3 generally had some of the lowest performing cover crops in terms of ecosystem service delivery. The success of weeds at this farm indicates that site productivity

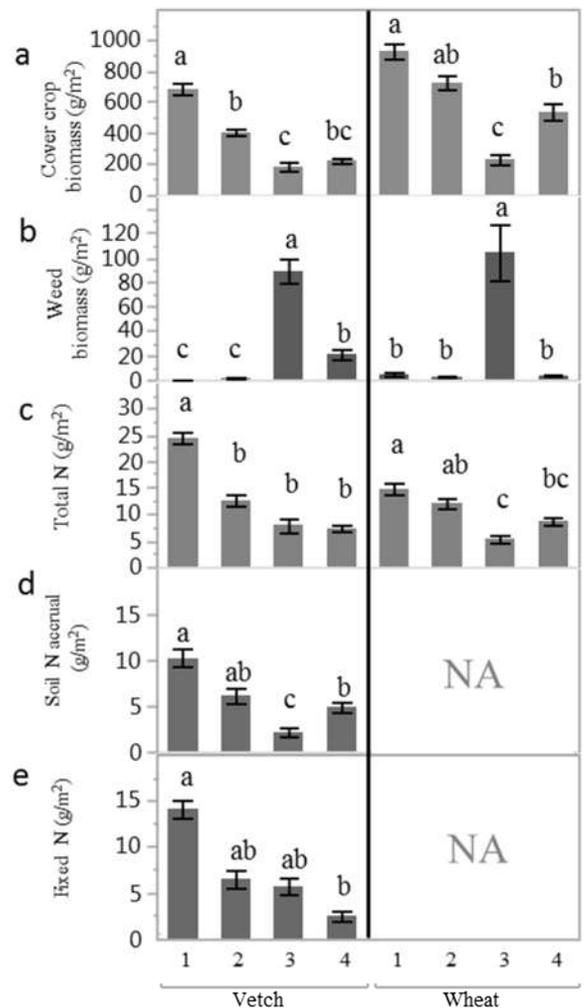


Figure 1 Average a) aboveground cover crop biomass b) weed biomass c) total N in biomass d) soil N accretion and e) fixed nitrogen (mean +/- SE) by farm (1-4) for all vetch and wheat monocultures. Different lower case letters within a species and metric indicate significant differences based on Tukey's HSD test ($p < 0.05$. Analysis on transformed data (square root for a,c,d,e, natural log for b), but untransformed data presented). See Table 9S for more details.

did not broadly inhibit cover crop performance. While we did not measure this, the seed may have been incorporated too deep for optimal establishment due to the use of a field cultivator for seed incorporation. Additionally, the low pH may have inhibited the legume growth as well as N mineralization from the large stores of organic matter under cooler temperatures (Fu et al., 1987; Curtin et al., 1998). The research farm, Farm 4, was also slightly lower performing compared to Farms 1 and 2, perhaps also due to soil conditions, such as low P and high pH (Table 2). The substantial differences in cover crop and ecosystem services performance among the farms reiterates how important and influential site conditions can be for both practical and research outcomes.

Intraspecific diversity

While there were substantial differences between farms in terms of overall cover crop performance (Figure 1), there was no effect of either measure of intraspecific diversity, and consequently, no interaction with farm (Table 4S and 5S). Specifically, we did not detect any differences in the mean cultivar monoculture performance compared to the cultivar mixture, nor did any of the cultivars differ from one another or from the cultivar mixture at a given site.

As we did not detect any differences among the cultivars (Table 5S), it would suggest that the range of phenotypic and trait diversity in the cultivar mixtures may have been too narrow. As such it may not have been sufficient to result in a measurable effect on the target ecosystem services (Cadotte et al., 2008). The limited trait information available for the cultivars, especially for vetch, made this difficult. Additionally, as vetch is an outcrossing species, there is greater intraspecific diversity within a given line or cultivar (Yeater et al., 2004; Maul et al., 2011), making the relative increase in genotypic diversity less when lines are mixed.

The lack of discernable differences in cultivars or cultivar mixtures may also have been

due in part to the average or above-average biomass production at all sites (Clark, 2007). As noted previously, our seeding rates were very high and may have contributed to this high biomass. Vetch was seeded as a monoculture at over 200% of the recommended rate, while the wheat was seeded very close to recommended rates in terms of kg/ha. In spite of these higher seeding rates, both the vetch and wheat established at much lower populations (Figure 1S, Table 6S). Self-thinning and other biotic and abiotic pressures may have reduced the final stands from the high density at seeding (Park et al., 2003). However, the final densities in most plots were still very robust. In fact, the seeding rate corresponding to the final established plant density in vetch is nearly the same as the revised recommended rate for the region by Mirsky et al. (2017). These authors suggest that in general, optimal biomass may be attained with lower than recommended rates for vetch (Figure 1, Table 2S and 6S), but of course there is a variable relationship between the planted seeding rate (Mirsky et al., 2017). Ultimately, the final outcome of these cover crops, such as productivity, is highly dependent on seeding date and time of termination. Given that these plots were planted at the later end of the fall cover crop window, a somewhat higher seeding rate would be appropriate, though our seeding rate was still excessive from an agronomic perspective (Mirsky et al., 2017). Especially given the large difference between seeded density and final density, our experimental rates and associated costs are not justified for a working farm, as the biculture cost of \$330/ha would be prohibitive for most farms (Table 2S). While costs are an important concern, surveys suggest that cover crop use and adoption is not inhibited by the costs (Dunn et al., 2016). Instead, evaluating the goals from a cover crop and working to estimate return, such as new nitrogen from fixation instead of compost, is a logical step for determining appropriate seeding rates on-farm (Snapp et al., 2005).

Increased complexity in grass/legume mixtures

Overall, the grass/legume mixture was as good as the best monoculture for three of the five measured services, with no differences for the other two (Figure 2). These same three metrics, cover crop biomass, total N and soil N accrual, were each context dependent, with an interaction between the effect of grass/legume mixing and farm site. Additionally, there were often significant differences between the vetch monoculture, wheat monoculture, and the biculture at each farm for these metrics.

Even though the biomass response of grass/legume mixture was context dependent, there were some consistent patterns across the farms (Table 4). On three out of the four farms, the mixture had significantly greater biomass than the legume monoculture and at Farm 3 mixture biomass was also significantly greater than the grass monoculture (Figure 3a). It is very common to grow grass/legume mixtures in large part because of the potential biomass benefits of mixing a grass and a legume together (Sainju et al., 2005; Bedoussac et al., 2015). We generally did observe this biomass improvement, but there are clearly differences in the way the cover crop treatments responded to the site context. With grass/legume mixtures there is often a concern that in soils with greater levels of N availability the grass might outcompete the legume, substantially reducing its biomass (Staniforth, 1962; Brainard et al., 2011; Poffenbarger et al., 2015a). Across the four farms, the vetch consistently made up 50-60% of the mixture (by density and biomass) with the exception of Farm 4 where it was closer to 50% by density and 20% by biomass. This 50-60% proportion is consistent across the three farms with varying overall grass/legume mixture responses (Figure 3a). More specifically, when the weeds are included as part of the non-legume portion of the cover crop community this difference is more dramatic (non-legume: 60% by biomass for Farms 1-3, 83% at Farm 4). This distribution of legume to non-legume

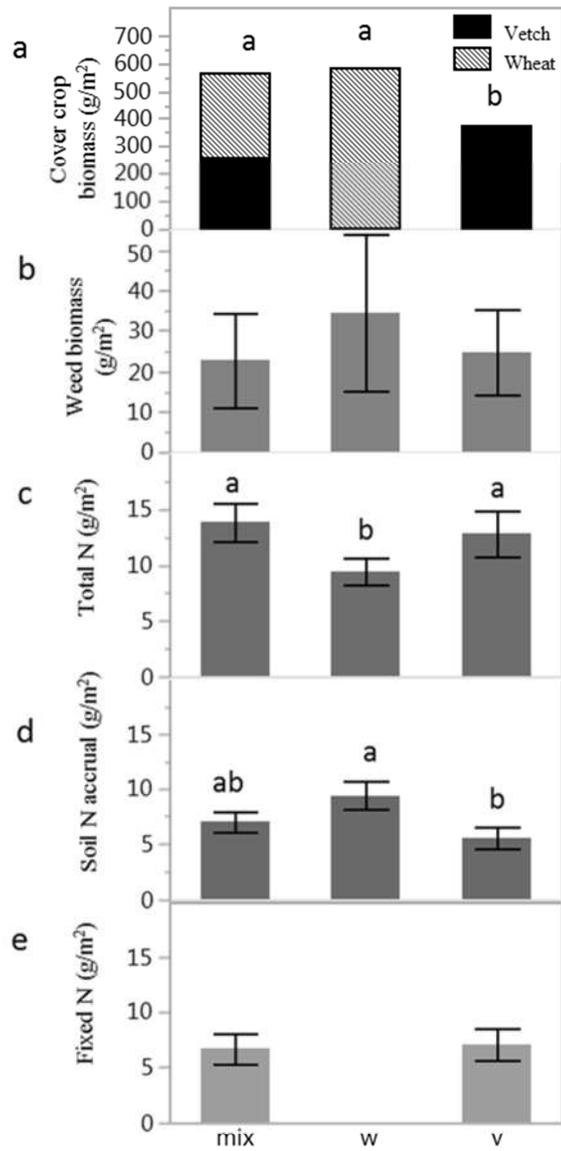


Figure 2 Effect of grass/legume mixture (mix) compared to wheat (w) and vetch (v) monocultures on a) aboveground cover crop biomass b) weed biomass c) total N in biomass d) soil N accrual and e) fixed nitrogen (mean \pm SE). Different lower case letters within a farm and metric indicate significant differences based on Tukey's HSD test ($p < 0.05$). Data was transformed for weed biomass (natural log) and soil N accrual (square root), but untransformed data presented. Also see Table 4.

Table 4 Mixed model results for effect of mixing grasses and legumes (mix) on five metrics of ecosystem service performance including interaction with farm site (also see Figure 2). When an interaction between mix and farm exists, farm sites with significant or nearly significant differences are presented. Data was transformed for weed biomass (natural log) and soil N accrual (square root). There was no interaction with farm and no diversity effect for either measure of intraspecific diversity (cultivar mixture effect or cultivar differences). See tables 1S and 2S for additional details. DF= Degree of Freedom in the numerator; DFDen= Degrees of Freedom in the denominator.

	Effect of grass/legume mixing (mix)				
Ecosystem service	<i>Biculture compared to wheat and vetch in monoculture</i>				
Cover crop biomass	Source	DF	DFDen	F Ratio	Prob > F
	Farm	3	9	19.074	0.0003*
	mix	2	18	23.2569	<.0001*
	Farm*mix	6	18	4.0732	0.0094*
	Farm 1	F(2,4)=50.8383, p=0.0014			
	Farm 2	F(2,4)=5.8129, p=0.0655			
	Farm 3	F(2,4)=38.3787, p=0.0025			
	Farm 4	F(2,6)=9.4535, p=0.014			
Weed biomass	Source	DF	DFDen	F Ratio	Prob > F
	Farm	3	9	77.6326	<.0001*
	mix	2	18	1.098	0.3549
	Farm*mix	6	18	2.4652	0.0643
Total N	Source	DF	DFDen	F Ratio	Prob > F
	Farm	3	9	17.2387	0.0005*
	mix	2	18	16.8699	<.0001*
	Farm*mix	6	18	4.3524	0.0069*
	Farm 1	F(2,4)=111.7145, p=0.0003			
	Farm 3	F(2,4)=25.1934, p=0.0054			
Soil N accrual	Source	DF	DFDen	F Ratio	Prob > F
	Farm	3	9	8.0648	0.0064*
	mix	2	18	25.0479	<.0001*
	Farm*mix	6	18	5.9612	0.0014*
	Farm 1	F(2,4)=16.26, p=0.012			
	Farm 2	F(2,4)=6.9661, p=0.0498			
	Farm 3	F(2,4)=20.3943, p=0.008			
	Farm 4	F(2,6)=13.9829, p=0.0055			
Fixed N	Source	DF	DFDen	F Ratio	Prob > F
	Farm	3	9	18.6615	0.0003*
	mix	1	9	0.032	0.862
	Farm*mix	3	9	0.4451	0.7267
	No interaction between mix and farm				

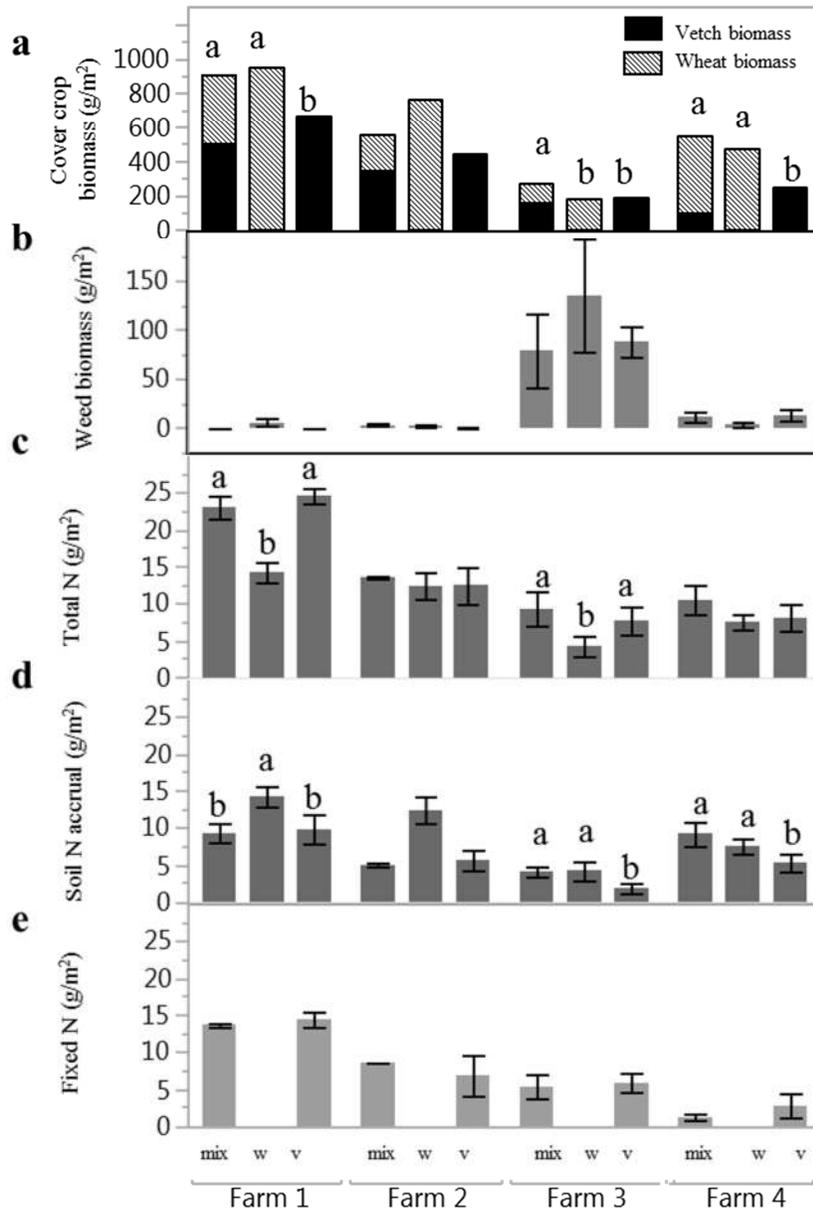


Figure 3 Mean effect of grass/legume mixture (mix) compared to wheat (w) and vetch (v) monocultures by farm on a) aboveground cover crop biomass b) weed biomass c) total N in biomass d) soil N accrual and e) fixed nitrogen. Different lower case letters within a farm and metric indicate significant differences based on Tukey's HSD test (p < 0.05). Data was transformed for weed biomass (natural log) and soil N accrual (square root), but untransformed data presented. Also see Table 4.

ratios, along with the significant farm interaction, suggests that the context under which the cover crop is propagated can influence the effect of mixing grasses and legumes even when the ratio of legume to non-legume in a mixture remains the same. With only four sites and a limited set of environmental variables measured it is difficult to isolate potential factors driving these different mixture responses given similar mixture compositions, but it is an area that demands additional research.

We found that total N in the cover crop biomass and soil N accrual was context dependent in the response to the mixing of a grass and a legume (Figure 3, Table 4), and closely followed the pattern of biomass as expected. Using the natural abundance method (Shearer and Kohl, 1986) we were able to partition total N into soil N and fixed N. For wheat, total N is equal to soil N, but the vetch and mixture treatments have both soil N and fixed N.

Although total N is an important consideration for the nutrition of subsequent crops, soil N accrual and fixed N each have critical roles in the long and short-term nutrient management strategy for a farm. By mixing the legume and grass we hoped to take advantage of both of their strengths and minimize tradeoffs (Ranells and Wagger, 1996; Aronsson et al., 2016). Though competition from wheat resulting in suppressed vetch growth and reduced fixed N is a typical concern, three of the four farms had close to 60% vetch in the mixtures, indicating that the vetch was not outcompeted by the wheat (Poffenbarger et al., 2015a; White et al., 2017). Additionally, at only one of the four farms was the mixture worse at soil N retention than the best monoculture. This was true even though the mixtures were generally evenly mixed with the vetch and wheat. In contrast to total N and soil N accrual, there was no interaction and no difference between vetch and the mix for the amount or rate of fixed N (Table 4, Figure 3e, Figure 2S). Consequently, for all farms we found that the amount of fixed N, as well as the % of

N from fixation was not different in the mixture compared to the vetch monoculture. While other research has shown how site conditions, such as soil fertility, can dramatically alter these N service outcomes, ultimately we found that the majority of the mixtures succeeded at balancing the tradeoffs of the legume and grass across the farms (West et al., 2005; Schipanski and Drinkwater, 2011). The mixture of the two functional groups maintained good soil N retention, while contributing new N equivalent to that of legume monocultures through SNF.

There was no interaction between farm and the effect of mixing grasses and legumes for weed suppression even given the large range of background weed pressure across the farms, as we might have predicted. (Table 4, Figure 3b). It appears that the cover crops were generally able to suppress weeds well, regardless of cover crop composition. The high seeding rate may have contributed to the smothering effect overall, and reduced any treatment differences. While we did not measure light penetration, it would be a useful parameter to consider to evaluate the mechanism for suppression in these cover crop compositions at these high seeding densities (Liebman et al., 2001).

Practical management implications

Across all the farms the grass/legume mixture was as good as or better than the best monoculture across ecosystem services, with only one exception, soil N accrual at Farm 1 (Figure 3d). Similarly, while we did not find a positive effect from increased intraspecific diversity as we expected, we also did not find any negative impacts. From our results, we can conclude that while increasing complexity, either through cultivar mixtures or grass/legume mixtures, may not regularly provide substantial benefits, the risk of a substantial downside to mixing grasses and legumes or cultivars is low. It does appear that the cover crop response to increasing composition complexity is context dependent, and as such a good practice might be to

try a mixture alongside the monocultures at a given site and observe any differences, where possible. Of course, weather and other variable site factors might influence the outcome in a given year.

Management considerations are also critical to include when making decisions about cover crop composition, as the outcome alone may not justify the costs or additional time and complexity of adding a cover crop into a rotation. When nitrogen fixation is a priority, the mixture of wheat and vetch is a better choice than vetch alone as it produced an equivalent amount of fixed N at half the seeding rate of vetch (Table 4, Figure 3e). This biculture is approximately \$25-50 per hectare cheaper than the monoculture of vetch (Table 5S). For farmers with a low-input approach to profitability, the additional cost of establishing a cover crop with any legume component could be prohibitive (Dunn et al., 2016). Identifying what other input costs could be offset from a biculture cover crop could help to justify these costs and make a cover crop more attractive. For instance, with higher rates of N fixation in mixtures, the cost of new nitrogen from fixation may be cheaper than off-farm sources of N such as manure. The mixture has the additional benefits of generally greater biomass and equivalent weed suppression compared to the vetch alone (Figure 3).

We only assessed a selection of ecosystem services from cover crops, and there may be other management considerations which might cause a farmer to make a certain decision about cover crop selection (e.g. pest or pollinator management). For instance, the preceding crop may make it difficult to establish a less cold tolerant legume compared to a more tolerant grass like rye. However, given these results, compared to their monocultures an even mixture of vetch and wheat appears to be an economical choice and provides good biomass production, N fixation and retention, and weed suppression across a range of on-farm contexts.

REFERENCES

- Appelgate, S.R., A.W. Lenssen, M.H. Wiedenhoef, and T.C. Kaspar. 2017. Cover Crop Options and Mixes for Upper Midwest Corn–Soybean Systems. *Agronomy Journal* 109(3): 968–984. doi: 10.2134/agronj2016.08.0453.
- Aronsson, H., E.M. Hansen, I.K. Thomsen, J. Liu, A.F. Øgaard, H. Känkänen, and B. Ulén. 2016. The ability of cover crops to reduce nitrogen and phosphorus losses from arable land in southern Scandinavia and Finland. *Journal of Soil and Water Conservation* 71(1): 41–55. doi: 10.2489/jswc.71.1.41.
- Asrat, S., M. Yesuf, F. Carlsson, and E. Wale. 2010. Farmers’ preferences for crop variety traits: Lessons for on-farm conservation and technology adoption. *Ecological Economics* 69(12): 2394–2401. doi: 10.1016/j.ecolecon.2010.07.006.
- Baba, M., R.A. Halim, A.R. Alimon, and I. Abubakar. 2011. Grass-legume mixtures for enhanced forage production: Analysis of dry matter yield and competition indices. *African Journal of Agricultural Research* 6(23): 5242–5250.
- Bedoussac, L., E.-P. Journet, H. Hauggaard-Nielsen, C. Naudin, G. Corre-Hellou, E.S. Jensen, L. Prieur, and E. Justes. 2015. Ecological principles underlying the increase of productivity achieved by cereal-grain legume intercrops in organic farming. A review. *Agronomy for Sustainable Development* 35(3): 911–935. doi: 10.1007/s13593-014-0277-7.
- Brainard, D.C., R.R. Bellinder, and V. Kumar. 2011. Grass-Legume Mixtures and Soil Fertility Affect Cover Crop Performance and Weed Seed Production. *Weed Technology* 25(3): 473–479. doi: 10.1614/WT-D-10-00134.1.
- Brainard, D., B. Henshaw, and S. Snapp. 2012. Hairy Vetch Varieties and Bi-Cultures Influence Cover Crop Services in Strip-Tilled Sweet Corn. *Agronomy Journal* 104(3): 629. doi: 10.2134/agronj2011.0360.

- Cadotte, M.W., B.J. Cardinale, and T.H. Oakley. 2008. Evolutionary history and the effect of biodiversity on plant productivity. *Proceedings of the National Academy of Sciences* 105(44): 17012–17017.
- Canali, S., L. Ortolani, G. Campanelli, M. Robacer, P. von Fragstein, D. D’Oppido, and H.L. Kristensen. 2017. Yield, product quality and energy use in organic vegetable living mulch cropping systems: research evidence and farmers’ perception. *Renew. Agr. Food Syst.* 32(3): 200–213. doi: 10.1017/S1742170516000314.
- Clark, A. (Ed). 2007. *Managing cover crops profitably*. 3rd ed. Sustainable Agriculture Research & Education (Program), College Park, MD.
- Clark, A.J., A.M. Decker, and J.J. Meisinger. 1994. Seeding Rate and Kill Date Effects on Hairy Vetch-Cereal Rye Cover Crop Mixtures for Corn Production. *Agronomy Journal* 86(6): 1065. doi: 10.2134/agronj1994.00021962008600060025x.
- Creamer, N.G., M.A. Bennett, and B.R. Stinner. 1997. Evaluation of Cover Crop Mixtures for Use in Vegetable Production Systems. *HortScience* 32(5): 866–870.
- Curtin, D., C.A. Campbell, and A. Jalil. 1998. Effects of acidity on mineralization: pH-dependence of organic matter mineralization in weakly acidic soils. *Soil Biology and Biochemistry* 30(1): 57–64. doi: 10.1016/S0038-0717(97)00094-1.
- Doltra, J., and J.E. Olesen. 2013. The role of catch crops in the ecological intensification of spring cereals in organic farming under Nordic climate. *European Journal of Agronomy* 44: 98–108. doi: 10.1016/j.eja.2012.03.006.
- Drinkwater, L.E. 2016. *Systems research for agriculture: innovative solutions to complex challenges*. Published by the Sustainable Agriculture Research and Education (SARE)

- Program, with funding from the National Institute of Food and Agriculture, U.S. Department of Agriculture, College Park [Md.].
- Dunn, M., J.D. Ulrich-Schad, L.S. Prokopy, R.L. Myers, C.R. Watts, and K. Scanlon. 2016. Perceptions and use of cover crops among early adopters: Findings from a national survey. *Journal of Soil and Water Conservation* 71(1): 29–40. doi: 10.2489/jswc.71.1.29.
- Elzebroek, A.T.G. 2008. *Guide to Cultivated Plants*. CABI.
- Fu, Y.-B. 2015. Understanding crop genetic diversity under modern plant breeding. *Theor Appl Genet* 128(11): 2131–2142. doi: 10.1007/s00122-015-2585-y.
- Fu, M.H., X.C. Xu, and M.A. Tabatabai. 1987. Effect of pH on nitrogen mineralization in crop-residue-treated soils. *Biol Fert Soils* 5(2): 115–119. doi: 10.1007/BF00257645.
- Gregory, A.S., J.A.J. Dungait, C.W. Watts, R. Bol, E.R. Dixon, R.P. White, and A.P. Whitmore. 2016. Long-term management changes topsoil and subsoil organic carbon and nitrogen dynamics in a temperate agricultural system. *Eur J Soil Sci* 67(4): 421–430. doi: 10.1111/ejss.12359.
- Grettenberger, I.M., and J.F. Tooker. 2015. Moving beyond resistance management toward an expanded role for seed mixtures in agriculture. *Agriculture, Ecosystems & Environment* 208: 29–36. doi: 10.1016/j.agee.2015.04.019.
- Helland, S.J., and J.B. Holland. 2001. Blend response and stability and cultivar blending ability in oat. *Crop science* 41(6): 1689–1696.
- Helland, S.J., and J.B. Holland. 2003. Genome-wide genetic diversity among components does not cause cultivar blend responses. *Crop Science* 43(5): 1618+.

- Kiær, L.P., I.M. Skovgaard, and H. Østergård. 2009. Grain yield increase in cereal variety mixtures: A meta-analysis of field trials. *Field Crops Research* 114(3): 361–373. doi: 10.1016/j.fcr.2009.09.006.
- Kruidhof, H.M., L. Bastiaans, and M.J. Kropff. 2008. Ecological weed management by cover cropping: effects on weed growth in autumn and weed establishment in spring. *Weed Res.* 48(6): 492–502. doi: 10.1111/j.1365-3180.2008.00665.x.
- Liebman, M., C.L. Mohler, and C.P. Staver. 2001. *Ecological management of agricultural weeds.* Cambridge University Press, Cambridge, U.K.; New York.
- Maul, J., S. Mirsky, S. Emche, and T. Devine. 2011. Evaluating a Germplasm Collection of the Cover Crop Hairy Vetch for Use in Sustainable Farming Systems. *Crop Science* 51(6): 2615. doi: 10.2135/cropsci2010.09.0561.
- Mengistu, N., P.S. Baenziger, L.A. Nelson, K.M. Eskridge, R.N. Klein, D.D. Baltensperger, and R.W. Elmore. 2010. Grain Yield Performance and Stability of Cultivar Blends vs. Component Cultivars of Hard Winter Wheat in Nebraska. *Crop Science* 50(2): 617. doi: 10.2135/cropsci2009.05.0280.
- Mirsky, S.B., V.J. Ackroyd, S. Cordeau, W.S. Curran, M. Hashemi, S.C. Reberg-Horton, M.R. Ryan, and J.T. Spargo. 2017. Hairy Vetch Biomass across the Eastern United States: Effects of Latitude, Seeding Rate and Date, and Termination Timing. *Agronomy Journal* 109(4): 1510. doi: 10.2134/agronj2016.09.0556.
- Mundt, C.C. 2002. Use of Multiline Cultivars and Cultivar Mixtures for Disease Management. *Annual Review of Phytopathology* 40(1): 381–410. doi: 10.1146/annurev.phyto.40.011402.113723.

- Murrell, E.G., M.E. Schipanski, D.M. Finney, M.C. Hunter, M. Burgess, J.C. LaChance, B. Baraibar, C.M. White, D.A. Mortensen, and J.P. Kaye. 2017. Achieving Diverse Cover Crop Mixtures: Effects of Planting Date and Seeding Rate. *Agronomy Journal* 109(1): 259. doi: 10.2134/agronj2016.03.0174.
- Myers, R., and C. Watts. 2015. Progress and perspectives with cover crops: Interpreting three years of farmer surveys on cover crops. *Journal of Soil and Water Conservation* 70(6): 125A–129A. doi: 10.2489/jswc.70.6.125A.
- Nyfelner, D., O. Huguenin-Elie, M. Suter, E. Frossard, J. Connolly, and A. Luescher. 2009. Strong mixture effects among four species in fertilized agricultural grassland led to persistent and consistent transgressive overyielding. *Journal of Applied Ecology* 46(3): 683–691. doi: 10.1111/j.1365-2664.2009.01653.x.
- Osman, A., and A. Diek. 1982. Effects of defoliation on yield and forage quality of some tropical grasses, legumes and their mixtures. *Experimental Agriculture* 18(2): 157–166. doi: 10.1017/S0014479700013636.
- Park, S.E., L.R. Benjamin, and A.R. Watkinson. 2003. The Theory and Application of Plant Competition Models: an Agronomic Perspective. *Ann Bot* 92(6): 741–748. doi: 10.1093/aob/mcg204.
- Poffenbarger, H.J., S.B. Mirsky, R.R. Weil, M. Kramer, J.T. Spargo, and M.A. Cavigelli. 2015a. Legume Proportion, Poultry Litter, and Tillage Effects on Cover Crop Decomposition. *Agronomy Journal* 107(6): 2083. doi: 10.2134/agronj15.0065.
- Poffenbarger, H.J., S.B. Mirsky, R.R. Weil, J.E. Maul, M. Kramer, J.T. Spargo, and M.A. Cavigelli. 2015b. Biomass and Nitrogen Content of Hairy Vetch–Cereal Rye Cover Crop

- Mixtures as Influenced by Species Proportions. *Agronomy Journal* 107(6): 2069. doi: 10.2134/agronj14.0462.
- Ramirez-Garcia, J., J.M. Carrillo, M. Ruiz, M. Alonso-Ayuso, and M. Quemada. 2015. Multicriteria decision analysis applied to cover crop species and cultivars selection. *Field Crop. Res.* 175: 106–115. doi: 10.1016/j.fcr.2015.02.008.
- Ranells, N.N., and M.G. Wagger. 1996. Nitrogen Release from Grass and Legume Cover Crop Monocultures and Bicultures. *Agronomy Journal* 88(5): 777. doi: 10.2134/agronj1996.00021962008800050015x.
- Ranells, N.N., and M.G. Wagger. 1997. Nitrogen-15 recovery and release by rye and crimson clover cover crops. *Soil Science Society of America Journal* 61(3): 943–948.
- Reiss, E.R., and L.E. Drinkwater. 2018. Cultivar mixtures: a meta-analysis of the effect of intraspecific diversity on crop yield. *Ecol Appl* 28(1): 62–77. doi: 10.1002/eap.1629.
- Ringuet, S., L. Sassano, and Z.I. Johnson. 2011. A suite of microplate reader-based colorimetric methods to quantify ammonium, nitrate, orthophosphate and silicate concentrations for aquatic nutrient monitoring. *J. Environ. Monit.* 13(2): 370–376. doi: 10.1039/C0EM00290A.
- Ritter, W.F., R.W. Scarborough, and A.E.M. Chirnside. 1998. Winter cover crops as a best management practice for reducing nitrogen leaching. *Journal of Contaminant Hydrology* 34(1): 1–15. doi: 10.1016/S0169-7722(98)00087-4.
- Ruffo, M.L., and G.A. Bollero. 2003. Modeling rye and hairy vetch residue decomposition as a function of degree-days and decomposition-days. *Agronomy Journal* 95(4): 900–907.

- Sainju, U.M., W.F. Whitehead, and B.P. Singh. 2005. Biculture Legume–Cereal Cover Crops for Enhanced Biomass Yield and Carbon and Nitrogen. *Agronomy Journal* 97(5): 1403. doi: 10.2134/agronj2004.0274.
- Schipanski, M.E., M. Barbercheck, M.R. Douglas, D.M. Finney, K. Haider, J.P. Kaye, A.R. Kemanian, D.A. Mortensen, M.R. Ryan, J. Tooker, and C. White. 2014. A framework for evaluating ecosystem services provided by cover crops in agroecosystems. *Agricultural Systems* 125: 12–22. doi: 10.1016/j.agsy.2013.11.004.
- Schipanski, M., and L. Drinkwater. 2011. Nitrogen fixation of red clover interseeded with winter cereals across a management-induced fertility gradient. *Nutrient Cycling in Agroecosystems* 90(1): 105–119. doi: 10.1007/s10705-010-9415-z.
- Schipanski, M.E., and L.E. Drinkwater. 2012. Nitrogen fixation in annual and perennial legume-grass mixtures across a fertility gradient. *Plant and Soil* 357(1–2): 147–159. doi: 10.1007/s11104-012-1137-3.
- Shearer, G., and D.H. Kohl. 1986. N₂ fixation in field settings: estimates based on natural ¹⁵N abundance. *Australian Journal of Plant Physiology* 13: 699–756.
- Sinebo, W. 2005. Trade off between yield increase and yield stability in three decades of barley breeding in a tropical highland environment. *Field Crops Research* 92(1): 35–52. doi: 10.1016/j.fcr.2004.08.009.
- Smithson, J.B., and J.M. Lenne. 1996. Varietal mixtures: a viable strategy for sustainable productivity in subsistence agriculture. *Annals of Applied Biology* 128(1): 127–158.
- Snapp, S.S., S.M. Swinton, R. Labarta, D. Mutch, J.R. Black, R. Leep, J. Nyiraneza, and K. O’Neil. 2005. Evaluating cover crops for benefits, costs and performance within cropping system niches. *Agron. J.* 97(1): 322–332.

- Staniforth, D.W. 1962. Responses of Soybean Varieties to Weed Competition. *Agronomy Journal* 54(1): 11–13. doi: 10.2134/agronj1962.00021962005400010004x.
- Tilman, D., D. Wedin, and J. Knops. 1996. Productivity and sustainability influenced by biodiversity in grassland ecosystems. *Nature* 379(6567): 718–720. doi: 10.1038/379718a0.
- Valkama, E., R. Lemola, H. Känkänen, and E. Turtola. 2015. Meta-analysis of the effects of undersown catch crops on nitrogen leaching loss and grain yields in the Nordic countries. *Agriculture, Ecosystems & Environment* 203: 93–101. doi: 10.1016/j.agee.2015.01.023.
- Vandermeer, J.H. 1989. *The Ecology of Intercropping*. Cambridge University Press.
- Wayman, S., L. Kissing Kucek, S.B. Mirsky, V. Ackroyd, S. Cordeau, and M.R. Ryan. 2016. Organic and conventional farmers differ in their perspectives on cover crop use and breeding. *Renewable Agriculture and Food Systems*: 1–10. doi: 10.1017/S1742170516000338.
- West, J.B., J. HilleRisLambers, T.D. Lee, S.E. Hobbie, and P.B. Reich. 2005. Legume species identity and soil nitrogen supply determine symbiotic nitrogen-fixation responses to elevated atmospheric [CO₂]. *New Phytologist* 167(2): 523–530. doi: 10.1111/j.1469-8137.2005.01444.x.
- White, C.M., S.T. DuPont, M. Hautau, D. Hartman, D.M. Finney, B. Bradley, J.C. LaChance, and J.P. Kaye. 2017. Managing the trade off between nitrogen supply and retention with cover crop mixtures. *Agriculture, Ecosystems & Environment* 237: 121–133. doi: 10.1016/j.agee.2016.12.016.

Yeater, K.M., G.A. Bollero, D.G. Bullock, A.L. Rayburn, and S. Rodriguez-Zas. 2004.

Assessment of genetic variation in hairy vetch using canonical discriminant analysis.

Crop Science 44(1): 185–189.

Supplemental material

Table 1S Additional site conditions at the four farm sites. Mean (standard deviation). Different upper case letters for a given variable indicate significant differences among farms based on Tukey's HSD test ($p < 0.05$). Farms 1-3 are small-scale working farms, and Farm 4 is a university-managed research farm.

	Farm 1	Farm 2	Farm 3	Farm 4
Mg (mg/kg)	237 (29) B	164 (18) C	107 (12) C	315 (26) A
Ca (mg/kg)	2020 (362) A	1330 (37) B	1284 (126) B	2399 (74) A
CEC	12.6 (2.1) AB	10.3 (0.3) B	12.0 (2.0) AB	14.8 (0.6) A
K (% saturation)	4.0 (0.3) A	2.8 (0.5) B	1.5 (0.3) C	1.3 (0.1) C
Mg (% saturation)	15.8 (0.8) A	13.3 (1.2) B	7.4 (0.5) C	17.7 (0.9) A
Ca (% saturation)	80.3 (0.7) A	64.5 (1.2) B	54.0 (6.5) C	81.0 (1.0) A
Sand %	40.7	37.2	36.7	44.6
Clay %	25.5	27.9	25.3	22.7
NRCS Soil Series	HnB: Honeoye silt loam	HnB: Honeoye silt loam	ErB: Erie silt loam	LtA: Lima silt loam
Field prep prior to planting	Rototill	Chisel plow	Mold board	Disk
Field treatment after seeding	Shallow rototill	None	Field cultivate	Cultipack
Crop history	20 yrs. of mixed vegetables and cover crops	Conventional corn/soy for 20+ years, then 5 yrs. of hay and mixed vegetables and cover crops	Hay for 10+ yrs. then 2 years of mixed vegetables	Conventional corn/soy/wheat rotation as part of University research farm

Table 2S Seeding rates and cost for vetch and wheat monocultures and biculture. Recommended rates per Clark 2007, Seeded rates based on 285 seeds/m²population, Actual established rates are calculated as the seed equivalent of harvested density of plants, i.e. one plant harvested = one seed planted. This does not account for germination, establishment losses etc. and should only be used as a comparison with recommended and seeded rates.

	Recommended rates		Seeded rates			Actual established rates	
	kg/ha	\$/ha	kg/ha	\$/ha	As % of recommended	kg/ha	\$/ha
Vetch monoculture	36	\$225	85	\$535	233%	20	\$124
Wheat monoculture	118	\$127	116	\$125	99%	26	\$28
Vetch/Wheat biculture	18/59	\$176	43/58	\$330	NA	14/13	\$100

Table 3S B values of vetch cultivars in study

Cultivar name	Shoot $\delta^{15}\text{N}$ (‰)
Vetch	
Purple Bounty	-0.72
Vetch VNS	-0.47
Ernst Vetch	-0.60
AL Vetch	-0.15
Purple Prosperity	-0.68

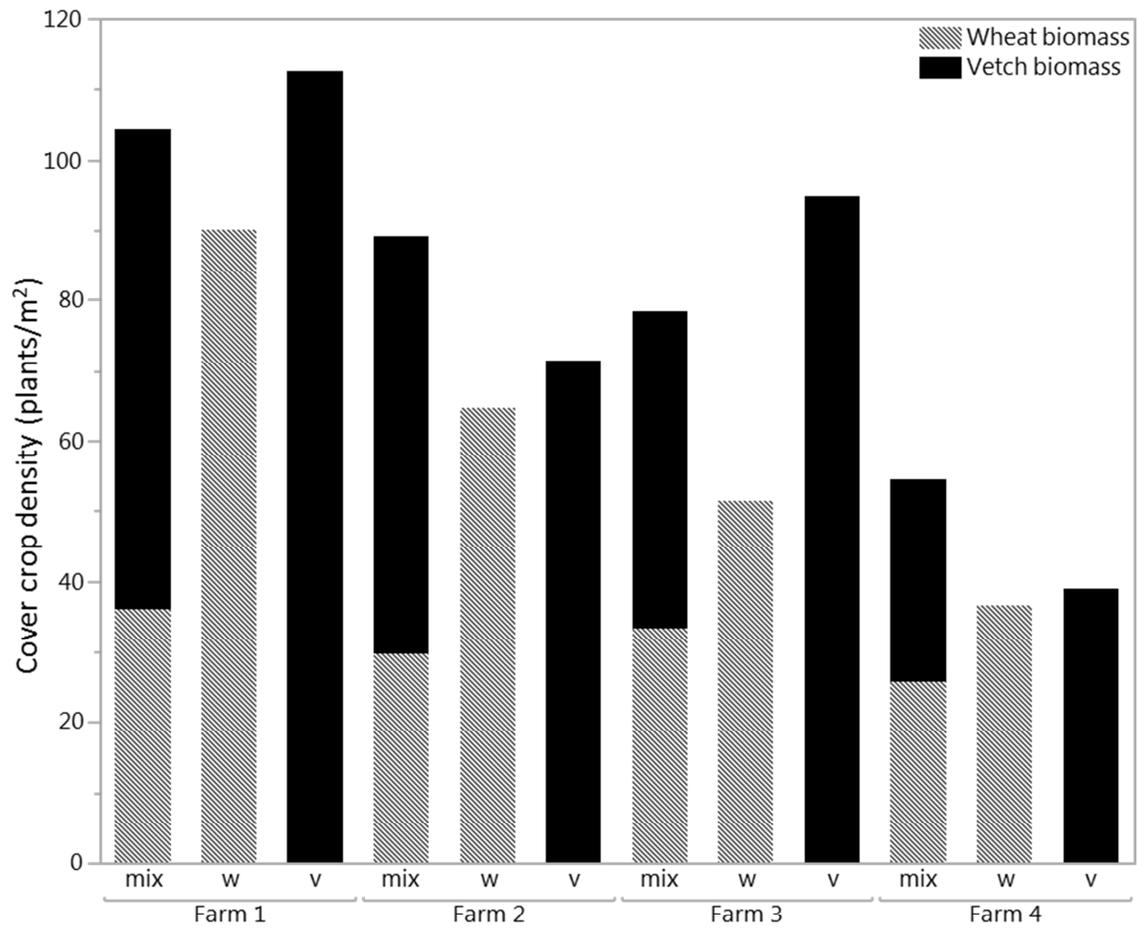


Figure 15 Mean density of vetch and wheat plants at biomass harvest by farm

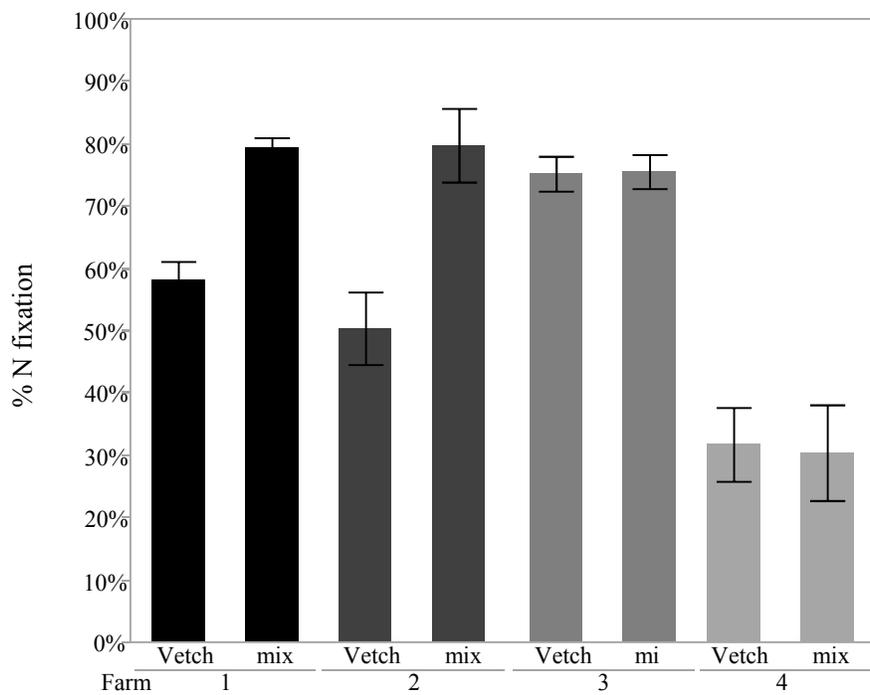


Figure 2S Percent of nitrogen from fixation (mean +/- SE) by farm (1-4) for vetch in monoculture (“Vetch”) and vetch in mixture with wheat (“mix”). No significant differences between Vetch and mix for any farm. For full model results see Table 2S.

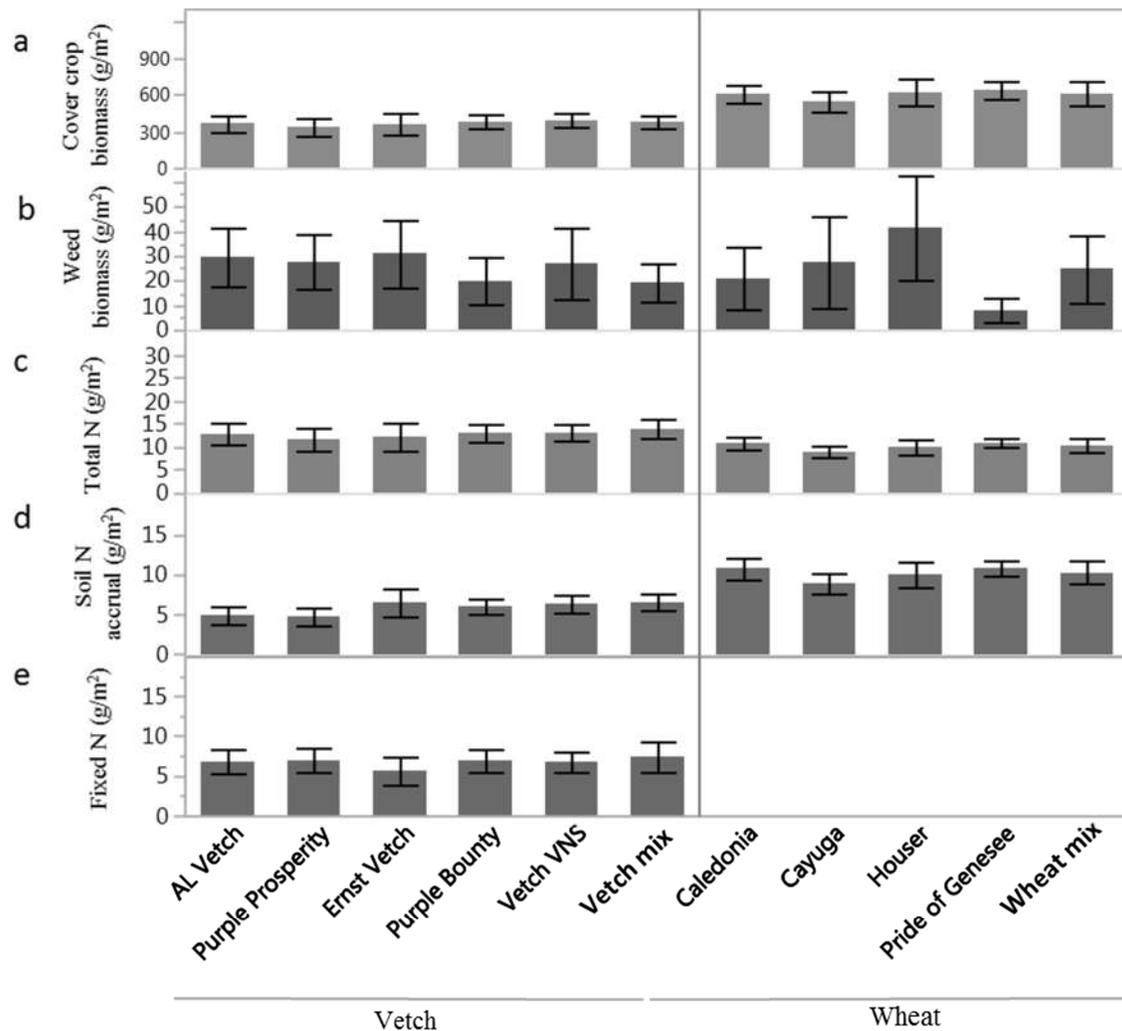


Figure 3S Average across all farms for a) aboveground cover crop biomass b) weed biomass c) total N in biomass d) soil N accrual and e) fixed nitrogen for each vetch and wheat cultivar and cultivar mixture (mean \pm SE).

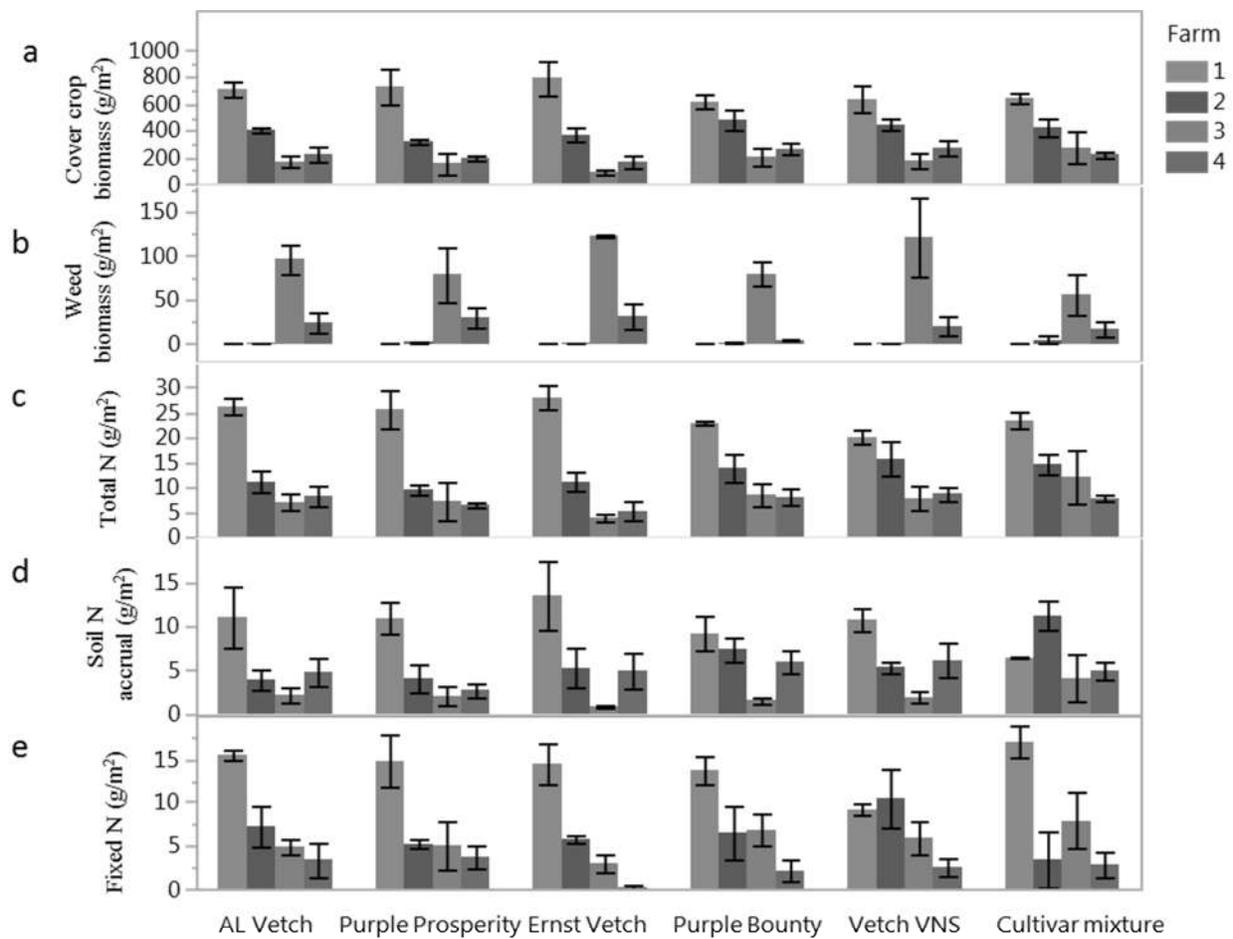


Figure 4S Average a) aboveground cover crop biomass b) weed biomass c) total N in biomass d) soil N accrual and e) fixed nitrogen for each vetch cultivar and cultivar mixture by farm (mean +/- SE).

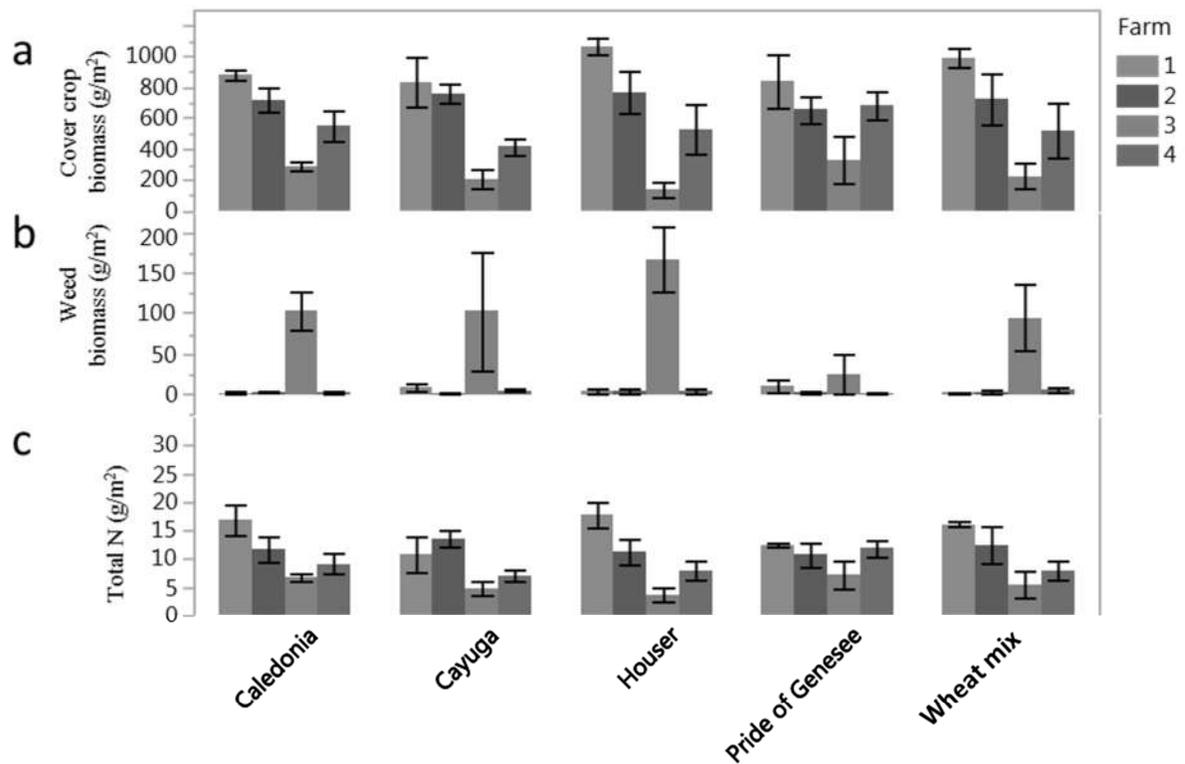


Figure 5S Average a) aboveground cover crop biomass b) weed biomass c) total N in biomass for each wheat cultivar and cultivar mixture by farm (mean +/- SE).

Table 4S Model results for effect of intraspecific diversity (ID) on five ecosystem services by species (Vetch and Wheat) at the four sites (“Farm”). Test of ID is the mean of the cultivar monocultures compared to cultivar mixture by species. Data transformed for noted variables. DF= Degree of Freedom in the numerator; DFDen= Degrees of freedom in the denominator

Ecosystem service	Vetch					Wheat				
Cover crop biomass	Source	DF	DFDen	F Ratio	Prob>F	Source	DF	DFDen	F Ratio	Prob>F
	Farm	3	9	22.5603	0.0002*	Farm	3	9	11.3664	0.0020*
	ID	1	9	0.1812	0.6803	ID	1	9	0.0193	0.8925
	Farm*ID	3	9	0.7909	0.5289	Farm*ID	3	9	0.1805	0.9069
Weed biomass (Natural log transformed)	Source	DF	DFDen	F Ratio	Prob>F	Source	DF	DFDen	F Ratio	Prob>F
	Farm	3	9	35.1874	0.0001*	Farm	3	9	19.1166	0.0003*
	ID	1	9	0.9301	0.3600	ID	1	9	1.3750	0.2711
	Farm*ID	3	9	0.6763	0.5881	Farm*ID	3	9	0.8662	0.4934
Soil N (Square root transformed)	Source	DF	DFDen	F Ratio	Prob>F	Source	DF	DFDen	F Ratio	Prob>F
	Farm	3	9	7.5164	0.0080*	Farm	3	9	8.505	0.0054*
	ID	1	9	0.5227	0.488	ID	1	9	0.01	0.9227
	Farm*ID	3	9	2.7737	0.1028	Farm*ID	3	9	0.4285	0.7375
Fixed N	Source	DF	DFDen	F Ratio	Prob>F	NA				
	Farm	3	9	10.0154	0.0032*					
	ID	1	9	0.4285	0.5291					
	Farm*ID	3	9	2.6814	0.1099					
Total N (Square root transformed)	Source	DF	DFDen	F Ratio	Prob>F	See soil N				
	Farm	3	9	14.3841	0.0008*					
	ID	1	9	1.1948	0.3028					
	Farm*ID	3	9	0.4918	0.6968					

Table 5S Model results for effect of cultivar identity (“cultivar”) on five ecosystem services by species (Vetch and Wheat) at the four sites (“Farm”). Details of cultivars and cultivar mixture are in Table 3. Data transformed for noted variables.

Ecosystem service	Vetch					Wheat				
	Source	DF	DFDen	F Ratio	Prob>F	Source	DF	DFDen	F Ratio	Prob>F
Cover crop biomass	cultivar	5	42.6	0.3753	0.8628	cultivar	4	31.85	0.336	0.8516
	Farm	3	8.279	37.3123	<.0001*	Farm	3	8.57	18.054	0.0005*
	cultivar*Farm	15	42.52	1.0016	0.4714	cultivar*Farm	12	31.73	0.5785	0.8428
Weed biomass (Natural log transformed)	Source	DF	DFDen	F Ratio	Prob>F	Source	DF	DFDen	F Ratio	Prob>F
	cultivar	5	43.76	0.4124	0.8376	cultivar	4	32.72	0.8587	0.4989
	Farm	3	8.957	102.1407	<.0001	Farm	3	9.858	15.5434	0.0005
	cultivar*Farm	15	43.63	0.5367	0.9049	cultivar*Farm	12	32.63	1.675	0.1191
Soil N (Square root transformed)	Source	DF	DFDen	F Ratio	Prob>F	Source	DF	DFDen	F Ratio	Prob>F
	cultivar	5	42.45	0.7193	0.6125	cultivar	4	29.73	0.9911	0.4276
	Farm	3	8.762	16.6418	0.0006*	Farm	3	7.028	9.0782	0.0082*
	cultivar*Farm	15	42.31	1.3446	0.2196	cultivar*Farm	12	29.62	1.2698	0.2865
Fixed N (Square root transformed)	Source	DF	DFDen	F Ratio	Prob>F	NA				
	cultivar	5	42.08	0.6382	0.6718					
	Farm	3	8.724	12.0063	0.0019*					
	cultivar*Farm	15	42	1.4137	0.1856					
Total N	Source	DF	DFDen	F Ratio	Prob>F	See soil N				
	cultivar	5	42.52	0.6794	0.6415					
	Farm	3	8.295	29.2734	<.0001*					
	cultivar*Farm	15	42.46	1.4813	0.156					

Table 6S Mean and standard error of density of vetch and wheat plants at biomass harvest by cover crop composition (pooled across farms)

	Vetch (plants/m ²)		Wheat (plants/m ²)	
	Mean	Std. Error	Mean	Std. Error
Biculture	49	5	31	2
Wheat	NA	NA	59	7
Vetch	76	9	NA	NA

Table 7S Model results for effect of species mixture on five ecosystem services. Three evaluations 1-2) performance of vetch or wheat in monoculture to the vetch or wheat component in mixture, and 3) the overall performance of the biculture mixture compared to the two monocultures. Data transformed for noted variables. DF= Degree of Freedom in the numerator; DFDen= Degrees of freedom in the denominator

Ecosystem service	1) Vetch mono compared to vetch in biculture					2) Wheat mono compared to wheat in biculture					3) Biculture compared to wheat and vetch in monoculture								
Cover crop biomass	Source	DF	DFDen	F Ratio	Prob > F	Source	DF	DFDen	F Ratio	Prob > F	Source	DF	DFDen	F Ratio	Prob > F				
	Farm	3	9	30.9674	<0.001*	Farm	3	9	15.4402	0.0007*	Farm	3	9	19.074	0.0003*				
	mix	1	9	19.2935	0.0017*	mix	1	9	58.2491	<0.001*	mix	2	18	23.2569	<0.001*				
	Farm*mix	3	9	1.424	0.2987	Farm*mix	3	9	12.0934	0.0016*	Farm*mix	6	18	4.0732	0.0094*				
Weed biomass (natural log transformed)	No interaction between diversity and farm																		
	Significant effect of mix with vetch mono>mix																		
	mix	1	12	18.6078	0.001	Farm 2	F(1,2)=28.9627, p=0.0328	Farm 3	F(1,2)=1.4115, p=0.3568	Farm 4	F(1,3)=0.1164, p=0.7555	Farm 1	F(2,4)=50.8383, p=0.0014	Farm 2	F(2,4)=5.8129, p=0.0655	Farm 3	F(2,4)=38.3787, p=0.0025	Farm 4	F(2,6)=9.4535, p=0.014
	Farm	3	9	30.9674	<0.001	Farm 4	F(1,3)=0.1164, p=0.7555												
Soil N (Square root transformed)	NA																		
	Source	DF	DFDen	F Ratio	Prob > F	Source	DF	DFDen	F Ratio	Prob > F	Source	DF	DFDen	F Ratio	Prob > F				
	Farm	3	9	5.8612	0.0168*	Farm	3	9	6.876	0.0105*	Farm	3	9	77.6326	<0.001*				
	mix	1	9	33.0358	0.0003*	mix	1	9	55.8119	<0.001*	mix	2	18	25.0479	<0.001*				
Fixed N	See mix																		
	Farm*mix	3	9	3.9276	0.0480*	Farm*mix	3	9	13.1205	0.0012*	Farm*mix	6	18	5.9612	0.0014*				
	Farm 1	F(1,2)=12.105, p=0.0736	Farm 2	F(1,2)=8.8431, p=0.0969	Farm 3	F(1,2)=0.6503, p=0.5046	Farm 4	F(1,3)=12.314, p=0.0392	Farm 1	F(1,2)=64.8357, p=0.0151	Farm 2	F(1,2)=22.3503, p=0.0419	Farm 3	F(1,2)=1.5808, p=0.3356	Farm 4	F(1,3)=0.0037, p=0.9553			
	Farm 4	F(1,3)=12.314, p=0.0392																	
Total N	NA																		
	Source	DF	DFDen	F Ratio	Prob > F	See soil N					No interaction between diversity and farm								
	Farm	3	9	24.5898	0.0001														
	mix	1	9	21.8087	0.0012														
% N fixation	See soil N																		
	Farm*mix	3	9	2.7496	0.1046														
	No interaction between diversity and farm																		
	Significant effect of farm and mix with vetch mono>mix																		
Farm	3	9	24.5898	0.0001*															
mix	1	12	16.1238	0.0017*															
Source DF DFDen F Ratio Prob > F																			
Farm	3	9	13.1826	0.0012															
mix	1	9	5.9638	0.0372															
Farm*mix	3	9	2.7138	0.1073															
No interaction between diversity and farm																			
Significant effect of farm, but not mix																			
Farm	3	9	13.1826	0.0012															
mix	1	12	3.5476	0.0841															

Table 8S Seed sources of cultivar or line by species

		Hairy vetch (<i>Vicia villosa</i>)	
		Cultivar or line	Source
Legumes		'AL Vetch'	Albert Lea Seed
		'Vetch VNS'	Albert Lea Seed
		'Ernst Vetch VNS'	Ernst Conservation Seeds
		'Purple Bounty'	Ernst Conservation Seeds
		'Purple Prosperity'	USDA NRCS National Plant Materials Center, Beltsville, MD
		Wheat (<i>Triticum aestivum</i>)	
		Cultivar	Source
Grasses		'Cayuga'	Cornell Small Grains Breeding & Genetics Program
		'Caledonia'	
		'Houser'	
		Pride of Genesee'	

Figure 9S Model results for vetch and wheat performance by five ecosystem services at the four sites (“Farm”). Also see Figure 1. Data transformed for noted variables.

Vetch monocultures by farm	Source	Nparm	DF	DFDen	F Ratio	Prob > F	Transformation
Cover crop biomass	Farm	3	3	8.455	21.3674	0.0003	square root
Weed biomass	Farm	3	3	8.769	97.2307	<0.0001	ln
Total N	Farm	3	3	8.497	16.9504	0.0006	square root
Soil N accrual	Farm	3	3	8.754	15.9928	0.0007	square root
Fixed N	Farm	3	3	8.796	11.9481	0.0019	square root

Wheat monocultures by farm	Source	Nparm	DF	DFDen	F Ratio	Prob > F	Transformation
Cover crop biomass	Farm	3	3	8.872	23.0688	0.0002	square root
Weed biomass	Farm	3	3	9.683	17.1145	0.0003	ln
Total N	Farm	3	3	7.584	14.4182	0.0017	square root
Soil N accrual	Farm	3	3	7.584	14.4182	0.0017	square root

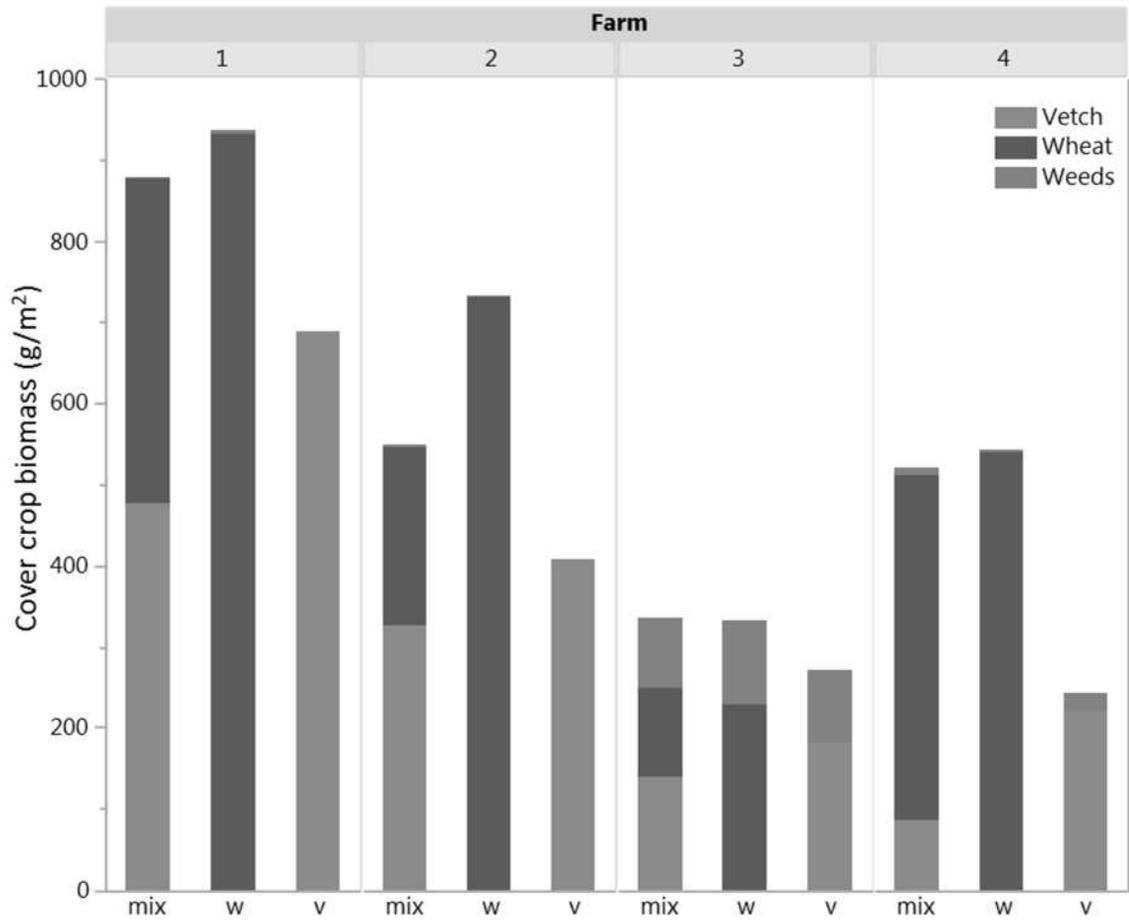


Figure 6S Mean biomass of vetch, wheat, and weed biomass by cover crop composition and by farm

APPENDIX 1

Figure 1A Field collected data for additional farm site with pea and rye monocultures and mixtures. NA= not applicable (e.g. a grass doesn't have %Ndfa), ND= No data collected or recorded

Plot ID	g/m2			plants/m2			Pea				Rye				Pea	
	Pea biomass	Rye biomass	Weed biomass	Pea density	Rye density	Weed density	%C	%N	dC13	dN15	%C	%N	dC13	dN15	B value	%Ndfa
1-10	0	520.24	238.44	0	100	448	NA	NA	NA	NA	ND	ND	ND	ND	NA	NA
1-11	0	611.4	186	0	68	440	NA	NA	NA	NA	44.28%	1.17%	-28.69413375	2.281429561	NA	NA
1-21	241	0	364	ND	0	668	43.36%	3.11%	-27.86070461	0.103664544	NA	NA	NA	NA	-0.825	70.1%
1-22	0	434.64	408.84	0	64	780	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
1-35	229.36	442.76	265.96	72	68	676	43.20%	3.12%	-27.88854216	0.176547493	NA	NA	NA	NA	-0.825	79.1%
1-50	0	0	532	0	0	816	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
1-8	337.36	0	438.96	84	0	620	43.61%	3.12%	-28.45463494	0.237346199	NA	NA	NA	NA	-0.879	64.7%
1-9	201.72	0	350	12	0	920	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
2-10	0	647.04	212.48	ND	0	144	NA	NA	NA	NA	ND	ND	ND	ND	NA	NA
2-11	0	378.76	175.76	0	48	408	NA	NA	NA	NA	44.26%	1.59%	-28.41407662	6.75957428	NA	NA
2-21	270.84	0	319.76	56	ND	ND	43.73%	3.47%	-28.73898162	0.124128557	NA	NA	NA	NA	-0.825	80.0%
2-22	0	458.36	248.16	0	32	476	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
2-35	273.2	437.8	290.88	40	20	228	ND	ND	ND	ND	NA	NA	NA	NA	ND	ND
2-50	0	0	288.28	0	0	544	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
2-8	265.92	0	307.08	52	0	328	43.19%	3.69%	-28.61472801	0.200741957	NA	NA	NA	NA	-0.879	80.9%
2-9	200.16	0	293.96	4	0	164	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
3-10	0	337.16	269.84	0	44	ND	NA	NA	NA	NA	43.79%	1.46%	-28.116719653	2.265534144	NA	NA
3-11	0	596.76	205.16	0	48	348	NA	NA	NA	NA	43.85%	0.87%	-29.37823856	1.694515975	NA	NA
3-21	215.84	0	487.76	12	0	588	43.35%	3.17%	-28.2671923	0.116516864	NA	NA	NA	NA	-0.825	71.5%
3-22	0	597.84	155.24	0	108	284	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
3-35	220.32	263.4	343.28	32	24	416	43.67%	3.35%	-28.87667295	0.088872014	NA	NA	NA	NA	-0.825	72.3%
3-50	0	0	448.6	0	0	588	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
3-8	385	0	495.76	92	0	380	42.81%	3.18%	-28.07412864	0.008612086	NA	NA	NA	NA	-0.879	73.6%